

Max Planck Research Group  
Epistemes of Modern Acoustics

---

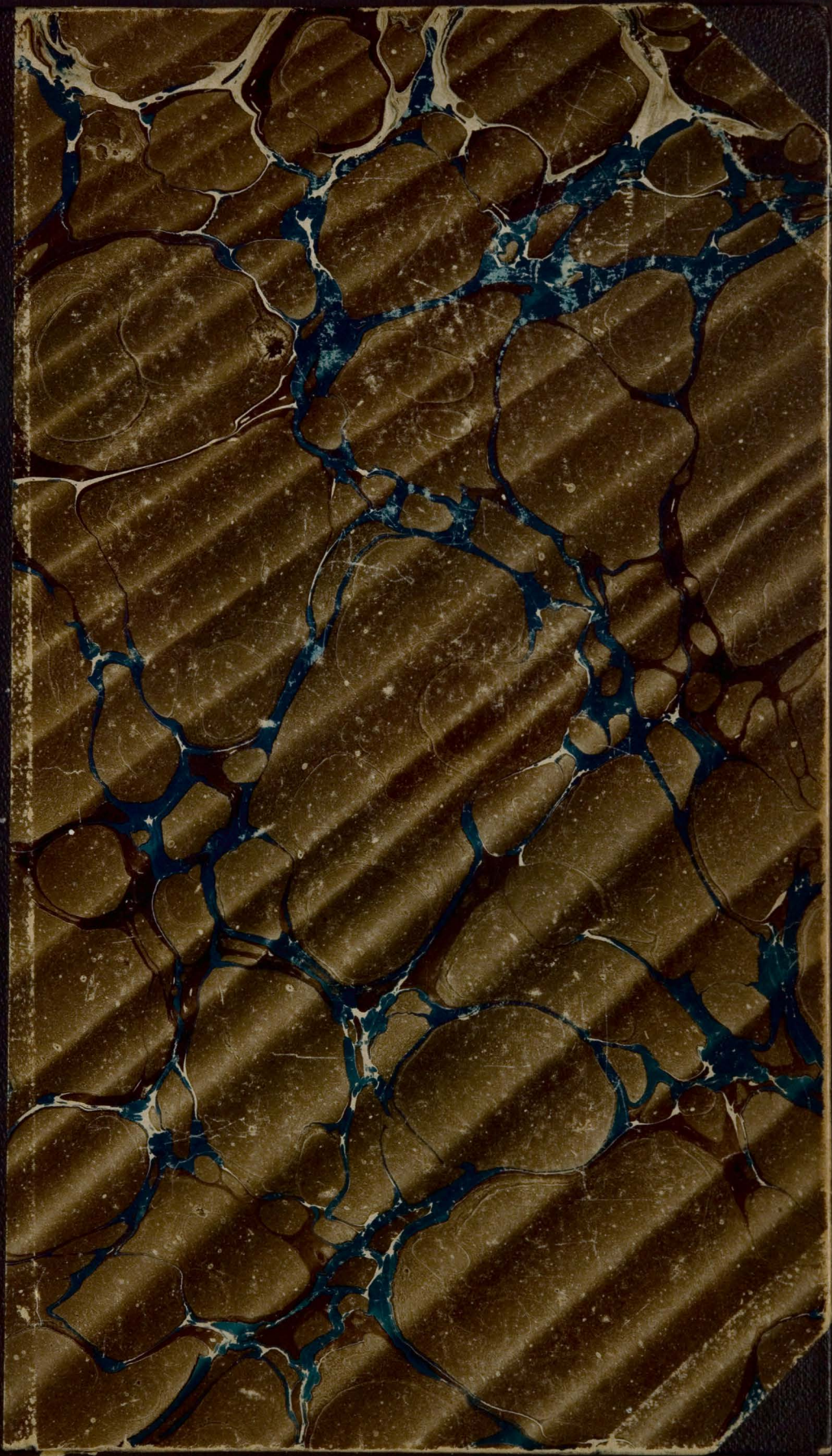
# Sound & Science: Digital Histories



Scan licensed under: [CC BY-SA 3.0 DE](https://creativecommons.org/licenses/by-sa/3.0/de/) | Max Planck Institute for the History of Science



**MAX PLANCK INSTITUTE  
FOR THE HISTORY OF SCIENCE**



393

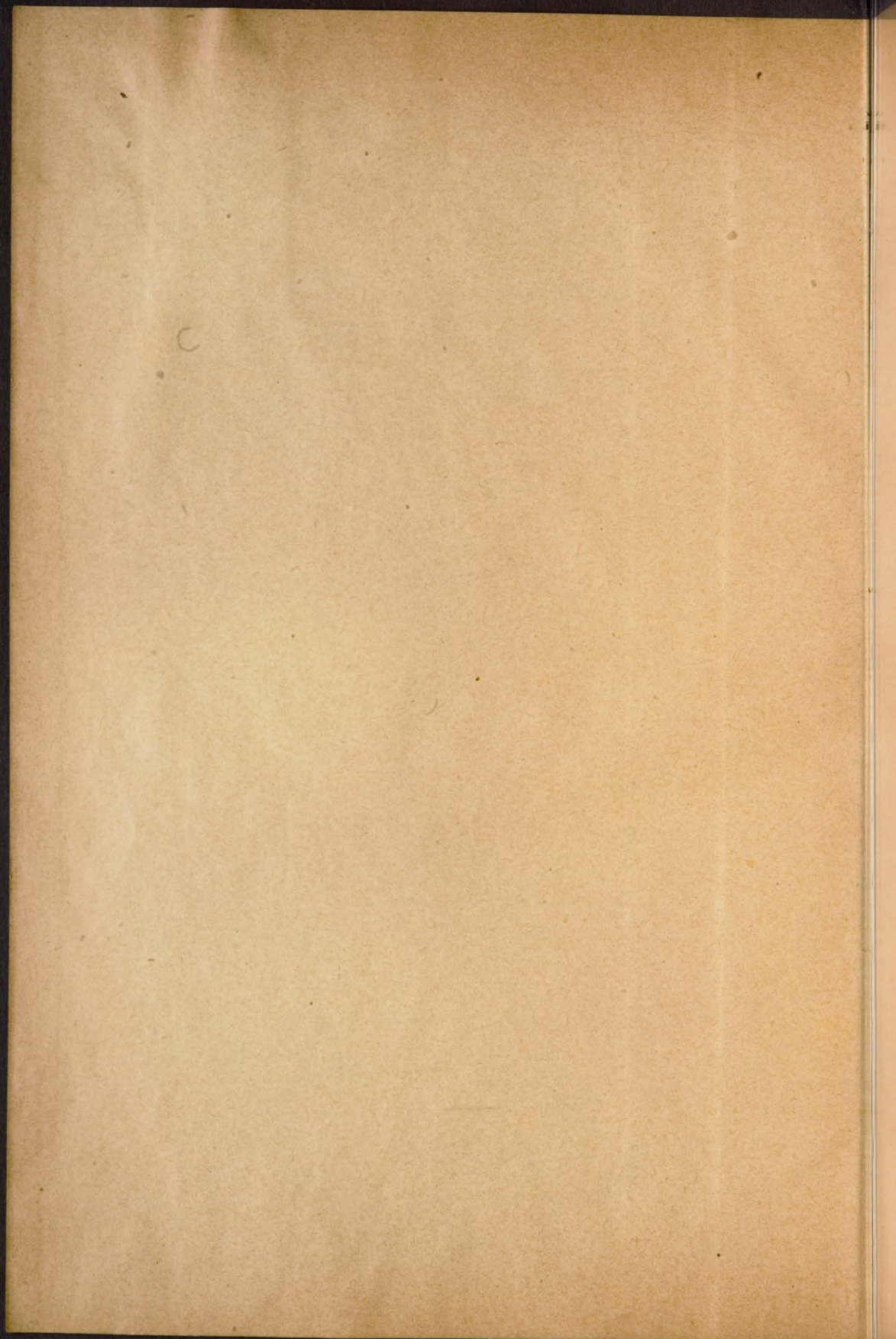


c

-5900- 27509



119E



# DIE SPRACHLAUTE

EXPERIMENTELL-PHONETISCHE  
UNTERSUCHUNGEN

NEBST EINEM ANHANG ÜBER  
INSTRUMENTALKLÄNGE

VON

**CARL STUMPF**

DR. PHIL., DR. MED. H. C.  
O. PROFESSOR AN DER UNIVERSITÄT ZU BERLIN

MIT 8 TEXTFIGUREN UND 8 NOTENBILDERN



BERLIN  
VERLAG VON JULIUS SPRINGER  
1926

534 59345s

C

ALLE RECHTE, INSBESONDERE DAS DER ÜBERSETZUNG  
IN FREMDE SPRACHEN, VORBEHALTEN.  
COPYRIGHT 1926 BY JULIUS SPRINGER IN BERLIN.

534x

MAX-PLANCK-INSTITUT  
FÜR WISSENSCHAFTSGESCHICHTE  
Bibliothek

07-907

DER  
MEDIZINISCHEN FAKULTÄT  
DER UNIVERSITÄT BERLIN  
IN DANKBARKEIT GEWIDMET



C

DR  
FACULTÄT DER MEDICIN  
DER UNIVERSITÄT BERLIN  
IN DANKBAREN GEBEN

## Vorwort.

Die experimentellen Arbeiten, deren Ergebnisse das vorliegende Buch zusammenfaßt, begannen 1913. In den Kriegsjahren, als draußen die Geschütze donnerten, war es um so stiller auf den Straßen und in den wissenschaftlichen Instituten Berlins. Fast wie in einem schalldichten Raume ließen sich akustische Beobachtungen durchführen und selbst die Flüsterlaute in ihre letzten Bestandteile zerlegen. So versuchte ich denn, wie wenig auch sonst die Stimmung dazu drängte, in dieser furchtbaren Zeit die wissenschaftliche Arbeit, wie sie gerade auf dem Gebiete der Phonetik seit HELMHOLTZ in Deutschland intensiv gepflegt worden war, weiterzuführen. Den Anstoß dazu hatten mir die Vokalstudien WOLFGANG KOEHLERS gegeben, über die ich auf dem Kongreß für experimentelle Psychologie zu Göttingen April 1914 vorzutragen hatte. Die Arbeit mußte aber oft längere Zeit unterbrochen werden, da die Herstellung und Reparatur der nötigen Einrichtungen während des Krieges große Schwierigkeiten machte. Über die einzelnen Stadien wurde der Berliner Akademie der Wissenschaften berichtet (28. V. 1914 Resonanzversuche, 22. VII. 1915 Interferenzversuche, 1. XI. 1917 Synthesen). Die erste Übersicht der experimentellen Ausbeute findet sich im akademischen Sitzungsbericht vom 4. IV. 1918. Es folgten weitere Monographien, die im Literaturverzeichnis dieses Buches zitiert sind.

Hier sind nun die Beobachtungen weit vollständiger zusammengestellt, die früher offen gelassenen allgemeinen Fragen erörtert, und das Ganze ist in einen gemeinschaftlichen Rahmen eingefügt. Inzwischen sind manche der bereits veröffentlichten Tabellen in Lehrbücher übergegangen, manche wurden für die ohrenärztliche Praxis, ja auch für die Telephonie und Radiophonie nützlich ge-

funden und dabei zugleich bestätigt. Während des letzten Dezzenniums haben sich aber auch amerikanische Physiker und Ingenieure vielfach mit denselben Fragen beschäftigt. Ihre auf gänzlich anderen Wegen durchgeführten Untersuchungen führten in den wichtigsten Punkten zu Ergebnissen, die sich mit den meinigen zu einem widerspruchslosen Gesamtbilde abrunden. Bei der absoluten gegenseitigen Unabhängigkeit ist dies ein erfreuliches Zeichen, daß wir in diesen schwierigen Fragen endlich dem Ziele der Erkenntnis, der objektiven Wahrheit näherkommen. In allem Wesentlichen hat sich HELMHOLTZENS vielumstrittene Vokaltheorie, die auch dem Verfasser keineswegs von vornherein unbezweifelbar erschien, doch zuletzt als siegreich erwiesen.

Das Interesse des Verfassers war nicht auf praktische Ziele gerichtet, sondern auf die rein theoretischen Fragen nach der Entstehung der sog. Komplexqualitäten und ihrem Zusammenhang mit den Eigenschaften der Elemente. Klangfarben und speziell Vokale liefern dafür die bekanntesten Beispiele und das geeignetste Untersuchungsmaterial. Die Ausführungen über Klangfarbe im Schlußparagraphen meiner „Tonpsychologie“ bedurften in dieser Hinsicht einer Revision und Ergänzung, weshalb die gegenwärtige Arbeit als unmittelbare Fortsetzung der früheren betrachtet werden kann. Von diesem Standpunkte bitte ich die Stellung der Probleme und die Richtung der ganzen Diskussion zu beurteilen.

Wie HELMHOLTZ neben den Vokalen auch Instrumentalklänge eingehend untersuchte, so hat sich auch dem Verfasser die Notwendigkeit solcher Parallelversuche aufgedrängt. Die auf die Vokale angewandten Methoden der Analyse und Synthese ließen sich in der Hauptsache darauf übertragen. Aber hier bin ich zu einer erschöpfenden Übersicht über die außerordentliche Mannigfaltigkeit der Erscheinungen nicht gelangt und kann nur sagen, daß HELMHOLTZENS Theorie hier allerdings nicht bloß ergänzt, sondern allem Anscheine nach auch modifiziert werden muß, daß sich zwischen Vokal- und Instrumentalklängen eine völlig scharfe Grenze nicht ziehen läßt, daß aber gerade dieser Umstand die Unterordnung der stimmhaften Vokale unter den Allgemeinbegriff der Klangfarben nur noch entschiedener rechtfertigt.

Linguistische, hauptsächlich auf die Erzeugungsweise der Sprachlaute gerichtete Fragen standen hier ebenso wie gesangstechnische nicht in erster Linie. Doch darf ich aus dem Interesse, mit dem auch Sprachforscher gelegentliche Demonstrationen aufgenommen haben, die Hoffnung schöpfen, daß die neue Durcharbeitung des akustischen Materials auch diesen Kreisen nicht unwillkommen erscheinen und daß der Nachweis fast völliger Übereinstimmung in den neuesten Untersuchungen das durch den zeitweiligen Stand der Forschungen hervorgerufene Mißtrauen beseitigen werde. Wieweit Sprachforschung und Gesanglehre positiven Gewinn aus dem Buche ziehen können, muß dem Urteil der Fachmänner überlassen bleiben.

Sollte dieses Buch philosophischen Kollegen in die Hände kommen, so werden wohl viele es bald kopfschüttelnd wieder aus der Hand legen. Es erhebt in dieser Richtung keine Ansprüche. Aber schließlich kann der Philosoph aus allen Blüten Honig saugen, und mir ist es nun einmal gewiß, daß auch die hohe Königin der Wissenschaften niemals anders wahrhaft fortschreiten wird, als auf dem Wege vom Einzelnen zum Allgemeinen.

Nicht unterlassen möchte ich, dem Verlage JULIUS SPRINGER für die außerordentlich bereitwillige Erfüllung aller meiner Wünsche in Hinsicht des Druckes und der Ausstattung dieses Buches meinen herzlichen Dank zu sagen.

Berlin, im Mai 1926.

CARL STUMPF.

## Inhaltsverzeichnis.

	Seite
Einleitung . . . . .	1
Richtungen und Forschungsmethoden der Phonetik. Antinomien der Vokalforschung.	
1. Kapitel: Analyse gesungener Vokale durch resonierende Gabeln . . . . .	9
I. Methode. . . . .	9
Wesen und Vorzüge der Gabelmethode. Keine multiple Reso- nanz. Resonanzbreite 9. Stärkeschätzungen. Vergleichung mit physikalischen Stärken 13. Gabeln und Sänger 18.	
II. Ergebnisse . . . . .	20
Zahl der Teiltöne 21. Harmonische Teiltöne 21. Charak- teristische Struktur 22. Zahl und Lage der Maxima 25. Ver- lauf der Stärkekurven. Stärke der tiefsten Teiltöne. Verände- rungen bei und über $c^2$ 27. Individuelle und Register-Unter- schiede 28. Schranken der Gabelmethode 31.	
Anhang: Resonanzversuche am Klavier 32.	
2. Kapitel: Analyse gesungener Vokale durch Interferenz- röhren . . . . .	36
I. Prinzipielle Schwierigkeiten und deren Lösung . . . . .	37
Gleichzeitige Auslöschung der ungeraden Multipla. Ab- und Aufbaureihen, Lücken-, Stich- und Isolierversuche 37. Ver- stärkung der geraden Multipla 39. Schwächung des Grund- tones 40. Bauch- und Knotenwirkung 41. Veränderung des Klanges durch die Hauptleitung 42. Interferenzbreite 43.	
II. Weiteres zur Technik und Methodik . . . . .	44
Leitung und Röhrensysteme 44. Einfluß der Röhrenweite auf die erforderliche Einstellung 45. Probe durch schwebende Hilfsgabeln 48. Wissentliches und unwissentliches Verfahren 49. Verschiedenheiten beim Auf- und Abbau 51. Analy- sierendes Verhalten 52. Beschreibung des Gehörten 52.	
III. Ergebnisse . . . . .	53
Das untersuchte Lautmaterial und die Versuchsreihen 53. Nur harmonische Teiltöne 54. Entwicklungsstadien 55. Die Formanten 62. Ihr langsames Hinaufrücken mit der Höhe	

	Seite
des Grundtones 65. Unwissentliche Aufbaureihe 68. Die Erscheinungen beim Grundton $c^2$ 70. Lücken- und Stichversuche 71. Isolierversuche 75.	
3. Kapitel: Das Unkenntlichwerden der Vokale in der höheren Sopranlage . . . . .	77
4. Kapitel: Das stimmhafte Sprechen und dessen Abbau durch das Interferenzverfahren . . . . .	86
Sprache und Sprachmaterial 86. Singen und Sprechen 88. Das Interferenzverfahren bei der gesprochenen Rede 91. Die BEZOLDSche „Sprachsext“ 95.	
5. Kapitel: Struktur der Flüstervokale und Konsonanten 98	
I. Klänge und Geräusche, Vokale und Konsonanten, stimmhafte und stimmlose Sprachlaute . . . . .	98
II. Analyse der Flüstervokale durch Interferenzröhren . . . . .	104
Ab- und Aufbauversuche 105. Lücken- und Stichversuche 108.	
III. Analyse von Konsonanten durch Interferenzröhren . . . . .	110
R; K, T, P; Ch gutt.; H; Sch; M, N, Ng; L; S; Ch pal.	
IV. Ergebnisse graphischer Methoden bei Konsonanten . . . . .	129
V. Übersicht der Veränderungen stimmloser Sprachlaute beim Abbau durch Interferenzröhren . . . . .	132
Anhang:	
I. Zur Analyse sonstiger Geräusche . . . . .	134
II. Die obere Hörgrenze 136. Ihr Sinken mit dem Alter 139.	
6. Kapitel: Die subjektiven Tonhöhen der stimmlosen Sprachlaute . . . . .	142
Tonhöhen von Klängen und Geräuschen 142. Tonhöhen der Flüstervokale 143. Frühere Beobachtungen 148. Hauch- und Pfeiftöne 154. Tonhöhen der stimmlosen Konsonanten 157. Verschiebung der Tonhöhen stimmloser Laute bei Interferenzversuchen 162. Deutung der gefundenen Tonhöhen 164.	
7. Kapitel: Synthetische Darstellung der stimmhaften Vokale . . . . .	167
I. Historisches 167. II. Die synthetische Einrichtung 171. III. Ergebnisse 175. IV. Unwissentliche Versuche zur Prüfung der Naturtreue 180. V. Der Grundton als Differenzton 185. VI. Unharmonische Teiltöne 188. VII. Vokale aus nur 2 Tönen 195. VIII. Besondere Vokaltypen 195 (russisches Y, Nasallaute). IX. Stimmfärbungen (Falsett, Knabenstimmen) 197.	
8. Kapitel: Die Entwicklung der Vokalforschungen und die Konvergenz ihrer Ergebnisse . . . . .	202
I. Hauptergebnisse der Vokalforschung seit WILLIS . . . . .	202
Resonanz- und Interferenzmethode. Subjektive Analyse 202. Graphische Methoden 207. Bestimmung des energetisch	

stärksten Teiltönen 213. Synthesen 214. Verknüpfung mehrerer Methoden 215. Die Analysen und Synthesen D. CH. MILLERS 217.	
II. Kritik einiger pseudo-synthetischer Versuche der neueren Zeit 220	
HERMANN'S Versuch mit der Doppelsirene 220. JAENSCH'S Versuche mit der Selsensirene 221. TER KUILES Versuche 226.	
9. Kapitel: Phonographische, telephonische und ohrenärztliche Beobachtungen . . . . .	228
I. Die Veränderungen der Vokale bei veränderter Umdrehungsgeschwindigkeit der Phonographenwalze . . . . .	228
II. Erfahrungen und Versuche in der Telephonie . . . . .	234
III. Ohrenärztliche Beobachtungen . . . . .	239
10. Kapitel: Systematik der Sprachlaute vom akustischen Standpunkte . . . . .	245
I. Vokale . . . . .	246
Sukzessiv ein- und mehrteilige 246. Simultan ein- und mehrteilige 246. Die 5 sog. Hauptvokale 247. Das Helligkeitsprinzip 249. Das Vokaldreieck 252. Freie und näselnde Vokale 257. Offenes und gedecktes Singen und offene und geschlossene Aussprache 258. Multiple Vokale 263. Vokalitäten außerhalb der menschlichen Stimme 267.	
II. Konsonanten . . . . .	268
11. Kapitel: Einheitliches und mehrheitliches Hören . . . . .	276
Allgemeinere Fragestellungen 276. Komplex- und Gestalteigenschaften 277. Günstige Bedingungen für das einheitliche Hören 280.	
12. Kapitel: Die gegenseitige Beeinflussung der Teiltonstärken und ihr Verhältnis zur Gesamtstärke des Klanges 290	
I. Gegenseitige Beeinflussung . . . . .	290
Ältere Beobachtungen 291. Neuere amerikanische Untersuchungen 292. Eigene Beobachtungen 297. Ergebnis für die Vokalstruktur 301. Die der phänomenalen Stärke entsprechende physikalische Intensität herausgehörter Teiltöne 302.	
II. Verhältnis der Teiltonstärken zur Gesamtstärke des Klanges 305	
13. Kapitel: Psychophysik der Sprachlaute . . . . .	310
Allgemeine Ergebnisse betreffs gesungener und stimmhaft gesprochener Vokale 310. Die Formanten des stimmhaften U 315. Die U-I-Linie des Vokaldreiecks 317. Die Vokale außerhalb der U-I-Linie. KOEHLER'S Vokalitäten 320. Zwei zentralphysiologische Prozesse 331. Besondere Gesetzmäßigkeiten hinsichtlich des Zusammenwirkens der Teiltöne 335. Ursprung des akzessorischen Prozesses 338. Formantverschiebung mit der Höhe des Grundtones und Unkenntlichwerden der Vokale jenseits $c^2$ 340. Vokalcharakter der Abbauprodukte 342. Vokalität der Flüster-	

vokale 343. Vokale und Geräusche 345. Zentralphysiologisches über die Konsonanten 347.

14. Kapitel: Zur Physik und Physiologie der Sprachlaute 349  
 I. Physikalisches . . . . . 349  
 II. Physiologisches . . . . . 352  
 Zur Erzeugung der Sprachlaute 352. Zur physiologischen Hörtheorie 357.

15. Kapitel (Anhang): Über Instrumentalklänge . . . . . 374  
 Erkennen instrumentaler Klangfarben 374. Hauptergebnisse bezüglich der Klangstrukturen 376. Resonanzversuche 378. Interferenzversuche 382. Synthesen 386. Allgemeines über das Wesen und die Unterschiede der Klangfarbe 389. Tiefere Erklärungsgründe 395. Klangmerkmale der Instrumente 402. Gruppierungen und Mischungen 407. Blick auf frühere und künftige Untersuchungen 409.

Nachträge 411. Druckberichtigungen 412. Literaturverzeichnis 413.

**Verzeichnis der Textfiguren.**

1. Schema der Einrichtungen für Interferenzversuche und Synthesen 44  
 2. Struktur stimmloser Vokale und Konsonanten . . . . . 107  
 3. Schema der 5 „Hauptvokale“ . . . . . 251  
 4. Das Vokaldreieck . . . . . 252  
 5. Schema der Vokalitäten nach KOEHLER . . . . . 321  
 6. Grundprozeß und akzessorischer Prozeß . . . . . 331  
 7. und 8. Zusammengesetzte Wellenformen . . . . . 360

**Tabellen.**

Ältere Formantbestimmungen (Notentabellen). . . . . 6  
 Amplituden und subjektive Stärken . . . . . 18  
 Teiltonstärken der Vokale nach der Resonanzmethode . . 23, 29, 30, 31  
 Interferenzeinstellungen (in cm) für die Töne der chromatischen Leiter von  $c^1$  bis  $c^7$  . . . . . 47  
 Umwandlungen stimmhafter Vokale beim Aufbau durch Interferenz 56, 69  
 Die Interferenzformanten auf Grundtönen von  $C$  bis  $c^2$  . . . . . 66  
 Gesamtumfänge und Zentralstrecken der If.-Formanten . . . . . 68  
 Statistik der Veränderungen der Vokale oberhalb  $c^2$ . . . . . 78  
 Abbau der stimmhaften Sprache durch If.-Röhren. . . . . 94  
 Aufbau der Flüstervokale durch If.-Röhren. . . . . 106  
 Formanten und Gesamtumfänge stimmloser Sprachlaute . . . . . 113  
 Abbau stimmloser Sprachlaute durch If.-Röhren . . . . . 133  
 Tonhöhen der Flüstervokale (Notentabellen) . . . . . 145, 147, 148  
 Tonhöhen stimmloser Konsonanten . . . . . 158  
 Teiltontabellen nach Synthesen der Vokale auf  $c, c^1, g^1, c^2$  176, 184, 197, 199, 200  
 Schema der Formantenzentren nach den Synthesen . . . . . 179  
 Formanten nach MILLER . . . . . 218  
 Orts- und Grenzbestimmungen des Tonreiches . . . . . 314  
 Formanten und Maxima bei Instrumentalklängen . . . . . 382  
 Teiltontabellen nach Synthesen von Instrumentalklängen auf  $c, c^1, c^2, c^3$  388



## Erläuterungen.

Bei Zitaten verweist die einem Namen beigefügte arabische Ziffer auf die entsprechende Nummer in Literaturverzeichnis am Schlusse des Buches. Bei HERMANN'S Arbeiten wird der Band des PFLÜGERSCHEN Archivs zitiert.

Die Notenbuchstaben beziehen sich bei den Resonanzversuchen und Synthesen (1., 7., 15. Kap.) auf die physikalische Stimmung ( $C = 64$ ,  $a^1 = 426\frac{2}{3}$  Schw.), in welcher die benutzten Gabeln und Pfeifen standen, im übrigen aber auf die heutige Orchesterstimmung und das temperierte System, für das die Schwingungszahlen unten angegeben sind. Doch ist der Unterschied gegenüber der physikalischen Stimmung nur gering und für die vorliegenden Zwecke so gut wie irrelevant.

Als „harmonische Teiltöne“ bezeichnen wir hier gemäß dem physikalischen Sprachgebrauch alle in der objektiven Schwingung durch Analyse nachweisbaren Teilschwingungen, deren Schwingungszahlen sich zu der des Grundtones wie ganze Zahlen zu 1 verhalten. Der Grundton selbst ist 1. Teilton. Der 1. „Oberton“ ist 2. Teilton. Wir verwenden hier ausschließlich die weit zweckmäßigere Zählung nach Teiltönen und sprechen von Obertönen nur gelegentlich, wenn Ordnungszahlen nicht in Betracht kommen. Der Begriff „harmonisch“ fällt in diesem Zusammenhange nicht mit dem in der Musik gebräuchlichen zusammen. Die ungeraden Teiltöne vom 7. an gelten dort nicht als harmonisch zum Grundton, während sie hier den geradzahligen gleichstehen.

Erläuterung der Stärkezahlen (in den Tabellen meist mit 4 multipliziert) 15.

Erläuterung der Lautbezeichnungen (AO, Ao, U + i usf.) 55.

Erläuterungen oft vorkommender Begriffe:

Farbigkeit der Gehörsempfindungen 100, 329.

Formant (Haupt- und Neben-, Ober- und Unterform.) 62, 377.

Komplexeigenschaft 277.

Musikalische Qualität 91 Anm., 142.

Abkürzungen: } D.T. = Differenzton.  
 } If. = Interferenz.  
 } Vp. = Versuchsperson.

Schwingungszahlen der Töne der temperierten zwölfstufigen  
 Leiter für  $a^1 = 435$  Schwingungen (v. d.)

	C	Cis Des	D	Dis Es	E	F	Fis Ges	G	Gis As	A	As B	B
<i>C—H</i>	65	69	73	77	81	86	91	97	103	109	115	122
<i>c—h</i>	129	137	145	154	163	172	183	194	205	218	230	244
<i>c<sup>1</sup>—h<sup>1</sup></i>	259	274	290	308	326	345	366	388	411	435	461	488
<i>c<sup>2</sup>—h<sup>2</sup></i>	517	548	581	615	652	691	732	775	821	870	922	977
<i>c<sup>3</sup>—h<sup>3</sup></i>	1035	1096	1161	1230	1304	1381	1463	1550	1642	1740	1843	1953
<i>c<sup>4</sup>—h<sup>4</sup></i>	2069	2192	2323	2461	2607	2762	2926	3100	3285	3480	3687	3906
<i>c<sup>5</sup>—h<sup>5</sup></i>	4138	4385	4645	4921	5214	5524	5853	6201	6569	6960	7374	7812
<i>c<sup>6</sup>—h<sup>6</sup></i>	8277	8769	9290	9843	10428	11048	11705	12401	13139	13920	14748	15625

## Einleitung.

### 1. Richtungen und Forschungsmethoden der Phonetik.

Zwei Grundrichtungen gehen in der Phonetik nebeneinander her, zum großen Teil gegenseitig unabhängig, vielfach aber auch sich ergänzend und unterstützend. Die eine untersucht die Erzeugungsweise der Sprachlaute, die Rolle der Atmung, des Kehlkopfs, der Mund- und Nasenhöhle, die Stellung der Sprachwerkzeuge bei den einzelnen Lauten; die andere die hervorgebrachten Laute selbst, wie sie dem Hörenden erscheinen, und ihre physikalischen, physiologischen und psychologischen Vorbedingungen.

Physiologen betätigten sich von jeher nach beiden Richtungen. Die akustische Zusammensetzung der Laute ist ihnen für die allgemeine Theorie des Hörens von Bedeutung, die Umstände ihrer Erzeugung für die der Sprachwerkzeuge.

Die Ziele der Sprachforscher sind zunächst vorwiegend organogenetischer (artikulatorischer) Art. Ihnen kommt es vor allem darauf an, die richtige Wiedererzeugung der Laute in die Gewalt zu bekommen, zu wissen, wie man die Kiefern, die Zunge einzustellen habe. Die Kunst des Linguisten in der Nachbildung der verschiedensten Laute beruht auf einer gut eingeübten willkürlichen Beherrschung dieses Instrumentariums. Freilich ist die Klangvorstellung das Primäre an dem ganzen Vorgang und ein scharfes Ohr und ein treues Gedächtnis für Klangeigentümlichkeiten muß die Ausführung kontrollieren. Und wenn auch eine Zerlegung in die letzten Bestandteile dazu nicht erforderlich ist, so werden doch Linguisten bei fortschreitender Vertiefung und Differenzierung ihrer Untersuchungen auch aus einer solchen nicht selten Nutzen ziehen; wie es denn heute schon möglich ist, feine Unterschiede in der Aussprache eines Vokals durch Angabe seines Formantenzentrums genauer als durch irgendwelche phonetische Transkription oder durch Schilderung der Mundstellung festzulegen.

Unter den deutschen Sprachforschern betonten dies besonders TRAUTMANN und BREMER. Dieser sagt S. 173 (vgl. 151ff.): „Die vollkommenste Unterscheidung würde erreicht werden, wenn man statt von einem Vokal

A, E, U usw. von einem Vokal  $b^2$  oder einem Vokal  $d^2 g^3$  sprechen dürfte.“ TRAUTMANN, der die Tonhöhen des Flüsterns vortrefflich beschrieben hat, sagt auch von den Konsonanten nach Untersuchung ihrer Artikulationsstellen (S. 103ff.): „Die Hauptsache bleibt immer der Klang: Angabe der Tonhöhe, also Bestimmung des Grundgeräusches, ist das mindeste, was verlangt werden muß, und zugleich das, was geleistet werden kann, und das letzte Ziel der Bestimmung der Konsonanten muß, wie das Ziel der Bestimmung der Vokale, die vollständige Zerlegung ihrer Klänge sein, so daß man von einem gewissen Laut sagen kann, er besteht aus dem Grundgeräusch  $x$  und den Nebengeräuschen  $x, z, w$ .“ Bezüglich eines von STORM genetisch beschriebenen skandinavischen L meint er (S. 106): „Wir wären sofort im reinen, wenn Storm auch die Tonhöhe angegeben hätte.“

Aber auch STORM selbst, der norwegische Verfasser einer englischen Phonetik, urteilt (S. 156, vgl. S. 343): „Der Laut selbst ist das Bleibende, das Wesentliche, der eigentliche Gegenstand der Ergründung; das akustische Prinzip ist das oberste. Die Kenntnis der Organstellungen, des Mechanismus der Laute, ist für die Wissenschaft unerläßlich, jedoch nur als Mittel, nicht als letztes Ziel zu betrachten.“ Ebenso unter den russischen Linguisten ALEXANDER THOMSON. ROUSSELOT und PIPPING machten von akustischen Hilfsmitteln bereits für ihre Dialektstudien Gebrauch. Der letztere nimmt (3, S. 11; vgl. 2, S. 164ff.) sehr energisch gegen rein genetische Systeme Stellung: „Alle Theoretiker, deren Systeme sich auf Beobachtung der Zungen- und Lippenstellungen gründen, haben stillschweigend angenommen, daß dieselbe Artikulationsform auch denselben Laut erzeugen müsse. Sowie diese nie bewiesene und in der Tat falsche Voraussetzung beseitigt wird, stürzen die Systeme krachend zusammen . . . Ein strenges System . . . kann nur von den konstanten Elementen bei der Vokalbildung ausgehen . . . Das konstante Element . . . zeigt sich unzweideutig in den Luftvibrationen, welche die Botschaft des Mundes zum Ohre befördern.“

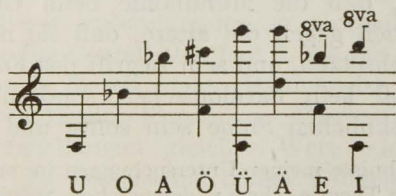
Das Interesse des Psychologen (zu dessen Aufgaben wir hier auch die Beschreibung und Theorie der sinnlichen Erscheinungen als solcher, die Phänomenologie, rechnen) ist in erster Linie ein akustisches. Die allgemeinen Probleme der Klangwahrnehmung und noch tiefer liegende prinzipielle Fragen über sog. „Komplexe“ können an diesem Material mit Vorteil untersucht werden. Gewiß hat auch die Lautproduktion beim Sprechen und Singen ihre psychologische Seite. Doch sind in dieser Hinsicht weniger die Unterschiede in der Erzeugung der einzelnen Laute als die allgemeinen psychologischen Bedingungen des Singens und Sprechens überhaupt wichtig, insofern diese Funktionen hervorragende Beispiele willkürlicher und unwillkürlicher Bewegungen darstellen. In diese Fragen gedenken wir uns hier nicht eingehender zu vertiefen. Vielmehr ist unser Ziel vor allem die Zergliederung der Sprachlaute selbst und der zu ihrer Perzeption führenden Vorgänge; wobei wir die Analyse erst dann als vollständig ansehen dürfen, wenn es gelingt, die Laute auch wieder aus ihren letzten Elementen oder Faktoren naturgetreu zusammensetzen.

Daß der Verfasser hierbei zunächst, ja fast ausschließlich, die Laute seiner deutschen Muttersprache berücksichtigt, wird man begreiflich finden. Mit dem Nächstliegenden und Bestbekanntesten muß jeder beginnen. Aber die Methoden, die hier angewandt, und die Strukturgesetzmäßigkeiten, die hier gefunden werden, lassen sich selbstverständlich auch auf andere Sprachen übertragen, und wir werden im einzelnen der Vergleichung halber dahinzielende Versuche und Hinweisungen einschalten.

## 2. Antinomien der Vokalforschung.

Die Vokale als lautlicher Grundstock der Rede ziehen auch die Aufmerksamkeit des Theoretikers zunächst auf sich. Um sie drehen sich vorzugsweise die Anstrengungen und Streitigkeiten der akustisch gerichteten Phonetiker seit HELMHOLTZ. Ein flüchtiger Überblick der Hauptstreitpunkte und Gegensätze, um die es sich dabei handelt, sei hier vorausgeschickt.

HELMHOLTZ betrachtete die Vokale als besondere Fälle von Klangfarben neben denen der Instrumente. Das Gemeinsame fand er in dem Aufbau aus Teiltönen, das Unterscheidende darin, daß eine instrumentale Klangfarbe gegeben sei durch ein bestimmtes Stärkeverhältnis von Teiltönen bestimmter Ordnungszahl, einerlei, welche absolute Höhe diese Teiltöne besitzen, eine Vokalfarbe dagegen durch die Anwesenheit und überwiegende Stärke eines oder zweier charakteristischer Teiltöne von bestimmter absoluter Höhe für jeden Vokal. Diese charakteristischen Töne müssen aber nach HELMHOLTZ als harmonische Teiltöne in dem Vokalklang enthalten sein, d. h. ihre Schwingungszahlen müssen ganzzahlige Vielfache von der Schwingungszahl des Grundtones sein, auf dem gesungen oder gesprochen wird. Sie sind in folgendem Schema, nach ihrer Tonhöhe geordnet, zusammengestellt:



HELMHOLTZ kam zu diesen Aufstellungen hauptsächlich durch die Untersuchung der Resonanzeinstellung der Mundhöhle. Gegenüber dem naheliegenden Einwand, daß doch nicht jeder Grundton alle diese Teiltöne als Multipla mit sich führe, während man alle

Vokale auf allen Tönen innerhalb des Stimmumfanges singen könne, antwortete er, daß man einen bestimmten Vokal tatsächlich nicht gleich gut auf jeder Tonhöhe singen könne, sondern am besten in der Nähe des charakteristischen Tones oder auf seiner tieferen Oktave, die diesen Ton stark enthalte.

Aber die Einwendungen wiederholten und mehrten sich. LUDIMAR HERMANN stellte neue ausgedehnte Untersuchungen an, für die er die photophonographische Methode mit glänzendem Erfolg einführte. Er photographierte zunächst die Schwingungen einer Membran, gegen die ein Vokal gesungen wurde, später die winzigen, in die Phonographenwalze eingegrabenen Vertiefungen, beide auf ein Spiegelchen übertragen, das die Bewegungen vergrößert auf die photographische Platte warf. So erhielt er Kurven, aus denen mit Hilfe der FOURIER-Analyse oder durch bloße Auszählung der Zacken ihre Bestandteile berechnet wurden. Er konnte dadurch nicht bloß den stärksten Teilton, sondern eine große Anzahl von Teiltönen ihrer relativen Stärke nach bestimmen und so die ganze Struktur der untersuchten Vokale in einer Tabelle ihrer Teiltonstärken klarlegen. Aber die Ergebnisse führten ihn zu ganz anderen Vorstellungen als HELMHOLTZ. Der charakteristische Ton — für den er den bequemen Ausdruck „Formant“ einführte — braucht nach ihm keineswegs ein harmonischer Teilton zu sein und entsteht nicht durch die Resonanzverstärkung eines solchen in der Mundhöhle. Diese wird vielmehr durch den im Rhythmus der Grundschwingungen intermittierenden Luftstrom aus dem Kehlkopf wie eine Art Pfeife angeblasen. Doch stimmte HERMANN seinem Vorgänger darin bei, daß der Formant für jeden Vokal eine feste absolute Lage in der Tonreihe habe, einerlei, auf welchem Grundton zufällig gesungen werde. Auch sind die Abweichungen der charakteristischen Töne in beiden Theorien nicht allzu groß<sup>1)</sup>.

HERMANN'S Vokallehre stieß aber gleichfalls auf Widerspruch. Die Vorstellung, daß die Mundhöhle beim Gesang angeblasen würde, konnte sich gegen die ältere, daß sie nur als Resonator wirke, nicht durchsetzen, und sein Begriff des Formanten, wonach dieser nicht bloß kein harmonischer, sondern überhaupt kein Teilton im gewöhnlichen Sinne sein sollte und auf keine Weise

<sup>1)</sup> Da die Ergebnisse meiner Untersuchungen in prinzipiellen Punkten der HERMANN'Schen Theorie scharf widersprechen, möchte ich vorab meiner Bewunderung für seine unermüdliche und musterhaft sorgfältige Experimentalarbeit Ausdruck geben. Er hat die graphische Methodik, die mit dem SCOTT-KOENIGSchen Phonautographen (1859) ihren Anfang nahm, um ein gewaltiges Stück weitergebracht. Die Originalkurven, die ich 1904 im Königsberger Physiologischen Institut sehen durfte, sind noch weit schöner als die veröffentlichten Abbildungen.

isoliert oder verstärkt zu Gehör gebracht werden könnte, wurde im Laufe seiner Darstellungen und Er widerungen immer unfaßlicher.

Verwandte Methoden sind seitdem zumeist benutzt worden, sei es, daß man sich der HERMANNschen Technik oder des HENSENschen „Sprachzeichners“ oder des Oszillographen und ähnlicher Einrichtungen bediente. Es wurde auch versucht, die Luftschwingungen ohne dazwischentretende Membran, bei der immer die Eigentöne gefährlich sind, zu photographieren (RAPS, GEHLHOFF). Sehr feine Bilder erhielt STRUYCKEN durch ein leichtbeweglich aufgehängtes, von den Schallschwingungen gedrehtes Spiegelchen. W. KOEHLER hat sogar durch ein an seinem eigenen Trommelfell befestigtes Spiegelchen gute Kurven erhalten, die den großen Vorzug haben, daß sie der am Hören unmittelbar beteiligten Membran entstammen. Andere vergrößerten auf mechanischem Wege die durch den Stift des Grammophons in die Wachsmasse eingezeichneten Kurven (SCRIPTURE, LIORET) oder unterzogen diese einer mikroskopischen Ausmessung (BOEKE, STEVANI). Aber die Ergebnisse zeigten zunächst keineswegs die erwünschte Übereinstimmung, weder im einzelnen, noch in den prinzipiellen Auffassungen, zu denen sie führten.

In letzterer Hinsicht fand HELMHOLTZ einen ausgezeichneten Verteidiger in PIPPING. Auf HERMANNs Seite traten (wenigstens im Prinzip) SAMOJLOFF, NAGEL, SCRIPTURE, GARTEN u. a.

Was aber die Bestimmung der Formanthöhen im einzelnen betrifft, so möge eine Übersicht der für U und A von verschiedenen Forschern angegebenen charakteristischen Töne das weite Auseinanderklaffen der Angaben veranschaulichen<sup>1)</sup>.

Die mit × bezeichneten Angaben stammen aus graphischen Methoden. Die durch Striche verbundenen Noten bedeuten, daß die Versuchsergebnisse innerhalb dieser Strecken variieren.

Der Anblick dieser untereinander so abweichenden Bestimmungen (die Werte für U umfassen fast 3, die für A  $1\frac{1}{2}$  Oktaven, und jeder Leiterton in diesen weiten Räumen ist vertreten) kann einen entmutigenden Eindruck machen. Aber es soll damit nicht allen diesen Untersuchungen gleicher Wert oder Unwert beigemessen werden. In wissenschaftlichen Dingen gibt es keine Volksabstimmung. Nur das Bedürfnis einer neuen Untersuchung auf neuen Grundlagen sollte durch diese Übersichten gerechtfertigt werden.

<sup>1)</sup> Die Tabellen gründen sich hauptsächlich auf die Zusammenstellungen bei NAGEL, GUTZMANN, ZWAARDEMAKER, JESPERSEN, VICTOR. Die neuesten Angaben sind darin noch nicht enthalten.

U: 1 HELMHOLTZ, 2 KOENIG, 3 PIPPING, 4 GUTZMANN, 5 AUERBACH (Perkussion), 6 DONDERS, 7 ABRAHAM, 8 AUERBACH (Resonanz), 9 HERMANN 1890, 10 BOEKE, 11 SAMOJLOFF, 12 HERMANN 1905, 13 RAPS, 14 DELSAUX, 15 STEVANI, 16 VERSCHUUR.

A: 1 DONDERS, 2 HERMANN 1890, 3 AUERBACH (Perkussion), 4 HERMANN 1905, 5 AUERBACH (Resonanz), 6 RAPS, 7 SAMOJLOFF, 8 GUTZMANN, 9 HELMHOLTZ, 10 KOENIG, 11 VERSCHUUR, 12 STEVANI, 13 ABRAHAM, 14 BOEKE, 15 PIPPING, 16 BEVIER.

Manche wurden durch die Divergenz der Ergebnisse sogar wieder an dem Grundgedanken irre, den HERMANN mit HELMHOLTZ teilte, daß nämlich der „charakteristische Ton“ überhaupt eine feste Lage besitze, und kehrten zu der Relativtheorie zurück, wonach Teiltöne von bestimmter Ordnungszahl maßgebend wären, so daß z. B. beim U jedesmal der 1., beim O der 2., beim A der 3. Teilton am stärksten wäre, einerlei, auf welche Tonhöhe sie fallen. Es wäre also nicht wesentlich anders als nach HELMHOLTZ bei den Klangfarben der Instrumente. (So H. GRASSMANN, LLOYD, zum Teil auch AUERBACH.)

Wir werden aber nach der positiven Darstellung unserer Ergebnisse in einem kritischen Rückblick auf die früheren (8. Kap.) zeigen, daß die Abweichungen zum großen Teil nur scheinbar sind, daß die methodisch einwandfreien Versuche tatsächlich in allem Wesentlichen übereinstimmen, und daß die neueren, unter denen auch die radiophonischen eine zunehmende Rolle spielen, immer weitere Bestätigungen liefern.

Die Einwände gegen die graphische Methode richteten sich, abgesehen von den Schwierigkeiten ihrer exakten Durchführung (Eigentönen der Membranen usf.), hauptsächlich gegen die Schlußkraft der FOURIER-Analyse. Vor allem erschien es unbegreiflich, daß dabei der Grundton fast regelmäßig sehr schwach oder gar nicht herauskam, während wir ihn doch so stark zu hören glauben, daß er die Höhe des ganzen Klanges bestimmt. Es wird sich zeigen, daß gerade dieser Einwand (den besonders SCRIPTURE in den Vordergrund stellt) in eine Rechtfertigung umschlägt. Aber

zwei allgemeinere Gründe lassen es erwünscht erscheinen, zunächst noch andere und direktere Methoden zu versuchen:

1. liefert die FOURIER-Analyse ihrer Natur nach nur harmonische Teiltöne, während man auf unharmonische nur etwa (mit ziemlicher Unsicherheit) aus der gleichmäßigen Verstärkung benachbarter harmonischer Teiltöne schließen kann. Man wird daher über diesen wichtigen Punkt, der auch mit den organo-genetischen Streitfragen zusammenhängt, auf diesem Wege nicht zur Entscheidung kommen.

2. kann man dadurch nur die physikalische Stärke, d. h. die Amplitude der Luftschwingungen und ihre kinetische Energie (Quadrat der Amplitude  $\times$  Quadrat der Schwingungszahl) messen, nicht die physiologische, wie sie im Gehörorgan und zuletzt in der Hirnrinde vorhanden ist, geschweige denn die psychologische (phänomenale), wie sie sich in der Empfindung selbst darstellt. Die letzteren Intensitäten dürfen aber keineswegs ohne weiteres den physikalischen proportional gesetzt werden. Dies ist ein Punkt, auf den bis zur neuesten Zeit in der Phonetik fast niemals geachtet wurde.

Als den leitenden Grundgedanken der vorliegenden Arbeit nach der analytischen Seite hin möchte ich den bezeichnen: den Klang so direkt als möglich, unter Einschaltung möglichst weniger Zwischenglieder, zu untersuchen. Daß allzu kostspielige und komplizierte Apparate nicht zu Gebote standen, hat zu diesem Bestreben mit beigetragen. Um so mehr aber waren der Verfasser und seine Mitarbeiter auf die anhaltendste Übung im direkten Beobachten des Klangmaterials, seiner Unterschiede und Veränderungen angewiesen. Zahlreiche jüngere Kräfte haben mir geholfen, denen ich auch hier den herzlichsten Dank ausspreche. Besonders bei den mühsamen Interferenzversuchen wirkten sie sämtlich abwechselnd mit, je nachdem es jedem der Kriegsdienst gestattete. Es waren die Herren Dr. ABRAHAM, Dr. v. ALLESCH, Dr. BLUMENFELD, Dr. FRIEDLÄNDER, Dr. A. GUTTMANN, Prof. Dr. v. HORNPOSTEL, Dr. KOFFKA, A. KREICHGAUER, Dr. LEWIN, Dr. RIEFFERT, Prof. Dr. RUPP. Dr. v. ALLESCH hat nicht nur bei den Beobachtungen, sondern auch bei der äußeren Einrichtung der Interferenzversuche und dem Aufbau des synthetischen Apparats unermüdlich geholfen.

Die „subjektive“ Methode, bei der die direkte akustische Beobachtung die Hauptrolle spielt, gilt vielen als weniger vertrauenswürdig gegenüber der „objektiven“, bei der die Struktur eines Sprachlautes in Kurvengestalt vor Augen liegt. Ohne im geringsten die Wichtigkeit dieser graphischen Hilfsmittel zu ver-



kennen, auf deren Leistungen wir des öfteren zur Ergänzung oder Bestätigung zurückgreifen werden, müssen wir doch nachdrücklich betonen: 1. daß Vokale als solche überhaupt etwas Subjektives sind und wir vom Standpunkte des Gesichtssinnes niemals Anlaß gehabt hätten, sog. Vokalkurven von anderen zu unterscheiden, 2. daß in akustischen Dingen das Ohr erster und letzter Richter sein muß, 3. daß die „objektive“ Methode, je feiner sie durchgebildet wird, um so kompliziertere äußere Einrichtungen erfordert, daß aber jedes Glied einer zwischengeschalteten äußeren Einrichtung eine besondere Fehlerquelle mit sich bringt. Dies ist Physikern und Physiologen selbst am besten bekannt, und viel Scharfsinn und Arbeit ist auf die Ausschaltung dieser Fehlerquellen verwandt worden. Gleichwohl stellt sich das Vorhandensein einer solchen nicht selten erst heraus, wenn die Untersuchung vorbei ist. Den Nachteil der subjektiven Methode, daß man auf das gute Ohr und die Übung des individuellen Beobachters angewiesen ist, haben wir durch die erwähnte Vervielfältigung der Beobachter und andere Kontrollmaßregeln auszugleichen gesucht. Auch werden die äußeren Umstände, unter denen die Beobachtungen gemacht wurden, und ihre Einzelheiten im folgenden so genau angegeben, daß sie jederzeit von anderen nachgeprüft werden können. Im übrigen haben auch GRASSMANN, DONDERS, HELMHOLTZ, AUERBACH und andere bei der Aufsuchung der Resonanzmaxima, der Flüsterhöhen usw. den direkten Weg eingeschlagen. Endlich ist überhaupt der Unterschied von objektiver und subjektiver Methode vielfach ein fließender. Denn zur Durchführung erfolgversprechender und unter gleichen Umständen wiederholbarer direkter Beobachtungen sind doch auch äußere Einrichtungen unentbehrlich, und wir werden namentlich bei der Interferenzmethode auf die Fehlerquellen dieser Einrichtungen einzugehen haben. Ob man also das Auge oder das Ohr benütze: der Vorsicht und Geduld bedarf es da wie dort.

## 1. Kapitel.

# Analyse gesungener Vokale durch resonierende Gabeln.

Von Analyse und von einer Struktur der Vokale wird hier zunächst in dem hergebrachten Sinn aller bisherigen Untersuchungen gesprochen, d. h. im Hinblick auf die dem Ohre zugeführten Schwingungen. Diese selbst nennen wir im Anschluß an Ohm einfach, wenn sie als pendelförmige oder Sinusschwingungen erfolgen, zusammengesetzt aber, wenn ihre Schwingungsform sich mathematisch und physikalisch in eine Anzahl superponierter Sinusschwingungen zerlegen läßt. Erst im Laufe der Darstellung werden wir auch die Fragen nach der physiologischen und psychologischen Struktur der Vokale aufwerfen und den Sinn und die Berechtigung dieser Begriffe erläutern<sup>1)</sup>.

## I. Methode.

### 1. Wesen und Vorzüge der Gabelmethode.

HELMHOLTZ hatte die Resonanz der Mundhöhle bei den verschiedenen Vokalen durch die Verstärkung festgestellt, die eine tönende Stimmgabel erfährt, wenn sie vor die zum Aussprechen des Vokals geöffnete Mundhöhle gehalten wird (Verstärkungsmethode, auch von R. KOENIG angewandt). Er hatte ferner die Verstärkung von Teiltönen während des Singens durch ans Ohr gehaltene Resonatoren ermittelt (Resonatorenmethode, auch von AUERBACH benützt). Er hatte endlich das Nachklingen bestimmter Klaviersaiten bei aufgehobener Dämpfung infolge eines hineingesungenen Vokals beobachtet (Klaviermethode). Andere suchten durch Anblasen der zum Singen oder Sprechen geöffneten Mundhöhle, durch Beklopfen verschiedener Teile des Schädels oder des Kehlkopfs die charakteristischen Töne zu ermitteln (Anblase- und Klopfmethoden: AUERBACH, ABRAHAM).

Merkwürdigerweise ist aber bisher ein Resonanzverfahren fast gar nicht angewandt worden, das, systematisch durchgeführt,

<sup>1)</sup> Siehe S. 12, ferner Kap. 11, 13, 14.

ausgiebigere und eindeutige Ergebnisse als alle diese anderen liefert: die Feststellung der in einem Klang enthaltenen Teiltöne und ihrer Stärkeverhältnisse durch mit- und nachschwingende Stimmgabeln (Resonanzmethode i. e. S. oder Gabelmethode). Auf dieses Verfahren wurde ich vor langer Zeit durch den verstorbenen Professor der Gesangskunst GUSTAV ENGEL aufmerksam gemacht, der sich 1888 zur Untersuchung der Obertöne der menschlichen Stimme von R. KOENIG eine schöne Serie von Gabeln auf Resonanzkästen hatte anfertigen lassen und mir damals sogleich mitteilte, daß es ihm gelungen sei, damit den 24. Teilton seiner Stimme, nämlich  $c^4$  beim Singen auf  $F$  (der Vokal war nicht erwähnt) festzustellen. Ich habe dann 1896 das Verfahren zur Bestimmung der Obertöne überhaupt empfohlen und angewandt<sup>1)</sup>.

Vielfach ist nun zwar früher behauptet worden (so von WHEATSTONE, später u. a. von WUNDT, der es durch besondere Versuche bestätigt haben wollte, von LLOYD und noch 1913 von MUSEHOLD<sup>2)</sup>), daß ein einfacher Ton höhere Gabeln erregen könnte, ohne daß entsprechende hohe Teiltöne im Klange vorhanden zu sein brauchten („multiple Resonanz“). Aber die Unrichtigkeit dieser Lehre ist von mir (3) durch besondere Versuche erwiesen: wurden die betreffenden Obertöne durch Interferenzröhren ausgeschlossen, so blieb auch die Resonanzwirkung aus<sup>3)</sup>. Man kann daher aus dem

<sup>1)</sup> Nur STEFANINI (Physiker in Pisa) hat seit 1911 wiederholt damit experimentiert, aber die Methode wegen zu geringen Gabelvorrates und anderer Bedenken (s. u.) wieder aufgegeben.

<sup>2)</sup> Selbst ein so hervorragender Akustiker wie F. AUERBACH scheint sich hiervon nicht freigemacht zu haben. Wenigstens kann ich seine Darstellung in 4, S. 125 und in „Grundlagen der Musik“ 1911, S. 80 nicht anders verstehen. HELMHOLTZ dagegen bemerkt ausdrücklich (S. 76): „In diesen Fällen (wenn ein Resonator auch auf tiefere Klänge anspricht, zu deren harmonischen Obertönen sein Eigenton gehört) tönt der Resonator durch einen der harmonischen Obertöne des im äußeren Luft-raum angegebenen Klanges.“ Er lehrt also ausschließlich unisono Resonanz. Gegen LLOYD hat PIPPING nachdrücklich die richtige Ansicht vertreten.

Über multiple Resonanz in einem ganz anderen, einwandfreien Sinne s. 14. Kap.

<sup>3)</sup> Theoretisch wurde das nämliche kurz darauf (Ann. d. Phys. N. F., Bd. 59, 1896) durch JOHANNESSEN abgeleitet. Auch BOSANQUET hat schon 1881 gegen R. KOENIG die mathematische Unmöglichkeit geltend gemacht (Philosophical Magazine Bd. 12, S. 280). Die beliebte Analogie eines Pendels, das bei jedem 2. oder 3. Hin- und Hergang angestoßen wird, läßt sich nicht auf die Einwirkung einer in stetigem Phasenwechsel schwingenden Luftmasse anwenden.

An Streichinstrumenten hat F. RITZ (1883) beobachtet, daß jeder Oberton einer Saite unisono Obertöne anderer Saiten in Mitschwingung versetzt, niemals aber deren Multipla.

Mit- und Nachschwingen einer Gabel durchaus eindeutig auf das Vorhandensein eines entsprechenden Teiltones schließen.

Gute Gabeln auf Resonanzkästen können durch bloßes Mitklingen auf einen kräftigen Oberton unter Umständen so stark erregt werden, als würden sie mit einem Bogen gestrichen. Wenn ich auf dem Tone *gis* ein O singe, schwingt die Gabel *gis*<sup>1</sup> (2. Teilton) noch in 2 Metern Entfernung merklich mit. Wenn ich auf *c*<sup>1</sup> ein A singe (allerdings aus voller Lunge), schwingt die Gabel *g*<sup>2</sup> (3. Teilton) sogar noch in 6 m Entfernung. In beiden Fällen handelt sich's eben um Formanttöne der bezüglichen Vokale.

Diese Methode hat den großen Vorzug vor den meisten anderen Resonanzmethoden, daß die Gabeln nachschwingen und man also den einzelnen Teilton bequem und rein für sich beobachten, auch seine verschiedene Stärke bei verschiedenen Vokalen vergleichen kann<sup>1</sup>). Das Verfahren ist unabhängig von der Fähigkeit eines Beobachters zum Heraushören, da die Analyse von den Gabeln selbst besorgt wird. Insofern kann man die Gabelmethode geradezu als eine objektive bezeichnen. Nur die Schätzung der Stärke des Mitschwingens (s. u.) bleibt subjektiv. Bei der Klaviermethode ist das zwar ähnlich, aber es schwingen doch immer, wenn nicht besondere Maßregeln ergriffen werden, mehrere Saiten auf einmal nach und sind noch Geräusche beigemischt. Auch schwingen die Saiten der besonders wichtigen 3- und 4-gestrichenen Oktaven nicht mehr deutlich mit. Bei den übrigen Resonanzmethoden aber erhält man, wenn es gut geht, überhaupt nur einen einzigen Ton, kein Bild der ganzen Klangstruktur. Außerdem ist es bekanntlich schwer, die Mundstellung nach dem Aussprechen oder Singen eines Vokals noch eine Weile ganz unverändert zu halten; es treten leicht unwillkürliche, ja unbemerkte Veränderungen ein, die den Punkt stärkster Resonanz verschieben. Bei der Verstärkungs- und der Resonatorenmethode besteht noch der Übelstand, daß eine Klangquelle auch vor jedem Hohlraum verstärkt wird, der ihren Eigenton als Multiplum des seinigen enthält, indem seine Luftmasse dann eben in die entsprechenden Teilschwingungen gerät, besonders wenn es sich um ein ungerades Multiplum handelt. An der „Flaschenorgel“ des Berliner Psychologischen Instituts, die aus kugelförmigen Resonatoren besteht,

<sup>1</sup>) Bei einer Einrichtung, die EDELMANN in München auf Anregung der Frau WEBER-BELL ausgeführt hat, ist das Nachklingen auf elektrischem Wege noch verlängert. Ich habe diese Einrichtung 1912 kennengelernt, nachdem ich längst die Gabelmethode zur Klanganalyse verwendet hatte, habe sie aber nicht benutzt, weil die Zahl der dadurch zu ermittelnden Obertöne für wissenschaftliche Zwecke viel zu gering ist.

läßt sich dies mit davorgehaltenen losen Gabeln leicht beobachten. Liegt der Eigenton des Resonators tief unter dem der Gabel, so wird man hier überhaupt immer eine, wenn auch geringe, Verstärkung bemerken, weil der Gabelton dann immer einem Multipulum entspricht oder naheliegt<sup>1)</sup>. Man kann bei diesem Verfahren also nicht ohne weiteres aus einer Verstärkung auf einen bestimmten Teilton des Vokalklanges schließen, sondern muß erst sorgfältig die Lage der maximalen Verstärkung aufsuchen, was in Fällen, wie den letztgenannten, keineswegs leicht ist.

Eine gewisse Resonanzbreite gibt es freilich auch bei den mitschwingenden Resonanzgabeln. Sie ist nach Umständen verschieden, beträgt aber in der Regel nicht mehr als einen Halbton nach oben und unten<sup>2)</sup>. Immerhin muß dieser Umstand berücksichtigt werden, wenn man nicht zu Fehlschlüssen kommen soll. Es muß also in Zweifelsfällen versucht werden, ob nicht durch eine kleine Erhöhung oder Erniedrigung des gesungenen Tones die Gabel noch stärker zum Mitschwingen gebracht werden kann.

Legt man die HELMHOLTZsche Theorie des Hörens zugrunde — und sie gibt, wenn nicht den realen Tatbestand selbst, doch jedenfalls das bisher beste Bild desselben —, so ist diese unsere Methode der Klangzerlegung keine andere als die des Ohres selbst. In der Tat fallen die Teiltöne, die wir dadurch finden, mit denen zusammen, die ein geübtes Ohr durch rein subjektive Analyse findet. Bald reicht das Ohr weiter, bald die Gabeln. Bei einer konstant tönenden Metallzunge von 50 Schwingungen pro Sekunde konnte

<sup>1)</sup> Es kommt aber auch vor, daß eine höhere Gabel, vor den Resonanzkasten einer tieferen gehalten, geschwächt wird, offenbar durch Interferenzwirkungen. Hält man die lose Gabel  $c^3$  vor den Kasten der Resonanzgabel  $c^2$ , so wird sie an gewissen Stellen verstärkt, an anderen geschwächt. Ich hörte (mit Dr. BAILEY) eine  $as^1$ -Gabel bei O-Einstellung des Mundes direkt vor der Mundöffnung schwächer, während sie, als im O-Formant liegend, hätte verstärkt klingen müssen. Diese dem Berliner Physikalischen Institut gehörige Gabel war von RUD. KOENIG ausdrücklich durch ein eingepprägtes „O“ als diesem Vokal entsprechend bezeichnet worden.

<sup>2)</sup> So bei Erregung durch eine kräftige menschliche Stimme aus unmittelbarer Nähe. Wenn der erregende Klang selbst eine Stimmgabel oder eine Zunge des Tonmessers ist, so erstreckt sich das Mitschwingen nur auf wenige Schwingungen auf- und abwärts; weshalb gewöhnlich Stimmgabeln, auch solche auf Resonanzkästen, von Physikern (HELMHOLTZ, AUERBACH, WAETZMANN) als Beispiel einer starken, aber engbegrenzten Resonanz angeführt werden. Aber die menschliche Stimme ist eben eine viel kräftigere Klangquelle und darum von breiterer Resonanzwirkung. Wenn z. B. eine Männerstimme den Vokal O auf dem Tone *gis* singt, so ist die Gabel  $a^1$  damit noch recht merklich in Schwingung zu setzen, obgleich sie um einen Halbton von dem in dem Stimmklang enthaltenen *gis*<sup>1</sup> abweicht.

v. HORNPOSTEL noch den 56. Teilton heraushören, der Verfasser noch den 48.<sup>1)</sup>). Bei den Vokalen, die vom Sänger nicht so lang ausgehalten werden können, kommt man wieder mit den Gabeln weiter.

Für die Beweisführung haben die Gabeln den unschätzbaren Vorzug vor dem Ohre, daß ihr Mit- und Nachklingen von jedem, auch dem Ungeübtesten, vernommen wird, während das direkte Heraushören subjektiv bleibt und nicht demonstriert werden kann, überdies von Übung und Anlage im höchsten Maße abhängig ist.

Setzen wir voraus, daß der Bestand des Klanges, wie er in der Schnecke des Ohres existiert, auf diesem Wege festgestellt werde, so dürfen wir wohl auch weiter annehmen, daß in den letzten oder vorletzten Prozessen in der Hörspäre der Hirnrinde eine ähnliche Zusammensetzung vorliege. Ob wir auch die Klangerscheinung selbst, wie sie in der Empfindung gegeben ist, aus ebenso vielen Teilen zusammengesetzt zu denken haben, mag hier dahingestellt bleiben. Jedenfalls wird man ihr in dem Sinne Teile zuschreiben müssen, daß es prinzipiell möglich ist, durch geschärfte Aufmerksamkeit unter günstigen Bedingungen die Teiltöne, einen nach dem anderen, aus dem Ganzen herauszuhören.

Diese Bemerkungen sollten nur die mögliche Tragweite der Methode erläutern. Was aber daraus zunächst erschlossen werden soll und kann, ist nur der objektive Klangbestand, und dieser Schluß ist von jeder Hypothese unabhängig.

## 2. Stärkeschätzungen.

Wichtig ist nun weiter die Abschätzung der Stärke des nachschwingenden Tones. In dieser Hinsicht hielt ich mich unter Verzicht auf physikalische Bestimmungen, die doch nur die Stärke der äußeren Reize messen, an die Empfindungsstärken selbst. Freilich gestatten diese keine mathematischen Bestimmungen. Niemals können wir auf Grund direkter Vergleichung behaupten, eine Tonempfindung sei 2- oder 3 mal so stark als eine andere. Aber wir können auf der stetigen Linie der Stärkegrade gewisse Typen

<sup>1)</sup> Man hört nicht alle Teiltöne bis zu diesen Grenzen, sondern unter den höheren, wie es scheint, nur die Oktaven der ersten 9 oder 10. Die Gründe und der Hergang dieses auswählenden Verhaltens wären noch zu untersuchen. Wahrscheinlich spielt das Entgegenkommen der Erfahrungsvorstellungen, die man sich von den tieferen Teiltönen auf Grund der bekannten musikalischen Intervalle bilden kann, dabei eine Rolle. Denn solche entgegenkommenden Vorstellungen sind für das Heraushören von Teiltönen überhaupt von Bedeutung. Nur der Mechanismus wäre zu finden, durch den mit der Vorstellung eines tieferen Teiltones zugleich dessen höhere Oktaven subjektiv verstärkt werden.

oder Kategorien unterscheiden, unter die wir einen gegebenen Schall einreihen. Ich benutzte 5 solcher Typen, die den musikalischen Unterscheidungen des *pp*, *p*, *f*, *ff*, *fff* entsprachen und durch die Ziffern 1 bis 5 bezeichnet wurden<sup>1)</sup>. Bald sah ich mich aber auch zu Zwischenstufen geführt. Namentlich mußten in den unteren Graden noch Halbe und Viertel, vor der Nullgrenze sogar eine Achtelstufe unterschieden werden.

Zu ähnlichen subjektiven Abschätzungen von Stärkeklassen findet man sich nicht bloß im Leben, sondern auch in der Wissenschaft oft genug veranlaßt. Bekannt sind die Sternklassen der Astronomen, die (ursprünglich wenigstens) auf Helligkeitsschätzungen beruhen. Zu diesen sind in neuerer Zeit die Intensitätsschätzungen der feinen Linien auf astronomischen Photogrammen getreten, bei denen gelegentlich 10 Klassen unterschieden werden. In der Farbenlehre haben PLATEAU, DELBOEUF, CHEVREUL, HERING, HILLEBRAND, EBBINGHAUS solche rein subjektiven Skalen angewandt. In der Meteorologie ordnet man die Bewölkung und die Windstärken in Klassen. Für die letzteren stellte BEAUFORT eine 12-stufige, von eben merklicher Bewegung bis zum Orkan reichende Skala auf, in welche geübte Segelfahrer eine gegebene Stärke mit hinreichender Bestimmtheit einordnen. Ebenso ist es mit den musikalischen Tempi, wie sie durch die gebräuchlichen 5—6 Hauptausdrücke bezeichnet sind. Die Tonstärken selbst ordnete HELMHOLTZ gerade anläßlich des Mitschwingens in 10 Stufen nach bloß subjektivem Gehörseindruck<sup>2)</sup>. AUERBACH (1), BOSANQUET, ROUSSELOT operierten gleichfalls mit solchen. DWELSHAUVERS ordnete die Erregung KOENIGScher Flammen durch Resonatoren, v. KEMPELEN die Weite der Lippenöffnung und des Raumes zwischen Zunge und Gaumen beim Sprechen in 5 Stufen. Und so wird man uns nicht ein wissenschaftlich fragwürdiges Verfahren vorwerfen. Schließlich ist man ja bekanntermaßen selbst bei genauesten wirklichen Messungen mit eingeteilten Maßstäben zuletzt vielfach auf bloße Schätzung der noch zwischen zwei Teilstrichen liegenden Dezimalen angewiesen.

<sup>1)</sup> In meiner Abhandlung 10 sind die Ziffern 1 bis 4 = *po*, *mf*, *fo*, *ffo* gesetzt. Aber die obige Gleichsetzung scheint mir jetzt den Eindruck noch besser wiederzugeben. *mf* ist dann eben =  $2^{1/2}$ .

<sup>2)</sup> S. 203. HELMHOLTZ brauchte diese Schätzungen, um den Phasenunterschied eines verstimmtten Resonators gegenüber der Klangquelle da; nach zu bestimmen. Die dazu dienende Tabelle ist reines Rechnungsprodukt, aber die Einordnung einer beobachteten Stärkeabnahme in die 10 Stufen erfolgte nach subjektiver Schätzung: „Wenn man das Verhältnis, in welchem diese Stärke abgenommen hat, ungefähr zu beurteilen weiß, findet man aus der Tafel den Phasenunterschied.“

Unsere Klassifikation unterscheidet sich aber von vielen anderen dadurch, daß sie nicht darauf ausgeht, eine Skala von gleich weit abstehenden Graden herzustellen. Die Helligkeitsunterschiede aufeinanderfolgender Sternklassen sind schätzungsweise gleich groß (FECHNER). Von unseren Stärkeunterschieden soll dies aber nicht behauptet werden. Vielmehr würde ich sagen, daß die Abstände zwischen 2 benachbarten Stufen von der untersten zur obersten hin immer kleiner werden. Ebendarum wurden bei den untersten Klassen auch noch vielfach Zwischenstufen eingeschaltet, namentlich gegen die Nullgrenze hin, während von 1 ab fast immer nur eine Zwischenstufe genügt; und zwar kamen solche zwischen 1 und 2, 2 und 3 noch häufig vor, zwischen den höheren Graden aber immer seltener. Unsere Zahlen sind daher gewissermaßen nur Ordnungszahlen. Die Stärke 4 soll nicht das Doppelte der Stärke 2, der Abstand von 4 zu 2 nicht das Doppelte des Abstandes von 2 zu 1 bedeuten.

Die Anwendung dieser Kategorien erfolgte ebenso wie die der entsprechenden musikalischen (pp, p usw.) nach sozusagen impressionistischen Gesichtspunkten. Die Intensitäten unserer Sinnesempfindungen bilden zwar in sich selbst eine durchaus homogene stetige Reihe, sind aber an bestimmten absoluten Punkten ihres Verlaufes mit charakteristischen Wirkungen auf das Bewußtsein und das Nervensystem verknüpft, durch die diese Punkte sich auszeichnen und wiedererkannt werden. pp oder 1 ist ein sehr schwacher Eindruck, der aber bei äußerer Ruhe schon gut und ohne Anstrengung der Aufmerksamkeit hörbar ist. p oder 2 eine Stärke, die erheblich darüber, aber noch im Gebiet des „Leisen“ liegt, mf oder  $2\frac{1}{2}$  ist schon kräftig, aber in keiner Weise angreifend. Hier liegt die Grenze der beiden großen Stärkezonen „Leise“ und „Laut“, in die wir die Schälle ordnen<sup>1)</sup>. Dagegen f oder 3 ist schon der erste Anfang des „Gellens“, ein Stärkegrad, der auf die Dauer dem Ohre doch lästig fallen, es angreifen würde. 4 und 5 endlich bedeuten weitere Schritte auf dieser Bahn, die sprachlich als „sehr stark“ bzw. „enorm stark“ auszudrücken wären. Zuweilen notierte ich sogar 6, d. h. extrem stark. Unterhalb der Stärke 1 bedeutet  $\frac{1}{8}$  einen eben (aber unzweifelhaft) merklichen,  $\frac{1}{4}$  einen eben übermerklichen Ton,  $\frac{1}{2}$  und  $\frac{3}{4}$  weitere Stufen zu dem gewöhnlichen pp hin.

<sup>1)</sup> BOSANQUET rechnete 5 seiner 10 Grade zu den „lauten“, 5 zu den „leisen“.

Unter den Tempi bildet das Andante, das Tempo des gemächlichen Gehens, die neutrale Mitte zwischen den schnellen und langsamen. Solche subjektive Gradbestimmungen scheinen überall absolute Mitten zwischen entgegengesetzten Polen einzuschließen.



Man erwirbt sich bald Übung in der Anwendung dieser Bezeichnungen, wenn man auch natürlich oft zwischen den feineren benachbarten Nuancen schwanken kann. Nicht leugnen möchte ich, daß bei den tiefsten und höchsten der angewandten Gabeln, die nicht so empfindlich waren wie die übrigen, die obigen Kriterien etwas verschoben wurden und z. B. mit „3“ schon ein Stadium bezeichnet wurde, das für das Ohr noch nicht gerade bedrohlich erschien. Da wurde eben die mangelhafte Resonanzfähigkeit mit in Rechnung gezogen.

Auch in anderen Sinnesgebieten sind übrigens Gradbestimmungen, die nicht in erster Linie nach Distanzgleichheit, sondern nach charakteristischen Bewußtseinswirkungen abgestuft werden, längst gebräuchlich. So unterscheiden wir unter den Temperaturen: „überschlagen (leicht angewärmt), lau, warm, heiß, glühend-heiß“, und könnten auch da die 5 ersten Ordnungszahlen einsetzen. Hier liegt wieder zwischen dem 2. und 3. Grade die behagliche Mitte. Vermutlich sind auch bei manchen der oben angeführten Einteilungen in ähnlicher Weise mehr qualitative Merkmale als reine Distanzkriterien angewandt worden.

Besonders muß noch das Mißverständnis ausgeschlossen werden, als ob Teiltöne von sehr verschiedener Höhe, z. B.  $c^1$  und  $c^4$ , wenn sie die gleiche Stärkeziffer, sagen wir 3, erhalten, damit als subjektiv gleich stark bezeichnet sein sollten. Es ist bekanntlich unmöglich, Töne so verschiedener Höhe in Hinsicht ihrer subjektiven Stärke direkt zu vergleichen. Die identische Stärkeziffer bedeutet also in unserem Falle nur, daß jeder in seiner Art, für sich beurteilt, die mit dieser Ziffer ausgedrückte Eigenschaft besitzt. Ganz anders liegen die Dinge aber bei gleicher oder nur wenig verschiedener Tonhöhe. Die deutlichen Stärkeveränderungen, die ein und derselbe Teilton bei verschiedenen auf dem gleichen Grundton gesungenen Vokalen erfährt, drängen sich in überzeugendster Weise auf. Es ist z. B. ein stets gelingender Vorlesungsversuch, wenn ich die Resonanzgabel  $h^1$  (480 Schw.) oder auch  $a^1$ ,  $b^1$  zuerst durch ein auf ihrer tieferen Oktave gesungenes O, dann durch ein ebenso kräftiges A oder I auf demselben Grundton zum Tönen bringe: im 1. Falle spricht sie laut an, im 2. erheblich weniger. Dagegen spricht wieder die Gabel  $g^2$  sehr laut an, wenn ich auf  $c^1$  ein A singe, lauter als etwa bei Ö oder I auf demselben Grundton.  $g^3$  reagiert besonders laut, wenn Ö auf  $c^1$  gesungen wird usf. Die Ursachen werden wir später in den Formanten dieser Vokale entdecken.

Aber auch wenn man von einem Teilton eines Vokals zu dem nächsthöheren oder tieferen desselben Vokals übergeht

(vorausgesetzt, daß der Schritt nicht zu groß ist, wie die Oktave zwischen dem 1. und 2. Teilton), kann man mit aller Bestimmtheit über Gleichbleiben, Zu- oder Abnahme der Stärke entscheiden. Wir erkennen z. B. direkt, daß beim Singen des A auf  $c^1$  vom 2. zum 3. Teilton eine Verstärkung, vom 4. zum 5. eine Schwächung eintritt. Man kann also die Wendepunkte, die Richtungsunterschiede der Intensitätskurve als solche mit Sicherheit erkennen. Nur die Größenverhältnisse der Maxima, zu denen die Kurven emporsteigen, lassen sich nicht durch direkte Vergleichung untereinander kontrollieren.

Zu leichterem Ausschluß von Mißverständnissen und Einwänden obiger Art könnte man statt der arabischen Ziffern römische, die heute zumeist als Ordnungsziffern dienen, oder Buchstaben setzen. Oder es wären, um ungefähr gleiche Abstände zu bekommen, die im folgenden unter II angegebenen Werte statt der unter I stehenden zu setzen:

I	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	1	$1\frac{1}{4}$	$1\frac{1}{2}$	$1\frac{3}{4}$	2	$2\frac{1}{2}$	3	$3\frac{1}{2}$	4	$4\frac{1}{2}$	5
II	$\frac{1}{2}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14

Doch hat sich unsere Bezeichnungsweise, die sich unmittelbar an die obigen Kriterien anschließt, als für den Gebrauch handlicher erwiesen und werden obige Bemerkungen hoffentlich genügen, um Mißverständnisse auszuschließen, auch wenn die Zahlen später, um die lästigen Brüche im Druck zu vermeiden, mit 4 multipliziert werden.

Von großem Interesse bleibt natürlich die Frage, wie sich die so bestimmten subjektiven zu den physikalischen Stärken verhalten. Die Schwierigkeiten der physikalischen Tonstärkemessung sind bekannt. Es ist nun kürzlich (Frühjahr 1922), lange nach Abschluß meiner experimentellen Arbeiten, durch eine von Dr. K. LEWIN ausgearbeitete Einrichtung, bei der die Schwingungen einer Membran durch ein darauf befestigtes Spiegelchen in enormer Vergrößerung sichtbar gemacht werden, möglich geworden, die zu den obigen Stufen gehörigen physikalischen Stärken (bzw. deren gegenseitiges Verhältnis) bei dem Tone  $c^2$ , auf den die Membran abgestimmt war, zu messen. Aus 4 Reihen, an deren Durchführung sich Frl. EBERHARDT beteiligte, ergab sich, daß die Amplituden etwas stärker, die Intensitäten (Quadrate der Amplituden) also viel stärker wachsen als die subjektiven Stärkezahlen, wenigstens bis zur Stärke 3. Darüber hinaus konnten Messungen nicht gemacht werden.

Das Ohr wurde in eine solche Stellung gebracht, daß die Entfernung des Trommelfells von der Öffnung der Schalleitung möglichst gleich war derjenigen der Membran bei der objektiven Messung. Die Stellung des Kopfes wurde durch einen Kopfhalter fixiert. Jeder der objektiven Stärkegrade wurde mehrmals gegeben und mit ihrer Aufeinanderfolge unregelmäßig gewechselt. Die fehlerfreie Subsumtion so wenig verschiedener Intensitäten unter bestimmte Kategorien ist nicht leicht. Doch kamen nur ganz wenige und geringe Abweichungen vor.

Die folgenden Tabellen geben ein Bild der 3 Versuchsreihen. Daß die Stufe 1 als Einheit auch für die objektiven Amplitudenverhältnisse genommen wird, ist natürlich willkürlich. Bei der 2. Reihe war der Abstand des Ohres und der Membran um 1 cm geringer als bei der 1. und 3. (dort 5, hier 6 cm). Aber auch zwischen der 1. und 3. Reihe, die durch 8 Tage getrennt waren, mögen geringe unvermeidliche Unterschiede in der Klangzuführung stattgefunden haben, auf denen es beruhen wird, daß gleichen subjektiven Stärken in der 3. durchweg größere Amplituden als in der 1. zugeordnet sind. Daher darf man die Werte nur innerhalb ein und derselben Reihe vergleichen. Die Versuche waren natürlich „unwissentlich“, die Reihenfolge der dargebotenen Stärkegrade bunt und willkürlich. Daher ist damit zugleich eine Probe für die Sicherheit der Einordnung in die subjektiven Stärkegrade gegeben.

1. Reihe		2. Reihe		3. Reihe	
Relative Amplituden	Subjektive Stärken	Relative Amplituden	Subjektive Stärken	Relative Amplituden	Subjektive Stärken
0,04	$\frac{1}{8}$	0,3	$\frac{1}{2}$	0,15	$\frac{1}{8}$
0,1	$\frac{1}{4}$	1	1	0,2	$\frac{1}{4}$
0,22	$\frac{1}{4} - \frac{1}{2}$	1,7	$1\frac{1}{2}$	0,3	$\frac{1}{2}$
0,4	$\frac{1}{2}$	2,7	fast 2	1	1
1	1			2	$1\frac{1}{2}$
1,25	$1\frac{1}{2}$	4,9	$2 - 2\frac{1}{2}$	2,5	2
3,4	3	5,6	$2\frac{1}{2} - 3$	4	$2\frac{1}{2}$
		6	3	4,5	knapp 3

Hier fügen sich nur der Übergang von der drittletzten zur vorletzten Stufe in der 1. und der von der 5. zur 6. Stufe in der 3. Reihe nicht der obigen Regel. Doch wollen diese Versuche nur als erste Probe einer Verhältnisbestimmung angesehen sein.

### 3. Gabeln und Sänger.

Die von mir benutzten Gabeln reichen von  $c$  bis  $c^5$  und umfassen fast sämtliche Obertöne für die Grundtöne  $c$ ,  $c^1$  und  $c^2$  bis zu dieser Grenze. Von  $c^2$  bis  $c^4$  waren es die obenerwähnten KOENIGSchen Gabeln<sup>1)</sup>. Bis zu dieser Höhe waren sie mit Resonanzkästen verbunden. Für die 4-gestrichene Oktave hatte ich sehr massive KOENIGSche Gabeln aus dem Breslauer Physikalischen Institut, die genau auf die Obertöne des  $c^1$  vom 8. bis zum 16. abgestimmt sind. Sie stehen nicht auf Resonanzkästen, sind aber

<sup>1)</sup> Der Umstand, daß die Oktave  $c^2 - c^3$  in dieser Serie durch die chromatische Leiter vertreten ist, also nicht überall genau mit den Obertonhöhen stimmt, mußte bei den Versuchen berücksichtigt werden. Galt es z. B., bei dem gesungenen Grundton  $c$  die Stärke des 5. Teiltones zu bestimmen, so mußte der Grundton ein wenig höher als sonst intoniert werden, weil erst dann die Gabel für  $c^3$  das genaue Multiplum des Grundtones war. Für geübte Sänger macht dies keine Schwierigkeit.

so empfindlich, daß sie schon durch leises Blasen oder rasche Bewegung in Schwingung kommen<sup>1)</sup>).

Für den Ton *c* war zwar auch eine Resonanzgabel vorhanden, die einen lang anhaltenden Ton gibt. Aber sie resoniert nicht auf ein noch so starkes Ansingen, war also für unseren Zweck nicht zu brauchen. Infolgedessen muß in den unten folgenden Tabellen bei den auf *c* gesungenen Vokalen die Stelle des Grundtons unausgefüllt bleiben.

Außer den Obertönen des *c* sind aber auch noch zahlreiche andere Töne (besonders die Multipla von 50) in der Gabelsammlung des Berliner Psychologischen Instituts vertreten, mit denen Obertöne anderer Grundtöne als *c*, *c*<sup>1</sup>, *c*<sup>2</sup>, wenn auch nicht so vollständig, geprüft werden können. Insgesamt standen gegen 60 Gabeln zu Resonanzversuchen zur Verfügung.

Immer wurde besonders darauf geachtet, daß die Gabel zu Beginn des Versuchs, wenn sie vor dem Munde des Sängers stand, sich in völliger Ruhe befand. Während des Singens aber wurde ein Karton zwischen dem Mund und der Gabel gehalten, um direktes Anblasen auszuschließen.

Von Seite des Sängers ist möglichst gleiche Expirationsstärke innerhalb einer Versuchsreihe, möglichst genaue Intonation des Grundtons und möglichst deutliche Nuancierung des Vokals erforderlich. Sänger ohne hinreichende Ausbildung (manchmal aber auch Geübte) pflegen den Ton etwas zu tief einzusetzen und dann, wenn auch sehr rasch, auf die richtige Höhe zu steigen. Dadurch kann man zu Fehlschlüssen kommen. Beispielsweise wird bei genauer, sicherer Intonation eines Vokals auf *d*<sup>1</sup> niemals die Gabel *c*<sup>3</sup> mitklingen. Wohl aber kann dies bei der angegebenen Singweise geschehen: denn wenn ein A auch nur für einen Moment zunächst auf *cis*<sup>1</sup> eingesetzt wird, so enthält dieser Klang den Oberton *cis*<sup>3</sup>, der bereits in der Resonanzzone der *c*<sup>3</sup>-Gabel liegt und sie in leise Mitschwingung versetzen kann. Läßt man in solchem Falle *d*<sup>1</sup> mit einem Vorschlag von oben her singen, so bleibt *c*<sup>3</sup> völlig in Ruhe.

In bezug auf die Nuancierung muß man darauf halten, daß bei solchen Versuchen Deutlichkeit über Schönheit geht. Kunstsänger verändern bekanntlich vielfach die natürliche Vokalisierung. Hier-

<sup>1)</sup> Diese schöne Serie habe ich für das Berliner Psychologische Institut nacharbeiten lassen und die Reinstimmung selbst mittels der Differenztonmethode (St. 4) überwacht. Die Abweichungen von den vorgeschriebenen Schwingungszahlen betragen, wenn die KOENIGSche Gabel 512 als genau vorausgesetzt wird, kaum mehr als eine Schwingung. Man kann dies mit Differenztönen in vielfacher Weise kontrollieren. Der Berliner Serie sind auch noch die Töne *f*<sup>4</sup> und *a*<sup>4</sup> der reinen diatonischen Leiter eingefügt.

von muß bei wissenschaftlichen Versuchen abgesehen werden. Da gilt es, die Vokale aufs schärfste zu charakterisieren und gegeneinander abzugrenzen. U muß daher sehr dunkel gehalten werden, O hell (offen), aber nicht schon A-haltig, A hell, aber nicht Ä-haltig, E geschlossen, wie in „Rede, Ebbe“, nicht Ä-haltig, wie in „fett“, I so hell und scharf als möglich.

Es ist erstaunlich, wie fein die Resonanzgabeln auf kleine Veränderungen der Nuancierung reagieren. Beispielsweise erzielte Frl. OHLHOFF durch Veränderungen der Mundstellung, die für mein Ohr noch kaum einen Unterschied des Vokals bedingten, wenn sie beim E die Lippen etwas emporzog, um mehr Luft herauszulassen, schon eine erhebliche Verschiebung in der Resonanzstärke einzelner Gabeln.

Als Sänger wirkten bei den systematischen Reihen außer mir selbst mit: Herr Dr. ABRAHAM (= Ab, Bariton), Frl. BEHREND (= Be, Sopran), Herr BELIAN (= Bl, Baß), Herr HARTMANN (= Ha, Bariton), Herr KRAUSE (= Kr, Tenor, etwas nach Bariton), Herr KREICHAUER (= Kg, Tenor), Frau MELANIE KURT (Sopran), Frl. OHLHOFF (= Oh, Sopr.), Frl. A. und H. PLANCK (= AP, HP, Sopr.) und ich selbst (= St, Bariton). Die Damen BEHREND, KURT und OHLHOFF sind rühmlich bekannte Konzert- bzw. Opernsängerinnen, die Herren BELIAN und KRAUSE ausgebildete Sänger, die übrigen Dilettanten<sup>1)</sup>. Allen bin ich für ihre geduldige und sorgfältige Mitwirkung zu lebhaftem Danke verpflichtet.

## II. Ergebnisse.

Untersucht wurden auf diesem Wege vorzugsweise die 5 sog. Hauptvokale A E I O U, gelegentlich aber auch die Umlaute Ö Ä Ü. Als Grundtöne wurden bei den systematischen Reihen *c*, *c*<sup>1</sup>, *c*<sup>2</sup> gewählt. Auf *c* wurden 6 Reihen mit 5 Sängern, auf *c*<sup>1</sup> 9 Reihen mit 6 Sängern und 2 Sängerinnen, auf *c*<sup>2</sup> 6 Reihen mit 5 Sängerinnen durchgeführt.

<sup>1)</sup> Versuchsreihen mit einer vierten Konzertsängerin von vorzüglicher Stimme mußte ich hier außer Betracht lassen, weil sie den Wohlklang so sehr über die Charakteristik der Vokale stellte, daß dadurch empfindliche Undeutlichkeiten in die Aussprache und demzufolge auch in die Tabellen kamen. Auch in der Stärke nahm sie die verschiedenen Vokale außerordentlich ungleich, namentlich I viel schwächer als die übrigen, A besonders kräftig.

Auch mehrere Versuchsreihen mit 2 anderen Personen, Alt und Bariton, sind nicht in die unten folgenden Tabellen aufgenommen, weil es die ersten Reihen waren, an denen die Methode ausprobiert wurde. Doch zeigen sie bereits alle wesentliche Züge der späteren Reihen.

Im folgenden erwähnen wir zunächst unter 1.—3. die Ergebnisse bei den Grundtönen *c* und *c*<sup>1</sup>, dann unter 4. die Modifikationen bei *c*<sup>2</sup>, unter 5. individuelle Unterschiede.

### 1. Zahl der Teiltöne.

Gesungene Vokale besitzen bei starker Stimmgebung in der Nähe der Klangquelle eine unerwartet große Zahl von Teiltönen. Beim *A* auf *c* lassen sich Teiltöne bis zum 32., bei dem auf *F* gesungenen *I* einer kräftigen Baßstimme bis zum 36. nachweisen; und zwar ist in diesen Fällen die Reihe, soweit nur immer unsere Gabeln sie verfolgen lassen, lückenlos. In anderen Fällen sind Lücken vorhanden; aber selbst bei dem dunkelsten Vokal *U* finden sich in manchen Stimmen noch Teiltöne von den Ordnungszahlen 20—22 (beim Grundton *c* die Töne *c*<sup>4</sup> und *fis*<sup>4</sup>).

Schon dieses Ergebnis ist theoretisch darum bedeutsam, weil gegen die Definition der Vokale durch harmonische Teiltöne der Einwand erhoben wurde, daß man schon den 14. Teilton in der menschlichen Stimme nicht mehr nachweisen könne, während diese Theorie viel höhere verlangt<sup>1)</sup>.

### 2. Harmonische Teiltöne.

Nur harmonische Teiltöne sind in der menschlichen Stimme enthalten. Gabeln, die von solchen um mehr als Resonanzbreite abweichen, bleiben vollkommen stumm. Liegen sie innerhalb der Resonanzbreite, so klingen sie zwar mit — und man kann dann zunächst zu dem Schlusse kommen, daß auch unharmonische Teiltöne vorhanden seien —, aber es zeigt sich regelmäßig, daß sie noch stärker mitschwingen, wenn sie mit benachbarten harmonischen Teiltönen in Einklang gebracht werden, wenn also der Sänger seinen Ton um soviel höher oder tiefer nimmt. A. STEFANINI ist bei ähnlichen Versuchen einem solchen Fehlschluß unterlegen<sup>2)</sup>,

<sup>1)</sup> HERMANN, Bd. 58, S. 274: „Wie kann man auch nur einen Augenblick annehmen wollen, daß der 28. oder 31. Partialton des Stimmklanges noch so weit vertreten sei, um durch Resonanz des Mundes verstärkt zu werden.“ NAGEL, S. 783: „Wollte man den Formanten *f*<sup>4</sup> des *I* als verstärkten Partialton betrachten, so käme man beim Singen des *I* mit einer Baßstimme auf die Note *G* auf den 28.—29. Partialton und selbst bei *g* noch auf den 14. Partialton. Das Vorhandensein von Obertönen solcher Ordnungszahlen ist überhaupt im Stimmklange nicht mehr nachweisbar; und doch ist der *I*-Klang auch auf den tiefen Noten deutlich zu singen.“

Dagegen schon KOEHLER I, II, S. 75.

<sup>2)</sup> STEFANINI behauptet ausdrücklich, daß die Resonanzgabeln keine Resonanzbreite besäßen, sondern nur auf ihre genaue Schwingungszahl ansprächen. Dies ist ein Irrtum. Übrigens hatte er auch bei weitem zu wenig Gabeln zur Verfügung. Seinem Bedenken, daß innerhalb einer Versuchs-

und ich selbst wurde durch die ersten Eindrücke der Resonanzversuche zu der Annahme unharmonischer Teiltöne verleitet, habe aber bald genug das angegebene Verhalten ausnahmslos bestätigt gefunden.

### 3. Charakteristische Struktur.

Jeder Vokal hat eine gewisse charakteristische Struktur, eine Intensitätskurve seiner Teiltöne. Dies ist aber nicht im Sinne der „Relativtheorie“ zu verstehen. Denn wenn man den Grundton wechselt, bleiben, abgesehen vom U, die Maxima auf ihrer absoluten Höhe liegen<sup>1)</sup>. Infolgedessen streckt sich die Intensitätskurve beim Sinken des Grundtons immer mehr in die Länge, wenn sie auch ihre allgemeine Gestalt beibehält. Man vergleiche z. B. unter den unten folgenden Tabellen die des A auf  $c$  und auf  $c^1$ . Das 1. Maximum bleibt auf  $g^2$ , das 2. auf  $fis^4$  liegen. Wo Verschiebungen vorkommen, sind sie doch nur geringfügig und verstehen sich leicht teils aus unvermeidlichen zufälligen Urteilschwankungen des Beobachters, teils aus den ebenso unvermeidlichen kleinen Unterschieden in der Nuancierung des Vokals und in der Stimme des Sängers. Um solche Zufälligkeiten und individuellen Unterschiede möglichst auszugleichen, habe ich versucht, die Stärkeziffern aller Versuchspersonen, die einen bestimmten Vokal auf bestimmter Tonhöhe gesungen haben, zu summieren. Die Gesetzmäßigkeiten sind dabei in der Tat besonders deutlich hervorgetreten. Dennoch möchte ich hier auf die Unterstützung solcher Summen- oder Durchschnittstabellen verzichten, da die Berechtigung der Summierung wegen der obenerwähnten qualitativ mitbedingten Schätzungskriterien bestritten werden könnte.

Wir geben für die Grundtöne  $c$ ,  $c^1$ ,  $c^2$  eine Anzahl von Einzeltabellen, die zugleich die individuellen Verschiedenheiten erkennen lassen, bei  $c^1$  und  $c^2$  auch je 2 Reihen der nämlichen Person zu verschiedenen Zeiten, die begreiflicherweise auch nicht ganz zusammenfallen. Zur Vergleichung mag man dann immerhin auch Summen- oder Durchschnittszahlen daraus bilden.

reihe nicht die nötige Garantie für die gleiche Nuancierung seitens des Sängers gegeben sei und daher eine Menge von Beobachtern auf einmal, jeder für einen der Teiltöne, zusammenwirken müßten, kann ich kein Gewicht beimessen. Der Beobachter muß eben achtgeben, daß die Nuancierung konstant bleibt, und Versuche mit zufälliger merklicher Abweichung sofort durch andere ersetzen.

<sup>1)</sup> Wir werden unten aus den Ergebnissen der Interferenzversuche, die eine unbeschränkte Veränderung des Grundtons gestatten, ein langsames Hinaufrücken der Maxima mit steigendem Grundton erschließen. Aber bei den Resonanzversuchen auf  $c$  und  $c^1$  kam es nicht zum Vorschein.

Bei den Notenbuchstaben bedeutet hier und in späteren Tabellen ein Strich über dem Buchstaben eine kleine Erniedrigung gegenüber dem damit bezeichneten Leiterton. Bekanntlich fallen die höheren Multipla des Grundtones nicht immer mit Leitertönen zusammen.

Das Fehlen von Zahlenwerten für den Grundton in der Tabelle für *c* beruht auf dem Fehlen der Resonanzgabel *c* (s. o.). Der Strich ist also nicht mit 0 zu verwechseln.

Man kann zur graphischen Veranschaulichung auch Diagramme zeichnen, in denen die Stärkegrade durch die Höhe schmaler Rechtecke wiedergegeben sind, wobei aber wieder nur das Wachsen und Abnehmen überhaupt, nicht genaue Maßverhältnisse, ausgedrückt sein sollen. 12 Tafeln solcher Diagramme nach meinen Resonanzversuchen besitzt das Berliner Psychologische Institut.

F. TRENDLENBURG hat kürzlich (2) aus 2 meiner Zahlenreihen für den Vokal A (Kr auf *c*<sup>1</sup> und der ersten HP auf *c*<sup>2</sup>) in der physikalisch gebräuchlichen Weise durch Verbindungslinien zwischen den diskreten Werten Kurven gebildet und diese mit 2 seiner eigenen, aus objektiven Schwingungsaufnahmen abgeleiteten Amplitudenkurven zusammengestellt. Da es sich in unserem Falle um subjektive Stärken handelt, bei denen die Umbildung der Tonprozesse im Ohr und im Nervensystem mitspielt, so kann man eine genaue Übereinstimmung nicht erwarten. Gleichwohl ist der Verlauf im allgemeinen überall ziemlich derselbe. Insbesondere haben das untere und das obere Maximum in allen 4 Kurven die gleiche absolute Lage. Und jedenfalls zeigt der Versuch, daß man auch diese Darstellungsweise anwenden und zu interessanten Vergleichen benutzen kann.

Grundton *c* = 128 Schw.

Teil-ton	Note	Ab	St	Ha	Bl	Ab	St	Ha	Bl	Ab	St	Ha	Bl	Ab	St	Ha	Bl				
32	<i>c</i> <sup>5</sup>									4	2	1		4	2	3		4	2	2	
30	<i>h</i> <sup>4</sup>									4	3	1		6	2	2		8	2	2	
28	<i>b</i> <sup>4</sup>									4	3	2	1	8	6	3	1	8	3	2	1
26	<i>a</i> <sup>4</sup>									8	3	2	2	6	4	3	4	8	4	3	2
24	<i>g</i> <sup>4</sup>								1	8	4	4	3	8	4	3	3	10	4	4	3
22	<i>f</i> <sup>4</sup>				1			2	2	10	6	2	4	8	4	4	2	8	4	4	3
20	<i>e</i> <sup>4</sup>				2		1	2		8	6	2	4	4	4	6	4	8	4	4	3
18	<i>d</i> <sup>4</sup>				0		1	1		4	4	2	2	8	2	4	1	8	2	1	3
16	<i>c</i> <sup>4</sup>				0		1	0		4	2	1	1/2	8	2	3	2	4	2	1	2
15	<i>h</i> <sup>3</sup>				0		0	0		4	3	1	0	4	1	0	0	3	1	0	0
12	<i>g</i> <sup>3</sup>				0		0	0		8	4	4	6	6	2	0	1	4	2	0	1/2
10	<i>e</i> <sup>3</sup>				1/2		1	1/2		10	6	4	8	6	0	0	1	3	0	0	1
9	<i>d</i> <sup>3</sup>				1/2		2	0	1/2	12	6	8	8	4	0	0	1	2	0	0	1
8	<i>c</i> <sup>3</sup>	2		1/2	2		4	1/2	4	2				4	0	0	1	3	0	0	1
7	<i>b</i> <sup>2</sup>	6	3	4	4		8	4	8	8				2	1/2	4	4	8	1/2	2	4
6	<i>g</i> <sup>2</sup>	12	10	6	16		20	12	12	20				10	1	2	4	6	1	0	4
5	<i>e</i> <sup>2</sup>	12	8	0	12		16	10	4	12				3	2	1/2	8	8	2	0	8
4	<i>c</i> <sup>2</sup>	16	12	6	20		20	16	14	20				10	10	6	12	8	6	1/2	12
3	<i>g</i> <sup>1</sup>	16	16	8	16		12	8	18	18				12	12	12	8	12	12	8	4
2	<i>c</i> <sup>1</sup>	16	16	14	12		12	16	10	14				8	8	10	12	6	14	12	12
1	<i>c</i>																				
		U				O				A				E				I			



Grundton  $c^1 = 256$  Schw.

Teilton	Note	Ab	Ab	St	Kg	Kr	Ha	Ab	Ab	St	Kg	Kr	Ha	Ab	Ab	St	Kg	Kr	Ha	Ab	Ab	St	Kg	Kr	Ha							
16	$c^5$				1			$1/2$	2	2							4	4	6			2	2	2	$1/2$							
15	$h^4$			$1/2$				2	4	4				6	4	6	0	0		3	3	10			4	3	3	1	0			
14	$\bar{b}^4$	1		$1/2$				6	5	4				8	4	4	4	1		4	6	8	10			4	4	4	12	2	1	
13	$\bar{a}^4$	1	1	$1/2$			1	6	5	3	2	2		2	10	8	4	6	3	8	10	10	8	2	3	12	10	6	12	1	4	
12	$g^4$	1	2	2	4	3		3	6	3	12	4		8	12	10	10	12	4	12	10	12	12	4	4	16	12	8	12	8	6	
11	$\bar{f}is^4$	2	2	4	2	1		4	10	8	14	4		4	14	14	10	16	4	12	7	14	16	12	4	16	10	8	16	10	6	
10	$e^4$	2	2	2	12	0		8	4	1	16	0		8	14	8	6	16	2	12	16	12	12	16	2	8	18	10	8	8	4	
9	$d^4$	2	3	3	8	1		2	4	1	12	0		6	8	6	4	8	1	12	12	10	8	20	4	4	14	12	10	12	4	
8	$c^4$	1	3	3	8	0		2	3	2	4	1		4	8	3	3	6	2	12	14	8	10	10	8	6	8	6	6	8	8	
6	$g^3$	4	3	2	1	2	2	4	3	4	3	4	2	8	12	10	8	12	4	8	10	8	4	12	4	6	8	4	3	8	2	
5	$e^3$	2	2	4	1	4	2	2	4	4	4	8	3	16	16	12	12	12	12	8	6	6	3	12	4	4	4	4	3	4	2	
4	$c^3$	4	4	4	4	6	4	4	4	8	8	12	3	16	16	12	16	16	12	12	6	6	3	16	6	8	4	4	4	4	4	
3	$g^2$	20	20	20	16	16	16	16	20	16	16	20	20	20	22	22	22	22	24	12	10	8	12	22	8	8	8	8	4	8	12	8
2	$c^2$	16	16	10	14	12	16	20	24	20	20	22	24	16	16	12	16	16	16	16	14	16	16	16	12	8	14	4	10	16	4	
1	$c^1$	20	22	22	16	20	18	16	16	16	16	12	18	12	12	12	10	16	10	16	16	16	10	8	12	16	18	18	14	20	14	
		U						O						A						E						I						

Grundton  $c^2 = 512$  Schw.

Teilton	Note	Be	Oh	AP	HP	HP	Be	Oh	AP	HP	HP	Be	Oh	AP	HP	HP	Be	Oh	AP	HP	HP	Be	Oh	AP	HP	HP					
8	$c^5$				2		$1/2$					2	4	8	8	6	1	2	4	4			$1/2$			4	8				
7	$\bar{b}^4$	6		2	4		8	4	4	6	2	8	6	8	8	6	12	4	4	8	6			12	4	10	10	4			
6	$g^4$	4		2	2	1	8	4	8	4	2	14	4	8	6	8	16	4	10	10	4			16	4	10	10	4			
5	$e^4$	2	3	2	2	8	2	6	4	4	6	4	4	2	4	4	8	8	12	16	12			4	12	16	16	8			
4	$c^4$	$1/2$	4	0	0	1	3	3	4	1	2	8	6	4	6	4	4	16	14	12	8			2	18	12	16	8			
3	$g^3$	6	4	4	4	2	6	6	4	4	2	16	12	6	6	10	6	10	10	10	8			6	10	10	10	8			
2	$c^3$	14	12	8	8	16	14	20	12	12	20	12	16	10	8	12	8	12	4	8	8			8	12	4	4	4			
1	$c^2$	24	24	20	16	22	24	20	18	18	16	16	18	18	12	12	16	22	18	16	16			20	24	18	20	20			
		U						O						A						E						I					

Folgendes läßt sich aus diesen Tabellen deutlich erkennen; wobei wir vorerst unter  $a-b$  nur die Verhältnisse für die Grundtöne  $c$  und  $c^1$  ins Auge fassen.

a) Zahl und Lage der Maxima.

Die Lage des ersten Maximums versinnlicht dieses Schema:

	Grundton $c$					Grundton $c^1$				
$g^2$			*					*		
$e^2$										
$c^2$		*				*			*	
$g^1$				*						
$c^1$	*				*	*				*
$c$										
	U	O	A	E	I	U	O	A	E	I

Wie man sieht, liegt es für beide Grundtöne in gleicher absoluter Höhe. Die Haupttabellen zeigen nur seltene und unbedeutende Rückungen, die als zufällig oder individuell angesehen werden müssen. Das des E rückt nur darum von  $g^1$  auf  $c^2$ , weil  $g^1$  eben nicht als harmonischer Teilton im  $c^1$ -Klange vorhanden ist. Für den Grundton  $c$  muß allerdings mindestens beim U und I die Frage noch offen bleiben, ob das Maximum hier nicht auf dem Grundton selbst liegt. Wir hatten ja dafür keine ansprechende Resonanzgabel. Diese Frage werden wir später zu bejahen finden.

Ferner ist immer im Auge zu behalten, daß die Verhältnisse bei schwächer oder in größerer Entfernung angegebenen Vokalen nicht genau dieselben zu sein brauchen. O. WEISS hat nach graphischen Befunden bereits darauf hingewiesen, daß gerade in Hinsicht der relativen Grundtonstärke zwischen stark und schwach angegebenen Vokalen ein Unterschied bestehe: in schwachen sei der Grundton relativ stärker vertreten.

Außer dem 1. Maximum weisen unsere Tafeln fast alle noch weitere auf. Zunächst ist für U und O fast immer ein solches bei  $g^2$  bemerkbar (auch in den Durchschnittszahlen deutlich ausgeprägt). Beim O auf  $c^1$  bildet es die Fortsetzung des 1. Maximums, da  $g^2$  unmittelbar auf  $c^2$  folgt. In hellen oder metallreichen Stimmen folgt aber, wie besonders aus der Tabelle für  $c^1$ , teilweise aber auch aus der für  $c$  hervorgeht, beim U und O der Bruststimme sogar noch ein 3. Maximum in der 4-gestrichenen Oktave.

In dieser Gegend haben auch A, E und I ein 2. Maximum, A aber wieder nur bei metallischen Stimmen im Brustregister.

Die oberen Maxima für U, O, A haben neuerdings F. TRENDELENBURG und gleichzeitig die Amerikaner CRANDALL und SACIA auf graphischem Wege nachgewiesen (vgl. 8. Kap.). Auch einige Kurven HERMANN'S und KOEHLER'S scheinen sie schon zu enthalten. In meinen Versuchen wären sie vielleicht auch noch öfter aufgetreten, wenn ich nicht in der ersten Zeit über diese hohe Region bei U und O etwas rasch hinweggegangen wäre, in der Annahme, daß hier keine Reaktion mehr zu erwarten sei. Übrigens vermuten die genannten Forscher mit Recht, daß diese schwachen hohen Teiltöne der dunklen Vokale nur die individuelle Klangfarbe der Stimme, nicht den Vokalcharakter beeinflussen. Für meine eigene Stimme habe ich auch bei gelegentlichen direkten Beobachtungen des U ohne Resonanzgabeln mehrmals notiert: „auch ganz hohe Töne der 4-gestrichenen Oktave, die aus der Nase zu kommen scheinen“. Und GRASSMANN hat, wie ich nachträglich sehe, diese Töne beim A schon 1854 direkt beobachtet (I, S. 14). Auch HELMHOLTZ erwähnt (S. 169), daß bei etwas angestrengtem Forte der Stimme auf höheren Noten deutlicher als bei allen (?) anderen Instrumenten hohe Obertöne aus der Mitte der 4-gestrichenen Oktave erscheinen, deren Stärke aber individuellen Verschiedenheiten unterworfen und bei scharfen, hellen Stimmen größer sei. Er beobachtete ferner (S. 188) ein eigentümliches Rasseln bei Chorvorträgen im Forte, das er auf die Schwebungen dieser hohen, durch die Resonanz des Gehörganges verstärkten Obertöne zurückführt. Dieses Rasseln ist auch mir in Konzerten stets aufgefallen; es ist sehr störend, sobald man seine Aufmerksamkeit darauf richtet, wird aber sonst überhört. Es erscheint mir gegenüber den Stimmen selbst näher lokalisiert, was mit der Verstärkung im Gehörgange zusammenhängen mag. Bei Orchesterinstrumenten hat es HELMHOLTZ niemals vernommen. In der Tat scheint es da nur ausnahmsweise aufzutreten und ist jedenfalls bedeutend weniger auffällig. Der Grund dürfte darin liegen, daß die hohen Teiltöne, die auch in Instrumenten kräftig vorhanden sein können, hier nicht, wie bei den Vokalen, durch leere oder sehr schwache Strecken von den tieferen getrennt sind, sich daher innerhalb der Klangmasse nicht so als gesonderte Gruppen geltend machen; wozu noch kommt, daß die Tonbewegung in der Orchestermusik durchschnittlich rascher, umfangreicher, mehr figuriert und auf abwechselnde Instrumente verschiedenster Klangstruktur verteilt ist. Bei den totenerweckenden Posaunenfanfaren in BERLIOZ' Requiem kann man ein scharfes Stechen wahrnehmen, an dem noch höhere Partialtöne beteiligt sind, aber es erscheint nicht ebenso abgelöst vom Gesamtklange.

Die genauere Lage des E- und I-Maximums läßt sich aus den Resonanztabellen nicht eindeutig erkennen; nur in den Durchschnittszahlen tritt ein deutlicher Unterschied hervor, indem sie für E bei  $e^4$ , für I bei  $fis^4-g^4$  ein Maximum aufweisen.

Auch die Umlaute Ö, Ä, Ü, die bei Bl und Kr mituntersucht wurden, ergaben auf den beiden Grundtönen  $c$  und  $c^1$  ähnliche Zahlenkurven mit einem unteren und einem oberen Maximum. Das erste fällt für Ö und Ä mit dem des O, für Ü mit dem des U zusammen. Das zweite liegt in der oberen Hälfte der 3- und der unteren der 4-gestrichenen Oktave.

## b) Verlauf der Stärkekurven.

Die Ab- und Zunahme der Stärken innerhalb der einen Vokal zusammensetzenden Teiltonreihe erfolgt im allgemeinen nicht ruckweise, so daß ein Maximum beiderseits von schwächsten Tönen oder einer Nullstrecke umgeben wäre, sondern allmählich. Man hat sich früher vielfach vorgestellt oder wenigstens so ausgedrückt, daß die Anwesenheit eines einzelnen bestimmten Tones, sei es eines harmonischen oder unharmonischen, schon den Vokalcharakter bedinge. Selbst HELMHOLTZ scheint (nach einer mündlichen Äußerung M. PLANCKS) eine solche Vorstellung gehegt zu haben, wenn er auch recht wohl wußte, daß die dem charakteristischen Tone naheliegenden Teiltöne etwas mitverstärkt werden. In Wahrheit ist aber deren Mitwirkung, wie wir immer deutlicher erkennen werden, fast überall durchaus wesentlich.

## c) Stärke der tiefsten Teiltöne.

Eine bemerkenswerte Regelmäßigkeit besteht sowohl für den Grundton  $c$  wie für  $c^1$  in Hinsicht der Stärke des tiefsten Tones der Tabellen (d. h. des Grundtons bei  $c^1$  und des 2. Teiltons bei  $c$ ): seine Stärke nimmt von U nach O und A ab, dann wieder nach E und I zu. Dies hängt offenbar mit der Verschiebung des unteren Stärkemaximums zusammen: je mehr dieses in die Höhe rückt, um so mehr wird dem tiefsten Ton an Stärke entzogen. Wir werden dies später bei  $c$  auch bezüglich des Grundtones selbst bestätigt finden.

Vergleicht man ferner die unteren Teiltöne bei  $c$ ,  $c^1$ ,  $c^2$ , so zeigt sich, daß ihre Stärke im allgemeinen mit der Höhe des Grundtones zunimmt. Die Gesamtenergie des Schalles verschiebt sich nach unten hin. Wir werden diesem aus allgemeineren Gründen wohlverständlichen Zuge noch öfter, auch bei den Instrumenten, wieder begegnen.

4. Veränderungen bei und über  $c^2$ .

Bei dem Grundton  $c^2$  verändert sich das Bild merklich. Die einzelnen Vokalstrukturen nähern sich einander. Die ersten Maxima liegen sämtlich auf dem Grundton. Doch ist die soeben erwähnte Gesetzlichkeit auch hier deutlich: der hier überall sehr starke Grundton nimmt von U bis A ab, dann wieder bis I zu. Ein 2. Maximum ist hier überall in der 4-gestrichenen Oktave vorhanden. Doch sinken die Teiltöne des U und O nach der Höhe hin alsbald zu niedrigen absoluten Werten herab, während die höheren Teiltöne von U bis I im allgemeinen immer mehr hervortreten. Da nun die ersten Maxima hier überall auf demselben

Ton, die zweiten auf nur wenig verschiedenen Tönen liegen, ist zu erwarten, daß die Eigentümlichkeiten der Vokale weniger als bei tieferen Grundtönen zur Geltung kommen, insbesondere können U und O, I und E untereinander nach den Tabellen nur wenig verschieden sein (besonders bei den 2 Konzertsängerinnen sind die bezüglichlichen Zahlen so gut wie identisch). Und so ist es auch in Wirklichkeit, wie wir im 3. Kap. eingehender dartun werden. Auch daß A in dieser Lage noch am deutlichsten erkennbar bleiben muß, läßt sich aus den Resonanztabellen verstehen: hier sind die 5 Zahlenreihen am gleichförmigsten und auch die höchsten Teiltöne überall noch ziemlich kräftig ausgebildet.

Bei Grundtönen, die über  $c^2$  hinaus liegen, tritt eine fortschreitende Vereinfachung und Angleichung der Vokalstrukturen ein. Mit der Resonanzmethode habe ich nur beispielshalber das A auf  $gis^2 = 800$  Schw. untersucht, wofür mir 6 Teiltongabeln zur Verfügung standen:  $gis^2$ ,  $gis^3$ ,  $dis^3$ ,  $gis^4$ ,  $his^4$ ,  $dis^5$ . Sängerin war Frau MELANIE KURT, die ein besonders helles, glänzendes A hat. Die ersten 5 Gabeln wurden alle, und zwar in regelmäßig abnehmender Stärke, in Mitschwingung versetzt. Das Maximum lag auf dem Grundton, ein zweites schien hier nicht mehr vorhanden zu sein.

##### 5. Individuelle und Registerunterschiede.

Während die auf gleicher Tonhöhe mehrfach aufgenommenen Reihen desselben Sängers (Ab, HP) in ihrem Verlaufe wesentlich übereinstimmen und die in der Klangfarbe ähnlichen Stimmen der Schwestern Pl. auch in den Stärkekurven diese Ähnlichkeit zeigen (nur etwas voller klingt die Stimme AP's, und auch dies läßt sich in den Zahlentabellen verfolgen), offenbaren sich in anderen Fällen gewisse Stimmeigentümlichkeiten. Dahin gehören die obersten Maxima für U und O (s. o.). Bei Kr scheint besonders  $e^4$  im Stimmklange begünstigt. E, das er nach eigener Angabe auf  $c^1$  schon „etwas gedeckt“ singt, hat fast dieselbe Zusammensetzung wie Ö. Bei Frau KURT fanden sich ungewöhnlich hohe Teiltöne, auch ihr U hatte noch ein Nebenmaximum in der 4-gestrichenen Oktave. Ähnliches bei Frl. BEHREND, bei der auch die zweiten Maxima des E und I höher als sonst liegen. Sie kann gleichsam in den Obertönen höher hinauf singen. Bei Frl. OHLHOFF haben die Grundtöne des U und I eine außerordentliche Stärke. Auch Ü singt sie besonders voll. Beim Übergang Ü—U erreicht die Resonanz der  $c^2$ -Gabel eine Stärke, die ich mit 7 bezeichnen würde. Auch Frl. BEHREND hat auf  $c^2$  ein U mit einer Grundtonstärke über 6 (24).

Die geringere Stärke meiner eigenen Stimme (darüber hat man selbst kein direktes Urteil) gegenüber der Dr. ABRAHAMS zeigt sich in den absoluten Zahlen, die fast überall kleiner sind. Noch geringer sind die absoluten Teiltonstärken bei Ha. Aber der ganze Aufbau der Vokale, die wesentlichsten Stärkeverhältnisse unter den Teiltönen werden dadurch nicht berührt. Bei Kg erkennt man die ungewöhnlich helle Stimme, die auch im Sprechen auffällt, besonders an den Zahlen für A und E. HP muß das U in der 2. Reihe etwas dunkler gegeben haben usw.

Schlechte, dünne Stimmen scheinen Lücken in der Ober-tonreihe, gleichsam zerschlis-sene Struktur, zu haben<sup>1)</sup>.

Im Anschluß an die indi-viduellen Unterschiede möge noch eine kleine Versuchsreihe mit Dr. ABRAHAM erwähnt werden, bei der es darauf ankam, ein helles A, wie es in diesen Versuchen stets ver-langt wurde, mit einem dunk-leren zu vergleichen. Die ver-schiedene Aussprache des A gehört ja auch unter die ty-pischen individuellen Verschie-denheiten im gewöhnlichen Leben.

Das Maximum bleibt in allen Fällen auf  $g^2$ . Der Un-terschied des hellen vom dunk-  
 len A zeigt sich aber darin, daß beim dunklen die unteren Teiltöne stärker sind als beim hellen, während bei diesem die hohen das Übergewicht über die des dunklen haben. Beim dunklen sind die Töne bis  $g^2$ , beim hellen die von  $c^3$  bis  $c^5$  stärker (nur einige Male gleich). Es sind meist kleine Unterschiede, doch die der tiefsten Töne sind bedeutend, und alle zusammenwirkend genügen, die Klangfarbe merklich zu verändern<sup>2)</sup>.

	A auf $c$		A auf $c^1$	
$c^5$	0	0	$1/2$	0
$h^4$	$1/2$	$1/2$	1	$1/2$
$\bar{b}^4$	1	1	2	1
$\bar{a}^4$	2	1	6	4
$g^4$	6	4	6	4
$\bar{j}is^4$	4	3	3	3
$e^4$	4	1	4	2
$d^4$	2	1	2	1
$c^4$	$1/2$	$1/2$	2	1
$\bar{h}^3$	1	0	—	—
$\bar{b}^3$	2	$1/2$	6	1
$g^3$	4	2	8	6
$e^3$	8	6	8	8
$d^3$	10	8	—	—
$c^3$	12	8	16	16
$\bar{b}^2$	16	16	—	—
$g^2$	18	20	22	24
$e^2$	12	14	—	—
$c^2$	12	16	16	20
$g^1$	4	10	—	—
$c^1$	3	8	4	8
$c$	—	—	—	—
	hell	dunkel	hell	dunkel

<sup>1)</sup> Auch SOKOLOWSKY glaubt aus seinen graphischen Aufnahmen schließen zu dürfen, daß gequetschte und gaumig klingende Stimmen ärmer an Obertönen seien (nach NADOLECZNY S. 655).

<sup>2)</sup> Die vorher mitgeteilten Zahlen für das ABRAHAMSsche A auf  $c$  und  $c^1$  unterscheiden sich im einzelnen nicht ganz unbedeutend von denen des

Zwischen Männern und Frauen zeigen unsere Tabellen, wenn die Vokale auf dem gleichen Grundton ( $c^1$ ) gesungen wurden, keine charakteristischen Verschiedenheiten. Doch liegen mir nur wenige Reihen auf  $c^1$  von Frauenstimmen zur Vergleichung vor.

Wegen der eigentümlichen Klangfarbe der Knabenstimme untersuchte ich den Vokal A bei 4 Knaben (Alter von R. 13, P. 10, N. und J. 12 Jahre) auf den Grundtönen  $f^1$  und  $c^2$ . Die besten Stimmen hatten N. und J., die aus dem Knabenchor der staatlichen Bildungsanstalt in Lichterfelde ausgewählt waren; besonders schön und charakteristisch schien mir die Stimme von N. Alle Stimmen waren aber schwächer als die gewöhnlichen Frauenstimmen. Die folgenden Tabellen zeigen die Ergebnisse.

Knabenstimmen auf  $f^1$  und  $c^2$ 

	R.	P.	N.	J.		R.	P.	N.	J.
$c^5$	1	$1/2$							
$a^4$	2	3							
$g^4$	2	2	2	3	$c^5$	4	4	$1/2$	$1/2$
$f^4$	4	4	3	3	$\bar{b}^4$	4	2	$1/2$	2
$c^4$	3	3	0	0	$g^4$	2	4	3	4
$a^3$	3	3	4	$1/2$	$e^4$	8	6	4	6
$f^3$	8	6	6	4	$c^4$	4	6	8	6
$c^3$	6	4	6	2	$g^3$	8	4	4	5
$f^2$	12	14	12	10	$c^3$	8	6	8	6
$f^1$	12	12	8	10	$c^2$	10	12	11	10

A

Das 2. Maximum beim Grundton  $f^1$  ist sehr schwach und bei N. und J. durch einen Nullpunkt (mehrmals geprüft) von den tieferen Teiltönen getrennt. Auffällig überall die Schwäche des  $c^3$  (KOEHLERS A-Zentrum). Die Tabelle für  $c^2$  gestattet Vergleiche mit denen der Frauenstimme für denselben Ton (S. 24). Die Teiltöne sind fast überall schwächer als dort, und das 2. Maximum liegt etwas tiefer (hier  $c^4$ — $e^4$ , dort  $g^4$ ). Ich möchte aber aus diesen wenigen Versuchen (bisher sind Knabenstimmen meines Wissens überhaupt noch nicht geprüft) noch keine sichere Schlußfolgerung über die Ursachen der Klangverschiedenheit ziehen.

Auch den Unterschied der Stimmregister endlich kann man mit der Resonanzmethode untersuchen. DR. ABRAHAM, der auch im Falsettsingen bewandert ist, sang die 5 „Hauptvokale“ in dieser Form auf  $f^1$ . (Dieser Ton wurde gewählt, weil bei  $c^2$  schon

hellen A in der gegenwärtigen Liste: dort ist der Gesamtklang auf  $c$  offenbar stärker gewesen, was sich besonders in den tiefsten und höchsten Teiltönen zeigt. Aber die Struktur des Klanges, die Verteilung und Lage der Maxima ist gleichwohl dieselbe.

der alterierende Einfluß der Höhenlage beginnt, der sich mit dem des Falsetts vermischt hätte, und weil für  $f^1$  noch eine größere Anzahl von Teiltönen in Frage kamen, für die auch die nötigen Gabeln vorhanden waren). Die Vokale sind in dieser Lage der Fistelstimme noch ganz wohl erkennbar, aber alle mehr oder weniger verdunkelt, am meisten O, das stark nach U, und I, das etwas nach E hin verschoben ist. Sie ergaben folgende Stärkereihen:

	Falsett auf $f^1$ .					Bruststimme	
						Heldentenor	Dilettant
$a^4$					2	2	4
$f^4$				8	8	10	8
$c^4$			$\frac{1}{2}$	2	0	6	6
$a^3$	$\frac{1}{2}$	$\frac{1}{2}$	3	4	3	4	4
$f^3$	3	2	6	6	4	10	8
$c^3$	$\frac{1}{2}$	4	6	4	4	6	6
$f^2$	16	18	12	10	6	20	16
$f^1$	16	14	10	12	12	10	12
	U	O	A	E	I	A	

Die Teiltöne gehen also nicht so hoch hinauf wie bei dem Brustregister in allen Lagen. Bei den dunkelsten Vokalen fehlen die hohen Teiltöne, dagegen sind die tiefen überall relativ verstärkt. Daher die weiche und etwas verdunkelte Farbe des Falsetts. Die Lage der Maxima ist aber unverändert.

Bei graphischen Aufnahmen von Falsettönen fand KATZENSTEIN (2), daß die Wellen fast alle wie Sinuskurven aussahen; auch die Analyse ergab nur winzige Obertonstärken. Da man aber auf Vokalen jodeln kann — duliöh, holdrio! —, so waren eben die Aufnahmen ungenügend. Besser ist schon die neuerdings von NADOLECZNY (S. 644) als Probe der FRANKSchen Methode aufgenommene Kurve eines auf  $g$  falsettierten O.

Zur Vergleichung wurde das A auf  $f^1$  auch bei einem jungen Heldentenor der staatlichen Hochschule für Musik mit äußerst kräftiger, metallreicher Stimme untersucht. Aus der obigen Tabelle, die auch mit der für den gleichen Ton bei Knabenstimmen verglichen werden kann, geht hervor, daß das 1. Maximum ( $f^2$ ) außerordentlich stark, ein 2. bei  $f^3$  und ein 3. ebenso starkes bei  $f^4$  vorhanden war.  $c^3$  war auch hier auffallend schwach. Die Teiltonreihe geht aber nicht etwa besonders hoch:  $a^4$  war schon fast unhörbar und  $c^5$  bei wiederholten Versuchen überhaupt nicht zur Resonanz zu bringen.

## 6. Schranken der Gabelmethode.

Bei der Deutung und Bewertung aller vorstehenden Ergebnisse darf niemals außer acht gelassen werden, daß sie unter speziellen



Bedingungen gewonnen sind: 1. bei großer Klangstärke, unmittelbar vor dem Munde des Sängers, 2. hauptsächlich auf den Grundtönen  $c$ ,  $c^1$ ,  $c^2$ . Wenn es richtig ist, daß nur harmonische Teiltöne die Vokale zusammensetzen, so folgt gerade daraus, daß das Maximum eines Vokals nicht immer auf genau demselben Tone liegen kann, wenn es auch innerhalb einer gewissen Zone verbleiben muß. Dafür liefert auch schon die Resonanzmethode selbst die Belege. So ergibt z. B. ein auf *gis* gesungenes A ein sehr stark resonierendes  $gis^2$ , ein auf *b* gesungenes ein sehr starkes  $b^2$ . Die gesamte Strecke, innerhalb deren das 1. Stärkemaximum des A liegt, muß also mindestens bis  $b^2$  reichen. Das Genauere hierüber ist auf diesem Wege, wenn nicht eine noch viel größere Menge von Gabeln angewandt werden, nicht festzustellen.

Überhaupt hat die Methode hieran ihre natürlichen Schranken. Aber zur Vorbereitung weiterführender Methoden, zur ersten Übersicht über die Anzahl und die Stärkeverhältnisse der Teiltöne und zur Analyse individueller Unterschiede wird sie immer nützlich bleiben. Ja, auch wenn nur wenige Gabeln vorhanden sind, behält sie ihren Wert als überzeugendes Demonstrationsmittel.

Vgl. hierüber o. S. 16.

Ein guter Demonstrationsversuch ist es auch, wenn man eine Reihe von Gabeln, die die harmonischen Teiltöne eines Stimmklanges enthalten, vor sich hinstellt und auf dem Grundton dagegensingt. Man hört dann einen ganzen Akkord mit verschiedener Stärke der einzelnen Komponenten nachklingen. Gleich zu Anfang meiner Resonanzversuche ließ ich einen Sänger die Vokale auf  $c^1$  vor der Gabelreihe  $c^1 c^2 g^2 c^3 e^3 g^3$  angeben und beobachtete, daß bei U  $c^1$ , bei O  $c^2$  und  $g^2$ , bei A  $g^2$  am stärksten erklangen, bei E wieder  $c^2$ , bei I  $c^1$ . Dies stimmt vollkommen mit den durchgeführten Resonanzreihen. Aber natürlich würde es nicht zur Erkenntnis der Vokalstrukturen genügen.

## Anhang.

### Resonanzversuche am Klavier.

1. Singt man in das Klavier bei aufgehobener Dämpfung einen Vokal, so hört man ihn bekanntlich mehr oder weniger deutlich nachklingen. Es werden eben die Teiltöne gleich oder nahezu gleich gestimmter Saiten in einigermaßen entsprechenden Stärkeverhältnissen erregt. Also zugleich eine Analyse und eine Synthese; freilich beide nur sehr unvollkommen, weil die höheren Saiten dabei den Dienst versagen. Aber auch alte und neue Instrumente und mehr oder weniger klangreiche unter den heutigen antworten ungleich. Schon HELMWAG sagt in einem hinterlassenen Manuskript vom Jahre 1780: „Wenn man die Vokale A, O, U in die Saiten eines Klaviers hineinruft, so tönen sie den Vokal wieder, aber nicht die Vokale Ä, E, I.“ HELMHOLTZENS Klavier war schon gefügiger, nur das I gelang „weniger

gut“ (S. 105). Doch dürfte HELMHOLTZ hier etwas nachsichtig geurteilt haben. Denn selbst über einen modernen Blüthner-Konzertflügel ersten Ranges (mit 4 gleichgestimmten Saiten für jeden Ton von  $as^2$  an) berichtet mir Herr A. KREICHGAUER: „A sehr gut, E und I unkenntlich und ununterscheidbar; die übrigen Vokale (auch Ä, Ö, Ü wurden untersucht) einigermaßen, doch keiner außer A so, daß ein völlig unbeeinflußter Beobachter, der den gesungenen Vokal nicht vorher hört und den Zusammenhang kennt, die Klavierwiedergabe für einen bestimmten Vokal halten würde . . . Er wird, befragt, vielleicht ebenso leicht auf Musikinstrumente tippen als auf Vokale.“

2. Bedeutsamer als das Nachklingen der ganzen Vokale ist für unsere Zwecke die Feststellung der einzelnen mitklingenden Saiten und die Stärke ihres Mitklingsens. Von solchen Versuchen scheinen HELMHOLTZENS Analysen überhaupt ausgegangen zu sein, da er sie allein in seinem Schreiben an DONDERS 1857 erwähnt, um darauf die These zu stützen, daß die Vokale sich durch ihre Obertöne unterscheiden. A enthält danach neben dem 1. (Grundton) deutlich den 3. und 5., schwächer den 2., 4., 7. O den 3. etwas schwächer als bei A, sehr schwach den 2. und 5. U fast nur den Grundton, schwach den 3. E sehr kräftig den 2., die höheren kaum hörbar. I schwach den 1., daneben den 2. und 3., schwach auch den 5.

Leider gibt HELMHOLTZ die Höhe des gesungenen Grundtones selbst nicht an. Es hat fast den Anschein, als habe er damals noch auf dem Standpunkt der Relativtheorie gestanden. Lag aber der Grundton, wie es wahrscheinlich ist, in der unteren Hälfte der kleinen Oktave, so lassen sich die Ergebnisse nach dem Obigen wohl verstehen; bei E und I resonierten eben nur die unteren Maxima. In der „Lehre von den Tonempfindungen“ erwähnt er die Methode S. 65, sieht aber von näheren Angaben ab. In der Tat reicht dieses Verfahren zur systematischen Analyse bei weitem nicht aus. Doch seien einige Versuche erwähnt, die jeder leicht nachprüfen kann.

Man hört die Teiltöne dabei leichter als in dem natürlichen Vokal heraus, weil die Saiten in verschiedenem Tempo abklingen und ungleich lokalisiert sind. Man muß nach aufgehobener Dämpfung die tiefen, unterhalb des gesungenen Tones liegenden Saiten durch ein Kissen bedecken, weil sich sonst darin zahllose Partialschwingungen bilden, die untereinander ein Geräusch geben, das namentlich beim A störend wird. Wenn dann zunächst die Stärke des Mitschwingens für den Grundton festgestellt ist, bedecke man auch dessen Saite mit der Hand, um den nächsthöheren Teilton deutlicher zu hören, usf. Das Dämpfungspedal muß vorsichtig, langsam niedergetreten werden, weil sonst wieder ein langnachhallendes Geräusch entsteht. Man macht die Beobachtungen am besten in der Nacht bei völliger äußerer Stille.

Ich habe so an einem älteren Flügel von Ritmüller die Vokale U, O, A, I (zwischen E und I sind auf diesem Wege keine deutlichen Unterschiede zu erwarten) auf verschiedenen Grundtönen untersucht und folgende Teiltöne beobachtet:

1. Grundton A (der großen Oktave). U: Grundton und  $e^1$  stark. O: Grundton weniger stark,  $e^1$  schwach,  $a^1$  kräftig. A: Grundton hörbar, Obertöne undeutliche Masse, doch  $e^2$  und  $g^2$  erkennbar. I: Grundton und  $e^1$  deutlich; von hohen Obertönen nichts erkennbar.

2. Grundton c. U: c sehr stark. O:  $g^1$  merklich, auch  $e^2$  scheint vorhanden. A: c schwach,  $b^2$  hervortretend. I: c stark. Hohe Töne?

3. Grundton *d*. U: *d* stark, auch  $a^1$ . O: *d* stark,  $a^1$  stärker. A: *d* schwach, die folgenden Obertöne, auch noch der 7. ( $c^3$ ), merklich. I: *d* stärker, unter den Obertönen höchstens ein schwaches  $d^4$ .

4. Grundton *f*. U: *f* und besonders stark  $f^1$ . O: *f* stärker, dazu  $c^2$ . A: wieder eine Menge Teiltöne, *f* und  $f^1$  ziemlich gut, außerdem  $a^2$  und  $c^3$  hervortretend. I: *f* vorhanden,  $f^1$  stärker. Hohe Teiltöne?

5. Grundton *a*. U: *a* stark,  $e^2$  ganz schwach. O: *a* schwächer,  $a^1$  stärker,  $e^2$  ganz schwach. A: Geräusch mit vielen Tönen, erkennbar *a*,  $a^2$ ,  $g^3 + a^3$  (diese nicht recht trennbar, schwebend),  $h^3$ . I: *a*, außerdem etwa schwaches  $d^4$ .

6. Grundton  $d^1$ . U:  $d^1$  stark, sonst nichts. O:  $d^1$ ,  $d^2$ . A: Geräusch mit vielen Tönen, darin  $a^2$ ,  $c^3$  kenntlich, ein anderes Mal auch  $e^3$ ,  $h^3$ . I:  $d^3$  und deutliches  $d^4$ , ein anderes Mal auch  $fis^4$ .

Weitere Beobachtungen zeigten, daß beim A immer Töne der 2-gestrichenen Oktave (verschieden je nach dem Grundton) besonders verstärkt waren.

Neuerdings hat auf meinen Wunsch Herr A. KREICHGAUER an dem obenerwähnten Konzertflügel die Versuche auf den Tönen *A*, *c*, *d* wiederholt. Der Hauptunterschied war, daß beim Vokal A auf dem Ton *A* auch der Grundton sehr stark war (vielleicht infolge zufällig besonders starker Tongebung) und daß hier noch über  $g^2$  hinaus schwache Teiltöne bis  $dis^3$  erklangen. Auch auf dem Tone *c* reichten die Teiltöne des A bis  $d^3$ , auf *d* sogar bis  $fis^3$ . Beim U machte sich noch das 2. Maximum um  $g^2$  ( $a^2$ ) geltend. Aber bei den hellen Vokalen (hier wurden auch die Umlaute geprüft) fand sich trotz der 4 Saiten der hohen Region auch hier nur das untere Maximum, höchstens ein ganz schwaches isoliertes  $e^3$  bei Ä und Ö auf *d*.

Das Klavier hat außer seiner Allgegenwart als Saloninstrument auch den Vorteil, daß man jeden beliebigen Grundton der chromatischen Skala wählen und so vor allem die annähernde Konstanz der Maxima in ihrer absoluten Höhenlage prüfen kann, wenigstens bei den dunkleren Vokalen. In dieser Hinsicht wird man die aus unseren Resonanztabellen ersichtliche und im folgenden weiter zu besprechende Gesetzmäßigkeit daran bestätigt finden; ebenso, daß niemals eine Saite zum Mitschwingen zu bringen ist, die nicht einem harmonischen Oberton unmittelbar nahe liegt (so auch KREICHGAUER). Die kleinen, durch die temperierte Stimmung verursachten Abweichungen liegen noch in der Resonanzbreite der Saiten. Immerhin bedingen sie Modifikationen in der Stärke des Mitschwingens. Überhaupt würde ich aus Klavierversuchen allein diesen Schluß auf nur harmonische Teiltöne noch nicht mit solcher Bestimmtheit ziehen.

STEFANINI beschreibt (1914) Klavierversuche, bei denen er einzelne Saiten auf ihr Mitschwingen untersuchte, indem er sie mit einem Mikrophon verband und den Ton durch ein Telephon abhörte. So antwortete die Saite  $gis^3 = 1600$  Schwingungen, wenn der Vokal U auf  $es = 155$  Schwingungen gesungen wurde. Dies wäre ein Ton, der 50 Schwingungen über dem 10. harmonischen Teilton läge. Dieselbe Saite antwortete bei O, A, E, I, wenn sie auf 170 Schwingungen gesungen wurden; was einem Ton entspräche, der 70 Schwingungen über dem 9. harmonischen Teilton läge. STEFANINI schließt daraus wieder auf unharmonische Obertöne. Aber wenn etwa im 1. Falle der Sänger auch nur um 5 Schwingungen, im 2. Fall um 9 Schwingungen zu hoch gesungen hat, was bei den besten Sängern

vorkommen kann<sup>1)</sup>, würde es auch mit den harmonischen Teiltönen genau stimmen. Außerdem kommt die Resonanzbreite der Saite in Betracht. Es wäre also erst durch kleine Höhenverschiebungen des gesungenen Tones der Punkt stärkster Resonanz der Saite aufzusuchen, ehe man auf unharmonische Teiltöne schließen könnte. Im übrigen ist das Vorhandensein eines so hohen Teiltones in den dunklen Vokalen bemerkenswert (s. o. S. 26). Vielleicht würde sich doch die weitere Ausbildung dieser Methode, wenigstens für gewisse Fragepunkte, lohnen.

Wer eine Harfe oder Zither zur Hand hat, kann natürlich auch da ähnliche Beobachtungen wie am Klavier machen. Hier lassen sich die Saiten auch sehr leicht in reine Stimmung zu einem gewählten Grundton bringen. Die hellen Vokale bzw. deren obere Maxima kommen auch bei der Zither nicht heraus, recht gut aber O, A und deren untere Maxima. Selbst die Violine kann zu gewissen Proben auf unsere Resonanzversuche dienen. Dämpft man die 3 oberen Saiten mit dem Finger und singt nun auf dem Tone  $g$  den Vokal A, so klingt  $g^2$  nach. Bei O klingt  $g^1$ , ebenso bei E (1. Maximum). Bei U  $g^2$  (2. Maximum). Diese Töne entstehen durch Partialschwingungen der  $g$ -Saite. Die Saite als Ganzes brachte ich durch Singen ihres Grundtones nicht zum Mitschwingen, auch nicht mit U. Mit I war überhaupt nichts zu erreichen, weil eben  $g$  nicht erregt wird, das obere Maximum aber schon zu hoch liegt.

<sup>2)</sup> Eine geübte Sängerin machte beim Nachsingen nach SOKOLOWSKYS Messungen (NADOLECZNY S. 653) Fehler bis zu 3,52%. Dies wären im 1. Falle bis  $5\frac{1}{2}$ , im 2. bis 6 Schwingungen. Es kommen aber noch größere Fehler vor; vgl. KLÜNDER im Arch. f. Anat. u. Physiol. 1879. Die Kurve eines routinierten Sängers zeigte derartige Schwankungen, daß sie zur Messung gar nicht zu brauchen war. Hierüber wären also in den obigen Fällen zunächst Feststellungen erforderlich gewesen.

## 2. Kapitel.

### Analyse gesungener Vokale durch Interferenzröhren<sup>1)</sup>.

Bekanntlich läßt sich eine Schallwelle dadurch vernichten, daß man sie durch eine Röhre mit einer Zweigleitung schickt, aus der sie mit einer Verlängerung ihres Weges um eine halbe Wellenlänge zur Hauptleitung zurückkehrt. Dann kommen Berg und Tal zusammen und heben sich gegenseitig auf. So kann man nun auch Teiltöne eines Klanges vernichten. GRÜTZNER benützte dieses Prinzip 1891 zur Vokalanalyse und gab zugleich der Zweigleitung die bequemste Form, indem an einer Stelle der Hauptleitung eine auf  $\frac{1}{4}$ -Wellenlänge des auszuschließenden Tones eingestellte Seitenröhre rechtwinklig angesetzt wurde, so daß der Ton, darin hin- und zurücklaufend, um eine halbe Wellenlänge verschoben wurde. GRÜTZNERS Analysen wurden durch SAUBERSCHWARZ 1895 fortgesetzt. Aber beide Forscher führten die Methode nicht systematisch genug durch. Sie beschränkten sich fast nur auf die Auslöschung des Grundtones und der damals angenommenen „charakteristischen Töne“. Später hat W. KOEHLER in seinen Vokalstudien davon Gebrauch gemacht und wesentliche Schlüsse gegen die HERMANNSCHE Theorie daraus gezogen. Gewisse Schwierigkeiten aber, die dieser Methode anhaften, mögen ihre allgemeinere Anwendung und konsequente Durchführung verhindert haben. Im folgenden sollen zunächst diese Schwierigkeiten besprochen werden. Gerade sie führten zu einer Ausgestaltung des Verfahrens, wodurch es für die Erkenntnis der Vokalstrukturen in hohem Maße fruchtbar wird<sup>2)</sup>.

<sup>1)</sup> „Interferenz“ wird im folgenden zumeist durch „If.“ wiedergegeben.

<sup>2)</sup> Zur Herstellung einfacher Töne habe ich bereits 1896 die If.-Methode angewandt und empfohlen und später namentlich in der Untersuchung über Kombinationstöne (1903–1909) davon systematischen Gebrauch gemacht. If.-Einrichtungen gehören zu den unentbehrlichsten Bestandteilen eines akustischen Instituts.

## I. Prinzipielle Schwierigkeiten und deren Lösung.

### 1. Gleichzeitige Auslöschung der ungeradzahigen Multipla.

(Ab- und Aufbaureihen, Lücken-, Stich- und Isolierversuche.)

Mit einem bestimmten Tone von der Schwingungszahl  $n$  müssen theoretisch auch die Schwingungen  $3n, 5n \dots$  ausgelöscht werden, weil auch bei diesen entgegengesetzte Phasen des Wellenzuges zusammentreffen. Hat man also beim Ausschluß eines Teiltone  $n$  eine bestimmte Veränderung des Vokalklanges bemerkt, so kann sie zunächst ebensogut von dem gleichzeitigen Ausschluß des Teiltone  $3n, 5n$  usw. herrühren. Nun trifft freilich die theoretische Folgerung so allgemein in Wirklichkeit nicht zu (s. u.). Aber eindeutige Schlüsse erscheinen eben zunächst doch unmöglich.

Dieser Schwierigkeit wurde in der Weise vorgebeugt, daß der Vokal von oben herab systematisch durch Einfügung immer längerer Seitenleitungen abgebaut wurde, bis endlich nur der Grundton selbst als einfacher Ton übrigblieb<sup>1)</sup>. Die dabei eintretenden Klangveränderungen wurden Schritt für Schritt beobachtet. Dann wurde der Laut auf dem umgekehrten Wege, durch Einschleiben der Röhren, wieder aufgebaut und die Veränderungen in umgekehrter Folge beobachtet.

Man braucht dazu natürlich ein System vieler Seitenröhren und muß zunächst, solange man noch nicht sicher weiß, ob außer den harmonischen nicht auch unharmonische Teiltöne vorhanden sind, in möglichst kleinen Schritten vorgehen. Aber bald bestätigte sich auch auf diesem Wege, daß unharmonische nicht da waren<sup>2)</sup>.

<sup>1)</sup> Immer noch findet man die Behauptung, es sei noch nicht gelungen oder überhaupt unmöglich, völlig einfache Töne herzustellen (z. B. bei ZIEHEN, Leitfaden der Physiolog. Psychologie, 12. Aufl. 1924, S. 151). Dies ist aber seit Jahrzehnten auf dem angegebenen Wege geschehen. Im übrigen braucht man auch nur zu pfeifen: denn Pfeiftöne sind tatsächlich einfache Töne (s. 6. Kap.).

<sup>2)</sup> KOEHLER fand bereits, daß beim Ausschluß aller harmonischen Teiltöne überhaupt nichts übrigblieb, und schloß daraus, daß unharmonische nicht vorhanden seien. S. GARTEN bestreitet diesen Schluß mit Hinweis auf die Hf.-Breite, vermöge deren mit einem harmonischen auch ein benachbarter unharmonischer Ton vernichtet werden könne. An sich ist dies auch richtig, wenn sie einander nahe genug liegen. Aber da bei den zahllosen Einzelversuchen und Versuchsmodifikationen, die meine Untersuchungen mit sich brachten, niemals ein solcher Ton eindeutig aufgetreten ist, auch nicht innerhalb eines großen Zwischenraums zwischen harmonischen Tönen, so wird man meine zuversichtliche Stellungnahme verstehen.

und nun konnten von vornherein die Röhreneinstellungen auf harmonische beschränkt und damit Röhren gespart werden.

Da bei solchem Vorgehen stets sämtliche über einem gewissen Punkte der Tonlinie liegenden Teiltöne ausgeschlossen bleiben, so kann eine beobachtete Klangveränderung nur auf die zuletzt eingefügte bzw. abgestellte Zweigleitung bezogen werden. Die Deutung auf ungeradzahlige Multipla ist ausgeschlossen.

Beim Ab- und Aufbau kann man Längs- oder Querschnitte machen, d. h. entweder einen Vokal nach dem anderen vollständig ab- und aufbauen, oder bei jeder Einstellung sämtliche oder mehrere Vokale durchprüfen und dann erst zur nächsten Einstellung übergehen. Meistens wurden Längsschnitte gemacht, Querschnitte aber vielfach zur Kontrolle eingefügt. Das Querschnittsverfahren für sich allein eignet sich besonders zu unwissentlichen Versuchen, d. h. solchen, bei denen der Hörende nicht weiß, um welchen Vokal es sich handelt. Es werden dann die untersuchten Vokale in beständig unregelmäßig wechselnder Reihenfolge gegeben.

Außer diesen Ab- und Aufbauversuchen lassen sich unter gewissen Voraussetzungen auch Lücken- und Stichversuche mit eindeutigen Erfolg anstellen, bei denen nur eine bestimmte Zone oder einzelne Teiltöne mitten aus dem Klangkomplex herausgenommen werden. Man kann dies tun, sobald sich auf dem obigen Wege herausgestellt hat, daß die oberhalb einer gewissen Höhengrenze liegenden Töne keinen Einfluß mehr auf die Natur eines Vokals haben. Denn dann kann man unterhalb dieser Grenze bis zu einer Duodezime (1:3) beliebig Töne herausnehmen, deren ungerade Vielfache in die darüberliegende einflußlose Region fallen.

In gewissen Fällen lassen sich auch umgekehrt alle Töne außer einem einzigen ausschalten, so daß man dessen Vorhandensein und relative Stärke feststellen kann (Isolierversuche). Eine solche vollkommene Isolierung durch If. ist aber nur möglich beim Grundton und seinen Oktaven (Oktavenversuche), weil nur die Potenzen von 2 die Eigenschaft haben, sich nicht als ungerade Vielfache irgendeiner anderen Zahl (1 inbegriffen) darstellen zu lassen.

Bei diesen Oktavenversuchen wird man, um die volle Stärke des Teiltöns zu erhalten, in Fällen, wo die bezügliche Oktave zwischen naheliegende Teiltöne fällt (z. B. 8 zwischen 7 und 9), auch diese freigeben, weil infolge der If.-Breite der Ausschluß eines Tons auch die Nachbartöne schädigt. Man wird dann gleichwohl nur den gewünschten Ton selbst hören, nicht seine Nachbarn, die wieder von benachbarten Interferenzen geschwächt sind.

Aus der im Text erwähnten Zahlengesetzlichkeit folgt auch, daß, wenn man nur den Grundton und seine Oktave ausschließt, sämtliche Teiltöne, theoretisch wenigstens, mit ausgeschlossen sind, also der ganze Klang vernichtet werden muß.

Man kann so z. B. jeden Bassisten mit Bruststimme das hohe  $c^3$  singen lassen. Er braucht nur den Vokal A auf  $c$  zu singen, während in der Leitung alle Teiltöne außer  $c^3$  ausgeschaltet sind. Am Ende der Leitung, im Beobachtungszimmer, hört man dann nur ein kräftiges  $c^3$  — ein sehr ergötzlicher Versuch. Wenn er aber U singt, hört man unter denselben Umständen nichts, weil  $c^3$  eben nicht darin enthalten ist.

Ist der Grundton nicht zu stark, so kann man immerhin in Verbindung mit ihm auch den 3., 5. oder 7. Teilton durch Vernichtung aller übrigen so gut wie isoliert zu Gehör bringen, da ein schwacher Grundton die Beobachtung und Stärkeschätzung so weit davon entfernter Töne nicht nennenswert beeinträchtigt. Analog auch den 6., 10., 14., wenn man den 2. daneben bestehen läßt, usf.

## 2. Verstärkung der geradzahligen Multipla.

Nach AUERBACH und GARTEN<sup>1)</sup> können durch If.-Einstellung auf einen Ton seine geraden Multipla infolge der Reflexion der Schwingungen in den Seitenröhren verstärkt, ja es kann nach GARTEN auf diesem Weg eine gar nicht im Klange vorhandene höhere Oktave erzeugt werden.

Dies mag nun bei der Prüfung auf objektivem Wege unter besonderen Umständen der Fall sein. Für das Ohr habe ich es und hat es auch Prof. v. HORNBOSTEL niemals bestätigt gefunden. Wenn ich z. B.  $c^2$  als einfachen Ton herstelle, es dann auf möglichste Schwäche bringe (am einfachsten wieder durch annähernde If.), sodann If.-Röhren auf die nicht vorhandenen Töne  $c^1$ ,  $c$ ,  $F$  einstelle, so wird dadurch für mein Ohr  $c^2$  nicht im geringsten stärker; noch weniger wird es, wenn es ganz ausgelöscht war, dadurch wiedererzeugt, auch nicht, wenn ich diese Einstellungen alle auf einmal und jede mit mehreren Röhren vornehme.

Auch dieses physikalische Bedenken braucht uns also nicht zu beunruhigen. Alles kommt ja doch hier gerade auf das menschliche Ohr an. Übrigens hat K. LEWIN (2, S. 328) auch bei seinen physikalischen Messungen keine solche Verstärkung gefunden.

Ich leugne nicht, daß auch in meinen Versuchen zuweilen Paradoxien vorkamen. Gewisse Einstellungen hatten unerwartete, ja zunächst unerklärbare Wirkungen. Ähnliches erwähnt LEWIN. Ein Physiker äußerte: „Die Akustik ist das seltsamste Gebiet. Da kommt alles anders als man berechnet hat.“ Und wenn sich noch alles berechnen ließe! Sehr unbedeutende, unvorhergesehene Faktoren können eine Rolle spielen, Reflexions- und Resonanzwirkungen, vielleicht sogar solche zwischen den einzelnen Mün-

<sup>1)</sup> AUERBACH 5, S. 598; GARTEN 3, Nr. VII, S. 20ff.



dungsstellen der Seitenröhren, usf. Einige Faktoren (Bauch und Knoten, Röhrenweite) werden wir noch besprechen. Aber im großen und ganzen sind die Ergebnisse meiner zahlreichen und mit vielen Variationen angestellten Versuchsreihen so durchsichtig und übereinstimmend, daß ich einen prinzipiellen Fehler nicht besorge. Wo wirklich ein Verdacht sich ergab, ist es im folgenden erwähnt.

### 3. Schwächung des Grundtons.

Rein physikalisch betrachtet, muß ein Ton durch jede beliebig lange Seitenröhre mehr oder weniger geschwächt werden, ausgenommen nur, wenn ihre Länge genau mit einer halben Wellenlänge oder einem ganzzahligen Vielfachen davon zusammenfällt, weil dann die Verzögerung eine oder mehrere ganze Wellen ausmacht. Durch Einstellung von Seitenröhren auf  $\frac{1}{4}$ -Wellenlängen höherer Töne muß daher ein tieferer ausnahmslos geschwächt werden, da ja die bezüglichen Seitenleitungen weniger als  $\frac{1}{4}$  seiner Wellenlänge haben. Die Einstellung auf Obertöne muß daher immer den Grundton schwächen, zumal wenn auf zahlreiche Obertöne zugleich eingestellt wird.

So liegen die Dinge mathematisch. So haben sie sich diesmal auch physikalisch bei den LEWINSchen Intensitätsmessungen bestätigt gefunden: wurden zahlreiche If.-Röhren auf die Obertöne eines bereits obertonfreien Tones eingestellt, so konnte er dadurch um  $\frac{1}{4}$ — $\frac{1}{3}$  physikalisch geschwächt werden. Aber eine andere Frage ist es, wie weit eine solche Schwächung des Grundtones noch durch ein geübtes menschliches Ohr wahrgenommen werden kann. Hier ist das Ohr nun doch im Nachteil gegen eine so empfindliche, auf den bezüglichen Ton ausschließlich abgestimmte Membran. Ich habe mit If. einfache Töne verschiedener Höhe hergestellt (namentlich Töne weiter Flaschen, die ohnedies nur ganz schwache, durch wenige Röhren auszuschaltende Obertöne haben) und dann jedesmal Seitenröhren in zunehmender Anzahl (bis zu 50) auf das darüber liegende Tongebiet eingestellt, dabei aber keine irgend erhebliche Schwächung bemerkt. Ein Ton von der Stärke 2 hatte dann noch mindestens  $1\frac{3}{4}$ , konnte aber auch noch ebensogut mit 2 taxiert werden.

Dieser Nachteil des Ohres gegenüber einer auf einen Ton speziell abgestimmten Membran ist aber für unsere Untersuchung ein Vorteil. Denn wir brauchen nun nicht zu besorgen, daß die Stärkeschätzungen in bezug auf einen durch If. isolierten Ton, insbesondere den Grundton, erheblich andere werden, als wenn wir den Ton irgendwie ohne das Mittel der If.-Einrichtungen isoliert hätten.

Auch die bloße Wegverlängerung eines Teiles der Energie in den Seitenröhren reicht in unserem Falle nicht zu einer wahrnehm-

baren Schwächung. Wenn man die Hauptleitung um ebensoviel verlängert, als die sämtlichen Seitenleitungen zusammen ausmachen, läßt sich dies leicht bestätigen.

#### 4. Unterschied der Bauch- und Knotenwirkung.

Es macht einen Unterschied, an welchem Orte der in der Hauptleitung vorhandenen stehenden Wellen man eine If.-Röhre einschaltet. Die Wirkung ist naturgemäß bei den Bäuchen größer. Darauf hat MAX MEYER zuerst aufmerksam gemacht<sup>1)</sup>. Er betonte es für die ältere If.-Methode (Zweigleitungen von  $\frac{1}{2}$  Wellenlänge). Bei der hier angewandten Methode der einfachen Seitenröhren ist der Unterschied kürzlich von Frl. EBERHARDT und Dr. LEWIN physikalisch gemessen worden. Die Amplituden des Tones  $e^2$  verhielten sich bei Bauch- und Knoten-Interferenz wie 38 : 202, die Intensitäten also wie die Quadrate dieser Zahlen. In meinen subjektiven Intensitätsstufen ausgedrückt, fand sich der Ton, der ohne If. mit der Stärke  $2\frac{1}{2}$  aus der Leitung kam, bei Bauch-If. auf  $\frac{1}{8}$ , bei Knoten-If. auf  $\frac{3}{4}$  vermindert. Immerhin ist auch diese letzte Verminderung schon sehr beträchtlich, und in beiden Fällen wurde der Ton durch Hinzufügung einer zweiten gleich langen Seitenleitung an beliebiger Stelle vollends vernichtet.

Bei Ab- und Aufbauversuchen ist es nun völlig undurchführbar, jedesmal und für jeden auszuschließenden Teilton einen Bauch auszusuchen. Aber es ist auch nicht nötig, wenn nur für kräftigere Teiltöne stets 2 oder mehr Röhren angewandt werden. Eine darunter liegt dann eben dem Bauch näher<sup>2)</sup>. Auch wirken ja zufolge der If.-Breite die Einstellungen auf die benachbarten Teiltöne bei Tönen höherer Ordnungszahl noch mit. Nur bei Stichversuchen könnte der Unterschied Bauch—Knoten einmal von Bedeutung werden. Doch begnügte ich mich auch da damit, mehrere Röhren auszuziehen, und glaube nicht, daß wesentliche Ungenauigkeiten dadurch in die Ergebnisse gekommen sind.

Daß der Unterschied Bauch—Knoten sich nicht stärker geltend macht, wird wohl darauf beruhen, daß außer den stehenden Wellen

<sup>1)</sup> Zeitschr. f. Psychol. Bd. 11, S. 191. 1896.

<sup>2)</sup> KOEHLER fand es in seinen If.-Versuchen nützlich, bei Einstellung mehrerer Seitenröhren eines Röhrensystems auf denselben Ton die ausziehenden durch eine nichtgebrauchte zu trennen. Bei meinen Einrichtungen, wo ich namentlich bei den Flüstervokalen auf diesen Punkt achtete, konnte ich keinen Unterschied in der Wirkung finden. Aber bei sehr dichter Anordnung der Röhren kann es damit wohl seine Richtigkeit haben. Es dürfte auf dem Umstande beruhen, daß auf diese Weise mehr Chance besteht, einen Wellenbauch zu treffen. Natürlich kommt es auch auf die jeweilige Tonhöhe an.

auch fortschreitende durch die Hauptröhre gehen. Daß dies der Fall ist, geht schon daraus hervor, daß man den Ton fast ebenso gut aus der Leitung hört, wenn das Ohr sich an einem Knoten als wenn es sich an einem Bauch befindet. Befestigt man am Röhrende einen ausziehbaren Ansatz, den man langsam verlängert, so hört man bei gespannter Aufmerksamkeit allerdings ein minimales An- und Abschwellen an gewissen Stellen, aber von einem Verschwinden ist nicht die Rede.

##### 5. Veränderung des Klanges durch die Hauptleitung.

Laute, die durch eine längere Röhrenleitung dem Ohre zugeführt werden, sind gewissen Alterationen ausgesetzt, schon ehe man die seitlichen If.-Einstellungen vornimmt. Besonders werden die tieferen Bestandteile durch die Reibung an den Röhrenwänden geschwächt; dadurch wird der Klang heller, schärfer als in freier Luft. Ein Instrument kann so seinen ganzen Charakter ändern. Bei den Vokalen hatte ich zwischendurch eine Anzahl besonders sorgfältig angestellter Versuche mit sehr langer Leitung gemacht, um zum Zweck „unwissentlicher“ Beobachtungen jedes direkte Herüberdringen der von einem vorzüglichen Sänger stark angegebenen Vokallaute auszuschließen (s. u. III, 3, Schluß). Die Folge war aber, daß die Formanten der helleren Vokale sich ein wenig in die Höhe zogen. Das heißt: die Vokale müssen ein wenig heller geklungen haben, obgleich diese Veränderungen nicht auffällig und die Vokale durchaus noch gut kenntlich waren. Aber auch bei kürzerer Hauptleitung kommen solche Alterationen vor, und manchmal in recht auffälliger Weise.

Es ist daher vor jeder Versuchsreihe genau darauf zu achten, ob die Laute natürlich und unentstellt aus der Röhrenöffnung herausklingen. Wenn dies nicht der Fall ist, muß zunächst die Leitung möglichst verkürzt werden. Bei Versuchen, in denen der Beobachter weiß, um welchen Vokal es sich handelt, schadet es nicht wesentlich, wenn der Laut auch direkt etwas in das Beobachtungszimmer dringt. Auch kann man das nichtgebrauchte Ohr verstopfen. Sodann kann im Schallzimmer durch Trichter nachgeholfen werden, die durch ihre Resonanz bestimmte Teiltongruppen verstärken. Bei meinen Versuchen wurden nach Bedarf Trichter verschiedenster Größe, von einem mächtigen Grammophontrichter (59 cm Durchm.) bis zu winzig kleinen, angewandt. Sie sind aber nur im Notfall und mit Vorsicht zu gebrauchen, sonst können sie den Klang noch mehr verändern als die freie Röhrenleitung. Immer muß das Ohr des Beobachters entscheiden, ob die Vokale deutlich unterscheidbar herauskommen. Steht dies nicht

völlig außer Zweifel, so sind Reihen unwissentlicher Erkennungsversuche den If.-Reihen vorzuschicken, um festzustellen, ob die aus der Leitung kommenden Vokale eine einwandfreie naturgetreue Beschaffenheit haben. Dies ist der Fall, wenn ausschließlich richtige Urteile erfolgen<sup>1)</sup>.

Wichtig ist auch, daß die Leitung keine scharfen Ränder und keine unvermittelten Übergänge zwischen Röhrenstücken von größerer zu solchen kleinerer Weite enthalte, da sonst durch Reflexion und Resonanz Veränderungen des Klanges entstehen. Die Übergänge müssen daher immer konisch geformt sein. Auch wende man möglichst wenig Schlaucheinsätze an, da sie den Ton viel mehr als Metallröhren schwächen, was namentlich bei den Flüsterlauten in Betracht kommt.

Endlich ist auch das Verfahren beim Abhören des aus der Leitung kommenden Lautes nicht unwesentlich. Man kann sich dabei einer Olive oder eines Schlauchstückchens bedienen, die mit der Leitungsröhre verbunden werden, oder das Ohr direkt an diese selbst halten. Letzteres wird in den meisten Fällen vorzuziehen sein. Hauptsache ist aber, daß man das Leitungsende, welches es auch sei, nicht in unmittelbare Berührung mit den Gehörränden bringe, da hierdurch störende Resonanzwirkungen entstehen.

#### 6. Interferenzbreite.

Wie die Resonanz, so hat auch die If. eine gewisse Breite. Sie erstreckt sich in abnehmendem Grade auf die Nachbarn des direkt ausgeschlossenen Tones. Besondere Versuche hierüber ergaben, daß (wenigstens bei den von mir benutzten Apparaten) eine merkliche Schwächung in den mittleren Lagen des Tonreiches nur etwa auf eine kleine Terz nach unten und oben sich erstreckt. In höherer Lage wächst sie etwas und kann von  $c^4$  ab etwa auf eine große Terz, von  $c^6$  ab auf eine Quart, bei  $c^7$  auf eine halbe Oktave angesetzt werden. Man kann diese Versuche mit einzelnen einfachen Tönen anstellen, sie aber auch bei der Analyse von Vokal- oder Instrumentalklängen kontrollieren. Es zeigt sich regelmäßig bei den Aufbauversuchen, daß ein durch Zurückstellung eines Röhrenstempels freigegebener Ton erst dann seine volle Stärke erlangt, wenn auch noch die weiter nach oben folgende Region des Klanges im Umfang einer kleinen Terz usw. freigegeben ist.

Die If.-Breite wächst, wie es scheint, mit abnehmender Röhrenweite, hängt aber sicher auch noch von anderen Umständen, namentlich der Abdichtung der Stempel, ab. Die Stärke der Töne scheint insofern keinen Unterschied zu machen, als die Zone der Schwächung dadurch

<sup>1)</sup> So glaube ich dem an sich nicht unberechtigten Bedenken S. GARTENS betreffs der Vorschaltung von Trichtern (3, VII, S. 15) zuvorgekommen zu sein. Eingehend hat D. C. MILLER auf physikalischem Wege die Wirkung von Trichtern untersucht.

nicht verändert wird (wohl wegen größerer Unterschiedsempfindlichkeit für schwächere Töne). Dagegen ist die Zone des Verschwindens natürlich für sehr schwache Töne breiter als für starke. Für diese bildet sie einen Ausschnitt aus der Schwächungszone, für die ganz schwachen dagegen fällt sie mit dieser zusammen, da ein solcher eben durch jede Schwächung schon verschwindet. Dieser Unterschied der Vernichtungsbreite zwischen starken und schwachen Tönen wird uns bei den Flüsterlauten wichtig werden.

Die so festgestellte If.-Breite ist bei der Ausdeutung der bei If.-Versuchen beobachteten Klangveränderungen immer im Auge zu behalten. Man hat z. B. bei Abbauprobungen damit zu rechnen, daß der Klang nicht bloß bis zu der der letzten Röhreneinstellung entsprechenden Tongrenze vernichtet, sondern noch eine kurze Strecke weiter hinab geschwächt ist. Bei Stichversuchen ist er nach beiden Seiten geschwächt.

## II. Weiteres zur Technik und Methodik.

### 1. Leitung und Röhrensysteme.

In dem folgenden Leitungsschema bedeuten die römischen Ziffern Zimmer des (alten) Berliner Psychologischen Instituts, die teilweise durch einen längeren Korridor getrennt sind. Die If.-Leitung *J* beginnt in *I* bei *S*<sub>1</sub> und endet in *V* bei *B*<sub>1</sub>. Die von *P* bis *V*<sub>*p*</sub> reichende synthetische Einrichtung bleibt hier außer Betracht. Die gewöhnlich benutzte kürzere If.-Leitung beginnt bei *S*<sub>2</sub>, die für Flüsterlaute erst bei *FL*. Die in die Hauptleitung einfügbaren If.-Systeme mit Seitenröhren sind auf *IV* und *V* verteilt.

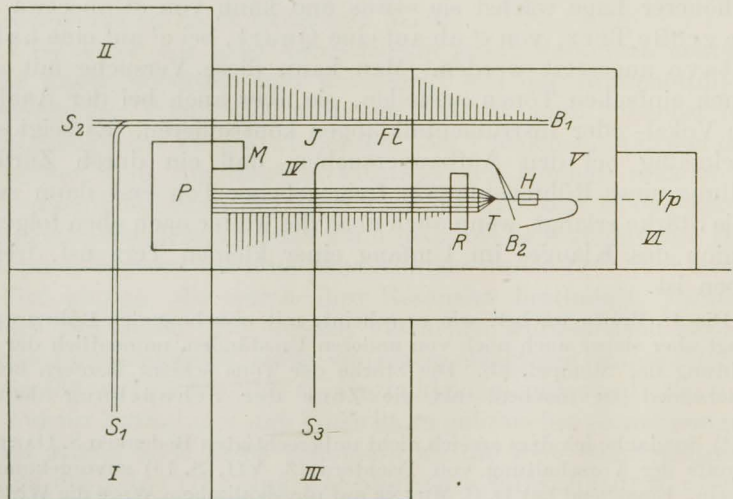


Abb. 1. Leitungsschema.

Es wurden dazu (außer Röhren von allergeringster Weite für die höchste Tonregion, deren Stempel nur die Dicke von Stricknadeln hatten, die aber bei den Hauptversuchen nicht mehr zur Anwendung kamen) 9 Röhrensysteme mit insgesamt 71 Seitenröhren, die mit ausziehbaren Stempeln versehen waren, benützt. Der Stempelabschluß war kreisrund und gut abgedichtet. Bei den 4 großen Systemen mit zusammen 39 Röhren war die Hauptleitung 1,8—2,1 cm, die Nebenleitung 1,8 cm weit, die Stempel teils 124, teils 54 cm lang. Bei 2 kleineren Systemen mit zusammen 20 Röhren waren Haupt- und Nebenleitungen 1 cm weit, die Stempel 18 cm lang, bei dem kleinsten mit 12 Röhren waren die Leitungen 0,5 cm weit, die Stempel 12 cm lang. Alle Stempel waren graduiert, die längeren (der weiteren Röhren) in  $\frac{1}{2}$  cm-, die kürzeren in mm-Skalen. Röhren von 2,1 cm Weite kann man noch bis  $c^5$  gut verwenden, die engeren am besten von  $c^4$  an. Die Zahl der Röhren zur vollständigen Auslöschung eines Vokals muß darum so groß sein, weil für jeden einigermaßen kräftigen Teilton 2—3 Röhren erforderlich sind. Er kommt aus der Seitenleitung nicht bloß mit entgegengesetzter Phase, sondern auch etwas geschwächt zurück, so daß er die stärkere Hauptwelle nicht sogleich ganz vernichtet. Außerdem ist die Vermehrung nützlich, um nicht immer erst die Bäuche aufsuchen zu müssen.

## 2. Einfluß der Röhrenweite auf die erforderliche Einstellung.

Die Wellenlänge eines Tones ist bekanntlich gleich der Raumbstrecke in Metern, die er in einer Sekunde durchläuft (Fortpflanzungsgeschwindigkeit), dividiert durch die Anzahl der Schwingungen in der Sekunde. Die Fortpflanzungsgeschwindigkeit bei  $0^\circ\text{C}$  in freier trockener Luft ist für Knalle neuerdings durch v. ANGERER und LADENBURG =  $330,8 \pm 0,1$  m/sek. bestimmt worden (übereinstimmend mit einem früher von REGNAULT gefundenen Werte), für Töne durch GRÜNEISEN und MERKEL = 331,57 m/sek. (extrapoliert auf unendlichen Röhrendurchmesser)<sup>1)</sup>. Für uns kommt der letzte Wert in Betracht. Die Schallgeschwindigkeit bei höheren Temperaturen kann daraus berechnet werden. Sie ist für  $18^\circ\text{C}$  = 341,6 m/sek.<sup>2)</sup>. Die gewöhnlichen Schwankungen der Zimmer-

<sup>1)</sup> Die beiden Abhandlungen in den Ann. d. Physik (4) Bd. 66. 1921.

<sup>2)</sup> Über die Berechnung vgl. K. L. SCHAEFER 3 und 4 S. 322ff. Den Tabellen SCHAEFERS ist als Schallgeschwindigkeit bei  $0^\circ\text{C}$  der Wert 331,8 zugrunde gelegt, woraus sich für  $18^\circ\text{C}$  342,6 ergibt. Dadurch werden die Viertelwellenlängen bei Tönen mittlerer Höhe um etwa 0,1 cm größer; bei höheren Tönen ist in der ersten Dezimale kein Unterschied.

temperatur machen nur geringe Unterschiede, namentlich bei hohen Tönen. Bei  $24^{\circ}\text{C}$  würde z. B. die Viertelwellenlänge für 384 Schwingungen um 0,2, für 3840 um 0,1 cm größer sein.

Nun findet ein Einfluß der Röhrenweite in doppelter Richtung statt:

a) Die Schallgeschwindigkeit wird nach KUNDT mit abnehmender Schwingungszahl des Tones und abnehmender Weite der Röhre vergrößert. Diese auch von Späteren bestätigte Gesetzmäßigkeit wird nur bei tieferen Tönen und bei Röhren von weniger als 2,6 cm Weite von Bedeutung. Nun liegt das Lumen unserer Röhren für die tieferen Töne etwas unter dieser Grenze. Aber die resultierenden Abweichungen sind wieder so gering, daß sie hier vernachlässigt werden können. Denn sie sind viel kleiner als die If.-Breite. Wenn diese bei Tönen von der Höhe etwa des *c* auch nur einen Ganzton beträgt, so bedeutet dies schon einen Spielraum der Einstellungen von 2 cm. Da wir, um der Auslöschung sicher zu sein, stets eine Mehrzahl von Röhren (bei tieferen Tönen 3—4) von etwas verschiedener Länge nach oben und unten von der berechneten einstellten und die einzelnen Einstellungen infolge der If.-Breite sich überdecken und in ihrer Wirkung gegenseitig unterstützen, so war damit dieser Einfluß der Röhrenweite außer Kraft gesetzt.

b) Dagegen trat bei hohen Tönen, die für uns weit wichtiger sind, ein anderer Einfluß zutage, auf den man bisher nicht aufmerksam geworden ist, der aber auch nur bei If.-Versuchen durch die Weite der Seitenröhren ausgeübt wird. Eine von mir (17) gemeinschaftlich mit Dr. v. ALLESCH durchgeführte Versuchsreihe ergab, daß von  $f_{is}^2$  an die zur völligen Auslöschung eines Tones erforderlichen Einstellungen der Seitenröhren um wachsende Beträge von den berechneten abweichen. Diese Abweichungen sind viel größer als die, welche sich für hohe Töne aus den Formeln für die unter a) erwähnte Abweichung ergeben würden, und sie bewegen sich in beiden vorhin erwähnten Beziehungen in umgekehrter Richtung: sie nehmen zu mit zunehmender Röhrenweite und zunehmender Schwingungszahl. Und zwar wachsen sie von  $g^3$ — $c^7$  von einem Halbton bis zu einer großen Sexte, d. h. man muß bei  $g^3$  auf den für  $gis^3$ , bei  $c^7$  auf den für  $a^7$  in freier Luft geltenden Wert einstellen<sup>1)</sup>.

<sup>1)</sup> Der Grund für dieses Verhalten dürfte darin liegen, daß die aus der Seitenröhre zurückkommende Schwingung zwar an dem Punkte, wo die Seitenröhre in die Hauptröhre mündet, die von der Rechnung verlangte entgegengesetzte Phase besitzt, sich aber von da in den Querschnitt der Hauptröhre weiterverbreitet und innerhalb dieser Strecke, wenn sie größer ist als etwa  $\frac{1}{10}$  der Wellenlänge, verschiedene Phasen annimmt, unter Umständen sogar eine mit der Hauptwelle übereinstimmende, also sie verstärkende (F. STUMPF mündlich).

In der folgenden Tabelle stehen unter I die für Viertelwellenlängen in freier Luft berechneten Werte von  $c^1$  an für die temperiert-chromatische Leiter und  $a^1 = 435$  Schw. Für die tieferen Oktaven sind sie durch Multiplikation mit Potenzen von 2 aus den Werten für  $c^1 - c^2$  abzuleiten<sup>1)</sup>. In der Spalte II stehen die korrigierten Werte für eine Röhrenweite von 10 mm. Von diesen Werten beruhen 17, die sich auf die Tonstrecke von  $f^2 - c^7$  gleichmäßig verteilen, auf wiederholt kontrollierten Beobachtungen, die übrigen auf Interpolation. Bei Röhren von 20 mm Weite ergeben sich teilweise um 1 Dezimale niedrigere, bei Röhren von 5 mm um 1 Dezimale höhere Werte.

Viertelwellenlängen für die chromatische Leiter,  $a^1 = 435$  Schw., Temp. 18° C, Röhrenweite 10 mm.

Note	n	l/4		Note	n	l/4		Note	n	l/4	
		I	II			I	II			I	II
$c^1$	259	33,0		$c^3$	1035	8,3	8,0	$c^5$	4138	2,1	1,8
$cis^1$	274	31,1		$cis^3$	1096	7,8	7,6	$cis^5$	4385	2,0	1,7
$d^1$	290	29,4		$d^3$	1161	7,4	7,2	$d^5$	4645	1,8	1,6
$dis^1$	308	27,7		$dis^3$	1230	6,9	6,8	$dis^5$	4921	1,7	1,5
$e^1$	326	26,2		$e^3$	1304	6,6	6,4	$e^5$	5214	1,6	1,4
$f^1$	345	24,7		$f^3$	1381	6,2	6,0	$f^5$	5524	1,6	1,3
$fis^1$	366	23,3		$fis^3$	1463	5,8	5,6	$fis^5$	5853	1,5	1,2
$g^1$	388	22,0		$g^3$	1550	5,5	5,2	$g^5$	6201	1,4	1,1
$gis^1$	411	20,8		$gis^3$	1642	5,2	4,9	$gis^5$	6569	1,3	1,0
$a^1$	435	19,6		$a^3$	1740	4,9	4,6	$a^5$	6960	1,2	1,0
$ais^1$	461	18,5		$ais^3$	1843	4,6	4,3	$ais^5$	7374	1,2	0,9
$h^1$	488	17,5		$h^3$	1953	4,4	4,1	$h^5$	7812	1,1	0,8
$c^2$	517	16,5		$c^4$	2069	4,1	3,9	$c^6$	8277	1,0	0,7
$cis^2$	548	15,6		$cis^4$	2192	3,9	3,7	$d^6$	9290	0,9	0,6
$d^2$	581	14,7		$d^4$	2323	3,7	3,5	$e^6$	10428	0,8	0,5
$dis^2$	615	13,9		$dis^4$	2461	3,5	3,3	$g^6$	12401	0,7	0,4
$e^2$	652	13,1		$e^4$	2607	3,3	3,1	$c^7$	16554	0,5	0,3
$f^2$	691	12,4	12,4	$f^4$	2763	3,1	2,9				
$fis^2$	732	11,7	11,6	$fis^4$	2926	2,9	2,7				
$g^2$	775	11,0	10,8	$g^4$	3100	2,8	2,5				
$gis^2$	821	10,4	10,1	$gis^4$	3285	2,6	2,3				
$a^2$	870	9,8	9,5	$a^4$	3480	2,5	2,1				
$ais^2$	921	9,3	9,0	$ais^4$	3687	2,3	2,0				
$h^2$	977	8,7	8,5	$h^4$	3906	2,2	1,9				

Man darf aber nicht etwa, um den für freie Luft berechneten Werten möglichst nahe zu bleiben, mit der Verkleinerung des

1) Für die reine Stimmung, in welcher bei  $a^1 = 435$   $c^2 = 522$  ist, würden die Werte in den tieferen Oktaven um 1-3 mm niedriger, für die „physikalische“ Stimmung ( $a^1 = 431$ ,  $c^2 = 512$ ) um 1-2 mm höher sein. Von  $c^3$  ab fallen die Werte der 3 Stimmungen mit denen unserer Tabelle bei einer Dezimale fast überall zusammen.



Lumens beliebig weit gehen. Das Lumen der Seitenröhren muß dem der Hauptröhre ungefähr gleich sein, diese aber kann nicht viel unter 10 mm heruntergehen, ohne daß die Fortleitung überhaupt geschädigt würde. Zu Beginn meiner Versuche gebrauchte ich für die höchste Tonregion Röhren bis herab zu 1 mm, bemerkte aber, daß dabei die ohnehin schwachen Töne schon in der Hauptleitung verschluckt wurden. Selbst ein sonst vortrefflich gebauter Apparat von 5 mm Weite der Hauptleitung ließ schwache Töne der höchsten Oktaven nicht mehr durch, wenn die Leitung länger als 1 m war.

Diese Abweichungen haben nun aber für unsere Zwecke zuweilen ihre gute Seite. Eine Folge davon ist nämlich, daß die ungeraden Multipla eines Tones nur bis zu einer gewissen Höhe mitausgeschlossen werden, so daß man z. B. bei *ff*-Einstellung auf den Grundton *c* doch den 11. Teilton isoliert zu Gehör bringen kann, weil ihm bei der benutzten Röhrenweite diese Einstellung nichts mehr schadet. Besonders gelingt dies, wenn der Grundton in einem Klange nur schwach vertreten ist, was (gegenüber einer verbreiteten Meinung) bei vielen obertonreichen Klängen tatsächlich zutrifft. Dann wird er schon durch eine nur annähernde Einstellung auf eine berechnete Viertelwellenlänge vernichtet, und man kann die ungeraden Teiltöne gleichwohl etwa vom 7. oder 9. ab hören, wenn man auf ihre isolierte Beobachtung Gewicht legt. Außerdem muß man sie eben durch besondere, ihrer eigenen Höhe entsprechende Einstellungen vernichten.

### 3. Probe durch schwebende Hilfsgabeln.

Die schärfste Probe, ob ein Ton wirklich vernichtet ist, geschieht durch annähernd auf diesen Ton gestimmte Hilfsgabeln, die mit einer ungeheuren Empfindlichkeit auf jede Spur des Tones durch Schwebungen reagieren (Vgl. m. Abh. 3 und 7).

Nur beim 2. Teilton, der Oktave des Grundtones, darf man sich nicht ohne weiteres darauf verlassen. Hier werden infolge der Bildung eines Differenztones zwischen der Hilfsgabel und dem Grundton „tiefe Schwebungen“ (auf dem Grundton) auch dann herauskommen, wenn der 2. Teilton bereits völlig aus dem Klange verschwunden ist (3, S. 669; 7, S. 110). Man kann sich nun zwar auf die Unterscheidung der „tiefen“ und „hohen“ Schwebungen einüben, aber dies ist nicht leicht. Die Probe ist daher, statt mit der Hilfsgabel, hier so zu machen, daß man auch den Grundton selbst vernichtet. War der 2. Teilton vorhanden, so wird er dann isoliert hörbar. Und man wird oft erstaunt sein, wie stark er ist, obschon man ihn direkt gar nicht herausgehört hatte.

Daß ein Klang gänzlich von Obertönen gereinigt sei, darf man nicht etwa aus der Tatsache schließen, daß bei Einstellung von If.-Röhren auf den Grundton überhaupt nichts mehr zu hören ist (so z. B. LEWIN 2, S. 328): denn es könnten ungerade Multipla vorhanden gewesen sein, die durch den Grundton mitausgeschlossen wurden. Eine Prüfung mit schwebenden Hilfsgabeln ist daher unerläßlich.

#### 4. Wissentliches und unwissentliches Verfahren.

Die Experimentalpsychologen pflegen das unwissentliche Verfahren, bei dem die „Versuchsperson“ (Vp.) womöglich nicht einmal über den Zweck der ganzen Untersuchung unterrichtet ist, im allgemeinen höher zu stellen. Vielfach mit Recht, besonders wenn die Vpn. nur so gebraucht werden, wie die Hühner oder die Frösche in der Physiologie, um aus ihren Reaktionen, die hier auch sprachlicher Art sein können, auf bestimmte Wahrnehmungen oder sonstige innere Vorgänge zu schließen. Anders aber, wenn es sich gar nicht um Vpn., sondern um Beobachter handelt, deren Aufgabe die Beschreibung einer wahrgenommenen Erscheinung ist. Einem solchen wird es nicht notwendig schaden, wenn ihm gesagt wird, um was es sich handelt; nicht einmal immer, wenn ich ihm sage, was ich selbst daran wahrgenommen habe. Er weiß dann genau, worauf er seine Aufmerksamkeit zu richten, was er zu prüfen hat. Freilich kann unter Umständen auch eine Suggestion eintreten; ein geschulter Beobachter verläßt sich aber mehr auf sein eigenes Ohr als auf das eines anderen<sup>1)</sup>.

Extrem wissentlich wäre das Verfahren in unserem Falle, wenn der Beobachter nicht bloß wüßte, daß ein Vokal und welcher Vokal aus der Röhre kommt, sondern auch, welche Veränderung er im Augenblick zu erwarten hat. In diesem Sinne wissentlich waren die Versuche nur, wenn ich selbst gegen das Ende meiner Studien hin beobachtete: denn dann wüßte ich im voraus, was eintreten würde.

Dagegen habe ich wissentliche Versuche in dem beschränkten Sinne, daß dem Beobachter bekannt war, welcher Sprachlaut untersucht wurde, im weitesten Umfange, besonders bei Längsschnitten (s. u.), nicht nur als zulässig, sondern als das Verfahren erprobt, das am besten vorwärtsbringt. Der Beobachter weiß dann noch immer nicht, welche Veränderungen eintreten werden, und jedenfalls nicht, ob gerade die Röhreneinstellung vorliegt,

<sup>1)</sup> Vgl. über verkehrte Heranziehung von Vpn., wo man die geübtesten Beobachter gebraucht hätte, m. Abh. 7, S. 12ff.

bei der eine bestimmte Veränderung einzutreten pflegte. Er kann aber nunmehr seine Aufmerksamkeit ganz und gar auf die engste Fragestellung konzentrieren.

Extrem unwissentlich wäre das Verfahren, wenn der Beobachter nicht einmal wüßte, ob ein Vokal oder ein Konsonant, ja nicht einmal, ob ein Sprachlaut oder ein sonstiger Gehörseindruck gegeben wird. Versuche dieser Art habe ich gelegentlich eingeschaltet, und sie sind nicht ohne Interesse. Aber systematische Versuchsreihen dürfen so nicht angestellt werden. Man muß mindestens wissen, daß es sich um Sprachlaute handelt, meist auch, um welche Gattung, Vokale, Konsonanten, gesungene, gesprochene, geflüsterte. Je weitere und vielfältigere Möglichkeiten man offen läßt, je unbestimmter die Instruktion lautet, um so größer wird auch der Spielraum zufälliger subjektiver Einstellungen, die dann oft auch noch sich fortsetzen (Perseveration) und eine ganze Reihe schädigen können<sup>1)</sup>.

In der lebendigen Sprache hat man beständig richtungweisende Einstellungen. Schon indem wir uns auf das Deutsche, Französische, Italienische einstellen, sind eine große Zahl von Möglichkeiten der Auffassung des Gehörten ausgeschlossen, andere nähergerückt. Eine weitere Determination bringt der Zusammenhang. Wie schlecht isolierte Laute, ja ganze Silben, bei den geringsten Hindernissen verstanden werden, zeigen bekannte Versuche (GUTZMANN 3). Es ist, als wäre der Laut selbst ein anderer, fremdartiger. Auch meine Versuche brachten hierfür Belege. Es ist sogar vorgekommen, daß einer meiner besten Beobachter, als ihm nach vielen Vokalreihen unwissentlich ein S dargeboten wurde, es nicht erkannte, obgleich es ausgezeichnet aus der Leitung kam, nur weil er eben augenblicklich nicht an die Möglichkeit dachte, daß auch Konsonanten gegeben würden. Sobald ihm dies gesagt wurde, erschien es wie verwandelt und wurde als gutes S anerkannt. Es wäre daher zweckwidrig, in Versuchen über Strukturveränderungen von Vokalen bei Ausschaltung von Teiltönen Unwissentlichkeit in diesem weitgehenden Umfange zu verlangen.

Bei den unwissentlichen Versuchen im obigen beschränkten Sinne ist es eine wichtige Vorschrift, die man selbst psychologisch geschulten Beobachtern immer wieder einprägen muß, daß es sich hier nicht darum handelt, aus dem aus der Röhre kommenden Laute den im Schallzimmer hineingesungenen richtig zu erschließen, sondern lediglich darum, die Beschaffenheit

<sup>1)</sup> So hatte sich ein Beobachter einmal eine Zeitlang auf A eingestellt und hörte ein geflüstertes Ö, als es beim Aufbau schon ganz deutlich sein mußte, als A.

des gehörten Lautes selbst als eines akustischen Phänomens so genau wie möglich zu beschreiben.

Der Anfänger hält sich zunächst oft an die aus dem Leben übernommene Gewöhnung, den vom Sprechenden oder Singenden intendierten Laut zu erkennen oder zu erraten. Man klammert sich dann an allerlei sekundäre Kriterien, und dies um so mehr, je stärker der Gehörseindruck von den gewohnten abweicht; während es doch gerade dieser Gehörseindruck als solcher ist, auf den alles ankommt. Es war dies einer der Gründe, weshalb ich das anfängliche unwissentliche Querschnittverfahren (s. u.) später meistens durch ein wissentliches Längsschnittverfahren ersetzte.

Bemerkungen über psychologische Einflüsse der hier erwähnten Art schon bei KEMPELEN und WILLIS, denen auch HERMANN (Bd. 91, S. 156 Anm.) zustimmt, besonders treffend und eingehend bei ROUSSELOT (I, 34ff. „L'éducation de l'oreille“).

Erstaunlich, ja fast unbegreiflich ist zuweilen die Unfähigkeit zum unbefangenen Beobachten bei nicht naturwissenschaftlich geschulten Personen. Ich demonstrierte eines Tages 7 Mitgliedern eines Universitätsseminars für neuere Sprachen, darunter einem Lektor, die Veränderungen der Vokale durch Interferenzröhren. Es wurde ein Ö abgebaut. Der erste Beobachter, eine Dame, gab aber fortwährend an, Ö zu hören, als es längst in ein reines, ja sogar in ein dunkles O übergegangen war. Und diese Beurteilung übertrug sich auf ihre sämtlichen Nachfolger, die ihre Aussagen mitangehört hatten. Ich begann fast an meinem eigenen Ohre zu zweifeln, bis ein herbeigerufener bewährter Beobachter, Dr. WERTHEIMER, der unwissentlich zu urteilen hatte, im ersten Augenblick auf O erkannte.

##### 5. Verschiedenheiten beim Auf- und Abbau.

Es gibt aber einen Punkt, in dem auch der geübteste Beobachter einem konstant wirkenden psychologischen Einfluß ausgesetzt ist: die Ergebnisse bei Ab- und bei Aufbauversuchen weichen fast ganz regelmäßig etwas voneinander ab, insofern alle Stadien der Umwandlung beim Aufbau etwas tiefer liegen als beim Abbau. So erscheint z. B. ein Laut beim Abbau an einer Stelle bereits alteriert, wo er beim Aufbau schon fertig und tadellos ist; oder er hat bei einer anderen Einstellung bereits ganz seinen spezifischen Charakter verloren, während beim Aufbau an der gleichen Stelle schon etwas von diesem Charakter zu bemerken ist. Der Unterschied kann natürlich nur psychologische Ursachen haben; wie ja auch bekanntermaßen bei der Bestimmung der Hörschärfe ein kleiner Unterschied des Schwellenwertes auftritt, je nachdem man eine Klangquelle dem Ohre nähert oder sie von ihm entfernt. Die Ursache liegt hier offenbar darin, daß man beim Aufbau für die erste Spur des spezifischen Charakters besonders empfindlich ist und mit der Erreichung einer gewissen

Vollkommenheit und Deutlichkeit den Laut schon für fertig hält, während man beim Abbau, wo der Laut in seiner natürlichen vollkommenen Gestalt kurz vorher noch gehört wurde, wieder für Alterationen empfindlicher ist und den gänzlichen Verlust des spezifischen Charakters zu früh konstatiert. Diesem verschiedenen Verhalten unterliegt man auch bei großer Übung immer wieder, selbst wenn Ab- und Aufbau unmittelbar aufeinanderfolgen.

#### 6. Analysierendes Verhalten.

Im gewöhnlichen Leben hören wir die Vokale als einheitliche Lautgebilde. Der analysierenden Einstellung des Linguisten und Phonetikers offenbaren sich aber schon beim direkten Hören gewisse Gliederungen, z. B. im I die dunkle Unterlage und der helle Oberbau. Diese Neigung zur Analyse, die ja überhaupt jedem wissenschaftlichen Beobachter eigen sein muß, verstärkt sich noch, wenn man weiß, daß während einer Beobachtungsreihe eine allmähliche Umbildung der Laute erfolgt, deren einzelne Stadien wahrgenommen und beschrieben werden sollen. So kommt es namentlich, daß in einem gewissen Stadium 2 Elemente, ein tiefes dunkles und ein neu hinzukommendes helles, im Gehörseindruck unverbunden nebeneinander liegen, wie wir dies z. B. in der Formel  $Uü + i$  ausdrücken. Der akustisch Geübte nimmt dann auch wohl geradezu einen oder zwei mit dem Ganzen noch nicht verschmolzene Obertöne wahr, die durch die Einschiebung der letzten Röhren frei geworden sind. In diesem Fall ist regelmäßig zu bemerken, daß die Teiltöne beim Aufbau ihre volle Stärke erst erlangen, wenn man mit der Einschiebung der Röhren schon um eine Terz höher gelangt ist (o. S. 43). In dieser Weise drängten sich oft bei den hellen Vokalen besonders starke Teiltöne auf, sowohl in den unteren Regionen (Unterformant) als in der 3- und 4-gestrichenen Oktave, z. B. beim Ä auf dem Grundton  $e$  die Obertöne  $gis^3$  und  $ais^3$ , beim I auf demselben Grundtone  $gis^4$ .

Solche Wahrnehmungen sind natürlich immer wertvoll, und die analysierende Einstellung ist so lange festzuhalten, bis der Eindruck selbst beim weiteren Aufbau wieder so einheitlich geworden ist, daß er dem natürlichen Laut auch in dieser Hinsicht gleichkommt.

#### 7. Beschreibung des Gehörten.

Hierbei vermeide man, von außen hereingetragene Kategorien anzuwenden (wozu geistreiche Beobachter neigen), gebe sich vielmehr völlig unbefangen dem Lauteindrucke selbst hin. Ein ausgezeichnetes Mittel, sich diesen klarzumachen, ist die Nach-

erzeugung, sei es in Wirklichkeit, sei es wenigstens in anschaulicher Vorstellung. Die aktuelle Nachbildung hat auch den Vorteil, daß der Versuchsleiter sie wahrnehmen und die Bezeichnungsweise des Beobachters dadurch kontrollieren kann. In Zweifelsfällen ist es auch nützlich, auf solche Weise eine Mehrzahl ähnlicher Laute durchzuprobieren, um den zutreffendsten zu finden. Man bemerkt dann geradezu ihre verschiedenen Abstände von dem vorliegenden.

In der Bezeichnungsweise gibt es individuelle Unterschiede. So nannte ein Beobachter das Nämliche, was ich ein dunkles A nannte und A<sub>o</sub> schrieb, stets ein „helles (oder offenes) O“. Für ihn fiel der Laut noch unter die Sphäre des O, was für meine Auffassungsgewohnheiten ganz ausgeschlossen war. Möglicherweise wirken hier auch Erziehungseinflüsse und Unterschiede der nord- und süddeutschen Aussprache mit. Ein anderer Beobachter (Dr. RIEFFERT) wurde in seinen Ausdrücken durch Photismen mitbestimmt. A nannte er dunkler als O, weil er bei A schwarze Photismen hat. Bei einer If.-Einstellung, durch die A in A<sub>o</sub> übergeht, nannte er es daher gerade „heller“. Nachdem sich dies herausgestellt hatte, wurde vereinbart, immer nur zu sagen: „A nach O hin, A nach E hin“ usf.

Durch diese Maßnahmen (1.—7.) und andere, die nicht alle hier aufgezählt werden können, läßt sich bei If.-Versuchen die Rolle des Zufalls und des Subjektiven auf ein Minimum herabdrücken. Man kann dann mit fast physikalischer Sicherheit darauf rechnen, daß jeder gute Beobachter unter gleichen Bedingungen gleiche Resultate haben wird.

### III. Ergebnisse.

Das untersuchte Lautmaterial und die Versuchsreihen.

Mit der If.-Methode wurden nicht bloß die sog. 5 Hauptvokale untersucht, sondern auch die in der deutschen Sprache gebräuchlichen Laute Ö, Ä, und Ü, die akustisch ganz die gleichen Probleme darbieten und ebenso gleichmäßig andauernde Klänge darstellen; während in den Diphthongen, wie Au, Ei, die analysierende Aufmerksamkeit leicht eine Aufeinanderfolge zweier durch einen raschen Übergang verbundener Vokale erkennt. Über die Aussprache der untersuchten Laute gilt das bereits bei der Resonanzmethode Gesagte.

Alle diese Vokale wurden in der beschriebenen Weise ab- und aufgebaut und die Teiltöne nach ihrer Stärke und ihrem Einfluß auf den Vokalcharakter auch mit Lücken- und Stichversuchen geprüft. In erster Linie wurden Aufbaureihen durchgeführt, da sie den wichtigsten Punkt, die untere Formantgrenze, am deutlichsten erkennen lassen. Abbaureihen dienten als Vorbereitung, Lücken- und Stichversuche als Ergänzungen.

22 vollständige Aufbaureihen wurden mit sämtlichen Vokalen ausgeführt, hauptsächlich auf den Grundtönen *C*, *Ges*, *c*, *ges*,  $c^1$ ,  $ges^1$ ,  $c^2$ , die je um eine halbe Oktave voneinander absteigen. Vollständige Reihen wurden aber auch auf *Gis* und *e* gemacht, unvollständige (besonders für den Vokal *A*) auch noch auf anderen Tönen. Es ist ein großer Vorzug der If.- vor der Resonanzmethode, daß man jeden beliebigen Grundton verwenden kann. Diesen Vorzug teilt sie mit der objektiven Aufnahme und Zerlegung der Wellen. Beim synthetischen Verfahren, das sonst alle anderen an Klarheit und Erkenntniswert seiner Ergebnisse weit übertrifft, waren ebenfalls bestimmte Grundtöne vorgeschrieben.

Wir beschreiben zuerst die Ergebnisse von 16 Reihen, die mit kürzester Leitung und unter Ausschluß der engsten Röhrensysteme gemacht wurden, damit der Laut am sichersten unversehrt hindurchkam. Die Schallgebung erfolgte hier im Zimmer II (s. Schema S. 44). Die Länge der Leitung betrug zwischen  $4\frac{1}{2}$  und 8 m. Verschiedene Sänger dienten zur Lautgebung, bei  $ges^1$  und  $c^2$  Frauen und Kinder, bei *C* ein früherer Domsänger, der diesen für die menschliche Stimme ganz ungewöhnlich tiefen Ton mit großer Kraft während der  $1\frac{1}{2}$ -stündigen Versuche wiederholen konnte. Die Vokale kamen auch da vollkommen deutlich heraus.

Wie bei den Resonanzversuchen, wird zunächst (1.—3.) von den Ergebnissen bei dem Grundton  $c^2$  abgesehen, da sich hier wesentliche Abweichungen zeigten.

### 1. Nur harmonische Teiltöne.

Es bestätigte sich, daß gesungene Vokale ausschließlich aus harmonischen Teiltönen bestehen. Niemals ist in den Versuchsreihen etwas von dem Laut übriggeblieben, wenn die sämtlichen harmonischen Teiltöne ausgeschlossen waren; auch während der Ab- und Aufbaureihen ist nirgends eine Spur anderer Beiltöne zum Vorschein gekommen. Für sich allein würde ich dies zwar in bezug auf die höheren Teiltöne wegen der If.-Breite nicht als voll ausreichenden Beweis ansehen, aber als Bestätigung können diese Erfahrungen immerhin gelten. Das synthetische Verfahren wird uns die Möglichkeit bieten, die Frage an der Wurzel zu fassen.

W. KOEHLER zeigte bereits (I, II, S. 71 ff.), daß bei U, O und A nach Vernichtung aller harmonischen Teiltöne nichts übrigbleibt, und folgerte mit Recht, daß wenigstens zwischen den niedrigen Teiltönen, die weit auseinanderliegen, keine unharmonischen liegen können, und daß U, O, A, bei denen diese niedrigen Teiltöne den Ausschlag geben, nicht durch solche von unharmonischen Verhältnissen charakterisiert sein können. Für O führte er noch den besonderen Nachweis, daß die Berufung auf die If.-Breite zur Entkräftung dieses Argumentes nicht genüge.

S. GARTEN machte nun folgendes Gegenexperiment (3, VII, S. 35ff.): er ließ ein A durch einen Resonator von 462 Schw. hindurchgehen, der dem O-Formanten entsprach, und fand es dadurch nach O hin verschoben. Darauf schaltete er den Grundton durch If. aus und beobachtete an der graphisch aufgenommenen Kurve eine Veränderung, die er auf den Wegfall des dem A künstlich beigefügten unharmonischen Resonatortones deutete. Daraus schloß er, daß durch die Vernichtung des Grundtones dieser letztere mitvernichtet worden sei, obgleich er von harmonischen Teiltönen um mehr als If.-Breite entfernt lag.

Da GARTEN nichts über die Höhe des gesungenen Grundtones angibt, kann man den ganzen Versuch schwer beurteilen. Überdies scheint es physikalisch unverständlich, wie eine solche Mitauslöschung vor sich gehen sollte. Wahrscheinlich war durch den Resonator, der ja auch in einer gewissen Breite wirkt, ein ungeradzahligter harmonischer Teilton verstärkt worden, der dann durch If. auf den Grundton mitausgelöscht wurde. Es gibt aber einfachere und exaktere Mittel, unharmonische Teiltöne einem Klang beizumischen. Die Wirkungen werden wir im 7. Kapitel beschreiben.

## 2. Entwicklungsstadien.

Jeder Vokal durchläuft beim Ab- und Aufbau bestimmte Stadien. Die folgenden Tabellen geben ein Bild dieser Veränderungen beim Aufbau. Die beiden ersten Kolumnen der Tabellen (von unten nach oben zu lesen) bedeuten das Hinzutreten des betreffenden Teiltönen zum Grundton und den vorangehenden Teiltönen. Jede der Vokalbezeichnungen gilt solange, bis sie durch eine andere abgelöst wird. Also z. B. beim E auf dem Grundton C bleibt der Eindruck des O vom Auftreten des Teiltönen  $g^1$  bis zum Hinzukommen von  $c^3$  unverändert. U auf demselben Grundton bleibt von  $d^2$  an überhaupt unverändert U<sub>0</sub>, O bleibt von  $c^2$  an O<sub>a</sub>, auch wenn sämtliche Röhren hineingeschoben, also alle Interferenzen beseitigt und alle Obertöne eingefügt sind.

Ou bedeutet ein dunkles O, etwas nach U hin; OU und UO einen ungefähr in der Mitte liegenden Laut (wobei der vorangehende Buchstabe eine vorwiegende Auffassung als O oder als U bedeutet). Üu + i ein Ü, dem Spuren sowohl von U als von I beigemischt sind, die aber untereinander nicht zu einer Einheit verschmelzen. Analog bei den übrigen Vokalen. Die Tabelle für c besagt also beispielsweise, daß ein auf c gesungenes Ö, solange alle Obertöne ausgeschaltet sind, als U erscheint, daß es beim Hinzutreten des 3. Teiltönen ( $g^1$ ) zu einem dunklen O, weiter bei  $c^3$  zu einem guten O wird, daß bei  $g^3$  die erste Spur eines Ö, etwas Helles, sich beimischt, und daß es mit  $d^4$  fertig ist. Die Stelle des ersten Auftauchens der eigentümlichen Färbung eines Vokals, bei den helleren die des ersten Auftauchens einer hellen Nuance oder einer hellen Beimischung zu dem dunklen Untergrund, ist durch Unterstreichen hervorgehoben.



Grundton C.

36	<i>d</i> <sup>4</sup>				Ä noch schneidender	Ü, immer noch mit etwas ö	E gut	I gut	
34	<i>des</i> <sup>4</sup>				Ä ausgezeichnet		Eä	Iu	
32	<i>c</i> <sup>4</sup>				Ä gut	Üö	Eä dunkel	U + i (hoher Beiton)	
30	<i>h</i> <sup>3</sup>				Äa (dunkel)		ÖÄ dunkel		
28	<i>b</i> <sup>3</sup>				Aa sehr markig		Oöä	Uü unklar	
26	<i>a</i> <sup>3</sup>				Aä (ebenmerklich)	Ö	Oö + ä	U, aber nicht mehr schön	
25	<i>as</i> <sup>3</sup>			Öä heller					
24	<i>g</i> <sup>3</sup>			Öä	A hell und klar	Oö			
23	<i>ges</i> <sup>3</sup>			ÖÄ dunkel		Uö	O etwas blökend		
21	<i>f</i> <sup>3</sup>			Oa + ö		U nicht mehr ganz so rund	O nicht mehr so rund		
20	<i>e</i> <sup>3</sup>		A gut	O sehr hell					
18	<i>d</i> <sup>3</sup>			Oa	A heller				
16	<i>c</i> <sup>3</sup>								
14	<i>b</i> <sup>2</sup>				A (gut)				
12	<i>g</i> <sup>2</sup>		A zieml. gut		Ao				
10	<i>e</i> <sup>2</sup>		Ao		AO				
9	<i>d</i> <sup>2</sup>	Uo							
8	<i>c</i> <sup>2</sup>		Oa	AO	O markig	OA			
7	<i>b</i> <sup>1</sup>					Oa	U		
6	<i>g</i> <sup>1</sup>	UO	O	O	O sehr kräft.	O	UO	O	
5	<i>e</i> <sup>1</sup>	Überall UO, etwas stärker und sehr markig, <i>e</i> <sup>1</sup> heraushörbar							
4	<i>c</i> <sup>1</sup>	Immer noch überall U, etwas stärker							
3	<i>g</i>	Überall noch U, aber markiger, dröhnender, wie schöner Instrumentalklang							
2	<i>c</i>	Überall majestätisches dumpfes U, tonal, C + c							
1	<i>C</i>	Unhörbar							
		U	O	A	Ö	Ä	Ü	E	I

Grundton Ges.

40	<i>b</i> <sup>4</sup>							I sehr gut	
32	<i>ges</i> <sup>4</sup>							I gut	
30	<i>f</i> <sup>4</sup>							E gut	
28	<i>e</i> <sup>4</sup>							E	
26	<i>es</i> <sup>4</sup>							ÖE	
25	<i>d</i> <sup>4</sup>			Ä gut		Ü gut		Ö	
24	<i>des</i> <sup>4</sup>			Ä dunkel				Üu	
22	<i>c</i> <sup>4</sup>							Uü	
21	<i>h</i> <sup>3</sup>			Ö fertig	AoÄ blökend		Üu	Öou	
20	<i>b</i> <sup>3</sup>				Ao + ä		Uü	OÖ	
18	<i>as</i> <sup>3</sup>			Oö		U nicht mehr so voll			
16	<i>ges</i> <sup>3</sup>			O sehr markig			O sehr markig	U markig	
14	<i>e</i> <sup>3</sup>								
12	<i>des</i> <sup>3</sup>				Ao				
9	<i>as</i> <sup>2</sup>		A gut	O					
8	<i>ges</i> <sup>2</sup>		Ao		AO		O		
7	<i>e</i> <sup>2</sup>		AO						
6	<i>des</i> <sup>2</sup>	O	OA		OA				
5	<i>b</i> <sup>1</sup>		O markig		O				
4	<i>ges</i> <sup>1</sup>	U	Ou	Ou	Ou	U	Ou	U	
3	<i>des</i> <sup>1</sup>	Immer noch überall U, aber kräftig, metallisch . . . . .							
2	<i>ges</i>	U stärker und voller . . . . .							
1	<i>Ges</i>	U äußerst dumpf und schwach . . . . .							
		U	O	A	Ö	Ä	Ü	E	I

Entwicklungsstadien.

Grundton *c*.

26	$\bar{a}^4$						E	I	
24	$g^4$							Iu	
22	$\bar{f}is^4$					Ü	ÖE	Ui	
20	$e^4$							Üu + i	
18	$d^4$				Ö	Ä	Ö	Üu	
16	$c^4$					Üu		Uü	
15	$h^3$				Öo	Äao	Öu		
14	$\bar{b}^3$								
12	$g^3$				Oö	AO			
10	$e^3$								
8	$c^3$			A	O				
7	$\bar{b}^2$			Ao					
6	$g^2$			OA		OA			
5	$e^2$			Oa		Oa			
4	$c^2$	O gut		Ou		O			
3	$g^1$	O		OU	Ou	Ou	O		
2	$c^1$								
1	$c$	U	U	U	U	U	U	U	
		U	O	A	Ö	Ä	Ü	E	I

Grundton *ges*.

20	$b^4$							I	
18	$as^4$						E		
17	$g^4$						Eö		
16	$ges^4$					Ü	Eo	Iu	
15	$f^4$						Eoö		
14	$e^4$				Ö gut	Ä gut		Uü + i (hoher Beiton)	
13	$\bar{es}^4$						Öoe		
12	$des^4$					ÄÄ	Öo	Uü	
11	$c^4$				Ö	Aoä	OÖ	U metallisch	
10	$b^3$				OÖä	Ao	Uü		
9	$as^3$				O me- tallisch		O me- tallisch		
6	$des^3$			A					
5	$b^2$			Ao					
4	$ges^2$	U me- tallisch	O heller	AO		AO			
3	$des^2$		O	Oa		Oa	O		
2	$ges^1$		Ou	Ou	O	O	Ou		
1	$ges$	U	U	U	U	U	U	U	
		U	O	A	Ö	Ä	Ü	E	I

Grundton *c*<sup>1</sup>.

13	<i>ā</i> <sup>4</sup>								I
12	<i>g</i> <sup>4</sup>								ÜI
11	<i>fis</i> <sup>4</sup>						Ü	E	
10	<i>e</i> <sup>4</sup>					Ä			Ü
9	<i>d</i> <sup>4</sup>				Ö gut			ÖE	Ü
8	<i>c</i> <sup>4</sup>				Ö	AoÄ	Uü	Öo	Ü
7	<i>b</i> <sup>3</sup>				Öo				
6	<i>g</i> <sup>3</sup>				Oö				
5	<i>e</i> <sup>3</sup>			A gut		Ao		O	
4	<i>c</i> <sup>3</sup>			A		OA			
3	<i>g</i> <sup>2</sup>			Ao				Ou	
2	<i>c</i> <sup>2</sup>		O	O	O	O		UO	
1	<i>c</i> <sup>1</sup>	U	U	U	U	U	U	U	U
		U	O	A	Ö	Ä	Ü	E	I

Grundton *ges*<sup>1</sup>.

9	<i>as</i> <sup>4</sup>					Ä	Ü	E	I
8	<i>ges</i> <sup>4</sup>					Äa		Eö	Iu
7	<i>e</i> <sup>4</sup>				Ö	Äao		Ö	IÜU
6	<i>des</i> <sup>4</sup>				Öo	AOä	Üu	O-E	
5	<i>b</i> <sup>3</sup>				Oö				
4	<i>ges</i> <sup>3</sup>			A gut					
3	<i>des</i> <sup>3</sup>	Uo	Oa	A(o)					
2	<i>ges</i> <sup>2</sup>		Ou	Oa	O	AO		O	
1	<i>ges</i> <sup>1</sup>	U	U	U	U	U	U	U	U
		U	O	A	Ö	Ä	Ü	E	I

Grundton *c*<sup>2</sup>.

Ohne If.	Ou	OA	Asehgut	Die helleren Vokale unvollkommen und einander vielfach ähnlich.					
7	<i>b</i> <sup>4</sup>			Öä	Ä ziemlich gut	Ü leidlich	Eä	Iu	
6	<i>g</i> <sup>4</sup>		Asehgut		Aä	Üu		Ü	Ü
5	<i>e</i> <sup>4</sup>	Ou		ÖÄ	A sehr hell		Ö heiser		
4	<i>c</i> <sup>4</sup>		A gut	O hell	Oa	Uü?			
3	<i>g</i> <sup>3</sup>		OA	A					
2	<i>c</i> <sup>3</sup>	UO	O stark	Ao	O	O	Uo	O	Uo
1	<i>c</i> <sup>2</sup>	Alle identisch; eigentlich kein Vokal, am ehesten noch ein sehr helles U.							
		U	O	A	Ö	Ä	Ü	E	I

Wenn hier vom Hinzutreten der Teiltöne gesprochen wird, ist immer gemeint: in der relativen Stärke, wie sie in dem betreffenden Vokal enthalten sind. Aber auch in dieser Hinsicht ist in acht zu behalten, daß sie ihre volle natürliche Stärke erst erreichen, wenn der Aufbau etwa um eine Terz höher gediehen ist (If.-Breite). Nur die 4 ersten Teiltöne, die größere Intervalle untereinander bilden, treten sogleich mit voller Stärke auf.

Beim Abbau wiederholen sich natürlich dieselben Stufen in umgekehrter Ordnung; nur treten die Wendepunkte meist ein wenig früher ein (o. S. 51).

Der Übersichtlichkeit halber nehmen wir hier auch schon die erst später zu besprechende Tabelle für  $c^2$  mit auf.

Sämtliche Laute beginnen mit einem leisen Ton von U-Charakter. Wird jedoch auf dem Grundton  $C$  gesungen, so ist dieser für sich allein überhaupt nicht zu hören. Liegt der Grundton zwischen  $Ge^s$  und  $c^1$ , so wird er in der Regel als ein sehr dumpfes U bezeichnet. Wäre der Beobachter nicht auf das Hören von Vokalen eingestellt, so würde er vielleicht nur angeben, einen leisen, mehr oder weniger dunklen, höchst weichen Ton zu hören. Irgendein Unterschied unter den Vokalen besteht hier nicht, abgesehen von gewissen regelmäßigen Stärkeverschiedenheiten (s. u.).

Ist der gesungene Vokal selbst ein U, so wird er, wenn ein Oberton nach dem anderen freigegeben wird, nur noch etwas voller, sozusagen körperlicher, auch wohl metallischer, ist aber längstens mit dem Auftreten der Teiltöne der 2-gestrichenen Oktave fertig.

O geht von seiner U-Grundlage aus durch  $Uo$ ,  $UO$  oder auch direkt in O über; letzteres ist der Fall, wenn der Grundton zwischen  $c^1$  und  $ges^1$  liegt. In diesem Falle genügt das Hinzukommen des 2. Teiltons (der Oktave) in der Stärke, wie er eben in dem gesungenen O enthalten ist, um den O-Charakter herzustellen, der dann durch weiter hinzutretende höhere Teiltöne nicht mehr wesentlich verändert wird.

A durchläuft die Hauptstufen U, O, A (nebst den etwaigen Übergängen) und wird weiterhin nur noch etwas heller und stärker.

Ö geht von U durch O in Ö über.

Ä wird aus U zu O,  $Ao$  und einem eigentümlichen Blöklaut  $Aoä$ , der an das französische *cœur*, *sœur* oder an das in braunschweigischen und hannoverschen Gegenden gesprochene A erinnert. Dieser wandelt sich dann durch Zurücktreten seiner dunkleren Bestandteile in Ä.

Ü bleibt lang ein U. Dann tritt plötzlich eine minimale Erhellung ein, und es geht durch  $Uü$  in Ü über. Analog ist es beim I. E dagegen geht zunächst aus U in O über, dann durch  $Oö$ , evtl. Ö, in E.

Außer U selbst hat also jeder Vokal seine nächste Grundlage in einem anderen Vokal, O in U, A in O, Ä in AO, E in O, I und Ü in U. Darauf baut sich erst das Eigentümliche des bezüglichen Vokals auf.

Es ist eine bemerkenswerte Tatsache, daß überhaupt alle wesentlichen Umwandlungsprodukte beim Ab- und Aufbau selbst Vokalcharakter tragen. Dies ist durchaus nicht selbstverständlich. Denn es gibt zahllose Kombinationen von Teiltönen in bestimmten Stärkeverhältnissen, die instrumental klingen. In einzelnen Fällen sind uns denn auch solche mehr instrumental gefärbte, klarinetten- oder fagottähnliche, auch wohl tierstimmenähnliche Laute begegnet; aber nur bei den Lückenversuchen sowie bei einzelnen Konsonanten wie R („Gurren“), niemals bei den gewöhnlichen Abbauversuchen mit Vokalen. Über die Erklärung dieser Tatsache s. 13. Kap.

Besonders hervorzuheben sind die langen toten Strecken bei den helleren und hellsten Vokalen. Sind beim Aufbau des Ü, E oder I die ersten Stempel zurückgestellt, also die tiefsten Teiltöne freigegeben, so verändert sich der Vokalcharakter bei weiteren Einschiebungen lange Zeit überhaupt nicht mehr, bis der erste Schimmer von „etwas Hellem“ auftritt, das den Anfang des spezifischen Ü-, E- oder I-Charakters darstellt.

Der nun folgende status nascendi selbst ist auch sehr interessant. Der Laut hat dann etwas Zwiespältiges, Brüchiges, wie eine mutierende Stimme; es streiten gewissermaßen ein dunkles und ein helles Element in ihm um die Vorherrschaft, man hört auch wohl z. B. ein O und ein E geradezu nebeneinander, das dunkle Element wird aber durch jeden neu hinzukommenden Oberton mehr und mehr zurückgedrängt, bis es zuletzt nicht mehr bemerkt wird.

Der geschilderte Entwicklungsgang ist nun aber von verschiedener Länge je nach der Höhe des Grundtons. Er umfaßt beim Grundton *C* bis zu 6 Oktaven, verkürzt sich aber mit steigendem Grundton. Dies ist, wie man leicht sieht, begründet in der festen (bzw. nur wenig verschiebbaren) Lage der „Formanten“ und liefert einen Beweis dafür. Aber wir kommen zugleich auf dem *If*-Wege zu einer näheren Bestimmung dieser entscheidendsten Regionen.

Zuvor noch einige Bemerkungen über die Erscheinungen beim Grundton *C*, der in der Musik wohl niemals (abgesehen vielleicht von russischen Kirchenhören) von einem Sänger verlangt wird, wenn er auch als Schlußeffekt zuweilen gratis beigegeben wird. Obgleich auch hier alle Vokale durchaus kenntlich und unter-

scheidbar sind, zeigen sich doch gewisse bemerkenswerte Eigenheiten. Der Grundton scheint hier überall nur als subjektiver Differenzton der Obertöne im Ohre zu entstehen, objektiv aber zu fehlen, ausgenommen vielleicht in der unmittelbaren Nähe des Sängers. Mit dem Hinzutreten des 2. Teiltöns, der Oktave *c*, wird der Klang hörbar, ist aber für alle Vokale noch identisch, ein dumpfes U von majestätischem Charakter. Mit dem 3. Teilton (*g*) wird er kräftiger, etwas dröhnend, ist aber immer noch ein unzweifelhaftes U, nur mit kleinen Stärkeunterschieden unter den einzelnen Vokalen. Erst mit dem 5. Teilton, *e*<sup>1</sup>, wandeln sich alle Vokale in ein sehr markiges UO, und dann beginnen die Differenzierungen. Das Markige kommt in den Klang durch die Schwebungen aller Teiltöne untereinander, die zwischen je zwei benachbarten 64 in der Sekunde betragen, sich also als Rauigkeit geltend machen müssen.

U und O sind zunächst, sobald die erforderlichen Teiltöne hinzukommen, gut, werden aber durch weitere, höhere Teiltöne geradezu alteriert. U bekommt vom 5. Teilton (*e*<sup>1</sup>) an einen Stich nach O und behält ihn. Ebenso O vom 7. Teilton (*b*<sup>1</sup>) an einen Stich nach A. Die durch Interferenz ihrer höheren Teiltöne beraubten Vokale sind also hier besser, charakteristischer als die im Schallzimmer gesungenen: ein Fall, der übrigens auch sonst nicht selten vorkommt. Beim natürlichen Singen auf *C* reicht offenbar die Resonanzeinstellung für U bzw. O nicht hin, um die nächsthöheren starken Teiltöne so weit zu unterdrücken, daß sie nicht etwas von ihrer Vokalvalenz dem Klange beimischen. Durch die If.-Vorrichtung wird dieser Mangel getilgt. Der Wegfall von Schwebungsrauigkeiten trägt zur Verbesserung noch bei.

Ferner ist hervorzuheben, daß sich, wenn auf *C* gesungen wird, bei dem Ö zwischen das gewöhnliche O-Stadium und das erste Auftreten einer Ö-Spur ein Übergang in Gestalt eines Oa einschiebt, und beim Ä an der analogen Stelle ein langes A-Stadium, das bei höheren Grundtönen fehlt (nur einmal, bei einer besonders metallreichen Stimme, kam es auf dem Grundton *e* zum Vorschein). Der Beginn des Ä selbst wird dadurch in abnormer Weise hinaufgeschoben. In unserer Tabelle sind die Anfänge dieser beiden Zwischenstadien durch Punktierung bezeichnet, da man sie schon als Vorformanten betrachten kann.

### 3. Die Formanten.

Die If.-Methode gibt nun auch Mittel an die Hand, um die relative Bedeutung der einzelnen Strukturelemente eines Vokals für seinen spezifischen Vokalcharakter, insbesondere die Tonlage der „Formanten“ zu bestimmen. Wir verstehen unter dem

Formanten nicht einen einzelnen Ton, sondern im allgemeinen eine Strecke des Tongebietes, die zur Charakteristik eines Vokals in besonderem Maße beiträgt. Und zwar nennen wir Hauptformant oder Formant schlechtweg die Strecke, die den überhaupt ausschlaggebendsten Ton eines Lautes einschließt. Dieser bildet, wie der Vergleich mit den Resonanztabellen lehrt, bei den Vokalen ein Stärkemaximum, das nach beiden Seiten von stufenweise abfallenden Teiltonstärken umgeben ist. Es soll aber nicht behauptet werden, daß er zugleich immer der absolut stärkste Ton des ganzen Klangkomplexes sei.

In meinen früher veröffentlichten Einzelabhandlungen bezeichnete ich die maßgebende Strecke als Formantregion, das Formantzentrum als Formant, betonte aber nicht minder wie jetzt das Zusammenwirken der ganzen Formantregion. Die jetzt vorgeschlagene Bezeichnungsweise ist kürzer und praktischer. Den Ausdruck „Formant“ selbst hat HERMANN für die von ihm behaupteten, im allgemeinen unharmonischen, maßgebenden Bestandteile der Vokale eingeführt. Er ist aber so zweckmäßig, daß ich mir erlaubt habe, ihn auch für unsere Darstellung des Sachverhaltes zu verwenden.

Außer dem Hauptformanten unterscheiden wir Neben-(Unter- und Ober-) Formanten, d. h. weitere relative Maxima, die für den spezifischen Vokalcharakter von Bedeutung sind. Die hellen Vokale von Ö bis I haben sämtlich untere Maxima, die von den oberen durch eine mit der Helligkeit des Vokals wachsende leere Strecke getrennt sind und den Gesamtcharakter nicht unwesentlich beeinflussen. Werden sie durch Abbau isoliert, so fragen sie selbst Vokalcharakter, und zwar ist für Ö und E der Unterformant O, für Ä AO, für Ü und I U.

Wir bezeichnen aber nicht alle relativen Maxima, die sich in den Resonanzversuchen fanden, als Formanten oder Nebenformanten. Manche, wie die obersten Maxima der dunklen Vokale, sind ohne Einfluß auf den Vokalcharakter als solchen, wie eben die If.-Methode zeigt.

If.-Versuche geben nun wertvolle Aufschlüsse über die wichtigste Frage der ganzen Vokaltheorie, die nach der Lage der Hauptformanten. Hierbei müssen in 1. Linie die Aufbau-reihen zugrunde gelegt werden, weil man bei ihnen für die untere Grenze, für das erste Auftauchen einer Spur des spezifischen Vokalcharakters empfindlicher ist. Es ist daher bei allen Versuchsreihen auch besonders auf diesen Punkt geachtet worden. Er wird als Formant anfang betrachtet. Bei den hellen Vokalen Ö bis I ist gleichwohl damit zu rechnen, daß der Formant in Wirklichkeit schon 1—2 Töne tiefer beginnt und nur infolge der If.-Breite, die ja nach der Höhe hin zunimmt, nicht eher merklich



wird. Damit stimmt überein, daß regelmäßig schon 1—2 Töne vor jenem Punkt eine gewisse Veränderung eintritt. Ö, Ä werden metallischer, Ü, E, I eigentümlich heiser und schwächer. Dann erst wird ein hohes Element für sich bemerkbar und beginnt jener oben geschilderte zwiespältige Zustand, aus dem das definitive Entwicklungsprodukt hervorgeht. Jene Heiserkeit und Schwächung dürfte darin ihren Grund haben, daß durch den bereits vorhandenen Formantenanfang der Unterformant im Klangganzen geschwächt wird, während doch der diese Wirkung ausübende höhere Bestandteil infolge der If.-Breite noch verdeckt ist. Sie ist also, wenn diese Erklärung zutrifft, ein Produkt des If.-Verfahrens selbst, wenn auch als Phänomen von allgemeinerem Interesse.

Die obere Grenze des Formanten ist ihrer Natur nach etwas schwankend. Mancher findet ein A schon gut, wenn es noch nicht die größte zulässige Helligkeit erreicht hat usf. Auch hier gilt aber, daß die Grenze mit Rücksicht auf die If.-Breite eher etwas tiefer als höher zu legen ist.

Die Abweichungen bei Wiederholung einer Reihe auf einem bestimmten Grundton sind bei hinreichender Übung des Beobachters und Einhaltung der vorgeschriebenen Maßregeln sehr gering; es kann sich da nur um Schwankungen etwa innerhalb einer kleinen Terz handeln. Diese Abweichungen kommen aber auch nur zum kleineren Teil auf Rechnung des Beobachters, zum Hauptteil sind sie objektiv begründet, da die Sänger natürlich auch nicht genau gleichmäßig vokalisieren. Bei einer nur ganz wenig dunkleren Nuancierung des A z. B. wird an der nämlichen Stelle, wo sonst bereits „reines A“ geurteilt wird, noch ein Ao angegeben. Mehr als einmal bezeichnete der Sänger selbst nach Beendigung einer Reihe einen Fall, in dem ihm eine kleine unwillkürliche Abweichung in der Nuance begegnet war, und es zeigte sich zugleich im Protokoll des Beobachters eine entsprechende Abweichung.

Prinzipiell könnte man noch fragen, in welchem Sinne wir eigentlich bei Auf- und Abbauversuchen sagen: „hier beginnt A, hier hört A auf.“ Wir schreiben OA mit zwei Buchstaben, aber es erscheint dem nicht theoretisierenden Beobachter ebenso einfach und wohlcharakterisiert wie A oder O. Was heißt nun also: „reines A“? Wir werden diese Frage nach der Existenz sog. Grundvokale erst später (10. Kap.) besprechen. Hier genügt es, darauf zu verweisen, daß wir die 8 gegenwärtig gebräuchlichen deutschen Vokale zunächst als gegeben hinnehmen und demgemäß zwar A und O, nicht aber OA als eine der Kategorien ansehen, auf die ein gegebener

Eindruck zu beziehen ist. Wir lassen daher A beim Aufbau schon an dem Punkte beginnen, wo O zuerst einen schwachen Anklang an A aufweist, und lassen es beim Abbau dort aufhören, wo ein reines O, ohne jeden Anklang an A, beginnt. Die ideale Aussprache aber ist für uns durch die Forderung maximaler Unterscheidbarkeit bestimmt (o. S. 20).

Aus den folgenden Tafeln lassen wir U hinweg, weil es keinen eigentlichen Formanten oder nur einen beweglichen besitzt. Der Grundton selbst trägt eben U-Charakter.

Von der Regel, daß der Formant aus mehreren Tönen besteht, durch deren Hinzukommen der Vokal von der ersten Spur seiner Eigenart bis zur fertigen Form gelangt, gibt es nur wenige Ausnahmen. So ist für O bei den Grundtönen  $c^1$  und  $ges^1$  mit dem Auftreten des 2. Teiltons Anfang und Ende des Formanten zugleich gegeben. Jeder weiter hinzukommende Teilton würde hier schon die Region überschreiten, die für O-Färbungen noch in Betracht kommt. Der 2. Teilton ist aber in diesen Fällen so kräftig, daß er allein schon im Zusammenwirken mit dem Grundton zur Färbung hinreicht.

Wir sprechen in der umstehenden Tabelle von „Interferenzformanten“ wegen der obigen Klausel betreffs der If.-Breite, die (besonders bei den helleren Vokalen) eine Abweichung der realen Formanten nach unten bis zu einer Terz als möglich erscheinen läßt. Hierüber werden uns erst die synthetischen Versuche Definitives lehren. Aber die relativen Lagen der Formanten treten doch schon hier deutlich zutage.

Die eingeklammerten Tonhöhen bedeuten Abweichungen einzelner Versuchsreihen, wo die untere Grenze etwas tiefer oder die obere etwas höher lag.

Die auffallendste Erscheinung in diesen Tabellen ist das langsame Hinaufrücken der Formanten mit der Höhe des Grundtons. Sie sind also doch nur relativ fest, nämlich im Vergleich mit dem Hinaufrücken der Grundtöne: ihre Bewegung erfolgt in gleicher Richtung, aber weit langsamer, sie umfaßt (abgesehen vom Grundton  $c^2$ ) nur wenige Töne. Die Erklärung dieser Erscheinung wird uns erst bei den allgemeineren Betrachtungen (13. Kap.) beschäftigen. Daß sie bei den Resonanzversuchen nicht zum Vorschein kam, liegt daran, daß dort nur Grundtöne innerhalb eines geringeren Umfangs untersucht wurden. Nur an gewissen Stellen wird man jetzt rückblickend auch dort die Tatsache finden (Vokal E auf  $c^1$  gegenüber  $c$ , S. 25).

## Tafel der Interferenzformanten.

## Vokal O

Grundton	Formant
<i>O</i>	$e^1 - g^1 (c^2)$
<i>Ges</i>	$ges^1 - des^2$
<i>c</i>	$g^1 - c^2$
<i>ges</i>	$ges^1 - des^2$
<i>c</i> <sup>1</sup>	$c^2$
<i>ges</i> <sup>1</sup>	$ges^2$
<i>c</i> <sup>2</sup>	$c^3$

## Vokal A

Grundton	Formant
<i>O</i>	$c^2 - g^2 (e^3)$
<i>Ges</i>	$des^2 - as^2$
<i>c</i>	$e^2 - d^3$
<i>ges</i>	$des^2 - e^3$
<i>c</i> <sup>1</sup>	$g^2 - e^3$
<i>ges</i> <sup>1</sup>	$ges^2 - ges^3$
<i>c</i> <sup>2</sup>	$c^3 - c^4$

## Vokal Ö

Grundton	Formant
<i>O</i>	$f^3 - as^3$
<i>Ges</i>	$as^3 - h^3$
<i>c</i>	$g^3 - d^4$
<i>ges</i>	$as^3 - e^4$
<i>c</i> <sup>1</sup>	$g^3 - d^4$
<i>ges</i> <sup>1</sup>	$b^3 - e^4$
<i>c</i> <sup>2</sup>	$c^4 (e^4) - b^4$

## Vokal Ä

Grundton	Formant
<i>O</i>	$\bar{a}^3 - des^4$
<i>Ges</i>	$b^3 - d^4$
<i>c</i>	$h^3 (c^4) - e^4$
<i>ges</i>	$c^4 - e^4$
<i>c</i> <sup>1</sup>	$c^4 - e^4$
<i>ges</i> <sup>1</sup>	$des^4 - as^4$
<i>c</i> <sup>2</sup>	$e^4 (g^4) - b^4$

## Vokal Ü

Grundton	Formant
<i>O</i>	$ges^3 - d^4$
<i>Ges</i>	$b^3 - d^4$
<i>c</i>	$h^3 - \bar{f}is^4$
<i>ges</i>	$b^3 - ges^4$
<i>c</i> <sup>1</sup>	$c^4 - ges^4$
<i>ges</i> <sup>1</sup>	$des^4 - as^4$
<i>c</i> <sup>2</sup>	$c^4 - \bar{b}^4$

## Vokal E

Grundton	Formant
<i>O</i>	$as^3 - d^4$
<i>Ges</i>	$b^3 - f^4$
<i>c</i>	$h^3 - g^4$
<i>ges</i>	$c^4 - f^4$
<i>c</i> <sup>1</sup>	$c^4 - \bar{f}is^4$
<i>ges</i> <sup>1</sup>	$des^4 - as^4$
<i>c</i> <sup>2</sup>	$e^4 - b^4$

## Vokal I

Grundton	Formant
<i>O</i>	$\bar{a}^3 - d^4$
<i>Ges</i>	$c^4 - ges^4$
<i>c</i>	$des^4 - \bar{a}^4$
<i>ges</i>	$des^4 - b^4$
<i>c</i> <sup>1</sup>	$d^4 - \bar{a}^4$
<i>ges</i> <sup>1</sup>	$e^4 - as^4$
<i>c</i> <sup>2</sup>	$g^4 - \bar{b}^4$

Schon hier aber müssen wir uns vergegenwärtigen (wie schon a. a. O.), daß gewisse kleinere Verrückungen der unteren Formantengrenze, selbst in umgekehrter Richtung wie die des Grundtons, die notwendige Folge sind, wenn die Vokale nur aus harmonischen Teiltönen bestehen<sup>1)</sup>. Beim Grundton *ges* kann z. B. A nicht schon mit  $e^2$  anfangen, sondern frühestens mit  $ges^2$ ; ebenso beim Grundton  $c^1$  erst mit  $g^2$ . O kann, auf *c* gesungen, nicht mit  $c^1$ , sondern erst mit  $g^1$  beginnen, auf  $c^1$  gesungen erst mit  $c^2$ . Der betreffende Ton ist eben immer der erste harmonische Teilton innerhalb des Formantbereiches. Da diese Verschiebungen sich nun wirklich ganz regelmäßig herausgestellt haben, so liegt darin ein sehr starker Beweis, daß in der Tat nur harmonische Teiltöne vorhanden sind. Besonders beweisend aber sind die Fälle, wo mit hinauf-rückendem Grundton der Formantbeginn hinabrückt. Wenn z. B. A auf *e* gesungen wird, beginnt der Formant bei  $e^2$ , wenn es aber auf *ges* gesungen wird, schon bei  $des^2$ . Ganz notwendig, da eben  $des^2$  nicht im *e*-Klang,  $e^2$  nicht im *ges*-Klang enthalten ist. So lösen sich scheinbare Anomalien der Tabellen in Gesetzmäßigkeiten auf.

Daß die Formanten für die helleren Vokale nur so wenig (öfters nur einen Ton oder einen Halbton nach oben oder unten) verschieden sind, könnte wundernehmen. Aber hier ist zu bedenken: 1. daß innerhalb der Formantstrecken Unterschiede in der Intensitätsverteilung sein können und nach unserer weiteren Untersuchung auch wirklich vorhanden sind, 2. und hauptsächlich, daß die Formanten erst im Zusammenwirken mit den übrigen Bestandteilen, namentlich den (in dieser Tabelle nicht angegebenen) Unterformanten ihre Wirksamkeit erlangen und daß diese charakteristische Verschiedenheiten zeigen. Der des E z. B. ist ein O, der des I aber ein U.

Aus der Haupttabelle ergeben sich 2 weitere, die die Sachlage noch etwas näher erläutern. Die 1. gibt den Gesamtumfang der Tonregion, innerhalb deren der Formant eines Vokals bei Grundtönen zwischen  $C$  und  $g^1$  ( $c^2$ ) liegt. Die oberen Grenzen beim Grundton  $c^2$  sind aber wegen der hier schon beginnenden Deformation der Vokale in Klammern gesetzt. Diese Gesamtumfänge sind darum von Bedeutung, weil innerhalb ihrer auf der Tonlinie die „färbenden“ Elemente oder Valenzen liegen müssen, die zu dem betreffenden Vokalcharakter besonders beitragen.

Die 2. Tabelle gibt diejenigen Strecken, die für die Grundtöne *Ges* bis  $g^1$  gemeinschaftlich gelten. Zwischen *Ges* und  $g^1$  liegen die

<sup>1)</sup> Dies hat auch SCHOLE (S. 14) richtig hervorgehoben.

normalen Grundtöne des männlichen und weiblichen Sprechens; daher erscheinen die bezüglichen Formantstrecken besonders wichtig. Zugleich ist die Formantverschiebung mit der Höhe hier nur gering. Diese gemeinschaftlichen Formantstrecken kann man als „Zentralstrecken“ bezeichnen. Sie sind aber nicht die vollständigen Formanten, diese gehen vielmehr, je nach dem Grundton, nach der einen oder anderen Seite darüber hinaus.

Beim O ist hier zu erinnern, daß  $ges^2$  nur für den Grundton  $ges^1$  in Betracht kommt, und daß es den Klang bereits zu einem Oa macht. Ein vollkommenes O ist auf dieser Grundtonhöhe schon nicht mehr möglich, weil eben der erste Oberton bereits in die Gegend des A-Formanten fällt. Deshalb ist diese Note hier eingeklammert.

Vokale	Gesamtumfänge der If-Formanten für die Grundtöne $C-ges^1$ ( $c^2$ )	Zentralstrecken der If-Formanten, gemeinschaftlich den Grundtönen $Ges-ges^1$
O	$e^1-ges^2$ ( $c^3$ )	$g^1-c^2$ ( $ges^2$ )
A	$c^2-ges^3$ ( $c^4$ )	$ges^2-d^3$
Ö	$f^3-e^4$ ( $b^4$ )	$b^3-d^4$
Ä	$b^3-as^4$ ( $b^4$ )	$des^4-es^4$
Ü	$ges^3-as^4$ ( $b^4$ )	$des^4-es^4$
E	$ges^3-as^4$ ( $b^4$ )	$des^4-f^4$
I	$c^4-as^4$ ( $b^4$ )	$e^4-g^4$

Außer den eben besprochenen Aufbaureihen wurden noch 6 andere für jeden Vokal durchgeführt, bei denen die Leitung, um jede Spur des direkten Hörens auszuschließen und dadurch unwissentliche Versuche zu ermöglichen, durch 4 Wände in ein weit entferntes Zimmer geführt wurde (s. das Schema S. 44 von  $S_1$  bis  $B_1$ , Zimmer I—V). Von den 6 Reihen wurden je 2 auf den Grundtönen  $c$ ,  $ges$  und  $c^1$  ausgeführt. Als Sänger wirkte Dr. ALFRED GUTTMANN, der eine hervorragende künstlerische Ausbildung besitzt (er kann z. B. 5 deutlich verschiedene U hervorbringen), als Beobachter Dr. RIEFFERT. Die Versuche wurden mit besonderer Sorgfalt durchgeführt. Ein Beispiel einer solchen unwissentlichen Reihe möge zeigen, wie genau erkennbar doch auch hierbei die oben festgestellten Umwandlungen zutage traten. Es handelte sich hier um Querschnittreihen, d. h. es wurden bei jeder Einstellung des Röhrensystems alle Vokale, aber in unregelmäßig wechselnder Anordnung gegeben. Jeder Einzelfall wurde, um dem Urteil möglichste Sicherheit und Eindeutigkeit zu geben, mehrere Male wiederholt. Die Abweichungen waren dabei nur ganz gering. In der Tabelle ist nur das dominierende Urteil angegeben, doch sind in zwei besonders interessanten Fällen, bei dem Blöklaut in der Entwicklung des Ä, auch Beispiele für die Varianten der Aussage angeführt. Eine wirkliche Abweichung gegenüber den wissentlichen Versuchen fand nur in einem einzigen Falle statt, beim Hinzukommen des 8. Teiltons im Aufbau des Ä. Sie ist offenbar zufälliger Art; der hier noch sehr schwache helle Bestandteil war unbeachtet geblieben.

Unwissentliche Aufbaureihe auf dem Grundton  $c^1$ .

$c^5$								
$h^4$							E	
$\bar{b}^4$						Ü	EÖ	Iu
$\bar{a}^4$							ÖEö	Ui, IU
$g^4$							EÖ	
$\bar{f}is^4$				Ö	Ä		ÖE	
$e^4$					Oä <sup>2)</sup>	ÜÜ	ÖOu	
$d^4$				Öü	ÄÖ <sup>1)</sup>			
$c^4$					Ao			
$\bar{b}^2$				Aö	AOö			
$g^3$				Ao				
$e^3$			A	Oa				
$c^3$			AO		Oa			
$g^2$				OU			OU	
$c^2$		Ou (O)	O	UO	O		Uo	
$c^1$	U	U	U	U	U	U	U	U
Teilton	U	O	A	Ö	Ä	Ü	E	I

I behielt hier immer, auch beim Hineinschieben aller Seitenröhren, etwas Dunkles (zuletzt Iu). Auch O kam meistens nicht über ein Ou hinaus.

Bezüglich der Formanten weichen die unwissentlichen Versuche begrifflicherweise etwas mehr als die wissentlichen untereinander ab, da die Unwissentlichkeit hier eine außerordentliche Erschwerung bedeutet. Aber die Gesamtumfänge der Formanten stimmen doch recht gut mit denen bei kurzer Leitung überein. Die Zentralstrecken sind, wenn wir hier noch die Bedingung stellen, daß sie mindestens 4 von den 6 Reihen gemeinschaftlich sein müssen, folgende: für O  $g^1-c^2$  (hier wieder mit der Maßgabe, daß ein einziger Ton dieser Strecke schon genügen kann, bei  $g$   $g^1$ , bei  $c^1$   $c^2$ ), für A  $ges^2-e^3$ , Ö  $a^3-d^4$ , Ä  $h^3-e^4$ , Ü  $es^4-g^4$ , E  $es^4-g^4$ , I  $ges^4-b^4$ . Sollen alle 6 Reihen die gemeinschaftlichen Strecken enthalten, so verengern sich diese noch. Verglichen mit den Zentralstrecken bei kurzer Leitung zeigt sich im allgemeinen auch wieder Übereinstimmung, doch liegen sie bei den 3 hellsten Vokalen hier um einen Ganzton höher. Dies könnte an der kunstvollen Vokalisierung des Sängers gelegen haben. Wahrscheinlicher ist es mir aber, daß infolge der Schwächung aller Teiltöne durch die lange Leitung der untere Abschnitt der Formantzone, der die ersten Spuren des betreffenden Vokals enthält, nicht zur Wirksamkeit kam und erst noch weitere Teiltöne des Formanten hinzutreten mußten, um das Charakteristische merklich zu machen. Für ein abnorm feines Ohr würde der Formant vielleicht auch unter diesen Umständen an derselben Stelle liegen wie bei der kurzen Leitung. Immerhin bewog mich diese Überlegung, wieder zur kurzen Leitung zurückzukehren.

<sup>1)</sup> Einzelurteile: AOe, Aeo, OeA, ÄÖ, ÄÖ.

<sup>2)</sup> Einzelurteile: ÄO, Oä, Oä, Oä, OÄ.

4. Die Erscheinungen beim Grundton  $c^2$ .

Geht man mit dem gesungenen Grundton über die Zone  $c$ — $ges^1$  nach oben hinaus, so erfolgen immer umfassendere Zerstörungen der Vokalstruktur. Es wurde in dieser Hinsicht allerdings nur  $c^2$  mit der If.-Methode untersucht; aber die Versuche des nächsten Kapitels lassen keinen Zweifel, daß die Rückbildung von da ab enorm zunimmt und daß von  $c^3$  ab die Vokale überhaupt nicht mehr voneinander unterschieden werden können.

Schon bei  $ges^1$  ist U nicht mehr recht dunkel, etwas nach O hin verändert, ebenso O nach A hin. Der O-Formant ist ja auch auf  $ges^2$  hinaufgerückt, wo zugleich der A-Formant beginnt (siehe Tabelle S. 66). Im übrigen ist aber hier noch kaum etwas Abnormes zu finden und sind die Vokale noch deutlich ausgeprägt.

Anders bei  $c^2$ . Hier verschieben sich fast alle Formanten stark nach oben, sie behalten auch nicht mehr ganz ihre frühere relative Lage gegeneinander, und die helleren Vokale werden, soweit überhaupt, eigentlich erst mit dem Hineinschieben sämtlicher Röhren fertig. Selbst dann sind sie nicht mehr so vollkommen wie in tieferen Lagen, sondern einander vielfach ähnlich, namentlich Ö und E dem Ä. Auch Ü und I lassen an Deutlichkeit zu wünschen, sie behalten etwas Dunkles, I nähert sich dem E oder Ü. Die If.-Versuche sind darum hier viel schwerer. Es scheint hier allerdings auch die Röhrenleitung mehr als bei tieferen Grundtönen alterierend zu wirken. Aber daß sie nicht die Hauptursache ist, lehrten unwissentliche Erkennungsversuche ohne Röhrenleitung, bei denen der Singende nur durch eine halboffene Türe vom Hörenden getrennt war: auch da wurde auf  $c^2$  z. B. Ö leicht mit Ä verwechselt. Doch ließ sich selbst eine unwissentliche Querschnittreihe mit teilweise sogar klareren Ergebnissen gegenüber den wissentlichen durchführen.

Was den qualitativen Verlauf der Aufbaureihen betrifft, so sind zunächst, solange alle Obertöne ausgeschlossen sind, natürlich wieder sämtliche Vokale einander gleich, aber man kann zweifelhaft sein, ob man den Laut, den einfachen Ton  $c^2$ , noch als ein gutes U bezeichnen soll. Immerhin steht er dem Eindruck eines sehr hellen U am nächsten.

Das gesungene U selbst geht dann beim Hinzukommen des 2. Teiltons in einen O-verwandten Laut über und bleibt so bei allen weiteren Einstellungen.

O geht in ein OA über und bleibt dabei. Der Formant schrumpft hier, wie schon auf dem Grundton  $ges^1$ , in einen Punkt zusammen, der Vokal ist mit dem Auftreten des Teiltone  $c^3$  in der Form

da, die er weiterhin behält. A wird kenntlich mit dem 2., zu einem guten A mit dem 4., zu einem sehr guten mit dem 6. Teilton.

Die Entwicklung der übrigen sei nicht weiter geschildert. Nur soviel, daß die einander benachbarten hellen Vokale auf lange Strecken hin schlecht unterscheidbar bleiben. Allgemein kann man sagen: die dunklen Vokale rücken gegen die hellen hin und diese rücken näher aneinander.

Die Versuchsreihen boten infolgedessen den Beobachtern Schwierigkeiten und Verlegenheiten, und die Ergebnisse waren lange nicht so übereinstimmend wie sonst. Es traten unverkennbare „Perseverationen“ (Fälle des Beharrens auf einem einmal gefällten Urteil auch bei veränderten Umständen) u. dgl. ein. Bezeichnend auch, daß bei den helleren Vokalen fast regelmäßig um Wiederholung, auch mehrfache, gebeten wurde. Immerhin blieben die Vokale nach Hineinschieben aller Seitenröhren im wesentlichen erkennbar und unterscheidbar.

#### 5. Lücken- und Stichversuche.

Lassen die Ab- und Aufbaureihen die Vokalstrukturen nach den Grundzügen und besonders nach der Lage der Formanten erkennen, so gewähren Lückenversuche, bei denen mitten heraus einzelne Teiltöne weggenommen werden, weitere Einblicke.

a) Zuerst sei eine Klasse erwähnt, die ich „Brückenversuche“ nannte. Sie bestehen darin, daß die untersten Teiltöne und der If.-Formant erhalten bleiben, dagegen eine dazwischliegende Strecke (die „Brücke“) ausgeschaltet wird. Damit die beobachtete Veränderung nicht auf mitausgeschaltete Multipla jenseits des Formanten bezogen werden kann, wird von vornherein der ganze über dem Formanten liegende Teil ausgeschaltet. Die Laute werden dadurch, wenn überhaupt, nur wenig geschädigt, bleiben jedenfalls durchaus kenntlich. Und nun vergleicht man den Eindruck ohne und mit Ausschaltung der Brücke. Man muß hier aber die Grundtöne und die Vokale selbst so wählen, daß durch Wegnahme der Brücke nicht auch Formantbestandteile als ungerade Multipla mit weggenommen werden. Es eignen sich dazu besonders die Vokale Ö und Ä auf Grundtönen aus der Mitte der kleinen Oktave.

Bei der Deutung der Beobachtungen ist jedoch damit zu rechnen, daß der wirkliche Formant wegen der If.-Breite um 1—2 Töne tiefer hinabreicht als der If.-Formant. Es können daher die tiefsten 1—2 Töne des realen Formanten als Bestandteile der „Brücke“ mitweggenommen sein.



α) Ö auf *ges*. Vorhanden die Teiltöne  $1-4 = ges-ges^2$  und der If.-Formant  $as^3-es^4$  (10-14), ausgeschlossen  $5-8 = b^2-ges^3$  und die Töne über  $es^4$ . Man hört O. Die Brücke ist also ganz wesentlich.

β) Ä auf *ges*. Vorhanden  $1-5 = ges-b^2$  und der Formant  $c^4-e^4$  (11-14), ausgeschlossen das übrige. Man hört AO. Wird auch der Formant selbst oder nur dieser ausgeschaltet, nicht wesentlich anders. Die Ausschaltung der Brücke wirkt also hier ebenso wie die des Formanten, während ihr Hinzukommen bei Aufbaureihen den Vokal noch lange nicht fertig macht.

γ) Ö auf *as*. Vorhanden  $1-4 = as-as^2$  und der Formant  $b^3-e^4$  (9-13), fehlend das übrige. Man hört ein gutes O. (Hier war allerdings das weggenommene  $as^3$  schon Formantton; ich hatte damals den Formanten höher angesetzt.)

Setzt man hier die aus 4 Tönen  $c^3-as^3$  bestehende Brücke Ton für Ton wieder ein, so erhält man ein sehr merkwürdiges Bild: mag man sie in der Folge von unten nach oben oder umgekehrt einfügen, immer tritt bei den 3 ersten nur eine geringe Veränderung auf, dagegen erscheint beim Hinzutreten des letzten fast sprunghaft ein gutes Ö. Besonders auffallend ist dies bei Einfügung von unten herauf, also in der Reihenfolge  $c^3, es^3, ges^3, as^3$ . Bei  $c^3$  und  $es^3$  hört man noch O, bei  $ges^3$  wird es metallisch, bei  $as^3$  plötzlich ein Ö.

Offenbar ist also Vollständigkeit und Kontinuität der Teiltonreihe innerhalb einer bestimmten Region erforderlich. Erst in dem Moment, wo sie hergestellt ist, erscheint der ausgesprochene Vokalcharakter. Daß dies bei Einfügung von unten nach oben besonders auffällt, erklärt sich daraus, daß man sich dabei eben dem Formanten nähert oder ihn sogar erreicht. Man kann hier auch einen einfachen Stichversuch ausführen, indem man aus der vollständigen Reihe 1-13 nur 8 ( $as^3$ ) herausnimmt: sofort geht der Laut in O über.

Allerdings ist zu bedenken, daß bei Stichversuchen mit Obertönen von höherer Ordnungszahl, die nahe beisammenliegen, niemals nur ein einziger ausgeschlossen wird, sondern infolge der If.-Breite auch seine nächsten Nachbarn, also mit  $as^3$  zugleich  $ges^3$  und  $b^3$ . Zum mindesten werden sie bedeutend abgeschwächt. Der wirkliche Unterschied vor und nach dem Einstich ist also doch auf eine, wenn auch kleine, Zone zu deuten. Gleichwohl bleibt die Erscheinung lehrreich, zumal da sie auch bei umgekehrter Richtung der Einfügungen, wenngleich in geringerem Grade, auftritt.

δ) Ä auf *as*. Vorhanden  $1-4 = as-as^2$  und der If.-Formant  $c^4-f^4$ , fehlend das übrige. Man hört ein dunkles, gedämpftes A, genau so wie wenn die ausgeschlossenen Töne vorhanden sind, der Formant aber fehlt.

Man kann hier die Brücke Ton für Ton einfügen: dasselbe Ergebnis wie bei γ). Erst mit dem letzten Ton von unten nach oben ( $b^3$ ) erscheint plötzlich Ä (nur etwas dunkel, wegen der fehlenden höheren Töne, die hier noch etwas beitragen). In umgekehrter Folge wieder nicht so unvermittelt: mit  $b^3$  ist es noch Ao, mit  $as^3$  schon dunkles Ä, das dann noch besser wird. Aber paradox ist doch zunächst auch dies, daß der Laut durch Hinzukommen immer tieferer Töne immer heller wird. Es versteht sich nur daraus, daß eben der ganze Komplex zusammenwirken muß und daß seine Kontinuität auch so erst allmählich hergestellt wird.

Bei Ü, E, I ist die Wirkung der Brückenversuche viel weniger auffallend, weil hier unterhalb der Formanten die ausgedehnten „toten Strecken“ liegen. Doch ist noch ein Unterschied zu bemerken, wenn man z. B. bei

dem I auf *as* die If. auf die Töne *ges*<sup>3</sup>, *as*<sup>3</sup>, *b*<sup>3</sup> einstellt. Die Strecken scheinen also doch nicht ganz leer an schwachen Teiltönen zu sein (vgl. die Resonanzversuche).

b) Lückenversuche über die hohen Teiltöne des A.

Wenn man bei dem A einer Männerstimme alle Teiltöne bis dicht unter *c*<sup>4</sup> ausschaltet, so hört man noch ein eigentümliches feines Knistern, Wispern oder Zwitschern, auf welches bereits KOEHLER (I, II, S. 73) aufmerksam geworden ist. Es kann nur von den hohen Teiltönen der 4-gestrichenen Oktave herrühren, die nicht als Multipla mitausgeschaltet sind (vgl. o. S. 48). KOEHLER fand es auch bei den dunklen Vokalen. Es verschwindet, sobald man die If. auch auf die 4-gestrichene Oktave erweitert. So führt auch die If.-Methode auf diese hohen Teiltöne, die aber keine wesentliche Rolle spielen können, da man beim Abbau, wenn die 4-gestrichene Oktave von oben her abgetragen wird, noch keine Alteration dieser Vokale beobachten kann.

c) Stich- und Lückenversuche über die Wirkung einzelner Teile des A-Formanten.

α) Grundton *c*<sup>1</sup>. Vorhanden 1–7 = *c*<sup>1</sup>–*b*<sup>3</sup>; alles Höhere ausgeschlossen. Man hört ein gutes A. Wird nun hier auf den einzelnen Ton *g*<sup>2</sup> If. eingestellt, so ist der Unterschied enorm, A geht in ein helles O oder AO über. Kein Stich an einer anderen Stelle des Formanten hat eine so eingreifende Wirkung. Allerdings ist auch hier nicht zu vergessen, daß gleichzeitig *e*<sup>2</sup> und *b*<sup>2</sup> bedeutend geschwächt werden.

β) Grundton *as*. Vorhanden 1–9 = *as*–*b*<sup>3</sup>; fehlend alles Höhere. Man hört ein gutes A. Wird If. auf 3 und 4 (*es*<sup>2</sup> und *as*<sup>2</sup>) eingestellt, so entsteht ein Oö (ein Beobachter schreibt Öo + ü). Durch den Wegfall dieser beiden Töne ist A zerstört, aber in der Mischung machen sich jetzt an Stelle der A-haltigen die angrenzenden O- und Ö-haltigen Teiltöne, *as*<sup>1</sup> und *as*<sup>3</sup> + *b*<sup>3</sup>, geltend. Die Gesamtstruktur entspricht hierbei in der Tat einem dunklen Ö, die Lücke bei *es*<sup>2</sup>–*as*<sup>2</sup> ist die erforderliche tote Strecke.

γ) Grundton *as*. Vorhanden 1, 2, 4 = *as*, *as*<sup>1</sup>, *as*<sup>2</sup>. Man hört ein gutes O. Dies ist merkwürdig, weil *as*<sup>2</sup> im Formantzentrum des A liegt, das in die Leitung gesungen wird. Aber es zeigt, daß eben ein einzelner Ton der Formantgegend im allgemeinen nicht genügt. Sobald 3 = *es*<sup>2</sup> hinzutritt, entsteht AO, obgleich 1, 2 und 3 allein auch nur ein O, höchstens Oa geben. Natürlich sind auch die Intensitätsverhältnisse hier speziell die in einem gesungenen A gegebenen und darf man das Ergebnis nicht darüber hinaus verallgemeinern.

δ) Grundton *ges*. Vorhanden 1–8 (*ges*–*ges*<sup>3</sup>). Gutes A. Werden *des*<sup>2</sup> und *ges*<sup>2</sup> ausgeschaltet, so entsteht ein eigentümlich nasaliertes OA, das ein Beobachter mit dem französischen „en“ verglich. Wird noch *b*<sup>2</sup> ausgeschaltet, so wird es zu einem „tiefen, gewürgten Ö“.

ε) Grundton *es*. Vorhanden 1–8 (*es*–*es*<sup>3</sup>). Wird If. auf 3–5 (*b*<sup>1</sup>–*g*<sup>2</sup>) eingestellt, so entsteht ein etwas blökendes und gedämpftes O.

ζ) Grundton *c*. Vorhanden 1–3, 7–10 (also *c*, *c*<sup>1</sup>, *g*<sup>1</sup>, *b*<sup>2</sup>, *c*<sup>3</sup>, *d*<sup>3</sup>, *e*<sup>3</sup>). Eigentümlich nasal verhülltes Ae (A mit E-Einschlag). Kein bekannter Vokal, höchstens dem schnarrenden A ähnlich, das man gelegentlich in

Norddeutschland zu hören bekommt. Wird jetzt  $g^2$  eingefügt, so wird ein gutes, aber gellendes A daraus. Durch  $e^2$  und  $c^2$  wird es noch verbessert („Brücke“). Die Hauptwirkung übt also wieder  $g^2$ .

d) Umwandlungen des A durch Lückenversuche.

Hier wurden die über dem Formanten liegenden Bestandteile nicht von vornherein ausgeschaltet; es wurden also durch Wegnahme tieferer Teiltöne deren ungerade Multipla mitweggenommen, während die übrigen höheren Teiltöne erhalten blieben. In dieser Weise kann man die wunderlichsten Metamorphosen erzielen, z. B. aus einem A ein  $\ddot{O}^1$ , ja ein E und I machen.

Grundton *as*. If. auf  $3-8 (es^2-as^3)$ . Man hört zunächst ein frappantes E; doch kann man denselben Laut auch als ein O mit sehr aufdringlichen hohen Obertönen hören. Dr. GUTTMANN schrieb ihn Oö. Bei If. auf  $2-8 (as^1-as^3)$  hört man I mit starkem dunklen Grund. Dr. G. schrieb Uo + i.

Berechnet man hier die mitausgeschalteten ungeraden Multipla, so findet sich, daß im 1. Falle doch 7, im 2. 5 Töne von  $c^4$  bis  $h^4$  erhalten blieben. Sind es auch infolge der If.-Breite einige weniger, so bleiben doch jedenfalls mehrere Töne dieser Zone übrig. Es muß daher, da der A-Formant weggenommen und dadurch eine tote Strecke gebildet ist, in der Tat ein ähnlicher Komplex entstehen, wie bei einem E und I. Der Unterschied der beiden Fälle rührt offenbar daher, daß der Untergrund durch Wegfall des 2. Teiltones aus einem O in ein U übergeht, das die Grundlage des I-Klanges bildet. So lassen sich selbst diese seltsamen Wandlungen aus dem Vorangegangenen gut begreifen. Sie werden uns später noch verständlicher werden.

Schaltet man aus dem A auf *as* die Töne  $es^2$  und  $as^2$  nebst ihren Multiplis aus, so entsteht ein Oö. Nimmt man auch noch  $as^1$  heraus, so resultiert ein wunderlich blökender Laut, „U + etwas Hohem, undefinierbarem“. Hier wird der Charakter auch durch den Wegfall der Multipla von  $as^1 (es^3, c^4)$  zerstört und im unteren Abschnitt des Klanges der Zusammenhang zerrissen; man hört die beiden Gruppen nebeneinander, der Laut ist gespalten.

So kann man aus einem A durch If.-Versuche geradezu alle sog. Hauptvokale sowie Umlaute und allerlei interessante Neubildungen erzeugen. Gelegentlich erhält man mit Lückenversuchen auch Klänge von instrumentalem, klarinetten- oder fagottähnlichem Charakter. Auch aus einem  $\ddot{O}$  auf *ges* erhielt ich durch If. auf den Teiltönen 2,4 und 6 einen schönen leisen Klarinettenklang. Solche Wirkungen erhält man aber niemals bei den systematischen Auf- und Abbaureihen.

e) Stichversuche zur Ermittlung der Formantzentren.

Man kann, nachdem die Lage der Formanten bekannt ist, innerhalb derselben die einzelne Röhreneinstellung aufsuchen, durch welche die größte Veränderung des Vokals erfolgt. Einiges aus dem vorangehenden gehört schon dahin. Der Ausdruck „Formantzentrum“ soll aber nicht eigentlich im lokalen, sondern im dynamischen Sinne verstanden werden. Es bleibt denkbar, daß ein Ton, der nicht genau in der Mitte der Formantstrecke liegt, doch

<sup>1)</sup> Darauf hat bereits KOEHLER hingewiesen (I, II, S. 70).

die entscheidendste Strukturbedeutung besäße. Im allgemeinen freilich ist anzunehmen, daß es sich um eine mittlere Partie des Formanten handeln wird.

Man stellt zweckmäßig immer 2 Stempel in gleicher Länge ein, um der Auslöschung sicher zu sein, und zieht bei jedem Einzelversuch den unveränderten Vokal durch abwechselndes Heraus- und Hereinschieben der Stempel immer zur Vergleichung heran.

Eine mit einem guten Sänger durchgeführte Versuchsreihe auf den Grundtönen  $c$  und  $c^1$  ergab für alle 8 Vokale Zentren, die mit den später bei den synthetischen Versuchen zu erwähnenden übereinstimmen. Nur das für A lag hier auf  $b^2-c^3$ , der Vokal muß ungewöhnlich hell gegeben worden sein. Nebenbei war interessant, daß beim U auf dem Grundton  $c$ , das am meisten durch Herausnahme von  $c^1$  geschädigt wurde, der Gesamtklang etwas Hohles, Klarinettenartiges erhielt; ferner daß beim U auf  $c^1$  die Wegnahme eines einzelnen Teiltones keine eigentliche Schädigung, sondern nur eine Schwächung bewirkte. Der Grundton wird eben hier bei objektiver Ausschaltung subjektiv als Differenzton der Obertöne erzeugt.

Bei diesen feinsten Stichversuchen kann nun aber der Unterschied in der If.-Wirkung, je nachdem die Röhre in einem Bauch oder einem Knoten der Welle einsetzt, eine wesentliche Rolle spielen. Gewisse Unregelmäßigkeiten schienen darauf hinzudeuten. Die Röhren wären daher durchweg nur an Bäuchen der betreffenden Wellen einzustellen. Ich habe diese Versuche aber nicht weitergeführt, da die synthetische Methode einfachere Mittel zur Bestimmung der Formantzentren bietet.

## 6. Isolierversuche zur Bestimmung von Teiltonstärken.

Die relative Stärke des Grundtons gesungener Vokale ist im allgemeinen gering; aber gewisse Unterschiede sind regelmäßig zu bemerken. Insbesondere ist unverkennbar, daß sie in der Vokalreihe U, O, A, E, I von U bis A ab-, dann wieder bis I zunimmt und bei I etwa ebenso groß ist wie bei U. Dies entspricht ganz den Befunden bei der Resonanzmethode (S. 27).

Beispielsweise waren die Stärkezahlen (hier wieder mit 4 multipliziert) für den Grundton *gis* (200 Schw.) die folgenden:

U	O	A	E	I
8	7	5	7	8

Ö hatte die Grundtonstärke 7, Ä 6, Ü 8.

Wurde die Stärke des Grundtones eines auf *ges* gesungenen U mit der Grundtonstärke einer gleich hohen losen Stimmgabel verglichen, so war von dieser sogar bei starkem Anschlag<sup>1)</sup> über-

<sup>1)</sup> Um konstante und wiederherstellbare Stärkegrade zu erzielen, benutze ich seit langer Zeit Holzstücke von nebenstehender Form, in welche die Zinken der Gabel eingeklemmt werden. Je nach der Enge der Klammer ist dann der Ton beim Abziehen von verschiedener Stärke. Sehr starke Schwingungen kann man aber bei Resonanzgabeln nur durch Streichen mit einem Baßbogen erzielen.



hauptsächlich nichts aus der Leitung zu hören, der Grundton des U also viel stärker. Wurde dagegen ein U auf  $c^1$  gesungen und mit einer aus gleicher Entfernung vor die Leitung gehaltenen kräftig angestrichenen Resonanzgabel  $c^1$  verglichen, so waren die Grundtöne bei Ausschaltung aller Obertöne ungefähr gleich stark und die beiden Klänge überhaupt nicht zu unterscheiden, wenn die Leitung erst nach dem Beginn freigegeben wurde. Erst mit dem 2. und besonders dem 3. Teilton ( $g^2$ ) trat der Unterschied hervor, indem der Vokal als das U einer menschlichen Stimme erkannt wurde, während der Gabelklang unverändert blieb.

Außer dem Grundton kann man prinzipiell auch seine Oktaven durch If. isolieren (o. S. 38) und ihre relative Stärke bestimmen. Praktisch ließ sich dies wenigstens für die beiden ersten Oktaven noch gut ausführen; die höheren wurden bei Vokalen, in denen sie nur schwach vorhanden sind, durch benachbarte Einstellungen ausgelöscht. Solche Versuche wurden für alle Vokale auf den Grundtönen  $c$ ,  $gis$ ,  $c^1$ ,  $ges^1$ ,  $c^2$  gemacht. Im ganzen stimmten auch hier die Ergebnisse mit denen der Resonanzversuche. Aber wir verzichten darauf, die Tabellen wiederzugeben, da die Oktaven doch nur Bruchstücke der ganzen Teiltonreihen sind, und die Synthese auch über die Stärkeverhältnisse genauere Aufschlüsse gibt.

### 3. Kapitel.

## Das Unkenntlichwerden der Vokale in der höheren Sopranlage.

Diese Tatsache, der wir sowohl bei den Resonanz- wie den Interferenzversuchen begegneten, verdient wegen ihrer theoretischen und praktischen Wichtigkeit eine besondere Untersuchung, die sich aber zunächst nur auf die genaue Feststellung und Beschreibung richten soll, während die Erklärung uns erst im 13. Kap. beschäftigt wird.

Es gibt Sängerinnen, die sich mit Leichtigkeit in der 3-gestrichenen Oktave bewegen. Aber Komponisten pflegen für Arien in dieser Lage keine Texte, sondern nur Koloraturen auf irgendeiner Silbe vorzuschreiben (Königin der Nacht). Dennoch scheint die Tatsache, daß in hoher Sopranlage die Vokale ununterscheidbar werden, vielen Musikern unbekannt. Selbst einem berühmten Musikforscher, der sich auch als Komponist und Dirigent vielfach betätigt hatte, war sie fremd, ja er wollte nicht daran glauben. Gewöhnlich kennt man eben den Text schon einigermaßen, weiß wenigstens, um was es sich handelt; auch die Sprache selbst gibt durch die Konsonanten und die Akzéntverteilung Anhaltspunkte, weiteres läßt der Zusammenhang erraten, endlich hilft auch der Anblick der wechselnden Mundöffnungen und der Gebärden mit.

HELMHOLTZ ist aber unsere Tatsache nicht ganz entgangen. Er folgerte sie für U, O und A sogar aus seiner Vokallehre und fand die Deduktion durch Beobachtung bestätigt<sup>1)</sup>. HERMANN hat sie

<sup>1)</sup> S. 184: „Oberhalb des  $f^1$  wird die Charakterisierung des U . . . unvollkommen werden. Solange es aber der einzige unbestimmt klingende Vokal ist und die anderen noch merkliche Verstärkung gewisser Regionen ihrer Obertonreihe hören lassen, wird dieser negative Charakter das U auszeichnen. Für Soprane wird dagegen in der Gegend des  $f^2$  die Unterscheidung des U, O und A undeutlich werden müssen, was meiner Erfahrung nach in der Tat der Fall ist.“ Von den helleren Vokalen sagt hier H. nichts.

dann zuerst statistisch untersucht<sup>1)</sup>. Wie er richtig bemerkt, ist es nicht zweckmäßig, Sängerinnen selbst darüber zu fragen, da sie sich bewußt sind, die Vokale richtig zu intendieren. Man läuft Gefahr, die Künstlerinnen zu kränken, ohne sie zu überzeugen. Doch empfehlen nach HERMANN manche Gesanglehrer, in der höchsten Lage immer nur ein A zu intendieren. Er hat nun an sehr geschulten Sängerinnen, darunter 2 ausgezeichneten Opernsängerinnen, Beobachtungen angestellt und dazu auch gelegentlich Mitbeobachter herangezogen.

Die zu singenden Töne waren  $g^1$ ,  $e^2$ ,  $g^2$ ,  $a^2$ . Die Vokale wurden in bunten Reihen zu je 15 unwissentlich dargeboten. Bei 3 Sängerinnen sind die Ergebnisse für die 3 letzten Noten tabellarisiert<sup>2)</sup>. Wir geben die Tabellen in der Form wieder, wie wir nachher unsere eigenen darstellen werden, beschränken uns aber auf die 5 „Hauptvokale“, deren jeder in der 15-gliedrigen Reihe in der Regel 2—3mal vorkam. A wurde öfter als die übrigen gegeben. Die linke Vertikalreihe gibt die von der Sängerin gesungenen bzw. intendierten Vokale, rechts davon stehen die Anzahlen der Urteile, die auf jeden der in der Horizontalreihe stehenden Vokale entfallen. Die Diagonale dieser Zahlen von links oben nach rechts unten gibt die richtigen Urteile, alle übrigen sind Verwechslungen. Die Gesamtzahl der Urteile ist jedesmal 12<sup>3)</sup>. Die Sängerin IV war Sopran, V Mezzosopran, VI Alt.

Tonhöhe  $e^2$ .

	U O A E I	U O A E I	U O A E I
U	1 1	1	1
O	2	2	3
A	2	3	4
E	3	1 2	1
I	2 1	2 1	1 2
Sängerin IV: $\frac{r}{n} = \frac{9}{12}$		Sängerin V: $\frac{r}{n} = \frac{9}{12}$	
			Sängerin VI: $\frac{r}{n} = \frac{10}{12}$

<sup>1)</sup> Bd. 141, S. 11 ff. 1911; vgl. Bd. 91, S. 155 ff. 1902.

<sup>2)</sup> Über den Ausfall bei  $g^1$  ist nichts berichtet. Leider hat H. die einzelnen Vokale nicht immer gleich oft dargeboten.

<sup>3)</sup> Bei  $g^2$ , Sängerin VI, stehen nur 11, weil E 1 mal mit Ö verwechselt wurde, was nicht in unsere Tabellen, aber in die Summenzahl der Urteile aufzunehmen war.

Tonhöhe  $g^2$ .Tonhöhe  $a^2$ .

	U O A E I	U O A E I	U O A E I
U	0 2	1 1	0 1 1
O	2	2 0	1 1
A	3	1 2	3
E	2	1	2 0
I	3 0	3 0	3 0
Sängerin IV: $\frac{r}{n} = \frac{7}{12}$		Sängerin VI: $\frac{r}{n} = \frac{4}{12}$	
			Sängerin V: $\frac{r}{n} = \frac{4}{12}$

Die Verschlechterung von  $e^2$  zu  $g^2$  und  $a^2$  ist bei allen 3 Sängerrinnen offenbar. Hauptsächlich handelt sich's um Verwechslungen zwischen I und E, U und O. Von allen U und I wurden überhaupt auf den höheren Tönen nur ein einziges U richtig verstanden. Beim I zeigte sich eine Neigung zur Verwechslung mit E schon auf  $e^2$ . Bei der Sängerin V erschienen auf  $a^2$  fast alle Vokale als A. Auch Ä und Ö (hier nicht aufgenommen) wurden als A gehört. Bei IV und VI hingegen wurde niemals ein anderer Vokal mit A verwechselt. Individuelle Unterschiede sind hier ja auch zu erwarten.

HERMANN findet die Ergebnisse paradox, und nach seiner Theorie mit Recht. Bei I, das den meisten Verwechslungen ausgesetzt ist, liege der Formant ganz sicher im Anfang der 4-gestrichenen Oktave, könne also theoretisch nicht durch Intonation von  $e^2$  oder  $a^2$  beeinflusst werden. Umgekehrt solle man erwarten, daß A in dieser Lage schwer oder gar nicht hervorzubringen wäre; denn da bei  $g^2$  der Formant mit der Stimmnote zusammenfalle, müsse der Vokal seine Charakteristik fast ganz verlieren. Diese Schwierigkeiten werden für HERMANN besonders dadurch unlösbar, daß seine Formanten unharmonische Bestandteile des Klanges sind, die nicht wie die harmonischen in Abhängigkeit vom Grundton stehen. Daher sieht man nicht ein, warum z. B. bei  $e^2$  der A-Formant  $g^2$  nicht ebensogut wie bei jedem anderen Grundton vorhanden sein könnte.

Ohne noch HERMANN'S Statistik zu kennen, habe ich selbst 1914 solche Studien angestellt und berichte darüber im folgenden.

Zunächst ließ ich die beiden schon früher erwähnten musikalischen Schwestern A. und H. PLANCK, die hohe, etwas feine Stimmen von sehr ähnlicher Klangfarbe hatten (A. nur etwas



voller) und rein intonierten, ohne jedoch kunstmäßig geschult zu sein, die 5 Töne



auf den 5 „Hauptvokalen“ in verschiedener Ordnung singen. Es waren 6 Folgen zur Verfügung gestellt, von denen eine beliebige ohne Wissen des Beobachters gewählt wurde. Der Beobachter, ich selbst, stand natürlich mit dem Rücken gegen die Sängerinnen. Es war mir nun nicht möglich, unter diesen Umständen irgend etwas über die Reihenfolge herauszubekommen. Ein Vokal klang wie der andere, alle ziemlich neutral, allenfalls einem etwas dunklen A ähnlich. Dasselbe schien mir auch der Fall zu sein, als die Vokale alle auf  $a^2$  gesungen wurden.

Dann wurde die obige Tonphrase um eine große Terz tiefer gelegt:  $f^2 a^2 g^2 e^2 f^2$ . Hier traten mir nach mehreren Versuchen gewisse Unterschiede in der Klangfarbe entgegen, die ich auf die Vokale deutete, und so erzielte ich zuletzt eine richtige Reihe. Aber in sich selbst war der spezifische Charakter eines jeden sehr unklar. Es war sozusagen nur ein relatives U, A usf.

12 Reihen, in denen je eine Sängerin auf dem konstanten Tone  $fis^2$  die 5 Vokale in einer beliebigen, mir unbekanntem wechselnden Folge zu singen hatte, wurden in den meisten Fällen richtig beurteilt, wenn auch nicht ohne Schwierigkeit. U und O, I und E standen sich sehr nahe. Einzeln gehört, wären sie auch in dieser Höhe sicher vielen Verwechslungen ausgesetzt; nur innerhalb der Reihe, von der ich wußte, daß sie diese 5 enthielt, konnte ich sie auseinanderhalten.

Mit denselben Sängerinnen wurde dann eine längere systematische Versuchsreihe angestellt, an der als Beobachter gleichzeitig 11—12 Herren und Damen aus dem Psychologischen Institut teilnahmen, die ihre Urteile aufzuschreiben hatten. Die Sängerinnen hatten, um Ermüdung zu vermeiden, abwechselnd je einen Vokal auf einem bestimmten Ton zu singen. Sie stellten sich wieder beliebige Reihen aus den 5 Vokalen zusammen, die nur die Bedingung erfüllten, daß jeder einmal darin vorkam. In jeder Gesamtreihe erschien also jeder 2mal. Die Beobachter wußten aber davon nichts. Ihnen war nur gesagt, daß es sich um Vokale handle. Auch daß die Umlaute ausgeschlossen waren, war ihnen unbekannt.

Als Tonhöhen wurden zuerst  $a^2$ , dann  $fis^2$ , dann  $c^2$  benutzt. Die Ergebnisse sind (unter Weglassung der wenigen Verwechslungen mit Umlauten) folgende:

	$c^2$					$fis^2$					$a^2$				
	U	O	A	E	I	U	O	A	E	I	U	O	A	E	I
U	16	6				3	8	11			2	8	12		
O		20	2			1	7	14			1	10	9		
A			3	19			2	19					22		
E				19	1			2	18				10	9	
I					10	10	1	1	12	2		1	5	10	4
$\frac{r}{n}$	$\frac{84}{110} = 76,4\%$					$\frac{49}{110} = 44,5\%$					$\frac{47}{110} = 42,7\%$				

Man sieht den starken Abfall von der Mitte der 2-gestrichenen Oktave an. Aber auch schon auf  $c^2$  sind nur  $\frac{3}{4}$  der Urteile richtig. Die Verwechslungen betreffen wieder überall besonders U und O, E und I, und zwar einseitig: U wird sehr häufig für O, I für E gehalten, aber fast niemals umgekehrt. Die Verwechslungen gehen eben nach der Mitte hin. Vielfach lautete das Urteil auch direkt: „U nach O verschoben, I nach E verschoben.“ Bei allen Vokalen kommen auf  $fis^2$  und  $a^2$  Verwechslungen mit A vor. Bei  $a^2$  fällt auf A kein einziges falsches Urteil; es ist offenbar am wenigsten verändert. Die Beobachter äußerten aber, daß es in den höheren Lagen mehr ein A-ähnlicher Laut als ein richtiges A sei. Man nannte es auch wohl ein „etwas trübes A“ oder „nach A hin“.

Man sieht aber auch, wenn man die Ergebnisse für  $fis^2$  mit den vorher erwähnten vergleicht, welchen Unterschied es schon macht, je nachdem der Hörer weiß, daß ausschließlich „Hauptvokale“ vorkommen und jeder 1 mal, oder nur weiß, daß Hauptvokale oder Umlaute in irgendwelcher Anzahl vorkommen.

Die vorwiegende Richtung der Verwechslungen ergibt sich auch aus den vielen Fällen, wo notiert wurde: „anfangs O, dann U“, „anfangs I, dann E“ oder: „E mit I-Einsatz“. In diesen Fällen rechnete ich nur das Schlußurteil. Der Einsatz selbst wurde übrigens in diesen Versuchsreihen vorwiegend falsch beurteilt, was im Widerspruch mit der gewöhnlichen Meinung steht, daß man den Vokal am Einsatz besonders deutlich erkenne. Ob dies an individuellen Eigenheiten der Sängerinnen lag, kann ich nicht entscheiden. Es wäre aber denkbar, daß der Einsatz bei so verwandten Vokalen wie U und O, I und E in kritischen Regionen des Tonreiches den ablenkenden Einflüssen leichter als sonst unterläge.

Endlich wurden Versuche gleicher Art mit der berühmten Opernsängerin Frau MELANIE KURT angestellt (21. V. 1914), der für ihre freundliche Bereitwilligkeit auch hier bestens gedankt sei. Sie war gebeten, die Vokale besonders deutlich und mit nicht zu starker Tongebung zu singen. Sie ist gewohnt, nur schwachen Glottisschlag anzuwenden, was auch erwünscht war. Unter den 9 Beobachtern, die, wie immer, mit dem Rücken gegen die Sängerin saßen, befand sich auch ihr Gatte, der trotz seiner besonderen Vertrautheit mit ihrer Stimme und Aussprache doch nicht weniger Verwechslungen als die übrigen unterlag. Den Beobachtern außer mir war wieder nur bekannt, daß es sich um Vokale handle. Die Sängerin entwarf jede Reihe so, daß jeder der 5 Vokale 2mal in jeder Reihe vorkam, außerdem die Umlaute Ä und Ö je 1mal. Letztere Maßregel war eingeführt, um jede Art von Schlußfolgerung unmöglich zu machen, da man sonst etwa aus der 10-Zahl das doppelte Vorkommen jedes Hauptvokals hätte erraten können. Ich selbst wußte von dieser Zusammensetzung, aber natürlich nicht von der Verteilung im einzelnen Falle.

Wie schon früher bemerkt, sind Kunstsänger für Vokalversuche nicht in jedem Falle die beste Tonquelle, da es ihnen vor allem auf Klangschönheit ankommt. Auch Frau KURT scheint manchmal weniger genau vokalisiert zu haben, als es ihr an sich noch möglich gewesen wäre. Im allgemeinen schien eine gewisse Neigung zu dunkler Vokalgebung zu walten. Ä wurde einige Male stark nach E hin verändert, vielleicht wieder unter dem Einfluß ästhetischer Motive.

In solchen Fällen zeigte es sich nun, daß Massenversuche neben dem Vorteil, schnell große Zahlen zu liefern, den Nachteil haben, daß zufällige Abweichungen eben auch sogleich das Urteil aller Hörer bestimmen. So sind einige Anomalien in die folgenden Tabellen gekommen. Sollten diese Versuche fortgesetzt werden, so würde es sich darum empfehlen, eine größere Anzahl von Kunst- und Natursängerinnen abwechselnd zu verwenden, wogegen 3—4 gute Beobachter genügen würden. Dann würden sich die Zufälligkeiten besser ausgleichen und zugleich individuelle Unterschiede der Sängerinnen herauspringen. Doch zeigen unsere Versuche zusammen mit den HERMANNschen schon deutlich, daß und in welcher Richtung überhaupt Veränderungen der Vokale in dieser Lage stattfinden, und genügen damit für unsere Zwecke.

Die Versuche wurden angestellt auf den Tönen  $c^2$ ,  $e^2$ ,  $gis^2$ ,  $gis^1$  in dieser Folge der Reihen. Dann noch eine Reihe auf  $e^2$ , bei der jeder Vokal mit L, und eine, bei der jeder mit T eingeleitet wurde. Dieser Umstand war den Beobachtern bekannt, sie hätten ihn ja auch sofort bemerkt.

Wir geben nun die Diagramme in der Reihenfolge von  $gis^1$  aufwärts. Die Urteile über Ä und Ö, die nur aus den erwähnten methodischen Gründen eingefügt waren, sind weggelassen; doch

sei erwähnt, daß sie ohne besondere Regelmäßigkeit vielfach miteinander und mit E verwechselt wurden<sup>1)</sup>.

	<i>gis</i> <sup>1</sup>					<i>e</i> <sup>2</sup>					<i>e</i> <sup>2</sup>					
	U	O	A	E	I	U	O	A	E	I	U	O	A	E	I	
U	18					13	3				14	4				
O	4	14				3	15				9	9				
A			18				8	10				5	13			
E				8	10				5	11				11	7	
I					18				1	16					4	13
$\frac{r}{n}$	$\frac{76}{90} = 84,4\%$					$\frac{59}{90} = 65,5\%$					$\frac{60}{90} = 66,7\%$					

	<i>gis</i> <sup>2</sup>					<i>e</i> <sup>2</sup> , mit L					<i>e</i> <sup>2</sup> , mit T				
	U	O	A	E	I	U	O	A	E	I	U	O	A	E	I
U	9	6	3			11	7				14	4			
O	5	9	4			3	15				1	17			
A	2	4	12				5	13					18		
E	1			16	1				14	4				15	
I				9	9				3	15				9	9
$\frac{r}{n}$	$\frac{55}{90} = 61,1\%$					$\frac{68}{90} = 75,5\%$					$\frac{73}{90} = 81,1\%$				

Hiernach sind sogar schon auf *gis*<sup>1</sup> Verwechslungen nicht ganz ausgeschlossen, wenn die Vokale so isoliert gegeben werden. Allerdings sind hier die Zahlen für E eine Anomalie, die durch eine zufällige Verschiebung der Vokalisierung nach I hin in einem einzelnen Falle zustande gekommen sein muß. Ich zweifle daher nicht, daß sich Versuchsreihen auf *gis*<sup>1</sup> erzielen lassen, die kein einziges falsches Urteil enthalten. Aber Tatsache bleibt doch, daß selbst

<sup>1)</sup> Urteile wie „zwischen U und O“ u. dgl. habe ich stets in günstigem Sinne gerechnet, also sowohl wenn U als wenn O angegeben war, als richtige. Bei der Schwierigkeit genauer objektiver Abgrenzung und den zufälligen Schwankungen der Lautgebung muß man wohl eine solche Milde walten lassen. Sonst würden noch viel kleinere Zahlen für richtige Urteile herauskommen.

eine so hervorragende Sängerin bei ausdrücklichem Streben nach genauer Charakteristik der Vokale des Erfolges nicht so sicher ist, wie in bezug auf die genaue Intonation der Tonhöhe, obschon natürlich auch in dieser Hinsicht den besten Sängern gelegentlich ein kleines Malheur zustoßen kann.

Bei  $c^2$  findet sich unter E dieselbe Anomalie und sicher aus demselben Grunde. Vielleicht erstrebte die Sängerin möglichst helle Vokalisierung, um dem Ä desto sicherer aus dem Wege zu gehen. Möglicherweise wurden aber auch die Urteilenden durch die ungewöhnlich silberne, d. h. an hohen Obertönen reiche Stimme (s. o. S. 28) irreführt, so daß ihnen E wie I vorkam, während das E der Schwestern PLANCK in dieser Lage noch fast ausnahmslos richtig erkannt wurde. Ohne diesen zufälligen Fehler wäre die absolute Zahl der  $r$  bei  $c^2$  sicher noch über 70% gewesen, aber auch sicher kleiner als bei  $gis^1$ . Bei  $e^2$  zeigt die Tabelle die gleichen, bei  $gis^2$  noch stärkeren Abfall. Hier fällt auch eine bedeutend größere Streuung der Urteile auf. Vorher erstrecken sich die Verwechslungen nur auf die unmittelbar benachbarten (der Helligkeit nach ähnlichsten) Glieder, jetzt auch auf entferntere.

Eine Bevorzugung des A ist hier weder in seinen  $r$  noch bei den Verwechslungen zu erkennen. Die Verähnlichung aller Vokale mit A nach der Höhe hin scheint also nicht bei allen Sängerinnen einzutreten. Dagegen ist bei Frau KURT auffallend die Zunahme der  $r$  für E. Dieses produziert sie auf  $gis^2$  weit sicherer und kenntlicher als auf  $c^2$  und  $gis^1$ .

Bemerkenswert ist die Steigerung der richtigen Urteile bei vorgesetztem Konsonanten, besonders bei dem Explosivlaut T. Man sieht hier schon den Einfluß der sprachlichen Verknüpfung, auch selbst bei sinnlosen Silben. Bei Doppelsilben wie Tota, Tite wären wahrscheinlich die  $r$  noch weiter gestiegen; der Kontrast hätte sowohl die Sängerin zur schärfsten Charakterisierung als die Hörer zur schärfsten Beachtung der vokalen Unterschiede genötigt. Bei isolierten Silben dürfte ein Grund der verbesserten Beurteilung darin liegen, daß der vorgesetzte Konsonant bekanntlich vielfach in Vorausnahme des Vokals modifiziert, diesem angeglichen wird, so daß der Hörende einen Anhaltspunkt mehr für die Erkennung gewinnt. Die Hauptursache der Verbesserung liegt aber wohl darin, daß der Sänger (Sprecher) gewissermaßen einen Anlauf nimmt, durch den ihm die genaue Charakterisierung des Vokals besser gelingt, ebenso wie der Springer mit Anlauf besser über den Graben oder das Seil kommt. Besonders beim T scheint diese Analogie überredend. Doch handelt es sich nicht so sehr (wie beim Springer) um die physische Vorbereitung, die Ansammlung

eines hohen Energiebetrages, als um die geistige, die Einstellung auf die besondere Vokalqualität, die mit dieser Hilfe sicherer erfolgen mag.

Wie mir Herr Prof. RUPP nach seinen Studien über die Technik des Telephonierens mitteilt, wird ein in ein Wort eingefügter Konsonant, wenn er auch gar nicht dahin gehört, leichter verstanden und weniger verwechselt als im isolierten Zustand. Wenn man z. B. in der für „31“ gebräuchlichen Modifikation „einsunddreißig“ (mit eingefügtem *s*) statt des *s* einen beliebigen anderen Konsonanten setzt, *f*, *t*, *l*, so werden auch diese, hier ganz sinnlosen, leichter erkannt als im isolierten Zustand. Man wird wohl annehmen müssen, daß die Hervorbringung der Sprachlaute in Verbindung miteinander allgemein vollkommener, charakteristischer erfolge. Daneben könnte auch erwogen werden, ob nicht durch den Zusammenhang an sich, selbst den sinnlosen, die Aufmerksamkeit des Hörenden stärker angespannt werde, und ob nicht jedes Element eines irgendwie gestalteten Komplexes als solches schon gegenüber dem isolierten Zustande für das auffassende Bewußtsein im Vorteil wäre. Diese allgemeineren Fragen mögen hier auf sich beruhen.

Einem Jahresberichte v. HORNBOSTELS entnehme ich, daß auch STEWART 1923 eine hierhergehörige Beobachtung gemacht hat: I, welches durch Ausschaltung aller Teiltöne über 2400 Schw. unkenntlich geworden war, blieb gleichwohl im Zusammenhang des Wortes „meal“ kenntlich. v. HORNBOSTEL weist darauf hin, daß schon physikalisch bei dem stetigen Übergang von einem Laut zum anderen kontinuierliche Umformungen der Wellen entstehen müssen, die zum Erkennen beitragen mögen, bisher aber nicht näher erforscht sind. Er fügt aber schließlich gleichfalls hinzu: „Wahrscheinlich eignet der Gesamtstruktur (eines Wortes) eine Schärfe der Charakteristik, die den einzelnen Lauten schon physikalisch immer fehlt.“

#### 4. Kapitel.

### Das stimmhafte Sprechen und dessen Abbau durch das Interferenzverfahren.

#### 1. Sprache und Sprachmaterial.

Die menschliche Sprache ist als Äußerung lebendiger Persönlichkeiten gewissermaßen selbst etwas Lebendiges und Beseeltes. Das Gesprochene bedeutet etwas, wenn es auch nicht immer der Rede wert ist. Daher ist ein Wort, ja ein grammatisch abgeschlossener Satz in seinem Tonfall und Akzent bis in die Einzelheiten durch den ganzen Zusammenhang und durch die Situation bestimmt. Der nämliche lexikalische Wortbestand in der nämlichen syntaktischen Ordnung kann dadurch eine akustisch verschiedene Erscheinungsform gewinnen. Außer den durch den Sinn der Rede bedingten akustischen Modifikationen gibt es aber auch solche, die in der Mechanik und Technik des zusammenhängenden Sprechens selbst wurzeln. Dahin gehören besonders die Wechselwirkungen, die benachbarte Teile, sowohl Konsonanten als Vokale, aufeinander ausüben. Der „Berührung benachbarter Laute“ widmete nach dem Vorgange WINTELERS (1876) SIEVERS ein langes Kapitel seiner Phonetik, und sie wurde auch von späteren Phonetikern gebührend berücksichtigt. Hier wirkt nicht nur das Vorangehende auf das Folgende, sondern auch umgekehrt, was natürlich nur dadurch möglich ist, daß das Folgende bereits vor seinem Eintritt psychisch und physiologisch vorbereitet wird. Spricht man Bu, Bl, so werden die Sprachorgane beim B schon mit Rücksicht auf den folgenden Laut verschieden eingestellt und das B darum verschieden gesprochen. Selbst die sich objektiv nicht berührenden Vokale verschiedener aufeinanderfolgender Silben können so aufeinanderwirken („Vokalharmonie“ östlicher Sprachen).

Aber diese inneren Bezüge und Bänder der menschlichen Rede, die gegenseitigen Beeinflussungen ihrer Teile und deren Abhängigkeit vom Ganzen scheiden, so hohes Interesse sie auch dem Psychologen bieten, für unsere Untersuchung ihrer Anlage nach aus. Diese geht nur auf die isolierten Bausteine in ihrer Isolierung.

Doch selbst in bezug auf dieses Material legen wir uns bestimmte Schranken auf: die Diphthongen und alle analogen Gebilde bleiben ausgeschlossen. Als echte Diphthongen bezeichnen wir die mit einem Laut anfangenden, mit einem anderen endigenden, dazwischen aber nicht scharf absetzenden Laute gleicher Gattung (vokalische oder konsonantische), mögen sie mit mehreren Zeichen oder nur mit einem (z. B. Z oder das lange O des Englischen, wie in note) geschrieben werden. Sie sind insofern zusammengesetzt, als mehrere verschiedene Laute in zeitlicher Folge darin erkennbar sind, während z. B. Ng, obgleich mit mehreren Zeichen geschrieben, kein Diphthong ist, da in keinem Augenblick ein N oder G darin erkennbar ist und die Lautbeschaffenheit während der ganzen Dauer der Aussprache unverändert bleibt. Es kommt zwar vor, daß beide Buchstaben getrennt ausgesprochen werden, wie in den Fremdwörtern Angora, Kongo<sup>1)</sup>, auch gelegentlich auf der Bühne oder bei Dichtern, die „Hang“ auf „sank“ reimen: aber hier wird man die Aussprache wie den Reim nicht loben können.

Ein zusammengesetzter Laut ist aber auch wieder nicht die bloße Summe oder Aufeinanderfolge der beiden Teillaute. Pf, Ps, Bl sind ebenso wie Ai, Eu (= AÖ-Ü) eigentümliche Einheiten, für die auch der Übergang zwischen den beiden Teilen wesentlich und charakteristisch ist. Diese die Diphthongen charakterisierenden flüchtigen Übergangserscheinungen bleiben hier ebenso wie alle sonstigen Übergänge zwischen den letzten Bauelementen, die während des Sprechens zum Vorschein kommen, unberücksichtigt. Ihnen ist mit den hier angewandten Mitteln nicht beizukommen. Graphische Methoden mögen darin weiterführen und lassen schon manches erkennen<sup>2)</sup>; aber eine Theorie ist auch darauf noch nicht zu bauen.

Von sehr zusammengesetzten Erscheinungen geht jede Forschung aus, da die unmittelbare Erfahrung uns eben nicht sofort die letzten Elemente, sondern zunächst die verwickeltesten Komplexe darbietet. Wir verkennen auch nicht, daß die Forschung ihres Ausgangspunktes eingedenk bleiben und auf sein Verständnis zuletzt hinarbeiten muß. Aber wie sehr man dies betonen mag: zur exakten Beschreibung eines Ganzen muß doch immer auf die

<sup>1)</sup> Genauer gesagt folgt hier (und zwar nach fachmännischer Belehrung auch in den Originalsprachen) auf ein leicht nasalisiertes ng noch ein besonders artikuliertes g. Es wäre fehlerhaft, n dabei wie in „ankommen“ zu sprechen.

<sup>2)</sup> Beispielsweise vgl. über Feinheiten des Einsatzes bei Vokalen und Konsonanten ISSERLIN II.



Teile, ja auf die Teile der Teile zurückgegangen werden. Indem wir also unsere Aufgabe in bezug auf die Sprache genau begrenzen, lehnen wir zugleich den Einwand ab, als hätte diese Art von Zergliederung mit dem Wesen der Sprache überhaupt nichts zu schaffen. Wie weit sie Sprachphilosophen nützlich sein kann, bleibt abzuwarten. Daß sie aber möglich und nach anderen Richtungen hin nützlich ist, soll das Folgende zeigen.

Da wir bisher nur gesungene Vokale behandelten, soll zunächst das Verhältnis des Sprechens zum Singen charakterisiert werden.

## 2. Singen und Sprechen<sup>1)</sup>.

Dieser vielverhandelte Unterschied liegt zum Teil auf psychologischem Gebiete, nämlich in den Einstellungen, den vorschwebenden Vorstellungen, den leitenden Intentionen, den zugrunde liegenden Gefühlen und ähnlichen Imponderabilien. Soweit nur das dabei zur Verwendung kommende Lautmaterial in Frage kommt, liegt er hauptsächlich darin, daß beim Singen, wie in der Musik überhaupt, im Prinzip nur feste Tonhöhen und bestimmte transponierbare Intervalle gebraucht werden, beim Sprechen dagegen auch schwankende Tonhöhen und Tonabstände.

Dies ist natürlich nicht so zu verstehen, als ob das Singen eines Vokals unter stetiger Tonveränderung schon ein Sprechen wäre: zu diesem gehört von seiten des Materials auch ein viel stärkeres Hervortreten der Konsonanten und ein bunter, durch keine „Tonleitern“ geregelter Wechsel der Tonhöhen.

Es ist auch kein Widerspruch gegen diese Abgrenzung, begreift sich vielmehr aus ihr, daß Übergänge und Mischungen vorkommen (Sprachgesang, singendes Sprechen). Auch in der gewöhnlichen Sprache fehlt es keineswegs an Tonhöhen, die für die kurze Dauer einer Silbe festgehalten werden, sowie an mehr oder minder genauen musikalischen Intervallen zwischen aufeinanderfolgenden Silben. Es ist nur die Aufmerksamkeit des Hörenden wie des Sprechenden in der Regel nicht den Tönen und Tonverbindungen als solchen zugewandt, sondern dem Sinne der Rede. Gerade die stetige Tonbewegung und die unregelmäßig wechselnden und ungenauen Intervalle begünstigen diese Richtung der Aufmerksamkeit, ebenso wie sie andererseits durch sie herbeigeführt werden; sie verhalten sich dazu sowohl als Ursache wie als Wirkung. Überdies werden viele Silben so tief, leise und rasch gesprochen, daß ihre Tonhöhe auch bei scharfer Aufmerksamkeit nicht deutlich wird.

<sup>1)</sup> Ausführliches hierüber s. in m. Abh. 13, auch in 2, S. 264ff. und in „Anfänge der Musik“ S. 14ff. 1911.

Beim Singen darf kein Ton, auch der leiseste nicht, so „unter den Tisch fallen“, daß seine Tonhöhe undeutlich wird. Aber wer es darauf anlegt, kann auch bei Gesprochenem, wenigstens an akzentuierten Stellen, Tonhöhen und Intervalle jederzeit feststellen (Sprachmelodie). Ein prinzipieller und durchgreifender Unterschied in den akustischen Grundeigenschaften als solchen existiert nicht.

Daß auch in zeitlicher und rhythmischer Hinsicht das gewöhnliche Prosasprechen sich weniger an feste Formen bindet als das Singen, verschärft den Unterschied in den extremen Formen. Doch gibt es auch hierin Annäherungen von beiden Seiten her.

Selbst in der Klangfarbe liegen gewisse Unterschiede. HELMHOLTZ vermutet (S. 183), und wohl mit Recht, daß beim Sprechen durch stärkeren Druck der Stimmbänder gegeneinander eine knarrende Klangfarbe, d. h. stärkere (und höhere, rasch miteinander schwebende) Obertöne als beim Singen gegeben werden. Die Folge hiervon muß sein, daß auch die Formanten der Vokale kräftiger als beim Singen hervortreten, zugleich aber auch, daß die Vokale außer U alle durchschnittlich etwas heller werden. PIPPING (4, S. 173) hat dies durch Kurvenaufnahmen bei finnischen Vokalen bestätigt. Zum Teil beruht der Unterschied der Klangfarbe aber auch auf beigemischten Geräuschen, mögen sie auf dem von HELMHOLTZ angedeuteten Wege oder sonstwie zustande kommen. Bei vielen Sprechstimmen erreichen diese Geräusche eine außerordentliche Stärke, die das Singen ganz unmöglich machen würde, und doch bleiben sie noch verständlich.

Neuerdings hat ISSERLIN auf objektivem Wege, durch Auszählung der Zacken von Vokalkurven, gezeigt, daß in der lebendigen Sprache mit der Tonhöhe auch die Klangfarbe der Vokale stetigen Veränderungen unterliegt, indem der Formant ein und desselben Lautes während des Aussprechens beim Fragesatz steigt, bei der Affirmation sinkt. Natürlich ändert sich hierbei mit dem Formanten auch der Vokal selbst; es ist eben, genau gesprochen, nicht mehr dasselbe A oder O, sondern ein sich stetig erhellendes oder verdunkelndes. Dagegen fanden sich solche gleitenden Veränderungen der Klangfarbe wieder bei gesungenen Vokalen nicht. Also auch in dieser Beziehung der Unterschied der schwankenden, stetig bewegten, von der festen Tongebung. Immerhin wird man die Festigkeit beim Gesang auch nur *cum grano salis* verstehen dürfen.

Wegen des entscheidenden Gewichtes der festen Tonhöhen und Intervalle spielen beim Singen die Vokale eine weit größere Rolle als die Konsonanten. Im Gesange werden Vokale oft lange

gleichmäßig ausgehalten: Dehnungen, zu welchen auch das affektivste Sprechen keinen Anlaß bietet. Einen bemerkenswerten Unterschied ergibt dabei die Behandlung der vokalischen Diphthongen: es ist bei solcher Dehnung unmöglich, ihre beiden Hauptbestandteile gleichmäßig zum Rechte kommen zu lassen, und auch ihr Übergang ineinander kommt zu kurz. Wenn etwa der Müllerbursche bei Schubert singt: „Die geliebte Müllerin ist mein!“ und das „mein“ bei mäßiger Geschwindigkeit  $1\frac{1}{2}$  Takte lang aushält, so kann er dabei nur den Vokal A singen und wird erst ganz am Schlusse mit einem äußerst verkürzten I zum N übergehen. Und so wird überhaupt das für das Sprechen so wesentliche Überschleifen der Sprachlaute ineinander beim Singen zugunsten der einfachen Vokale reduziert. Aber da wir die Diphthongen überhaupt außer Betracht lassen, brauchen uns auch diese Modifikationen hier nicht zu kümmern.

HERMANN formuliert das Verhältnis zwischen Singen und Sprechen einmal kurz so (Bd. 58, S. 258): „Der Unterschied zwischen gesungener und gesprochener Artikulation liegt ganz ausschließlich darin, daß Notenhöhe, Intensität und Dauer der Silben, genauer der Vokale, beim Gesang durch Melodie und Takt, beim Sprechen durch die Betonungsgesetze des Sinnes und der Konstruktion beherrscht werden.“ Bis auf das „ganz ausschließlich“ ist diese Grenzziehung richtig und geht sogar in gewissem Sinne noch tiefer auf das Wesen der Sache. Aber die verschiedenen Gesetzlichkeiten bringen eben den von uns und schon von den alten Griechen<sup>1)</sup> hervorgehobenen sinnenfälligen Unterschied in der Haltung und Veränderungsweise der Klänge mit sich. Voll zustimmen können wir HERMANN, wenn er fortfährt: „Bei einem einzelnen Vokal kann also absolut kein Unterschied zwischen Gesungenem und Gesprochenem gesucht werden. Ein lautes A hat, ob gesungen oder gesprochen, stets irgendeine Notenhöhe, nur mit dem Unterschiede, daß dieselbe im ersteren Falle bestimmt beabsichtigt, im zweiten dem Zufall oder der Gewohnheit überlassen wird. Jeder kann sich leicht überzeugen, daß er, wie er auch den Vokal A ausgesprochen habe, sofort mit demselben in einer Tonleiter fortfahren kann; d. h. das gesprochene A erweist sich zugleich als gesungenes.“

<sup>1)</sup> So von ARISTOXENUS, PTOLEMAEUS, BACCHIUS, GAUDENTIUS. Vgl. BELLERMANN: Die Größe der musikalischen Intervalle, 1873, S. 3ff., und des Verfassers „Geschichte des Konsonanzbegriffes“ (Abh. d. bayrischen Akad. d. Wiss. 1897).

Unter den neueren Phonetikern hat auch AUERBACH (4, S. 157) die Unterscheidung in gleichem Sinne besprochen.

Abweichende Anschauungen haben W. KOEHLER (1, III, S. 103ff.) und E. R. JAENSCH (1, S. 255ff.) ausgesprochen. Danach kämen gesprochenen Vokalen keine Tonhöhen im Sinne der „musikalischen Qualitäten“ zu, nach JAENSCH sogar auch keine oder nur undeutliche Tonhöhen im gewöhnlichen Sinne<sup>1</sup>). Hierüber vgl. m. Abh. 13, S. 17ff. Hier sei nur hervorgehoben, daß KOEHLERS These mit seiner negativen Stellungnahme zu den unbemerkten Teilinhalten unserer Sinneserscheinungen zusammenhängt. Er will den Empfindungen nur solche Eigenschaften zusprechen, die wir tatsächlich im Augenblick bemerken. Da nun die musikalischen Qualitäten im gewöhnlichen Gebrauche des sprachlichen Verkehrs unbemerkt bleiben, so sind sie nach ihm nicht vorhanden. Dasselbe muß dann freilich auch von den Tonhöhen im gewöhnlichen Sinne (Helligkeiten) und allen anderen akustischen Eigenschaften gelten. Wir kommen auf diesen Standpunkt, den ich nicht teilen kann, weiter unten noch einmal zu sprechen.

Von den Konsonanten ist hier noch nicht die Rede. Aber es liegt auf der Hand, daß auch sie beim Singen und Sprechen identisch sind, jedenfalls nicht verschiedener als auch schon innerhalb des Sprechens, wo ja ein und derselbe Konsonant in verschiedenem Zusammenhang gewissen Modifikationen unterliegt.

Im folgenden werden wir nun zeigen, daß die If.-Methode auf Gesprochenes in derselben Weise wie auf Gesungenes anwendbar ist und auch zu den nämlichen Ergebnissen führt. Wir wollen nicht sagen, daß dies für sich allein bereits ein Beweis dafür wäre, daß musikalische Qualität und Tonhöhe gesprochenen Lauten zukommen. Aber soviel ist sicher, daß sich die zu erwähnenden Tatsachen einer Anschauung leichter einfügen, die in jeder Beziehung gleiche akustische Grundeigenschaften beiderseits voraussetzt.

### 3. Das Interferenzverfahren bei der gesprochenen Rede.

Spricht man einen einzelnen Vokal oder eine Silbe oder auch ein mehrsilbiges Wort mit gleichen Vokalen in allen Silben, wie „Handschlag, Abraham, Abradacabra, Mahabharata“, und hält

<sup>1</sup>) Wir verstehen hier unter „musikalischer Qualität“ die Eigenschaft der Töne, welche bei Verdoppelung der Schwingungszahl identisch wiederkehrt und zu der gleichen Buchstabenbezeichnung Anlaß gegeben hat. KOEHLER nennt sie a. a. O. Tonhöhe, v. HORNBOSTEL neuerdings (2) Tonigkeit. Aber den ersten Ausdruck gebrauchen wir für die den Schwingungszahlen einfach parallel veränderliche Eigenschaft (= Helligkeit), mit dem zweiten dürfte passender das bezeichnet werden, was Töne von Geräuschen unterscheidet (s. Kap. 5). Leider ist über die zweckmäßigste Bezeichnungsweise in der Phänomenologie des Tongebietes noch keine Einigung erzielt. Das hängt aber auch teilweise an sachlichen Differenzen der Auffassungen.

der Sprechende dabei zunächst eine bestimmte Tonhöhe fest, so wird durch die Einstellung der Röhren auf den Formanten des auf dieser Tonhöhe gesungenen A auch das gesprochene A in ein dunkles O verwandelt. Sprechen und Singen ist ja dann objektiv so gut wie identisch. Werden solche Versuche unwissentlich angestellt, so sind die einzelnen Vokale am Röhrende durchaus unkenntlich. Wörter können allerdings auch dann oft aus den Konsonanten und dem Rhythmus der Aussprache erraten werden.

Schwankt sodann die Tonhöhe in der Weise des gewöhnlichen Sprechens, so braucht man nur die Formantgrenzen ein wenig weiter zu nehmen, indem man die Höhen der Grundtöne, zwischen denen sich der Sprechende bewegt, berücksichtigt. Nun dürfte sich die europäische Männerstimme bei ruhigem Sprechen durchschnittlich (besonders dunkle oder helle Stimmen ausgenommen) in der Oktave  $G-g$  bewegen, die Frauenstimme eine Oktave höher<sup>1</sup>). Wenn man daher den Gesamtformanten des A für die Grundtöne zwischen  $G$  und  $g^1$ , also  $d^2-g^3$ , ausschaltet, muß der Vokal

<sup>1</sup>) Nach HELMHOLTZ (S. 183) gebrauchen Männer in der Regel die obere Hälfte der großen, Frauen die obere Hälfte der kleinen Oktave. Er statuiert also beide Male nur die tiefere Abteilung der obigen Zonen. Es scheint mir aber, daß an etwas akzentuierten Stellen gerade die höheren Abteilungen, die Töne dicht über  $c$  bzw.  $c^1$ , benutzt werden. Nach den Untersuchungen von PAULSEN und GUTZMANN liegt der Stimmton bei Männern durchschnittlich zwischen  $A$  und  $e$ , bei Frauen und Kindern zwischen  $a$  und  $e^1$  und ist der Hauptgesprächston, gewissermaßen die Tonika der Sprachmelodie,  $c$  bzw.  $c^1$ . Von da weicht sie nach beiden Richtungen um etwa eine Terz ab (im Affekt natürlich weit mehr). Diese Bestimmungen dürften den Tatbestand richtiger als die HELMHOLTZschen wiedergeben. Sie gelten aber auch nur innerhalb unseres Kulturkreises. Bei exotischen Völkern finden sich starke Verlagerungen.

In Zusammenhang mit der mittleren Höhe der affektlosen Sprechstimme steht der von MERKEL so bezeichnete, von BARTH, KOEHLER u. a. näher untersuchte „phonische Nullpunkt“, bei dem weder Anspannung noch Entspannung der Stimmlippen stattfindet. Nach dem unmittelbaren Gefühl wurde er von KOEHLER zwischen  $H$  und  $dis$ , bei weiblichen Personen zwischen  $h$  und  $e^1$  gefunden, am häufigsten in der Mitte dieser engen Zonen. Für mich liegt er zwischen  $A$  und  $c$ , also nur wenig unter der Tonika des affektlosen Sprechens, was sich auch psychologisch leicht begreift. Nach R. SCHILLING, der 48 Personen untersuchte (Med. Wochenschr. 1924, 1, S. 631), geht er bei den meisten mit der Helligkeit des gesprochenen Vokals mehr oder weniger in die Höhe, variiert aber auch etwas unter dem Einfluß des (akustischen, motorischen) Vorstellungstypus. M. GIESSWEIN (2) fand auch die Resonanz des „Bronchialbaumes“ (des reich verästelten Röhrensystems der Lunge) sowie das Auftreten des „Pektoralfremitus“ (der Erschütterung der Brustwandungen beim Sprechen) in Übereinstimmung mit dem phonischen Nullpunkt.

wieder unkenntlich werden. Dann kann jemand, Frau, Mann oder Kind, ein überraschtes Ah! mit beliebig großen Stimmexkursionen sprechen, es wird immer als stöhnendes Oh! aus der Leitung kommen. Und erweitert man die ausgeschaltete Strecke noch bis  $b^4$ , so muß sich alles Sprechen, es mag noch so affektiv betont sein, in ein dumpfes Murmeln verwandeln, worin nur noch U und O erkennbar bleiben.

Man muß jedoch hier, um sicher zu sein, daß alle Obertöne, die von beliebigen Grundtönen herrühren können, innerhalb einer solchen Strecke vernichtet sind, eine große Anzahl fein abgestufter Einstellungen aneinanderreihen. Man braucht also mehr Röhren als bei einem auf einem konstanten Ton gesungenen oder gesprochenen Vokal.

Endlich lassen sich auch systematische Abbaureihen mit der Sprache überhaupt anstellen, indem man von oben herab immer größere Strecken des Tonreiches abträgt, auch dies in der eben erwähnten Weise durch zahlreiche nur wenig verschiedene Einstellungen. Es genügt aber, mit Einstellungen von 2 cm an zu beginnen und dann mit Differenzen von 0,3 cm, zuletzt von 0,5 cm immer längere Seitenleitungen hinzuzufügen. Dann sind bis zur völligen Unkenntlichmachung der Sprache etwa 44 Röhren erforderlich. Zwischen Schall- und Beobachtungszimmer muß dabei unbedingt ein Zimmer in der Mitte liegen, um jedes direkte Herüberdringen der Laute zu verhindern. Die Hauptleitung muß also durch 2 Wände geführt werden.

Bei den von mir angestellten Versuchen wurde als Text ein beschreibender Katalog des Berliner Museums für Völkerkunde mit vielen ethnologischen, geographischen und technischen Ausdrücken benutzt. Er verlangte also eine gewisse Spannung der Aufmerksamkeit, bot aber den Hörenden (Dr. v. ALLESCH, Frl. stud. EBERHARDT und mir selbst) an sich keine Schwierigkeiten des Verständnisses. Bei jeder der unten verzeichneten Einstellungen wurden außer dem allgemeinen Sprachverständnis auch die einzelnen Vokale und die empfindlichsten Konsonanten (S, F, Ch palatale) geprüft. Solche Versuchsreihen sind übrigens, wenn einmal die äußeren Einrichtungen zweckmäßig getroffen sind, sehr leicht und schnell zu machen; eine Reihe dauert nur eine Viertelstunde.

Die Ergebnisse, wie sie hier zusammengestellt sind, sind fast genau dieselben, wenn der Text auf der schwankenden Tonhöhe des natürlichen, ungezwungenen Sprechens und wenn er auf einer der Tonhöhen  $c$ ,  $gis$ ,  $c^1$  gesprochen wird. Sie gelten daher für jeden Grundton innerhalb der Sprechzone.

## Veränderungen der stimmhaften Sprache beim Abbau durch Interferenzröhren.

Einstellung in cm von 0,5 bis	Herabgesetzte obere Tongrenze	
3,35	$es^4$ = 2461 Schw.	Sprache noch ganz gut verständlich. Einzel: I und Ü nach U, E nach O hin alteriert.
4,8	$as^3$ = 1642 Schw.	Ebenfalls noch alles verständlich, doch etwas nebelhaft; schärferes Aufmerken erforder- lich. Einzel: I und Ü = U, E = O, Ö fast O, A fast AO.
6,7	$es^3$ = 1230 Schw.	Vieles unverständlich, doch öfters einige Worte nacheinander bei günstigem Zusammenhang verstanden. Einzel: Ö = O, Ä = AO.
9,6	$a^2$ = 870 Schw.	Nur selten noch ein Wort zu verstehen. Einzel: A stark verdunkelt.
13,4	$e^2$ = 652 Schw.	Alles unverständlich, kein Laut auch nur zu erraten. Dunkles U-artiges Lallen. Einzel alle Vokale wie U oder dunkles O.

Über die Veränderungen der Konsonanten ist weiter unten (5. Kap.) ausführlich zu berichten.

Bei Frauen und Kindern als Sprechern sind die Stadien und Stellen der Zerstörung dieselben wie bei Männern. Doch scheint eine Frauen- oder Kinderstimme bei gleicher Zerstörungsstufe (z. B. bei Einstellung bis 6,7 herab) etwas stärker alteriert. Dies würde sich leicht daraus begreifen, daß *c* weit mehr Teiltöne innerhalb einer gegebenen Tonstrecke enthält als *c*<sup>1</sup>, weshalb ja überhaupt bei tieferem Grundton die Vokale deutlicher sind. Dies muß sich auch bei Zerstörungen geltend machen.

Immer tritt zuerst eine leichte Verdunkelung ein, es legt sich gleichsam ein Nebel über die Sprache. Dann ist's, als hätte der Sprecher irgendeinen Sprachfehler. Dann werden einzelne Wörter unverständlich, viele nur aus dem Zusammenhang erschlossen; es ist ähnlich, wie wenn man eine fremde Sprache hörte, die man nicht genügend beherrscht; man muß sich bemühen, den Faden festzuhalten. Dann gelingt auch dies nicht mehr, es tauchen bei größter Aufmerksamkeit nur wenige verstandene Bruchstücke auf. Zuletzt ist es ein seltsames dunkles Lallen. Die entscheidendste Wendung fällt in die obere Abteilung der 2-gestrichenen Oktave, in die Gegend des A-Formanten. Mit seiner Zerstörung ist alles vorbei.

Außerordentlich auffallend ist bei solchen Versuchen in jedem Stadium der Zerstörung, wie lange doch noch die zusammenhängende sinnvolle Rede verständlich bleibt, nachdem sehr maß-

gebende Bestandteile für sich allein bereits stark alteriert, ja unkenntlich geworden sind. Man kann sich diesen, im allgemeinen ja bekannten<sup>1)</sup> Einfluß des Zusammenhanges auf die richtige Auffassung gar nicht schlagender als durch solche Versuche vergegenwärtigen und ist immer wieder davon überrascht. Bei der Einstellung 6,7 wurde selbst ein Wort wie „still“ noch erkannt, obgleich so gut wie alle Bestandteile einzeln unkenntlich waren. Ebenso noch bei Einstellung 9,6 „insbesondere“, wo höchstens o und r für sich erkennbar sein konnten. Die Gesamtstruktur, zumal die Akzentverteilung, mag hier beigetragen haben.

Schon hier sei darauf hingewiesen, daß diese Veränderungen auch ihre praktische Bedeutung haben, sowohl bei pathologischen Affektionen des Gehörorganes als auch im gewöhnlichen Leben bei zunehmender Entfernung der Schallquelle und anderen Schallleitungshindernissen, die die hohen Töne mehr als die tiefen schädigen.

#### 4. Die BEZOLDSche „Sprachsext“.

Der verdiente Ohrenarzt und Forscher v. BEZOLD glaubte aus der Vergleichung zahlreicher Fälle partieller Hördefekte, die er mit Hilfe seiner „kontinuierlichen Tonreihe“ untersuchte, schließen zu müssen, daß speziell die Sexte  $b^1-g^2$  von fundamentaler Bedeutung für das Sprachverständnis sei. Diese Lehre hat beträchtlichen Einfluß auf die Praxis erlangt (sie wurde in den bayrischen Taubstummenanstalten zum offiziellen Gebrauch eingeführt, indem danach entschieden wurde, welche Kinder mit Gehördefekten noch zur Erlernung der Lautsprache heranzuziehen seien). Man hat sie aber sogar als Grundlage für hirnpathologische Theorien verwendet<sup>2)</sup>. v. BEZOLD hatte zwar niemals einen Fall beobachtet, in dem nur diese Sexte und zugleich das Sprachverständnis erhalten war oder nur diese Sexte und zugleich das Sprachverständnis fehlte. Aber den zahlreichen Fällen von Hörresten, bei denen Sprachverständnis noch möglich war, war diese Strecke gemeinsam, und den zahlreichen Fällen partieller Taubheit, bei

<sup>1)</sup> Vgl. bes. H. GUTZMANN (3). Die Zahl der nichtverstandenen sinnlosen Silben erwies sich größer als die der Verwechslungen bei Schwerhörigen.

<sup>2)</sup> WERNICKE schloß daraus, daß die Töne  $b^1-g^2$  in der ersten linken Schläfenwindung lokalisiert seien, während er die übrigen Töne auf die sonstige Schläfenrinde verteilte; eine Anschauung, die allerdings schon von psychiatrischer und hirnpathologischer Seite (BONHOEFFER, HENSCHEN, FLECHSIG) zurückgewiesen wurde. Ich entnehme dies HENSCHENS Werk „Klinische und anatomische Beiträge zur Pathologie des Gehirns“ Teil VII, 1922, S. 301.



denen das Sprachverständnis nicht mehr möglich war, war das Fehlen dieser Strecke gemeinsam, wenn auch in den Fällen der ersten Art immer noch andere Reste und in denen der zweiten Art noch andere Lücken da waren. Allerdings fand er selbst später einen Fall, bei dem von der Sprachsext sogar nur eine Quart übrig war (außer anderen Resten) und doch die Sprache verstanden wurde, hätte daher eigentlich jetzt von einer Sprachquarte reden müssen. Aber die Schlußweise war überhaupt logisch nicht ganz einwandfrei. Zudem liegen, was v. BEZOLD natürlich selbst sehr wohl wußte, zwischen vollem Verständnis und vollem Unverständnis der Sprache zahllose Übergänge.

Wir können jetzt aus unseren Feststellungen über die Lage der Vokalformanten die Störungen, die bei Ausfall dieses Bezirkes entstehen müssen, voraussagen. Wäre nur diese Sext-Zone für einen Patienten unhörbar, so müßte O für sein Ohr nahezu in U übergehen, A in einen Mischlaut aus UO und Ö (vgl. o. S. 73). Ö, Ä und E müßten durch den Wegfall ihres Unterformanten gleichfalls alteriert sein. Aber das Sprachverständnis wäre nicht ganz vernichtet, zumal da die Konsonanten im wesentlichen erhalten blieben.

Es ist aber auch mit unseren Interferenzeinrichtungen die Probe in der Weise gemacht worden, daß man eben diesen Tonbezirk ausschaltete (FRANKFURTHNER und THIELE). Hierbei wurden freilich die ungeraden Multipla mitausgeschaltet, also eine große Zahl höherer Töne jenseits  $f^3$ , deren Wegfall zu der Alteration beitragen muß. Aber da tatsächlich nur eine geringe Erschwerung, keineswegs ein Verlust des Sprachverständnisses beobachtet wurde, so war a fortiori der Beweis gegen die BEZOLDsche Lehre erbracht<sup>1)</sup>.

Im einzelnen wurden von den beiden Autoren wesentlich nur die Konsonanten untersucht. Ich machte selbst eine Probe mit Vokalen, indem ich aus einem gesungenen  $c^1$  die in die Sprachsext fallenden Teiltöne  $c^2$  und  $g^2$  ausschaltete. Es fanden sich O, A, Ö, Ä, E tatsächlich in dem Sinn alteriert, wie es oben erschlossen wurde (wobei mitausgeschaltete Multipla nur bei den 3 letzten in Betracht kommen); U und Ü hingegen blieben, wie gleichfalls vorauszusehen, unversehrt. Aber die verbleibenden Reste sowohl der Vokale wie der Konsonanten ermöglichten eben doch noch ein leidliches Sprachverständnis bei sinnvoller Rede.

<sup>1)</sup> Vielleicht hätte das Sprachverständnis doch etwas mehr gelitten, wenn mehr als 15 Seitenröhren gebraucht worden wären (die Einstellungen standen um je eine Halbtonstufe voneinander ab). So mögen immerhin die Töne der hohen Region nicht vollständig mit ausgelöscht worden sein.

Exakt läßt sich die Prüfung mit Hilfe des im 7. Kap. zu beschreibenden synthetischen Verfahrens vornehmen. Man braucht nur einen Vokal aus seinen Teiltönen herzustellen und dann die in die „Sprachsext“ fallenden wieder durch Verstopfung ihrer Leitungen zu beseitigen. Es zeigte sich auch so, daß O und A besonders leiden, aber auch Ä und E; die übrigen nur wenig oder gar nicht.

Zugeben muß man v. BEZOLD, daß die 2-gestrichene Oktave von besonderer Bedeutung ist. Sie ist es, weil sie den A-Formanten enthält, und weil A, der wichtigste aller Vokale (princeps vocalium nennt es schon HELLWAG), gewissermaßen als Vokalzentrum gelten kann. Aber wir würden aus diesem Grunde, wenn schon von einer Sprachsext die Rede sein soll, die höherliegende Sext  $e^2$ — $c^3$  dafür in Anspruch nehmen, nur nicht im BEZOLDSchen, sondern in einem viel eingeschränkteren Sinne. Besonders wäre nicht zu vergessen, daß auch die Gegend um  $c^4$  eine hervorragende Bedeutung besitzt, da sie die Formanten der 3 hellsten Vokale enthält. Für das volle Sprachverständnis aber, bei dem kein einziger Bestandteil der gesprochenen Rede in seinem spezifischen Charakter geschädigt sein soll, ist die Unversehrtheit der gesamten Tonregion von  $c^1$  bis  $c^5$  erforderlich. Näheres darüber im 6. und 9. Kapitel.

## 5. Kapitel.

### Struktur der Flüstervokale und Konsonanten.

Zwei Fragen tauchen jetzt für unsere Untersuchung auf: 1. Wenn die stimmhaften Vokale durch Teiltöne von fester Höhe charakterisiert sind, was charakterisiert geflüsterte Vokale, und was ist ihnen mit den gesungenen gemeinsam? 2. Was unterscheidet stimmlose Konsonanten von den stimmlosen Vokalen, und was ist ihnen mit den stimmhaften Konsonanten gemeinsam, von denen manche doch wieder den stimmhaften Vokalen so nahe stehen wie das M dem U?

Hier liegen überall Paradoxien, die zu etwas tieferem Eindringen in die Phänomenologie der Gehörsempfindungen zwingen.

#### I. Klänge und Geräusche, Vokale und Konsonanten, stimmhafte und stimmlose Sprachlaute.

1. Klänge und Geräusche sind spezifisch verschiedene Gehörerscheinungen, die darum rein erscheinungsmäßig nicht definiert, sondern nur durch Beispiele erläutert werden können. Unsere Sinnesempfindungen zerfallen zunächst in Gattungen, wie Licht- und Schallempfindungen, innerhalb dieser Gattungen aber wieder in Arten. So haben wir unter den Gesichtsempfindungen die farbtonlose Schwarz-Weiß-Reihe und die getönten oder im engeren Sinne farbigen Empfindungen, wie Rot, Blau. Diesem Unterschied entspricht im Gehörssinne nach vielen (nicht allen) Seiten hin der von Geräuschen und Klängen. Wie man für den Unterschied der Grauererscheinungen von den farbigen in erster Linie das Auge als Zeugen hat, so für den von Geräuschen und Klängen das Ohr. Aus Anschauungen muß man sich diese Elementarbegriffe bilden und auf Grund der Anschauung auch wieder die Einzelercheinung dem Begriffe unterordnen. Die Erfahrung zeigt allerdings auch in den Gesetzlichkeiten der so unterschiedenen Teilgebiete bemerkenswerte Verschiedenheiten, die den wissenschaftlichen Begriff vervollständigen. Aber Grundlage der Unterscheidung bleibt auch in wissenschaftlicher Hinsicht immer das einfache Gehör.

Auch die Frage nach den physikalischen und physiologischen Ursachen dieser Unterschiede ist scharf von der bloßen Beschreibung der Erscheinungen zu trennen. Auf diese Frage kommen wir am Schlusse (14. Kap.) zurück. Die erscheinungsmäßigen Feststellungen sind davon durchaus unabhängig<sup>1)</sup>.

Es dürfte keinen eigentlichen Übergang zwischen Klängen und Geräuschen geben. Aber Mischungen gibt es in Fülle. Und zwar sind Geräuschen in der Regel mehr oder weniger Töne beigemischt<sup>2)</sup>. Dagegen kommen vollkommen geräuschfreie Töne häufig vor und ist WILHELM BUSCHS berühmter Vers: „Musik wird oft nicht schön gefunden, weil sie stets mit Geräusch verbunden“ eine poetische Übertreibung. Nur darin könnte er recht haben, daß für besonders unmusikalische Individuen selbst die reinsten Töne stets mit Geräuschen verbunden wären und daß darin vielleicht eine der Ursachen dieses auffallenden Typusunterschiedes läge.

2. Den Unterschied von Vokalen und Konsonanten definieren heißt: eine sachlich zu rechtfertigende und hinreichend wichtige Unterscheidung unter den Sprachlauten aufzeigen, die sich mit diesen hergebrachten, uns geläufigen Klassen so gut als möglich deckt. Ihre Unterscheidung als „Selbstlauter“ und „Mittlauter“, worauf der Ausdruck „Konsonant“ beruht, war offenbar keine vollgenügende. Denn man kann Konsonanten auch für sich aussprechen und sogar zu ganzen Wörtern wie „Pst!“, ja zu dem ganzen tschechischen Satze „strč prst skrz krk“ (wo r die Funktion eines Vokals übernimmt) verbinden. Die Frage ist nur, welche bessere Definition an die Stelle zu setzen wäre<sup>3)</sup>.

<sup>1)</sup> Vgl. Tonpsych. II, S. 497ff. Wenn K. L. SCHAEFER in seiner sonst trefflichen Darstellung des Gehörsinnes (2, S. 583) gegen die spezifische Verschiedenheit der beiden Empfindungsgruppen einwendet, daß man Geräusche aus rasch wechselnden Impulsen von verschiedener Schwingungszahl zusammensetzen könne, so kann ich die Beweiskraft dieses Arguments nicht zugeben. Man muß eben die deskriptive von der genetischen Frage scheiden. Impulse sind nicht Töne, sondern nur deren äußere Ursache.

<sup>2)</sup> Den Satz, „daß jedem nicht allzu leisen Geräusch objektiven Ursprungs Töne beigemischt sind“ (a. a. O. S. 500), würde ich auch in dieser eingeschränkten Fassung heute nicht mehr aufrechterhalten. Das Regengeräusch z. B. scheint ganz tonfrei zu sein. Man verwechsle nicht die in Geräuschen, z. B. in allen Konsonanten, erkennbaren Tonhöhen mit Tönen. Diese sog. Tonhöhen, die ich selbst früher auf beigemischte Töne deutete, dürften vielmehr reine Geräuschhöhen sein. Vgl. das folgende Kapitel.

<sup>3)</sup> Die schärfste Kritik an den überlieferten Kategorien hat SIEVERS geübt. Aber seine eigene Einteilung ist mir nicht klar genug geworden, um dazu Stellung zu nehmen. Für uns muß wieder das rein Akustische maßgebend sein.

Nun könnte man versuchen, den Unterschied einfach durch den von Klängen und Geräuschen zu definieren und zu sagen: „Vokale sind Sprachklänge, Konsonanten Sprachgeräusche.“ Aber wo blieben dann die geflüsterten Vokale? Man kann sie doch nur den Geräuschen zuzählen. Der Verfasser hat früher die Definition dadurch zu retten versucht, daß er die Vokalität dieser Geräusche auf beigemischte leise Töne zurückführte, die durch das Anblasen der Mundhöhle auch bei ruhenden Stimmlippen hervorgebracht würden (12, S. 168). Aber mit dieser Anschauung läßt sich die Tatsache nicht vereinigen, daß die Tonhöhe der Geräusche bei den geflüsterten U, O und A um eine Oktave höher ist als die Formanten der gesungenen Vokale, obgleich die Mundstellung beim Flüstern dieselbe ist wie beim Singen (s. 6. Kap.). Die beigemischten Töne müßten also einen anderen Vokalcharakter bedingen als er tatsächlich vorhanden ist.

Daher erscheint mir folgende Definition jetzt als die richtige: „Vokale sind sprachlich herstellbare Klänge oder Geräusche mit ausgeprägter Färbung, Konsonanten aber sprachlich herstellbare Geräusche ohne ausgeprägte Färbung.“

Zur Erläuterung folgendes. Das, was Töne von Geräuschen unterscheidet, möge als Grundqualität oder Substanz dieser Empfindungen bezeichnet werden. Sie ist verschieden für beide Klassen, dieselbe für alle Glieder einer Klasse. Jede einzelne Empfindung einer dieser Klassen hat aber gewisse Eigenschaften oder Attribute, durch die sie näher charakterisiert werden kann, wie Höhe, Stärke, Volumen usw. Vergleichen wir in dieser Hinsicht Klänge mit Geräuschen, so zeigt sich bei genauer Analyse, daß den Geräuschen die nämlichen Attribute zukommen wie den Klängen. Außer den zeitlichen und räumlichen Eigenschaften besitzen sie offenbar Stärke-, ebenso auch Höhen- oder Helligkeitsunterschiede, ja sogar solche der „musikalischen Qualität“ (über diesen Begriff s. o. S. 91). Denn auch Oktaven und andere Intervalle lassen sich in bloßen Geräuschen erkennen und dürften nicht bloß auf beigemischten Tönen beruhen.

Außer diesen Attributen muß aber auch „Farbigkeit“ als gemeinsames Attribut von Klängen und Geräuschen anerkannt werden, wenn anders geflüsterte Vokale als echte Vokale zu gelten haben. Aber sie ist bei Geräuschen in sehr verschiedenem Maße ausgeprägt: ganz deutlich bei den Flüstervokalen, dagegen nur sehr schwach bei den stimmlosen Konsonanten. Zwar hat ein dunkles breites Sch Verwandtschaft mit U, ein Ch palatale Verwandtschaft mit J, da bei diesen Vokalen sich die Vokalitäts- mit Helligkeits-

unterschieden decken. Aber Konsonanten mit ausgeprägtem O-, A-, Ä-Charakter wird man nicht finden. Die Farbigkeit im engeren Sinne also, das Analogon der getönten Farben des Gesichtssinnes, fehlt den Konsonanten im echten, vollen Sinne dieses Begriffes.

Daß wir uns hier nicht in bloßen Tautologien bewegen, wird später noch klarer werden, wenn von den Vokalitäten der einfachen Töne zu sprechen sein wird (13. Kap.). Es handelt sich darum, daß das Moment der Farbigkeit ebenso als ein wesentliches Attribut von Gehörsempfindungen anerkannt wird, wie das der Tönung längst als Attribut von Gesichtsempfindungen anerkannt ist, daß aber dieses Attribut nicht allen Empfindungen desselben Sinnes in gleich ausgeprägtem Maße zukommt.

Mit diesem Unterschiede hängt die alte Entgegensetzung von Selbstlautern und Mitlautern, die sich auf die Funktion der beiden Klassen im Zusammenhang des Sprechens bezieht, augenscheinlich zusammen. Man versteht, daß im großen und ganzen die farbigen Laute als die vollwichtigeren gelten, wenn auch in verschiedenen Sprachen sich das Gewicht der beiden Bestandteile verschieden verteilt und gelegentlich echte Konsonanten oder „Halbvokale“ vokalische Funktionen übernehmen können. Überdies gehen ja tatsächlich immer Färbungen von Vokalen auf begleitende Konsonanten über. Man kann daher die Bedeutung der alten Unterscheidung sehr wohl anerkennen, aber nur eben als konsekutive, nicht als konstitutive.

Legen wir obige Definitionen zugrunde, so ergibt sich ohne weiteres, daß der Unterschied von Vokalen und Konsonanten kein ganz scharfer sein kann, da eben die Färbung auch bei manchen im populären Gebrauche zu den Konsonanten gerechneten Lauten doch so deutlich sein kann, daß sie den Vokalen naherücken, wenn nicht gar in sie eingeordnet werden müssen. Tatsächlich sind denn auch längst „Halbvokale“ als Übergang dazwischengestellt worden.

Die Linguisten pflegen sich jetzt zumeist mit genetischen Unterscheidungen zu helfen. Am schärfsten tritt dieser Standpunkt wohl in TECHMERS Definitionen hervor: „Vokale sind Laute mit größter oraler Apertur, Konsonanten Laute mit oraler Enge oder oralem Verschuß.“ Zu deutsch: bei den Vokalen macht man den Mund weit auf, bei den Konsonanten weniger oder gar nicht. Übergänge ergeben sich auch für diesen Standpunkt. Wir bestreiten nun nicht, daß sich so oder ähnlich eine für den Sprachforscher brauchbare und bequeme Scheidung vollziehen läßt. Aber der Akustiker wird sofort fragen, ob und worin denn die so hervorgebrachten Laute sich auch als Gehörserscheinungen voneinander unterscheiden.

Gleichfalls genetisch, aber nicht physiologisch, sondern rein physikalisch ist die Grenzziehung F. TRENDELENBURGS (II): „Vokale sind rein periodische Klänge, stimmhafte Konsonanten sind Klanggemische“ (von periodischen und unperiodischen Schwingungen, letztere durch das Anblasen der Mund-

höhle zum Stimmton hinzukommend). Die Frage nach dem phänomenalen Unterschiede läßt auch diese Abgrenzung offen. Außerdem entsteht aber auch die Frage, wie sich die Flüstervokale ihr einordnen. Vielleicht sind sie wirklich, physikalisch betrachtet, rein periodische Klänge, aber dann jedenfalls so ungeheuer zusammengesetzt, daß ihre Zerlegung in eine begrenzte Zahl von Sinusschwingungen aussichtslos und die reine Periodik schwer nachweisbar wäre. Sie würden sich von den stimmlosen Konsonanten darin kaum unterscheiden. Wenn ich recht sehe, müßte aber auch ein stimmhafter Vokal durch Beimischung unharmonischer, zum Grundton in keinem rationellen Verhältnis stehender Teiltöne ohne weiteres in einen stimmhaften Konsonanten (ein Klanggemisch) übergehen. Sofern es sich um bloße Definitionen handelt, kann man dem natürlich nicht widersprechen. Aber mit den Phänomenen und dem darauf bezüglichen allgemeinen Sprachgebrauch würde dies schwerlich stimmen. Vgl. zur Beschreibung der so entstehenden Erscheinungen das 7. Kapitel unter VI.

3. Der Unterschied der stimmhaften Sprachlaute von den stimmlosen deckt sich akustisch nicht, sondern kreuzt sich mit dem der Vokale und Konsonanten. Bei den Vokalen haben wir bereits in der Definition diesem Unterschiede Rechnung getragen: die stimmhaften Vokale sind Klänge, die stimmlosen Geräusche. Bei den Konsonanten aber sind eigentlich nur die stimmlosen wahre Konsonanten, da nur sie reine Geräusche sind. Stimmhafte aber, wie sie im zusammenhängenden Singen und Sprechen beständig vorkommen, aber auch isoliert erzeugt werden können, sind eben nicht reine Konsonanten, sondern mit gesungenen oder gesprochenen Klängen vermischt, die bei ein und demselben Konsonanten je nach der augenblicklichen Tonhöhe der Stimme sehr verschieden sein können. Was konsonantisch darin ist und unabhängig von der augenblicklichen Tonhöhe bei einem bestimmten Konsonanten regelmäßig vorhanden ist, das ist geräuschiger Art.

Bei den Reibe- oder Zischlauten kommt übrigens die stimmhafte Aussprache, obgleich sie für Sch und S an sich möglich ist, in der normalen deutschen Sprache kaum vor, und für F und palatales Ch ist sie überhaupt nur unvollkommen herzustellen, sie leiden darunter. Stimmhaftes H, dessen Vorkommen bereits PURKINJE und CZERMAK behaupteten, soll nach E. A. MEYER<sup>1)</sup> nicht bloß in außerdeutschen (slawischen, finnisch-ugrischen) Sprachen, sondern auch im Deutschen vorkommen, und zwar regelmäßig zwischen stimmhaften Lauten, z. B. in „Daheim, Freiheit“. Aber der Beweis wird nur auf graphischem Wege, mit dem Kehltonschreiber geführt und der akustische Standpunkt wird ausdrücklich abgelehnt<sup>2)</sup>.

<sup>1)</sup> Stimmhaftes H. Zeitschr. „Neuere Sprachen“ Bd. 8 (Phonetische Studien, N. F. Bd. 14), 1900.

<sup>2)</sup> S. 269: „Nicht vom Gehörseindruck muß man ausgehen. . . . Was der eine deutlich hört, ist dem anderen unhörbar, und was ein und derselbe heute nicht hört, wird er vielleicht morgen hören.“

In der Definition: „Stimmhaft nenne ich einen Laut, während dessen Hervorbringung die Stimmbänder Schwingungen von solcher Geschwindigkeit und Regelmäßigkeit vollführen, daß sie unter günstigen Umständen vom menschlichen Ohr als Ton empfunden werden“, wird nun freilich das Gehör zuletzt doch angerufen; aber nur unter günstigen Umständen soll sein Zeugnis gelten. Wir meinen dagegen, daß es in Fragen der Stimmhaftigkeit durchaus und allein ausschlaggebend sein müsse, und möchten von diesem Standpunkte das Vorkommen der stimmhaften Form im Deutschen bezweifeln. Freilich wird man niemals ein H finden, das nicht eine gewisse Tonhöhe, richtiger Geräuschhöhe, besäße, und diese ist im Zusammenhange immer bestimmt durch einen angrenzenden Vokal. Aber eine Geräuschhöhe hat auch jeder andere stimmlose Konsonant, ohne dadurch zum stimmhaften zu werden<sup>1)</sup>.

Die Verschlußlaute vollends können, wenn sie rein gesprochen werden sollen, nur stimmlos angegeben werden. Die Vorsetzung eines leisen stimmhaften M oder N vor die Mediae, wodurch bei norddeutscher Aussprache der Anschein eines stimmhaften D, B, G entsteht, ist im Grunde eine akustische Unsitte, da sie aus den einfachen Konsonanten konsonantische Diphthongen macht<sup>2)</sup>. Bei der gewöhnlichen Aussprache eines isolierten stimmhaften deutschen W wird eine Art U vorausgeschickt, das gleichfalls nicht zum W selbst gehört. Dieses ist, stimmlos und im isolierten Zustand gesprochen, fast unhörbar leise.

Es wäre verkehrt, zu schließen, daß im Zusammenhange des gewöhnlichen stimmhaften Sprechens infolge der Einwirkung der umgebenden Vokale auch die Konsonanten durchweg stimmhaft gesprochen würden. Von den sog. „Resonanten“ (M, N, Ng) und L gilt dies allerdings, aber da ist auch fraglich, ob man sie noch den Konsonanten zurechnen dürfe (s. u.). J ist zweifellos ein Vokal. Bei den übrigen findet im stimmhaften Sprechen ein kurzer gleiten-

<sup>1)</sup> Über die Rolle des stimmhaften H im Slawischen s. BROCH S. 19, 66ff., 186. Ein im vollen Wortsinne stimmloses H findet sich hiernach im Slawischen überhaupt selten. Andererseits ist beim stimmhaften H keine Vollstimme zu hören, sondern eine „Murmelstimme, wie beim Stöhnen“. In seiner Tonhöhe richtet es sich weniger als das stimmlose nach den Nachbarvokalen. Übrigens bemerkt auch BROCH gegen MEYER, daß man von einem stimmhaften H nur reden dürfe, wenn ein geschultes Ohr wirklich den Stimmklang wahrnimmt.

Der Lektor für slawische Sprachen, Herr LANE, hat mir ein solches vorgesprochen, wie es im Ukrainischen gebräuchlich sein soll. Es enthielt in der Tat einen tiefen Stimmtton, ähnelte aber stark einem Ch gutturale.

<sup>2)</sup> Ebenso urteilt NAGEL, der hier (mit Bremer) von der Vorsetzung von „Blählauten“ spricht.



der Übergang zwischen Vokalen und Konsonanten statt und wird die Aussprache des Konsonanten durch den vokalischen Zusammenhang mitbedingt (man spreche z. B.: rusch, rasch, risch, Lust, Last, List, ku, ka, ki, uk, ak, ik); aber sie verlieren dadurch nicht ihren wesentlich ton- und farblosen Charakter.

## II. Analyse der Flüstervokale durch Interferenzröhren.

Man kann das Interferenzverfahren bei geflüsterten ebenso wie bei gesungenen Vokalen anwenden<sup>1)</sup>. Nur muß man statt der Einstellung der Seitenröhren auf bestimmte Teiltöne ganze Regionen des Tongebietes durch zahlreiche Röhren von fein abgestuften Stempellängen ausschließen. Denn Geräusche erstrecken sich immer stetig oder in unmerklich kleinen Abständen, wie ein Tonstaub, über ziemlich große Teile des Tonreiches. Bei einigen Versuchsreihen gebrauchte ich zur völligen Vernichtung 59, bei anderen 49 oder 34 Seitenröhren; manchmal reichten bei Beachtung gewisser Ersparnismaßnahmen 26. Man kann bei Flüstervokalen und schwachen Konsonanten allenfalls auch ohne durchbohrte Wände auskommen, wenigstens in wissentlichen Versuchen; nur muß das der Leitung abgewandte Ohr verschlossen werden. Ich habe jedoch nur bei den allerschwächsten Lauten, wo möglichste Kürze der Leitung angestrebt werden mußte, auf die Zwischenwand verzichtet. Vor Beginn der Einstellungen ist es hier besonders wichtig, festzustellen, ob der Laut unverändert und natürlich, mindestens noch gut erkennbar, aus der Leitung kommt. Eventuell können Anweisungen an den Flüsternden, Verkürzung der Leitung, Vorschaltung eines geeigneten Trichters helfen. Im übrigen gilt auch hier, was im 2. Kap. von den Ab- und Aufbaureihen, den Lücken- und Stichversuchen, Längs- und Querschnitten, wissentlichen und unwissentlichen Versuchen gesagt ist.

Gute Hörschärfe des Beobachters, tadellose Aussprache des Flüsterers sind selbstverständliche Voraussetzungen. Manche unterscheiden z. B. nicht sehr scharf zwischen geflüstertem Ö und Ä, sowohl wenn sie selbst flüstern als wenn sie beobachten (ich glaubte bei Österreichern diese Eigentümlichkeit zu finden). Im folgenden sind nur Ergebnisse benutzt, denen auch kleinere Abnormitäten in dieser Hinsicht nicht anhafteten. Die Zahl der Beobachtungsreihen über Flüstervokale, deren eine oft mehrere Stunden füllte, betrug über 150. Mitarbeiter waren die schon o. S. 7 Genannten.

<sup>1)</sup> Meine Untersuchungen nach der If.-Methode nahmen sogar von den Flüstervokalen und den höchstgelegenen Konsonanten (S, F, Ch) ihren Ausgang.

## 1. Ab- und Aufbauversuche.

Das allgemeinste Ergebnis dieser Reihen ist, daß alle Flüstervokale innerhalb des Tonbezirkes von  $c^1$  bis zur Mitte der 5-gestrichenen Oktave, d. h. von etwa 250 bis 6000 Schw., enthalten sind. Sie füllen auf dieser Linie mehr oder minder große Strecken aus. Der Umfang dieser Strecken wächst vom U bis zum I; aber vom Ö an tritt eine Zerreiung ein: es liegen zwischen einer ganz schwachen tiefen Abteilung und der Formantregion Nullstrecken. I enthlt die grte derartige Lcke. Wir erschließen dies aus dem Umstande, da beim Abbau, nachdem der Vokal seinen Charakter verloren hat, ein dunkler Rest brigbleibt, der erst erheblich weiter unten verschwindet, und da ebenso beim Aufbau, nachdem eine dunkle Grundlage entstanden ist, eine lngere Strecke hindurch bis zum Eintritt in die Formantregion keine wesentliche Vernderung stattfindet. In diesen toten Strecken bleiben denn auch Stichversuche, die sonst bedeutende Vernderungen erzielen knnen, wirkungslos. Sie entsprechen den bei gesungenen Vokalen gefundenen. Sie scheinen bei den Flstervokalen nicht absolut leer zu sein, da beim Aufbau eine kleine Erhellung der Unterlage im Lauf der toten Strecke doch zu bemerken ist. Bei den Resonanzversuchen mit gesungenen Vokalen fanden wir ja gleichfalls in unmittelbarer Nhe der Klangquelle schwache Teiltne in diesen Strecken.

Die qualitativen Umwandlungen sind die nmlichen wie bei gesungenen Vokalen. I geht beim Abbau in U, E in O,  in den Blklaut, dann in AO ber usf. Die Umwandlungen treten im allgemeinen auch in denselben Gegenden auf, nur etwas hher, weil die If.-Breite sich bei diesen schwachen Lauten darin geltend macht, da die Schwchungsbreite mit der Vernichtungsbreite zusammenfllt.

Die folgende Tabelle stellt diese Umwandlungen in Analogie zu denen der stimmhaften Vokale bersichtlich dar. Die 1. Kolumne bedeutet, da die If.-Rhren von der oberen Grenze (0,3 oder 0,5) bis zu der betreffenden Stempellnge in mglichst kleinen Abstnden eingestellt sind. Die 2. gibt an, bis zu welcher Tonhhe von oben herab die Vernichtung reicht, wenn von der If.-Breite abgesehen wird, also nach der Tabelle S. 47<sup>1)</sup>. Die 3. gibt die unter Bercksichtigung der If.-Breite (o. S. 43) reduzierten Hhen, also die Grenze, bis zu welcher herab Vernichtung der leisen Flstergerusche anzunehmen ist.

<sup>1)</sup> Gegenber m. Abh. 11, S. 243 sind hier von der oberen Tongrenze bis herab zu  $g^3$  die durch Abh. 17 geforderten Korrekturen eingesetzt.

## Aufbau der Flüstervokale mit Interferenzversuchen.

1,5	<i>cis</i> <sup>5</sup>	<i>a</i> <sup>4</sup>								I gut
2	<i>a</i> <sup>4</sup>	<i>f</i> <sup>4</sup>							reines Ü	E gut
2,5	<i>fis</i> <sup>4</sup>	<i>d</i> <sup>4</sup>					Ä tadellos	Ü gut	E	I deutlich
3	<i>e</i> <sup>4</sup>	<i>c</i> <sup>4</sup>					Ä	Ü dumpf	O + E; Eö	I etwas pelzig
3,5	<i>d</i> <sup>4</sup>	<i>b</i> <sup>3</sup>					AoÄ	Üü	Oö	Blasen mit Ü; Ui
4	<i>c</i> <sup>4</sup>	<i>a</i> <sup>3</sup>					Aoä	U etwas bedeckt	O etwas heller	mehr Blasen
4,5	<i>b</i> <sup>3</sup>	<i>g</i> <sup>3</sup>						U blasend		
5	<i>gis</i> <sup>3</sup>	<i>f</i> <sup>3</sup>			A gut					
5,5	<i>fis</i> <sup>3</sup>	<i>es</i> <sup>3</sup>						geräuschiges Ö		
6	<i>f</i> <sup>3</sup>	<i>d</i> <sup>3</sup>						Oö		
7	<i>es</i> <sup>3</sup>	<i>c</i> <sup>3</sup>			Recht gut			O		
8	<i>cis</i> <sup>3</sup>	<i>b</i> <sup>2</sup>		Osehr gut	Ao heller			Ao		
9	<i>h</i> <sup>2</sup>	<i>gis</i> <sup>2</sup>						Ao		
10	<i>a</i> <sup>2</sup>	<i>fis</i> <sup>2</sup>	U fertig	volles O	Ao			AO		
12	<i>fis</i> <sup>2</sup>	<i>es</i> <sup>2</sup>	kräftiges U		AO					
14	<i>es</i> <sup>2</sup>	<i>c</i> <sup>2</sup>	gutes U	Ou	leises O	UO hauchend		helleres U		
16	<i>cis</i> <sup>2</sup>	<i>b</i> <sup>1</sup>	U	UO	leises UO				U-artiges Hauchen	
18	<i>h</i> <sup>1</sup>	<i>gis</i> <sup>1</sup>	U-ähnliches Geräusch . . . . .							
22	<i>g</i> <sup>1</sup>	<i>e</i> <sup>1</sup>	leisestes charakterloses Geräusch			unhörbar	unhörbar	unhörbar	unhörbar	unhörbar
cm	Un-reduziert	Reduziert	U	O	A	Ö	Ä	Ü	E	I

Auch die Formanten wiederholen sich in analoger Weise. Da kein wechselnder Grundton vorhanden ist, ist ihre feste Lage hier selbstverständlich. Auch U, das bei gesungenen Vokalen keinen festen Formanten hat, sondern mit dem jeweiligen Grundton beginnt, hat hier einen festen Anfang in der Tonreihe. Es

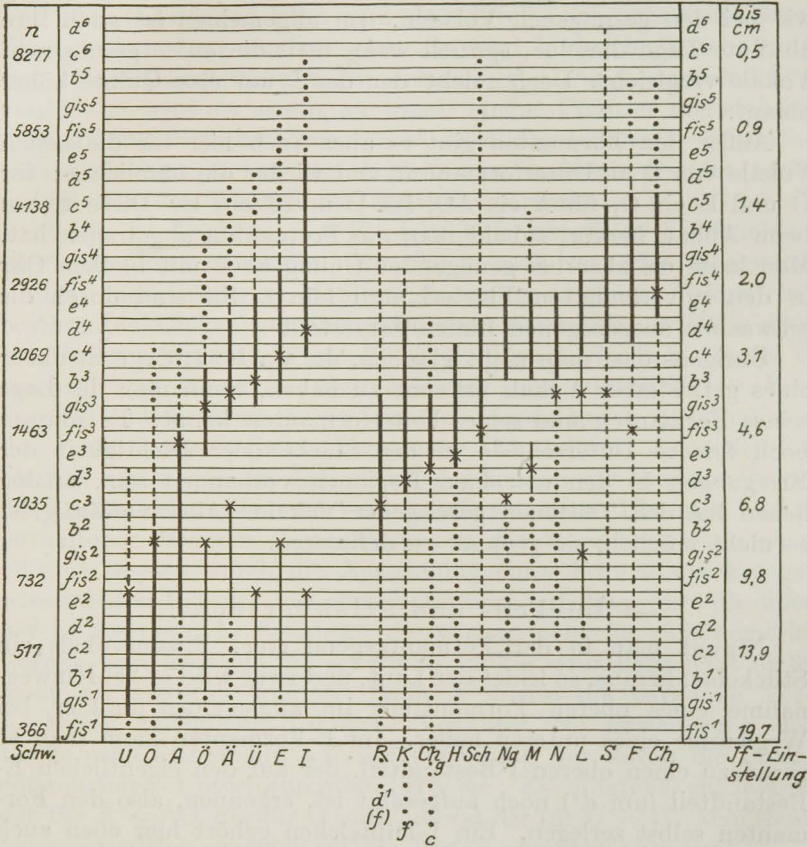


Abb. 2. Struktur stimmloser Vokale und Konsonanten.

wird als U kenntlich, sobald beim Aufbau  $b^1$  erreicht ist, also mit Abrechnung der If.-Breite bei etwa  $g^1$ .

Für die Bestimmung der Formanten gelten die S. 63 aufgestellten Regeln; nur muß auf die If.-Breite, also die reduzierten Höhen, Rücksicht genommen werden.

In dem vorstehenden graphischen Schema, in das auch die Konsonanten im voraus mitaufgenommen sind, bedeuten die stark ausgezogenen Linien die Formantstrecken. An sie schließen

sich aber hier nach oben hin noch gestrichelte und punktierte an, in deren Bezirk sich die Vokale noch etwas vervollkommen, ohne daß aber diese Gegenden für den spezifischen Charakter ausschlaggebend wären. Die  $\times$ -Zeichen bedeuten die später zu besprechenden Flüsterhöhen.

Die Formanten folgen sich von U bis I in derselben Ordnung wie bei den gesungenen Vokalen. Im allgemeinen ist auch ihre absolute Lage dieselbe, speziell wenn man die auf *c* gesungenen Vokale vergleicht. Doch reicht der des I um eine Quinte höher hinauf.

Außer den Formanten gibt es aber auch hier für die hellen Vokale von Ö an Unterformanten, und wieder die nämlichen: für Ö und E ein O, für Ä ein AO, für Ü und I ein U. Diese treten beim Abbau hervor, sobald man die Formanten abgetragen hat. Man kann sie aber bei genügender Übung auch mit bloßem Ohr in den unveränderten Flüsterlauten hören. Sie sind durch die schwachen ausgezogenen Linien dargestellt<sup>1)</sup>.

Übrigens darf man nicht glauben, die vollständige Struktur eines geflüsterten Vokals erkannt zu haben, wenn man die Lage seines Formanten und seines Unterformanten kennt. Es müssen noch feinere Unterschiede in der Stärke oder Dichtigkeit der Erregungen in den einzelnen Regionen vorhanden sein, analog denen bei den Teiltönen gesungener Vokale. Aber vorläufig ist es nicht möglich, sie genauer zu definieren.

## 2. Lücken- und Stichversuche.

Nimmt man in der Formantengegend eines Flüstervokals ein Stückchen heraus, so leidet der Laut, und zwar wird er bei Hinwegnahme eines oberen Formantenteils im allgemeinen dunkler, bei Wegnahme eines unteren heller. Im E-Formanten kann man so geradezu einen oberen I-Bestandteil, der auf den eigentlichen E-Bestandteil (um *d*<sup>4</sup>) noch aufgesetzt ist, erkennen, also den Formanten selbst zerlegen. Ein I-Tüpfelchen gehört hier eben auch zum E. Das ohnedies schwache U geht jedoch durch jeden solchen Lückenversuch in ein bloßes Blasen über.

<sup>1)</sup> Nur die Ganztonstufen der chromatischen Leiter sind in den Notenbuchstaben des Schemas angegeben, die dazwischenliegenden Töne aber durch die entsprechenden Horizontallinien angezeigt.

Einige Differenzen von einer halben oder ganzen Tonstufe zwischen diesem Schema und dem des Aufbaues S. 106 rühren daher, daß hier die aus den einzelnen Versuchsreihen ersichtlichen äußersten Punkte mitberücksichtigt sind, während dort nur die durchschnittlichen Ergebnisse wiedergegeben wurden.

Bei Herausnahme eines mittleren Formantstückes werden alle Flüstervokale unkenntlich. Die drei hellsten, Ü, E, I, werden gemeinsam zerstört, wenn man auch nur auf 2,4 bis 2,6 cm einstellt, weil zufolge der If.-Breite die Wirkung sich von dem dadurch direkt ausgeschlossenen  $g^4$  noch um eine Terz nach beiden Seiten ausdehnt.

Unter Umständen kann man aber durch Lückenversuche einen Flüstervokal sogar verbessern. So wird U durch Ausschaltung von  $c^3-f^3$  (Einstellung 6—8) verschönert. Durch Ausschaltung von  $f^2-f^3$  wird es schon zu tonal, einem Pfeifton angenähert. Ä wird durch Herausnahme von  $c^3-gis^3$  (Einstellung 5—8) gleichfalls verbessert, indem das begleitende Geräusch zerstört wird, wobei auch der Wegfall der Multipla mitwirkt.

Auch die Tonhöhe der Flüstervokale unterliegt bei Lückenversuchen Veränderungen, von denen noch zu reden sein wird.

Die merkwürdigen Metamorphosen gesungener Vokale bei bestimmten Lückenversuchen wiederholen sich hier. So kann man aus U durch Einstellung 13—25 ein dunkles A, durch 8—14 ein Ö machen, aus einem A durch 6—8 ein AoÄ, durch 6—10 Äao, durch 6—14 Ö, durch 6—16 leises Ü. Schon durch Ausschaltung des  $c^3$  mit seiner näheren Umgebung (Einstellung 7—8 oder 7,6 bis 8,6) erhält man ein ÄÖ. Auch hier können wir aber auf Grund des bereits Erkannten zum Verständnis gelangen. Die im U und A enthaltenen höheren Bestandteile aus der 3- und 4-gestrichenen Oktave werden durch die Ausstoßung eines Formantstückes gewissermaßen frei, die Verschmelzung hört auf, man hört die über der Lücke liegende Vokalität, die vorher nicht für sich bemerkt worden war, wogegen die schwache dunkle Grundlage jetzt nicht mehr zur Geltung kommt. Beim A dürfte die paradoxe Verschiebung zu immer hellerer Vokalität bei Erweiterung der Lücke nach unten hin auf der Vertiefung der durch die Lücke entstehenden Unterformanten beruhen. Es entsteht sukzessive der für Ä = AO, der für Ö = O, zuletzt der für Ü = U. Man sieht hieraus zugleich, wieviel auf die Unterformanten ankommt.

Auch bloße Stichversuche sind wieder lehrreich. Man kann dadurch einen Flüstervokal abtasten und die mehr oder weniger schädigende Wirkung der Wegnahme einzelner kleinster Abschnitte untersuchen. Dabei muß man von oben nach unten, von kleineren zu größeren Einstellungen fortgehen, um bei der Beurteilung der Wirkungen die vorher ermittelte Wirkung der Multipla mit berücksichtigen zu können. A wird z. B. schon durch eine einzige Einstellung in seiner Formantmitte zu einem „sonderbaren Gemisch, etwas Ö enthaltend“. Bei Ö kann man durch Einstich E

und O nebeneinander hören, beide allerdings nur schwach gefärbt. Ü wird gleichfalls durch Stich in die Formantmitte ( $3,5 = d^4$ ) zu einem Nebeneinander eines dunklen und eines hellen Geräusches.

### III. Analyse von Konsonanten durch Interferenzröhren.

Auch die meisten Konsonanten lassen sich dem If.-Verfahren unterwerfen<sup>1)</sup>. Ich habe zahlreiche Versuchsreihen darüber durchgeführt bei den Zisch- und Reibelauten Sch, S, F, Ch, den Explosivlauten K, T, P, sowie den stimmlos angegebenen Lauten R, M, N, Ng, L und H. Konsonantische Diphthongen, wie Z, blieben natürlich ausgeschlossen. Obige Laute kommen, mit Ausnahme der 5 letzten, aus einer nicht zu langen und nicht zu engen Röhrenleitung, solange noch keine seitlichen Interferenzröhren eingestellt sind, mindestens so kräftig wie beim Flüstern innerhalb eines Zimmers und sind bei zweckmäßiger Versuchseinrichtung qualitativ wohl erhalten, wenn auch minimale Veränderungen nicht absolut vermieden werden können. Man muß bedenken, daß auch in der Praxis des gewöhnlichen Lebens die Umstände, unter denen wir sie hören und verstehen, äußerst verschieden sind, und daß beim Flüstern aus größerer Entfernung wahrscheinlich sogar stärkere Modifikationen auftreten als bei einer kurzen Röhrenleitung, ohne daß der Laut unverständlich zu werden braucht. Auch individuelle Verschiedenheiten der Aussprache in bezug auf Deutlichkeit und Charakter der einzelnen Konsonanten sind, wie bei den Vokalen, selbst unter normal Sprechenden noch bedeutend genug. Wenn man darauf achtet, wird man immer wieder davon überrascht. Bei wissenschaftlichen Versuchen müssen manche Individuen, die man im gewöhnlichen Verkehr ohne jede Schwierigkeit, auch beim Flüstern, versteht, für bestimmte Laute geradezu ausgeschieden werden.

Bei M, N, Ng und L, die sich ihrer Schwäche wegen auf dem Interferenzwege in stimmloser Form nur unvollkommen untersuchen lassen, wurden die stimmhaften Formen zu Hilfe genommen und daraus gewissermaßen durch Subtraktion des Stimmtones Rückschlüsse auf die stimmlosen gezogen, die mit den direkten Ergebnissen verglichen werden konnten. Auch das stimmlose R ließ sich so mit dem stimmhaften vergleichen.

<sup>1)</sup> Bereits HERMANN hat dies versucht, aber unbegreiflicherweise keine entschiedene Deformation damit erzielt (Bd. 83, S. 24). Besser gelang es KOEHLER (1, III, S. 24ff., 74ff.; vgl. Vorläuf. Mitt. S. 92ff.). Durchsichtige Ergebnisse sind aber auch hier nur durch ganz systematischen Ab- und Aufbau zu gewinnen.

Für die Aussprache galt im allgemeinen, wie bei den Vokalen, die Regel, die Laute so charakteristisch als möglich, d. h. möglichst verschieden von den benachbarten, verwandten Lauten anzugeben. Jedoch wurde Sch, das sehr dunkel und sehr hell gesprochen werden kann, mittelhell gegeben, S dagegen so scharf als möglich (Ss), Ch in zweifacher Form: als vorderes (palatales), möglichst hell, fein und dünn, wie in „Ich“ oder „Chi“, und als hinteres (gutturales, velares), wie in „Ach“. R als gutes Zungen-R. K, das großen Verschiedenheiten je nach dem darauffolgenden Vokal unterliegt, mit der Mundstellung, wie sie der Silbe kö (doch nach ka hin verdunkelt) entsprechen würde. Es ist dies wieder, wie bei Sch, eine mittlere Helligkeit. Die Explosivlaute immer hart und kräftig. Ng selbstverständlich als einfacher Laut, nicht als konsonantischer Diphthong. Alle Laute durften nicht übertrieben stark genommen werden, da sonst gewisse Veränderungen eintreten.

Man kann zur Charakteristik der jeweiligen Nuance eines Konsonanten auch seine im Geräusch erkennbare Tonhöhe benutzen, wenn das Ohr dafür geschult ist. So wurde Sch mit der Tonhöhe  $fis^3$  gegeben, hinteres Ch mit der Höhe  $es^3$ , vorderes mit  $f^4$ , K mit  $d^3$ .

Betreffs der äußeren Einrichtungen gilt dasselbe wie bei den Flüstervokalen. Für die tieferen Teile der Konsonanten wurden Röhren von 21 mm und 18 mm lichter Weite, für die höheren und höchsten solche von 10 mm und 5 mm verwandt; zuerst sogar noch engere. Aber Röhren von 3 mm und darunter haben den Nachteil, daß sie die höchsten Konsonanten für feine Ohren eben merklich abstumpfen. Hier sind also etwas weitere zu benutzen.

Beim Abbau wurden zunächst immer die allerkleinsten Einstellungen ausprobiert, wie 0,3 cm oder 0,5 cm oder die ganze Zone 0,3—0,5. Wenn dann solche Einstellungen nicht die geringste Veränderung bewirkten, konnte bei den weiteren Versuchen über den betreffenden Laut von diesen Einstellungen abgesehen und der Abbau sogleich etwas tiefer begonnen werden, da man immer auf Röhrenersparnis bedacht sein muß.

Die angegebenen Konsonanten umfassen zusammengenommen den Tonbezirk von etwa  $c$  (die meisten beginnen aber erst in der 1- oder 2-gestrichenen Oktave) bis zum Anfang der 6-gestrichenen Oktave. Keiner geht über die jetzt gewöhnlich angenommene Hörgrenze  $f^7$  hinaus<sup>1)</sup>, die meisten bleiben erheblich darunter.

<sup>1)</sup> Dies zu bemerken ist nicht überflüssig, da KOEHLER einige Konsonanten (Ch, F) diese Grenze weit überschreiten läßt. Vgl. Anhang I dieses Kapitels.



Nach unten hin könnten sie jedoch noch etwas weiter reichen, da die tiefsten Teile durch die If.-Einstellungen auf die höheren etwas geschwächt, also, wenn sie ohnedies schwach sind, vernichtet werden können (o. S. 40).

Innerhalb dieses Bezirkes entwickeln sie sich beim Aufbau aus schwächsten charakterlosen Geräuschen, die der gemeinschaftlichen U-Unterlage der Vokale zu vergleichen sind, und erlangen erst allmählich ihre Bestimmtheit und Unterscheidbarkeit. Der Abschnitt der Tonlinie zwischen dem Punkte, wo die erste Spur des spezifischen Lautcharakters eintritt, und dem, wo er fertig erscheint, möge auch hier der Formant heißen. Für seine Bestimmung gelten dieselben Regeln wie früher, speziell bei den Flüstervokalen. Aber er ist hier nicht so scharf und bestimmt markiert wie dort. Lücken- und Stichversuche dienen wieder zur Kontrolle.

Qualitative Umwandlungen der Laute während des Ab- und Aufbaues zeigen sich hier nicht in gleicher Weise wie bei den Vokalen. Die Veränderungen gehen in einer gleichbleibenden Richtung vor sich; das Eigentümliche der betreffenden Konsonanten prägt sich nur graduell immer deutlicher aus. Sch, S, F, Ch sind beim Aufbau zunächst längere Zeit identisch. Sie differenzieren sich erst von einem bestimmten Punkte der Tonlinie aus. Macht man Querschnitte, d. h. vergleicht man sie bei einer und derselben Einstellung, so ist dies besonders deutlich. Man tut auch gut, bei Längsschnitten, d. h. beim systematischen Ab- und Aufbau eines einzelnen Lautes, gelegentlich einen solchen Querschnitt einzuschalten. Durch die Verbindung beider Methoden stellt man am genauesten die fortschreitenden Veränderungen und ihre Orte in der Tonlinie fest.

Mit der gleichbleibenden Art der Veränderungen hängen hier auch etwas größere Schwankungen und Verschiedenheiten der einzelnen Beobachter zusammen. Rein graduelle Veränderungen gestatten und verlangen eben keine so festen Klassenbegriffe. Ob man z. B. ein S bereits scharf genug findet oder noch etwas schärfer wünscht, ist einigermaßen subjektiv. Doch sind die im folgenden mitgeteilten Ergebnisse auf übereinstimmende, regelmäßige Aussagen der besten Beobachter gegründet.

Die Hauptergebnisse in bezug auf die Struktur der untersuchten Konsonanten sind in dem graphischen Schema o. S. 107 zusammengestellt. Die starken ausgezogenen Linien bedeuten wieder die Formanten, die schwächeren die Unterformanten, die gestrichelten und punktierten Teile weniger wesentliche Abschnitte des Gesamtumfanges auf der Tonlinie. Die Anordnung der Konsonanten ist,

wie die der Vokale, bestimmt durch den Beginn ihrer Formantregionen, bei gleichem Beginn durch die Erstreckung nach oben.  $Ch_g$  bedeutet gutturales (hinteres),  $Ch_p$  palatales (vorderes)  $Ch$ . Die Kreuzzeichen auf den Linien bedeuten die Tonhöhen der Laute, wie sie meinem Ohre bei den hier untersuchten Nuancen erscheinen. Diese sollen uns im folgenden Kapitel näher beschäftigen. Die If.-Einstellungen sind hier sogleich mit Berücksichtigung der If.-Breite angegeben, weichen also von der Tabelle S. 47 um die entsprechenden Beträge ab (für  $fis^1$  Einstellung auf  $a^1$  usw.).

Beschränkt man sich auf die Formanten und die Gesamtumfänge, so lassen sich diese in Notenbuchstaben so zusammenstellen:

Geflüsterte Vokale	Formanten	Gesamtumfänge	Stimmlose Konsonanten	Formanten	Gesamtumfänge
U	$g^1-f^2$	$f^1-es^3$	R ling.	$a^2-a^3$	(f) $d^1-b^4$
O	$a^1-b^2$	$f^1-des^4$	K (P, T)	$des^3-es^4$	$f-g^4$
A	$es^2-ges^3$	$f^1-des^4$	Ch gutt.	$d^3-a^3$	$c-es^4$
Ö	$e^2-h^3$	$g^1-b^4$	H	$es^3-des^4$	$g^1-des^4$
Ä	$as^3-es^4$	$g^1-d^5$	Sch	$f^3-es^4$	$a^1-c^6$
Ü	$a^3-f^4$	$g^1-es^5$	Ng, M, N	$b^3-f^4$	$c^2-f^4(c^5)$
E	$b^3-g^4$	$g^1-gis^5$	L	$c^4-f^4$	$c^2-g^4$
I	$c^4-e^5$	$g^1-c^6$	S	$des^4-c^5$	$a^1-d^6$
			F	$des^4-des^5$	$a^1-b^5$
			Ch pal.	$es^4-des^5$	$a^1-b^5$

Im einzelnen hat sich folgendes ergeben:

### R linguale<sup>1)</sup>.

Ein stimmloses gutes Zungen-R wird beim Abbau mit Einstellung 3 cm<sup>2</sup>) eben merklich geschwächt, bei 4 deutlich schwächer und tiefer; bei 6 nähert es sich dem Gaumen-R; bei 8 ist es kaum mehr als R überhaupt zu bezeichnen, zeigt aber noch scharfe Intermittenz. Bei 10 eine Art Schnarren oder Gurren, ähnlich wie beim Streichen über eine gerippte Oberfläche, nur tiefer. Bei 12 ein gurgelndes Hauchen, das Intermittieren noch deutlich; aber schwächer. Bei 20 nur ganz schwaches dunkles Geräusch, noch intermittierend. Das tonale Element überwiegt jetzt. Weiterhin immer schwächer, zuletzt auch nicht mehr intermittierend, nur wie ein Hauch. Verschwindet je nach der Stärke des Angebens und der Hörschärfe des Beobachters bei 26 bis etwa 50.

Beim Aufbau erscheint nach dem Anfangsstadium bei 30 ein ganz dunkles, tiefes, schon etwas rauhes Geräusch, bei 22 leises Gurren, bei 16

<sup>1)</sup> R kann stimmlos, d. h. ohne Innervation der Stimmlippen, aber nicht tonlos hervorgebracht werden, da der tiefe sog. „Unterbrechungston“ auch durch periodische Geräuschunterbrechungen erzeugt wird.

<sup>2)</sup> Wir geben hier überall die Einstellungswerte, weil danach die Beobachtungen am einfachsten zu kontrollieren sind. Die zugehörigen Tonhöhen sind aus der Tabelle S. 47 zu ersehen, wobei aber niemals zu vergessen ist, daß die Wirkung der If. um das Intervall der If.-Breite tiefer hinabreicht.

heller geworden, aber noch kein R, bei 12 und 10 stärker intermittierend, aber noch zu tonal, noch ohne das Schnatternde und Zischende des R. Dieses beginnt bei etwa 8. Der Laut ist auch erheblich höher geworden. Bei 6 besseres, bei 5 ziemlich gutes, bei 4 gutes R. Dann nur noch etwas heller und zischender.

Die entscheidende Zone demnach etwa 8—4, d. h. die 3-gestr. Oktave, genauer  $a^2—a^3$ . Doch ist hier besonders zu beachten, daß die Zonen nicht so bestimmt wie bei den Vokalen abzugrenzen sind.

Den tiefen „Unterbrechungston“, etwa *Fis* = 23 Schw., kann man, da er zu dem Charakter auch des stimmlosen R beiträgt, als eine Art Unterformanten ansehen.

Stichversuche ergeben von 2 an geringe Schwächung und Verdunkelung, ein Leerwerden, auch gelegentlich Erhellung, wenn man gerade auf den höheren der beiden durch den Stich getrennten Bestandteile achtet, aber keine ausgezeichneten Punkte. Die Schädigung scheint sich ziemlich gleichmäßig von 2 bis 25 zu erstrecken; nur bei Stichen zwischen 12 und 16 ( $c^2—ges^2$ ) ist sie etwas stärker. Aber es bleibt bei so kleinem Defekt immer ein gutes R. Mehr erreicht man natürlich durch Lückenversuche; doch wird selbst durch Ausschaltung der ganzen Zone 4—8 nur die Helligkeit stark vermindert, kenntlich bleibt der Laut immer noch.

Beim stimmhaften, tonalen R sind die Verhältnisse ganz dieselben und kommt nur eben der Stimmton selbst (nebst Obertönen) noch hinzu.

### K (mittelhell), T, P.

Die Explosivlaute untersucht man am besten in Verbindung miteinander, um sie auf ihre Unterscheidbarkeit bei den verschiedenen Einstellungen zu prüfen (Querschnitte). Aber nur für K ließ sich eine Art Formantregion feststellen; die beiden anderen, namentlich P, sind schon bei der Nullstellung der Röhren, ohne jede Interferenz, nicht so leicht wie K erkennbar und verlieren beim Abbau sehr bald ihre Charakteristik, verwandeln sich in indifferente, nur immer schwächer werdende Geräusche.

Wegen der bedeutenden Stärke dieser Laute und ihrer großen Erstreckung in der Tonlinie ist eine zweckmäßige Verteilung der Röhren besonders notwendig, um auszureichen. Ich wandte zuletzt folgende an: von 2 bis 12 cm jeweilige Differenz der Stempellängen 0,3; von 12,4 bis 20 Differenz 0,4; von 20,5 bis 28 Differenz 0,5; weiter bis 32 Differenz 1 bis 2 cm. Noch größere Stempellängen schwächen wohl den übriggebliebenen Rest noch ein wenig; aber er ist bei 32 schon fast Null.

Abbau: Bei 1 noch kaum eine Schädigung. Bei 2 beginnende Schwächung und Verdunkelung. Bei 3 wird K den beiden anderen schon wesentlich ähnlicher, bei 4 in gleicher Richtung erheblich geschädigt, bei 5 überhaupt nur noch ein trockenes Klopfgeräusch. Die drei Laute sind jetzt nur graduell verschieden, namentlich der Stärke nach (K am stärksten). Unwissentlich werden sie verwechselt; doch wird K noch gelegentlich als „hartes Q“ bezeichnet. Bei 7 alles nur wie H, ohne Schärfe des Ansatzes und ohne irgend deutlichen Unterschied. Bei 13 ist T recht schwach, P noch schwächer. Bei 18 auch K nur ein mattes kurzes Geräusch. Bei 30 ist K selbst für v. ALLESCHS scharfes Ohr äußerst schwach, T minimal, P verschwunden. Bei 40 von K nur eine Spur, von den beiden anderen nichts mehr.

Aufbau: Nach den minimalen Anfängen bei 20 eine Art Glucksen, bei 11 schon ein gewisser Impetus des K, aber noch kein eigentlicher Ansatz; alle drei haben etwas Knallähnliches. Bei 8 könnte man sie allen-

falls wesentlich schon voneinander unterscheiden. Bei etwa 6,5 erste Anfänge eines deutlichen K; T setzt hier weniger scharf an, noch weniger P. Bei 5 K schon ziemlich gut<sup>1)</sup>, bei 4 noch erheblich deutlicher, bei 3 wohl fertig, bei 2,6 vollkommen deutlich. T dagegen könnte hier auch als Pf, P auch als T verstanden werden. Diese beiden werden erst gegen die Nullstellung hin deutlicher.

Hiernach kann man als entscheidende Interferenzregion für K etwa 6,6–2,6 bezeichnen, demnach als Formant *des<sup>3</sup>—es<sup>4</sup>*.

Stichversuche: Der erste deutliche Unterschied gegen die Nullstellung ist bei 2,5 oder 3: etwas dunkler und schwächer. Der stärkste Unterschied bei 5 (Formantmitte): entschieden heller und schwächer, nähert sich dem T; „wie ein leises Spucken“. Hier kommen offenbar die hohen, vorher ausgeschlossenen Bestandteile wieder zur Geltung. Daß nicht etwa der Ausschluß ungerader Multipla an der Veränderung schuld ist, geht daraus hervor, daß deren direkter Ausschluß (Einstellung auf  $\frac{1}{3}$  oder  $\frac{1}{5}$  von 5) keine so wesentliche oder gar keine Veränderung bewirkt. Einstellungen über 5 hinaus schädigen den Laut wieder weniger.

### Ch gutturale (hinteres, wie in Ach).

Abbau: Bei 3 eben merklich schwächer, bei 4 etwas weicher und dunkler, bei 5 wesentlich dunkler und schwächer. Bei 5,4 ist der Laut um eine große Sexte tiefer (= *a*<sup>2</sup>). Dieser Übergang vollzieht sich aber, so scheint es, nicht durch Senkung, sondern so, daß ein höherer Bestandteil (von der scheinbaren Tonhöhe *es<sup>3</sup>—f<sup>3</sup>*) schwächer wird und verschwindet und nun der tiefere hervortritt. Bei 6 wird die Alteration sehr groß, der Laut ist fast unkenntlich. Bei 7 nur noch ein kurzes dunkles Blasen, bei 8 noch milder und schwächer. Bei 9 ist es wieder um einige Töne tiefer geworden und mehr vokalisch, wie ein blasendes O. Bei 13 noch deutliches U-ähnliches Hauchen. Bei 27 höre ich selbst nichts mehr, junge Mitbeobachter hören noch ein ganz tiefes minimales Geräusch. Die letzte Spur ließ sich hier bis 70 verfolgen.

Aufbau: Um 60 erscheint (bei starker Aussprache) die erste Spur eines tiefen Geräusches. Bei 10 ganz dunkles U-ähnliches Keuchen. Bei 8 ist es etwas höher und zugleich schraubender geworden. Bei 7 noch kein Ch. Bei 6 dagegen Ch bereits erkennbar; bei 5 besser, bei 4 schon gut, fast besser als ohne Interferenz, da es ohne solche zwar lauter, aber etwas Sch-haltig erscheint.

Entscheidende Region hiernach 6–4, also Formant *a<sup>3</sup>—a<sup>3</sup>*.

Stichversuche: Bei 4 geringer, bei 4,5 aber erheblicher Unterschied gegen Null, auch Vertiefung um einen Ganzton. Bei 5–6 stärkster Unterschied, Vertiefung um eine Sexte (oberer Teil des Formanten). Bei 6,5 wieder weniger alteriert.

### H<sup>2)</sup>.

Beim H lassen sich Ab- und Aufbauversuche natürlich wieder nur mit kürzester Leitung, und auch da nur unvollkommen anstellen. Da es in der gewöhnlichen Rede nur als Aspiration eines darauffolgenden Vokals

<sup>1)</sup> Hier tritt wieder besonders der Unterschied gegenüber dem Abbau zutage, der nur psychologisch zu verstehen ist (s. o. S. 51).

<sup>2)</sup> Durch objektive Methoden zeigte ISSERLIN (1, S. 77), daß jedem Aussprechen eines (isolierten) Vokals ein kurzes Hauchen des Mundes vorausgeht, auch wenn er nicht ausdrücklich aspiriert wird. Dies ist der Spiritus lenis der Alten, auch wohl das hebräische Aleph. Was hier untersucht wird, ist der Spiritus asper.

vorkommt (der Buchstabe als solcher auch als Dehnungszeichen), so muß es, um möglichst rein für sich untersucht zu werden, so ton- und vokallo als eben möglich gegeben werden, als reines Hauchen; sonst wird man statt der Veränderungen des H durch Interferenz die des betreffenden Vokals untersuchen. Die Mundstellung etwa wie beim Ö, aber ohne daß dieses selbst intendiert und im Laute merklich wird.

Es läßt sich hier nur sagen, daß der Laut beim Abbau mit etwa 3 dunkler und schwächer wird, und daß dies stetig weitergeht bis 18, wo nur noch ein ganz schwacher Rest vorhanden ist. Da nun die Tonhöhe eines solchen H, soweit ich urteilen kann,  $e^3$  ist, so kann man vielleicht den Formanten unter der mehrfach erwähnten Voraussetzung auf  $es^3$ — $des^4$  ansetzen. Aber zu viel Gewicht möchte ich auf diese Bestimmung nicht legen.

### Sch (mittelhell)

wird beim Abbau schon bei 0,5 eben merklich dunkler; deutlichere Verdunkelung bei 0,8. Dann immer dunkler, tiefer und schwächer. Bei 4 schon sehr charakterlos, bei 5 nur noch ein schwaches indifferentes Geräusch. An diesem Punkte sind Sch, S, F, Ch pal. beim Querschnitt fast identisch, keiner von ihnen erkennbar. Geht man mit der Abtragung weiter bis 17, so bleibt nur noch ein ganz leises tiefes Hauchen oder nichts mehr.

Beim Abbau kann man nach dem anfänglichen dunklen, leisen Blasen zuerst bei 5 und 4 ein dem Sch sich näherndes Fauchen konstatieren, bei 3 schon ein stumpfes und bei 2,6 ein gutes Sch, das sich aber weiterhin immer noch etwas verbessert.

Die entscheidende Zone liegt also etwa zwischen 5 und 2,6; woraus sich als Formant  $f^3$ — $es^4$  ergibt.

Stichversuche schädigen den Laut von 1 an immer etwas im Sinne der Verdunkelung und Schwächung, aber ein stärkerer Unterschied gegen die Nullstellung (mit der hier immer verglichen werden muß) tritt erst bei 5 und 6 auf: er wird leerer, dem Ch ähnlicher. Von 9 an wird er kaum mehr beeinflusst. Maximale Schädigung also bei 5—6 =  $f^3$ — $as^3$ , an der unteren Grenze des Formanten.

### Die Nasalkonsonanten M, N, Ng.

Diese drei Laute pflegt man jetzt „Resonanten“ zu nennen, weil der Luftstrom an die drei Verschlüßstellen gewissermaßen anpralle und zurückgeworfen werde (H. GUTZMANN 5, S. 177). Aber die Bezeichnung erscheint unzweckmäßig, da sie verleitet, an Resonanz im gewöhnlichen Sinne der Akustik zu denken, wovon doch hier nicht die Rede ist, da sie bei diesen Lauten nicht mehr und nicht anders als bei allen übrigen in Betracht kommt. Fast möchte man sagen: „Resonantes a non resonando.“ NAGEL, der den Namen gleichfalls gebraucht, rechnet die Klasse der Resonanten zu der allgemeineren Gruppe der „kontinuierlichen phonischen Laute“ (S. 758, 765), wobei er ganz zu übersehen scheint, daß sie auch stimmlos, flüsternd gesprochen werden können, und daß die stimmhafte Aussprache nur eben den Stimmton nebst seinen Obertönen zum Gehörseindruck hinzufügt. In diesem Stimmklang kann aber, da er immerfort wechselt, das Wesentliche und Charakteristische nicht gefunden werden.

Richtiger ist die von GUTZMANN auch angewandte zweite Bezeichnung: Nasallaute. In der Tat sind sie, stimmhaft gesprochen, nichts anderes als mehr oder weniger nasalisiertes U (wenigstens bis in die Höhe des  $c^2$ ); stimmlos aber sind sie die Nasalierung selbst, das Näseln in Geräuschform.

Die Untersuchung dieser Laute als reiner Flüsterkonsonanten mit der Interferenzmethode stößt aber auf Schwierigkeiten wegen ihrer geringen Stärke. Sie sind schon in der Nähe des Sprechenden für sich allein sehr schwach (besonders M) und nicht leicht voneinander zu unterscheiden; um so weniger am Ende einer längeren Röhre. Doch hört man sie und kann ihr Verschwinden und Wiedertreten verfolgen, wenn die Leitung durch Verlegung der ganzen Einrichtung einschließlich der Schallgebung in das Beobachtungszimmer aufs äußerste Maß verkürzt wird; und es ist, wenn man das nicht beschäftigte Ohr zuhält, keine Gefahr, dabei durch das direkte Hören gestört zu werden. Beobachter war hier vorzugsweise Dr. v. ALLESCH. Auch für tadellose Lautgebung mußte gesorgt werden, da die Nasallaute, besonders N, von vielen Personen nur schlecht rein flüsternd angegeben werden.

Auf diese Weise war festzustellen, daß von oben herab die erste Schädigung für M mit 2,5 oder etwas vorher, für N und Ng etwas tiefer einsetzt, und daß die drei Laute bei 3,5 nicht mehr sicher unterscheidbar sind. Bei 5 nur noch Stärkeunterschiede, Ng etwas stärker als die anderen, aber alle drei nur ein schwaches charakterloses Blasen oder Hauchen. Der Rest erlischt bei 14, also in der Gegend des  $c^2$ .

Beim Aufbau differenziert sich das leise Geräusch zuerst mit 3,6, wo man den Beginn des M ansetzen kann. Bei 2,8 sind sie ziemlich unterscheidbar, bei 2,4 oder 2 ausgebildet, werden dann wohl noch etwas deutlicher, wenigstens stärker, aber nicht mehr wesentlich verändert. M erfuhr auch bei 1,6 einmal noch eine Besserung. Die charakteristische Zone wäre hiernach gemeinsam für die drei Laute etwa 3,6–2, und die Formantregion etwa  $b^3-f^4$ . Aber recht scharfe Bestimmungen lassen sich auf diesem Wege hier nicht gewinnen.

Einen Schritt weiter führen Beobachtungen über die gehörten Tonhöhen. Daß die 3 Laute gewisse Höhendifferenzen haben, erkennt auch der Ungeübte leicht. Bei wiederholter aufmerksamer Beobachtung in stillen Nachtstunden konnte ich folgende Tonhöhen feststellen: Geflüstertes M =  $des^3-f^3$  (je nach dem dunkleren oder helleren Charakter), N =  $as^3-h^3$ <sup>1)</sup>, Ng =  $c^3-des^3$ . Es muß also für diese Stellen des Tongebietes auch eine Resonanz durch den betreffenden Laut ausgelöst werden. Auf die Strukturbedeutung dieser Strecken kommen wir unten zurück.

Der Gesamtumfang der 3 stimmlosen Nasalkonsonanten erstreckt sich etwa von  $c^2$  bis  $f^4$ .

Ausgiebigere, wenn auch weniger direkte, Aufschlüsse erhält man, wenn man diese Laute als stimmhafte analysiert und sie mit dem gesungenen U, dem sie in dieser Form am nächsten stehen, vergleicht. Sie sind in stimmhafter Form in der Tat nichts anderes als ein in verschiedener

1) Diese Tonhöhen gelten für das zwanglos geflüsterte N. Ich kann hier aber willkürlich die Tonhöhe innerhalb der ganzen Oktave  $des^3-des^4$  stetig oder in der Tonleiter auf- und absteigen lassen. Es ist ein leisestes stimmloses Singen durch die oberen Nasenräume. Die Tonhöhen sind vollkommen klar ausgeprägt. Dasselbe ist auch mit geschlossenem Munde, also mit M-Einstellung, möglich. Dann kommen besonders die hohen Töne,  $c^4$ ,  $des^4$ , sehr deutlich heraus, die mit der Tonhöhe eines geflüsterten E zusammenfallen. Je leiser man in dieser Art flüstert, um so besser.

Weise genäselt<sup>1)</sup> U<sup>1)</sup>. Auf diese Art ist daher zugleich die akustische Natur des Näsels zu studieren.

Man wolle sich hier nicht daran stoßen, daß wir die stimmhaften M, N, Ng als Vokale, die stimmlosen Laute mit gleicher Bezeichnung aber als Konsonanten in Anspruch nehmen. Wie man auch den Unterschied zwischen Konsonanten und Vokalen definieren möge, so läßt sich doch aus der Tatsache der gleichen Buchstabenbezeichnung nicht ohne weiteres ableiten, daß das stimmlose M, N, Ng notwendig derselben Klasse von Gehörerscheinungen zugeteilt werden müsse wie das stimmhafte. Es kann sich auch nur um Verwandtschaft durch Analogie oder durch gleiche Teile handeln. In der Tat bezeichnen wir die stimmlosen M, N, Ng als Konsonanten nur wegen ihrer überwiegend farblosen Natur. Durch diese unterscheiden sie sich von den stimmhaften Lauten von identischer Schreibweise. Gleich aber sind sie ihnen durch die farbigen Elemente.

Ich machte, wie bei den gewöhnlichen Vokalen, zunächst Vorversuche nach der Resonanzmethode mit der Resonanzgabel 200 (*as*) und ihren Multiplis. Der Grundton ergibt für M und Ng kräftige, für N schwächere Resonanz. Der 2. Teilton für Ng gleichfalls kräftige, für M und N schwächere. Der 4. für Ng nur noch schwache, für die beiden anderen minimale. Der 3. und 5. aber gar keine; und so auch die höheren Teiltöne. Dies beweist aber nicht, daß überhaupt keine solchen Teiltöne mehr vorhanden sind; sie könnten nur zu schwach sein, um noch Gabeln zu erregen. Wir werden in der Tat solche auf anderem Wege finden.

Mit U verglichen, zeigen die Nasallaute aber hier schon einen Unterschied: jenes bringt bei tiefen Grundtönen die Gabeln auch für den 3. Teilton und die in der Mitte der 2-gestrichenen Oktave liegenden Teiltöne (seinen Oberformanten) in Mitschwingung.

Auch die Interferenzmethode läßt sich hier gut verwenden. Verfährt man nach den für die Isolierung und Stärkeschätzung oben gegebenen Anweisungen<sup>2)</sup>, so ergibt sich folgendes: Das freie gesungene U ist gegenüber den auf gleicher Höhe nasalisiert gesungenen U-ähnlichen Lauten an tiefen Teiltönen reicher und führt außer diesen einen in der Höhe gleichbleibenden Oberformanten in der Gegend des  $g^2$  mit sich<sup>3)</sup>. Dagegen kommt

<sup>1)</sup> Ng kann man bekanntlich stimmhaft auch auf anderen Vokalen angeben, besonders gut auf O, AO, Ö, Ä (= franz. on, en, un, in). Wenn es aber „vokalfrei“ oder ohne nähere Angabe des Vokals gewünscht wird, wird jedermann das U-haltige Ng angeben. M und N können in tiefer und mittlerer Lage überhaupt nur so angegeben werden.

In hoher Stimmlage, von  $c^2$  an, wo U immer undeutlicher wird, bleiben diese Laute wohl noch eine Strecke weiter hinauf kenntlich, aber ihre Unterschiede voneinander und von einfachen Tönen werden doch auch immer kleiner, weil ihre charakteristische Teiltonstruktur sich verliert.

<sup>2)</sup> Näheres über deren Durchführung in diesem besonderen Fall in der Abhandlung 12, S. 169ff.

<sup>3)</sup> Man kann auch ein ungewöhnlich stark näselsndes M herstellen; es nähert sich dann vokalisiert mehr dem O und erhält eine gewisse metallische Färbung. Der Laut erscheint im oberen Teil der Nase lokalisiert. Ng kann auf diese Art zu einem förmlichen Schnarren verstärkt werden. Zweifellos werden dabei die hohen charakteristischen Obertöne zahlreicher und stärker.

bei den Nasalen eine dem reinen U fehlende (bzw. darin nur in der Nähe und minimal vertretene) Schär hoher schwacher Beitoné hinzu, die für Ng mit  $c^3$ , für M und N etwas höher beginnt und sich bis in die Mitte der 4-gestrichenen Oktave erstreckt<sup>1)</sup>.

Bestätigend möge erwähnt sein, daß ich auch beim Fagott, dem typisch näseldnen Instrument, den Teilton  $d^4$  als hauptsächlich beteiligt feststellen konnte (15. Kap.).

Nach diesen Ermittlungen an den stimmhaften M, N, Ng können wir nun auch den an ihren stimmlosen Vettern gemachten Beobachtungen eine bestimmtere Deutung geben. Denn diese haben eben doch die Nasalisierung mit jenen gemein, werden überhaupt als die gleichen Sprachzeichen durchs Gehör wiedererkannt, und so kann man nicht zweifeln, daß es sich in beiden Fällen um eine großenteils analoge Struktur handelt, nur daß im einen Fall eine diskrete Zahl von Teiltönen, im anderen Fall jene stetige, wenn auch mehr oder minder dichte oder intensive Ausfüllung der Tonlinie vorliegt, wie sie für Geräusche charakteristisch ist. Die durch Interferenzversuche an den stimmlosen Nasalkonsonanten gefundene Formantstrecke  $b^3-f^4$  stimmt in der Tat gut überein mit dem spezifischen Näselformanten für das feine, dünne Näseln, wie er durch die Teiltöne von  $c^4-g^4$  bei den gesungenen Nasallauten gegeben ist. Die beobachteten Tonhöhen der stimmlosen aber (s. o. S. 117) weisen darauf hin, daß der Gesamtformant doch auch für sie bereits in der 3-gestrichenen Oktave beginnt, und zwar mit den bezüglichen Tonhöhen. Dies entspricht einer Regel, die sich für die Tonhöhen der meisten Konsonanten zu bestätigen scheint. Daß bei Interferenzversuchen diese untere Formantstrecke sich nicht geltend macht, daß die Laute beim Abbau schon mit  $b^3$  unkenntlich sind und beim Aufbau erst mit  $b^3$  spurweise auftreten, kann man aus ihrer Schwäche wohl verstehen.

Man kann übrigens das stimmlose Ng auch so angeben, daß der tiefere Bestandteil deutlich gesondert und kräftig herauskommt. Er trägt den Vokalcharakter des geflüsterten AO und hat, wie dieses, die Höhe  $c^3$ . Die Mundöffnung ist auch dieselbe wie beim Aussprechen des AO. In meinen Interferenzversuchen wurde es aber nicht in dieser etwas unnatürlichen Weise, sondern möglichst hell, dünn und tonfarblos, als möglichst reiner Nasalkonsonant angegeben.

Versuchen wir nun, aus allen diesen Einzelheiten ein zusammenfassendes Bild der drei Laute zu entwerfen. Für die stimmlosen Laute M, N, Ng liegt der Formant in der 1. Hälfte der 4-gestrichenen Oktave. Außerdem haben sie einen Unterformanten. Der des Ng beginnt mit etwa  $c^3$ , der des M mit  $des^3$ , der des N mit  $as^3$ . Diese Unterformanten sind bei Ng

<sup>1)</sup> Nur in diesem Umstande kann ich eine gewisse Erklärung dafür finden, daß W. KOEHLER, aber auch schon HELMHOLTZ, vom M den Eindruck hatten, als läge sein Zentrum auf der Tonlinie unter dem des U. Das Fehlen des „Oberformanten“ des U ( $f^2-g^2$ ) läßt M milder, dumpfer erscheinen. Aber wie sollte es einen unter  $c^1$  liegenden Formanten haben, wenn es auf  $c^1$  oder  $g^1$  gesungen wird? Man kann nur zugeben, daß das stimmhafte M, d. h. die Stimme bei geschlossenem Munde, relativ dunkler ist als ein auf der nämlichen Tonhöhe gesungenes U; aber nicht, daß es absolut dunkler oder tiefer wäre als U. Selbst in der 2-gestrichenen Oktave, wo U mit steigendem Grundton immer unvollkommener wird, kann M noch deutlich sein, reicht also sogar höher hinauf.



und M von den Formanten durch eine relativ leere Strecke getrennt, die von Ng zu M kleiner wird und bei N verschwindet (vgl. das Schema S. 107).

Bei den gesungenen Lauten M, N, Ng sind gleichfalls die 3- und die 4-gestrichene Oktave entscheidend für ihren Charakter. Die Obertöne in der 4-gestrichenen, die hier bis etwa  $as^4$  reichen, sind auch bei ihnen speziell für die feinen, dünnen Elemente des Näsels verantwortlich. Außerdem trägt hier die besondere Stärke des 1. und 2. Teiltones gegenüber der Schwäche oder dem Ausfall der ihnen zunächst folgenden zu dem Gesamtcharakter bei.

Wir verweisen hier auch auf die oben (S. 73) erwähnten Lückenversuche mit einem auf *ges* und auf *c* gesungenen A, aus denen hervorgeht, daß es für die nasale Wirkung günstig ist, wenn die Teiltöne der 3-gestrichenen Oktave durch eine leere oder nur schwach ausgefüllte Strecke von den tieferen getrennt sind.

Nach alledem haben wir uns aber noch zu erinnern, daß diese Nasallaute im gewöhnlichen Gebrauch unseres europäischen stimmhaften Sprechens und Singens wohl niemals als Selbstlaute, sondern nur als Einleitung oder Abschluß eines Vokals gebraucht werden, und daß gerade der Übergang zwischen ihnen und dem Vokal ihnen ein besonders charakteristisches Gepräge gibt. Man kann bekanntlich, wenn man auf irgendeinem Wege künstlich einen A-ähnlichen Klang erzeugt hat, leicht ein „Mama“ daraus machen, indem man die Öffnung, aus der der Klang kommt, mit dem Finger oder dem Daumenballen abwechselnd schließt und öffnet. Auch bellende Hunde kann man so zum Sprechen bringen. Feinere Unterschiede zwischen M, N, Ng, auch W, sind allerdings kaum so herzustellen, und die genauere Definition des Verhaltens, das durch diese Prozedur dem Vokallaut in Hinsicht seines Ab- und Anschwellens auferlegt wird, dürfte auf Schwierigkeiten stoßen.

Beim Singen mit Brummstimmenbegleitung, wie es in den Männerchören der „Liedertafeln“ zuweilen ausgeübt wird, haben wir allerdings auch ein selbständiges Auftreten des gesungenen M, d. h. eines in dieser Weise schwach nasalisierten U. Auch der gregorianische Choral kannte eine eigentümliche Verzierungsform unter dem Namen der „Liquescenz“, bei welcher Liquida und Nasalkonsonanten als selbständige Träger von Tönen gebraucht wurden. Ähnliches findet sich in exotischen Gesängen und Sprachen<sup>1)</sup>.

### L<sup>2)</sup>.

L wird jetzt zumeist als „Halbvokal“ bezeichnet. GUTZMANN geht noch weiter: „Unser deutsches L gehört überhaupt nicht zu den Konsonanten seinem Klange nach . . . Wir hören bei dem L in der Tat keine Spur von irgendwelchem Reibegeräusch und keine Spur von irgendwelchem Geräusch überhaupt.“ Immerhin — es gibt doch auch ein geflüster-tes L, und dieses ist unleugbar ein Geräusch. Aber es ist ein stark farbiges Geräusch, und zwar von analogem Farbcharakter wie das gesungene und

<sup>1)</sup> Vgl. R. LACH: Über die Gesänge russischer Kriegsgefangener. Sitz.-Ber. d. Wiener Akad. 1917, philos.-histor. Kl., Bd. 183, 4. Abhandl., S. 42.

<sup>2)</sup> Hier sei noch besonders daran erinnert, daß nur von der deutschen Aussprache die Rede ist. SIEVERS hält 5 verschiedene L auseinander, unter denen eines das deutsche ist. Doch dürften die allgemeineren Thesen des Folgenden auch von den übrigen Formen gelten; nur die akustischen Konstanten werden kleine Unterschiede zeigen.

gesprochene L, deren Vokalnatur ich nicht bezweifle. Subsumiert man daher stimmlose Sprachlaute (Sprachgeräusche) mit stark ausgeprägtem Farbcharakter unter die Vokale, so ist in der Tat auch das stimmlose L als Vokal zu bezeichnen.

Da das stimmlose L nur schwach und nicht ganz ungeändert durch die Röhrenleitung kommt, habe ich nach einigen Vorversuchen hier sogleich das stimmhafte L herangezogen, um erst von da aus auf das stimmlose zurückzukommen. Hierbei stellte sich bald heraus, daß wir in diesem interessanten Laut einen leicht näselsnden Vokal, und zwar wesentlich wieder ein näselsndes U vor uns haben. Man kann es auch stärker näselsn machen, indem man es etwas schnarrend spricht oder singt. Schon bei dem gewöhnlichen L nähert sich ja auch die Erzeugungsweise der des N. Durch die Anlegung der Zunge an die Zähne wird ein Teil der Leitung in die Nase verlegt. Aber wir stützen uns hier nicht auf genetische, sondern auf rein akustische Beobachtungen.

Schon die direkte Beobachtung mit freiem Ohr lehrt, und zwar für ein geübtes Ohr auch bei dem gewöhnlichen, nicht forcierten L, daß außer den ersten Obertönen, besonders der Duodezime, höhere Teiltöne vorhanden sind. Näher konnte ich 2—3 unterscheiden: einen in der 3-gestrichenen Oktave zwischen  $e^3$  und  $h^3$  (je nach dem Grundton) und einen oder mehrere in der unteren Hälfte der 4-gestrichenen Oktave. Der aus der 3-gestrichenen bringt etwas Ü-artiges in den Laut, die höheren bedingen seinen näselsnden Charakter.

Auch hier glaube ich aus den Beobachtungen schließen zu dürfen, daß es sich nur um harmonische Teiltöne handelt. Singe ich z. B. ein L sukzessive auf *cis*<sup>1</sup>, *h*, *gis*, *e*, *c*, *H*, so höre ich in derselben Aufeinanderfolge *gis*<sup>3</sup>, *h*<sup>3</sup>, *gis*<sup>3</sup>, *e*<sup>3</sup>, *fis*<sup>3</sup> (11. Teilton des *c*), *fis*<sup>3</sup> (12. des *H*); also immer einen harmonischen Teilton, der sich aber in seiner Ordnungszahl der Bedingung fügt, in der bestimmten absoluten Tonregion zwischen  $e^3$  und  $h^3$  zu liegen, wie es Formantenart ist. Bei dem Grundton *e* glaubte ich auch noch *gis*<sup>3</sup> und  $h^3$  oder einen dieser beiden zu hören. Ferner hörte ich bei *cis*<sup>1</sup> und *h* die Töne *cis*<sup>4</sup> bzw. *dis*<sup>4</sup>. Der Beiton ging hinauf, während der Grundton abwärts ging: wiederum zur Wahrung der absoluten Lage. Doch sind diese Beobachtungen schwierig und nicht ganz sicher. Auch bleibt zu beachten, daß man bei Teiltönen von hoher Ordnungszahl nicht so bestimmt sagen kann, ob ihre Stimmung nicht doch ein wenig von der der harmonischen Teiltöne abweicht; ein unharmonischer würde sich von den umliegenden harmonischen nur äußerst wenig unterscheiden.

Genauerer lehrt uns wieder die Interferenzmethode, die auf das gesungene L ganz so wie auf jeden Vokal angewandt werden kann. Ein auf *c* und  $c^1$ , sowie auf der dazwischenliegenden Tonhöhe von 200 Schw. gesungenes L wurde auf diesem Wege ab- und aufgebaut. Es geht beim Aufbau aus einem anfänglichen U zunächst in ein Ou über, das immer markiger wird. Die entscheidendste Region liegt zwischen Einstellung 4,8 und 3,7, d. h. zwischen  $e^3$  und  $b^3$ <sup>1</sup>). Hier beim  $b^3$  angelangt, ist es schon ein gutes L, wird aber noch immer wesentlich besser bis 2,2, d. h. bis  $f^4$ . Wir müssen daher, wie beim S, eine engere und eine weitere Formantzone statuieren; zur letzten gehört noch die untere Hälfte der 4-gestr. Oktave.

<sup>1</sup>) Betreffs der Höhenbestimmungen ist immer die Bemerkung S. 113, Anm. 2 zu beachten.

Man hat bei diesen Ab- und Aufbauversuchen den Eindruck, daß die harmonischen Teiltöne des gesungenen L ziemlich lückenlos vertreten sein müssen, weil fast jedes Hineinschieben einer Röhre einen Unterschied macht. Dies bestätigt sich durch das Folgende.

Ich versuchte, auf dem oben S. 38ff. angegebenen Wege auch hier die einzelnen Teiltöne nach Möglichkeit zu isolieren und ihre relative Stärke zu bestimmen. Hierbei stellte sich heraus, daß die 4 tiefsten mit mäßiger Stärke vorhanden sind, der 2. stärker als die anderen, daß dann eine Abschwächung eintritt, einige (der 7. und 9.) auch wohl gar nicht vorhanden sind, daß aber in der 3-gestrichenen Oktave und der 1. Hälfte der 4-gestrichenen noch immer weitere Töne von merklicher Stärke hinzutreten. Wenn man die untere Abteilung des Lautes bis etwa zur Hälfte der 3-gestrichenen Oktave ausschließt (außer soweit ungerade Multipla mit ausgeschlossen würden), so vernimmt man ebenso wie bei den stimmhaften M, N, Ng die Gesamtheit der oberen Bestandteile, die Formantgegend, als ein Wispern, das offenbar (im Zusammenhang mit den tieferen Bestandteilen) dem Ganzen den leicht näselnden Charakter verleiht. Und man kann in diesem Wispern auch die einzelnen Bestandteile bis zu einem gewissen Grade isolieren. So habe ich alle Töne von  $c^4$  bis  $g^4$  gelegentlich für sich hören können.

Die Lage dieses nasalen Elements erweist sich hierbei als unabhängig von der Höhe des Grundtons. Es liegt also wieder ein Nasalformant im strengen Wortsinne vor, und zwar derselbe, den wir bei den nasalen Konsonanten gefunden haben. Das gesungene L ist akustisch ein näselnder Vokal.

Diese Ergebnisse wurden durch Versuche mit Resonanzgabeln bestätigt. Das auf  $c^1$  gesungene L ergab für die aufeinanderfolgenden 6 ersten Teiltöne folgende subjektive Stärken: 1, 2,  $1\frac{1}{2}$ , 1,  $\frac{1}{2}$ ,  $1\frac{1}{2}$ , dann für den 7. ( $b^3$ ) bis 12. ( $g^4$ ) je  $\frac{1}{4}$  bis  $\frac{1}{8}$ . Bei dem auf demselben Grundton gesungenen U ist der 3. Teilton  $g^2$  viel stärker, dagegen alle höheren schwach oder = 0. Es ist also sogar auf diesem einfachen Wege die Existenz der nasalen Formantregion beim L nachweisbar. Eine Senkung der L-Kurve zwischen den unteren Teiltönen und dieser Region läßt sich allerdings auf diesem Wege nicht feststellen; dazu ist die Reaktion der hohen Gabeln schon an sich zu schwach. Aber daß Beimischungen bis zu  $g^4$  vorhanden sind und den Unterschied gegen U wesentlich mitbedingen, geht klar daraus hervor.

Nach diesen Ergebnissen versuchte ich nun doch auch das geflüsterte L noch auf dem Interferenzwege zu analysieren. Hier ist besonders wichtig, sich vorher einer guten Aussprache zu versichern, da L zu den Lauten gehört, die vielfach nur schlecht geflüstert werden (dieser etwa nach dem palatalen Ch hin), ohne daß man es im Zusammenhange bemerkt. Möglichst kurze Leitung ist natürlich vonnöten, doch kann sie durch eine Wand geführt sein. Der Laut wird beim Abbau durch die Einstellung 1,8 (entsprechend dem  $as^4$ ) zuerst leicht gedämpft. Etwas Dunkles tritt darin auf. Bei 3,2 (entsprechend dem  $c^4$ ) ist dieses Dunkle herrschend geworden, das Helle ganz geschwunden und damit der L-Charakter zerstört. Bei 4 ist nur ein dunkles Blasen zu hören, das bei 8,5 schon fast unhörbar wird. Bei 13 nur noch ein minimaler Rest für scharfe Ohren vorhanden.

Beim Aufbau tritt die erste Spur des beginnenden L-Charakters mit 3,2 auf. Es erscheint da ganz plötzlich ein hohes Element. Bei 2,0 ein deutliches L. Diese Grenzen sind beiderseits bestimmt markiert und immer

wieder zu beobachten. Hiernach kommt man auf den Formanten  $e^4-f^4$  ( $g^4$ ), also auf denselben Näselformanten wie bei den vorher untersuchten Nasalkonsonanten.

Nun beachte man aber auch hier die Tonhöhe des Flüster-L, die besonders deutlich ist. Sie ist  $as^3-b^3$ , je nach der Helligkeit, mit der es gegeben wird. Daraus läßt sich schließen, daß auch beim L unter dem Näselformanten ein Unterformant liegt, der dicht unter der angegebenen Tonhöhe, also bei etwa  $g^3$ , beginnt. Er geht nach oben hin, wie beim N, unmittelbar in den Näselformanten über, kann daher ebensogut auch als untere Abteilung des Formanten bezeichnet werden, den man dann von  $g^3$  bis  $g^4$  zu rechnen hat. Das gesungene L ist beim Aufbau mit Interferenzröhren sogar schon mit dieser Unterformantregion, mit Erreichung des  $h^3$ , gut erkennbar; hier ist die entscheidende Gegend. Aber vollkommen wird es doch auch erst durch die darauffolgenden Teiltöne<sup>1)</sup>.

Außer diesem Unterformanten kann man aber im geflüsterten L ziemlich leicht noch einen zweiten, tieferen Unterformanten hören. Seine Tonhöhe schätzte ich bei meinem eigenen L zunächst auf  $f^2-as^2$ . Zur genaueren Bestimmung der Höhe läßt sich das Interferenzverfahren heranziehen. Baut man den Laut damit auf, so erscheinen nacheinander drei verschiedene Tonhöhen (Geräuschhöhen). Das minimale, nur eben hörbare dunkle Untergeräusch, mit dem der Laut überhaupt beginnt, scheint die Höhe  $es^2$  zu haben. Dann taucht, wenn man mit dem Hineinschieben der Röhren bis zur Einstellung 7 gelangt ist, noch eine neue Tonhöhe auf,  $as^2$ , und man hört zunächst eine simultane Geräuschquarte. Diese beiden Tonhöhen scheinen mir den Beginn und Schluß des tiefen Unterformanten zu bezeichnen, der freilich seiner Schwäche wegen nur eine sekundäre Bedeutung für den Charakter des Gesamtlautes haben kann. Weiterhin, bei den Einstellungen 6 und 4, wird  $es^2$  schwächer,  $as^2$  stärker. Bei 3 aber (entsprechend der Tonhöhe  $des^4$ ) tritt plötzlich das hohe Element des Lautes auf; auch schon bei 3,5 (entsprechend der Tonhöhe  $b^3$ ) ist es eben merklich. Und zwar hat es die konstant bleibende Tonhöhe  $b^3$ , die den Beginn der höheren Geräuschabteilung bezeichnet. Einen weiteren Wechsel der Tonhöhe, etwa zu Beginn des Näselformanten, habe ich nicht beobachten können.

So lassen sich alle Beobachtungen zu einem einheitlichen Strukturbilde zusammenfassen.

Zugleich zeigt der Laut die stärksten Analogien zu den vorher betrachteten Nasalkonsonanten. Wenn wir ihn dennoch vom rein akustischen Standpunkte als näselsnden Vokal bezeichnen, auch in geflüstertem Zustande, so sind dafür zwei Gründe maßgebend. Erstens tritt die Flüsterhöhe  $a^3$  bis  $b^3$  so klar hervor wie bei den Flüstervokalen und steht der Laut dem geflüsterten Ü, das dieselbe Höhe hat, überhaupt sehr nahe, wäre fast als näselsndes Flüster-Ü zu bezeichnen<sup>2)</sup>. Zweitens hat er einen deutlichen tiefen Unterformanten, ebenso wie die hellen Vokale; und sowohl der Formant wie der tiefe Unterformant liegen auch im allgemeinen

<sup>1)</sup> Zu dieser und der folgenden Ausführung vgl. das graphische Schema S. 107.

<sup>2)</sup> Bereits WENDELER und BREMER haben L als Vokal in Anspruch genommen. Auch ROUSSELOT fand (II, S. 208), daß das deutsche L dem Ä am nächsten stehe (dessen Flüsterhöhe mit der des Ü fast ganz übereinstimmt).

in der nämlichen Tongegend. Infolge dieser Eigenschaften nähert sich tatsächlich das stimmlose L im Gesamteindruck den stimmlosen Vokalen ebenso wie das stimmhafte L den stimmhaften Vokalen, und wird daher am natürlichsten zu ihnen gerechnet. Aber aller Wortstreit ist vom Übel, und der Fall zeigt wieder deutlich, daß die Einteilung in Vokale und Konsonanten, wenn auch begrifflich scharf, doch in Wirklichkeit nur fließend ist. Überdies gilt das Gesagte nur eben akustisch. Funktionell, nach seiner Verwendung innerhalb des sprachlichen Zusammenhanges, wird man dem L immer die Rolle eines Konsonanten zuerkennen müssen, wenn es als Voroder Nachlaut eines Vokals oder als Verbindungsglied zweier Vokale dient. Aber selbst funktionell übernimmt es in slawischen Sprachen (ebenso wie R) unzweifelhaft vokalische Funktion, wenn es von echten Konsonanten umgeben ist, wie im tschechischen „vlk“.

Bemerkenswert ist auch, daß HERMANN'S L-Kurven niemals den Einfluß eines vorhergehenden oder nachfolgenden Vokals aufweisen, wie ein solcher bei Konsonanten die Regel ist. Vgl. GUTZMANN 5, S. 126.

### S (Ss).

Obgleich fast alle meine Beobachtungen durch jüngere Ohren kontrolliert wurden, trug ich doch beim S peinlichste Sorge dafür, daß 5 junge Beobachter in besonderen Versuchsreihen eine solche Kontrolle übernahmen. Es ergab sich aber eine Differenz auch nur an der äußersten Spitze des Lautes.

S wird beim Abbau von 0,5 an (also mit Rücksicht auf die Interferenzbreite etwa von  $c^6$  an) mit Erweiterung der Einstellungszone immer stumpfer und schwächer. Bei 1,5 ist es schon dem Sch ähnlich, bei 2 wie ein etwas dünnes F. Weiterhin eine Strecke lang F-artig blasend, dann leise hauchend oder keuchend. Bei 7 sehr schwach, bei 16 minimaler Hauch.

Beim Aufbau zeigen sich schon bei 3 die ersten Anfänge eines, wenn auch noch sehr stumpfen, S-Charakters. Der Laut ist etwas streichend und dünner als Sch bei derselben Einstellung. Bei 2,6 und 2,2 wird schon ein schlechtes, stark lispelndes S daraus, bei 1,4 ist es gut, wird sogar mehrmals als „sehr gut, sehr schön“ bezeichnet, wird aber doch noch bis 0,5 oder 0,4 (für mich nur bis 0,6) immer besser. Fast jede Stufe dieser letzten Etappe pflegt der Beobachter mit „gut, fertig“ zu bewerten, und doch erscheint ihm dann die nachfolgende noch besser. Dies kommt wohl daher, daß man das allerschärfste Ss, das mit einem gewissen Elan herausgestoßen werden muß, im Leben relativ selten zu hören bekommt und so einen etwas dehnbaren Maßstab mitbringt. Man beurteilt daher als „gutes, scharfes S“ eines, das noch recht wohl eine Verschärfung zuläßt.

Auffällig ist hier wieder der Unterschied des Eindrucks bei ein und derselben Röhrenstellung, je nachdem man sich im Ab- oder Aufbau befindet (S. 51). Dasselbe, was beim Abbau bereits als „lispelndes S“ erscheint, gilt beim Aufbau schon als „gutes S“. Ich habe darum in der Formulierung der Ergebnisse hier zwar den Aufbau zugrunde gelegt, aber diesen Unterschied doch mit berücksichtigt, namentlich in der Bestimmung der obersten Abteilungen und der oberen Formantgrenze.

Die Hauptentwicklung liegt hiernach zwischen 3 und 1,4, woraus sich als engerer Formant  $des^4-c^5$  ergibt. Man hätte eine höhere Lage dafür erwarten können, da die aus dem S gelegentlich heraushörbaren Töne für die direkte Beobachtung nach ihrer Spitzigkeit der 6-gestr. Oktave anzugehören scheinen. Aber die Versuche, die bei diesem wichtigen Laut in

besonders großer Zahl und mit vielen verschiedenen Beobachtern angestellt wurden, sind immer wieder in dem genannten Sinn ausgefallen. Verbessert wird der Laut zwar immer noch bis zum Beginn der 6-gestr. Oktave. Aber die entscheidendste Gegend liegt zweifellos vor und bei  $c^5$ .

Die Angaben Dr. v. ALLESCHS wichen bei diesem Laut insofern etwas von denen der übrigen Beobachter ab, als er den Beginn beim Aufbau, die „erste Spur von etwas Scharfem“ manchmal schon bei 5, ja bei 6,5, andererseits an der oberen Grenze auch noch eine Art sprunghaften Überganges zur schärfsten Form bei 1,2 fand. Vielleicht hängt dies mit einer außergewöhnlichen Feinhörigkeit dieses (übrigens unmusikalischen) Beobachters für Helligkeitszuwüchse zusammen. Aber selbst 1,2 entspricht mit Berücksichtigung der Interferenzbreite erst dem  $d^5$ . Hier ist das S auch für v. ALLESCH schon ausgezeichnet gut.

Nach W. KOEHLER ist der Ton  $c^6$  (= 8400 Schw.) das „reine, optimale S“ (1, III, S. 26ff.). Dies könnte immerhin in dem Sinne richtig sein, daß unter den einfachen Tönen dieser die relativ größte Ähnlichkeit mit S hätte. Aber ein natürliches, ohne weiteres überzeugendes S kommt doch nur durch ein Geräusch zustande, in welchem diese Tonhöhe neben vielen anderen wesentlicheren enthalten sein mag<sup>1)</sup>. Auch ließe sich nach dem Obigen nicht festhalten, daß die weitere Umgebung des  $c^6$  nach oben und unten gleichwertig wäre: vielmehr liegen die wirkungsvollsten Bestandteile nach unten, und zwar bedeutend tiefer.

Bei Stichversuchen mit S genügt niemals eine einzelne Einstellung, um den Laut unkenntlich zu machen (obgleich doch immer auch die nächste Umgebung des ausgeschlossenen Tones mit ausgeschlossen bzw. geschwächt wird). Immer bleibt er ein zweifelloses S, geht nie durch Ausschaltung eines einzelnen engbegrenzten Tonbezirks etwa in ein Sch über. Doch wird er durch Stiche von 0,6 und besonders von 1 an stumpfer und dunkler. Bei Einstellungen über 1,8 kann man aber nicht sicher sein, ob nicht die Wirkung auf der Ausschaltung der ungeraden Multipla beruht.

Ich versuchte auch, einzelne Teile des S dadurch zu isolieren, daß nach Einschaltung der Röhren für die gesamte irgend in Betracht kommende Region, also etwa von 0,5 bis 4 (wobei nur ein Sch gehört wurde) eine einzelne Röhre oder wenige nebeneinanderliegende wieder hereingeschoben, also die bezügliche kleine Geräuschstrecke eingeschaltet wurde. Aber es gelang nur dann, eine merkliche Wirkung zu erzielen, wenn die eingeschaltete Strecke schon etwa eine halbe Oktave umfaßte, und in diesem Falle hörte man eben ein stumpfes S, aber nicht einen oder mehrere einzelne hohe Töne, auch wenn es sich um die charakteristische Region handelte.

Dagegen führte einmal der Zufall die Isolierung eines hohen Tones herbei. Ich hatte, um den Laut möglichst kräftig zu bekommen, ausnahmsweise nur weite Röhren (18 mm lichte Weite) genommen. Die Einstellungen reichten von 0,5 bis 5. Also war die Tonstrecke  $f^3-c^6$  ausgeschlossen. Es blieb aber ein einzelner sehr hoher Ton aus der Gegend der 6-gestr.

<sup>1)</sup> KOEHLER selbst hebt S. 33 hervor, daß  $c^6$  für sich allein als Nachahmung des gesprochenen S nicht völlig befriedige; man könne aber die vermißte Nuance nach dem „Geschlossenen, Trüben“ hin dazubringen, wenn man durch 2 Galtonpfeifen zugleich 2 Töne dieser Gegend angebe, die weit genug voneinander lägen, um nicht mehr deutliche Schwabungen zu geben. Unter diesen Umständen sei für ihn selbst gar kein Unterschied mehr. Hier wären doch unwissentliche Versuche wünschenswert.

Oktave zu hören, der scharf begrenzt war, wie ein Ton der Galtonpfeife, und sich auch ohne weiteres seiner Höhe nach mit denen der Pfeife vergleichen ließ. Die Unterschiedsempfindlichkeit ist in dieser Höhe allerdings nicht groß. Aber Prof. v. HORNBOSTEL und ich einigten uns rasch darüber, daß der Ton mit dem  $d^6$  des Galton zusammenfiel. Ein Ton dieser Höhenlage also war im S zweifellos enthalten. Da die Interferenzröhren auf die berechneten (statt auf die korrigierten) Längen eingestellt waren, mochte er nicht mehr in die If.-Breite des höchsten ausgeschlossenen Tones fallen und war zufällig auch durch Multiplawirkung tieferer Einstellungen nicht mit ausgelöscht. Dieser Ton könnte nun wohl besonders stark in dem Gemisch vertreten sein und ihm die letzte Schärfe geben. Aber ich würde nicht einräumen, daß der zischende Charakter, der doch vor allem für das S wesentlich ist, auf dem Vorhandensein eines einzelnen Tones dieser Höhe beruhte.

Diese gegen die 6-gestr. Oktave hin oder schon innerhalb ihrer liegenden Schwingungen werden es auch sein, die eine hohe, dünne Gasflamme beim S zusammenzucken lassen, während sie beim Ch pal. bedeutend weniger, beim Sch fast gar nicht mehr zuckt<sup>1)</sup>.

In einer unter R. WACHSMUTHS Leitung durchgeführten Untersuchung wurde zu speziellen physikalischen Zwecken der Ton 4000 =  $c^5$  des Galtonpfeifchens außerordentlich intensiv angegeben. Daraufhin trat für eine halbe Stunde S-Taubheit ein. KOEHLER erklärt dies (1, III, S. 34) als Ermüdungswirkung durch den 2. Teilton  $c^6$ . Aber vor allem müßte doch Ermüdung für den Grundton eintreten. Diese hätte sich nach KOEHLERS Vokallehre als I-Taubheit geltend machen müssen, wovon aber nichts erwähnt wird. Ist dagegen der Formant des S der oben angegebene ( $des^4-c^5$ ), so haben wir hier eine Stichprobe im wörtlichsten Sinne: der schärfste Teil des Formanten ist herausgestochen. Natürlich wird sich die Schädigung nicht haarscharf nur auf  $c^5$ , sondern auch auf seine nächste Umgebung erstrecken. Wenn sie außerdem, wie KOEHLER vermutet, auch für den 2. Teilton  $c^6$  eintritt, so wird dies die S-zerstörende Wirkung noch begünstigen; aber sie nur auf diesem Wege zu erklären, dürfte schwerfallen.

Das S weist auch noch eine tiefere Beimischung von ausgeprägtem Höhencharakter auf, die nur wenig variiert:  $as^3-b^3$  (man kann sie innerhalb dieser engen Grenzen leicht willkürlich verschieben). Sie dürfte der Resonanzhöhe des durch die Lippen vor den Zähnen gebildeten Hohlraumes entsprechen. Diese Komponente gehört integrierend mit zu dem Gesamteindruck des S, wenigstens wie es aus der Nähe vernommen wird, liegt aber isoliert unter seinem Formanten. Man kann sie als einen *U n t e r f o r m a n t e n* bezeichnen.

Der Gesamtumfang des S ist etwa  $a^1-d^6$ . Innerhalb dieses Tonbezirkes liegt nach unseren Versuchen zuunterst eine Region mit den allen Konsonanten gemeinsamen schwächsten dunklen Geräusch-Elementen, dann eine, die dem S mit den übrigen Zischlauten gemeinsam ist und den eben erwähnten Unterformanten enthält, dann folgt sein Formant  $des^4-c^5$  als die Strecke, die dem Laut sein unterscheidendes Gepräge gibt. Darüber aber noch in der 5- bis 6-gestr. Oktave verschärfende Bestandteile (vielleicht nur eine beschränkte Anzahl diskreter Töne), die zur Formantregion im weiteren Sinne zu rechnen sind.

<sup>1)</sup> LINDNER, R.: Zeitschr. Vox 1914, S. 90. Auf dem Hamburger Phonetikerkongreß 1914 konnte ich mich von diesem schönen Effekt überzeugen.

## F.

Der Laut zeigt im Abbau bei 0,5 merkliche Verdunkelung und Schwächung, bei 1,5 schon Annäherung an W. Er ist aber insofern reiner geworden, als der im F enthaltene S-Bestandteil verschwunden ist (so alle Beobachter). Bei 2 leiser und dunkler. An dieser Grenze sind F, S und palatales Ch nicht mehr sicher voneinander zu unterscheiden. Bei 3 wie ein dumpfes W, doch allenfalls mit gutem Willen noch als F zu bezeichnen. Bei 4 U-artig, ganz dunkel, kein F mehr. Bei 5 dunkles, keuchendes H. Von 6 an immer leiser, dunkler, tiefer, bei 9 ganz dumpfer Hauch, bei 11 nichts mehr.

Aufbau: Bei 14 dunkles Blasen, bei 5 etwas fauchend, bei 3,5 oder 3 erste Anfänge des F, wie mit weiter Lippenöffnung. Bei 2,6 leidliches F. Mit 1,6 kommt das Stoßende hinein. Bei 1,4 oder 1,2 gutes, nicht mehr zu dunkles F, bei 0,8 nur wenig oder nicht mehr anders.

Hiernach ist als wichtigstes Stadium etwa 3—1,2 anzusehen, also der Formant etwa *des<sup>4</sup>—des<sup>5</sup>*, fast derselbe wie für S.

Stichversuche ergeben größere Unterschiede gegenüber der Nullstellung bei 1,5 und 2. Es ist da zwar noch F, nicht W, hat aber etwas Dunkleres. Tiefer hinunter lassen Stichversuche, wie beim S, keine reine Deutung mehr zu.

Das Vorstehende habe ich mehrere Jahre nach den Hauptversuchen noch einmal mit Dr. v. ALLESCH nachgeprüft, da mir die Gleichheit der Formanten für so verschiedene Konsonanten wie S und F allzu paradox erschien. Wir machten mit kürzester Leitung Ab- und Aufbauversuche, bei denen beständig S und F verglichen wurden. Dieses sehr instructive Verfahren (in Lautgebung und Beobachtung wechselten wir ab) lieferte gleichwohl wieder fast genau die früheren Ergebnisse. Beim Aufbau begannen mit Einstellung 3 die vorher nur der Stärke nach verschiedenen Laute (F stärker) sich zu differenzieren. Bei 2,6 wurden die ersten Anfänge beider Laute merklich, bei 2,2 waren beide erkennbar, bei 1,4 beide gut. Aber sie verbesserten sich noch weiterhin, besonders das S. Bemerkenswert fand v. ALLESCH auch diesmal das Auftreten eines zischenenden, S-artigen Elements im F, wenn es in die gemeinsame Formantengegend eintrat.

F hat ebenso wie S eine Komponente mit ausgeprägtem Höhencharakter in der 3-gestr. Oktave, etwa eine Terz unter der des S:  $f^3 - g^3$ . Die beim S darüber gemachten Bemerkungen gelten auch hier.

## Ch palatale (vorderes, wie in Ich, Chi).

Abbau: Bei 0,5 und 1 geringe Schädigung, schwächer, aber ohne Charakterveränderung. Bei 1,5 sehr deutliche Abstumpfung, aber noch Ch. 1,8 schon bedenklich, 2 bereits mehr F-ähnlich. F selbst ist hier erheblich dunkler, voller, Sch noch voller als Ch. Bei 2,5 wie dunkleres F. Bei 3 ebenso, schon gegen W hin.

Aufbau: Der Laut wird nach dem langen Indifferenzstadium bei 3 zunächst F-artig, dann etwas schärfer, setzt mit seinem eigentümlichen schiebenden, gleitenden Charakter bei 2,6 ein und ist schon bei 2 recht gut, bei 1,5 noch besser, bei 1,2 fertig.

Hieraus folgt als entscheidende Einstellungszone 2,6—1,2, als Formant daher *es<sup>4</sup>—des<sup>5</sup>*.

Stichversuche: 0,5 kein deutlicher Unterschied. 0,8 und 1 vielleicht etwas stumpfer. 1,5 und 2 deutlich dunkler, fast „F mit I-Bestandteil“. 2,5 und 3 mehr wie stumpfes S, das Schiebende verschwunden. 3,5 wieder



besser, nur wenig von der Nullstellung verschieden. Größte Schädigung also zwischen 1,5 und 3,  $des^5$  und  $f^4$  ( $e^4$ ), den Grenzen der Formantzone, und besonders bei 2,5 bis 3, d. h. zwischen  $g^4$  und  $f^4$  ( $e^4$ ), am unteren Teile der Formantzone, der demnach (wie bei Sch) hauptsächlich entscheidend sein dürfte.

So läßt sich durch Interferenzversuche in Verbindung mit sonstigen Beobachtungen die Struktur der wichtigsten der gewöhnlich als Konsonanten bezeichneten Laute einigermaßen aufklären. Wie aber bei den Flüstervokalen, so sei auch hier betont, daß man nicht glauben darf, mit diesen Umrissen die ganze Struktur eines solchen Lautes zu durchschauen. Vor allem müssen mit den Formanten die übrigen Teile zusammenwirken. Aber es müssen auch innerhalb der Formanten noch feinere und gleichwohl den Charakter wesentlich mitbestimmende Unterschiede vorhanden sein. Dies ergibt sich schon daraus, daß für manche Konsonanten, z. B. S und F, aber auch vorderes Ch, die Formantengenden fast ganz zusammenfallen. Selbst die der Vokale überdecken sich vielfach mit denen der Konsonanten, z. B. schließt die des I die des vorderen Ch in sich ein. Dem entspricht auch eine gewisse Verwandtschaft der Laute für den unmittelbaren Eindruck. HERMANN sagt sogar: „Der Konsonant Chi scheint mir völlig identisch mit dem geflüsterten Vokal I“ (Bd. 83, S. 25). Das geht freilich zu weit; aber eine starke Verwandtschaft ist unleugbar. Und was S, F und vorderes Ch betrifft, so ist im F und Ch zweifellos etwas S-ähnliches enthalten. Aber es müssen doch eben feinere Unterschiede bestehen.

Auch gelegentliche Aussagen meiner Beobachter, daß sie bei gleichmäßig fortschreitendem Ab- oder Aufbau, manchmal auch noch innerhalb des Formanten, sprunghafte Änderungen beobachteten (z. B. v. ALLESCH beim S), deuten auf solche feinere Strukturunterschiede. Ferner kann ein geübtes Ohr manchmal in einem Konsonanten, wieder selbst innerhalb seiner Formantzone, ein tieferes und ein höheres Element auseinanderhalten, ähnlich wie man Tongemische durch das bloße Ohr analysieren lernt. Nicht selten äußerten sich meine Mitbeobachter unaufgefordert in dieser Richtung, so z. B. der akustisch sehr geübte Herr KREICHGAUER beim S und Sch. Mir selbst schien beispielsweise bei Stichversuchen am vorderen Ch durch die Einstellung 1,5 ein S-artiges, durch 2,5 ein Sch-artiges Element geschädigt zu werden. Man kann ja auch in sonstigen Geräuschen oft mehrere Elemente sondern. Ich vermochte nach gehöriger Übung bei gewissen künstlichen Geräuschen (s. Anhang I dieses Kap.) selbst die Tonlage der Bestandteile annähernd nach dem bloßen Gehörseindruck zu bestimmen.

Schließlich darf nicht übersehen werden, daß die Unterschiede der Konsonanten überhaupt nicht bloß auf ihren im engeren Wortsinne akustischen Bestandteilen beruhen. Die Unterschiede des Einsatzes, des Verlaufes, der Dauer, der Stärke sind ja bekannt und

von jeher in erster Linie zu Klassifikationen benützt worden. Immerhin bleibt die Bestimmung ihrer Lage, Ausdehnung und Verteilung auf der Tonlinie ein theoretisch wie praktisch besonders interessantes Problem.

#### IV. Ergebnisse graphischer Methoden bei Konsonanten.

Bei der großen Bedeutung graphischer Methoden wird es nicht überflüssig sein, daß wir uns mit ihren Ergebnissen in bezug auf die besonders schwierige Konsonantenlehre auseinandersetzen (über die bei Vokalen, auch Flüstervokalen s. 8. Kap.). Dabei muß vor allem in acht behalten werden, daß auf diesem Wege zumeist stimmhafte Konsonanten geprüft wurden, daß also der gesamte Stimmklang einschließlich seiner Obertöne mit in den Kurven enthalten sein muß. Die Bestimmung der als Formanten aufgefaßten Töne geschah teils durch die Fourier-Analyse, teils und meistens durch bloße Auszählung der der Grundperiode überlagerten kleinen Zacken. Aus diesen kann man aber nur entnehmen, daß die Töne im Gesamtklang vorhanden sind und irgendeine Rolle spielen, aber nicht, welche. Die Ergebnisse bedürfen daher sorgfältiger Diskussion. Ja sie erscheinen zunächst nicht einmal ohne weiteres unter sich vereinbar.

Wir geben zuerst eine Zusammenstellung der „Konsonantformanten“ HERMANN'S (Bd. 58 u. 83). Sie liegen zumeist in der 3-gestrichenen Oktave, nur für Ch pal. und W in der 1. Hälfte der 4-gestrichenen. Als Formanten können wir sie freilich nicht anerkennen, schon darum nicht, weil die für den Lautcharakter entscheidendsten Strecken bei Konsonanten noch weniger als bei Vokalen auf eine einzige Tonhöhe beschränkt sind. Sie entsprechen, wie wir im nächsten Kapitel hören werden, zu nicht geringem Teile den Tonhöhen, die man durch das bloße Ohr in den Konsonantengeräuschen wahrnehmen kann, und tragen als regelmäßig vorhandene Bestandteile zum Charakter bei, sind aber nicht die entscheidenden Teile. HERMANN gibt sie übrigens auch nur mit starker Zurückhaltung: „Alle diese Bestimmungen selbstverständlich sehr ungenau.“ In einigen Fällen hat er mit M. GILDEMEISTER auch versucht, einen Konsonanten künstlich nachzubilden und die Höhe an dem künstlichen Laut zu ermitteln.

Für Sch und S hat dann O. WEISS mit seinem „Phonoskop“ (Seifenblasen-, später Goldblattmembran) Kurven aufgenommen. Er fand durch Zackenausählung bei Sch Schwingungen zwischen 300 (*es*<sup>1</sup>) und 4500 (*des*<sup>5</sup>), bei dem scharfen S zwischen 150 (*es*) und mehr als 6000 (*ges*<sup>5</sup>). Die rascheren Frequenzen waren den

langsameren aufgesetzt. Der Gesamtumfang ist hiernach in beiden Fällen von ähnlicher Ausdehnung wie nach der If.-Methode, beim Sch etwa 4, beim S über 5 Oktaven; und die von uns als Formanten angegebenen Strecken liegen in beiden Fällen in der oberen Hälfte dieses Umfanges. Jedenfalls also besteht kein Widerspruch.

Konsonanten × stimmlos.	„Formanten“ nach HERMANN.
B	ähnlich D
× Ch gutt.	$b^2-des^3$ , vielfach auch $e^3-f^3$ beigemischt
× Ch pal.	wenig befriedigend; tiefe und hohe Töne, darunter $d^4, f^4$ . Künstliches Ch = $f^4-g^4$
D	$h^3-c^4$
× F	$f^3-g^3, a^3-c^4$
G	$g^3-a^3$
K	$f^3-fis^3$
L	$f^3-g^3$
M N Ng	$h^3-cis^4$ ; kein sicherer Unterschied der 3 Laute
P	$a-ais$ , einige Male $d^1-e^1$
R ling.	$h^2$ , sowie obere Hälfte der 3-gestr. und Anfang der 4-gestr. Oktave
× S	$gis^3-h^3$ , zuweilen auch etwa $g^4$ . Künstliches S etwa $gis^3$
× Sch	$ais^3-h^4$ ; meist um $h^3$ . Künstliches Sch $b^3-c^4$
T	$fis^3-g^3$
W	$e^4-d^4$ , auch $fis^3-ais^3$

S. GARTEN (1) erhielt mit einer verbesserten Form des Phonoskops für Sch durchschnittlich 3243, für S gleichfalls über 6000 Schw.

F. TRENDELENBURG fand in Kurven, die mit dem „Kondensatormikrophon“ erzeugt wurden, für scharfes Sch als mittlere Frequenz 3500 ( $a^4$ ), für sehr scharfes Sch 4500 ( $des^5$ ), für S die nämliche, für Ch 4000 ( $c^5$ ). Die Werte für Sch liegen bedeutend über den unserigen; aber unser Sch war ein mittleres, während jenes offenbar heller, mehr dem S angenähert war. Die Werte für Ch und S liegen an der oberen Grenze unserer Formanten. Vielleicht hängt dies mit der Ablesung und Zählung bei den abklingenden Wellenzügen dieser Geräusche zusammen, wobei die höchsten Schwingungen am meisten hervortreten mögen.

Für stimmhaftes R ling. fand TRENDELENBURG als „Tonhöhen der Feinstruktur“ Töne zwischen 2700 und 4000 Schw. ( $f^4-c^5$ ). Diese sind wieder ganz erheblich höher als die unserigen; sie würden den If.-Einstellungen 2,9—1,8 entsprechen. Aber bei diesen Einstellungen findet sich immer nur eine eben merkliche Schwächung, keine irgend wesentliche Veränderung oder gar Zerstörung. Ich vermute daher, daß hierbei nur hohe Teiltöne des Stimmklanges, nicht des Konsonanten als solchen festgestellt sind. In der als Beispiel mitgeteilten Gesamtreihe der Teiltonamplituden

für den Grundton 454 (sie wurde nach FOURIER berechnet) kommt jedoch die weitaus größte Amplitude auf den Teilton 1816 =  $ais^3$ , erst die nächstgrößte fällt auf 3178 =  $gis^4$  (II S. 12).  $ais^3$  wäre unsere obere Formantgrenze.

Die stimmhaften Nasalkonsonanten wurden zuerst von KATZENSTEIN (2), dann von GUTZMANN (1) graphisch aufgenommen. KATZENSTEIN ließ A, E, I, O, U frei und nasaliert auf  $g$  singen und analysierte die Kurven bis zum 9. Teilton =  $a^3$ . Dies ist aber erst der Anfang des Nasalformanten. GUTZMANN, dessen eingehende Untersuchung auch über die Geschichte und den damaligen Stand der Anschauungen gut orientiert, hat in Fällen von Rhinolalie Kurven eines nasalierten  $\ddot{U}$  sowie der stimmhaften M, N, Ng aufgenommen und durch Fourier-Analyse deren Teiltöne (in einigen Fällen bis zum 19.) ihrer Stärke nach bestimmt. Er hebt bei dem auf  $c$  gesungenen  $\ddot{U}$  hervor (S. 55), daß die tiefen Teiltöne un- gemein verstärkt seien. Indessen ist der 2. vielmehr sehr geschwächt, und dies ist nicht ohne Bedeutung für den Nasalcharakter (s. o.). Nicht minder auffallend erscheint aber in der Tabelle, daß, während die späteren Teiltöne zunächst geschwächt werden, unter den 4 letzten, die der 4-gestrichenen Oktave angehören, 3 wieder stärker sind als bei dem nichtnasalierten  $\ddot{U}$ . Gerade mit  $c^4$  kehrt sich das Verhältnis plötzlich um. Dies ist bezeichnend für die höhere Näseregion. In beiden Beziehungen also eine schöne Übereinstimmung mit unseren Befunden<sup>1)</sup>.

Aus weiteren Versuchen mit nasalierten U, O, A, E, I schließt GUTZMANN, daß bei Nasalierung hohe Teiltöne zwischen  $e^3$  und  $h^3$  hinzutreten, die er als Eigentöne des suprapalatalen Resonanzraumes betrachtet. Beim E und I fielen dafür die ihnen sonst eigentümlichen Formanten aus dieser Gegend hinweg, so daß ihre Kurven sich viel mehr vereinfachten. Bei M, N, Ng zeige die Kurvenanalyse bedeutende Stärke des Grundtones, derart, daß das Klangbild sich mehr oder weniger dem einer Sinuskurve nähere. Ein charakteristischer Formant sei dabei niemals aufgetreten, da eben der Grundton stärkster Teilton sei. Diese Definition des Formanten würde ich nun allerdings nicht unterschreiben, sondern als Formant nur die für den unterscheidenden Charakter eines Lautes hauptsächlich maßgebende Zone relativer Verstärkung (gegenüber der Umgebung) bezeichnen. In dieser Beziehung aber stimmt die Verstärkung in der 3-gestr. Oktave gut mit unseren Ergebnissen betreffs des unteren Teiles der Formantgegend überein.

<sup>1)</sup> Interessant ist in diesen Tabellen auch die bedeutende Verstärkung des Teiltones  $b^3$  gegenüber seiner Umgebung, sowohl beim nasalierten als beim nichtnasalierten  $\ddot{U}$ . Dieser Ton ist in der Tat das Formantzentrum des  $\ddot{U}$ .

Die Analyse der Kurven, die die Nasales bei Verstopfung der Nase ergaben (GUTZMANN S. 61ff.), läßt eine Verstärkung in der 4-gestr. Oktave im allgemeinen nicht (nur in einzelnen Fällen) erkennen. Aber hier sind die Obertöne überhaupt viel schwächer.

Daß auch bereits GRÜTZNER solche Verstärkung hoher Beitöne als Folge der Nasenresonanz vermutete, ist von GUTZMANN hervorgehoben. Auch HELMHOLTZ sprach von einer „größeren Zahl von Teiltönen“ bei näselsnden Instrumentalklängen. Daß dabei die ungeradzahligen eine ausschlaggebende Rolle spielten, kann ich allerdings dem großen Forscher mit GUTZMANN nicht einräumen.

Neuerdings haben F. TRENDELENBURG und in Amerika J. B. CRANDALL L, M, N, Ng mit den Hilfsmitteln der Radiophonie untersucht. TRENDELENBURGS Befunde stimmen, wie er selbst bemerkt, mit unserem Formantschema recht gut überein. Doch gehen M und N in seinen Kurven noch höher hinauf, bis etwa 4000 ( $c^5$ ), was nach den If.-Versuchen nur für M, und auch da nicht für den Formanten, sondern nur für den Gesamtumfang zutrifft. CRANDALL erzielte ähnliche Ergebnisse. Er findet auch eine bemerkenswerte Ähnlichkeit zwischen L und I und bei allen 4 Lauten eine doppelte Resonanzregion, wodurch sie unter sich und mit den Vokalen vergleichbar („Halbvokale“) werden. Zuletzt untersuchte CRANDALL (3) noch 16 Konsonanten, teils stimmhaft, teils stimmlos, aber immer mit nachfolgendem A verknüpft. Unter den Ergebnissen scheint am bemerkenswertesten, daß die Schwingungszahlen beim S bis 7800 ( $h^5$ ) hinaufreichen. Dies ist auch fast genau die Grenze seines Gesamtumfanges nach den Interferenzversuchen; die Formantgrenze freilich liegt eine Oktave tiefer; die darüberliegenden Frequenzen tragen nichts Wesentliches mehr zu dem Lautcharakter bei.

## V. Übersicht der Veränderungen stimmloser Sprachlaute beim Abbau durch Interferenzröhren.

Wir können nunmehr als Seitenstück zu der Tabelle S. 94 auch die wesentlichsten Veränderungen der stimmlosen Sprache beim Abbau durch If.-Röhren in einer Tabelle zusammenstellen. Doch berücksichtigen wir hier nur die einzelnen Laute, nicht das zusammenhängende Sprechen. Das Verhalten der einzelnen Laute hat besonderes Interesse, weil bei ohrenärztlichen Gehörprüfungen mit geflüsterten Zahlwörtern die darin vorkommenden Vokale und Konsonanten den Unterschied bedingen. Aus dem Zusammenhang als solchem darf dabei nichts erraten werden. Über den Einfluß des Zusammenhanges wären die S. 94ff. gemachten Bemerkungen auch nur zu wiederholen.

Die folgende Tabelle ist mit Querschnittbeobachtungen noch besonders durchgeprüft. Es sind nur die größeren Veränderungen aufgenommen. Bei ohrenärztlichen Prüfungen macht es natürlich auch einen Unterschied, wie genau der Patient das Gehörte auf-

Veränderungen der stimmlosen Sprachlaute beim Abbau durch Interferenzröhren.

Einstellung in cm von 0,5 bis	Herabgesetzte obere Hörgrenze		1. bei den geflüsterten Vokalen I, E, Ü, Ä, Ö, A, O, U	2. bei den stimmlosen Konsonanten Ch pal., F, S, L, N, M, Ng, Sch, H, Ch gutt., K, T, P, R ling.
	für Töne	für leise Geräusche		
0,8	<i>h</i> <sup>5</sup>	<i>g</i> <sup>5</sup>		S abgestumpft.
1,2	<i>fis</i> <sup>5</sup>	<i>d</i> <sup>5</sup>		S stark abgestumpft. Ch etwas stumpfer und dunkler.
1,6	<i>d</i> <sup>5</sup>	<i>b</i> <sup>4</sup>	I und E etwas verdunkelt und geschwächt	S sehr unscharf. F abgestumpft.
2,3	<i>gis</i> <sup>4</sup>	<i>e</i> <sup>4</sup>	I und E etwas heiser und blasend	S und F nicht sicher unterscheidbar. Ch <sub>p</sub> einem stumpfen S ähnlich.
3,3	<i>es</i> <sup>4</sup>	<i>h</i> <sup>3</sup>	I = U. E = Oö dunkel. Ü = Uü. Ä = AOä. Ö = Öo	S, F, Ch <sub>p</sub> ununterscheidbares Hauchen (Blasen). Ch <sub>g</sub> stumpfer. M, N, Ng kaum unterscheidbar.
4,9	<i>gis</i> <sup>3</sup>	<i>f</i> <sup>3</sup>	E = Ou. Ü = leises U. Ä = AO. Ö fast = O. A verdunkelt	Ebenso T und P. L undeutlich, mehr ein Blasen. M, N, Ng, L nur ein dunkles Hauchen. Sch un- kenntlich. K mehr wie T. R ein schwaches intermittierendes Gaumen-R.
7,0	<i>d</i> <sup>3</sup>	<i>h</i> <sup>2</sup>	Ä = Oa. A = Ao	S, F, Ch <sub>p</sub> , Ch <sub>g</sub> , Sch gleichförmiges dunkles Hauchen; nur Stärkeunterschiede. K, T, P ununterscheidbares dunkles Stoßgeräusch. Auch R dunkles, nur schwach intermittie- rendes Geräusch. Sämtliche Konsonanten haben die Tonhöhe <i>b</i> <sup>2</sup> .
9,8	<i>gis</i> <sup>2</sup>	<i>f</i> <sup>2</sup>	I und Ü = ganz leises U. Ä = O. Ö = Ou. A fast = O. O = Ou	R ganz mattes Gurren. Allgemeine Tonhöhe <i>fis</i> <sup>2</sup> .
13,9	<i>es</i> <sup>2</sup>	<i>c</i> <sup>2</sup>	I, E, Ü = U. Ä = Uo. Ö = OU. A = U. O = OU. U schwach	Wie vorher. Alle = äußerst schwaches, dumpfes Geräusch; höchstens noch R erratbar. All- gemeine Tonhöhe <i>c</i> <sup>2</sup> .
19,7	<i>a</i> <sup>1</sup>	<i>fis</i> <sup>1</sup>	A, O, U höchstens minimales dumpfes Geräusch; die übrigen unhörbar	Alle nahezu oder ganz unhörbar.

zufassen und wiederzugeben imstande ist. Aber größere Veränderungen werden sich unabhängig davon geltend machen.

Bei der jeweiligen oberen Hörgrenze ist berücksichtigt, daß die Vernichtung sehr leiser Töne um den Betrag der If.-Breite tiefer hinabreicht, als es der Einstellung entspricht, was auch für die Flüstergeräusche anzunehmen ist. Es sind daher in der 2. Abteilung der 2. Kolumne diejenigen Tonhöhen beige geschrieben, die sich aus den S. 43 aufgestellten Regeln für die If.-Breite ergeben; in der 1. Abteilung aber stehen die den If.-Einstellungen direkt entsprechenden, unreduzierten Höhen; diese geben die oberen Hörgrenzen an, wie sie vom Ohrenarzt mit den gebräuchlichen Instrumenten (der Galtonpfeife, dem Struyckenschen Monochord) ermittelt werden. Die den Noten entsprechenden Schwingungszahlen sind aus der Tabelle S. 47 abzulesen. Im 9. Kap. werden wir die ohrenärztlichen und sonstigen Erfahrungen zum Vergleiche heranziehen.

Für die in einer Horizontalzeile nicht angeführten Laute gilt das in der vorhergehenden Horizontalzeile Gesagte; es sind überall nur die neu hinzukommenden Veränderungen aufgeführt.  $Ch_p = Ch$  palatale,  $Ch_g = Ch$  gutturale.

Verglichen mit der Tabelle für die stimmhafte Sprache zeigt diese dieselben Veränderungen und im allgemeinen auch an denselben Stellen. Doch reagieren die geflüsterten Laute begrifflicherweise noch etwas empfindlicher auf den Abbau. Vielleicht würde es sich lohnen, gemeinschaftliche Querschnitte für beide Gruppen mit genau identischen Einstellungen auszuführen. Ein wesentlicher Unterschied ist natürlich der, daß bei den stimmhaften Sprachlauten nach dem Schwinden aller Unterschiede immer noch der gemeinschaftliche Rest des gesungenen oder gesprochenen Klanges in Gestalt des Grundtons und bei tieferen Grundtönen auch eines oder mehrerer Obertöne übrigbleibt. Dieser Rest trägt in der Hauptsache U-Charakter und verschwindet erst bei entsprechend tieferen Einstellungen.

## Anhang.

### I. Zur Analyse sonstiger Geräusche.

Es scheint, daß, prinzipiell gesprochen, geradezu alle Arten von Geräuschen durch Interferenz ausgelöscht werden können. Wegen der theoretischen Bedeutung dieser Frage berichten wir im folgenden noch über dahin zielende Versuche.

1. Um eine Vermutung KOEHLERS nachzuprüfen, wonach das Geräusch eines Galtonpfeifchens die jetzt allgemein angenommene obere Hörgrenze von ca. 22000 Schw. weit überschritte, wurde ein solches Geräusch seiner Angabe gemäß mit Einstellung des Pfeifchens auf Teilstrich 3,5 und

Maulweite 2 durch komprimierte Luft aus einer Bombe konstant angeblasen. Dabei war kein Ton mehr, sondern nur ein Geräusch zu hören. Wurden dessen tiefere Bestandteile durch If.-Einstellungen ausgeschieden, so blieb ein sehr feiner Bestandteil übrig, der durch den Mitausschluß der ungeraden Multipla nicht betroffen war. Dieser Bestandteil verschwand jedoch gleichfalls, sobald weitere Röhren auf 0,3–0,5 cm eingestellt wurden. Fing man mit den Einstellungen von 0 an und ging in Millimeterabständen vor, so wurde von den Besitzern der schärfsten und jüngsten Ohren, Dr. v. ALLESCH und Dr. BALEY, bei höchster Anspannung der Aufmerksamkeit die erste Schädigung des Geräusches bei 0,3 beobachtet; d. h. das Geräusch reichte höchstens bis etwa  $c^7 = 16500$ , mit Berücksichtigung der If.-Breite aber sogar nur bis etwa  $g^6 = 12400$  Schw.

Auch mit anderen Pfeifenstellungen und Maulweiten sowie mit einem anderen Atmosphärendruck wurde das Geräusch untersucht, aber niemals eine höhere Grenze gefunden.

2. Bereits HERMANN und GILDEMEISTER versuchten ein künstliches S herzustellen, um es bequemer als das gesprochene untersuchen zu können. Ich habe gleichfalls ein möglichst S-artiges, hohes und feines, konstantes Geräusch erzielt, indem ich den Luftstrom des elektrischen Ventilators, der zum Anblasen von Pfeifen gebraucht wird, durch eine Düse mit horizontal stehendem Spalt leitete, hinter welcher zwei kleine Kartons senkrecht übereinander so befestigt waren, daß sie nur einen feinen Schlitz zwischen sich ließen. Dieses Geräusch wurde direkt durch einen Schlauch (nicht erst durch einen Trichter, der es vertieft hätte) aufgefangen und ins Beobachtungszimmer geleitet. Dann wurden durch If.-Einstellungen von 15 bis 3,1 cm seine tieferen Bestandteile ausgelöscht. So blieb nur ein sehr hohes Zischen übrig. Es zeigte sich nun, daß dieses S-artige Zischen gleichfalls auszulöschen war, wenn noch Röhren auf etwa 2 cm eingestellt wurden. Verlängerte oder verkürzte man die Einstellungen bis 2,3 bzw. 1,5, so kam es wieder. Hiernach entsprach seine Höhe etwa 4000 Schw. Dies ist in der Tat die obere Formantgrenze des S.

Genau dasselbe ergab sich, als vor dem Spalt der Düse ein Karton in horizontaler Lage angebracht wurde, so daß der Luftstrom gegen seine Kante schlug. Auch so entstand ein scharfes Geräusch, das durch dieselbe Einstellung vernichtet wurde. Prof. v. HORNPOSTEL, der an diesen Beobachtungen teilnahm, variierte selbst unwissentlich die Einstellungen und fand zunächst als Grenzen, jenseits deren das Geräusch wiederkam, 1,5 und 2,4, also die nämlichen wie im vorigen Falle. Aber bei genauester Beobachtung zeigte sich doch noch ein ganz schwacher hoher Bestandteil, gewissermaßen ein Oberton, in dem Geräusch. Wir kamen beide zu dem Ergebnis, daß der Hauptteil etwa 5000, der höhere aber etwa 10000 Schw. entspreche. Das Geräusch hatte nach unserer gemeinschaftlichen Ansicht etwas von S und etwas von F.

3. Auch das Uhrtickern und ähnliche hohe Geräusche, wie das Knistern eines Quecksilberkontaktes oder das Knipsen mit den Fingernägeln, kann man leicht ihrer Höhe nach auf diesem Wege bestimmen. Das kräftige Ticken einer Glashütter Taschenuhr wurde beim Abbau zuerst bei 1,5 geschwächt und verschwand ganz bei 3. Es füllte also die Tonstrecke  $c^4-h^4$  (unter Berücksichtigung der If.-Breite). Das schwache, hohe Ticken einer kleinen Armbanduhr lag fast um eine Oktave höher, es füllte die Strecke  $b^4-b^5$ .

4. Endlich wurden kräftige Knalle, insbesondere der einer Kinderpistole (Luftknallbüchse) untersucht. Auch solche lassen sich durch eine



genügende Anzahl von Seitenröhren vernichten. Zunächst zeigten sich, solange die Einstellungen noch nicht dicht genug waren oder nicht weit genug in die Tiefe reichten, eigentümliche Umwandlungen in ein Klatschen, Klirren u. dgl. Völlige Vernichtung wurde erreicht durch 59 Röhren von den Stempellängen 4 bis 64 cm, wobei die Differenzen benachbarter Einstellungen von 2,5 mm bis zu 4 cm wuchsen. Das Geräusch, wie es aus der Leitung kommt, reicht also in die Tiefe bis etwa zu  $c = 128$  Schw., in die Höhe bis mindestens  $c^4$ . Schiebt man die obersten Röhren bis 5 cm hinein, so wird ein leises hohes Ticken vernehmbar, das durch die Multipla der tieferen Einstellungen nicht mit ausgeschlossen ist. Es müssen also in der Gegend von  $c^4$  noch Geräuschelemente dieses Knalles liegen. Baut man dann nach völliger Vernichtung den Knall durch allmähliches Hereinschieben der Stempel von den längsten an wieder auf, so entsteht zunächst ein dumpfes, leises Pochen, dann wird es immer stärker, besonders zwischen  $c^1$  und  $c^3$ . Hier wächst es zu einem starken Glucksen oder Klatschen, wie wenn ein großer Wassertropfen in einen gefüllten Behälter fiel.

Man kann bei solchen Knallen auch eine Tonhöhe ohne Schwierigkeit, am besten durch Vergleichung mit Klaviertönen, feststellen. Der Pistolenknall lag je nach der Art des Abschießens zwischen  $g^2$  und  $d^3$ . Der Ton entspricht der durch den Stempel und den Pfropfen abgegrenzten Luftsäule im Moment des Herausfahrens des Pfropfens. Baut man nun den Klang mit Interferenz ab, so geht die Tonhöhe herunter und liegt jedesmal einige Töne unter der tiefsten If.-Einstellung, also am oberen Ende der noch vorhandenen Erregungszone. Der Abstieg der Tonhöhe ließ sich bis  $a^1$  verfolgen, dann wurde das Geräusch zu schwach. Wir werden dieselbe Gesetzmäßigkeit bei den Konsonanten im nächsten Kapitel wiederfinden. Doch möchte ich diese Beobachtungen nur als gelegentliche und nicht abschließende anführen. Es ist auch zu beachten, daß sie in einem geschlossenen Raume angestellt sind, wo die Reflexion durch die Wände mitwirkt. Auch in der Röhrenleitung selbst treten Reflexionen auf, und nicht minder kommen, wie S. GARTEN objektiv gezeigt hat (3, VII, S. 24ff.), durch die Seitenröhren Echowirkungen zustande. Die obigen Beobachtungen gelten daher nur für den Knall, wie er aus der Leitung kommt; aber da erscheint er eben doch auch als ein momentanes Geräusch von durchaus echtem Knallcharakter, und ein Nachhall ist nicht wahrnehmbar<sup>1)</sup>.

## II. Die obere Hörgrenze.

1. Mit der Frage nach der Natur der Konsonanten steht in einem gewissen Zusammenhange die nach der oberen Hörgrenze des menschlichen Ohres; wie denn Ohrenärzte seit langer Zeit die pathologische Herabsetzung dieser Grenze an der Perception der Flüsterlaute und Konsonanten zu

<sup>1)</sup> Über die Tonhöhe der leisen Knalle bei den sehr kurzdauernden Geräuschen einer Lochsirene vgl. ABRAHAM, Ann. d. Phys. (4) Bd. 60. 1919. Er fand die Höhe eines „Einlochknalles“ abhängig von seiner Dauer, d. h. der Länge der einzigen hierbei vorhandenen Knallwelle. Es gelang ihm in solchen Fällen nicht, den Knall durch Interferenz ganz auszulöschen. Über Knalle aus einer Schwingung vgl. auch S. GARTEN, daselbst (4) Bd. 48. 1915.

Die vielen Untersuchungen über Geschoßknalle in freier Luft haben noch zu keinen abschließenden Ergebnissen, wohl aber zu der Erkenntnis geführt, daß die Form der Wellen auch hier infolge der Reflexionen in der Atmosphäre und an der Erde nichts weniger als sinusartig ist.

messen pflegen. Es ist jetzt fast allgemein angenommen, daß die Grenze für das normale jugendliche Hörvermögen nicht höher als etwa 22000 Schw. ( $f^7$ ) liegt. Die Behauptung EDELMANNS, des verdienten Herstellers vielgebrauchter Galtonpfeifchen, daß man Töne dieses Instrumentes noch höre, die mit KUNDTschen Staubfiguren = 50000 Schw. bestimmt worden seien, hat sich als irrig erwiesen. Zwar die Eichung des Pfeifchens auf Schwingungszahlen war richtig. Aber was man hörte, waren nicht die so bestimmten Töne, sondern „Schneidentöne“ von etwa 11000 Schw., die bei nicht genügend starkem Anblasen herauskommen<sup>1)</sup>.

Gegen diese Communis opinio hat aber KOEHLER Einsprache erhoben (I, Vorl. Mitt. u. III). Er dehnt die Hörfähigkeit noch um mehrere Oktaven höher aus, was bei der geometrischen Zunahme der Schwingungszahlen viel heißen will. Die These steht in Verbindung mit seiner Lehre, daß die Oktaven des Tones  $c$  ausgezeichnete Punkte für bestimmte Sprachlaute seien, und daß über dem  $I$ , welches sein Zentrum in  $c^5$  habe, noch die Zentren für  $S$ ,  $F$  und  $Ch$  pal. bei  $c^6$ ,  $c^7$  und  $c^8$  lägen, von denen sich wenigstens die ersten beiden experimentell nachweisen ließen. Es sei demnach „wohl mit einiger Wahrscheinlichkeit zu erwarten, daß nähere Untersuchungen die Hörgrenze zwischen 34000 ( $c^8$ ) und 68000 festlegen würden, falls keine Qualität mehr folgt und das Oktavengesetz hier noch gelten sollte“.

Wir würden hiernach unterscheiden müssen zwischen der Tongrenze, wie sie durch die üblichen Mittel bisher bestimmt wurde, und der Hörgrenze. Die Geräusche würden weit über die Töne hinausreichen. Wie man sieht, zieht KOEHLER die Folgerung nur mit Zurückhaltung. Er erkennt auch nicht, daß  $F$  einen erheblich dunkleren und breiteren Eindruck macht als  $Ch$  pal., das doch eine Oktave tiefer liegen soll, hält es aber für möglich, daß mit zunehmenden Schwingungszahlen die Gehörsempfindungen wieder tiefer, nämlich eben dunkler und breiter werden.

Nun kann man zwar rein a priori diese Möglichkeit wohl kaum leugnen. Bei den Farben findet wirklich eine solche Umkehr statt, indem sie sich mit Erhöhung der Schwingungszahl wieder dem am anderen Ende des Spektrums stehenden Rot nähern. Bekanntlich bildet das im Spektrum nicht vorhandene Purpur den Übergang. Auch die sog. spezifische Helligkeit der Farben läuft nicht einfach parallel ihren Schwingungszahlen, sondern nimmt zuerst zu, dann wieder ab und würde bei Rückkehr des Farbentons zu Rot vermutlich auch wieder zunehmen. Bei den Tönen freilich deutet sonst nichts auf eine solche Umkehr hin und ist die Veränderung der Höhe wie des Volumens mit den Schwingungszahlen eine so gleichmäßige, daß eine solche Annahme doch äußerst schwerfallen würde.

Was aber die Tatsachen selbst betrifft, so stehen die im vorausgehenden mitgeteilten Untersuchungen zu KOEHLERS These in unlösbarem Wider-

<sup>1)</sup> Schon ehe dies von CH. MYERS und HEGENER nachgewiesen wurde, habe ich bei einem Besuch in EDELMANNS Laboratorium die Überzeugung ausgesprochen, daß es sich um tiefere Töne handeln müsse, und damit Anstoß bei dem Fabrikanten erregt, der eine Anzweiflung seiner Messungen darin erblickte. Es bildet sich nach vielfacher Beschäftigung mit so hohen Tönen eine Art absolutes Tonbewußtsein dafür aus, das aber nicht wie bei den tieferen auf die „musikalische Qualität“ gegründet ist (die den höchsten Tönen fehlt), sondern auf den Helligkeits- oder Spitzigkeitsgrad, und das bei genügender Übung wohl nicht mehr als eine halbe Oktave fehlt.

spruch. Weder F noch Ch werden wesentlich beeinträchtigt, wenn man die sämtlichen über  $c^6$  liegenden Schwingungen beseitigt, soweit nur überhaupt Interferenzen wirksam sind. Selbst für S beginnt beim Abbau von oben her erst mit diesem Punkte eine eben merkliche, ganz leise Schädigung; sein Zentrum liegt gleichfalls tiefer. Gerade mit Rücksicht auf jene These sind sehr zahlreiche Versuche mit jüngeren Beobachtern angestellt worden, die immer wieder zu denselben Ergebnissen führten. Daß auch das Blasegeräusch des Galtonpfeifchens nicht über der Tongrenze, sondern tief darunter liegt, ist oben erwähnt. Es darf also meines Erachtens als sicher gelten, daß Geräusche in der Linie der Schwingungszahlen nicht höher hinaufreichen als Töne.

Nur eine Möglichkeit möchte ich nicht ausschließen: daß nämlich subjektive Geräusche sich noch in die 8-gestr. Oktave erstrecken könnten. Ich höre bei nervöser Abspannung regelmäßig ein im Innern des Kopfes (mehr rechts) lokalisiertes, überaus feines, dünnes Zischen, dessen Höhe vielleicht noch über die höchsten objektiven Geräusche hinaus liegen könnte. Aber dieses läßt sich eben nicht durch Interferenz oder eine andere objektive Methode untersuchen.

Eine theoretische Bemerkung M. GILDEMEISTERS sei hier noch kurz besprochen. Er weist darauf hin (1, S. 188; 2, S. 253ff.), daß, physikalisch gesprochen, ein Resonator nicht nur durch Schwingungen von der Höhe seines Eigentones, sondern durch jede beliebige Schwingung, nur eben entsprechend schwächer, erregt werde, daß also unter Voraussetzung der HELMHOLTZschen Theorie des Hörens bei hinreichender Verstärkung des Reizes jeder noch so hohe objektive Ton gehört werden könnte. Das ist gewiß richtig (soweit es das Ohr aushielte); aber es folgt nicht daraus, daß die gehörte Tonhöhe entsprechend steigen müßte. GILDEMEISTER hat dies wohl selbst auch nicht gemeint, aber andere könnten seine Ausführungen in diesem Sinne mißverstehen, und so möchte es nicht überflüssig sein, die Unzulässigkeit einer solchen Folgerung besonders hervorzuheben. Unsere Empfindungen hängen ja nicht direkt von den Reizen ab, sondern von den durch sie im Zentralorgan ausgelösten Prozessen. Wenn nun für diese das Gesetz der „spezifischen Energien“ als gültig angenommen wird, und zwar in der Form, daß die Energien auch innerhalb eines und desselben Sinnes noch weiter spezifiziert sind, ja daß jedem Ton eine besondere spezifische Energie entspricht, so wird die nämliche Faser oder Fasergruppe der Basilmembran, einerlei durch welche objektive Schwingung sie erregt wird, im Gehirn immer dieselbe spezifische Energie, denselben Endprozeß, also auch im Bewußtsein denselben Ton auslösen. Die Zahl der Fasern ist nun aber in allen Fällen eine begrenzte. Wenn also Schwingungen, die weit über die Eigenfrequenz des höchsten Resonators der Schnecke hinausgehen, bei genügender Verstärkung wirklich noch auf das Ohr wirkten, so würden sie gleichwohl nur eben die kleinsten dort vorhandenen Fasern erregen, und diese würden dem Bewußtsein keinen höheren Ton liefern können, als es die ihnen zugeordneten spezifischen Energien in der zentralen Hörsphäre zulassen. Die Hörgrenze in dem Sinne, daß man über den Ton von der Spitzigkeit des  $f^7$  hinaus immer weitere noch spitzere hörte, würde also dabei doch nicht erweitert. Erweitert würde nur die Grenze des Hörbaren, der auf das Ohr noch wirkenden, es erregenden Schwingungsfrequenzen. Man würde mit zunehmender Reizstärke 50 000, 100 000 Schwingungen hören, aber sie würden als der nämliche unveränderliche Ton, etwa als  $f^7$ , gehört werden. Freilich

kommt es auch darauf an, ob man das Gesetz der spezifischen Energien überhaupt und ob man es speziell auch für die Eigenschaft der „Spitzigkeit“ und noch spezieller für deren äußerste Grenzen zu gibt.

2. Man hat vielfach das Sinken der oberen Hörgrenze mit dem Alter untersucht. MELDES Klangplatten, KOENIGS Klangstäbe und besonders die Galtonpfeife, nach zuverlässigen Methoden geeicht und richtig gehandhabt<sup>1)</sup>, sowie STRUYCKENS Monochord, neuerdings endlich elektrische Schwingungen bieten sichere Bestimmungsmittel.

Auf dem letzteren Wege hat GILDEMEISTER (a. a. O.) die Hörgrenze bei 51 Personen von  $6\frac{1}{2}$  bis  $47\frac{1}{2}$  Jahren, auch bei einigen höheren Alters, bestimmt. Es ergab sich eine Kurve, die bis zum 47. Jahre von 20 000 Schw. ( $es^7$ ) bis etwa 13 000 ( $gis^6$ ) herabsinkt, allerdings mit erheblicher Streuung, d. h. beträchtlichen individuellen Unterschieden. Ein 77er hörte nur bis 9000 ( $d^6$ ).

In diese Kurve ordnen sich Beobachtungen gut ein, die ich 1899 gelegentlich mit der sicher richtig bestimmten KOENIGSchen Gabel  $c^7$  und mit MELDES Stimmlatte für denselben Ton anstellte. Der Ton war nicht bloß für mich (damals 51 Jahre) und den Physiker KOHLRAUSCH (59 Jahre) unhörbar, sondern auch für 2 etwa 40jährige Teilnehmer meiner Übungen, während alle übrigen 17 Teilnehmer, Studierende im Alter zwischen 20 und 30 Jahren, ihn noch hören konnten (4, S. 115).

Vermutlich würde sich aber bei weiterer Vermehrung des Beobachtungsmaterials, namentlich an älteren Personen, die Durchschnittskurve GILDEMEISTERS doch weniger senken. Fälle, wie der des 44jährigen Schneiders, der noch über 18 000 Schw. ( $d^7$ ) hörte, dürften nach dem folgenden nicht gar selten sein. Jedenfalls muß die Kurve für das spätere Alter nicht mehr geradlinig, sondern langsamer abfallen. Wenn auch die Angaben früherer Forscher zumeist unter der Unzuverlässigkeit der damaligen physikalischen Höhenbestimmungen leiden, so müssen einige darunter doch als exakt anerkannt werden.

So bedutzte KOENIG die bei richtiger Anwendung durchaus sichere<sup>2)</sup> Differenztonmethode, die er virtuos handhabte, F. A. SCHULZE eine mit konstantem Druck angeblasene und nach 6 verschiedenen Methoden übereinstimmend geeichte Galtonpfeife. KOENIG berichtet nun über sich selbst, daß er mit 48 Jahren noch  $f^7 = 21846$  Schw., mit 58 Jahren noch  $d^7 = 18204$  Schw. gehört habe<sup>3)</sup>. Mit 67 J. war auch dieses verschwunden,

<sup>1)</sup> Des öfteren wurden Eichungen zur Prüfung der von den Verfertignern beigegebenen Tabellen vorgenommen und diese bei den neueren Pfeifchen recht genau gefunden, sobald der genügende Druck und eine konstante Anblasevorrichtung angewandt werden (STUMPF und MEYER, SCHWENDT, F. A. SCHULTZE, KOEHLER, HEGENER, SCHAEFER u. a.). Um die richtige Anwendung des Pfeifchens hat sich besonders K. L. SCHAEFER verdient gemacht (vgl. Passows u. Schaefers Beiträge z. Anat. usw. d. Ohres usw., bes. Bd. 20 u. 21).

<sup>2)</sup> Vgl. m. Abh. 4. Daß die von F. KRUEGER behaupteten Abweichungen der Differenzöne von ihren berechneten Höhen auf Irrtümern beruhen, ist von mir zahlenmäßig nachgewiesen. Die Differenzöne stimmen auf das genaueste mit der Berechnung überein (7, S. 69ff.).

<sup>3)</sup> Ann. d. Phys. (3) Bd. 69. 1899. S. 636. Eine KOENIGSche  $f^7$ -Gabel wurde auch bei einigen Versuchen SCHWENDTS (Verhandl. d. Baseler Naturf. Ges. Bd. 12, S. 23) von Normalhörenden noch gehört. Das Alter ist nicht angegeben, aber die Tatsache wird wohl richtig sein. Dagegen muß

dagegen (wie aus dem Zusammenhang zu erschließen)  $c^7 = 16\ 284$ ) noch erhalten. Dies mag ungewöhnlich sein. Dennoch ist kaum anzunehmen, daß der hochbegabte Akustiker gerade auch im Punkte der oberen Hörgrenze eine einzigartige Begabung besessen haben sollte<sup>1)</sup>.

F. A. SCHULZE war, als er die Hörgrenze an sich selbst = 20000 bis 21000 Schw. bestimmte (Ann. d. Phys. (4) Bd. 24, S. 785ff. 1907), 35 Jahre alt. Mit 47 Jahren bestimmte er sie nach brieflicher Mitteilung aufs neue mit einer ganz ähnlichen Methode, wie sie GILDEMEISTER gebrauchte, dessen Arbeiten er noch nicht kannte, und erhielt 14000—15000 Schw., was in GILDEMEISTERS Durchschnittskurve etwa 33 Jahren entsprechen würde. HEGENER fand an sich selbst mit 38 Jahren 18000, mit 50 Jahren 16000 als Grenze (briefl. Mitt.). 18000 würde bei GILDEMEISTER 20 Jahren, 16000 28 Jahren entsprechen.

Es sei noch erwähnt, daß auch einige meiner jüngeren Mitarbeiter bei den vorliegenden Untersuchungen von GILDEMEISTER geprüft wurden und daß sie fast alle, auch der damals 42jährige Prof. v. HORNBOSTEL,  $c^7 = 17200$  Schw. noch gut hören konnten. Auch diese Befunde liegen über den entsprechenden Punkten der GILDEMEISTERSCHEN Kurve. Übrigens hatte GILDEMEISTER bei allen seinen Prüfungen absichtlich nur eine mittlere Stärke angewandt, um des einohrigen Hörens sicher zu sein. Er bemerkt, daß man bei 25facher Intensität um einen Halbton höher komme, was in diesen Höhen gegen 1000 Schw. ausmacht.

Ich selbst hörte mit 46 Jahren bis gegen das Ende der 6-gestr. Oktave, also etwa 15000 Schw., mit 72 Jahren noch bis  $g^6 = 12200$ , und dies war selbst bei der letzten Prüfung mit 76 Jahren, die die Herren Prof. K. L. SCHAEFER und Dr. CLAUS vornahmen, noch der Fall.  $g^6$  war auf beiden Ohren eben noch hörbar, wenn das genau geeichte Galtonpfeifchen mit 8 cm Quecksilberdruck (= 100 cm Wasserdruck) angeblasen wurde. Für den schwächeren Ton des STRUYCKENSCHEN Monochords lag die Grenze bei  $e^6 = 10200$  (also bewirkte Verstärkung hier Erhöhung um eine kleine Terz oder 2000 Schw.).

Soviel läßt sich nach alledem jedenfalls sagen, daß das normale Ohr in der Jugend bis  $e^7$ , höchstens  $f^7$  hört, daß vom 40. Lebensjahr an die Grenze unter  $e^7$ , im Greisenalter bis  $g^6$  herabsinkt und da stehenbleibt. Dies stimmt recht schön mit ZWAARDEMAKERS „Senilitätsgesetz“ überein, der mit KOENIGSCHEN Klangstäben folgende Grenzen fand: Kindesalter  $e^7$  ( $f^7$ ), Jugendliche noch ebenso, 32 J.  $es^7$ , 40 J.  $d^7$ , 43 J.  $cis^7$ , 51 J.  $e^7$ , 54 J.  $b^6$ , Greisenalter etwas oberhalb oder unterhalb  $g^6$ . Wenigstens für die höchsthörenden Individuen dürften diese Grenzen bei mäßig starker Tongebung

die Angabe, daß sogar das  $a^7$  einer EDELMANNSCHEN Galtonpfeife gehört wurde, stark bezweifelt werden. Daß etwas gehört wurde, bezweifeln wir nicht, aber es wird ein tiefer liegender „Schneidenton“ gewesen sein (s. o.), an den man damals noch nicht dachte.

<sup>1)</sup> Daß musikalische Menschen höher hinauf hören könnten als unmusikalische, ist bis jetzt unerwiesen. Wenn auch unter den 16 kürzlich von PRATT (Americ. Journ. of Psychol. Bd. 31, S. 403) untersuchten Individuen die 3 Höchsthörenden zugleich die musikalisch Begabtesten waren, so hat doch prinzipiell das gute Gehör im musikalischen Sinne mit den höchsten Tönen, die für die Musik ganz gleichgültig sind, nichts zu tun. Die Höchhörigkeit könnte nur etwa als häufiger Nebeneffekt musikalischer Anlage oder als korrelative Eigenschaft im Sinne einer bloßen Begleiterscheinung auftreten. Hierüber fehlt es noch an Material.

gelten<sup>1</sup>). Auch die von STRUYCKEN aufgenommene Statistik (Passows u. Schaefers Beiträge Bd. 6, S. 289ff.) zeigt bei starken individuellen Verschiedenheiten doch einen im allgemeinen regelmäßigen Abfall in diesem Sinne zwischen 10 und 80 Jahren. Die Grenzen liegen auch hier wohl durchgängig etwas höher als bei GILDEMEISTER. Ebenso fand GRADENIGO (1) bei Prüfung vieler Personen zwischen 10 und 85, ja 90 Jahren eine Senkung von  $g^7$  ( $f^7$ ?) bis  $d^6$ .

Es geht also normalerweise nur eine Sexte, höchstens eine Oktave, verloren, und wenn nicht noch eine allgemeine Gehörsschwächung dazu tritt, so ist für das Verständnis der Sprache damit fast nichts, für die Musik überhaupt nichts verloren. Aber freilich dürfte in den meisten Fällen eine solche mehr oder minder hinzutreten.

Das Verhältnis zwischen der allgemeinen Hörschärfe und der oberen Hörgrenze ist nicht einfache Parallelität. Bei einer 74jährigen Dame, deren eines Ohr geradezu schwerhörig ist, so daß leises Sprechen in der Nähe ihr auf dieser Seite unverständlich bleibt, fand sich die Hörgrenze bei Prüfung mit dem Monochord durch die Herren SCHAEFER und CLAUS auf beiden Ohren noch in der Gegend von  $c^6$ , auf dem schwerhörigen sogar einen halben Ton höher. Hieraus scheint hervorzugehen — was auch STRUYCKEN betont hat —, daß man die obere Hörgrenze im gewöhnlichen Sinne nicht ohne weiteres als Maß der Hörfähigkeit benutzen kann. Das Gehör kann für Flüstervokale und Konsonanten oder ferne Töne schon öfters versagen, wenn die Hörgrenze für starke Töne in der Nähe noch recht hoch liegt<sup>2</sup>).

Auf pathologische Herabsetzungen der oberen Hörgrenze und des Verständnisses der Sprachlaute werden wir im 9. Kapitel näher eingehen.

<sup>1</sup>) Mit sehr starken, durch Kondensatoren erzeugten Tönen fand C. E. LANE (Physical Rev. (2) Bd. 19. 1922) zuweilen bei Kindern unter 15 J. 25000 =  $g^7$  als Grenze. Solche Ausnahmen mögen vorkommen; aber für die gewöhnlichen Stärkegrade und das durchschnittliche normale Ohr dürfte  $f^7$  das Äußerste sein.

<sup>2</sup>) Auffallend bleibt, daß viele Personen die Feldgrille nicht hören, deren Zirpen nach den zuverlässigen Messungen phonographisch aufgenommener Proben durch KREIDL und REGEN (Sitzungsber. d. Wiener Akad. 1905) um  $c^5 = 4200$  Schw. liegt und in Hinsicht der Höhe bei den untersuchten Exemplaren nur ganz wenig variierte. Vermutlich sind da Gehörsschwächungen, vielleicht solche partieller Art gerade in dieser Gegend, mit im Spiele. Ich selbst höre es noch, aber auch nicht immer, wahrscheinlich nur bei starker Tongebung. Noch kürzlich hörte ich eine auf  $g^4$  singen; der Ton war leicht durch das bloße Gehör zu bestimmen. Sollte es nicht doch Baß- und Diskantgrillen geben?

## 6. Kapitel.

# Die subjektiven Tonhöhen der stimmlosen Sprachlaute.

### 1. Tonhöhen von Klängen und Geräuschen.

Wenn der Musiker die Tonhöhe eines gegebenen Klanges als  $e^2$  bestimmt, so benennt er den Klang seiner „musikalischen Qualität“ nach als ein  $\mathcal{C}$ , dieses aber seiner Höhenlage nach als ein 2-gestrichenes. Er unterscheidet die Eigenschaft, die sich oktavenweise wiederholt und mit einem bestimmten Buchstaben bezeichnet wird, von der Eigenschaft, die mit den Schwingungszahlen einfach parallel läuft und durch die Zahlenindizes sowie durch den Unterschied der großen und kleinen Buchstaben, am genauesten durch die Schwingungszahlen selbst ausgedrückt wird. Die Höhe in diesem engeren Sinne kann man auch Helligkeit nennen; Höhe in weiterem Sinn ist Helligkeit und musikalische Qualität zusammengenommen.

Nun ist, was gewöhnlich Ton genannt wird, im akustischen Sinne vielmehr ein Tonkomplex oder Klang. Darin hat jeder Teilton seine besondere Höhe<sup>1)</sup>; gleichwohl schreiben wir auch dem Klang als Ganzem eine Höhe zu, nämlich die des stärksten oder wenigstens für unsere Auffassung maßgebendsten 1. Teiltons. Genau gesprochen haben wir es hier mit einer „Komplexhelligkeit“ zu tun. Aber Klänge von analoger oder mit der Schwingungszahl des Grundtones nur langsam veränderlicher Zusammensetzung lassen sich eben ihrer Komplexhelligkeit nach gleichfalls in eine Reihe ordnen, die mit den Schwingungszahlen der Grundtöne parallel läuft. Die Einordnung der seltener gehörten einfachen oder nahezu einfachen Töne in diese uns gewohnte Reihe der Klänge, also die Bestimmung der absoluten Höhe einfacher Töne ist nicht immer leicht. Aber das Gehör korrigiert sich bei genauerer Vergleichung selbst, und man kann zuletzt doch einen Klang und einen Ton in Hinsicht ihrer Höhe einander zuordnen.

<sup>1)</sup> Denkt man sich den Teilton für die Empfindung erst durch das Heraushören selbst entstehend, so wächst ihm eben auch seine Höhe erst durch diesen Akt zu.

Geräusche, wie es die Flüstervokale und stimmlosen Konsonanten sind, haben unstreitig gleichfalls Helligkeits-, also Höhenunterschiede; aber die Höhen lassen sich meist noch weniger leicht als die der einfachen Töne den gewohnten Klanghöhen zuordnen. Sie sind meist weniger ausgeprägt und haben etwas Diffuses, wenn das Geräusch, wie z. B. das des Regnens, sich annähernd gleichmäßig über eine größere Schwingungszone erstreckt. Aber vielfach treten doch markiertere Stellen in Geräuschen hervor, die von stärkerer oder dichter Erfüllung kleiner Teile der Tonlinie mit Schwingungen herrühren mögen und genauere Tonhöhenbestimmungen durch das bloße Ohr zulassen.

Geräusche haben aber auch, wie schon bemerkt, musikalische Qualitäten. Nicht bloß bei den tonhaltigen, sondern auch bei den reinen Geräuschen sind Intervalle erkennbar, durch welche auch die Bestimmung der absoluten Tonhöhe außerordentlich erleichtert wird. Freilich kommen auch da, wie bei den Tönen, Oktaventäuschungen vor, bei denen die musikalische Qualität richtig, die absolute Höhenlage aber falsch bestimmt wird: ein deutlicher Beweis, daß auch Geräuschen eine von der bloßen Höhe noch zu unterscheidende musikalische Qualität zukommt.

Die in den Sprachgeräuschen beobachtbaren Höhen, die als gleich mit bestimmten Tonhöhen erkannt werden, wollen wir im folgenden kurz als die Tonhöhen dieser Geräusche bezeichnen. Wir bestimmen sie rein nach dem Gehör, ohne Zuhilfenahme graphischer Messungen, fügen aber auch Beobachtungen hinzu, die sich bei den If.-Versuchen über die dabei auftretenden Veränderungen der Tonhöhen stimmloser Sprachlaute machen ließen.

## 2. Tonhöhen der Flüstervokale.

Hierüber liegen zahlreiche Beobachtungen vor, da man in den beobachteten Höhen zugleich die Formanthöhen der Vokale, auch der gesungenen und gesprochenen, zu finden glaubte. Sind doch die Flüstervokale geradezu als „Vokale in Reinkultur“ bezeichnet worden, weil hier die Formanten allein ohne die variablen Stimmtöne zu beobachten seien. Die Beobachtungen selbst zeigen leider noch nicht die wünschenswerte Übereinstimmung. Sie erfordern eben Übung und vielfache Wiederholung, wenn man zur Sicherheit kommen will<sup>1)</sup>. Der Verfasser hat sich seit Jahren in Augenblicken der Muße immer wieder damit beschäftigt und

<sup>1)</sup> GUTZMANN sagt (5, S. 73), die Höhen seien durch ein geübtes Ohr leicht zu bestimmen, NAGEL dagegen (S. 751), im Flüstergeräusch sei überhaupt keine bestimmte Tonhöhe zu erkennen. Die Wahrheit liegt in der Mitte.



glaubt die folgenden Angaben in der Zuversicht machen zu können, daß andere, die sich in ähnlichem Maße darein vertiefen, unter gleichen Beobachtungsumständen im wesentlichen zu den gleichen Ergebnissen gelangen werden.

Um konstante Werte zu bekommen, muß man natürlich einen Vokal immer in der gleichen Nuance flüstern. Aber auch die Stärke darf nicht zu sehr variieren, jedenfalls nicht zu groß sein, da sonst das Flüstern in ein Blasen oder Fauchen übergeht und sich die Höhe etwas verschiebt. Manche finden es vorteilhaft, den Vokalen Konsonanten vorzuschicken („wo, pi“ usw.). Äußerste Stille der Umgebung ist nötig, besonders empfehlenswert die Nachtzeit. Vergleichung mit Klaviertönen hilft trotz der Verschiedenheit der Klangfarbe außerordentlich, auch solchen, die mit absolutem Gehör begabt sind, besonders zur Feststellung der Oktavenlage. Steht ein Flötenpfeifenregister zur Verfügung, das mindestens die Region  $c^2$ — $g^4$  umfaßt, so dient es natürlich noch besser. Hat man z. B. erkannt, daß es sich um ein  $\mathcal{G}$  handelt, so probiere man  $g^2$ ,  $g^3$ ,  $g^4$  zur Vergleichung. Einer dieser Töne deckt sich genauer als die übrigen mit dem Vokal und wird bei gleichzeitigem Angeben mit ihm zu einer unanalysierbaren Einheit. Man schätzt Flüstervokale, wie Geräusche überhaupt, bei nicht sehr vorgeschrittener Übung zunächst regelmäßig um 1, ja 2 Oktaven zu tief.

Mit Normalstimmgabeln oder Stimpfpfeifchen können solche, die kein absolutes Gehör, wohl aber ein gutes Intervallurteil haben, die musikalische Qualität des Flüstertones feststellen, aber nicht die Oktavenlage. Auch kommt es vor, daß man in Hinsicht der musikalischen Qualität einer Autosuggestion bzw. Assimilation unterliegt: hat man das  $a^1$ -Pfeifchen vorher oder zwischendurch angegeben, so hört man leicht die  $\mathcal{M}$ -Qualität auch in den Flüsterton hinein. Man muß also ein Instrument haben, bei dem man rasch und bequem alle möglichen Töne nacheinander angeben kann.

Die musikalischen Qualitäten der Flüsterlaute springen auch ohne Vergleich mit Instrumentaltönen leichter hervor, wenn man mehrere verschiedene Vokale nacheinander angibt, ähnlich wie man bei dem bekannten Versuch mit hingeworfenen oder beklopften Holzstücken leichter die Melodie heraushört als einen einzelnen Ton. Man kann mit Flüstervokalen leicht Intervalle, gebrochene Akkorde und kleine melodische Phrasen bilden.

Man kann aber nicht auf ein und demselben Vokal eine Melodie flüstern. Dies ist schon öfters bemerkt worden (so von GRABOW und NAGEL) und ist wichtig. Denn gerade daran

zeigt sich, daß für jeden Flüstervokal von bestimmter Nuance nur eine bestimmte charakteristische Tonhöhe existiert. Selbst eine kleine Veränderung der Tonhöhe ist nur möglich, wenn zugleich der Vokal anders nuanciert, z. B. ein A heller oder dunkler, mehr nach Ä oder nach O hin, ausgesprochen wird. Versucht man gleichwohl die Nuancierung beizubehalten, so geht es nicht ohne Gewaltigkeit und nur in engen Grenzen. Ich kann z. B. auf  $as^3$  statt Ö auch Ä flüstern, aber es ist weniger klangvoll als Ö auf dieser Höhe und als Ä auf seiner normalen Höhe ( $a^3$ ). Nur insofern greifen die Höhenzonen ineinander über, als, namentlich bei den hellen Vokalen, die hellste Nuance des tieferen sich mit der dunkelsten oder einer mittleren des höheren in der Tonhöhe deckt. So kann man auf  $b^3$  ein gutes Ü, aber auch ein helles Ä und ein unnatürlich helles Ö flüstern.

Im folgenden sind die Tonhöhen, wie sie sich meiner Beobachtung darstellen, in das HELLWAGSche Vokalschema eingetragen,

da man daran zugleich die möglichen Übergänge anschaulich vor sich hat (über das Schema selbst vgl. 10. Kap.). Bei den rechts stehenden helleren Vokalen sind außer den Haupttönen noch leise Untertöne bemerklich, deren Höhe aber gegen E und I hin immer schwerer erkennbar wird, da sie immer schwächer werden. Sie sind mit Viertelnoten bezeichnet. Sämtliche Töne sind aber um eine Oktave höher, als es den Noten im Violinschlüssel entspricht.

Die absolute Höhe des U ist am leichtesten zu bestimmen, da es mit ganz geringfügiger Änderung der Mundstellung auch gepfeifen werden kann, und für Pfeiftöne die richtige Höhe am

Klavier oder an Labialpfeifen, auch mit Schwebungen und anderen objektiven Methoden sicher festgestellt werden kann. Von da aus kann man sich dann schrittweise auch von der richtigen Oktavenlage der übrigen Töne überzeugen, zu denen stetige Übergänge hinführen.

Solche Übergänge führen zunächst von U über O, OA und AO nach A, dann über Ä und E nach I, wobei die Tonhöhe stetig hinaufrückt. Man kann aber auch in horizontaler Richtung von einem Grenzpunkt zum anderen gehen. Der streng stetige Übergang von O nach Ö fällt allerdings etwas schwer und der von U nach Ü noch schwerer, weil auf diesen Strecken uns ungewohnte, wenn auch mögliche, Laute liegen<sup>1)</sup>. So kann man wohl auch von A aus senkrecht zur U-I-Linie kommen, indem man von dem Blöklaut AOä statt nach Ä nach einem Ö-ähnlichen, auf der Linie O—E liegenden Laut und von da zu einem dunklen Ü gelangt. Und so dürften überhaupt zwischen allen Punkten in horizontaler und vertikaler Richtung Übergänge möglich sein, die zugleich Höhenübergänge sind.

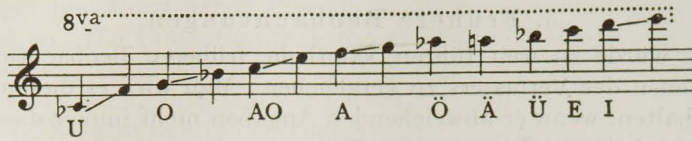
Für jeden Vokal gibt es eine gewisse Zone der Tonhöhen, in der er trotz Änderung seiner Nuance erkennbar bleibt, weil der Begriff des U oder des A selbst diese Nuancen zuläßt und einschließt. Diese Höhen- und Vokalitätsbreiten sind durch die beigefügten Punkte und die darüber stehenden Noten angezeigt. So entspricht z. B. dem dunkelsten zulässigen A  $e^2$ , dem hellsten  $g^{22}$ ). Ein U mit der Tonhöhe  $c^1$  ist nur äußerst schwer und schwach zu flüstern, es ist mehr ein Hauchen. Das gewöhnliche hat die Höhe  $f^2$ , ist aber keineswegs ein ideales U, sondern stark dem O genähert. Umgekehrt ist I von Erwachsenen nur schwer oder gar nicht in der Helligkeit herzustellen, die es unserer Intention nach haben müßte. Kinder flüstern es leichter und besser.

Verzichtet man auf die Darstellung aller möglichen Extreme, Übergänge und Zwischenstufen, sowie auf die Untertöne, so läßt

<sup>1)</sup> In den skandinavischen Sprachen sind sie gebräuchlich. STORM unterscheidet (S. 131) von dem deutschen U das „gemischte U“ im norwegischen Hus (Haus), das Ausländern wie ein Zwischenlaut von U und Y vorkomme. Das schwedische U in Hus, Sju (sieben) liege dem Y noch näher.

<sup>2)</sup> BREMER gibt an (S. 153), daß das geflüsterte Ü in: „Lügst du?“ (drohend) und: „Du lügst“ (strafend) verschiedene Tonhöhen habe. Im ersten Falle könne sein Eigenton bis  $f^3$  herunter-, im zweiten bis  $d^4$  hinaufgehen. Die Beobachtungen können richtig sein, nur muß dann das intendierte Ü in E (I) bzw. ÖÄ übergehen. Den Einfluß des sprachlichen Zusammenhanges, besonders der umgebenden Konsonanten, auf die Flüsterhöhen hat THOMSON (2, S. 563 ff.) eingehend untersucht.

sich die durch Flüstervokale herzustellende Tonreihe übersichtlicher so schreiben:



Zur Kontrolle der angegebenen Höhen mag man außer der Beobachtung der einzelnen Laute etwa folgende Intervalle, gebrochene Akkorde und melodische Phrasen benützen. Dabei ist derselbe Vokal in den verschiedenen Beispielen entsprechend dem verschiedenen Notenwert verschieden anzugeben, z. B. U in Nr. 5 heller als in 1; sonst kommt A nicht als reine Oktave heraus oder man muß ein besonders dunkles A nehmen. Überall aber bleiben die Vokale kenntlich. Vielleicht gewährt das Aufsuchen solcher Tonfolgen, bei denen Text und Melodie sich gegenseitig beinahe eindeutig festlegen, manchem Leser eine Unterhaltung.

## Durchweg höhere Oktave

1. U O OA A Ö E 2. U O OA A Ö E I 3. OA Ao Ö Ä E

4. U (O) OA (Ö) E 5. U A O Ö O Ü OA E 6. E Ä Ö A Ö Ä E

(Ü) (Ü)

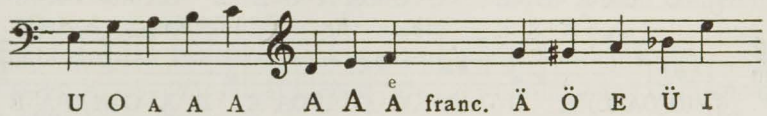
Der Verfasser hat, außer an sich selbst, auch an anderen Individuen — 8 Männern, 5 Frauen und 2 Kindern — die Flüsterhöhen untersucht. Die Fähigkeit zu deutlichem Flüstern ist, wie schon erwähnt, verschieden, sowohl überhaupt als betreffs einzelner Vokale. Aber die Tonhöhen zeigen nur geringe Verschiedenheiten. So kommt es vor, daß Ü und Ä ihre Tonhöhen vertauschen. Daraus geht schon hervor, daß ihr Vokalcharakter nicht ausschließlich durch diese Tonhöhen bestimmt sein kann, denn sie bleiben dabei gut kenntlich. U und O überschreiten zuweilen nach oben hin etwas die angegebenen Grenzen. Kinder flüstern ein U noch auf  $g^2$ , ebenso wie sie mit dem I um einige Töne höher gehen. Auch bei Frauen fand ich U immer höher als bei Männern, etwa =  $g^2$ . Aber immer findet sich dann auch die Farbe etwas geändert, obschon sie kenntlich bleibt. O ist in solchen Fällen nur ganz wenig höher als U, aber der Farbe nach genügend verschieden. Erheblichere Differenzen in

der Tonhöhe zeigt A, das ja auch sonst am verschiedensten gesprochen wird. Hiervon weiter unten.

### 3. Frühere Beobachtungen.

Es würde zu weit führen, sämtliche früheren Beobachtungen mit denen des Verfassers zu vergleichen. Man wird es diesem zu gute halten, wenn er abweichenden Angaben nicht immer dasselbe Gewicht beilegt. In manchen Fällen handelt es sich aber auch nur um Oktaventäuschungen, in anderen um Modifikationen der Aussprache, die teils in dialektischen Gewohnheiten, teils in individuellen Verhältnissen der Mundhöhle wurzeln. An einer kurzen Übersicht seien indessen sowohl Übereinstimmungen als Abweichungen erläutert. Die Ergebnisse objektiver Methoden müssen aber hier, wo es sich nur um die gehörten Tonhöhen als solche handelt, vorerst beiseite bleiben.

In dem von veralteter Gelehrsamkeit strotzenden Buche des Kieler Professors Samuel Reyher: „Mathesis Mosaica“ 1679 finden sich S. 432ff. gewiß zuerst die Flüsterhöhen in Notenbuchstaben und Noten verzeichnet. Er führt sie als eine nebensächliche Kuriosität an, scheint aber, abgesehen von der Tieferlegung um 2 Oktaven, vollkommen richtig beobachtet zu haben. Er notiert so (nur die Schlüssel und die Anordnung sind hier geändert):



A in verschiedener Größe deutet augenscheinlich die wachsende Mundweite an, mit der es vom dunkelsten bis zum hellsten gesprochen wird. In der Beschreibung nennt R. das erste „longissimum sive obscurissimum“, das letzte „brevissimum sive clarissimum“, wobei „lang und kurz“ Unterschiede der Dauer bedeuten müssen. Aus diesen Flüsterhöhen läßt sich schließen, daß das lange holsteinische A damals, wie heute noch, viel dunkler als das hochdeutsche gesprochen wurde, bis in die O-Gegend hineinreichend<sup>1)</sup>. Unter dem „A franconicum“ ist aber sicherlich nicht das A des heutigen Franken verstanden, das damals ebenso wie heute sehr dunkel gesprochen wurde (HANS SACHS reimt sogar a auf o), sondern aller Wahrscheinlichkeit nach das sehr helle rheinfränkische, näher „mittelfränkische“, wie es etwa heute in Köln gesprochen wird. Dem an dunklere Nuancen gewöhnten Ohr REYHERS mochte dieses schon als dem Ä genähert erscheinen.

<sup>1)</sup> REYHER stammte aus Thüringen, lebte aber seit 14 Jahren in Kiel.

Er gibt ihm dieselbe Flüsterhöhe, wie wir sie unserem hochdeutschen A zuschreiben. In den Erläuterungen rechnet er es zu den „vocalis mixtae“ und führt neben ihm als 2. Beispiel noch das hebräische Saegol an. Die Umlautnatur deutet er durch ein darübergesetztes Zeichen an, das in dem mir vorliegenden alten Druck zwar ebenso wohl als o wie als e aufgefaßt werden kann, sicherlich aber e sein soll<sup>1)</sup>.

E ist bei REYHER in der Tonhöhe gleich unserem Ä, wurde also ganz „offen“ gesprochen, wie heute noch im holst. Del (Diele), Hek (Hecht)<sup>2)</sup>.

Auch für einige Diphthongen: ou, au, ei, eu gibt REYHER Noten an, wobei er die Höhen der darin enthaltenen Vokale zugrunde legt. Daraus ist zu ersehen, daß diese Diphthongen damals ebenso wie heute ausgesprochen wurden, z. B. eu = AO—Ü.

Es dürfte für die Sprachforschung nicht ohne Belang sein, daß man auf solche Art durch akustische Feststellungen Kriterien schaffen kann, die auch in späteren Jahrhunderten noch Aufschluß über die Aussprache geschriebener Lautbuchstaben geben. Dabei sind allerdings auch Unterschiede der musikalischen Stimmung zu berücksichtigen. Diese war im 17. Jahrh. höher. Aber die Orgel in dem holsteinischen Städtchen Glückstadt hatte nach A. ELLIS (History of Musical Pitch) im J. 1645 doch nur ein  $a^1 = 453$  Schw., also nur um einen kleinen Halbton höher als das unsrige. Setzt man REYHERS Noten<sup>3)</sup> soviel höher, so wird die Übereinstimmung mit den unsrigen fast noch besser.

HELLWAG (1781) stimmt REYHERS Aufstellungen im allgemeinen zu, gibt aber seinerseits keine Notierungen.

In neuerer Zeit hat DONDERS aus den Flüsterhöhen auf die Formantöne geschlossen, ja sogar den vokalischen Charakter der gesungenen und stimmhaft gesprochenen Vokale aus den beigemischten Flüstergeräuschen hergeleitet und diese darum aufmerksam beobachtet. Aber seine Angaben enthalten auffallende Inkongruitäten. Wie weit die holländische Aussprache daran beteiligt ist (HELMHOLTZ), kann ich nicht beurteilen. Er setzt

<sup>1)</sup> Meine anfängliche Vermutung, daß das französische A in madame gemeint sei, läßt sich nach den Aussagen fachmännischer Berater (Prof. J. PETERSEN, Dr. HOCHSTETTER) nicht mit dem damaligen Sprachgebrauch vereinigen. Franconia bedeutet niemals Frankreich, sondern immer das deutsche Franken, zu dem aber das Rheinland gerechnet wird. Du Cange: „Franconia, Francia Germaniae, unde franconicus.“ REYHER selbst hat einen nachgelassenen Thesaurus linguae latinae seines Vaters herausgegeben, worin das nämliche steht. Zwar „franciscus“ findet sich für „französisch“ wie für „fränkisch“ (KÖRTINGS lat.-rom. Wörterbuch 1907), „franconicus“ aber allem Anscheine nach nur für das letztere.

<sup>2)</sup> Vgl. MÜLLENHOFFS Glossar zu KL. GROTHS „Quickborn“.

$U = f^1$ ,  $O = es^1$  (tiefer als  $U!$ ),  $A = b^1$ ,  $\ddot{U} = a^2$ ,  $E = c^3$  + einem tieferen Ton,  $I = e^3$ <sup>1)</sup>. Überall also zu tiefe Oktaven, aber auch sonst Merkwürdiges; doch die 3 letzten mit unseren Angaben übereinstimmend.

HELMHOLTZ setzt DONDERS' Beobachtungen die seinigen wie folgt gegenüber:  $U = f$ ,  $O = b^1$ ,  $A = b^2$ ,  $\ddot{O} = cis^3$ ,  $\ddot{U} = g^3$  bis  $as^3$ ,  $E = b^3$ ,  $I = d^4$ . Die 3 letzten stimmen ziemlich gut zu unseren Angaben, die übrigen freilich weichen sowohl von denen seines Vorgängers wie von den unsrigen ab,  $U$  ist sogar um 2 Oktaven zu tief, wie bei REYHER<sup>2)</sup>.

MERKEL (1866) und v. ZAHN (1871), der außer sich selbst 45 Personen beiderlei Geschlechts untersuchte, geben die Töne aller 8 Vokale wieder 2 Oktaven zu tief, sonst nur bei  $A$  und  $\ddot{O}$  mit erheblicheren Abweichungen, möglicherweise infolge der sächsischen Aussprache.

ENGELS fragmentarische Angaben (1869) für  $U$ ,  $A$ ,  $\ddot{U}$ ,  $I$  treffen mit den unsrigen zusammen, nur sind sie wieder um 1 Oktave zu tief, obgleich er für  $U$  und  $\ddot{U}$  die Überführung in Pfeiftöne benützte und dabei einen Irrtum für unmöglich hielt (2, S. 316). Er ist eben auch beim Pfeifen, trotz seines ausgezeichneten Gehörs, der Oktaventäuschung unterlegen. Auch KRÖNIGS Angaben (1876) sind durchweg 1 Oktave zu tief, sonst in leidlichem Einklang mit den unsrigen:  $U = c^1 - gis^1$ , bei weiterer Erhöhung in  $\ddot{U}$  übergehend,  $O = d^1 - a^1$  ( $A$  nicht angegeben). „Auf  $cis^3$  kann man ziemlich bequem die 4 Vokale  $\ddot{A}$ ,  $E$ ,  $\ddot{O}$ ,  $\ddot{U}$  flüstern.“

GRABOWS Notierungen (1879) stimmen gut zu den meinigen, auch in der Oktave (daß die höhere Oktave aller notierten Töne gemeint ist, geht aus erläuternden Bemerkungen hervor). Nur bezüglich der beiden mittleren Vokale,  $A$  und  $\ddot{O}$ , ist wieder eine Differenz: sie sind eine kleine Terz tiefer angegeben.

Vorzüglich harmonieren die Notenwerte des Englischen TRAUTMANN (1884), der Hunderte von Personen untersuchte, mit unseren:  $U = g^2$ ,  $O = h^2$  (diese beiden nur wenig höher),  $A = f^3 - g^3$ ,  $E = h^3 - d^4$ ,  $I = f^4$ . Er erkennt darin einen Septimenakkord<sup>3)</sup>.

<sup>1)</sup> DONDERS gibt keine Notenbezeichnungen, sondern nennt die Intervalle, welche die verschiedenen Vokale mit der Normalstimmgabel bilden. Die Wiedergabe seiner Bestimmungen bei HELMHOLTZ ist nicht ganz identisch mit der obigen.

<sup>2)</sup> Nach AUERBACH (1878, bei NAGEL S. 774) bekannte sich HELMHOLTZ später zu DONDERS'  $U$ . In der noch von ihm selbst herausgegebenen 4. Aufl. der „Tonempfindungen“ 1877 steht noch die erste Notierung.

<sup>3)</sup> Auch HELMHOLTZ zieht einen Akkord zur Erläuterung heran: er hört scharfes  $A$ ,  $\ddot{A}$ ,  $E$ ,  $I$  als Quartsextakkord  $d^3 g^3 b^3 d^4$ . Ich würde diesen Noten die Vokale  $OA$ ,  $A$ ,  $\ddot{U}$ ,  $I$  unterlegen.

Auch was BREMER (1893) an Tonhöhen angibt, können wir nur durchweg bestätigen. Gleiches gilt von HEINRICH (1905), der beispielsweise  $\ddot{A} = gis^3 - b^3$ ,  $E = c^4$  setzt.

LLOYD (1890ff.), als Theoretiker bedenklich, als Beobachter aber sehr beachtenswert, gibt an (vgl. besonders 3, S. 241):  $U = d^1$  bis  $e^1$ , offenes  $O = gis^2$ , hinteres  $A$  (father) =  $c^3$ , vorderes (man) =  $f^3$ , offenes  $E$  ( $\ddot{A}$ ) =  $fis^3$ , geschlossenes =  $c^4$ ,  $I = dis^4 - f^4$ : alles bis auf die Oktaventäuschung bei  $U$  zutreffend; „father“ muß dabei allerdings sehr dunkel, mehr nach  $OA$ , gesprochen worden sein (vgl. das schottische father von derselben Höhe sogleich bei STORM).

Auch der Norweger STORM (1892) hat die Flüsterhöhen aufmerksam beobachtet, und was sich von seinen Angaben (S. 98) nach den Beispielen eindeutig auf die deutsche Aussprache beziehen oder übertragen läßt, stimmt, wenn man es 2 Oktaven höher legt, bis auf das  $U$  sehr gut mit den unsrigen überein:  $U$  (wie in *Du*) =  $a$ ,  $O$  (wie in *So*) =  $f - a$ ,  $A = d^1 - fis^1$  (in französischen Beispielen: *lâche* bis *madame*, während das breite  $A$  im Schwedischen und im schottischen father — das gewöhnlichen englischen Ohren nach S. 129 wie *fawther* klingt — sogar bis  $c^1$  hinuntergehe),  $\ddot{O} = a^1$ ,  $\ddot{A} = b^1$ ,  $\ddot{U} = b^1 - h^1$ ,  $E = h^1 - c^2$ ,  $I = cis^2 - d^2$ . Seltsam ist aber die Bestimmung des  $U$ . Der ausgezeichnete Sprachforscher behauptet es ebenso wie die deutschen Gelehrten auszusprechen, vermutet aber, daß er es auf andere Weise bilde. Es gäbe eben einen „Polymorphismus“, demzufolge derselbe Vokal ohne wesentliche Änderung seiner Qualität verschiedene Höhen haben könne; übrigens hätten auch der Däne THOMSEN, der Schwede WULFF und der Engländer SWEET höhere Töne für  $U$  als für  $O$  bekommen. Wir hörten das nämliche soeben von dem Holländer DONDERS. Aber hier muß doch unbedingt eine wesentliche akustische Verschiedenheit der Aussprache obwalten, und es dünkt mich äußerst unwahrscheinlich, daß es Skandinaviern und Holländern physisch unmöglich sein sollte, ein dunkleres  $U$  herzustellen, dessen Tonhöhe unter die des  $O$  hinabginge. Daß das  $A$  bei STORM weiter in die Tiefe reicht, als wir es zulässig fanden (ebenso wie LLOYDS „hinteres  $A$ “), gründet natürlich auch in Verschiedenheiten der Aussprache. Mit Recht sagt STORM: „Es gibt keinen Laut, der einen größeren Spielraum hätte, als gerade  $A$  oder was als  $A$  gilt.“

Für die 15 französischen Vokale gibt MONOYER (C. R. 1898, 1, S. 1637) Tonhöhen, aber ohne alle methodischen Erläuterungen und Garantien, was das Gewicht seiner Behauptungen beträchtlich vermindert.  $U$  (franz. *ou*),  $\ddot{U}$ ,  $I$  sollen in Oktaven übereinanderliegen:  $b^1, b^2, b^3$ . Für  $\ddot{U}$  stimmt  $\mathfrak{B}$  genau, wenn es 1 Oktave höher



notiert wird. Unter derselben Bedingung könnte vielleicht ein extrem spitzes I hergestellt werden, also auf  $b^4$ . Ein U aber scheint mir weder auf  $b^1$  noch auf  $b^2$  möglich, auf  $b^1$  wenigstens nur als äußerst schwacher Pfeifton, wie er rein flüsternd nicht mehr gebildet werden kann. Unter den übrigen Vokalen stimmt nur A =  $f^2$ — $fis^2$  unter der Bedingung der Oktaventransposition mit unserem Befunde. Aber auch auf die Übereinstimmungen können wir bei so dogmatisch vorgetragenen Beobachtungen nur wenig Wert legen.

Sorgfältig hat ROUSSELOT (1901 ff.) in Verbindung mit THOMSON die Höhen der französischen Flüstervokale und dann THOMSON selbst (1906 ff.) die der russischen untersucht. Die Angaben beider Forscher stimmen unter Berücksichtigung der Ausspracheverschiedenheiten vollkommen mit den unsrigen. ROUSSELOT setzt U je nach der Färbung =  $d^2$ — $gis^2$ , O =  $h^2$ — $d^3$ , A =  $dis^3$  (bras) bis  $a^3$  (quart), Ö =  $g^3$  (peur),  $gis^3$  (feu),  $a^3$  (queue), Ä (tête) =  $b^3$ , Ü =  $h^3$ — $c^4$ , E =  $d^4$ , I =  $d^4$ — $g^4$ . THOMSON setzt U =  $es^2$ — $g^2$ , O =  $gis^2$ — $h^2$ , A =  $e^3$ — $g^3$ , Ä („palatales A“) =  $g^3$ — $gis^3$ , offenes E =  $gis^3$ — $ais^3$ , geschlossenes E =  $h^3$ — $cis^4$ , I =  $d^4$ — $dis^4$  (nur selten  $c^4$  oder  $e^4$ ). Diese Höhen erleiden nach THOMSONS eingehenden Beobachtungen kleine Modifikationen durch den Zusammenhang, insbesondere gehen sie vor palatalisierten Konsonanten etwas hinauf, in unbetonten Silben einen Halbton herunter. Außerdem fand er kleine individuelle Unterschiede<sup>1)</sup>.

Die Angaben ABRAHAMS (1916) weichen in einigen Punkten von allen zuletzt genannten wie von den unsrigen um eine Terz bis Quarte ab: O =  $e^2$ — $f^2$ , AO =  $f^2$ — $g^2$ , A =  $c^3$ — $d^3$ , I =  $gis^4$ — $a^4$  (die übrigen übereinstimmend).

Bei dem Gewicht, das ich den Angaben dieses Beobachters beilege, möchte ich dazu folgendes bemerken. Beim I kommt eine solche Höhe gewöhnlich nur vor, wenn es sehr leise und mit Vorsetzung eines Ch pal. gegeben wird. Ich selbst kann so ein I auf  $f^4$ — $g^4$  erzeugen.

Ein A mit der Tonhöhe  $c^3$ — $d^3$  fanden wir schon bei STORM und LLOYD. Ich kann ein solches erzeugen, wenn ich den Luftstrom mehr auf den weichen Gaumen hin lenke und ein wenig nasaliere (LLOYDS „hinteres A“). Leite ich ihn aber, wie gewöhnlich, nach vorn, hauptsächlich gegen den Oberkiefer, so entsteht unweigerlich  $f^3$  oder  $fis^3$ . THOMSON berichtet (2, S. 569), daß er zuerst sein eigenes A bei lautem Vortrag und weiter Mundöffnung =  $cis^3$  gefunden, später aber das A der gewöhnlichen Flüstersprache an sich selbst wie an anderen =  $e^3$ — $g^3$  erkennt habe. Im übrigen muß zugegeben werden, daß beim A wegen der weiten Mundöffnung keine

<sup>1)</sup> Obige Angaben hauptsächlich in der lehrreichen Abh. 2, S. 568ff. Wie mir gesagt wird, stehen sie auch in seinem russisch geschriebenen Buche „Allgemeine Sprachwissenschaft“ 1906, S. 183 und hat THOMSON auch in seinen Arbeiten über russische Dialekte das Hilfsmittel der Tonhöhenbestimmungen vielfach verwendet.

so scharfe Resonanz vorhanden ist wie bei den übrigen Vokalen. Eine ganze Zone des Tongebietes wird diffus auf einmal verstärkt, innerhalb deren bei einzelnen mit einer gewissen Konstanz andere Töne als  $f^3-g^3$  hervortreten. An mir selbst hörte ich gelegentlich bei besonders lautem Flüstern (auf der Straße)  $c^3$ . Bei zwei Damen mit sehr überzeugendem Flüster-A fanden sich sowohl  $c^3$  als  $g^3$  gleichzeitig vertreten. Bei Dr. v. ALLESCH, der ein ungewöhnlich helles A flüstert, macht sich außer  $fis^3$  sogar noch eine höhere Tonpartie,  $c^4-d^4$ , geltend, dieselbe, die bei den Resonanzversuchen mit gesungenen Vokalen ein zweites Maximum darstellt.

Diese Umstände machen die einhellige Bestimmung der Flüsterhöhen beim A schwieriger als sonst. Die beste Hilfe ist wieder Vergleichung benachbarter Vokale und Beobachtung der so entstehenden Intervalle, z. B. zwischen Ö und A oder O und A. A ist im 1. Fall um 1–2 Töne tiefer, im 2. um eine Quinte bis Sexte höher. Unmöglich kann ich in dem Übergang von A zu E eine Oktave erkennen, wie man es nach der ABRAHAMSCHE Bestimmung erwarten sollte, sondern nur eine Quinte oder Quarte. Ebenso urteilten Dr. SCHÜNEMANN und andere darüber befragte hochmusikalische Personen.

Dennoch ist nicht zu zweifeln, daß ABRAHAM seine eigenen Flüsterhöhen richtig beobachtet hat. Diese weichen nur eben von den gewöhnlichen für O, OA und A erheblich ab. Es bleibt aber noch die Paradoxie, daß er diese Laute besonders hell flüstert, während seine Tonhöhen tiefer als sonst im Hochdeutschen liegen. Darüber werden wir weiter unten bei den genetischen Betrachtungen eine wahrscheinliche Vermutung äußern. In anatomisch-physiologischer Hinsicht wäre es wohl denkbar, daß der Unterschied von Brachy- und Dolichocephalie hier von Einfluß würde (v. HORNBOSTEL, an dem er ähnliche Tonhöhen festgestellt hat, ist zwar dolichocephal, gehört aber — das gereicht ihm keineswegs zur Unehre — überhaupt nicht zu den guten Flüsterern).

Zuletzt hat R. PAGET (1922) vollständige Tabellen der Flüsterhöhen veröffentlicht<sup>1)</sup>, die er, wie DONDERS, TRAUTMANN, LLOYD u. a. als Grundlage der ganzen Vokallehre zu betrachten scheint. Er schreibt jedem Vokal, auch U, O, A, zwei durch eine Lücke getrennte Töne zu, einen tieferen „Resonatornton“ und einen höheren „Pfeifton“, die den beiden durch die Lage der Zunge im Munde gebildeten Hohlräumen entsprechen<sup>2)</sup>. Der obere steigt von U bis I in die Höhe, der untere rückt zunächst in die Höhe, dann wieder herunter (vgl. unsere Tafel S. 145). Der 1. liegt zwischen  $dis^2$  und  $e^4$ , der 2. zwischen  $dis^1$  und  $ais^2$ . PAGET findet sie ebenso bei gesungenen Vokalen; jeder Vokal habe aber eine gewisse Variationsbreite seiner Töne.

<sup>1)</sup> Die mir bekanntgewordenen Artikel (1) sind äußerst knapp gehalten und gestatten keinen näheren Einblick in die Einzelheiten der Methode und Ergebnisse. Hauptsächlich hat P. mit dem bloßen Ohr gearbeitet, verlangt aber mit Recht für die Bestimmung der Flüsterhöhen ein hervorragend gutes musikalisches Gehör; später benutzte er auch (für gesungene Vokale) den „Kehltonschreiber“.

<sup>2)</sup> Dieselbe Behauptung findet sich schon bei LLOYD (3), der aber die tieferen Resonanzen nicht näher angibt.

Die oberen Komponenten PAGETS decken sich gut mit den von uns angegebenen Höhen (soweit der Unterschied der Aussprache eine Vergleichung gestattet). Auch in der Oktavenlage ist kein Unterschied. Dagegen ist es mir nicht möglich, aus den geflüsterten U, O, A noch je 2 Töne herauszuhören und im U einen Ton zu erkennen, der unterhalb  $b^1$  läge. Bei gesungenen Vokalen liegen die Verhältnisse anders: da hat U einen Oberformanten und lassen sich auch in einem starken O und A in unmittelbarer Nähe noch mehrere Verstärkungszonen unterscheiden (1. Kap.). Aber man darf nicht ohne weiteres für geflüsterte und gesungene Vokale dieselben Strukturen verlangen, wenn auch gewisse wesentliche Züge gemeinschaftlich sein müssen. Und in beiden Gebieten ist mit den sich dem Ohr aufdrängenden Tönen noch bei weitem nicht die Gesamtstruktur gegeben.

Im ganzen lehrt dieser Überblick doch, daß die Beobachter seit mehr als einem Vierteljahrhundert fast ausnahmslos übereinstimmen, die Fehler in Hinsicht der Oktavenlage ganz verschwunden sind und die übrigbleibenden Verschiedenheiten wesentlich auf verschiedener Nuancierung desselben Vokals beruhen, wodurch sie den Zusammenhang der Flüsterhöhen mit der Vokalität nur aufs neue bestätigen.

#### 4. Hauch- und Pfeiftöne.

Unter Hauchtönen verstehe ich die mit Aspiration geflüsterten Vokale, wobei also keine Sprengung des Kehldeckels stattfindet, sondern die Luft durch die offene Glottis durchgesandt wird; wie in Hu, Ha. Die Hauchvokale bleiben gewissermaßen vom Geräusch des H durchzogen. Man kann dabei größere Stärke anwenden (Ha! gegenüber A), und dadurch wird auch die Tonhöhe etwas verändert. Im ganzen ist sie natürlich dieselbe.

U kann man noch gut und kräftig auf  $c^2$  hauchen, wobei allerdings das Hauchen in ein Pfeifen mündet. Auch nach oben kann seine Grenze noch gut bis  $g^3$  erweitert werden, ohne daß O entsteht. Die O-Zone selbst ist hier  $e^2-c^3$ , die A-Zone  $e^3-a^3$ .  $a^3$  kommt bei starkem Hauchen mit  $a^2$  zusammen heraus, so daß die Täuschung entstehen kann, als hätten hier O und A dieselbe Höhe. Aber auch in diesem Falle scheint mir  $f^3$  vorhanden und wesentlich. Ä kann man bis  $c^4$  treiben. E reicht in seinen Nuancen von  $b^3$  bis  $d^4$ , I von  $c^4$  bis  $d^4$ , dieses also etwas weniger hoch als beim reinen Flüstern.

Aus dem Hauchen wird ein Pfeifen, sobald die Lippen einen so engen Kanal bilden, daß ein Labialpfeifenton entsteht, der nebenher von Geräuschen begleitet ist. Bei U und Ü ist dazu nur

eine geringfügige Veränderung erforderlich, eine etwas größere bei I. Daher kommt es, daß die entstehenden Pfeiftöne eine unverkennbare Ähnlichkeit mit den genannten Vokalen haben. In ganz auffälliger Weise gilt dies von  $\ddot{U}$  und dem Pfeifton  $b^3$ . Man kann direkt die Aufforderung stellen, ein  $\ddot{U}$  zu pfeifen, und wird fast unfehlbar von einigermaßen geübten Pfeifern  $b^3$  bekommen<sup>1)</sup>. Dieser Ton ist eben im geflüsterten  $\ddot{U}$  (wie im stimmhaften und gesungenen) als Hauptbestandteil enthalten.

Bei den übrigen Vokalen ist die erforderliche Mundstellung eine für das Pfeifen ganz ungeeignete. Pfeift man aber die zwischen  $U = c^2-f^2$  und  $\ddot{U} = b^3$  liegenden Töne selbst, so stehen sie ihrer Vokalähnlichkeit nach jenen Nuancen am nächsten, die durch Flüstern beim stetigen Übergang von U nach I entstehen. Dies ist in Anbetracht der Einfachheit der Pfeiftöne (s. u.) für die allgemeine Theorie der Vokalität von Bedeutung.

Der Umfang des menschlichen Pfeifregisters deckt sich ziemlich genau mit dem des Flüsterns, doch überschreitet es die Grenzen des deutlichen und bequemen Flüsterns nach beiden Seiten um 1—2 Töne. BARTH gibt an  $c^2-c^5$ , AUERBACH  $a^1-c^4$ , ABRAHAM  $cis^2-a^4$ ; bei mir ist er, wenn ich mich auf die äußersten Grenzen einübe,  $a^1-fis^4$ , bequem aber  $c^2-c^4$ , bei einem daraufhin geprüften Sänger gleichfalls  $c^2-c^4$ , bei einer Kunstpfeiferin bis  $a^4$ . Bei Kindern scheint das ganze Register etwas nach der Höhe verschoben. In der Höhe macht indessen die Übung beträchtliche Unterschiede. Auf der Straße hört man pfeifende Kinder, auch gelegentlich Erwachsene, die sich noch in der oberen Hälfte der 4-gestrichenen Oktave mit Leichtigkeit bewegen; selbst ein  $d^5$  ist mir begegnet. Es ist übrigens begreiflich, daß man bei der Intensität der hohen Pfeiftöne (sie gehören zu den das Ohr am stärksten angreifenden Lauten) die Höhe weitertreiben kann als beim Flüster-I.

Pfeiftöne sind allem Anscheine nach völlig einfache Töne. Wenn man mit Interferenzröhren den Grundton ausschaltet, bleibt nichts als das Blasen übrig. Nun könnten zwar ungerade Teiltöne vorhanden gewesen sein, die mit dem Grundton zusammen ausgeschlossen würden. Aber für solche sind schwebende Hilfsgabeln

<sup>1)</sup> So auch NAGEL. Man kann sogar — ein nicht unbekannter Scherz — beim Aussprechen des Wortes „hübsch“ den Vokal pfeifen und trifft regelmäßig die Nähe von  $b^3$ . Die Franzosen sollen es ebenso machen mit der Frage: „Aimez-vous les confitures?“

AUERBACH sagt von den Pfeiftönen überhaupt (S. 690): „Die nächste Verwandtschaft unter den Vokalen ist die mit dem  $\ddot{U}$ .“ Dies geht aber zu weit.

von annähernd gleicher Höhe ein Prüfungsmittel von idealer Feinheit: und sie zeigen nichts davon. Wahrscheinlich hängt die außerordentliche Stärke und die angreifende Wirkung hoher Pfeiftöne gerade mit ihrer Einfachheit zusammen, indem sich die ganze Energie der Schwingungen auf diesen einzigen Punkt der Tonlinie, also des aufnehmenden nervösen Endgebildes, konzentriert.

Hier liegt nun ein weiterer Grund, warum man U, Ü und I pfeifen kann, aber nicht A. Denn die Reihe der einfachen Töne trägt, wie wir noch näher ausführen werden (13. Kap.), von den tiefsten bis in die Nähe der unteren Pfeiftongrenze U-Charakter, in der  $b^3$ -Gegend Ü- und weiter hinauf I-Charakter, während andere Vokalitäten in ihr nur mit fern anklingenden Ähnlichkeiten vertreten sind. Aus der völligen Einfachheit erklärt sich ferner zum größten Teil auch die bekannte Oktaventäuschung bei den tieferen Pfeiftönen. Im Grunde ist es verkehrt, zu sagen, daß ein Pfeifton um 1—2 Oktaven zu tief geschätzt würde: wir müßten sagen, daß gewöhnliche Klänge um soviel zu hoch erscheinen. Unser Urteil über solche bezieht sich auf den Grundton; diesen meinen wir zu hören, er erscheint aber infolge der Obertöne weit heller als es seiner Beschaffenheit an sich entspricht. Da nun die Klänge der Musik durchweg obertonhaltig sind, so hat sich unsere Vorstellung von den Tonlagen, dem 1-, 2-gestrichenen  $c$  usf. an ihnen gebildet, und wir messen nach ihnen die einfachen Töne.

Etwas dürfte wohl auch der Umstand mitwirken, daß bei allen durch unsere angeborenen Tonwerkzeuge hergestellten Tönen die jeweilige untere Grenze des betreffenden Registers uns wie eine absolute Tongrenze anmutet und daher die ihr nahe kommenden, auch nur mit einer einzigen Anstrengung zu erzeugenden Töne einen absolut „tiefen“ Charakter annehmen. Daß aber dieses Motiv nicht das entscheidende ist, geht daraus hervor, daß auch durch Interferenzröhren von ihren Obertönen befreite Töne leicht zu tief erscheinen. So hielt einer der gewiegtsten Instrumentenkennner, Prof. C. SACHS, ein solches  $c^1$  für  $c$  und einen nicht einmal ganz obertonfreien, nur besonders weichen  $c$ -Klang für  $C^1$ .

<sup>1)</sup> Auch der Kuckucksruf, der seiner Klangfarbe nach gleichfalls aus ganz oder nahezu einfachen Tönen bestehen muß, wird in der Regel eine Oktave zu tief geschätzt. Selbst BALDAMUS (Das Leben der europäischen Kuckucke, 1892) notiert ihn  $e^1-c^1$  statt  $e^2-c^2$ . Mit Recht findet er die Töne U-ähnlich, worauf ja auch schon der Name des Vogels hindeutet; die Tonlage ist genau die eines tiefen Flüster-U. Aber mit Unrecht meint er, man könne ihn nicht mit Instrumenten nachbilden: man braucht ihn nur in derselben Höhe (also die tiefstmögliche Terz) zu pfeifen, so glaubt man den Vogel zu hören.

Das Zutieferscheinen gilt aber auch nur für die untere Gegend des Pfeifregisters. Sobald man gegen die 4-gestrichene Oktave kommt, hört die Täuschung auf, weil eben in dieser Gegend auch die gewöhnlichen Klänge schon weit weniger obertonhaltig sind und alle Klangarten einander ähnlich werden. Infolge dieses verschiedenen Eindrucks der tiefen und hohen Pfeiftöne erscheint uns der Umfang des Pfeifregisters, der Abstand der höchsten von den tiefsten Tönen weit größer als 2 Oktaven zu sein. Es ist dieselbe Sache auch bei den subjektiven Tönen des Ohrenklingens, über die ich (5) zahlreiche Beobachtungen machen konnte. Da sie um eine Oktave tiefer als die Pfeif- und Flüstertöne hinabreichen, so nimmt der absolute Tiefeneindruck hier noch bedeutend zu. Die um  $c^1$  oder noch darunter liegenden erscheinen sozusagen unendlich tief. Es tritt da allerdings noch eine besondere Täuschungsquelle auf: die merkwürdige Zerflossenheit in räumlicher Hinsicht. Sie haben gegenüber den hohen, scharf umrissenen etwas Nebuloses, ähnlich wie die Töne gegen die wirkliche untere Tongrenze hin. Auch die Lokalisation im rechten und linken Ohr, die bei den höheren subjektiven Tönen ganz unzweideutig und unfehlbar ist, nimmt hier an Schärfe ab.

Auf die Tieferschätzung der Flüstervokale bei früheren Beobachtern können aber diese Erklärungsgründe nicht wohl Anwendung finden. Sie erstreckt sich ja auch auf die helleren Vokale. Es scheint hier vielmehr eine allgemeine Eigentümlichkeit der Geräusche von Einfluß. Auch wenn man über die Öffnung des Resonanzkastens einer Gabel hin bläst, hat man den Eindruck einer tieferen Oktave des Stimmgabeltones.

##### 5. Tonhöhen der stimmlosen Konsonanten.

Obgleich diese einen noch weniger tonalen Charakter tragen als die Flüstervokale, lassen sich doch auch bei ihnen durchweg gewisse Tonhöhen erkennen, und es sind auch bereits Angaben veröffentlicht, unter denen die neueren recht gut miteinander stimmen<sup>1)</sup>. Ein gutes Mittel, sich darüber klar zu werden, ist oft, bei unveränderter Mundstellung den dieser Stellung entsprechenden

<sup>1)</sup> Daß die von OSKAR WOLF früher angegebenen Tonhöhen der Konsonanten nicht nur von v. BEZOLD, sondern auch noch von URBANTSCHITSCH (Lehrb. d. Ohrenheilkunde) u. a. zugrunde gelegt wurden, war ganz ungerechtfertigt. Der APPUNNSche Obertonapparat, mit dem WOLF die Konsonantengeräusche verglich, eine Metallzungenreihe, war ein gar zu ungeeignetes Hilfsmittel; das schlechteste Klavier hätte bessere Dienste getan. Auch MERKELS Notierungen der Konsonanten sind sehr unzuverlässig.

Vokal zu flüstern. Die Höhe des Vokals ist dann auch die des Konsonanten. Dieser mündet eben in den ganz leise und kurz angehängten Vokal. Doch ist natürlich hier Vorsicht in der Ausführung nötig.

Das folgende Verzeichnis enthält die von mir beim eigenen stimmlosen Aussprechen beobachteten Tonhöhen der wichtigsten Konsonanten in alphabetischer Ordnung. Von vornherein sei aber betont, daß die Höhen je nach der Aussprache innerhalb gewisser Grenzen variieren. Namentlich bei K, Sch und Ch gutt., aber auch selbst bei S kann dies leicht eine Quinte und mehr ausmachen. Daher sind auch individuelle Unterschiede nur zu erwarten. Die angegebenen Höhen sind die meinigen bei bequemer mittelheller Aussprache.

Konsonanten	Tonhöhen	Konsonanten	Tonhöhen
B	$e^3 - fis^3$	M	$des^3 - f^3$
Ch insgesamt	$c^1 - g^4$	N	$as^3 - h^3$
Ch gutt.		Ng	$c^3 - des^3$
a) dunkelstes	$es^2$	P	$cis^3 - a^3 (f^3)$
b) wie in Ach	$b^2 - es^3$	R ling.	$Fis_1, e^2 - f^3 (c^3)$
Ch pal.	$d^4 - g^4$	S	$g^3 - b^3$
D	$g^3$	Sch	$g^2 - c^4 (fis^3)$
F	$f^3 - g^3$	T	$g^3 - a^3$
G	$a^3$	W	$d^3$
H	$e^3$		
K	$e^2 - a^3 (d^3)$		
L	$f^2 - a^2, as^3 - b^3$		

Das Überwiegen der 3-gestrichenen Oktave, und namentlich deren mittlerer Gegend, ist auffallend. Bei mehreren Konsonanten mit sehr weiten Grenzen der möglichen Tonhöhen ist in Klammern die mittlere Höhe beigefügt, wie man sie beim bequemen stimmlosen Aussprechen zu wählen pflegt. Genau betrachtet, beruhen die Variationen darauf, daß diese Konsonanten, auch wenn man sie so isoliert als möglich gibt, doch immer in einen, wenn auch nur kurz angedeuteten Flüstervokal auslaufen. Die mittlere Höhe entspricht eben einem mittelhellen Vokal. Schickt man den Vokal voraus, so ist die Wirkung dieselbe. Desgleichen macht es keinen wesentlichen Unterschied, ob die Konsonanten mit stimmhaften oder mit geflüsterten Vokalen verbunden werden. Deshalb gilt unsere Tabelle in der Hauptsache auch für Konsonanten im Zusammenhange des gewöhnlichen Sprechens. Natürlich kommt hier auch noch die augenblickliche Höhe des Stimmtones hinzu.

## 6. Bemerkungen über einzelne Konsonanten.

Ch hat die größte Variationsbreite. Vom dumpfsten „Uch“ bis zum spitzigsten „Ich“ umfassen seine Tonhöhen den Gesamtbereich der Flüstervokalhöhen, ja sie reichen noch über die des isolierten Flüster-I hinaus. Die beiden im vorigen Abschnitt untersuchten Formen, Ch gutturale und Ch palatale, sind noch besonders aufgeführt. Bei dem gewöhnlichen Ch gutt. sind 2 gleichzeitige Töne, neben dem höheren Hauptton ein tieferer Unterton, bemerkbar.

Für H ist die Höhe angegeben, die es bei möglichst vokalfreier Aussprache hat. Freilich ist es nie ganz vokalfrei. Schon PURKINJE hat (nach BRÜCKE) bemerkt, daß in das H die qualitativen Verschiedenheiten der sämtlichen Vokale und aller übrigen Laute eingehen.

Von den Explosivlauten hat K die größte Variationsbreite. Die des P dürfte eine Sexte, die des T einen Ganzton nicht viel überschreiten (man flüstere z. B. ut—et). Allerdings mögen hier, wie auch sonst bei den Konsonanten, individuelle Verschiedenheiten eine Rolle spielen. Doch gilt wohl allgemein, daß P und besonders T mehr als K gegenüber den damit verbundenen Vokalen ihre Selbständigkeit bewahren.

L, das wir akustisch richtiger den Vokalen zurechnen zu müssen glauben, hat neben dem hauptsächlich hervortretenden  $as^3—b^3$  einen leisen Unterton  $f^2—a^2$ , ähnlich den Untertönen der hellen Flüstervokale und identisch mit dem Ton eines geflüsterten O.

Zur Höhenbestimmung der Nasalkonsonanten vergleiche man sie nicht nur untereinander, sondern auch mit Flüstervokalen, z. B. N mit Ä, dann mit O: im 1. Falle sind die Tonhöhen identisch, im 2. gibt N die höhere Oktave von O. Ng, M, N kann man als aufgelösten Dreiklang  $c^3$ ,  $es^3$  ( $e^3$ ),  $g^3$  flüstern.

Die stark veränderliche Tonhöhe des stimmlosen Zungen-R ist nicht leicht zu erkennen. Man findet sie besser, wenn man ohne sonstige Veränderung der Mundhöhle ein Sch nachschickt. In der Tat ist im R ein leises Sch von seiner Tonhöhe enthalten. Zu dieser Geräuschhöhe kommt nun der tiefe Brummtton  $Fis^1$  hinzu, der vom Wechsel des hohen Tones unberührt bleibt. Er ist augenscheinlich ein Oberton des durch die Zungenschwingung direkt erzeugten, an sich unhörbaren Grundtons. Da die Zahl der Unterbrechungen nach HERMANN'S und späteren Messungen bei einem guten Zungen-R 20—25 beträgt, also der Grundton zwischen  $E_2$  und  $Gis_2$  liegen muß, so stimmt die gehörte Höhe gut mit dieser Deutung überein. Übrigens könnte der Ton auch  $Fis$  sein, wäre dann also



erst der 4. Teilton. Ihn als „Unterbrechungston“ zu bezeichnen, scheint mir nicht korrekt, auch abgesehen davon, daß die sog. Unterbrechungstöne nach den Untersuchungen von SCHAEFFER und ABRAHAM sowie von F. A. SCHULZE als objektive Töne anzusehen sind. Jedenfalls wird hier nicht ein Klang unterbrochen, sondern durch die Zungenschwingungen ein Klang erzeugt, wie von einer schwingenden Metallzunge oder dem Propeller eines Luftschiffes. Zu dem Gesamtcharakter des Lautes trägt dieser tiefe Ton entschieden bei, indem er etwas U-artiges hineinbringt.

Im S sind außer den klar erkennbaren Geräuschhöhen der 3-gestrichenen Oktave auch noch gelegentlich, besonders bei ganz leisem Flüstern, einzelne Töne der 6-gestrichenen Oktave zu hören. In einem besonderen Falle konnten wir einen solchen als  $d^6$  bestimmen (s. S. 125 ff.). Aber unter gewöhnlichen Umständen versagt in dieser Höhe nicht nur das absolute Tonbewußtsein, sondern auch die Vergleichung mit tieferen Tönen zur Intervallbestimmung. Innerhalb des eigentlichen Zischgeräusches, des die 4-gestrichene Oktave ausfüllenden Formanten, sind für mein Ohr weder einzelne Töne noch Geräuschteile von deutlich ausgesprochener Höhe erkennbar. Nur unter besonderen experimentellen Umständen scheint die wahrnehmbare Geräuschhöhe von  $a^3$  noch bis  $g^4$  oder  $d^5$  zu steigen.

Auch hier seien zur Vergleichung noch die Angaben anderer Beobachter beigefügt, die stimmlose Konsonanten nach dem unmittelbaren subjektiven Gehörseindruck untersuchten. Sie berücksichtigten aber fast nur die Reibelaute. Wo mittlere Werte angegeben wurden, sind sie wieder in Klammern beigefügt.

	MICHAELIS (1)	TRAUTMANN	LLOYD (3)
Ch insgesamt	—	—	$c^2 - fis^4$
Ch gutt.			
dunkelstes	$es^2 - e^2$	} $d^3 - f^3$	$fis^3$
helleres	$g^2 - h^2$		
Ch pal.	$b^4 - h^4$	$g^4$	$es^4$
F	$f^4$	$f^3$	$f^2 - f^4$ ( $f^3$ )
H	—	—	$d^3 - f^4$ ( $e^3$ )
R ling.	—	$f^3 + a^3$	—
S	$a^4 - b^4$	$a^3$ (mittleres)	} $g^2 - g^4$ ( $e^4$ ) ( $c^4$ )
Sch	$d^4 - es^4$	$a^3 + d^4$	
W	—	$e^3 + g^3$ (engl. = U = $g^2$ )	—

Die Angaben stimmen mit denen des Verfassers im ganzen recht gut überein. In einigen Fällen mögen Oktaventäuschungen

nach der Höhe zu vorliegen, so bei der Notierung des F, S, Sch durch MICHAELIS und des Ch durch LLOYD<sup>1)</sup>.

Neuerdings hat PAGET (2), wie von Flüstervokalen, auch von stimmlosen Konsonanten auf Grund direkten Hörens „Resonanzhöhen“ angegeben, um sie dann mit Hilfe darauf abgestimmter komplizierter Hohlräume künstlich nachzubilden. Er glaubt überall eine Mehrzahl von Tonhöhen zu beobachten. So bei L:  $b-as^1$ ,  $f^2$ ,  $as^3-h^3$  (nach den von ihm angegebenen Schwingungszahlen der chromatischen Leiter mit  $C = 64$ ); bei Ng:  $as-b$ ,  $cis^2-fis^2$ ,  $c^3-fis^3$ ; bei S:  $a-as^1$ ,  $b^2$ ,  $cis^3-d^4$ ,  $fis^5-a^5$ . Auch hier finden sich zumeist erfreulich genaue Koinzidenzen mit unseren Angaben: so bei L die 2. und 3., bei Ng und S die 3. der angegebenen Höhen. Aber die tiefsten Resonanzen in der kleinen bis 1-gestrichenen Oktave („throatresonances“) erscheinen mir sehr zweifelhaft. Einigermaßen kräftige Geräuschteile von dieser tiefen Lage sind nach den Interferenzversuchen in diesen Konsonanten nicht enthalten; schließt man alles darüber Liegende aus, so hört man überhaupt nichts mehr. Nun wäre zwar denkbar, daß sehr schwache tiefe Geräuschteile durch Röhreneinstellungen auf höhere vernichtet würden. Aber wenn sie so schwach sind, wird eben die Feststellung ihrer Tonhöhen innerhalb des Gesamtgeräusches durch direktes Heraushören kaum möglich sein. Man vermißt hier wie bei den Vokalen Aufklärungen über das methodische Vorgehen und die Kontrollmaßregeln. Schließlich könnte aber auch die englische Aussprache hier Unterschiede bedingen.

Weitere Vergleichungspunkte bietet die Tabelle der HERMANNschen „Formanten“ (o. S. 130), insbesondere für die in stimmloser Form geprüften: Ch, F, S, Sch. An diesen zeigt sich, daß die von HERMANN gefundenen Maxima mit den subjektiven Tonhöhen unserer Tabelle im wesentlichen zusammenfallen. Bei den stimm-

<sup>1)</sup> LLOYD fand die Bestimmung der Oktavenlage durch das bloße Ohr überhaupt unmöglich und benutzte einen physikalischen Umweg: Messung der Länge einer Röhre, die beim Darüberblasen ein möglichst gleich hohes Geräusch gibt. Aber bei der Vergleichung der beiden Geräusche kann man auch fehlgehen. Ich habe es nicht zu schwer gefunden, die Oktavenlage hier wie bei den Vokalen durch die Verschmelzung mit Klaviertönen zu kontrollieren.

Eigentümlich ist es, daß die Oktaventäuschungen, wenn es sich in den obigen Fällen um solche handelt, bei Konsonanten nach oben gehen, bei Vokalen aber regelmäßig nach unten. Der Unterschied könnte mit der verschiedenen absoluten Höhe zusammenhängen: die Vokale sind im ganzen tiefer, erscheinen auch voluminöser, besonders die dunkleren und zumeist das U, wo die Täuschung am stärksten ist. Aber zunächst liegen zu wenig sichere Fälle von Oktaventäuschung bei Konsonanten vor, um sich darüber Gedanken zu machen.

haften Lauten freilich sind nach dem dort Bemerkten Abweichungen verständlich.

Nach den mitgeteilten Höhen der stimmlosen Vokale und Konsonanten könnte man nun versuchen, die „Flüstermelodie“ eines ganzen Wortes, ja Satzes, in Noten aufzuschreiben, die hier schon durch den Text festgelegt ist. Aber wenn Notierungen lautgesprochener Sätze noch einen gewissen Zweck haben können: hier wären sie ganz nutz- und sinnlos, da diese Tonhöhen so schwach ausgeprägt sind, daß man sie bei dem Vorüberhuschen in der lebendigen Rede unmöglich im Zusammenhange wahrnehmen kann. So würde ein solches Notenbild uns nur den Wesensunterschied von Geräuschen und Tönen aufs neue zum Bewußtsein bringen. Außerdem sind ohne Zweifel auch beim Flüstern die relativen Ruhepunkte der Sprachmelodie unter sich durch stetige Übergänge verbunden. Die einzelnen Laute, Konsonanten wie Vokale, schwanken während ihrer kurzen Dauer in der Nuance, daher auch in der Höhe, wenngleich wohl in geringerem Maße als beim stimmhaften Sprechen. Diese Fluktuationen lassen sich aber in unseren Noten nicht wiedergeben, man müßte denn noch viele besondere Zeichen einführen, wofür es dann doch wieder an Anwendungsmöglichkeiten fehlen würde.

#### 7. Verschiebung der Tonhöhen stimmloser Laute bei Interferenzversuchen.

Bei Ab- und Aufbaureihen zeigen sich in Verbindung mit den Umwandlungen der Geräuschlaute selbst auch Veränderungen ihrer Höhe, die einem bestimmten Gesetze zu folgen scheinen. Diese Frage habe ich nicht systematisch untersucht, aber öfters von den Beobachtern Angaben erhalten, aus denen hervorzugehen scheint, daß die Klanghöhe der Geräusche beim Aufbau (wo also zunächst nur ein unterer Teil vorhanden, die ganze darüberliegende Strecke aber abgeschnitten ist) in der Regel nahe an der oberen Grenze der jeweilig vorhandenen Geräuschmasse liegt. Ist aber ein bestimmter kritischer Punkt erreicht, nämlich die in dem betreffenden Laut (Vokal oder Konsonant) beim interferenzfreien Hören wahrnehmbare Höhe, so bleibt er darauf liegen, auch wenn er sich durch Freigabe weiterer Strecken noch etwas vervollkommnet oder verstärkt. Doch scheint auch schon vorher die Tonhöhe nicht durchweg stetig zu steigen, sondern öfters sprungweise, indem sie etwa in einem Unterformanten eine Weile liegenbleibt (s. die Beobachtungen beim L, S). Bis zum kritischen Punkte liegt die Tonhöhe zunächst etwa 1 Ton unterhalb der der

Einstellungsgrenze entsprechenden Tonhöhe, dann steigt diese Differenz bis zu einer Quarte (infolge der wachsenden If.-Breite). Beim Abbau ist es natürlich umgekehrt: die Veränderung der Tonhöhe beginnt erst, wenn der kritische Punkt nach unten hin überschritten ist.

Beispielsweise war die Tonhöhe für U, O und A bei Einstellung bis 18 cm (wobei also das Tonbereich etwa von  $a^1$  aufwärts abgeschnitten war) gemeinschaftlich  $a^1$ , bei Einstellung bis 12 es<sup>2</sup>. Die Höhe eines hellen Sch reichte für den Beobachter ABRAHAM, der infolge seines hervorragenden absoluten Gehörs öfters von selbst Veranlassung nahm, sich darüber zu äußern, beim Aufbau bis  $h^3$  oder  $c^4$  (für mich lag sie immer etwas unter den von ihm angegebenen Höhen und blieb zuletzt auf  $b^3$  liegen). Beim Aufbau eines scharfen S war die Tonhöhe für Dr. ABRAHAM bei Einstellung (immer von oben herab) bis 4,6  $g^3$ , bis 2,6  $g^4$  und schien auf  $d^5$  liegenzubleiben, wo freilich auch für das beste musikalische Gehör das Höhenurteil unter so schwierigen Umständen unsicher wird. Ch pal. hatte bei Einstellung bis 13 die Höhe  $b^1$ , rückte dann bis  $fis^4$  hinauf und blieb dort liegen. Bei einer Einstellung bis 5 war die Tonhöhe für Sch, S, F, Ch gleichmäßig  $d^3$  usf.

Diese Beobachtungen fügen sich, wenn man die Tabelle S. 158 zum Vergleich heranzieht, der obigen Regel, ausgenommen die beim S, dessen Tonhöhe im unversehrten Zustand danach  $g^4$ , ja  $d^5$  sein müßte. Wenn nicht zufällig bei diesem Versuch ein überscharfes S gegeben wurde, könnte die Erklärung darin liegen, daß das unversehrte S mehrere Geräuschhöhen besäße, eine bei  $g^3-b^3$ , eine bei  $g^4-d^5$ , wo das eigentlich Zischende zustande kommt, und daß für manche Beobachter die tiefere, für andere die höhere hervorträte; ähnlich wie es beim geflüsterten A zu sein scheint. Bei Aufbauversuchen, wo die Aufmerksamkeit auf das neu Hinzukommende gerichtet ist, wäre es möglich, daß die höchsten Geräuschlagen, deren Tonhöhe sich sonst für die meisten nicht gesondert abhebt, bemerkbar werden.

Bei Stichversuchen mit Flüstervokalen wurden ebenfalls gelegentlich nicht uninteressante Tonhöhenänderungen beobachtet. So rückte die Höhe eines sehr dunkel geflüsterten U, wenn man mit Stichen von Einstellung 13 allmählich bis zu 20 übergang, von  $d^2$  bis  $g^2$  und  $a^2$  hinauf, ging dann aber von 22 ab wieder auf  $f^2$  und  $d^2$  zurück. Dies ist wohl so zu deuten, daß zunächst durch Herausnahme eines tieferen Geräuschteils aus dem Formanten eine Erhöhung eintritt, dann aber, nachdem die untere Formantengrenze nach unten überschritten ist und somit Stichversuche zunächst keine Veränderung mehr bewirken können, die Tonhöhe wieder auf den ursprünglichen Stand des unversehrten U zurückkehrt.

Stellt man auf die einem Flüstervokal zugehörige Tonhöhe selbst ein, so wird er keineswegs vernichtet, wie man erwarten müßte, wenn die Tonhöhen (gemäß der seit DONDERS herrschenden Anschauung) mit den Formanten oder Formantzentren identisch wären. Der Vokal wird nur ganz wenig verändert. Dies leitet uns zur nächsten Frage:

## 8. Deutung der gefundenen Tonhöhen.

Schon ein Blick auf das Schema S. 107, worin die Tonhöhen der mit Interferenz untersuchten stimmlosen Laute durch Kreuze bezeichnet sind, läßt die auffällige Tatsache erkennen, daß sie sich in einem verhältnismäßig engen Bezirk, zwischen  $f^2$  und  $f^4$ , zusammendrängen, und daß weitaus die meisten, besonders die der Konsonanten, der 3-gestrichenen Oktave angehören. In der Tabelle S. 158, die noch mehr Konsonanten umfaßt, tritt dies noch umfassender zutage. Viele Konsonanten können danach überhaupt in gleicher Höhe gesprochen werden. Ja es können auch Konsonanten mit Vokalen (wie N, L, S mit A) dieselbe Höhe haben. Unmöglich also, in diesen Tonhöhen die Formanten zu erblicken. Stehen sie nun ganz außer Beziehung zu diesen, und welches ist ihre positive Bedeutung?

a) Wir betrachten zuerst die Vokalhöhen im Anschluß an das Schema S. 107. Die Höhen liegen für U, O, A an ihren oberen Formantgrenzen, für Ö in seiner Formantmitte, für Ä bis I dicht an ihrer unteren Formantgrenze. Die Erklärung dieser anscheinend gesetzlosen Lagen ergibt sich, wenn ich recht sehe, aus den Formantzentren der gesungenen Vokale in Verbindung mit einer zunächst vielleicht gewagt erscheinenden Hypothese<sup>1)</sup>. Die Formantzentren der gesungenen U, O, A liegen ziemlich genau eine Oktave tiefer als die Töne der geflüsterten. Denn das gesungene U hat sein (bewegliches) Formantzentrum von  $g^1$  abwärts, O das seinige bei  $g^1-c^2$ , A um  $g^2$ , die Flüsterhöhen aber sind bzw.  $b^1-f^2$ ,  $g^2-b^2$ ,  $f^3-g^3$ . Da nun, wovon sich jeder leicht überzeugen kann<sup>2)</sup>, die Mundhöhle beim Singen und Flüstern eines Vokals genau in gleicher Weise eingestellt ist, also dieselbe Resonanz gibt, so müssen eben die Flüstertöne von U, O, A die ersten Obertöne, also sozusagen Überblasungsgeräusche der jeweiligen Mundhöhlengeräusche sein. Wir wissen, daß bei vielen Instrumenten nur Überblasungstöne zum Vorschein kommen. Wenn es nun auch zuerst seltsam klingt, beim Flüstern von Überblasung zu reden, so wird die Mundhöhle beim Flüstern doch wohl immer noch stärker angeblasen als beim Singen und sind wir über die Bedingungen des Überblasens unter diesen besonderen Umständen noch zu sehr im unklaren, als daß ein Einwand daraus hergeleitet werden könnte.

Von hier aus ergibt sich auch eine wahrscheinliche Erklärung für die von O. ABRAHAM beobachteten Flüsterhöhen des O, AO und A, für die

<sup>1)</sup> Die Frage, warum die Formanten der tiefen Flüstervokale gegenüber denen der gesungenen etwas nach oben verschoben sind, kann hier beiseite bleiben. Wir behandeln sie im 13. Kapitel.

<sup>2)</sup> Vgl. auch GARTEN 3, VIII, S. 18.

man, da er diese Vokale heller als die meisten meiner Versuchspersonen gibt, höhere Töne erwarten müßte, während er tiefere an sich beobachtet hat. Wahrscheinlich kommt es daher, daß er eben nicht überbläst, sondern das Grundgeräusch seiner Vokale flüstert. Seine Töne sind in der Tat einfach die Formanten, nicht deren höhere Oktaven. Und so dürfte auch bei v. HORNBOSTEL dasselbe der Fall sein.

Zugleich wird aber jetzt besonders deutlich, warum man die Formanten der stimmhaften Vokale nicht primär aus den Flüsterhöhen ableiten darf. Der Verfasser selbst geriet infolge dieser Annahme lange Zeit in ein peiniges Dilemma, da das gesungene A ein Resonanzmaximum bei  $g^2$  aufwies, während die Flüsterhöhe eine Oktave höher lag. Seltsamerweise scheint diese Paradoxie bisher niemand aufgefallen zu sein, obgleich die beiden Tatsachen vielen bekannt waren.

Bei den übrigen Vokalen, vom Ö an, decken sich die beobachteten Höhen ohne weiteres mit den Formantzentren der gesungenen Vokale. Hier handelt es sich also um die Grundschwingungen der durch das Flüstern angeblasenen Mundhöhle. Für die hellsten Vokale, Ä, Ü, E, I, ergibt sich so dieselbe Regel wie für die meisten musikalischen Klänge: wie dort der Grundton die wahrgenommene Tonhöhe bestimmt, so hier der unterste Teil des Geräuschformanten. Wahrscheinlich besitzt diese Stelle der Geräuschschwingungen eine besondere Stärke oder Dichtigkeit.

Ich habe auch versucht, Resonanzgabeln durch Flüstervokale zum Mitschwingen zu bringen, nicht (wie bei den gesungenen) zum Zwecke der Analyse, sondern nur, um die Ergebnisse der Formant- und Tonhöhebestimmungen dadurch zu kontrollieren. Geflüstertes U brachte  $f_{is}^1$ , ein sehr dunkles  $c^2$ ,  $dis^2$ ,  $e^2$  schwach zum Mitschwingen. Hier entspricht  $f_{is}^1$  wahrscheinlich direkt der Mundresonanz, die höheren Töne aber ihrer jeweiligen Oktave, wie sie in den Flüsterhöhen gehört wird. Auf O resonierte  $a^1$ , aber auch  $g^2$  und sogar (bei hellerem O)  $h^2$ ; was ebenso zu deuten sein wird. Auf offenes O (Oa) resonierte am stärksten  $f^2$ , auf Ao  $h^2$ . Auf A schwingen alle Gabeln von  $f^2$  bis  $f^3$ , aber alle recht schwach. Wurde besonders stark geflüstert, mehr gehaucht oder gefaucht, so resonierte am stärksten  $h^2$  (stärker allerdings auf Ao als auf ein helles A). Hier zeigt sich jene diffuse Resonanz, von der wir oben sprachen, wie sie auch dem unmittelbaren Eindruck entspricht; und zwar ist sie über die ganze Formantregion ausgedehnt, nicht bloß auf einen Grundton und seine Oktave. Bei Ö resonieren schwach  $a^3$  und  $b^3$ : die Flüsterhöhen in ihrer wirklichen Lage.

Man kann aus diesen Versuchen auch schließen, daß beim Flüstern der dunklen Vokale der der Mundresonanz entsprechende Grundton im Luft- raume stark genug vorhanden ist, um noch eine sehr empfindliche Gabel anzuregen, aber nicht stark genug, um selbst noch aus dem Geräusch als maßgebende Tonhöhe herausgehört zu werden. Auch beim A macht sich dies noch insofern geltend, als die für das Ohr (wenigstens das des Verfassers) maßgebende Tonhöhe eine Oktave über dem untersten Ton des Resonanzbereiches liegt.  $h^2$  als Resonanzton eines Ao bleibt allerdings dabei etwas befremdlich.

Im übrigen ist mit Resonanzversuchen bei den Flüstervokalen begreiflicher Weise nicht viel zu erreichen.

b) Bei den Konsonanten, soweit sie hier untersucht wurden, liegen die beobachteten Tonhöhen zum Teil wieder am unteren Ende ihrer Formanten. Es gilt also insoweit dieselbe Regel und Deutung wie bei den hellen Flüstervokalen. Bei anderen dagegen, nämlich den nasalierten und den Zischlauten F und S, liegen die für mich deutlich bemerkbaren Tonhöhen mehr oder weniger tief unter dem Formanten. Bei S und F dürfte es sich um „Schneidentöne“ handeln, die durch das Vorbeistreichen der Luft an den Zähnen entstehen, ähnlich wie solche bei zu schwachem Anblasen der Galtonpfeife auftreten. Bei den nasalierten Konsonanten dürfte der gehörte Ton durch Anblasen der Nasenhöhle, bei L durch Anblasen des infolge Vorlegung der Zunge gebildeten Hohlraums zustande kommen.

Das beschriebene Emporsteigen der gehörten Tonhöhen beim Aufbau der stimmlosen Laute versteht sich im allgemeinen, wenn man annimmt, daß die Tonhöhe durch das jeweilig höchstliegende in dem Geräusch vorhandene relative Maximum bestimmt wird. Mit der Erreichung des Formanten ist zugleich das letzte Maximum erreicht, und so wird die Tonhöhe da liegenbleiben.

Rückblickend sehen wir, daß die subjektiven Tonhöhen der stimmlosen Laute zwar nicht als primäre Kennzeichen für die Formanten dienen können, von denen sie in verschiedenen Fällen bedeutend abweichen, daß sie aber zumeist in verschiedenen Beziehungen zu den Formanten stehen und wahrscheinlich auch Stellen besonderer Stärke (Dichtigkeit) des Gesamtgeräusches entsprechen. Ist demnach ihre methodische und theoretische Bedeutung nicht eine so grundlegende, wie viele Forscher glaubten, so verdienen sie doch nach der praktischen Seite die Beachtung der Sprachforscher: denn sie ermöglichen es, die verschiedenen Nuancen des nämlichen Vokals oder Konsonanten kurz und eindeutig (wenn auch nicht erschöpfend) zu definieren. Darauf ist denn auch von Phonetikern und Linguisten bereits mehrfach (LLOYD, BREMER, TRAUTMANN, THOMSON u. a.) hingewiesen worden. Für die allgemeine Theorie der Formanten scheint aus den Betrachtungen dieses Kapitels hervorzugehen, daß die für den Charakter eines Sprachlautes entscheidendste Stelle nicht notwendig mit dem absoluten Maximum der physikalischen Amplitude oder Intensität zusammenzufallen braucht. Immer wird natürlich der Formant eine Strecke sein, die ein relatives Maximum enthält; aber daneben können andere Strecken in der Schwingungsmasse vorhanden sein, die gleichfalls relative Maxima, vielleicht sogar das absolute Maximum, enthalten, aber weniger bedeuten, weil ihre Tonlage nicht die ausschlaggebende ist.

## 7. Kapitel.

# Synthetische Darstellung der stimmhaften Vokale.

### I. Historisches.

Vollkommen durchsichtig und endgültig läßt sich die Struktur gesungener (und daher auch stimmhaft gesprochener) Vokale nur feststellen durch künstliche Zusammensetzung aus ihren letzten Elementen, den einfachen Tönen. Diesen Weg hat bereits HELMHOLTZ beschritten, indem er die Vokale durch 8 (später 12, bis  $b^3$  reichende) elektromagnetisch erregte Stimmgabeln, deren Stärke durch vorgesetzte Resonatoren von veränderlichem Abstände abgestuft werden konnte, nachbildete. Aber die Einrichtung, deren Bestandteile jetzt im Deutschen Museum zu München aufbewahrt werden, lieferte nach HELMHOLTZ' eigenen Bemerkungen sowie Berichten von Ohrenzeugen doch nur Annäherungen, besonders hinsichtlich der helleren Vokale. Überdies sind die Klänge elektrischer Gabeln keine ganz einfachen Töne, auch wenn Resonatoren vorgeschaltet werden<sup>1)</sup>.

Herr Prof. F. AUERBACH, der noch selbst mit HELMHOLTZENS Einrichtung gearbeitet hat, schreibt mir darüber: „Die Vokale kamen im allgemeinen nicht sehr überzeugend heraus; nur wenn man alle Faktoren ganz minutiös variierte, tauchte ganz plötzlich einmal der betreffende Vokalklang verblüffend auf, ohne daß es indessen gelungen wäre, ihn exakt zu fixieren und später wieder aufzufinden.“ Ähnlich lautete ein mündlicher Bericht Prof. HAGENS.

HELMHOLTZENS Originalgabeln standen mir vor etwa 30 Jahren zur Verfügung, aber ohne die große Unterbrechungsgabel; weshalb ich sie zu anderen Studien verwandte (s. 12. Kap.). 1922 sah ich eine Nachbildung mit dem Grundton  $c$  durch eine Berliner Firma; aber die Gabeln reichten nur bis  $e^3$ , so daß bestenfalls nur U, O, A zu erzielen waren, die Stärke-regulierung war schlecht, kein  $ppp$  zu erzielen, und die Gabeln hatten zahlreiche Obertöne (die  $c$ -Gabel mindestens 5), die teils mit bloßem Ohr, teils durch Schwebungen von Hilfgabeln konstatierbar waren. Bei  $c^1$  war, wenn es stark erklang, die Oktave sogar stärker als der Grundton. Mit einem solchen Apparat ist natürlich keine gute Synthese möglich.

<sup>1)</sup> Wie ich 1896 zeigte (3, S. 673), enthalten Klangquellen, die für einfach galten, wie Resonanzgabeln, elektromagnetische Gabeln, KOENIGS Wellensirene, noch mehr oder minder zahlreiche Obertöne.



Vermutlich hat ZAHM, der (nach MILLER S. 246) HELMHOLTZ' Synthese nachmachte und sie für mehr oder weniger „fanciful“ erklärte, einen derartigen Apparat benutzt. Denn immerhin: ein HELMHOLTZ läßt sich nicht durch bloße Phantasien betören. Es ist kein Zweifel, daß seine Einrichtung in seiner Hand gute, zur Bekräftigung seiner Theorie hinreichende Annäherungen ergab, ebensowenig wie daß er selbst sich über die noch verbleibenden Mängel keiner Täuschung hingab. Nach AUERBACHS Beschreibung scheint es wesentlich an der Verschmelzung zu einer Klangeinheit gefehlt zu haben, die oft durch winzige Verschiebungen der Stärkeverhältnisse bedingt ist. (Vgl. 11. Kap., woselbst auch über JAENSCHS Kritik der HELMHOLTZschen Synthesen.)

1880 hat DWELSHAUVERS eine ähnliche Einrichtung noch einige Töne höher hinaufgeführt. Merkwürdigerweise ist dann aber über ein halbes Jahrhundert lang kein wesentlicher Fortschritt erzielt worden. 1885 kombinierte LAHR einige gestrichene Resonanzgabeln (eine höchst unvollkommene Methode), 1909 machte v. WESENDONK ausgedehntere Versuche mit angeblasenen Flaschen, aber ohne Ausschluß der Obertöne und ohne genügende Stärke-regulierung. Doch erhielt er mehrfach bemerkenswerte Ergebnisse. HERMANN versuchte es einmal mit Pfeifen, fand aber die Synthese unausführbar. Desgleichen der Musikprofessor VOLBACH, der aus seinem Mißerfolg sogar auf die Unrichtigkeit der HELMHOLTZschen Vokal- und Klangfarbenlehre zurückschloß. Endlich erklärte auch JAENSCH 1914 auf Grund eigener Versuche, daß zwar eine Synthese mit Gabeln möglich sei, eine solche mit Pfeifen aber „im allgemeinen zu einem völligen Mißerfolge führe, indem sie ein rein musikalisches Schallphänomen hervorbringe“. Wenn es sich in beiden Fällen um einfache Töne handelt, ist freilich irgendein Grund für diese prinzipielle Gegenüberstellung nicht zu entdecken.

Man muß sehr wohl unterscheiden zwischen Synthese und bloßer Nachahmung. Seit KRATZENSTEIN (1780) und KEMPELEN (1791) sind zahlreiche Mittel erdacht, um Vokale mehr oder weniger täuschend nachzuahmen. Aber solche Apparate haben im allgemeinen mit Synthese nichts zu tun. NAGEL verkennt die Bedeutung des HELMHOLTZschen Unternehmens, wenn er es in eine Reihe damit stellt. Ebenso muß ich S. GARTEN widersprechen, wenn er neben der Synthese aus einfachen Tönen noch andere Wege für möglich hält. Eine Synthese hat nur Sinn und Beweiskraft, wenn sie auf einfache Töne (Sinusschwingungen) zurückgeht. Erst dann sind die Ergebnisse durchsichtig. Wenn WILLIS eine Feder von veränderlicher Länge an einem laufenden Zahnrad schleifen ließ, wobei von diesem der Grundton, von der Feder die Vokalität bestimmt wurde, oder wenn HERMANN (1890) eine Art von A erzeugte durch Verbindung zweier Klänge der HELMHOLTZschen Doppelsirene, deren 1. Differenzton dem gesungenen

Grundton entsprach, so liegen hier hochzusammengesetzte Tonkomplexe vor und wäre nun erst zu untersuchen, welche von allen diesen Tönen an dem Vokaleindruck beteiligt sind<sup>1)</sup>. Ähnliches gilt von seinen Versuchen mit der sinnreich konstruierten Telefonsirene (1902), die „genau so A sagte wie eine natürliche Männerstimme“: denn auch sie mußte obertonreiche Klänge geben. Reiner war schon sein einfacher Versuch (1911) mit 2 Stimmgabeln,  $e^2 + a^2$  oder  $f^2 + a^2$ , an Stelle der Doppelsirene; diese ergeben wieder einen Differenzton, der als Grund- oder Stimmton gehört wird, während die Gabeltöne den Formanten bilden. Aber dies war schließlich doch nur eine sehr verkürzte Neuausgabe der HELMHOLTZschen Stimmgabelsynthese. In seinen letzten „synthetischen“ Versuchen (1911) blies HERMANN durch eine Lochsirene intermittierend einen Resonator an, der den A-Formanten  $g^2$  gab; die Sirene selbst gab den wechselnden Grundton. „Die Vokalität war so eindringlich, daß ganz Unbefangene sie ohne Befragen erkannten.“ Diese letzten Versuche, durch die HERMANN seine Theorie bewiesen glaubt, sind in mancher Beziehung wertvoll (s. u. S. 194), aber den wahren Weg der Synthese weisen sie nicht. Der eigentliche akustische Tatbestand bleibt so lang unklar, als nicht die nebenher erzeugten Teiltöne festgestellt sind. HERMANN selbst betont des öfteren, daß die Formanten allein keineswegs zur Charakteristik der Vokale hinreichen (so Bd. 91, S. 161). In der Tat lassen sich aus seinen Synthesen nicht einmal die so wichtigen Unterformanten erschließen.

Gerade auch die Frage, auf die es HERMANN und seiner Schule besonders ankommt: ob neben harmonischen auch ebenso wirksam oder gar allein entscheidend unharmonische Teiltöne beteiligt sein können, läßt sich vollständig nur durch entsprechende Kombination einfacher Töne von bekannter Höhe entscheiden. Nach den Ergebnissen unserer Analysen werden wir allerdings zunächst harmonische Töne verwenden; aber nichts steht entgegen, auch unharmonische in den Komplex einzufügen und die Wirkung zu beobachten. Und so lassen sich überhaupt alle prinzipiellen Fragen über Klangfarben, auch instrumentale, erst auf diesem Wege zur vollen Lösung bringen.

Schon kurz nach dem Beginne meiner Vokalstudien, im September 1913, versuchte ich mit Labialpfeifen (die ich der Billig-

<sup>1)</sup> Wir werden im 8. Kap. durch Analyse dieses Versuches als eines Beispiels zeigen, welche Bestandteile dabei zu einem Vokaleindruck zusammenwirken konnten. Die Berichte HERMANNs über seine wiederholten sog. Synthesen siehe in Bd. 47, S. 385ff. 1890; Bd. 48, S. 574ff. 1890; Bd. 56, S. 467ff. 1894; Bd. 91, S. 135ff. 1902; Bd. 141, S. 33ff. 1911.

keit und leichten Handhabung wegen wählte, obgleich die Reinhaltung ihrer Stimmung viel Mühe macht) eine synthetische Einrichtung aufzubauen und hatte sie im Laufe des Jahres 1914 so weit fertiggestellt, daß ich von dem vollen Gelingen überzeugt sein durfte. Aber infolge mehrfacher Motordefekte und der während des Krieges immerfort wachsenden Schwierigkeiten der Reparatur wesentlicher Teile war erst im Herbst 1917 die systematische Durchführung sämtlicher Versuche möglich. Im November berichtete ich der Berliner Akademie darüber und veröffentlichte das Wesentlichste in den Sitzungsberichten vom 4. IV. 1918.

Inzwischen hatte 1916 der amerikanische Physiker D. C. MILLER nach Vorträgen, die er 1914 am Lowell-Institute gehalten, das schon oben (Einleitung) erwähnte Werk veröffentlicht, worin sowohl physikalische Analysen als auch ausgedehnte synthetische Versuche mit Pfeifen beschrieben sind<sup>1)</sup>. Für jeden Vokal baute MILLER eine besondere Zusammenstellung von Pfeifen, eine Art Orgelmixtur. Er hat aber keine Maßregeln ergriffen, um die Pfeifentöne von ihren Obertönen zu befreien. Er gibt an, daß seine gedackten Pfeifen nur minimale Obertöne enthalten hätten; bei geringem Winddruck seien auf den Grundton 99% der Gesamtintensität entfallen. Dies ist ein Punkt, in dem ich Zweifel nicht unterdrücken kann. Mir ist noch keine Pfeife vorgekommen, auch bei den mildesten Orgelregistern, die nicht merkliche Obertöne enthalten hätte (s. 15. Kap.). Ein A, das ich nach MILLERS Verfahren mit gedackten Pfeifen ohne Ausschaltung ihrer Obertöne, aber mit den entsprechenden Stärkeabstufungen der einzelnen Pfeifen herstellte, besaß auch lange nicht die vollkommene Naturtreue, wie sie mit einfachen Tönen erreicht werden kann. Aber ich vermag über MILLERS Einrichtungen nicht aus der Ferne zu urteilen und zolle der Umsicht und Sorgfalt seiner physikalischen Methodik größte Bewunderung.

MILLER erwähnt gelegentlich auch ein neues, in Amerika gebautes Instrument „Choralcelo“, das nahezu einfache Töne gebe (mit elektromagnetisch erregten Saiten) und sich zu Synthesen eigne. Ohne Zweifel sind Pfeifen nicht die einzige oder auch nur beste Tonquelle zu diesem Zweck. Gute elektromagnetische Stimmgabeln in der erforderlichen Anzahl und Höhe (wenn sich so hohe herstellen lassen), besonders aber die durch elektrische Schwingungen vermittelten Töne, die zu beliebiger Höhe geführt werden können, wären an sich wohl vorzuziehen. Aber immer müssen zuerst die Obertöne durch auslöschende Vorrichtungen beseitigt werden.

<sup>1)</sup> Eine Besprechung dieses Buches durch v. WESENDONK in den Berichten der Deutschen physikalischen Gesellschaft vom 30. VI. 1917 kam mir am Tage meines Akademievortrages (1. XI. 1917) zu Gesichte, das Buch selbst erst im Sommer 1918.

## II. Die synthetische Einrichtung.

Das Schema S. 44 gibt die Grundzüge der Einrichtung wieder, wie sie im Psychologischen Institut der Universität Berlin aufgebaut wurde<sup>1)</sup>. Die Abteilung  $S_1$ — $B_1$  kommt aber hier nicht in Betracht; sie ist a. a. O. erläutert.

Das Klangmaterial lieferten 28 Lippenpfeifen, welche die ersten 22 harmonischen Teiltöne des Grundtones  $c$  lückenlos und weiterhin die Teiltöne 24, 26, 28, 32, 40, 48 ( $g^4$ ,  $\bar{a}^4$ ,  $c^5$ ,  $e^5$ ,  $g^5$ ) angaben. Von  $c$  bis  $c^3$  sind sie gedackt, die folgenden offen. Mit diesem Vorrat kann man nicht nur Klänge auf dem Grundton  $c$ , sondern auch auf  $c^1$ ,  $g^1$ ,  $c^2$ ,  $g^2$  herstellen, wie man leicht aus den erforderlichen Ordnungszahlen erkennt<sup>2)</sup>.

Diese Pfeifen werden durch einen elektromotorischen Ventilator angeblasen. Der intermittierende Luftstrom geht zunächst in ein Magazin, von wo aus er unter konstantem Druck den einzelnen Pfeifen zugeführt wird. Der Motor (M) steht im Zimmer IV, die Pfeifen (P) im Zimmer II, so daß das Motorgeräusch sich nicht dem Klange beimischt. Die Pfeifen sind mit einziger Ausnahme der tiefsten ( $c$ ) gegeneinander möglichst isoliert, indem jede in einen Kasten eingefügt ist, der von einem zweiten Kasten umschlossen ist. Zwischen den beiden Kästen liegt Watte. Die Kästen haben nur Öffnungen für die Zu- und Abführung der Luft. Jede Pfeife ist innerhalb des Kastens mit einem entsprechenden Resonator zur Tonverstärkung verbunden. Von diesem aus wird jeder Ton in einer besonderen Röhre durch die Wand zum Zimmer IV geleitet. Hier sind in die sämtlichen Leitungen Interferenzröhren eingeschaltet, die auf die Obertöne der Pfeifen (bei den tieferen von  $c^1$  an auf die Oktave und Duodezime, bei den höchsten nur auf die Oktave) eingestellt sind. Auch in dem darauffolgenden Zimmer mußten noch If.-Röhren angefügt werden, um jede Spur auszulöschen. Bei der tiefsten Pfeife  $c$  waren 5 Obertöne, darunter die

<sup>1)</sup> Da das Institut im Herbst 1919 in das ehemals Königliche Schloß übersiedelte, mußte die Einrichtung abgebrochen und in Anpassung an die neuen Räume wieder aufgebaut werden. Daher trifft das Schema gegenwärtig nicht mehr genau zu; auch die Auswahl der Pfeifen ist etwas geändert. Aber die Beschreibung erfolgt hier nach der alten Einrichtung, mit der die Versuche gemacht wurden.

<sup>2)</sup> In der gegenwärtigen Einrichtung sind die Teiltöne des  $c$  nur bis zum 16. lückenlos, dann folgen die geradzahigen vollständig bis zum 32., dann noch 40 und 48. Die Töne 17, 19, 21 hatten sich als unnötig erwiesen, dagegen war 30 ( $h^4$ ) für die Grundtöne  $g$ ,  $g^1$ ,  $g^2$  erwünscht. Außerdem wurden 2 Leitungen leer gelassen, um beliebige andere Töne, seien es harmonische oder unharmonische, auch etwa Geräusche einzufügen.

bei gedackten Pfeifen sehr starke Duodezime, durch zahlreiche Seitenröhren in III und V (wo sie am wirksamsten sind) zu vernichten. Die Kontrolle über das Vorhandensein von Obertönen erfolgte durch schwebende Hilfsgabeln, bei der Oktave durch Auslöschung des jeweiligen Grundtones (s. o. S. 48). Auf diesen Punkt wurde die allergrößte Sorgfalt verwandt, so daß mit Sicherheit nur einfache Töne im Zimmer V aus den Leitungen kamen.

Ferner wurde die reine Stimmung immer von Zeit zu Zeit kontrolliert und wiederhergestellt. Bei den Pfeifen bis zum Ende der 3-gestrichenen Oktave ist dies unbedingt nötig, bei den tiefsten am nötigsten, während sich bei den höchsten kleine Abweichungen später als unschädlich erwiesen.

Nun war noch eine zweite, nicht minder wesentliche Bedingung zu erfüllen: eine bequeme und ausgiebige Stärkeregulierung. Da bei Lippenpfeifen mit den Änderungen der Anblasestärke zugleich Höhenänderungen eintreten, muß eben die Regulierung erst erfolgen, nachdem der Ton die Pfeife schon verlassen hat und auch schon von seinen Obertönen gereinigt ist (damit nicht bei Verstärkung diese etwa wieder hervorkommen), also am Ende der ganzen Leitung, im Zimmer V bei R. Hier sind die Zinnblechröhren, aus denen die Leitung bis dahin hauptsächlich besteht (in Schläuchen findet weit größerer Energieverlust statt), auf eine kurze Strecke durch Schlauchstücke ersetzt, die durch Schrauben, an denen kleine Metallplättchen sitzen, so zusammengedrückt werden können, daß der Ton von voller Stärke bis zum völligen Verschwinden abgeschwächt werden kann. Die Schlauchstücke müssen von Zeit zu Zeit erneuert werden. Die Schrauben, für jeden Ton eine, sind in einen gemeinsamen viereckigen Rahmen eingefügt.

Die Anblasestärke bleibt also hierbei vollkommen konstant, und jede Höhenänderung der Pfeifentöne ist vermieden<sup>1)</sup>.

Schließlich werden alle Leitungen in einem Gipsverband eng zusammengedrängt und durch einen trichterförmigen abhebbaren Ansatz (T) in eine gemeinschaftliche Mündung übergeführt, aus der der resultierende Klang abgehört werden kann. Wird der Ansatz abgenommen, so hat man die Mündungen aller Einzelleitungen vor sich und kann hier durch kleine Kork- oder Gummistöpsel auch einzelne, oder ganze Gruppen davon, verstopfen und so nach

<sup>1)</sup> MILLER hat sich dadurch geholfen, daß er für jeden Vokal ein besonderes Pfeifensystem baute, bei dem jede Pfeife ein für allemal nur Wind von bestimmter Stärke bekam. Damit ist aber das Ausprobieren aller möglichen Kombinationen und Stärkeverhältnisse, worauf es bei solchen Untersuchungen in erster Linie ankommt, ausgeschlossen; man erhält nur Demonstrationsapparate.

Belieben Lücken- und Stichversuche anstellen, um die Wirkung jedes einzelnen Teiltones auf den Vokalklang zu erkennen. Doch läßt sich natürlich auch schon im Schallzimmer (II) jeder einzelne Ton ausschalten, auch sind Ventile für ganze Gruppen vorhanden<sup>1)</sup>. Man kann ferner mit Hilfe eines in die einzelnen Leitungsmündungen eingeführten Röhrchens sich von der augenblicklichen Stärke eines Teiltone an dieser Stelle überzeugen.

Die Versuche wurden aber fast immer so eingerichtet, daß der vereinigte Klang aus dem Trichter zunächst noch durch einen Schlauch zu einem Zweiwegehahn (H) und von diesem erst durch einen weiteren Schlauch in das Ohr des Beobachters (B<sub>2</sub>) geleitet wurde. Der Hahn konnte durch Drehen um 180° so umgestellt werden, daß eine andere Leitung statt der vorigen geöffnet wurde, durch die ein natürlicher, gesungener Vokal zum Ohr gelangte. Der Sänger stand im Zimmer III bei S<sub>3</sub>, erhielt ein elektrisches Klingelzeichen, und seine Stimme kam durch eine sehr weite Röhrenleitung (Ofenrohr) zum Beobachter herüber. Dieses Verfahren wurde aus sogleich anzugebenden Gründen zumeist dem direkten Angeben des Lautes im Beobachtungszimmer vorgezogen. So ließ sich der künstliche beständig mit dem natürlichen Vokal vergleichen und so lange verbessern, bis die Gleichheit erreicht war. Die durch die Analyse gewonnenen Kenntnisse über die Vokalstrukturen dienten dabei als Leitfaden, sonst wäre des Probierens kein Ende gewesen. Durch rechtzeitige Drehung des Hahnes konnte man auch den Stimmeinsatz des Sängers, der sich durch den Apparat nicht nachbilden läßt, abschneiden und nur den zu vergleichenden Klangkörper selbst zu Gehör bringen.

Es ist bei solchen Vergleichen, aber auch bei der isolierten Darbietung eines künstlichen Vokals, von großem Vorteil, die Laute nur kurz, nicht länger als etwa eine Sekunde, anzugeben, indem man den Hahn oder die Schlauchleitung abwechselnd öffnet und schließt<sup>2)</sup>. Bei längerer Dauer macht sich die starre Konstanz der künstlichen Laute gegenüber den kleinen Schwankungen, denen selbst gut ausgehaltene gesungene Vokale in ihrer Stärke, Höhe,

<sup>1)</sup> Die Isolierung der Töne vor ihrer Vereinigung im Hörschlauche wird öfters etwas durch den Gipsverband beeinträchtigt, insofern starke Töne, auch wenn ihre Löcher verstopft sind, durch den Gipsverband und die Trichterwände in den Schlauch dringen können. Daher muß man, wenn es sich um sichere und vollständige Ausschaltung starker Teiltöne handelt, die bezüglichen Pfeifen selbst im Schallzimmer abstellen.

<sup>2)</sup> Dies hat auch schon HERMANN (Bd. 141, S. 36) und haben wohl alle bemerkt, die ähnliche Versuche machten. Man hört dann auch leicht bei einem A „Pa—pa“ oder „Ma—ma“, je nach der Raschheit des Schließens und Öffnens.

Farbe unterworfen sind, doch geltend. Es ist eben der Unterschied des Lebendigen vom Toten. Man könnte ja auch versuchen, dem künstlichen Laut Leben und Seele durch Nachahmung des Vibrando einzuhauchen. Aber dieser Kunstgriff der Orgelbauer wäre hier, wo es nur auf die Klangfarbe, die Vokalität als solche ankommt, zwecklos. Daher ist es das Richtige, die Laute nur ganz kurz anzugeben, wobei sowohl die Gleichheit wie die Unterschiede in dieser besonderen Richtung klar hervortreten.

Es gelang nun in der Tat, mit dieser Einrichtung innerhalb einiger Wochen alle 8 Vokale der deutschen Sprache naturgetreu nachzubilden. In nicht wenigen Fällen urteilten die beigezogenen Beobachter, daß der künstliche den natürlichen Vokal sogar an Charakteristik und Reinheit übertreffe. Dies war besonders der Fall beim I, das von Männerstimmen nur selten scharf und hell genug intoniert wird und meist eine starke Ähnlichkeit nach E hin hat. Auch die Unterschiede der Vokale auf  $c^2$  kann man oft besser, deutlicher hervorbringen, als es der menschlichen Stimme möglich ist, die ihrem Resonanzapparat eben nicht jede beliebige Einstellung erteilen kann.

Für die Stärkebestimmung der einzelnen in einen Vokal eingehenden Teiltöne bediente ich mich der schon bei den Resonanzversuchen angegebenen subjektiven Kategorien. An sich wäre auch eine physikalische Messung sehr erwünscht und wird hoffentlich später nachgeholt. Aber vorläufig sind die damit verbundenen Schwierigkeiten noch außerordentlich groß, auch standen mir die technischen Hilfsmittel nicht zu Gebote. Die LEWINSche Methode (o. S. 17) erfordert für jeden Ton eine besondere Membran, was zu den größten Weitläufigkeiten führen würde, und dann wäre immer noch die Frage, ob die Ergebnisse für die verschiedenen Töne vergleichbar wären. Indessen erwies sich für die Zwecke der Synthese die unmittelbare Schätzung nach dem Gehör als durchaus genügend, da man bei einiger Übung imstande ist, nach den so gewonnenen Stärketabellen einen Vokal in wenigen Minuten wiederherzustellen. In den folgenden Tabellen sind aber die Stärken nicht so angegeben, wie sie sich am Ende der Zinkröhrenleitung bei T mit eingeführtem Sondierungsröhrchen zeigen, sondern wie sie bei  $B_2$  aus der Schlauchleitung ins Ohr gelangen, also da, wo der Vokal mit dem gesungenen verglichen und als ihm gleich erkannt wurde. Zu diesen Bestimmungen wurde jeder Teilton bei unverändertem Stande der Regulierschrauben einzeln angegeben und auf seine Stärke geschätzt. Die absolute Stärke des Gesamtklanges wie seiner Teiltöne ist an dieser Stelle bedeutend geringer als bei T (weshalb in den Tabellen nirgends hohe Stärkegrade vorkommen); wahr-

scheinlich sind sogar gewisse kleinere Verschiebungen der Stärkeverhältnisse damit verbunden (die Regulierungsschrauben müssen für ein gutes A zuweilen etwas anders eingestellt werden, wenn es bei T als wenn es bei B<sub>2</sub> abgehört wird). Die folgenden Zahlen geben also diejenigen subjektiven Stärkeverhältnisse wieder, wie sie einem guten O, A, Ü usf. an dieser Stelle B<sub>2</sub> zukommen.

Zur Vergleichung sei an einem Beispiel erläutert, welche Teiltonstärken genommen werden müssen, wenn der Vokal bei T unmittelbar am Trichter beobachtet und die erforderlichen Teiltonstärken hier bestimmt werden. Ein gutes Ä auf  $c^1$  ergab da folgende Stärken (mit 4 multipliziert):  $c^1$  4,  $c^2$  10,  $g^2$  12,  $c^3$  12,  $e^3$  4,  $g^3$  4,  $\bar{b}^3$  8,  $c^4$  12,  $d^4$  12,  $e^4$  8,  $\bar{f}is^4$  8,  $g^4$  8. Die absoluten Zahlen sind überall größer als bei dem Ä auf dem gleichen Grundton in der unten folgenden Tabelle, aber ihr Steigen und Fallen befolgt dieselbe Regel. Die kleine Nullstrecke dort ist hier durch einen sehr schwachen Ton ( $e^3$ ) besetzt. Dies stimmt damit überein, daß in der unmittelbaren Nähe des Singenden nach dem Zeugnis der Resonanzversuche auf dem Grundton  $c^1$  selbst bei E und I keine Nullstrecken vorhanden sind<sup>1)</sup>.

### III. Ergebnisse.

Wir geben nun die Tabellen für 4 verschiedene Grundtöne (die Stärkezahlen durchweg mit 4 multipliziert). Innerhalb gewisser Grenzen sind natürlich in diesen Zusammensetzungen noch Variationen möglich, da man einen Vokal eben recht verschieden aussprechen kann, da ferner die absolute Stärke der Tongebung, die Entfernung der Hörenden und die Art der Überleitung (durch freie Luft oder Röhren) Modifikationen auch der natürlichen Stimme bewirken. Aber keine Stärkeveränderung eines Teiltönen in den entscheidenden Regionen läßt den Vokal ganz ungeändert, und jeder kleinsten Veränderung des Vokals entspricht eine Verschiebung in den Teiltönen. Nur in den Zwischenregionen, z. B. auf den toten Strecken des Ü, E, I kann man statt der Stärke 0 auch die Stärke 1 (hier also 4) einsetzen, ohne daß die Wirkung merklich geändert wird. Ebenso können beim A auf  $c$  und  $c^1$  über

<sup>1)</sup> In den Stärketabellen für die Vokale bei den unwissentlichen Versuchen u. S. 184, hat Ä auf  $c^1$  gleichfalls keine Nullstrecke und auch sonst teilweise höhere Stärkezahlen, obgleich es am entferntesten Punkte der Leitung, im Raume VI bei Vp beobachtet wurde. Dies kommt daher, daß, um möglichste Gesamtstärke des Klanges und namentlich des als Differenzton auftretenden Grundtones zu bekommen, die höheren Töne alle so stark genommen wurden, als es irgend mit einer guten Charakteristik des Vokales verträglich war.



c

$g^5$									1
$e^5$									1
$c^5$									1
$\bar{b}^4$									4
$\bar{a}^4$								2	4
$g^4$		1		2				2	3
$\bar{f}is^4$		4		3				4	4
$\bar{f}^4$				3				2	0
$e^4$		4		3				6	0
$dis^4$				3	2			4	0
$d^4$		4		5	3			7	0
$cis^4$				5	2			6	0
$c^4$			5	5	3			6	0
$h^3$			5	5	3			0	0
$\bar{b}^3$			5	5	4			0	0
$\bar{a}^3$			6	4	3			0	0
$g^3$		5	8	1	3			0	0
$\bar{f}is^3$		5	6	0	2			0	0
$e^3$		6	0	0	2			0	0
$d^3$		7	0	4	2			0	0
$c^3$		7	0	4	2			2	0
$\bar{b}^2$		7	0	5	2			2	0
$g^2$	5	10	0	6	0			4	0
$e^2$	6	5	0	6	0			2	1
$c^2$	4	6	3	0	6			0	5
$g^1$	5	10	3	6	4			0	6
$c^1$	5	4	1	5	3			0	5
$c$	6	3	0	0	3			4	4
	U	O	A	Ö	Ä	Ü	E	I	

$c^1$

$g^5$									1
$e^5$									1
$c^5$									1
$\bar{b}^4$									4
$\bar{a}^4$								2	4
$g^4$		1		2				2	3
$\bar{f}is^4$		4		3				4	4
$e^4$		4		3				6	4
$d^4$		4		5	3			7	2
$c^4$			4	5	3			0	0
$\bar{b}^3$			4	5	4			0	0
$g^3$		5	7	1	3			0	0
$e^3$	4	6	5	0	2			0	0
$c^3$	4	7	0	4	2			2	0
$g^2$	8	5	10	0	6			0	4
$c^2$	4	6	4	7	6			0	10
$c^1$	7	4	0	2	3			8	5
	U	O	A	Ö	Ä	Ü	E	I	

$c^2$

$g^5$									1
$e^5$									1
$c^5$									1
$\bar{b}^4$									2
$\bar{g}^4$									2
$e^4$				5		3			1
$c^4$			4	6	3	2			0
$g^3$		5	8	6	4	4			0
$c^3$	4	5	8	6	8	2			4
$c^2$	8	8	7	7	5	10			8
	U	O	A	Ö	Ä	Ü	E	I	

$g^1$

$\bar{b}^{4(1)}$									4
$g^4$				1					4
$f^4$				1		3			3
$d^4$			3	6	5	6			5
$h^3$		4	4	4	6	4			4
$g^3$		4	8	5	2	2			2
$d^3$	2	4	6	6	6	0			0
$g^2$	4	5	8	4	10	3			2
$g^1$	8	5	4	3	4	6			6
	U	O	A	Ö	Ä	Ü	E	I	

1) Für das fehlende  $h^4$  konnte hier ohne Schaden  $\bar{b}^4$  oder  $c^5$  eingesetzt werden.

dem Formanten noch die in den Tabellen durch kleingedruckte Stärkezeffern angedeuteten Töne in diesen geringen Intensitäten beigefügt werden. Sie würden dem in den Resonanztabellen gefundenen und dort so ausgeprägten 2. Maximum entsprechen. Aber nötig sind sie nicht und verändern den Vokal kaum merklich. Daher sprechen wir hier nicht von einem Oberformanten.

Vorstehende Tabellen sind unter vielen Synthesen als die besten, besonders typischen ausgewählt<sup>1)</sup>.

Man sieht aus den Tabellen, wie mit steigender Höhe des Grundtones die Zahl der Teiltöne immer mehr abnimmt, wie für den nämlichen Grundton die Intensitätskurven der Vokale sich charakteristisch voneinander unterscheiden, wie die dunkleren Vokale durch tiefere, die helleren durch höhere Teiltöne ausgezeichnet sind, wie bei den helleren vom Ö an zwei Stärkemaxima durch ein Minimum oder eine Nullstrecke getrennt sind, wie ferner die Maxima für verschiedene Vokale sich verschieden verteilen, aber für denselben Vokal auf verschiedenen Grundtönen ihre absolute Höhe beibehalten, wie fast niemals ein einzelner Ton allein formierend wirkt, sondern immer eine Gruppe, innerhalb deren allerdings 1 oder 2 Töne besonders stark sind — kurz, wie alle jene Züge, die im allgemeinen schon durch die Analyse gefunden waren, nun in bestimmterer Gestalt sich darstellen.

Die Unabhängigkeit der Formanten vom Grundton zeigt sich eklatant darin, daß man, um aus einer Einstellung auf *c* die Einstellung für denselben Vokal auf *c*<sup>1</sup> zu gewinnen, oft fast nichts zu tun braucht, als die ungeradzahligcn Teiltöne zu entfernen. Man muß nur noch den jetzt zum Grundton gewordenen Ton *c*<sup>1</sup> und allenfalls auch *c*<sup>2</sup> etwas verändern (meist verstärken). Die Formantregion dagegen kann bleiben, wie sie war. Dies entspricht ja auch der Tatsache, daß man die Mundstellung nicht wesentlich ändert, wenn man etwa ein A zuerst auf *c*, dann auf *c*<sup>1</sup> singt.

Die Lage der Formanten und Nebenformanten (des Oberformanten beim U, der Unterformanten bei Ö bis I) ist aus der Bewegung der Zahlen in den Kolumnen zu ersehen. U hat keinen oder nur einen mit dem Grundton beweglichen Hauptformanten (Näheres darüber im 13. Kap.). Bei den helleren Vokalen von O an haben wir wieder außer den Formanten die uns aus den Interferenz-Versuchen bekannten und schon in den Resonanzversuchen als erste Maxima gefundenen Unterformanten.

<sup>1)</sup> In der Abh. 9, wo diese Tabellen zuerst veröffentlicht wurden, sind versehentlich dem A auf *g*<sup>1</sup> die nämlichen Stärkezeffern zugeschrieben wie dem Ä. Ich bitte die Besitzer jener Abhandlung, die Zahlen nach der gegenwärtigen Tabelle zu berichtigen.

Die wesentliche Funktion der Unterformanten wird auch durch die Synthese bestätigt. Sie stellen gleichsam Untermalungen dar, ja sie könnten mit den chemischen Radikalen verglichen werden. Man braucht nur zu bedenken, daß Ä und Ü den gleichen Formanten haben; den Ausschlag gibt hier der Unterformant, nebenbei allerdings auch noch die sonstige Struktur. Schon eine ganz geringfügige Änderung in der Stärke oder Lage des Unterformanten bedingt eine wesentliche Änderung in der Klangwirkung. Gibt man z. B. dem Ä statt eines OA ein O als Unterformanten, so kann es nicht gelingen.

Aber auch auf den 1. und 2. Teilton kommt außerordentlich viel an; die kleinsten Verschiebungen hierin können den Charakter des Vokals bedeutend ändern, ja vernichten. So darf beim A auf  $c^1$  auch  $c^2$  nur schwach sein, wenn es sich nicht dem O nähern soll. Ein zu dunkles Ä kann man durch Schwächung oder Vernichtung des Grundtons sofort hell und scharf machen, ohne am Formanten etwas zu ändern, usw.

Bemerkenswert ist, daß die Teiltöne von hoher Ordnungszahl, wie sie für die Grundtöne von  $c$  bis  $c^2$  in der 4-gestrichenen Oktave und darüber liegen, nicht lückenlos im Klange vorhanden sein dürfen; wie dies schon v. WESENDONK bei seinen Synthesen beobachtete. Sie liegen zu nah aneinander und geben dem Klange zu viel Schärfe, namentlich auch durch die Schwebungsrauhigkeit (in dieser Höhe werden Schwebungen noch bis über 300 in der Sekunde als Rauigkeit empfunden). Daher sind die in der ersten Einrichtung noch gebrauchten Töne 17, 19, 21 später als unnötig fortgelassen worden<sup>1)</sup>.

Bei den Vokalen auf  $c^2$  ist es auch für die Synthese oft schwer, die Unterschiede herauszubringen, z. B. zwischen U und O. Und doch gelingt es zuweilen sogar besser als bei der natürlichen Erzeugung. So fügt man einem O den 3. Teilton,  $g^3$ , der im natürlichen O auf  $c^2$  enthalten ist, besser nicht bei. Es entsteht sonst ein OA, wie es das auf  $c^2$  gesungene in der Tat ist (vgl. die Tabelle S. 59).

Das langsame Hinaufrücken der Formanten mit steigendem Grundton, wie es die If.-Versuche ergaben, zeigt sich hier nur spurenweise (vgl. z. B. O auf  $c$  und auf  $c^1$ , A auf  $g^1$  und auf  $c^2$ ). Dies hängt teils mit dem viel geringeren Spielraum für die Wahl der Grundtöne infolge der disponiblen Pfeifen, teils mit gewissen Kompensationen zusammen, die bei der Synthese möglich sind.

<sup>1)</sup> Auch schon das A auf  $c$  wird fast besser, reiner, wenn man den 7. und 9. Teilton ( $b^2$  und  $d^3$ ) herausnimmt. Die umliegenden genügen zur Herstellung der Vokalität.

Das folgende Schema veranschaulicht am einfachsten die Lage der Formanten für Grundtöne der mittleren Oktaven, wie sie sich aus den synthetischen Versuchen in Übereinstimmung mit den analytischen ergeben. Diese Gegenden müssen bei der Synthese in erster Linie vertreten sein, wenn der betreffende Vokal herauskommen soll. Die Formantzentren sind durch 2 Sterne, die Nebenformantzentren durch 1 Stern angezeigt. Die Sterne mit darunterstehendem Pfeil bedeuten den nach unten beweglichen U-Formanten, der zugleich Unterformant für Ü und I ist.

Formantzentren

$b^4$							**
$e^4$							*
$d^4$						**	
$c^4$							
$b^3$				**	**		
$g^3$			**				
$c^3$							
$g^2$			**				
$f^2$	*						
$e^2$				*			
$c^2$							
$g^1$	**	**		*		*	*
$c^1$	↓				↓		↓
	U	O	A	Ö	Ä	Ü	I

Hierzu wie zu allen ähnlichen Übersichten ist aber zu erinnern, daß das Formantzentrum eines Vokals sich innerhalb einer gewissen Zone verschieben kann und muß. Es klingt paradox und leuchtet doch, wenn anders die Formanten nur aus harmonischen Teiltönen bestehen, ohne weiteres ein, daß die Verlegung des Grundtons um eine ganze Oktave den Formanten ungeändert lassen kann, daß er sich aber mit dem Steigen der Stimme um

einen halben oder ganzen Ton verschieben muß. Bei völlig starren Formanten würde man, da ein bestimmter Ton nicht zu beliebigen Grundtönen harmonisch sein kann, einen Vokal überhaupt nur auf bestimmten Grundtönen singen können, auf dicht danebenliegenden nicht. Also muß gerade für nur wenig verschiedene Grundtöne Beweglichkeit der Formantzentren innerhalb einer engen Zone, sei es auch nur einer kleinen Terz, zugestanden werden. Hätten wir statt  $c$  als Ausgangspunkt und Grundlage der Synthese  $d$  gewählt, so hätte für den Vokal A statt  $g^2$  eben  $a^2$  stärkster Teilton sein müssen. Und wenn HELMHOLTZ  $b^2$  angibt, so trifft auch dies für den Fall zu, daß der Vokal auf  $b$ ,  $b^1$ ,  $es$ ,  $es^1$  gesungen wird. Eine geringe Erhellung muß freilich damit verbunden sein; aber innerhalb der gesamten Verschiebung der Klangfarbe, wie sie mit jeder Änderung der Klanghöhe verbunden ist, pfllegt sie nicht bemerkt zu werden.

Wollte man die Formantzentren bis auf die Schwingungszahl genau definieren, so müßte man ja auch verschiedene Notenwerte je nach der gewählten Stimmung angeben. Derselben Schw.-Zahl 800 entspricht für die „physikalische“ Stimmung ( $a^1 = 426\frac{2}{3}$ )  $gis^2$ , für die ältere musikalische ( $a^1 = 440$ )  $g^2$ , für die heutige ( $a^1 = 435$ ) eine dazwischenliegende Note.

Auffallen könnte es, daß der bewegliche Hauptformant des U für den Fall, daß es auf  $g^1$  gesungen wird, zusammenfällt mit dem festen Hauptformanten des O. Tatsächlich ist es aber so: das nämliche  $g^1$ , das den Grundton und Hauptformanten eines U bildet, aber auch schon als einfacher Ton deutlichen U-Charakter hat, macht als Oberton eines tieferen Grundtons, z. B. des  $g$ , den Klang sofort zu einem O. Es wirkt vokalisch verschieden als Grundton und als Oberton. Wir werden auf diese und ähnliche theoretisch bedeutsame Tatsachen in allgemeinerem Zusammenhang zurückkommen.

Beim I ist in dem Schema nahe dem Hauptformanten noch ein zweiter Unterformant (Zwischenformant) eingefügt. Dies gründet sich auf die Ergebnisse von Stichversuchen, wie man sie auch mit dem synthetischen Apparat leicht (und ohne Rücksicht auf Nebenwirkungen wie If.-Breite und Mitausschaltung von Multiplis) ausführen kann. In den Zahlentabellen S. 176 tritt es kaum hervor.

#### IV. Unwissentliche Versuche zur Prüfung der Naturtreue.

Da bei synthetischen Versuchen leicht eine gewisse Neigung entsteht, die Naturtreue bereits als erreicht anzusehen, wenn auch nur eine Annäherung erzielt ist, so habe ich nicht nur beständig

andere Beobachter herangezogen<sup>1)</sup>, sondern auch systematisch unwissentliche Versuche angestellt und eine Statistik darüber aufgenommen. Zu diesem Zwecke wurde die Leitung noch in ein weiteres, durch eine türlose Wand getrenntes und gegen den gemeinsamen Korridor durch eine gepolsterte Doppeltür schalldicht abgeschlossenes Zimmer (VI) verlängert und hier dem Beobachter (Vp) immer abwechselnd ein natürlicher und ein künstlicher Vokal gleicher Art zugeführt, jeder nur eine Sekunde lang, aus dem S. 173 angegebenen Grunde. Der Stimmeinsatz des natürlichen Vokals wurde stets dadurch abgeschnitten, daß der Hahn erst aufgedreht wurde, wenn jener bereits intoniert war (was der Versuchsleiter im Zimmer V durch eine in der Zeichnung angegebene Zweigleitung kontrollierte).

Zu jedem Vokal wurden 5 verschiedene Beobachter benutzt, deren jeder 20 Urteile abzugeben hatte. Sie hatten zu urteilen, welchen Vokal sie bei jedem Versuche hörten, ob er vollkommen sei und was ihnen etwa daran mangelhaft erschiene, hatten aber keine Ahnung von der ganzen Einrichtung und ihrem Zwecke. Nur dies eine war ihnen gesagt, daß sie Vokale hören würden. Denn ein so kurzer Klang ohne das Kennzeichen des Einsatzes ist so

<sup>1)</sup> Oft ist auch die Einrichtung Physikern, Medizinern, Sprachforschern demonstriert worden; und hierbei waren mir nicht nur die zustimmenden, sondern auch gelegentliche kritische Bemerkungen wertvoll. Als z. B. ein I etwas Ö-haltig befunden wurde, ergab eine Nachprüfung, daß zufällig die Pfeife für  $g^3$  nicht ganz abgestellt war:  $g^3$  ist aber das Formantzentrum für Ö. Ein anderes Mal wurde das I auf  $c$  nicht recht einheitlich befunden, es ließ sich sowohl als U wie als I fassen. Hier war das Stärkeverhältnis des Unterformanten zum Formanten nicht genau getroffen; es war jenes duale Stadium, das wir schon von den If.-Versuchen her kennen.

Der gewöhnlichste Demonstrationsversuch für flüchtige Besucher bestand in der Herstellung des A auf  $c^1$  durch die Kombination  $c^1 = 0$ ,  $c^2 = 2$ ,  $g^2 = 5$ ,  $c^3 = 3$ ,  $e^3 = 2$  (so bei Untersuchung an der Stelle T mit dem Hörröhrchen. An der Schlauchöffnung bei  $B_2$ , wo die Vokale gehört wurden, sind dann die Stärken etwa 1, 3, 2, 1). Der Grundton  $c$  kommt als Differenzton hinzu. Hierbei war es immer frappant, wie das helle A durch eine leichte Drehung der Regulierschraube für  $c^2$ , wodurch dieses nur wenig stärker wurde, sogleich in AO, Oa überging. Dagegen spielt  $e^3$  nur eine geringe Rolle, kann sogar überhaupt ohne erhebliche Schädigung wegbleiben.

Übrigens kann man selbst im Schallzimmer, wo die Pfeifen noch nicht von ihren Obertönen befreit, aber durch den Einschluß in die Kästen und die Anfügung der Resonatoren schon reiner geworden sind, bei zufällig günstigen Stärkeverhältnissen überraschende Vokalbildungen hören. So ist A in obiger Zusammensetzung auch hier recht gut. Doch ist absolute Reinstimmung hier noch wesentlicher als im Beobachtungszimmer, weil die Obertöne der Pfeifen bei kleinen Abweichungen sofort scharfe Dissonanzen geben. Der Versuch läßt sich hier einer größeren Anzahl von Personen zugleich zeigen, ebenso wie es bei MILLERS Pfeifensystemen der Fall ist.

mehrdeutig, daß eine derartige Einstellung vorausgesetzt wird, wenn überhaupt eine Deutung möglich sein soll. Im gewöhnlichen Leben ist sie beim Sprechen und Singen von vornherein im Hörenden vorhanden, und es ist genugsam bekannt, daß sogar noch viel speziellere Einstellungen das Verständnis fortwährend unterstützen und daß ihr Wegfall sofort das wunderlichste Verhören herbeiführen kann. Um auch eine Übertragung des Urteils von dem natürlichen auf den künstlichen Vokal auszuschließen, wurde in jeder Versuchsreihe regelmäßig der künstliche zuerst gegeben. Die Anordnung war selbstverständlich im übrigen eine durchaus unregelmäßige.

Das Ergebnis war:

1. Der künstliche Vokal wurde fast immer sofort richtig erkannt.
2. Unter den 400 Doppelversuchen wurden in 109 Fällen beide als gleich vollkommen, in 43 Fällen beide als gleich unvollkommen, in 58 Fällen der natürliche als besser, in 190 Fällen der künstliche als besser beurteilt.

Man kann also sagen: in 3 Vierteln sämtlicher Fälle erschien der künstliche entweder gleich gut oder besser als der natürliche. Das letztere kann darum nicht so sehr wundernehmen, weil der natürliche doch durch die lange Leitung etwas leiden muß, wenn er auch erkennbar bleibt, während man den künstlichen eben so herstellen kann, daß er gerade am Ende der Leitung gut herauskommt.

Hiermit ist nun wohl der objektive Beweis geliefert, daß die synthetischen Vokale als naturgetreu bezeichnet werden dürfen.

Einige nähere Angaben über diese Versuche sind vielleicht in allgemeiner Beziehung, besonders in Hinsicht auf die Modalitäten des Erkennens von Vokalen, von Interesse.

Einer Aufforderung in meiner Psychologievorlesung, sich als Beobachter bei Vokalversuchen zu beteiligen, folgten 30 Damen und Herren. Sie wurden zunächst durch Unterscheidungs- und Erkennungsversuche bei so kurz dauernden natürlichen Vokalen auf ihre Beobachtungsfähigkeit geprüft. Dabei zeigte sich, daß nicht weniger als 12 ausgeschieden werden mußten, da sie nicht imstande waren, sicher und fehlerfrei zu urteilen. Manche beurteilten z. B. I als Ä oder Ö oder konnten überhaupt kein bestimmtes Urteil gewinnen. Wieder ein Beweis, wie wenig man ohne weiteres der Beobachtungsfähigkeit sonst intelligenter Personen vertrauen darf (vgl. o. S. 51). Bei länger dauernden oder stärkeren Eindrücken würden sie wahrscheinlich keine besondere Schwierigkeit gefunden haben. Die übrigen 18 aber erkannten sofort alle aus der Leitung S<sub>3</sub> kommenden Vokale richtig; nur selten kam, was ja bei den besten Beobachtern unvermeidlich ist, ein zufälliger Fehlgriff vor.

Die Tonhöhe war durchgängig c<sup>1</sup>. Jedem Vokal wurden 5 Versuchsreihen gewidmet, jede davon aus 10 Doppelurteilen (über die natürliche und die künstliche Form) bestehend. An diesen 5 Reihen beteiligten sich immer 5 verschiedene Beobachter, die aus den obigen 18 in verschiedener Kombination ausgewählt wurden, um individuelle Eigentümlichkeiten auszuscheiden. Die Zeitfolge des natürlichen (n) und des künstlichen (k)

Vokals wechselte unregelmäßig zwischen 2 Schematismen, innerhalb deren die Verteilung wieder unregelmäßig war, nämlich: „kn, nk, nk, kn, kn, nk, nk, kn, kn, nk“ und „kn, nk, nk, kn, nk, kn, kn, nk, nk.“ Die Instruktion lautete: „Sie werden Vokale von sehr kurzer Dauer hören. Beim ersten fragen Sie sich, was für ein Vokal es ist und ob er gut oder mangelhaft ist, und in welcher Beziehung letzteres; z. B. ‚E zu sehr nach Ä‘. Beim 2. fragen Sie sich, ob es derselbe, und wenn ja, ob er jetzt gleich gut, besser oder schlechter als vorher war und in welcher Beziehung. Dann folgen immer 2 unmittelbar aufeinander, die unter sich zu vergleichen sind. Alles irgend Bemerkenswerte ist aufzuschreiben.“ (Der 2. Teil der Instruktion wurde Neulingen immer erst nach Absolvierung des ersten Vokalpaares gegeben.)

Diese Instruktion sollte die Aufgabe erleichtern, indem sie darauf hinwies, daß man je 2 sich unmittelbar folgende Eindrücke miteinander zu vergleichen hatte. Aber in weitaus den meisten Fällen wurde dies nicht beachtet, sondern jeder Eindruck für sich beurteilt, wogegen ich nichts einzuwenden hatte, da auch so die Protokolle in gleichem Sinne zu verwerten waren und die Aufmerksamkeit durch die Aufeinanderfolge der 20 einander sehr ähnlichen Eindrücke doch immerzu auf die Beachtung ihrer Nuancen konzentriert blieb. Öfters äußerten sich die Vpn. am Schlusse einer Reihe deprimiert, daß sie fast gar keine deutlichen Unterschiede hätten finden können, daß sie immer nur denselben Vokal gehört hätten; was ich nicht ohne heimliches Vergnügen zur Kenntnis nahm. Aus allen Äußerungen ging hervor, daß auch während der Versuche keiner Vp eine Ahnung davon aufging, daß ein künstlicher mit einem gleichartigen natürlichen Vokal abwechselnd dargeboten wurde.

Da der gesungene Vokal durch die Röhrenleitung alteriert werden konnte, wurde vorher sorgfältig ausprobiert, auf welche Weise er möglichst gut herüberkam, namentlich wurden Trichter vorgeschaltet, bei den hellen Vokalen kleine, bei O ein großer, bei U ein mächtiger Grammophontrichter, bis sie bei Erkennungsversuchen sicher zu identifizieren waren. Auch überwachte ich die Vokalisierung beständig durch den Nebenschlauch.

Bei der Zusammenstellung der Ergebnisse bezeichnen wir als „kv“ die Fälle, wo innerhalb eines Versuchspaares der künstliche Vokal als vollkommener beurteilt wurde (sei es durch Vergleichung mit dem vorausgehenden oder nachfolgenden natürlichen, sei es durch absolute Prädikate, wenn z. B. beide Male ein Ü als solches erkannt, aber k als tadellos, n als etwas zu hell, oder wenn k als Ü, n aber als I bezeichnet wurde). Im gegenteiligen Falle wird „nv“ geschrieben. Mit gv (gleich vollkommen) bezeichne ich die Fälle, wo beide als der betreffende Vokal erkannt und als gut oder nur wenig alteriert beurteilt wurden (z. B. k als „Ü nach E hin“, n als „Ü nach I hin“). Als gu (gleich unvollkommen) endlich die Fälle, wo beide mit Nachbarvokalen oder gar (ein seltener Fall) mit entfernteren Vokalen verwechselt wurden.

Individuelle Urteilsneigungen zeigten sich auch hier. So schien einer Vp ein gutes klares A, das n wie das k, durchgehends nicht hell genug. Gerade in bezug auf A sind ja die Anforderungen verschieden. Eine andere Vp hörte das natürliche E immer etwas nach Ä hin, was übrigens vielleicht objektiv nicht ganz grundlos sein mochte, da E von männlichen Sängern leicht zu „offen“ angegeben wird. Eben wegen solcher kleinen individuellen Divergenzen wurde eine größere Anzahl von Vp herangezogen.

Die folgende Tabelle zeigt nun die Verteilung der Aussagen auf die 4 unterschiedenen Klassen. Sie schwankt freilich von Vokal zu Vokal. Viel-



leicht waren die k bei einigen Vokalen doch nicht so überzeugend wie sonst oder die n stärker als sonst alteriert, oder es spielten Zufälligkeiten mit. Immerhin läßt sich sagen, daß die n nirgends gegen die k überwiegen, sondern ihnen nur einmal — beim Ü — gleichstehen, bei den übrigen Vokalen dagegen mehr oder weniger hinter den k zurückbleiben, beim U, E und I sogar auf Null oder fast Null sinken. Es kann also, auch wenn man

	gv	<th>nv</th> <th>gu</th>	nv	gu
U	20	25	1	4
O	24	10	6	10
A	14	21	15	0
Ö	10	25	10	5
Ä	13	24	9	4
Ü	17	14	14	5
E	5	42	0	3
I	6	29	3	12
	109	190	58	43
	400			

nicht summieren will, kein Zweifel obwalten, daß der künstliche Vokal im allgemeinen weitaus den Vorrang hat.

Ich habe auch die Stärkeverhältnisse der Teiltöne, wie sie bei diesen Versuchen in den künstlichen Vokalen obwalteten, noch besonders festgestellt, und zwar am Ende der Leitung, wo der Vokal von den Vp gehört wurde. Da einzelne Vokale auf mehrere Versuchsreihen verteilt waren, für die sie immer neu aufgebaut wurden, ist ihre Zusammensetzung mehrfach angegeben. Dabei wurde auch gelegentlich ein anderer Sänger oder ein anderer Trichter benützt, so daß der

künstliche Laut, um dem natürlichen zu gleichen, in seiner Zusammensetzung etwas verändert werden mußte. Das zweite O und das zweite A gefielen besonders.

Tonhöhe  $c^1$ .

$g^5$												1
$e^5$												$1/2$
$c^5$												1
$b^4$												6
$a^4$												4
$g^4$							1				$1/2$	2
$fis^4$							2	3			2	4
$e^4$						3	4	4			3	4
$d^4$						2	6	5		4	5	1
$c^4$						3	6	5	4	6	5	2
$b^3$						5	6	7	8	6	$1/2$	0
$g^3$				1	4	8	4	6	4	5	2	0
$e^3$				4	8	4	8	6	3	0	1	3
$c^3$		4	5	6	8	5	8	8	2	0	3	3
$g^2$	6	8	6	10	10	2	9	8	3	$1/2$	3	4
$c^2$	4	10	10	4	3	6	6	6	3	0	8	3
$c^1$	8	3	2	0	0	5	0	3	6	7	5	6
	U	O	O	A	A	Ö	Ä	Ä	Ü	Ü	E	I

Verglichen mit der Tabelle S. 176 für den Grundton  $c^1$  läßt diese im allgemeinen dieselbe Struktur der Vokale erkennen. Ein völliges Zusammenfallen ist natürlich angesichts der kleinen Schwankungen der Vokalisierung wie der subjektiven Stärkeschätzung nicht zu erwarten. Auffallend ist nur der Mangel einer toten Strecke beim E. Möglicherweise ist es gerade dadurch gegenüber dem aus der Leitung kommenden natürlichen E so

verbessert worden, daß damit das unbedingte Übergewicht der  $k$  über die  $n$  (42 gegen 0) zusammenhängt.

Das von einem guten Sänger gesungene  $U$  habe ich auch durch Resonanzversuche analysiert, mit folgendem Ergebnis:  $c^1$  10,  $c^2$  3,  $g^2$  6–8,  $c^3$  1–2.

## V. Der Grundton als Differenzton.

Schon mehrfach ist darauf hingewiesen, daß der Grundton als subjektiver Differenzton (D. T.) aller benachbarten harmonischen Teiltöne zustande kommen oder wenigstens verstärkt werden kann; und so ist auch bei unseren Synthesen seine objektive Stärke häufig nahezu oder ganz = 0, während er dem Hörer kräftig vorhanden zu sein scheint. Dennoch sind mir zuweilen Zweifel gekommen, ob er auch nur subjektiv, im Ohre, real vorhanden war. Es ließe sich denken, daß dem Klange, sagen wir dem A-Klange aus den reellen Tönen  $c^2 g^2 c^3 e^3$ , nur die Höhe des  $c^1$  in der Auffassung des Hörenden irgendwie zuwüchse, ohne daß der Grundton  $c^1$  auch nur im Ohre dazu träte<sup>1)</sup>. Man könnte versuchen, eine Entscheidung dadurch herbeizuführen, daß man mit einer etwas verstimmtten  $c^1$ -Gabel auf Schwebungen prüfte. Aber gerade durch die Hilfgabel kann in Verbindung mit dem vorhandenen  $c^2$  ein auf  $c^1$  schwebender D. T. entstehen, außerdem kann auch der in der Gabel vorhandene schwache Oberton  $c^2$  mit dem vorhandenen  $c^2$  schweben (worauf hindeutet, daß die Schwebungen bei starkem Anschlag der Gabel gehört werden). Es müßten also Vorsichtsmaßregeln getroffen werden, um Fehlschlüsse zu vermeiden<sup>2)</sup>. Ein anderes Prüfungsmittel wäre die Untersuchung der Schwingungskurven des Trommelfells, wie sie KOEHLER durchgeführt hat. Denn das Trommelfell ist sicherlich der Hauptentstehungsort der subjektiven D. T. (nebenbei wohl auch das ovale Fenster). Aber die KOEHLERSchen Kurven reichen dazu nicht aus<sup>3)</sup>, und eine Wiederholung wird so schnell nicht stattfinden.

<sup>1)</sup> An eine solche psychologische Erklärung denkt auch PIPPING (2 gegen den Schluß).

<sup>2)</sup> Vgl. m. Abh. 7, S. 19ff.

<sup>3)</sup> Eine Stelle in KOEHLERS Beschreibungen (1, I) scheint dafür zu sprechen, daß auch hier objektiv kein oder nur ein sehr schwacher Grundton auftrat, wenn der Vokal in tiefer Lage gesungen wurde: „Aufnahmen des O von einer Baßstimme, die noch tiefer hinab in der Skala die erforderliche Stärke besaß, zeigten, daß, je tiefer der Vokal gesungen wurde, sein Wellenbild immer verwickelter ward, und daß Teiltöne, die an demselben in hoher Lage gar nicht zu bemerken waren, tiefer unten mit starken Zacken die Periode fast allein ausfüllten.“

Schließlich scheint mir doch die direkte Beobachtung für das reale Vorhandensein im Ohr in den meisten Fällen zu sprechen<sup>1)</sup>. Wenn man in der obenerwähnten A-Kombination das  $g^2$  abwechselnd wegnimmt und wieder hinzufügt, so kann man sich des Eindrucks nicht erwehren, daß eine wirkliche Veränderung in dem tieferen Klangbestande, im Träger der Tonhöhe, nicht bloß in der Tonhöhe des Klanges als solchen, Platz greift.

Außerordentlich lehrreich ist es übrigens bei diesem Versuch, die 3 Erscheinungen:  $c^1$  für sich allein als einfachen Ton, dann den A-Vokal auf dem D. T.  $c^1$ , endlich  $c^2$  wieder für sich allein, miteinander in Hinsicht ihrer Helligkeit zu vergleichen. Der A-Vokal auf  $c^1$  ist weder so dunkel wie  $c^1$  noch so hell wie  $c^2$ , liegt also gewissermaßen dazwischen, aber nicht so wie etwa *fis*<sup>1</sup> in der Mitte läge. Er hat eben eine Komplexhelligkeit, die nicht ohne weiteres in die Reihe der einfachen Helligkeiten eingeordnet werden kann. Genau kann sie nur eingereiht werden in eine Reihe von Komplexhelligkeiten gleicher oder verwandter Art, wie ein Klavierklang in die Reihe der Klavierklänge eingeordnet wird. So kann unsere Einrichtung auch für das Studium dieser prinzipiellen Fragen gute Dienste leisten.

Man kann die Frage aufwerfen, wie viele der tieferen harmonischen Teiltöne einem Klange fehlen können, ohne daß sich der Grundton für unser Gehör verändert, d. h. ohne daß er aufhört, als subjektiver D. T. hinzutreten. Natürlich hängt dies sehr von der Intensitätsverteilung innerhalb der Teiltonreihe ab. Für die Entstehung von D. T. ist am wirksamsten gleiche Stärke der beiden erzeugenden Primärtöne. Diese ist aber im allgemeinen bei den Teiltönen eines Vokals, auch bei den Instrumentalklängen, nicht vorhanden. Wird sie künstlich hergestellt, und hat der Klang zugleich in allen Teilen eine genügende absolute Stärke, so kann man überhaupt sämtliche Teiltöne bis auf 2 beliebige einander benachbarte wegnehmen: immer wird ein D. T. von der Verhältniszahl 1 entstehen. Also z. B. wenn nur der 15. und 16. Teilton übrig sind; denn das Intervall 15 : 16 liefert, etwa mit Stimmgabeln oder Pfeifen angegeben, den deutlichen Differenzton 1.

Kürzlich hat nun H. FLETCHER (3) dieser Frage unter Anwendung von elektrischen Schwingungen und Filtern, wie sie in der Radiotechnik jetzt üblich sind, eine experimentelle Untersuchung gewidmet und angegeben, daß man aus einem beliebigen Vokal- oder Instrumentalklang den Grundton und die 5—7 ersten

1) Über einen Ausnahmefall vgl. u. S. 193.

Obertöne ausschalten könne, ohne daß sich die Klanghöhe ändere. Nur die Klangfarbe erleide eine Veränderung. So auch bei einem auf  $d = 145$  Schw. gesungenen „ah“. Hier waren die sämtlichen Teiltöne von  $d-d^3$  ausgeschaltet, und dennoch hörte man den Vokal auf der Höhe des  $d$ . Wurden auch der 8. und 9. Oberton, also der 9. und 10. Teilton, noch weggenommen, so wurde die Tonhöhe ungewiß, die Klangfarbe aber zum Geräusch. Die Urteile wurden von 3 musikalischen Personen in Übereinstimmung miteinander abgegeben. Es wurden dann auch Synthesen mit 10 Röhrensender-Oszillatoren ausgeführt, welche Töne von 100—1000 Schw. mit je 100 Schw. Differenz lieferten. Hier waren je 3 benachbarte Töne genügend, um den Ton 100 im Ohre hervorzurufen.

Das letztere ist nun ohne weiteres verständlich, da es sich in diesem Falle, wie anzunehmen ist, um ungefähr gleich starke Primärtöne handelte. Dagegen stehe ich der Behauptung in Hinsicht der 8 (mit dem Grundton) ausgelöschten Teiltöne ganz ungläubig gegenüber und möchte sie, Genauigkeit der Einrichtung vorausgesetzt, als Ausfluß weitgehender „Perseveration“ bei den 3 urteilenden Personen auffassen. Der Grundton eines guten A verschwindet nach meinen Beobachtungen schon bei Wegnahme der 2 oder 3 untersten Teiltöne; z. B. geben die 3 Töne  $c^3 e^3 g^3$  in der Stärkeverteilung der Tabelle für  $c^1$  S. 167 kein A auf  $c^1$  mehr.

Es könnten aber auch Fehler der benutzten Apparatur mitwirken. Solche müssen ohne allen Zweifel angenommen werden, wenn behauptet wird, daß die Zusammenfügung der Töne 100, 300, 500, 700, 900 zwar den D. T. 200, aber keinen Klang von bestimmter Höhe und Klangfarbe, sondern nur eine Art Geräusch (like a noise) ergebe. Dasselbe sei der Fall, wenn die Töne um je 300 oder 400 Schw. voneinander abständen. Dies ist für jeden, der einmal die auf dieselben Schwingungszahlen abgestimmten Resonanzgabeln unseres Psychologischen Instituts hat zusammenklingen hören, ganz und gar unbegreiflich. Da man das normale Ohr der amerikanischen Beobachter nicht bezweifeln kann, muß es an der Einrichtung gelegen haben.

Auch was über die Klangveränderungen durch Filter bei den musikalischen Instrumenten berichtet wird (bei der Orgel wird das Register, auf das hier doch alles ankommt, nicht genannt), deutet auf solche Mängel hin. Die Tonhöhe soll bei Instrumenten durch solche Ausschaltungen nicht geändert werden, solange der filtrierte Klang überhaupt noch als ein musikalischer erkannt werde (S. 432). Warum aber sollen hohe Töne, wie sie in vielen Instrumenten vorhanden sind, beim Wegfall der tieferen nicht immer noch als Töne erkannt werden? Dies widerspricht aller Erfahrung. Es müssen also wohl durch die Apparatur Geräusche in den Klang hineingekommen sein, die die hohen feinen Töne zuletzt unterdrückten.

Zur Analyse der Kurven bediente sich FLETCHER, wie MILLER, statt des umständlichen Fourierrechens, eines „harmonischen Analysators“, der in 5 Minuten die Bestandteile einer Kurve, das „Tonspektrum“, auf

einer photographischen Platte auseinanderlegte. Auffallend ist, daß hier bei einem auf 2 verschiedenen Tonhöhen gesungenen A gerade der Grundton die größte Amplitude erhält (ebenso bei der Klarinette, wo versehentlich, wie aus den Abszissen hervorgeht,  $c$  statt  $c^1$  als Grundton angegeben ist).

## VI. Unharmonische Teiltöne.

Nummehr kann auch die Streitfrage über das Vorhandensein und die Bedeutung unharmonischer Teiltöne in den Vokalen allseitig geprüft werden. Man braucht a priori gar kein Vorurteil gegen solche zu haben, etwa pythagoreischen Zahlenspekulationen oder dem naturphilosophischen Prinzip der Einfachheit zuliebe. Es wäre ganz wohl möglich, daß in dem vielgliedrigen Mechanismus der Sprachwerkzeuge an irgendeiner Stelle dergleichen in den Klang hineinkäme. Nur rein tatsächlich haben wir in der Analyse nirgends etwas gefunden, was auf das Vorhandensein unharmonischer Beitäne in unseren gewöhnlichen Vokalen in den entscheidenden Gegenden deutete.

Mit dem synthetischen Verfahren läßt sich, wie ich glaube, eine definitive Aufklärung erzielen. Man kann z. B. durch bloße Erhöhung oder Vertiefung des Grundtons die sämtlichen übrigen Teiltöne zu unharmonischen machen, oder man kann nur eine oder einige höhere Pfeifen verstimmen oder neue einfügen. Bei der neueren Einrichtung (von 1919), die 2 Leitungen frei läßt, habe ich in diese gelegentlich 2 Pfeifen mit verschiebbaren Stempeln eingesetzt, mit denen man jede Tonhöhe zwischen  $g^1$  und  $c^3$ , auch 2 unharmonische Töne dieser Gegend gleichzeitig, dem Klange beimischen, ja auch eine stetige Tonbewegung innerhalb des sonst konstanten Klanges erzeugen kann. Aber schon vor diesen besonders darauf zielenden Versuchen hatten zufällige Verstimmungen einzelner Pfeifen mehr als genug Gelegenheit geboten, die Wirkungen unharmonischer Teiltöne zu erproben. Ebenso ist bereits bei den obigen Synthesen für den Grundton  $g^1$  ein unharmonischer Beiton  $\bar{b}^4$  statt  $h^4$  benutzt (s. die Tabellen).

Wie man schon aus diesem Beispiel sieht, tut der unharmonische Eindringling unter Umständen dieselben Dienste wie der fehlende harmonische Ton, wenn er nur der betreffenden Tonlage angehört und nicht zu schrille Schwebungen mit den Nachbartönen gibt. Überhaupt aber kommt es in dieser hohen Lage, in der 4-gestrichenen Oktave und darüber, nicht darauf an, ob der Ton zu der ganzzahligen Obertonreihe gehört. Es dürfen nur nicht zu viele und zu dicht benachbarte Töne eingefügt werden, mögen sie harmonisch sein oder nicht.

Verhängnisvoller wird die Einfügung tieferer, d. h. zwischen den harmonischen Teiltönen von niedrigen Ordnungszahlen liegender, unharmonischer Beitöne. Hier müssen aber zunächst gewisse Fälle ausgeschieden werden, in denen es sich nur scheinbar um unharmonische Beitöne handelt. Wenn man in den künstlichen Vokal A auf  $c^1$  die Töne  $g^1$  oder  $e^2$  oder beide einfügt, die in dessen Obertonreihe nicht enthalten, also zu  $c^1$  unharmonisch sind, so ist die Wirkung einfach die, daß der Grundton um eine Oktave hinabspringt: man hört wieder ein gutes A, aber auf  $c$ . Denn dieser Ton wird jetzt als subjektiver Differenzton gebildet. Zu ihm aber sind die neuen Teiltöne harmonisch; wir haben jetzt die Reihe  $c, c^1, g^1, c^2, e^2, g^2, c^3$  usw. Solche Fälle also müssen hier ausscheiden<sup>1)</sup>.

Fügt man nun beliebige andere unharmonische Teiltöne ein, so können Schädigungen in mehrfacher Richtung entstehen, und es ist kein Fall denkbar, wo nicht eine oder mehrere davon eintreten müssen:

1. Wenn der unharmonische in der Nähe von reell vorhandenen harmonischen liegt, so schwebt er mit diesen und bringt dadurch eine Unebenheit oder Rauhigkeit in den Klang, wie sie den natürlichen Vokalen bei normaler Stimmbeschaffenheit nicht eigen zu sein pflegt.

2. In allen Fällen entsteht ein mehr oder minder ausgeprägter Übelklang, der von der Rauhigkeit der Schwebungen noch zu unterscheiden ist und sich namentlich dem musikalischen Hörer fühlbar macht, auch ohne daß er den Klang analysiert. Beobachtungen haben gelehrt<sup>2)</sup>, daß dem Dur-Dreiklang und seinen Derivaten (Um- und Weitlagerungen) auch für Unmusikalische oder in homophoner Musik Aufgewachsene (Asiaten, Naturvölker) ein gewisser Reiz innewohnt, der ihn vor allen anderen Zusammenklängen auszeichnet. Die Zusammensetzung aus harmonischen Teiltönen, von denen die ersten 6 den reinen Dur-Dreiklang nebst tieferen Oktaven des Grundtons und der Dominante geben, erteilt daher den Vokalen auch im unanalysierten Zustande einen Vorzug, an den wir zwar gewöhnt sind, dessen Wegfall wir aber unliebsam empfinden würden. Natürlich bleibt ein Unterschied gegenüber einem Akkord („Konkord“), bei dem in der Regel die

<sup>1)</sup> An diesem Fehler leiden die schon von v. WESENDONK versuchten Kombinationen unharmonischer Teiltöne. Wenn man in seinen Beispielen, welche alle ein A geben sollen ( $f^1 f^2 b^2 f^3$ ;  $fis^2 a^2 d^3 fis^3$ ;  $a^2 d^3 a^3$ ;  $a^2 fis^3 a^3$ ), die Differenztöne  $h-t$  und  $2t-h$  berechnet, die stark herauskommen müssen, so ergeben sich regelrechte durchweg harmonische Teiltonreihen auf den Grundtönen  $B, d$  und  $d^1$ . Es ist also nichts damit zu beweisen.

<sup>2)</sup> Vgl. z. B. Tonpsych. II, S. 264 und „Musik der Siamesen“ in m. „Beitr. z. Akust. u. Musikwiss.“, Heft 3, S. 106. 1901.

einzelnen Töne gleiche Stärke haben. Aber etwas von der rein physiologischen Wirkung des Durklanges scheint doch auf die Sprachklänge, namentlich auf O und A, bei denen die Stärkeverteilung am günstigsten ist, übergegangen zu sein<sup>1)</sup>.

Man darf aber diese harmonische Zusammensetzung nicht als einen wunderbaren Zufall oder als eine exzeptionelle Einrichtung ansehen, die besonderer Erklärungsgründe bedürfte. Die Polsterpfeife des menschlichen Kehlkopfs hat darin nichts vor den meisten übrigen Pfeifen und vor den schwingenden Zungen und Saiten voraus. Instrumente mit unharmonischen Obertönen bilden Ausnahmen. Nun sind freilich alle musikalischen Instrumente mehr oder weniger Kunstprodukte und insofern teleologische Einrichtungen. Aber auch jeder Organismus ist eine solche. Und daß unser angeborenes Instrument nicht besondere Gesetzlichkeiten für sich verlangt, kann der Theorie nur willkommen sein.

Der S. 170 erwähnte Demonstrationsapparat, an dem die Obertöne nicht ausgeschaltet waren, gab bei der geringsten Verstimmung für mein Ohr statt eines Vokals einen abscheulichen Mißklang, woran die mitverstimmten Ober- und Differenzttöne nebst ihren Schwebungen Schuld waren — auch ein Grund, warum nur ganz einfache Töne zu Synthesen taugen. Bei diesen ist die üble Wirkung kleiner Verstimmungen weit geringer, noch geringer die der regelrechten Dissonanzen, wie solche ja auch unter den höheren harmonischen Teiltönen vorkommen. Aber wenn statt der tieferen, stark konsonierenden Teiltöne oder zwischen sie hinein stark dissonierende gesetzt werden, wird doch auch bei ganz einfachen Tönen die sinnliche Annehmlichkeit des Zusammenklanges vermindert.

3. leidet die Einheitlichkeit, die für den Vokaleindruck unentbehrlich ist. Die harmonischen Teiltöne bis zum 6. verschmelzen in hohem Grade zu einer Klangeinheit (vgl. 11. Kap.). Tritt ein dissonanter Ton zwischen sie oder an die Stelle eines von ihnen, so hört die Einheitlichkeit und damit der Vokaleindruck entweder ganz auf oder der Klang wird (so bei sehr unmusikalischen, zur Analyse unfähigen Personen) wenigstens als fremdartig, weniger rund und voll empfunden.

4. Der Grundton wird, wenn der unharmonische Teilton an die Stelle eines ausfallenden harmonischen tritt, nicht mehr genügend durch Differenztonbildung verstärkt. Dies wirkt besonders schädlich in den Fällen, wo er, wie beim A und Ä, fast

<sup>1)</sup> Vgl. HELMHOLTZ, S. 120: „Wenn wir die Reihe der 6 ersten Partialtöne eines zusammengesetzten Klanges überblicken, so können wir letzteren in musikalischer Beziehung als einen Durakkord mit überwiegend starkem Grundton betrachten, und wirklich hat auch ein solcher Klang, z. B. ein schöner Gesangston, neben einem einfachen Tone ganz deutlich etwas von der angenehmen Wirkung eines harmonischen Akkordes.“ In diesem Sinne verteidigte auch HENSEN (3) die „Harmonie in den Vokalen“.

nur auf diesem Wege zustande kommt. Außerdem entstehen unharmonische Differenztöne, die, wenn sie auch nicht für sich hörbar sind, die unter 1—3 genannten Übelstände vergrößern können.

Wird der Grundton selbst verstimmt oder durch einen anderen ersetzt, so müssen ähnliche Wirkungen auftreten: Übelklang, mangelnde Einheitlichkeit, unter Umständen Schwebungen mit dem aus der unverstimmten Teiltonreihe resultierenden Differenzton.

Die Reihe der ganzzahligen Multipla eines Grundtones ist eben dadurch, daß der Grundton zugleich Differenzton aller unmittelbar aufeinanderfolgenden Obertöne, der 2. Teilton zugleich Differenzton aller um 2 Einheiten differierenden Obertöne ist usf., ein ganz einzigartiges, in sich geschlossenes akustisches Gebilde, ein Organismus, der sich gegen jeden Eindringling wehrt und durch ihn in seinen Funktionen sofort geschädigt wird, gleich einem Staate, der in sich uneins, von Parteien zerrissen ist. Das sagen wir nicht aus apriorischen Theorien, sondern aus den Beobachtungen heraus. Die synthetischen Vokale wurden immer nur dann den natürlichen gleich, wenn (abgesehen von den höchsten Obertönen) harmonische Teiltöne, und zwar in ganz reiner Stimmung, verwendet wurden, entfernten sich dagegen bei Einfügung unharmonischer in den angegebenen Richtungen von diesem Ziele, wenigstens für feinerhörende Ohren. Damit ist, meine ich, der definitive Beweis geliefert, daß in den natürlichen Vokalen, mindestens unter ihren tieferen Bestandteilen, keine unharmonischen Bestandteile vorhanden sind. Speziell die Formanten des A und O, die in den Diskussionen über HERMANN'S Theorie der Vokalbildung eine entscheidende Rolle spielen, enthalten sicher nichts von solchen. Wenn unter den höchsten Obertönen sich noch irgendwelche schwache Bestandteile dieser Art gelegentlich oder selbst regelmäßig einmischen sollten — vielleicht bei den nasalierten Lauten —, so würde es sich kaum lohnen, darüber lebhaft zu streiten, da sie keine weitergreifende theoretische Bedeutung hätten.

In einem Punkt wäre es allerdings für die Deutlichkeit der Vokalisierung nützlich, wenn der Sachverhalt der von HERMANN behauptete wäre: die Formanten haben nach ihm mit dem Stimmton gar nichts zu schaffen, es wäre also die Identifikation eines Vokals auf verschiedenen Tonhöhen aufs einfachste gewährleistet. Ferner würde bei dem zwei- und mehrstimmigen Gesang auf demselben Text der Formant jedes Vokals durch die verschiedenen Stimmen verstärkt, der Vokal also noch kenntlicher werden, während nach uns gewisse Kollisionen eintreten müssen, wenn z. B. A auf



der Simultanterz  $c + e$  gesungen wird, indem die Formanttöne  $g^2$  und  $gis^2$ , sowie  $\bar{b}^2$  und  $h^2$  miteinander schweben müssen.

Aber es ist nun einmal tatsächlich anders, als es sich HERMANN vorstellte. Er selbst hat bereits zugegeben, daß das Unkenntlichwerden der Vokale von  $c^2$  an nicht mit seiner These stimmt, und seine Schüler sind auf das langsame Ansteigen der Formanten mit der Höhe aufmerksam geworden. Übrigens wären unharmonische Teiltöne auch nicht eo ipso unveränderliche Teiltöne: die unharmonische, falsche Duodezime der Stimmgabel z. B. steigt ebenso mit dem Grundton wie die harmonische reine Oktave. Die Unveränderlichkeit fließt also nur aus HERMANN'S physiologischer Hypothese über den Mechanismus der Vokalisierung, und diese Hypothese ist in sich selbst nicht einmal so klar, daß die Konsequenz zweifellos einleuchtet.

Die Sicherheit des Wiedererkennens aber wird in Wirklichkeit dadurch gewährleistet, daß das „reine“ A, O etc. immer noch eine gewisse Breite hat, innerhalb deren die kleinen, durch die Verschiebung des Formantenzentrums bedingten Änderungen der Klangfarbe liegen, daß ferner bei dem gleitenden Tonfall des gewöhnlichen Sprechens diese Zone beständig gewissermaßen abgestreift wird, und daß bei mehrstimmigem Gesang auf gleichem Texte die Obertöne der einzelnen Stimmen sich ergänzen, so daß der Formant durch mehrere davon gleichzeitig vertreten ist<sup>1)</sup>. Ihre Kollisionen aber müssen zwar den Gesamtklang unstreitig etwas rauher machen — eine in allen Fällen unvermeidliche Wirkung der Mehrstimmigkeit —, dürften aber die Vokalwirkung der nahe zusammenfallenden Formantzentren nicht beeinträchtigen, sondern doch auch nur verstärken<sup>2)</sup>. In welchem Sinne und welchem Grade tatsächlich die Deutlichkeit des Vokaleindrucks durch Mehrstimmigkeit beeinflusst wird, ist meines Wissens noch nicht untersucht.

Es verhält sich mit den Nachteilen der harmonischen Teiltöne für das Sprachverständnis ähnlich wie mit den Nachteilen der optischen Konstruktion des Auges. Dieses hat bekanntlich, als optisches Instrument angesehen, eine ganze Anzahl von Fehlern, aber sie sind nach seiner Entstehungsweise und seinen Ernährungsbedingungen unvermeidlich und halten sich, wie HELMHOLTZ gegen-

<sup>1)</sup> So liegen im A-Formanten beim Singen der Terz  $a^1 + c^2$  die Töne  $a^2, c^3, e^3$ , beim Singen der Quinte  $g + d^1$  die Töne  $g^2, a^2, h^2, d^3$  (dieser doppelt).

<sup>2)</sup> Beispielsweise haben in HÄNDELS „Halleluja“ die 4 Stimmen bei ihrem Forte-Einsatz in D-Dur auf dem „Ha“ folgende Teiltöne innerhalb des A-Formanten:  $gis^2$  doppelt,  $a^2$  doppelt,  $c^3, cis^3, d^3, e^3$ . Der Formant ist also ausgezeichnet stark vertreten. Für die englische Aussprache liegt es freilich wieder anders, aber doch wohl ähnlich.

über KARL VOGT betont hat, beim normalen Auge in solchen Grenzen, wie es die Lebensbedürfnisse des Organismus verlangen.

Wegen der Wichtigkeit der Sache soll das Gesagte noch durch die Beschreibung und Analyse einiger Versuche erläutert werden. Bei der Analyse muß allerdings einiges zum 13. Abschnitt Gehörige (über „Vokalitäten“) vorweggenommen werden.

1. Einfachste Anordnung:  $c^2$  (schwächer) +  $c^3$  (stärker) gibt schon ein deutliches A, wenn man auf Vokale im allgemeinen eingestellt ist. Nun werde  $gis^2$  eingefügt. Ich höre dann zunächst einen Mehrklang aus den 3 Tönen, aber keinen Vokal. Der Institutsgehilfe KAUMANN, der ganz unmusikalisch ist und nicht zur Analyse neigt, aber eben darum oft um seine Meinung gefragt wird und sich ein sehr sicheres Vokalurteil erworben hat — ich nenne ihn hier gern als Mitarbeiter — hört Ao. Ebenso, wenn statt  $gis^2$   $fis^2$  genommen wird. Mit einiger Mühe gelingt es auch mir, in beiden Fällen eine dem A nahekommende Vokalität, ein verdunkeltes A, wahrzunehmen. Aber woher diese Verdunkelung? Vermutlich stammt sie von den Differenztönen  $h-t$  und  $2t-h$  ( $h$  der höhere,  $t$  der tiefere von je 2 Tönen), die der kleinen und 1-gestrichenen Oktave angehören<sup>1</sup>).

Eigentümlich ist auch, daß mir bei nicht absolut scharfer Aufmerksamkeit die Tonhöhe des Ganzen als  $c^1$  erschien, obgleich dieses hier nicht als D. T. auftreten kann. Wahrscheinlich sind hieran kleine Rauigkeiten des Zusammenklanges schuld, die auf kollidierenden D. T. beruhen (bei  $c^2 + gis^2 + c^3$  entsteht eine solche Kollision auf  $dis^1$ , bei  $c^2 + fis^2 + c^3$  auf  $gis$ ). Da diese D. T. nicht selbst bemerkbar werden, so veranlaßt die Rauigkeit nur eine Oktaventäuschung bezüglich  $c^2$ .

Für mich macht sich außerdem (wenn nicht etwa statt  $gis^2$   $as^2$  oder statt  $fis^2$  das halbkonsonante  $\bar{fis}^2 = c^2 \cdot \frac{7}{5}$  gesetzt wird) hier schon das Gefühl des Mißklanges geltend.

2. Wird mit dem zwischen  $c^2$  und  $c^3$  eingeschalteten Ton auf  $f^2$  heruntergegangen, so findet KAUMANN schon etwas U-artiges darin; was wieder auf den D. T.  $f$  und  $f^1$  beruhen mag. Wird  $f^2$  durch  $e^2$  ersetzt, so hört K. ein schönes helles A. Er hört nämlich jetzt doch das hohe  $c^3$  heraus, das er auch für sich allein schon als A bezeichnet.

3.  $g^1 + g^2 + d^3$  in entsprechenden Stärkeverhältnissen (1:2:1) gibt ein gutes A. Ersetze ich nun  $g^2$  durch  $fis^2$ , so kann ich den Klang nicht recht einheitlich bekommen, doch scheint er mir wesentlich nach O zu liegen. K. hört OU. Wird  $g^1$  auf die Stärke  $\frac{1}{2}$  reduziert, so geht die Vokalität für uns beide nach OA.

Auch hier dürften D. T. schuld sein. Bei dem ursprünglichen Dreiklang wird nur  $g^1$  als D. T. verstärkt. Er wird zum Träger der durch die Obertöne bewirkten Vokaleigenschaften, wobei, wie immer, seine eigene Vokalität nicht zur Geltung kommt. Bei dem veränderten Dreiklang entstehen durch  $fis^2 + d^3$  die D. T.  $d^1$  und  $a^1$  (durch  $g^1 + fis^2$  auch ein sehr schwaches, wohl wirkungsloses  $f^1$ ), welche die Annäherung an O hineinbringen. Daß bei Schwächung des  $g^1$  sogar OU gehört wurde, könnte (wenn nicht eine Zufälligkeit des Urteils vorlag) darauf beruhen, daß dadurch die Schwebungen von  $g^1$  mit den nahen D. T.  $a^1$  und  $f^1$  merklicher wurden (Schwebungen sind bei Stärkegleichheit am merklichsten). Freilich sind diese Deutungen etwas hypothetisch.

<sup>1</sup>) Diese beiden Differenztöne sind bei Intervallen innerhalb der Oktave immer, und zwar mit annähernd gleicher Stärke vorhanden. Vgl. m. Abh. 7.

4.  $g^1, c^2, e^2, g^2, c^3, e^3$  gibt in richtigen Stärkeverhältnissen ein gutes A auf dem D. T. c. Ersetzen wir  $g^2$  durch  $as^2$ , so ist mir zunächst das Auffälligste der böartige Mißklang. Bemühe ich mich, einheitlich zu hören, so erscheint mir als Vokalcharakter etwa Oä. K. findet ein „unreines O, mit A versetzt“. Durch die zerstörte Einheit scheint die O-Vokalität des tiefen und die A- oder Ä-Vokalität der hohen Elemente frei geworden zu sein, wie die Atome eines aufgelösten Moleküls.

Füge ich  $g^2$  und  $as^2$  zugleich in den Komplex ein, so entsteht für mich ein unerträglicher Mißklang. K. findet den Vokal „etwas mehr nach A liegend“. Doch sei „noch etwas dazwischen“.

5.  $c^1 g^1 c^2 e^2 g^2 \bar{b}^2 c^3 e^3$  in der nötigen Stärkeverteilung ist wieder ein gutes A auf c. Wird  $g^2$  durch  $gis^2$  ersetzt, ist es für mich auch so noch ein A, aber abscheulich unrein und ziemlich rau, auch nicht mehr so einheitlich. Wird dann  $\bar{b}^2$  ausgeschaltet, ist es nicht mehr so rau (da die Schwebungen zwischen  $\bar{b}^2$  und  $gis^2$  wegfallen), aber noch weniger einheitlich.

Solche Mißklänge unanalysierter Tonkomplexe waren für den Verfasser, der von Jugend auf Akkorde als Tonmehrheiten hört und bei Unreinheiten sich stets den Sitz der Verstimmung vergegenwärtigt, etwas Neues und theoretisch Bemerkenswertes. Für solche, die nicht oder weniger analysierend zu hören pflegen, mögen sie gewohnte Erlebnisse sein. Doch dürfte ein mißklingender Vokal auch anderen noch nicht oft vorgekommen sein.

6. Veränderung des Grundtons bei unveränderten Obertönen:  $c^2 g^2 c^3 e^3$  in entsprechenden Stärkeverhältnissen gibt ein A auf D. T.  $c^1$ . Setzt man nun  $b$  oder  $cis^1$  oder  $d^1$  als realen Grundton dazu, so höre ich einfach einen Mißklang, und der Grundton hebt sich viel zu scharf von dem darüberliegenden Tonkomplex ab. K. hört in diesem Fall ein AO, ohne durch den unharmonischen Grundton bzw. die unharmonischen Obertöne gestört zu werden. Der Grundton für sich allein klingt ihm mehr wie U.

Die Zusammenstellung  $c^1 g^1 c^2 e^2 g^2 \bar{b}^2 c^3 e^3 g^3$  in den nötigen Stärkeverhältnissen hört er wieder als ein gutes A. Fügt man nun noch  $es$  hinzu, wodurch sämtliche Töne unharmonisch zu diesem Grundton werden, so fehlt dem Klange für mich wieder die nötige Einheit. K. hört ihn dagegen „O-artig“.

Für einen Unmusikalischen wird also der Vokalcharakter beide Male infolge des hinzukommenden tiefen Tones (der stärker als der Differenzton sein mag) verdunkelt, aber nicht vernichtet. Es können also sogar sämtliche Obertöne unharmonisch sein. Sie sind es nur tatsächlich in der wirklichen Menschenstimme nicht, und das ist günstig, da es dem Klang auch für Ohren, die mehr an das Analysieren gewöhnt sind, erst die nötige Einheit sichert.

In diesen Fällen waren nun freilich die Obertöne wenigstens unter sich harmonisch. Setzt man einen komplizierten Klang ganz aus Teiltönen zusammen, unter denen nirgends einfache Zahlenverhältnisse bestehen, so muß die Sache schlimm werden. Aber es schien nicht erforderlich, über solche Kunstgebilde besondere Versuche anzustellen.

Nach alledem können wir nun auch HERMANN'S synthetischen Versuch, auf den er das entscheidendste Gewicht für seine Auffassung der Vokale legt, würdigen. Es gelang ihm, einen Resonator für  $g^2$  durch eine Lochsirene intermittierend anzublase, und er erhielt so einen tiefen Sirenenton, der zwischen 60 und 200 Schwingungen variieren konnte, und den konstanten Resonator. Sie ergaben zusammen ein gutes A, wenn keiner der beiden sich vorwiegend aufdrängte, sondern das Ganze als Einheit

gehört wurde (Bd. 141, S. 52). Dabei war  $g^2$  im allgemeinen unharmonisch zum Grundton. Der Versuch war zwar nicht rein genug, sofern keine genügende Garantie für die Einfachheit der verwendeten Töne gegeben war. Aber wir haben keinen Grund, an dem Erfolg selbst zu zweifeln, und es bleibt eines von HERMANN'S Verdiensten, zuerst die Verwendbarkeit unharmonischer Komponenten erwiesen zu haben. Nur würden wir nicht die Folgerung ziehen, daß damit seine Theorie über die Entstehung und das Wesen der Vokale bewiesen wäre.

### VII. Vokale aus nur 2 Tönen.

Von theoretischer Bedeutung ist die Möglichkeit, einen guten Vokal unter Umständen aus nur 2 einfachen objektiven Tönen herzustellen. So kann man O durch  $c^1 + c^2$ , A durch  $g^1 + g^2$  oder  $c^2 + c^3$  unter entsprechender genauer Regulierung des Stärkeverhältnisses darstellen (z. B. für A:  $c^2 = 2^{1/2}$ ,  $c^3 = 3$ )<sup>1)</sup>, oder etwas dunkleres A:  $g^1 = 1^{1/2}$ ,  $g^2 = 3$ ). Aber der Hörer muß dabei auf einheitliches und vokalisches Hören eingestellt sein. Unter dieser Bedingung kann sogar die Quinte  $c^2 + g^2$  ein A (auf dem Differenzton  $c^1$ ) geben. Bei Unmusikalischen wird dies leichter vorkommen. Aber auch dem Verfasser gelang es einmal, nachdem ein vollständigeres A unmittelbar vorausgegangen war. Bald freilich drängten sich wieder die einzelnen Töne auf, und es war auch der D. T.  $c^1$  daneben hörbar, zunächst noch mit der Vokalfarbe des A, aber nicht mehr so überzeugend, dann ohne diese Färbung: lehrreiche Erfahrungen für den Einfluß des einheitlichen Hörens auf den Vokaleindruck. Wir werden aber auch in anderen Beziehungen noch auf diese binären Verbindungen zurückkommen (13. Kap.).

Auch Beobachtungen aus den If.-Versuchen können hier angezogen werden: wenn bei Lückenversuchen aus einem auf *as* gesungenen A alle Teiltöne außer *as* und  $as^2$  oder  $c^3$  ausgeschaltet waren, also sogar 2 Oktaven zwischen beiden Tönen lagen, konnte gelegentlich noch ein A gehört werden, allerdings nicht mehr einheitlich genug und „wie aus der Ferne“.

### VIII. Besondere Vokaltypen.

Leicht lassen sich durch die synthetische Einrichtung die nationalen und dialektischen Unterschiede in der Aussprache der Lautbuchstaben wiedergeben; so z. B. in bezug auf den wichtigsten und zugleich veränderungsfähigsten Vokal, das A. Die obigen Struk-

<sup>1)</sup> Der Stärkeunterschied muß hier größer sein als in der obigen Tabelle für den Grundton  $c^2$ , weil dort durch den 3. Teilton  $g^3$  noch eine Erhellung bedingt ist, die hier durch größere relative Stärke des  $c^3$  hervorgebracht werden muß.

turen dürften dem durchschnittlichen hochdeutschen A entsprechen. Wünscht man es aber heller (wie etwa das italienische), so braucht man für den Grundton  $c$  nur  $c^3$  von gleicher Stärke wie  $g^2$  zu nehmen. Nur stärker darf es nicht sein. Ebenso läßt sich das schwedische A = AO, das etwas näselnde, schneidige, dem Ä nahestehende A, oder jenes blökende Äao oder ÄÖao, von dem schon bei den If.-Versuchen die Rede war, erzeugen. Dieses wird dadurch hergestellt, daß von einem synthetischen Ä die über  $c^4$  liegende Partie abgeschnitten und die zunächst darunterliegenden Teiltöne etwas geschwächt werden, in der Weise, wie es bei If.-Versuchen mit Ä durch die Einstellungen oberhalb  $c^4$  infolge der If.-Breite der Fall ist. Das helle skandinavische U (s. o. S. 151) wäre vielleicht durch Erhöhung des Oberformanten zu gewinnen. Usf.

Eine kleine Studie wurde dem russischen Y (Jeri, Ы) gewidmet, dessen phonetische Eigentümlichkeit auch die Linguisten vielfach beschäftigt hat. Der Laut soll aus einem urslawischen U hervorgegangen sein, wird in den verschiedenen slawischen Sprachen ungleich gesprochen und scheint im Begriffe, sich dem I zu nähern<sup>1)</sup>. Nach dem bloßen Gehör erschien er mir in der Aussprache gebildeter Russen beiderlei Geschlechts aus Petersburg, Moskau, Kasan, Sibirien, Litauen im wesentlichen als ein zwischen Ö und Ū liegender Laut mit starkem, dunklem Unterformanten, in unserer Schreibweise etwa = UÖü. So aber nur als ganz kurzer, trockener Laut (was russische Phonetiker, glaube ich, einen „irrationalen Vokal“ nennen), wie in my (= wir). Dagegen als langer Vokal wie in byt' (bytj = sein) klang er wie Uöü-ī, also diphthongisch, wobei i nur äußerst kurz dem Hauptlaut angehängt wird<sup>2)</sup>. Von dem Assistenten des Moskauer

<sup>1)</sup> Vgl. u. a. BROCH S. 150ff., 163, 173, besonders aber THOMSON in seinen Abhandlungen 1—3 und in 2 Monographien, die 1926 in der Zeitschr. f. slawische Philologie erscheinen werden und mir durch den Herausgeber, Prof. VASMER, freundlichst im Manuskript mitgeteilt wurden. Unter den physiologischen Phonetikern hat sich GRÜTZNER S. 165 mit diesem Laut beschäftigt. Er wird gebildet, indem man die Lippen auf I, die Zunge nach oben gewölbt auf Ū einstellt.

<sup>2)</sup> Diese diphthongische Natur betont namentlich THOMSON für das Großrussische und findet sie um so klarer, je länger die Silbe. Im Kleinerussischen dagegen fällt nach ihm der Laut mit I zusammen. Auf die diphthongische Natur deutet er auch die aus dem Kirchenslawischen stammende Schreibung Ы und die altböhmische ui (wie in buistrice). Er hat auch die Flüsterhöhen des Y an sich und anderen bestimmt. Die des Diphthongen ist =  $gis^3$  ( $a^3$ ) mit auslautendem  $h^3$  ( $c^4$ ), individuell auch wohl  $\frac{1}{2}$  Ton tiefer. Die des einfachen Lautes =  $ais^3$  ( $h^3$ ). Diese Höhen stimmen gut mit den Formanten unserer obigen Synthesen, ebenso wie sie sich in das Schema unserer Flüsterhöhen o. S. 145 einfügen.

Psychologischen Instituts, Herrn LURIA, gesungen — wobei natürlich nicht die diphthongische Form, sondern die des kurzen Lautes oder der Hauptbestandteil des Diphthongen zur Anwendung kam —, trug er auf  $c$  mehr Ö-, auf  $c^1$  mehr Ü-Charakter. Ebenso bei dem aus Rußland stammenden Lektor der slawischen Sprachen, Herrn LANE, der eine schöne metallreiche Stimme besitzt. Bei LURIA war auf beiden Grundtönen  $e^4$  mit bloßem Ohr herauszuhören. Der Laut ließ sich zur Befriedigung der Sänger durch folgende Zusammensetzungen künstlich herstellen:

Russisches Y auf  $c$ :

	$c$	$c^1$	$g^1$	$e^2$	$e^2$	$g^2$	$\bar{b}^2$	$c^3$	$e^3$	$\overline{fis}^2$	$g^3$	$\bar{a}^3$	$\bar{b}^3$	$c^4$	$e^4$
LURIA	0	4	4	3	0	0	0	0	0	7	7	7	4	4	2
LANE	0	6	4	3	0	0	0	0	0	6	8	4	0	0	0

Russisches Y auf  $c^1$ :

	$c^1$	$c^2$	$g^2$	$c^3$	$e^3$	$g^3$	$\bar{b}^3$	$c^4$	$e^4$	$g^4$
LURIA	2	$\frac{1}{2}$	0	0	4	6	4	4	3	0
LANE	6	4	2	0	6	8	4	4	4	2

Auch Nasallaute, wie die französischen on, en, un, in lassen sich erzeugen. So wird aus einem guten Ä auf  $c^1$  ein französisches in, wenn man die Teiltöne  $g^2$  und  $c^3$  herausnimmt, dagegen noch hohe Töne bis zu  $c^5$  hinzufügt; wie dies den Analysen S. 116ff entspricht. (Bei dieser Gelegenheit war auch zu bemerken, daß die Nasalierung eines zum Vergleich gesungenen Lautes undeutlich wird, wenn er durch eine Röhrenleitung oder hinter einer geschlossenen Türe gesungen wird. Die für die Nasalierung wesentlichen leisen hohen Teiltöne werden eben dann verschluckt.)

Endlich sind Umwandlungen der Vokale ineinander, ja auch von Vokal- in Instrumentalklänge und umgekehrt, mit dieser Einrichtung leicht auszuführen. Überall zeigt sich dabei, wie der Klangcharakter ganz und gar von den Verschiebungen in den Stärkeverhältnissen der Teiltöne bedingt ist. Aber es erübrigt sich, auf diese Einzelheiten weiter einzugehen.

## IX. Stimmfärbungen.

Die Teiltonstruktur eines Stimmklanges ist nicht nur die Grundlage seiner Vokaleigenschaften, sondern auch die Hauptursache der individuellen Stimmqualität: der dünnen oder vollen (pastosen), der massiven oder ätherischen, der weichen oder metallischen, ja scharfen, schneidenden, der herben und süßen, und wie man sie sonst im Leben, in der Gesangstechnik und Konzertkritik charakterisieren hört. Allerdings nur die Hauptursache: denn vieles wirkt hier zusammen. Wenn z. B. eine Sprechstimme melodisch,

musikalisch genannt wird, so trägt oft schon eine besonders hohe Lage und ein reicher Wechsel zwischen hohen und tiefen Tönen, auch das Zurücktreten der Geräusche dazu bei. Bei Gesangsstimmen macht auch die durchschnittliche Stärke der Stimmgebung etwas aus. Dies alles muß zunächst ausgeschieden werden, wenn nach den maßgebenden Struktureigentümlichkeiten gefragt wird. Die Stärke kommt besonders insoweit in Betracht, als der Obertonreichtum und die Verschiebung der Intensitäten nach oben damit zusammenhängen.

Nun fragt es sich aber: kann man einem gegebenen Vokal ohne jede Änderung seiner Vokalität überhaupt noch eine veränderte Zusammensetzung erteilen?

Will man Haare spalten, so müßte man dies vielleicht in der Tat für unmöglich erklären. Aber wie die Dinge praktisch liegen, gibt es Unterschiede, die den Stimmklang beeinflussen können, ohne die Vokalität zu alterieren. Wir erwähnten solche schon im 1. Kap. Und die Synthese bestätigt, daß man an 2 Stellen die Teiltonreihe durch schwache Obertöne in diesem Sinne ergänzen kann: jenseits der Formanten aller Vokale und in den leeren Strecken der hellen Vokale. In ersterer Beziehung kann man namentlich beim A die 1. Hälfte der 4-gestrichenen Oktave noch mit Teiltönen von der Stärke 1 (4) ausfüllen, ohne daß es etwa in  $\ddot{A}$  übergeht. In den unangenehm scharfen Stimmen sind sogar stärkere hohe Beitäne vorhanden, die ein geübtes Ohr heraushören kann, und sie mögen auch wohl teilweise unharmonisch zum Grundton sein.

Mit der synthetischen Einrichtung ist auch die Klangverschiedenheit der Register, namentlich der Unterschied von Falsett- und Bruststimme, ebenso der von Frauen- und Männer- bzw. Knabenstimme näher zu untersuchen.

In der vielverhandelten Registertheorie ist schon die Definition des Grundbegriffes strittig. Ich würde unter Register eine Methode der Stimmgebung verstehen, durch die ein Sänger eine von mehreren auf demselben Tone möglichen Klangfarbengattungen erzeugt. Daß das Falsett im allgemeinen obertonärmer ist als die Bruststimme, ist bekannt, und die Resonanzversuche haben es im einzelnen erwiesen (o. S. 31). Dasselbe lehrt nun die Synthese, bei der mir der nämliche Sänger, Dr. ABRAHAM, behilflich war, indem er  $g^1$  auf den 5 „Hauptvokalen“ als Mustern für die Nachbildung falsettierte. Sie konnten durch folgende Zusammenstellungen nachgebildet werden<sup>1)</sup>:

<sup>1)</sup> Der Ton  $f\bar{i}s^4$  ist hier statt des in der neuen Einrichtung fehlenden  $\bar{f}^4$  eingefügt, aber in dieser Höhe tun unharmonische Teiltöne dieselben Dienste.

	$g^1$ Falsett					$g^1$ Brust- stimme
$h^4$						2
$g^4$						4
$\overline{f}is^4$			6			4
$d^3$			3	8		4
$h^3$			8	0		4
$g^3$			0	0		4
$d^3$			0	0		6
$g^2$	8	12	14	4	8	10
$g^1$	12	12	8	12	12	4
	U	O	A	E	I	A

Zur Vergleichung ist wieder das A des Helden Tenors in Bruststimme auf demselben Ton in synthetischer Nachbildung daneben gesetzt. Es erreichte freilich nicht die Stärke des Originals, selbst wenn dieses mit mäßiger Kraft in großer Entfernung angegeben wurde; aber die Qualität schien mir gut getroffen.

Am interessantesten ist, daß im Falsett alle 3 dunklen Vokale nur Grundton und Oktave in verschiedenen Stärkeverhältnissen erfordern. Daß O dem U näher rückt, ist schon a. a. O. erwähnt; aber hier gehört auch A zu derselben Gruppe, wie es denn in der Tat gegenüber dem A der Bruststimme, auch der von Frauen, dunkler klingt. In letzterem muß auf  $g^1$  überall mindestens noch  $d^3$  enthalten sein. Aber auch verglichen mit der Resonanztabelle des falsettierten  $f^1$  ist die gegenwärtige wieder bedeutend vereinfacht, weil es sich dort um die Vokale in unmittelbarer Nachbarschaft, hier um entfernter gesungene handelt.

Frauen- und Männerstimmen sind, wenn sie den gleichen Ton singen, wenigstens im Chorgesange an der Klangfarbe schwerlich zu unterscheiden. Dies ist mir in Oratorien aufgefallen, wenn sich die Tenor- und die Altstimme überschneiden, z. B. in HÄNDELS „Israel in Ägypten“ Nr. 8: „ein Dunkel, daß niemand sah“ (ppo); und so oft. Wirklich zeigt selbst das A des Helden Tenors in der synthetischen Nachbildung keinen wesentlichen Strukturunterschied gegen die Frauenstimme auf  $g^1$  S. 176, nur größere absolute Stärken.

Dagegen besteht nach den Erfahrungen der musikalischen Praxis, insbesondere des Chorgesanges, ein durchschnittlicher Unterschied zwischen Frauen- und Knabenstimmen. Ich habe darum die S. 30 erwähnten Knabenstimmen, besonders die des N., auch synthetisch nachzubilden versucht und noch während des Druckes zwei mir durch den Dirigenten empfohlene Knaben des Berliner Domchores geprüft. Schon bei der direkten Vergleichung, wenn der nämliche Vokal auf der nämlichen Tonhöhe abwechselnd von einer guten Frauen- und von einer guten Knabenstimme gesungen wurde, ließ sich leicht erkennen, daß er im 2. Falle schwächer, aber auch weicher und doch zugleich magerer klang. Vergleichende Proben mit der Resonanzmethode ergaben, daß die höchsten Teiltöne der Frauenstimme in der Knabenstimme schwächer waren oder fehlten. Der Grundton hingegen war stärker.



Beim A auf  $f^1$  war  $c^3$  auch diesmal auffallend schwach. Die synthetische Nachbildung eines A auf  $c^2$  gelang am besten mit folgenden Teiltonstärken<sup>1)</sup>:

	$c^2$	$c^3$	$g^3$	$c^4$
Frauenstimme:	6	8	6	4
Knabenstimme:	10	8	4	0

Ähnliches ergab sich für A auf der Tonhöhe  $g^1$ :

	$g^1$	$g^2$	$d^3$	$g^3$	$h^3$
Frauenstimme:	3	8	6	4	4
Knabenstimme:	6	10	5	4	0

Nun sind freilich Knabenstimmen unter sich recht verschieden und Frauenstimmen noch mehr. Es würde sich also fragen, inwieweit diese individuellen Stimmen als typische gelten können. Soweit ich aus dem durchschnittlichen Erfahrungseindruck urteilen kann, möchte ich die Frage bejahen. Aber zu bestimmter Aussagen wären natürlich umfassende Vergleichen erforderlich. Die Schwäche der einzelnen Knabenstimmen, mit welcher auch die Intensitätsverschiebungen der Teiltöne zusammenhängen dürften, wird im Chorgesange durch die Menge ausgeglichen, aber der Klangfarbenunterschied bleibt. Im übrigen ist in der ganzen Frage auch nicht zu übersehen, daß Knabenchöre in den bekanntesten Fällen auch räumlich weit gesondert aufgestellt werden, was zu der eigenartigen Wirkung nicht wenig beiträgt.

Mädchen gleichen Alters pflegen von den Dirigenten nicht zu den beabsichtigten Wirkungen (z. B. für den Choral im Eingangchor der Matthäuspassion oder in der Abendmahlsszene des Parsifal) herangezogen zu werden. Ich versuchte daher, um dem Unterschied auf den Grund zu kommen, die Methoden auch auf solche anzuwenden. Nach dem direkten Eindruck scheinen sie im allgemeinen noch etwas schwächer als die Knabenstimmen. Resonanzproben ergaben nur schwache Reaktionen. Die Synthese des A auf  $c^2$  lieferte bei einer angenehmen Mädchenstimme die Stärken 7, 9, 4, 2; auf  $g^1$  8, 7, 6, 4 („schönes helles A“). Aber allgemeinere Schlüsse lassen sich hier noch weniger ziehen, und es bleibt mir zweifelhaft, ob von weiteren Vergleichen ein deutlicher Unterschied mit einer gewissen Regelmäßigkeit zu erwarten wäre, es sei denn in der durchschnittlichen Stärke. Es wären hier zunächst Erkennungsversuche anzustellen, um zu erproben, ob von guten und musikerfahrenen Hörern die Knaben- von der Mädchenstimme bei gleicher Stärke überhaupt mit einiger Sicherheit unter-

<sup>1)</sup> Die größere Gesamtstärke der Frauenstimme tritt hier nicht zutage, weil die Dame in größerer Entfernung singen mußte, um den Stärkeunterschied zu tilgen und nur den qualitativen hervortreten zu lassen.

schieden werden kann, was beim Vergleich der obigen Frauen- und einer guten Knabenstimme entschieden der Fall war. Hier aber, beim Unterschied der Knaben- und Mädchenstimme, bin ich nicht dazu gekommen, da der durchschnittliche Unterschied so gering ist, daß man vielfältige Wiederholungen machen müßte, um einigermaßen beweiskräftige Zahlen zu bekommen.

Auch ob die Unterschiede der individuellen Stimmlage innerhalb desselben Geschlechts, wie zwischen Baß, Bariton und Tenor oder Alt, Mezzosopran und Sopran, sich in der Klangstruktur in einer regelmäßigen Weise geltend machen, scheint mir noch ungewiß. Man muß dazu natürlich den nämlichen Grundton wählen. Denn derselbe Vokal, auf  $G$  und auf  $g^1$  gesungen, muß unvermeidlich der Klangfarbe nach verschieden sein, auch wenn die Zahl und das Stärkeverhältnis der Teiltöne gleich sind (vgl. 15. Kap.). Es kann also nur gefragt werden, ob sich regelmäßige Strukturunterschiede zeigen, wenn ein Bassist und ein Tenorist den gleichen Vokal auf dem gleichen Ton, etwa  $g$  oder  $c^1$ , singen. Vielleicht sind bei einer ausgeprägten Baßstimme die tieferen Teiltöne, bei einem typischen Tenor die höheren relativ kräftiger und diese auch zahlreicher, unbeschadet gleicher Stärke innerhalb des Formanten. Ganz identisch könnte allerdings die Vokalität des A dann nicht sein, aber der daran gewöhnte Hörer würde die etwas dunklere, breitere Färbung bei dem Bassisten doch der Stimme und nicht dem Vokal zurechnen. Nun sind aber schon die individuellen Unterschiede innerhalb derselben Stimmgattung groß (Helden- und lyrischer Tenor). Darum bedürfte es wieder einer bedeutenden Zahl von Vergleichen, um über diese an sich nicht sehr wichtige Frage etwas Triftiges zu sagen.

In einer Göttinger Dissertation von O. INTRAU „Experimentell-statistische Singstimmenuntersuchungen“ (1924) ist sie behandelt, aber nicht in einwandfreier Weise. Abgesehen von technischen Mängeln des die Kurven aufnehmenden Apparates (zu tiefen Eigentons der Membran), wurde nicht konstant der nämliche Ton, sondern wurden „Töne von  $f$  aufwärts und abwärts“ gesungen; wie es scheint, bald dieser, bald jener. Seine Ergebnisse spricht der Verfasser so aus: „Die dunklere Klangfärbung der Baßstimme ist dadurch bedingt, daß nur wenige Obertöne, die der Bildung des Vokals dienen, in erheblichem Maße hervortreten; das Hinzufügen eines oberen Verstärkungsgebietes verursacht die hellere Klangfarbe, die dem Bariton eigen ist, während die gleichzeitige erhebliche Verstärkung auch der dem Grundton benachbarten Obertöne die noch schärfere und vollere, metallreichere Klangfarbe des Tenors bewirkt“. Ich kann diesen Ergebnissen aus obigen Gründen kein Gewicht beimessen.

## 8. Kapitel

# Die Entwicklung der Vokalforschungen und die Konvergenz ihrer Ergebnisse.

Nachdem unsere Analysen durch die Synthesen bestätigt sind, wird es sich lohnen, die früheren Ergebnisse, deren rein statistische Übersicht im 1. Kap. einen fast entmutigenden Anblick bot, noch einmal und etwas näher ins Auge zu fassen. Denn selbstverständlich muß in so vielen und vielfach ausgezeichneten Untersuchungen auch eine Fülle von Tatsachen stecken. Wirklich zeigt sich, daß die scheinbar so divergierenden Angaben zumeist doch nur Teile eines umfassenderen Sachverhaltes darstellen und daß die neueren Beiträge mehr und mehr, die neuesten so gut wie ganz unter sich zusammenstimmen. Wir gehen teils nach methodologischen Gesichtspunkten, teils (innerhalb derselben) chronologisch vor. Vollständigkeit ist nicht beabsichtigt, da uns nur sachliche, nicht historische Interessen leiten. Auf Berührungen im einzelnen ist auch im vorausgehenden öfters hingewiesen. Unter II. soll dann an mehreren Untersuchungen, die sich auf einen prinzipiell abweichenden Standpunkt stellen, gezeigt werden, daß sich alles Tatsächliche darin gleichfalls in den gemeinsamen Rahmen einfügt und daß auch so nur wieder die einzige wirklich durchführbare Anschauung, d. i. im wesentlichen die HELMHOLTZsche, bestätigt wird.

## I. Hauptergebnisse der Vokalforschung seit WILLIS.

### 1. Resonanz- und Interferenzmethode. Subjektive Analyse.

Den Anfang der heutigen experimentellen Vokalforschung bilden die praktischen Bestrebungen KRATZENSTEINS und v. KEMPELENS, die sich unabhängig voneinander seit etwa 1770 um eine Nachbildung der Vokale durch Apparate bemühten. Besonders näherte sich KEMPELEN dem natürlichen Vorgang, in dem er über einer schwingenden Zunge eine konische Röhre anbrachte, die verschieden abgestimmt werden konnte. Der eigentliche Vater der

gegenwärtigen Forschungsweise aber und zugleich der erste Vertreter der „Absolut-Theorie“ war R. WILLIS (1829). Anknüpfend an v. KEMPELEN baute er eine aus 2 übereinandergeschobenen Zylindern bestehende ausziehbare „Vokalröhre“, mit der er einer schwingenden Zunge von gleichbleibender Tonhöhe verschiedenen Vokalcharakter erteilen konnte. Er beobachtete ferner, daß Pfeifen von bestimmter Höhe auch schon allein einen gewissen Vokalcharakter hatten. Die Höhen, die er angibt, weichen nicht allzuviel von unseren Formantzentren ab, alle liegen aber etwas höher, und diese Abweichung wächst nach den helleren Vokalen hin. Nur U ist auch ihm unbestimmt;  $O c^2$ ,  $Ao es^2 - g^2$ , A im engl. „Part“  $des^3$ , in „Pad“  $f^3$ , E in „Pay“  $d^4$ , E nach I (in „Pet“)  $c^5$ , I  $g^5$ . Auf rein physikalische Gründe dieser Abweichungen hat bereits HELMHOLTZ (S. 190) hingewiesen. Erstaunlich bleibt, daß WILLIS überhaupt auf diesem Wege den wahren Werten so nahegekommen ist. Er hebt auch bereits hervor, daß Sängern in der Höhe kein richtiges U und O mehr erzeugen können. Auch seine psychologischen Beobachtungen sind nicht ohne Wert. Besonders betont er den Einfluß des Kontrastes beim Sprechen: sobald man einen Vokal lang aushalte, werde er unkenntlich.

CHARLES WHEATSTONE, der 1837 über diese Versuche berichtete<sup>1)</sup>, schloß daraus, daß die Vokaleigenschaft eines Stimmtones auf einem ihn begleitenden höheren Tone beruhe, der infolge des Mitschwingens der im Mund eingeschlossenen Luft zustande komme, wenn ihr Eigentone ein Vielfaches des an sich einfachen Stimmtones darstelle. Daß hier ein theoretischer Irrtum des großen Physikers vorlag, ist bereits S. 10 erwähnt. Nur dann tritt Mitschwingung ein, wenn in dem erregenden Klang ein entsprechender Oberton real vorhanden ist (d. h. wenn die Welle eine solche Gestalt besitzt, wie sie durch Übereinanderlagerung mehrerer Sinusschwingungen von entsprechenden Längen- und Höhenverhältnissen entsteht<sup>2)</sup>),

<sup>1)</sup> Da das mit C. W. unterzeichnete Referat in „London and Westminster Review“ allgemein, auch von HELMHOLTZ, WHEATSTONE zugeschrieben wird, nehme ich die Echtheit an, obgleich darin mehrmals in der dritten Person auf WHEATSTONE Bezug genommen wird (S. 35 bei Erwähnung der multiplen Resonanz — „a subject first investigated by Professor WHEATSTONE“ — und S. 39, wo eine Sprechmaschine erwähnt ist, die Prof. WHEATSTONE nach dem Vorbilde der KEMPELENSCHEN mit einigen Verbesserungen gebaut habe und die in der Instrumentensammlung des King's College in London stehe (wo WH. damals lehrte).

<sup>2)</sup> Die in Klammern stehende Erläuterung ist veranlaßt durch eine Bemerkung KOEHLERS (I, II, S. 75), der Mißverständnisse fürchtet, wenn man von einem zusammengesetzten Klange rede, der Obertöne als reale Teile enthalte.

und dann ist es eben Erregung durch einen gleichgestimmten Ton, was WHEATSTONE ausdrücklich in Abrede stellt. Daß nicht jeder Klang für multiple Resonanz gleich geeignet sei, hat er schon bemerkt (S. 36), aber nicht, daß dieser Unterschied von der ungleichen Anzahl der Teiltöne herrührt. Erst auf Grund der OHM-HELMHOLTZschen Vorstellungsweise über die Zusammensetzung der Klänge aus Sinusschwingungen war volle Klarheit auch über die Resonanzfunktion der Mundhöhle zu gewinnen<sup>1)</sup>. WHEATSTONES Verdienst liegt nicht in diesem theoretischen Unterbau zu WILLIS' Versuchen, sondern in seinen eigenen Beobachtungen über die Resonanz der Mundhöhle auf vorgehaltene Stimmgabeln, wodurch er die Resonanztheorie der Vokale vorbereitete.

DONDERS erkannte 1857 gleichfalls die Einstellung der Mundhöhle als maßgebend und zog zur Ermittlung der entsprechenden Tonhöhen Beobachtungen über Flüstervokale heran, über die schon im 6. Kap. gesprochen wurde. Der hochverdiente Physiolog hat hier nicht ganz den richtigen Weg beschritten, da die Flüstertöne als solche nicht einmal über das Wesen der geflüsterten, noch weniger über das der stimmhaften Vokale direkten Aufschluß geben können. Seine Beobachtungen selbst konnten wir nur teilweise anerkennen. Später (1864) nahm er Kurven für 14 verschiedene Vokale auf, welche die prinzipielle Richtigkeit der (von HELMHOLTZ damals bereits verfochtenen) Absoluttheorie bestätigten und, wenn sie auch natürlich noch an Vollkommenheit zu wünschen übrigließen, doch als erste systematische Anwendungen der graphischen Methodik auf Vokale einen wesentlichen Fortschritt bedeuteten.

Auf HELMHOLTZ' Leistung kann man nur mit immer erneuter Bewunderung zurückblicken. Indem er die Klangfarbenunterschiede auf die Zusammensetzung aus Teiltönen zurückführt, die Vokale unter die Klangfarben ordnet, ihre Eigentümlichkeiten und Unterschiede in den „charakteristischen Tönen“ findet, diese selbst mit Resonanzmethoden viel eingehender und sicherer als seine Vorgänger bestimmt, auch zum 1. Male bei den hellen Vokalen die unteren Maxima feststellt und eine Fülle feiner Einzelbeobachtungen über die menschliche Stimme einflicht, endlich durch seine geniale Synthese der Vokale aus wenigstens annähernd einfachen Tönen die Analyse bestätigt, stellt er die ganze Vokallehre

<sup>1)</sup> Vor OHM und HELMHOLTZ verstand man unter „einfachen Tönen“ Einzelklänge gegenüber Zusammenklängen, also z. B. den Klang einer Pfeife gegenüber dem dreier Pfeifen. Dieser Sprachgebrauch ist auch bei den Abhandlungen von WILLIS und WHEATSTONE im Auge zu behalten. Daß auch der Einzelklang Obertöne mit sich führt, war zwar nicht unbekannt, aber man zog nicht die Folgerung, daß die Teiltöne überhaupt den Klang ausmachen.

auf die Grundlage, von der sie sich nur zu ihrem Schaden zeitweilig wieder entfernt hat. Nur beachtete er zu wenig, daß fast durchweg nicht ein einzelner Ton, sondern eine Mehrheit gleichzeitiger Teiltöne formierend wirkt, und daß auch das Formantzentrum notwendig Verschiebungen innerhalb einer gewissen Zone erleidet.

Daß die Resonanzbreite der Mundhöhle mindestens eine Quinte beträgt, hebt er S. 183ff. hervor und leitet daraus die Möglichkeit ab, die Vokale zu unterscheiden, auch wenn der „charakteristische Ton“ selbst nicht zu den harmonischen Teiltönen des Stimmtones gehört. Aber er scheint den gleichzeitig mitresonierenden Teiltönen in der Nachbarschaft des charakteristischen keine Bedeutung für die Charakteristik des Vokals zugeschrieben, sondern diese ausschließlich auf den einzelnen Ton bzw. den diesem zunächstliegenden Teilton zurückgeführt zu haben. So bestätigt es mir sein Schüler MAX PLANCK. Doch ist anzunehmen, daß HELMHOLTZ einer These, die eigentlich in der Konsequenz der seinigen liegt, ohne weiteres zugestimmt haben würde.

H. GRASSMANN beobachtete die Teiltöne mit unbewaffnetem Ohre. Schon 1854 gab er an, daß ein aufmerksames Ohr leicht beim Übergange von U durch Ü zu I eine Reihe leiser, harmonischer Nebentöne höre, die von  $c^2$  bis  $c^5$  fortschreiten. Beim A könne man die Reihe noch bis zur 4-gestrichenen Oktave verfolgen. Bei den Konsonanten beobachtete er in den „Halbvokalen“ auch eine Menge unharmonischer, schwer voneinander unterscheidbarer Töne und in den Zischlauten ganz hohe Töne, wie sie keinem Vokal mehr angehörten. Er muß ein außerordentliches Gehör besessen haben<sup>1)</sup>. 1877 — nach HELMHOLTZ — sagt er im Vertrauen darauf sogar allzu kühn: „Das Ohr übertrifft bis jetzt alle künstlichen Hilfsapparate.“ Auch in dieser eingehenderen Studie konstatierte er eine mit der Helligkeit wachsende Reihe von harmonischen Teiltönen, kam aber dadurch fälschlich zu einer Relativ-Theorie, wenigstens für die mittleren Vokale.

F. AUERBACHS Ergebnisse nach der Resonatoren- und Klopfmethode (s. I. Kap.), die ihn gleichfalls zu Zugeständnissen an die Relativtheorie führten, scheinen in der Tat zunächst stark von den unsrigen abzuweichen. Aber genauer betrachtet zeigen sich sehr wesentliche Übereinstimmungen und ist die Diskrepanz öfters nur scheinbar. So ergaben die Resonanzversuche (1) eine von U bis A wachsende Zahl der Teiltöne und eine feste Lage des maximal starken Teiltons beim Hinaufrücken des Grundtones; z. B. beim hellen A ist es für den Grundton  $c$  der 6., für  $g$  der 4., für  $c^1$  der 3., für  $g^1$  der 2., mithin jedesmal  $g^2$ . In den Klopfversuchen (2) findet er: U  $f^1$ , O  $a^1$ , volles A  $f^2$ , scharfes A  $g^2 - b^2$ . Dies können wir

<sup>1)</sup> Auch AUERBACH erwähnt übrigens gelegentlich (4, S. 150), daß er in zahllosen Fällen die verschiedensten Obertöne mit Leichtigkeit aus Vokalen herausgehört habe, niemals aber unharmonische.

unterschreiben, nur würden wir  $f^1$  als einen speziellen Fall des an sich beweglichen U-Formanten betrachten. Wenn dann für Ö  $gis^1—a^1$ , für Ä  $c^2—d^2$ , für Ü  $e^1—f^1$ , für E  $g^1—a^1$ , für I  $f^1$  angegeben werden, so ist klar, daß hier einfach statt der Formanten die Unterformanten gefunden sind, diese wieder mit bemerkenswerter Genauigkeit.

Auch bei einigen späteren Untersuchungen sind für die helleren Vokale infolge der angewandten Methoden nur die Unterformanten herausgekommen. So gibt KATZENSTEIN bei der Analyse graphisch aufgenommener Kurven für E als Maximum  $g^1$ ; ebenso sagt RUBENS von der Reaktion seiner „Flammenröhre“ auf E, sie sei der eines O zum Verwechseln ähnlich. Ferner vgl. KAISERS oszillographische E- und I-Kurven bei GUTZMANN (5) S. 97 und die unten folgenden Ergebnisse von BENJAMINS mit KUNDTschen Staubfiguren.

In AUERBACHS Gesamtdarstellungen treten die obigen Befunde leider vor manchen starken Unrichtigkeiten zurück, und die Übersicht von „Durchschnittswerten“ der Formanten in seinen „Grundlagen der Musik“ (1911, S. 132) ist durchweg recht weit von der Wirklichkeit entfernt.

Hier seien sogleich die viel später (1916), aber gleichfalls nach der Klopfmethode erfolgten Beobachtungen O. ABRAHAMS abgeschlossen. Sie ergaben durch Beklopfen des Stirnbeins, der Backe und durch Plessimeterbeklopfen vor dem Munde an verschiedenen männlichen und weiblichen Personen bei Einstellung auf Vokale überall Resonanztöne, die in unsere Formantzonen fallen, allerdings mit gewissen Modifikationen. Für U =  $f^1—a^1$  gilt das schon Gesagte. O =  $c^2—d^2$  deutet auf eine besonders helle Aussprache (Oa). Ebenso A =  $h^2—d^3$ . Die Resonanztöne der helleren Vokale stimmen ohne weiteres. Über die von ABRAHAM beobachteten Flüsterhöhen ist bereits S. 152 ff. gesprochen.

Die Ergebnisse der If.-Versuche von GRÜTZNER (1891) und SAUBERSCHWARZ (1895) verstehen sich leicht aus den im 2. Kap. beschriebenen Erscheinungen, wenn man nur immer berücksichtigt, daß bei ihren Lücken- und Stichversuchen die ungeraden Multipla der direkt ausgeschalteten Töne mitwegfielen. Auch das Übergehen in nasalierte Formen durch solche künstliche Lücken, namentlich beim O und A, und das Übergehen aller hellen Vokale in ein tiefes, unbestimmtes Brummen bei Ausdehnung der If. bedarf keiner weiteren Erläuterung. Ebenso, daß die Wegnahme des Grundtons und seiner Oktaven (nebst deren ungeraden Vielfachen) alle Vokale in einfache, nur durch die Stärke unterschiedene „Pffiffe“ verwandelte. Was dabei Pffiffe genannt wird, sind wohl einzelne schwache hohe Teiltöne, die durch diese Einstellungen nicht hinreichend mitausgeschlossen wurden<sup>1)</sup>.

<sup>1)</sup> Vgl. o. S. 48. Rein rechnerisch müßten eigentlich alle harmonischen Teiltöne auf diese Art vernichtet werden; wenn man auch nur die Töne 1, 2, 4, 8, 16 ausschließt, müßten damit schon sämtliche Teiltöne bis zum

Die If.-Versuche KOEHLERS und SCHOLES s. unter 5.

Systematische Ab- und Aufbauversuche nach dem Vorbilde der von mir (11) beschriebenen hat 1922 R. PAPALE an den italienischen Flüstervokalen U, O, A, E, I angestellt und fast genau dieselben Formanten gefunden. Doch beginnt E beim Aufbau erheblich tiefer, und zwar als Öe an derselben Stelle, an der nach unserer Tabelle Ö beginnt, während die obere Grenze wieder mit unserer E-Grenze zusammenfällt. Als Unterformant des E erscheint statt O ein Ao, der Unterformant des deutschen Ä. Diese Modifikationen dürften, wie auch PAPALE hervorhebt, auf einer verschiedenen Aussprache des E im Italienischen beruhen. I wird beim Aufbau um eine kleine Terz höher „gut“ (bei *fis*<sup>5</sup>), „vollkommen“ aber erst bei *c*<sup>7</sup>. Eine Art von Vervollkommnung bis fast zur Tongrenze fand sich bei den Flüstervokalen auch in meinen Versuchen bei den helleren Vokalen häufig; aber ich rechnete diese letzte Strecke nicht mehr zum Formanten, da sich der Vokalcharakter selbst nicht mehr merklich veränderte und der Laut nur an Stärke und Klarheit noch zunahm. Auch „freier“ und namentlich „näher“ waren häufige Bezeichnungen in solchen Fällen<sup>1</sup>).

PAPALE findet Unterschiede auch bei U, O und A, sofern sie beim Aufbau an höheren Stellen der Tonreihe begännen und höhere Formanten hätten als unsere deutschen Flüstervokale. Dies wäre gewiß möglich und würde einen neuen Beitrag zur akustischen Festlegung von Aussprache-Verschiedenheiten liefern. Doch kann ich es in seiner Tabelle nicht deutlich bestätigt finden; seine Formanten fallen auch hier so gut wie vollständig mit den meinigen zusammen.

## 2. Graphische Methoden.

Diese sind seit dem SCOTT-KOENIGSchen „Phonautographen“ (1859) außerordentlich vervollkommnet worden. Bei der großen Zahl und Ausdehnung dieser Untersuchungen können wir nur an

31. vernichtet sein. Aber in Wirklichkeit können aus den früher angegebenen Gründen einige höhere Multipla, z. B. etwa die ungeraden Teiltöne vom 17. ab, übrigbleiben, und diese schwachen, unanalysierten Komplexe hat wohl SAUBERSCHWARZ als „Pfeife“ bezeichnet. Daß sie bei den helleren Vokalen stärker sind als bei den dunkleren, ist selbstverständlich. Ich würde allerdings statt von Pfeifen eher von Knistern oder Wispern reden.

<sup>1</sup>) Mit Rücksicht auf diese letzten geringfügigen Vervollkommnungen sind im Schema o. S. 107 noch punktierte Zonen oberhalb der Formanten für Ö—I eingetragen, die in den Abhandlungen 11 und 12 fehlen, obgleich die Erscheinungen selbst auch dort (11, S. 244, 246) erwähnt sind. Die Punktierungen sind jetzt bis zu den äußersten Grenzen geführt, die sich aus den Versuchsprotokollen ergeben.



einigen den Grad ihrer Übereinstimmung unter sich und mit unseren Ergebnissen aufzeigen.

HERMANN, dessen unermüdlicher Sorgfalt die Phonographie in erster Linie ihre Ausbildung verdankt, gab 2 Reihen von Formantbestimmungen, eine 1890, eine 2. 1905 (diese für lange Vokale). Die 1. ergab:  $U = c^2 - d^2$ ,  $O = d^2 - e^2$ ,  $A = e^2 - gis^2$ ,  $E = h^3 - e^4$ ,  $I = d^4 - g^4$ . Die 2.:  $U =$  Anfang der 1. und der 2. Oktave,  $O =$  Anfang der 2. Okt.,  $A =$  Mitte der 2. Okt. ( $f^2 - a^2$ ),  $E =$  Anfang der 2. und Ende der 3. Okt.,  $I =$  Mitte der 4. Okt. Diese letzten Angaben erscheinen, obwohl unbestimmter, doch richtiger. Beim  $U$  ist zu dem früher allein angegebenen (etwas tiefen) Oberformanten der Formant hinzugetreten,  $O$  muß wieder beide Male sehr hell angegeben sein, beim  $E$  ist das 2. Mal der Unterformant hinzugetreten, beim  $I$  fehlt er auch diesmal. Zuletzt (1911) setzt HERMANN für  $A$  „mindestens eine Quart, von  $e^2$  bis  $a^{2''}$ , ja, wenn man die undeutlicheren  $A$ -Laute mit gelten lasse, sogar eine Sext,  $des^2$  bis  $ais^2$ ; für  $E$  mindestens eine kleine Terz,  $c^4$  bis  $dis^4$ . Er betont hier (und schon vorher mehrfach), daß der Formant innerhalb dieser Grenzen variieren könne, ohne daß der Vokal wesentlich leide. Dies ist eine wichtige Erkenntnis. Aber ein gleichzeitiges Zusammenwirken mehrerer Formanttöne statuiert auch er nicht; es würde sich seiner Anschauung auch viel weniger leicht einfügen als der von HELMHOLTZ. Er behauptet ausdrücklich, daß für ein auf  $c$  gesungenes  $A$  nur 2 Schwingungszahlen, die des Stimmtones  $= c$  und die des Mundtones  $=$  etwa  $g^2$ , wesentlich und zur Charakteristik hinreichend seien (Bd. 48 S. 566). Wir fanden doch zu einer vollbefriedigenden Synthese hier ein Dutzend Teiltöne vonnöten; 2 können nur ausnahmsweise unter entgegenkommender psychischer Einstellung genügen.

Wenn man die (bei GUTZMANN 5, S. 104 abgedruckte) Tabelle der  $A$ -Formanten, wie sie von HERMANN nach der „Schwerpunktmethode“ für die Grundtöne der diatonischen Leiter zwischen  $G$  und  $d^1$  ausgerechnet wurden, durchsieht, so ist trotz mancher Schwankungen eine leichte Erhöhung des Formanten mit steigendem Grundton nicht zu verkennen; aber sie beträgt nur einen halben, höchstens ganzen, Ton. Diese geringe Erhöhung dürfen wir wohl auch nach den Erörterungen o. S. 65 ff. als Ausdruck der Tatsachen betrachten<sup>1)</sup>.

<sup>1)</sup> BOEKE, ein Anhänger der HERMANNschen Theorie, zog denn auch 1891 aus mikroskopischen Messungen an Phonographenglyphen die Folgerung, daß der  $A$ -Formant, wenn die Stimme von  $c$  bis  $d^1$  stieg, mit in die Höhe ging, und zwar ergab sich ihm bei einer bestimmten Berechnungsweise ein Steigen von 865 bis 1182 Schw., d. h. von  $a^2$  bis  $d^3$ . Das wäre sehr viel.

Ist es gestattet, einen Durchschnittswert zu berechnen, so wäre er 772,5 Schw. =  $g^2$ . Diesen Ton betrachtet man in der Tat meistens als den „HERMANNschen A-Formanten“, für welchen auch GUTZMANN eintritt. Er ist auch in unseren Versuchen immer als besonders ausschlaggebend hervorgetreten.

Sehr schöne Übereinstimmung mit den Ergebnissen unserer A-Synthesen zeigt die Tafel der Amplitudenverhältnisse für ein auf  $c^1$  gesungenes A<sup>1</sup>):

$$c^1 = 0, \quad c^2 = 0,41, \quad g^2 = 0,54, \quad c^3 = 0,40, \quad e^3 = 0,11.$$

Dies ist dasselbe Bild wie bei dem von mir ganz unabhängig davon immer wieder ausprobierten synthetischen A auf  $c^1$ .

Daß HERMANN die Formanten im allgemeinen unharmonisch zum Grundton sein läßt, ist ein Irrtum, tut aber dem Werte seiner experimentellen Bestimmungen keinen Eintrag. Von großem Interesse sind die starken Verschiedenheiten der Kurvenformen, die der nämliche Vokal bei aufeinanderfolgenden Tönen der Leiter ergibt<sup>2</sup>). Solche müssen herauskommen, wenn der Formant eine feste Lage hat, einerlei ob er harmonisch oder unharmonisch zum Grundton ist. Sie sind dafür ein schöner Beleg, aber nicht für die HERMANNsche Lehre.

Es ist unbegreiflich, wie HERMANN die Ähnlichkeit der Vokalcurven mit Schwebungskurven als Stütze dafür ansehen konnte. Kurven mit regelmäßig wechselnden Maximis und Minimis müssen bei jeder Kombination zweier oder mehrerer Töne entstehen, sowohl wenn sie einander in der Tonreihe dicht benachbart sind, wie bei den schwebenden Tönen, als auch bei weiteren Intervallen. Physikalisch ist hier gar kein wesentlicher Unterschied<sup>3</sup>). Daß wir im 1. Falle Diskontinuitäten und Rauigkeiten hören, hat physiologische Gründe (infolge der Resonanzbreite werden durch benachbarte Töne zwischenliegende Fasergruppen gemeinschaftlich erregt). Bei weiteren Intervallen fallen diese Gründe weg. Aber physikalisch müssen auch sie ganz ebenso als ein Auf und Ab der Wellengipfel erscheinen. PIPPING, AUERBACH, KOEHLER u. a.

Aber die Berechnungsweise und mehr noch die (auch vorher von JENKING und EWING, nachher von STEVANI benutzte) mikroskopische Messungsmethode sind Einwänden ausgesetzt; weshalb sie in neuerer Zeit aufgegeben bzw. durch Vergrößerungen auf mechanischem Wege (SCRIPTURE, LIORET) ersetzt wurden, die aber auch wieder zum Teil ihr Bedenkliches haben.

<sup>1</sup>) Bd. 47; abgedruckt von NAGEL S. 781 (der diese 1. Untersuchung als „klassische“ bezeichnet) und von GUTZMANN 5, S. 103.

<sup>2</sup>) Vgl. z. B. die bei AUERBACH 5, S. 697 abgebildeten U-Kurven auf den Tönen der diatonischen Leiter.

<sup>3</sup>) Vgl. sämtliche Kurven in m. Abh. 6.

haben dies bereits genügend hervorgehoben<sup>1)</sup>. Die schönen A-Kurven, von denen HERMANN sagt (Bd. 61, S. 175): „Deutlicher kann die Selbständigkeit des Formanten gegenüber der alten Lehre von der Verstärkung harmonischer Teiltöne gewiß nicht bewiesen werden“, sind also nichts weniger als beweisend.

Ohne Kenntnis von HERMANN'S Arbeiten hat der Amerikaner BEVIER 1900 ff. mit einer phonographischen Einrichtung Vokale auf Grundtönen von *B* bis  $d^2$  aufgenommen und durchgerechnet. Wenn man die amerikanische Aussprache berücksichtigt, dürften seine Werte ziemlich den Tatsachen entsprechen. Das A-Maximum gibt er wesentlich höher an (ungefähr 1150 Schw. =  $d^3$ ), bemerkt aber selbst, daß dieses A, wie in „father, papa“, zwar mit dem italienischen, aber keineswegs mit dem deutschen zusammenfalle. Damit stimmt, daß er auch noch ein unteres Maximum für A bei 675 =  $f^2$  fand, welches der Unterformant des Ä ist, während unser A keinen Unterformanten besitzt.

SAMOJLOFF, der HERMANN'S Untersuchungen fortsetzte, fand:  $U = c^1 - g^1, c^2 - e^2$ ;  $O = h^1 - des^2$ ;  $A = g^2 - a^2$ ;  $E = h^1 - des^2, h^3 - des^4$ ;  $I = c^1 - g^1, c^2 - e^2, d^4 - e^4$ . Alles in guter Übereinstimmung mit unseren Befunden, ausgenommen den Zwischenformanten  $c^2 - e^2$  beim I, der auf einer Zufälligkeit der Artikulation beruhen muß (SAMOJLOFF selbst versieht ihn ebenso wie die Unterformanten für E und I mit einem Fragezeichen; die letzteren sind aber zweifellos richtig bestimmt). Nicht unmöglich wäre hier ein Einfluß des slawischen Y (o. S. 196).

S. glaubte, wie BOEKE, aus seinen Tabellen schließen zu können, daß im allgemeinen mit steigendem Grundton auch der Formant langsam steige, was ja gleichfalls richtig ist. Aber in der mitgeteilten Tabelle für A auf den Tönen der diatonischen Leiter von G bis  $c^1$  steigt er nur von 758 bis 845 Schw., d. h. von  $g^2$  bis  $a^2$ . Bemerkenswert ist, daß selbst diese kleine Steigung durch ein Fallen unterbrochen ist: während die Stimme von H nach *c* geht, fällt der Formant von 804 auf 755 Schw. und steigt dann wieder langsam. Dies ist begreiflich, wenn es sich um einen harmonischen Teilton handelt (s. o. S. 67), aber nicht bei einem Formanten in HERMANN'S Sinne. Es würde sich wohl sogar zahlenmäßig ableiten lassen, wenn genaue Bestimmungen für die Höhe der Grundtöne gegeben wären. Aber diese wurden nur nach dem Gehör geschätzt.

<sup>1)</sup> AUERBACH S. 699: „Schwebungskurven entstehen auch bei Konsonanzen, z. B. bei der Kombination eines Tones mit seiner Duodezime, und man könnte hier erklären: das ist gar kein Akkord, sondern das ist die im Tempo des Grundtones an- und abschwellende Duodezime; daß man den Grundton zu hören glaubt, ist eine Täuschung. Niemand wird bezweifeln, daß diese Auffassung des konsonanten Akkords keinen Fortschritt, sondern einen Rückschritt bedeuten würde.“

PIPPING (Helsingfors), der HELMHOLTZ' Auffassung gegen HERMANN verteidigte, kam bei seinen zahlreichen sorgfältigen Versuchen mit HENSENS Sprachzeichner zu gleichen oder nur wenig abweichenden Verstärkungsgebieten. Er wählte, wie seine Vorgänger, verschiedene Grundtöne aus der kleinen und 1-gestrichenen Oktave, wobei die Formantzentren notwendig innerhalb gewisser Grenzen wechseln müssen. Die Kurven analysierte er nach FOURIER bis zum 15. Teilton. In manchen Fällen enthalten seine Tabellen aber nur die Unterformanten. Außer den Amplituden berechnete er auch die Intensitäten ( $a^2 \cdot n^2$ )<sup>1)</sup> und erhielt dadurch, namentlich bei den höheren Teiltönen, viel stärkere Ausschläge, zuweilen sogar beträchtliche Verschiebungen der Maxima. Im einzelnen bleiben manche Inkongruenzen in seinen Tabellen, die zum Teil auf Eigentönen der benützten Membran beruhen mögen (so liegt beim U auf  $c^1$  wiederholt ein ungeheures Maximum auf  $e^3$ ); aber im allgemeinen fügen auch sie sich in den nunmehr feststehenden Rahmen. Es sei von Einzelheiten nur angeführt, daß das A-Maximum in seinen ersten Versuchsreihen (Grundtöne 190, 224, 393, 557 =  $g$ ,  $ais$ ,  $g^1$ ,  $des^2$ ) immer in der Gegend  $des^3$ — $d^3$  liegt, in den späteren Reihen (Grundtöne  $c$  und  $c^1$ ) bei  $c^3$  bzw.  $e^3$ . Doch nehmen die Teiltonstärken überall schon von  $g^2$  an zu. Wir haben es offenbar wieder mit einem besonders hellen A zu tun<sup>2)</sup>. Ob etwa finnische Einflüsse hier und beim U mitspielen, kann ich nicht beurteilen. In einer späteren Arbeit (3, S. 30 ff.) zeigen die Tabellen bei Erhöhung des Grundtons von 128 bis 412 Schw. ein nicht von PIPPING hervorgehobenes, aber deutliches Ansteigen des Maximums von  $a^2$  ( $h^2$ ) bis  $c^3$ , ja  $dis^3$ .

O. WEISS und S. GARTEN, deren Ergebnisse für die Konsonanten S und Sch wir S. 129 ff. erwähnten, untersuchten mit dem „Phonoskop“ auch gesungene und besonders geflüsterte Vokale. Aus den erhaltenen Kurven bestimmten sie durch Auszählung der den längeren Wellen aufgesetzten Schwingungszacken die Höhen der Flüstertöne. WEISS erhielt folgende Werte:

<sup>1)</sup> In den späteren Abhandlungen gibt P. nur die Amplituden, findet aber (4, S. 159) beide Formeln nicht voll geeignet zum Ausdruck der gehörten Stärken, da die tieferen Teiltöne in Amplituden zu stark, in Intensitäten zu schwach herauskämen, und hält es mit LLOYD für möglich, daß  $n \cdot a$  die Verhältnisse am besten wiedergäbe. Diesen Ausdruck legte schon OHM zugrunde. Vgl. m. Tonpsych. I, S. 370.

<sup>2)</sup> PIPPING selbst statuiert (2, S. 558f.) für A ein zweigipfliges Verstärkungsgebiet mit  $gis^2$ — $a^2$  und  $cis^3$  als Gipfeln. Doch finde ich in den Tabellen des A auf  $c$  und  $c^1$  auch in der 2. Abhandlung nur wieder  $c^3$  als Maximum innerhalb der breiteren Verstärkungszone  $g^2$ — $c^3$ . P. hat wohl noch andere Tabellen im Auge.

	Schwing.	Noten	Mittel	Noten
U	400—600	$g^1-d^2$	450	$b^1$
O	550—710	$des^2-ges^2$	610	$es^2$
A	700—840	$f^2-as^2$	760	$g^2$
E	2200—2600	$des^4-e^4$	2500	$es^4$
I	2500—3160	$es^4-g^4$	2900	$ges^4$

Sowohl die Mittelwerte als die Gesamtzonen liegen durchgängig innerhalb unserer nach der If.-Methode bestimmten Flüsterformanten, die nur mehr oder weniger nach beiden Seiten darüber hinausgehen. S. GARTEN bestätigte im wesentlichen diese Ergebnisse. Für das Flüster-A schätzt er den Formanten = 981—1000 ( $h^2$ ), etwas höher als WEISS. Bei den gesungenen Vokalen erhielt er fast genau gleiche Kurven wie HERMANN. Auch er fand (1915), daß die Formantschwingungen des A sich mit steigendem Grundton etwas erhöhen.

Weitere Untersuchungen stellten GARTEN und KLEINKNECHT mit einem kugelförmigen Resonator an, dessen Größe in weniger als einer Sekunde soweit verkleinert werden konnte, daß sein Resonanzton gleichmäßig um 3 Oktaven oder mehr anstieg. Der Innenraum des Resonators war mit einem Schallschreiber verbunden, der die jeweilige Stärke des Mitschwingens automatisch registrierte. Der Mund des Singenden war etwa 6 cm von der Resonatoröffnung entfernt. Nur U, O und A wurden so untersucht, weil die Einrichtung nicht auf Teiltöne über  $h^3$  reagierte. Der Grundton erwies sich stärker als bei anderen graphischen Registrierungen, nahm aber von U nach A ab. In einer Versuchsreihe, die auch E und I enthielt, nahm er gegen I wieder stark zu, ganz entsprechend unseren Befunden (o. S. 27, 75). Die höchsten noch beträchtlich verstärkten Teiltöne lagen für U um etwa 600 ( $d^2$  = Oberformant), für O um 780 ( $g^2$ ), für A um oder über 1000 Schw. ( $c^3$ ).

Aber die beiden Forscher glauben auf diesem Wege auch unharmonische Teiltöne nachgewiesen oder wenigstens wahrscheinlich gemacht zu haben. So trat bei einem A auf tiefem Grundton eine resonatorisch verstärkte Schwingung auf, bei der  $9\frac{1}{2}$  Teilschwingungen auf eine Grundtonperiode kamen (S. 39, vgl. 24, 27, 34). Die Verfasser geben aber zu (S. 28), daß für strikte Anhänger der HELMHOLTZschen Theorie ihre bisherigen Versuche vielleicht noch nicht hinreichend überzeugend seien. Wir haben uns, ohne irgendeiner Theorie von vornherein anzuhängen, von den Tatsachen überzeugen lassen, daß jedenfalls die Vokalfärbungen nur von harmonischen Teiltönen herrühren. Wie aber die unharmoni-

nischen im vorliegenden Falle hineingekommen sind, kann man begreiflicherweise, ohne mit dem Apparat selbst gearbeitet zu haben, nicht sagen. Über einen anderen Versuch GARTENS in der gleichen Angelegenheit ist schon oben S. 55 berichtet.

Vor kurzem sind durch FERD. TRENDLENBURG im Berliner Siemenslaboratorium die Hilfsmittel der Radiotechnik in erfolgreicher Weise für die graphische Methode herangezogen worden. Durch ein Kondensator-Mikrophon wurden die akustischen Schwingungen in Frequenzmodulationen einer elektrischen Hochfrequenzwelle verwandelt, diese vergrößert und durch eine Verstärkerschaltung zum Steuern einer hochabgestimmten Oszillatorschleife benützt. Verzerrungen werden dabei so gut wie völlig vermieden. Aus den Klangbildern der 5 „Hauptvokale“ und der Zischlaute wurden die wichtigsten Schwingungszahlen durch Abzählung der Zacken erschlossen. Die Ergebnisse stimmen mit den unsrigen überein (Vgl. auch o. S. 26, 130).

Ähnlicher Art sind die gleichzeitigen Untersuchungen von CRANDALL und SACIA. Sie lieferten wieder die charakteristischen Maxima (wobei die Teiltonstärken ausgedrückt werden durch Multiplikation der Amplitude mit der Empfindlichkeit für die betreffende Tonhöhe).

Die Versuche von RAPS (Ann. d. Phys. 1893) und GEHLHOFF (Zeitschr. f. Phys. 1920), die Luftschwingungen ohne Vermittlung eines auf Eigentöne abgestimmten Zwischengliedes zu photographieren, sind leider in den Anfängen stecken geblieben. Doch hat RAPS einige Formanten angegeben, von denen wenigstens der des A ( $f^2 - a^2$ ) mit den zuverlässigsten sonstigen Ergebnissen übereinstimmt.

### 3. Bestimmung des energetisch stärksten Teiltones.

Mit der von H. RUBENS konstruierten „Flammenröhre“ (einer Manometerröhre mit Seitenlöchern), wodurch nur der stärkste Bestandteil eines Klanges festzustellen war, fanden RUBENS und KRIGAR-MENZEL, daß sie bei allen Vokalen auf den Grundton reagierte, wenn dieser in der 2-gestrichenen Oktave oder höher lag, ferner für O und E auf Grundtöne oberhalb  $e^1$ , endlich für U, Ü, I in allen Tonlagen. Dies entspricht ganz unseren Befunden, wenn man im Auge behält (was auch die beiden Forscher selbst betonen), daß es bei E, Ü und I nur die unteren Maxima sind, die auf diese Art herauskommen. Beim A zeigte die Flamme je nach der dunkleren oder helleren Vokalisierung  $f^2$  bis  $b^2$ , einmal sogar  $d^3$ . Im allgemeinen lag der wirksamste Teilton des A zwischen  $gis^2$  und  $c^3$ , am häufigsten um  $b^2$ . Für O, Ö und E lag er bei  $g^1$  (für Ö und E wieder Unterformant).

Auf dem Hamburger Phonetiker-Kongreß 1914 berichtete ZWAARDEMAKER über Untersuchungen, bei denen ein auf  $c^1$  gesungener Vokal zu 5 auf seine ersten 5 Teiltöne abgestimmten Resonatoren und weiter zu einem RAYLEIGH'schen Scheibchen geleitet wurde, das nur bei dem energetisch stärksten Ton erzitterte. Beim U auf  $c^1$  war dies der 1., beim O der 2., beim A der 3., beim I wieder der 1. Teilton. Die Ergebnisse (kurz berührt auch in 3, S. 11) stimmen genau mit den Resonanztabellen o. S. 25.

C. E. BENJAMINS, der den energetisch stärksten Ton („Hauptton“) von Vokalen auf den Grundtönen  $g$  bis  $g^1$  durch KUNDT'sche Staubfiguren aufsuchte, erhielt für die helleren Vokale wieder ihre Unterformanten, ohne jedoch deren Funktion zu erkennen. Im allgemeinen fand er, daß mit steigendem Grundton der Stimme der Hauptton seiner Ordnungszahl nach sinkt (was notwendig ist), aber doch seiner absoluten Höhe nach bei gleichem Vokal nicht fest bleibt (was gleichfalls innerhalb gewisser Grenzen aus der Absolut-Theorie folgt). Er schließt, daß der energetisch stärkste Teilton nicht zugleich der charakteristische Ton des Vokals zu sein brauche. Auch diese Folgerung kann zutreffen. Vor allem könnten die Unterformanten recht wohl physikalisch stärker sein als die Formanten: die physiologischen Stärken könnten sich dabei gleichwohl umgekehrt verhalten. Aber selbst wenn sie den physikalischen genau parallel gingen, wäre es nicht notwendig, daß der stärkste Teilton zugleich die stärkste färbende Kraft (Vokalvalenz) besäße. Gegen die HELMHOLTZ'sche Lehre würde daraus kein Einwand erwachsen; sie enthält keine Bestimmungen über die letzten Fragen der Gehirnmechanik.

#### 4. Synthesen.

Dieser Weg ist, wie erwähnt, nach HELMHOLTZ außer von MILLER und dem Verfasser nur von wenigen und nur bruchstückweise betreten worden, aber auch die so erzielten Ergebnisse lassen sich von den jetzt gewonnenen aus gut verstehen.

LAHR, der Resonanzgabeln verwendete, erhielt U am reinsten durch eine einzige Gabel, wenn sie nicht zu hoch war; O durch  $c + c^1$ , noch besser durch die 4 Gabeln  $c$ ,  $c^1$ ,  $g^1$ ,  $c^2$ . Bei Hinzufügung des  $g^2$  änderte sich der Klang nach A hin, und mit der Gabel 1000 (nahe  $c^3$ ) wurde dieses besser. Sobald man aber nur eine dieser Gabeln für sich anschlug, sei es auch die höchste, ging der Laut in U über. Dies ist alles richtig beobachtet, wenn auch das extrem helle U der Gabel mit 1000 Schw. kaum mehr diese Bezeichnung verdient. Daß eine so primitive Methode doch Ergebnisse lieferte,

die so gut mit den Tatsachen stimmen, ist wohl nur einem für Vokalitäten besonders feinen Ohr des Beobachters zu verdanken.

v. WESENDONK fand mit Flaschentönen, daß beim I ein tiefer Ton oder mehrere beteiligt sein und daß diese mit den hohen verschmelzen müssen, ferner daß es für ein gutes A nicht genüge, den Formantton  $b^2$  allein zum Grundton hinzuzufügen, daß Teiltöne, die für sich allein wie U klingen, z. B.  $d^1$  und  $d^2$ , in ihrer Verbindung ein O geben können, und andere wesentliche Tatsachen.

Kürzlich hat STRUYCKEN (4) wieder Synthesen mit Stimmgabeln, aber in exakterer Durchführung als LAHR, für U, O, A veröffentlicht. Die Stärken sind dabei nach subjektiver Schätzung in den ganzen Zahlen 1—3 angegeben. Da die Gabeln aber nur bis 860 Schw. ( $a^2$ ) reichten, begreift sich, daß in den Tabellen nur einmal ein reines A (mit hohem Maximum auf  $e^2$ ) vorkommt, und dieses dürfte auch mehr dem Kontrast zu noch dunkleren Färbungen seine Bezeichnung verdanken<sup>1</sup>).

Auch unter Benützung elektrischer Schwingungen sind bereits Synthesen versucht worden. So hat J. Q. STEWART 1922 durch Unterbrechungen eines elektrischen Stromes im Rhythmus des Stimmtones und Einschaltung zweier auf den Formanten und Unterformanten abgestimmten Schwingungskreise Vokale hergestellt, die trotz der von STEWART selbst hervorgehobenen Unvollkommenheit der Einrichtung doch in 50% der Fälle erkennbar waren. Die Formanthöhen waren die schon bekannten, für A in father 500 + 1000. Ähnliches berichtet PAGET (1) von Dr. ECCLES 1922. Freilich werden hierbei niemals, solange die elektrischen Schwingungen durch Telephone oder Lautsprecher in Töne übersetzt werden, rein sinusförmige Schwingungen kombiniert. Andererseits sind wieder zu wenig Töne beteiligt. Diese Synthesen gehören daher zunächst mehr in die Klasse der bloßen „Nachahmungen“ und können zur Formulierung der theoretischen Grundtatsachen vorläufig kaum etwas hinzufügen.

##### 5. Verknüpfung mehrerer Methoden.

KOEHLER ging von den o. S. 5 erwähnten graphischen Aufnahmen am Trommelfell, an denen zunächst mehr die Möglichkeit des Unternehmens überhaupt als das Detail der Kurven von Inter-

<sup>1</sup>) STRUYCKEN hat sich auch um die graphische Methode durch seinen vorzüglichen Aufnahmeapparat verdient gemacht und mir freundlichst Proben seiner Ergebnisse mitgeteilt, die gut zu den unserigen stimmen. Er ist brieflich (1919) auch auf die allgemeinsten, im 11. bis 13. Kap. zu behandelnden Fragen eingegangen, aber leider bisher nicht zur Veröffentlichung seiner Forschungen darüber gekommen.



esse war, zu If.-Versuchen über, beschritt dann aber einen fast neuen Weg mit Beobachtungen über den Vokalcharakter einfacher Töne, die ihn zuletzt auch zu elementaren synthetischen Versuchen führten. Seine If.-Versuche dienten, soweit sie Vokale betrafen, hauptsächlich der Entscheidung über das Vorhandensein unharmonischer Teiltöne (o. S. 37). Auf seine Lehre von der Vokalität einfacher Töne wird uns die Schlußbetrachtung (13. Kap.) zurückführen. In synthetischer Beziehung hebt KOEHLER mehrfach hervor, daß die von ihm als A, E, I bezeichneten einfachen Töne den wirklichen Vokalen ähnlicher werden, wenn leise tiefe Töne hinzutreten. Die tiefen Resonanztöne des E und I hat er auch bereits richtig mit O und U identifiziert. Später bemerkte er, daß alle Teiltöne, die eine A-Valenz besitzen, wozu alle zwischen  $c^2$  und  $c^4$  gehören sollen, für die A-Färbung mitverantwortlich seien. Darin liegt, daß die gleichzeitig gegebenen Vokalvalenzen innerhalb der Formantzone zusammenwirken, wenn auch diese selbst zu weit gezogen ist. Aus der Verbindung zweier einfachen Töne läßt KOEHLER eine Vokalfarbe entstehen, die in dieser Deutlichkeit an keinem von beiden zu bemerken ist; wenn z. B. durch ein *fis*<sup>1</sup> und ein *fis*<sup>2</sup> (bei bestimmtem Stärkeverhältnis) ein gutes O entstehe, während keiner der Komponenten für sich ein reines O darstelle. In soweit können wir überall nur beistimmen.

In der trefflich durchgeführten Dissertation H. SCHOLÉS (1917, mir seit 1921 bekannt) sind nur die Vokale U, O, Ä, A, A (unser Ää), diese aber auf mehreren Wegen untersucht: teils durch subjektive Beobachtung der durch die Mundstellung verstärkten Teiltöne, teils objektiv durch Verknüpfung der MARBESCHEN Rußmethode (graphischer Aufnahmen rußender Flammenbilder) und stroboskopischer Analyse von Flammenbildern, zugleich immer mit Hinzuziehung des If.-Verfahrens<sup>1</sup>). Die Arbeit erhält besonderes Gewicht dadurch, daß der Verfasser ursprünglich (auf Grund einer vorhergehenden Arbeit WITTMANN'S, die auf ungenauer Auslegung von Rußbildern beruhte) für die HERMANN'SCHE Lehre eingenommen war, aber durch die zwingenden Ergebnisse seiner eigenen Untersuchung zu HELMHOLTZ zurückgeführt wurde. Er fand ausschließlich harmonische Teiltöne. Für U war ein starker Grundton

<sup>1</sup>) Die Handhabung der If.-Methode ist im ganzen korrekt, insbesondere ist auf die Mitausschaltung der ungeraden Multipla geachtet. Doch scheint übersehen zu sein, daß in einem A noch höhere Teiltöne als  $e^3$  vorhanden sein können und daß diese durch Vernichtung gewisser tieferer mitwegfallen müssen. So ist in dem S. 40 erwähnten Falle mit dem 3. auch der 9., mit dem 2. der 10. Teilton vernichtet, also *fis*<sup>3</sup> und *gis*<sup>3</sup>, die bei einem so hellen A nicht ganz einflußlos sind.

erforderlich, außerdem Töne zwischen  $b^1$  und  $a^2$  (Oberformant). Für O mußten Teiltöne zwischen  $b^1$  und  $d^2$  stark sein, sekundär spielten noch Resonanztöne bis  $gis^2$  mit. Für A ergab sich die kurze Strecke  $b^2$ — $cis^3$  als entscheidend. Hier muß es sich wieder um ein sehr helles A gehandelt haben. Auch O scheint besonders hell gegeben zu sein. SCHOLE nennt sein A ein norddeutsches. Immerhin bezeichnen auch norddeutsche Forscher wie HERMANN, GUTZMANN u. a. ein A mit dem Formantzentrum  $g^2$ , höchstens aber  $b^2$ , als das beste d. h. von O und Ä gleich weit entfernte<sup>1)</sup>.

#### 6. Die Analysen und Synthesen D. CH. MILLERS.

Die ausgiebigste Bestätigung finden unsere Ergebnisse durch die 1916 veröffentlichten umfassenden Analysen und Synthesen D. Ch. MILLERS, über die hier eingehender berichtet werden soll, da das Buch in Deutschland wenig bekannt ist. MILLER registrierte die Schallkurven durch seinen „Phonodeik“, eine verfeinerte phonophotographische Einrichtung nach HERMANN'SCHEM Prinzip. Die erhaltenen Kurven wurden an Stelle der mühsamen Messung und Ausrechnung durch einen verbesserten mechanischen Analysator, wie solche mehrfach konstruiert sind<sup>2)</sup>, in ihre sinusförmigen Elementarschwingungen zerlegt. Diese konnten durch einen weiteren komplizierten Apparat („Synthesizer“) auch wieder zusammengesetzt und dadurch die Probe auf die Analyse gemacht werden, die sehr befriedigend ausfiel. Insoweit war das Gehör überhaupt nicht beteiligt.

Nun wurde die gehörte Tonstärke als proportional mit  $a^2 \cdot n^2$  angenommen und 61 Pfeifen zwischen 129 und 4138 Schw. nacheinander so angeblasen, daß sie gleichlaute Töne ergaben. Es zeigte sich, daß die Phonodeikkurven starker Korrekturen bedurften, um mit den Aussagen des Gehörs zu stimmen (infolge der Eigen-

<sup>1)</sup> Nicht verständlich ist mir die Angabe (S. 45), daß bei einem auf  $c^1$  gesungenen Ä (Aä), wenn  $e^3$  und  $e^3$  herausgenommen wurden und nur  $c^1$ ,  $c^2$ ,  $g^2$  (wahrscheinlich doch auch  $g^3$ ) übrigblieben, U gehört wurde. Auch wenn  $c^2$ ,  $g^2$  (und  $g^3$ ) nur schwach in dem Klang enthalten waren, konnten sie doch nicht so ohne Einfluß bleiben. So würde ich bei einigen Angaben die Genauigkeit der Beschreibung anzweifeln. Aber weitaus die meisten entsprechen sicher der Wirklichkeit.

<sup>2)</sup> Auf die älteren Einrichtungen dieser Art darf man nicht zu viel Vertrauen setzen. Wenn z. B. GARTEN (2, S. 290) mit dem MADERSCHEN Analysator von 1909 in einem Stimmgabelklang 9 Teiltöne fand, meist freilich sehr schwach und schon in das Gebiet der Messungsfehler fallend, den 7. aber doch so stark wie den 2., so müssen hier wohl Fehler vorliegen. Eine gewöhnliche Stimmgabel enthält, auch wenn sie stark gestrichen wird, außer dem bekannten unharmonischen Beiton nur eine deutlich ausgeprägte Oktave und eine minimale Duodezime.

töne des Trichters, der Membran usw.) Diese Korrekturen wurden ein für allemal ausgerechnet und danach die wirklichen physikalischen Energien der ermittelten Teilschwingungen in jedem Einzelfalle bestimmt.

Mit 32 gedackten Pfeifen, deren Stärke durch die Anblasevorrichtung reguliert wurde, bildete sodann MILLER die analysierten Klänge auch synthetisch nach.

Das allgemeinste Ergebnis seiner Untersuchungen ist, daß die Vokale aus harmonischen Teiltönen mit bestimmten Verstärkungsgebieten von fester absoluter Lage bestehen. Diese sind in folgender Tabelle angegeben (worin neben MILLERS Wortbeispielen, durch die er die englische, bzw. amerikanische Aussprache festlegt, unsere Schreibweise beige setzt ist).

Formanten nach MILLER.

	Vokal	Schwing.	Noten
1. moo, gloom	U	326 <sup>1)</sup>	<i>e</i> <sup>1</sup>
mow, no	O	461	<i>b</i> <sup>1</sup>
maw, raw	Ao	732	<i>fis</i> <sup>1</sup>
ma, father	A	922 <sup>2)</sup>	<i>b</i> <sup>2</sup>
2. mat, pat	Aä	800, 1840	<i>gis</i> <sup>2</sup> , <i>b</i> <sup>3</sup>
met, pet	Äe	691, 1953	<i>j</i> <sup>2</sup> , <i>h</i> <sup>3</sup>
mate, they	E	468, 2461	<i>h</i> <sup>1</sup> , <i>dis</i> <sup>4</sup>
meet, bee	I	308, 3100	<i>dis</i> <sup>1</sup> , <i>g</i> <sup>4</sup>

Formanten wie Unterformanten decken sich mit unseren Befunden so gut als man nur wünschen kann. Unser Ö und Ü hat MILLER nicht in seine Forschungen aufgenommen.

Von besonderem Interesse sind wieder die Ergebnisse über das von verschiedenen Stimmen auf verschiedenen Höhen gesungene A. Der verstärkte Ton liegt immer zwischen *f*<sup>2</sup> und *d*<sup>3</sup>, meistens ist er *b*<sup>2</sup>, also der HELMHOLTZsche A-Formant<sup>3)</sup>. Bei einem von einer Sopranstimme auf *c*<sup>2</sup> gesungenen A ist neben dem ziemlich starken

<sup>1)</sup> Hierfür steht S. 225 362 (= *fis*<sup>1</sup>), was auf einem Druckfehler beruhen dürfte. Freilich stimmen die Angaben S. 225, 237, 259 überhaupt nicht genau miteinander; vielleicht liegen verschiedene Versuchsreihen zugrunde.

<sup>2)</sup> An einigen Stellen steht statt des A in father = 922 ein anderes, das durch die Beispiele ma und pat erläutert und durch 1050 (*c*<sup>3</sup>), 950 (*h*<sup>2</sup>), 1240 (*dis*<sup>3</sup>) definiert wird. Vgl. die Noten S. 260. Aber ma gilt auch als Parallelbeispiel zu father. Die Darstellung S. 225–230 ist mir in dieser Hinsicht nicht ganz klar geworden.

<sup>3)</sup> In v. WESENDONKS o. S. 170 erwähntem Bericht ist eine Zusammenstellung gegeben. Natürlich ist aber nicht anzunehmen, daß in allen diesen Fällen genau dasselbe A (in derselben Helligkeitsnuance) gesungen wurde. Daß meistens gerade *b*<sup>2</sup> verstärkt war, darf also nicht zu sehr urgiert werden;

Grundton eine Verstärkung bei  $c^3$ , aber auch noch bei  $g^3$  deutlich bemerkbar, übereinstimmend mit unseren Angaben. Bei einem auf *fis* gesungenen A war  $fis^2$  am stärksten; woran man sieht, daß ein geringes Aufsteigen des Formanten doch auch in MILLERS Ergebnissen liegt, obgleich er es nicht erwähnt.

MILLER hat richtig erkannt, daß nicht ein einzelner Teilton den Vokal charakterisiert, sondern daß (im allgemeinen) ein ganzes Verstärkungsgebiet vorhanden sein muß, innerhalb dessen das Maximum liegt. In diese Region fällt nach seinen Intensitätskurven fast die ganze physikalische Energie des Klanges (oft etwa 90% nach S. 226). Zuweilen verteilt sie sich auf eine ziemlich weite Strecke, z. B. bei dem auf *f* gesungenen A auf  $c^2$  bis  $f^3$ . Aber zu beiden Seiten dieser Strecke ist die physikalische Klangstärke beim A fast Null. Auch dies findet in unseren Synthesen seine Bestätigung und weist darauf hin, daß die tieferen Töne durch Differenztonbildung subjektiv ersetzt oder verstärkt werden.

Auch von geflüsterten Vokalen hat MILLER mit dem Phonoidek 45 Kurven aufgenommen, die durch Auszählung der Wellengipfel im wesentlichen dieselben Formanten lieferten wie die gesungenen, nur ein wenig erhöht (S. 235). Auch dies entspricht unseren Beobachtungen, wenn wir bei den dunklen Flüstervokalen die gehörten Töne als Überblasungstöne ansehen (o. S. 164).

Wie weit die Naturtreue seiner Synthesen reicht, kann man, da unwissentliche Versuche fehlen, nicht beurteilen. Ein Mangel der Einrichtung war neben der nicht völligen Einfachheit der Komponenten die unzumutbare Stärkeregelung durch die Windzufuhr wegen der unausbleiblichen Verstimmungen. Daß die Intensitäten in physikalischen Maßen angegeben sind, ist auf der einen Seite ein Vorteil gegenüber bloß subjektiver Schätzung; auf der anderen Seite aber kommen wir durch diese den physiologischen Stärken, auf die es zuletzt ankommt, sicherlich näher<sup>1)</sup>. Die Tatsache, daß von  $c^2$  an eine Alteration der Vokale eintritt, scheint MILLER nicht, jedenfalls nicht in ihrem vollen Umfange, bekannt. Aber es liegt mir fern, an der mit bewunderungswürdiger Exakt-

und um etwa auf das durchschnittliche amerikanisch-englische A daraus zu schließen, wären die Fälle wieder zu wenig zahlreich. Übrigens sind hierzu auch BEVIERS Angaben (o. S. 210) zu vergleichen, dessen A in „father“ an der Grenze steht, wo mit dem Formanten  $d^3$  zugleich die Abspaltung eines Unterformanten beginnt.

<sup>1)</sup> Bei dem Korrektionsverfahren mit den 61 Pfeifen waren subjektive Stärkevergleichen doch unentbehrlich. Dabei ist auch wohl noch die Frage, ob nicht ein konstanter Fehler schon bei der Vergleichung je zweier unmittelbar benachbarter Töne mitwirken könnte, der sich von Paar zu Paar summierte.

heit und Geduld durchgeführten Untersuchung mäkeln zu wollen. Nicht häufig dürften ganz unabhängig voneinander und mit völlig verschiedenen Methoden ausgeführte Untersuchungen eines verwickelten Problemgebietes zu solcher Harmonie der Ergebnisse gelangt sein. Und man darf wohl daraus schließen, daß die Frage nach der akustischen Natur der Vokale als im wesentlichen gelöst zu gelten hat.

## II. Kritik einiger pseudo-synthetischen Versuche der neueren Zeit.

Die meisten der nach v. KEMPELEN immer wieder versuchten, mehr oder minder erfolgreichen Vokalnachahmungen, die nur mit Unrecht als Synthesen bezeichnet werden, übergehen wir hier. Nur auf 3 darunter soll eingegangen werden, die mit dem Anspruch auftreten, als experimentelle Beweise bestimmter theoretischer Grundanschauungen angesehen zu werden.

### 1. HERMANN'S Versuch mit der Doppelsirene.

An dieser von HELMHOLTZ konstruierten Lochsirene lassen sich 2 gleichzeitige Töne in beliebigem Intervall zueinander angeben. HERMANN gab nun z. B.  $e^2$  und  $g^2$  und hörte dann ein A auf dem Grundton  $c$ , der als Differenzton auftritt. Aber der Klangbestand ist hier nur sehr unvollständig angegeben. Nach unserer Kenntnis der Sirenenklänge und der Differenztonbildung (vgl. m. Abh. 7, S. 133) müssen eine große Menge von Tönen auftreten, unter denen als besonders stark folgende in Betracht kommen: 1. die Primärtöne selbst, 2.  $c$  und  $c^2$  als 1. und 2. Differenzton, 3.  $e^3$  und  $g^3$  als erste Obertöne, 4.  $c^1$  und  $c^3$  als deren Differenztöne. Man hat also mindestens (abgesehen von höheren Obertönen) mit dem Komplex  $c$ ,  $c^1$ ,  $c^2$ ,  $e^2$ ,  $g^2$ ,  $c^3$ ,  $e^3$ ,  $g^3$  zu rechnen. Das sind aber so ziemlich alle Teiltöne, die zu einem guten A auf  $c$  gehören (vgl. unsere Tabelle o. S. 176), zwar nicht genau in der nötigen Stärkeabstufung, doch wenigstens mit Bevorzugung der Formantregion.

Ob man nun in solchen Fällen, bei immerhin verschobenen Stärkeverhältnissen, wirklich ein A hört oder nur einen Mehrklang, das hängt sehr von der Einstellung und den musikalisch-akustischen Gewohnheiten des Beobachters ab. Für Unmusikalische ist die Wirkung jedenfalls leichter zu erzielen, weil diese von vornherein an einheitliches Hören gewöhnt sind. Ich habe mit mehreren meiner jungen Freunde diesen Versuch mit der Doppelsirene wiederholt; wir hörten aber bei der ihnen wie mir natürlichen Einstellung (doch ohne absichtliche Analyse) zunächst nur eben die beiden

schreienden Primärtöne und den brummenden 1. Differenzton, aber keine Spur von Vokalität.

Nimmt man eine Dissonanz, wie  $f^2 + g^2$  (8 : 9), so resultieren (wenn wieder nur die beiden ersten Obertöne berücksichtigt werden) die Töne 1, 2, 7, 8, 9, 16, 18 =  $F, f, \bar{e}s^2, f^2, g^2, f^3, g^3$ , also eine zwar auch harmonische, aber stark lückenhafte Teiltonreihe. Doch mag auch dann für Unmusikalische, zumal wenn sie auf Vokale vorbereitet sind, etwas A-Ähnliches entstehen; vermutlich wird es stark nach Ä hin liegen. Jedenfalls muß es infolge der raschen Schwebungen (87 p. Sek.), die auf den tiefen Grundton  $F$  übertragen werden, wie von einer rauhen Bierbaßstimme klingen.

Was also an Vokalähnlichkeit bei dem Doppelsirenenversuch herauskommt, ist leicht aus HELMHOLTZ' Lehre herzuleiten.

## 2. JAENSCH' Versuche mit der Selensirene.

Nach E. R. JAENSCH, der, von HERMANNschen Voraussetzungen ausgehend, eine besondere Theorie aufgestellt hat, ist ein Vokal dann gegeben, wenn sinusförmige Einzelschwingungen von etwas verschiedenen, aber um einen Mittelwert schwankenden Längen aufeinanderfolgen<sup>1)</sup>. Da nun unser Gehör auf unperiodische Schwingungen durch Geräuschempfindungen reagiere, so seien die Vokale eben Geräusche, und zwar — nach der ursprünglichen Fassung der Theorie — sogar die eigentlichen Qualitäten des Geräuschsinnes<sup>2)</sup>. Die Besonderheit des einzelnen Vokals hänge ab von der mittleren Länge, um welche herum die Wellenlängen schwanken. Liege diese z. B. in der Gegend des  $b^1$ , so höre man ein O, liege sie bei  $c^3$ , ein A. (In dieser Lehre von den Vokalitäten der einfachen Töne schließt sich JAENSCH an KOEHLER an.) Die verschiedene Höhe des Grundtons (Stimmtons) aber entstehe dadurch, daß solche „gemischte Sinuskurven“ auf reine Sinuskurven von der Länge des Grundtones aufgesetzt seien.

<sup>1)</sup> Auch wenn nur Amplituden- oder selbst bloße Phasenverschiedenheiten sich regelmäßig wiederholen, sollen Vokale auftreten, am besten, wenn mehrere dieser „Störungsfaktoren“ sich verbinden. Wir halten uns hier der Einfachheit halber an die Wellenlängen.

<sup>2)</sup> Hier lag freilich eine innere Unklarheit. Denn danach müßten die Vokale um so deutlicher und eindringlicher auftreten, je ausgeprägter der Geräuschcharakter einer Gehörsempfindung ist, während JAENSCH selbst öfters hervorhebt, daß der Vokalcharakter durch gewisse Versuchsänderungen in demselben Maße abnehme, in welchem der Geräuschcharakter zunimmt. Neuerdings hat er denn auch diese Formulierung zurückgezogen (3) und sich der von FRÖBES hypothetisch vorgeschlagenen angeschlossen, wonach die Vokale einem dritten Sinn neben Ton- und Geräuschsinn angehörten. Daß auch diese Fassung paradox genug ist, wird man zugeben.

Man sieht, daß hier von Obertönen, wenigstens harmonischen, nicht mehr die Rede ist. JAENSCH stellt sich, wie HERMANN, außerhalb der ganzen OHM-HELMHOLTZschen Hörtheorie.

Den experimentellen Nachweis glaubte er durch eine, bereits von O. WEISS für andere Zwecke eingeführte Einrichtung zu führen. Eine Selenzelle wird in dem gewünschten Rhythmus belichtet. Sie steht in einem Stromkreis, in den auch ein Telephon eingeschaltet ist. Die durch den Belichtungswechsel erzeugten Widerstandsschwankungen werden im Telephon als Ton (bzw. Geräusch) vernommen. Um den gewünschten Rhythmus herzustellen, schaltete JAENSCH zwischen der Lichtquelle und der Selenzelle rotierende Kreisscheiben ein, auf deren äußeren Teilen Sinuskurven ausgeschnitten waren. Waren die Wellenlängen untereinander gleich, und zwar =  $24^\circ$ , also 15 auf jede Umdrehung, so hörte er einen reinen Ton (Kurve I), standen sie aber in den obenerwähnten Beziehungen, so daß z. B. Sinuswellen von 20, 24, 22, 28, 26 usf. Bogenlänge periodisch aufeinanderfolgten („gemischte Sinuskurven“), so hörte er bei 65—70 Umdrehungen der ganzen Scheibe p. Sekunde ein „gutes A“ (Kurve III). Bei noch ungleicheren Wellenlängen verwandelte sich die Empfindung in ein bloßes Geräusch.

Daß nun auch bei „gemischten Sinuskurven“, die sich in gleichen Verhältnissen periodisch wiederholen, die FOURIER-Analyse eine Anzahl harmonischer Teiltöne liefern muß, ist sicher. Aber auch abgesehen davon darf man nicht schließen, daß bei dieser verwickelten Einrichtung die auf den Scheiben ausgeschnittenen Kurven unverändert zum Ohre gelangen<sup>1)</sup>. Schon in Anbetracht der Trägheit der Selenzelle ist anzunehmen, daß aufeinanderfolgende Wellen sich superponieren. Vor allem aber liefert das Telephon niemals Sinusschwingungen einfach weiter, sondern fügt Obertöne hinzu. Eine Stimmgabel, die direkt gehört nur die Oktave und allenfalls eine minimale schwache Duodezime erkennen läßt, hört sich durchs Telephon fast wie eine Klarinette. Bei der Resonanzgabel 250 hörte ich Teiltöne bis zum 6., v. HORNBOSTEL bis zum 12., bei der Gabel 512 bis zum 5. Diese Obertöne geben auch Schwebungen mit entsprechenden Hilfsgabeln.

<sup>1)</sup> Es wäre der nämliche Fehlschluß, dem KOENIG mit seiner „Wellensirene“ und HERMANN bei Versuchen mit derselben (Bd. 48) unterlagen. HERMANN setzte seine phonographische A-Kurve darauf und erhielt ein „ausgezeichnetes A“. Der Beweis für die Genauigkeit der Kurve (die ich an sich nicht bestreite) war aber damit nicht geliefert, weil die Wellensirene, wie ich 2, S. 177 nachwies, schon bei Aufsetzung einer reinen Sinuskurve dem Tone 4 Obertöne hinzufügt; was bei der Diskontinuität der Luftstöße auch begreiflich ist.

In der Tat überzeugte mich eine experimentelle Nachprüfung mit den mir freundlichst übersandten, sorgfältig hergestellten Originalscheiben, daß auf diesem Wege hochzusammengesetzte Klänge herauskommen, die unter Umständen einem Vokal ähnlich sein können, aber nur dann und darum, wenn und weil die Zusammensetzung aus Teiltönen, wie sie die alte Theorie verlangt, gegeben ist.

Die Versuche wurden im Berliner Physikalischen Institut mit Unterstützung der Herren Dr. WESTPHAL und Dr. v. ALLESCH angestellt. Die wichtigsten Beobachtungen, die das Vorhandensein zahlreicher Teiltöne betrafen, wurden auch durch die Herren Dr. ABRAHAM und Dr. v. HORN-BOSTEL bestätigt. Der die Scheiben drehende Motor wurde so reguliert, daß er etwa 70 Umdrehungen pro Sekunde machte. Das Telephon war von vorzüglicher Beschaffenheit. Der erzeugte Klang wurde in einen vom Apparatenzimmer durch einen Zwischenraum getrennten Beobachtungsraum geleitet. Von den Beobachtungen sei folgendes erwähnt:

a) Bei einer reinen, ununterbrochenen Sinuskurve auf der Scheibe wurde nicht ein „reiner Ton“ vernommen, sondern ein ziemlich zusammengesetzter Klang, worin allerdings der der Frequenz  $15 \cdot 70$  entsprechende Ton  $c^3$  besonders stark vertreten war, außerdem aber  $d^3$ ,  $g^3$  und noch eine Anzahl leiser Töne der 3-gestrichenen Oktave. Woher diese Töne, wage ich nicht bestimmt zu sagen. Vor allem kommt das Telephon in Betracht. Aber auch der Motorton und seine Obertöne, soweit sie ins Telephon gelangen, oder der damit zusammenfallende sog. Scheibenton, der durch die bloße Umdrehung der Scheibe entsteht und ebenfalls Obertöne hat (SCHAEFER und ABRAHAM Pflügers Arch. Bd. 88, S. 475). Ferner war ein feines rasselndes Geräusch bemerklich, an das auch ein Ton,  $c^2$ , geknüpft schien, und das irgendwie mit dem Strom zusammenhing, der auch bei Ausschaltung der Lampe immer durch das Telephon der Selenzelle geht.

b) Bei der Kurve III konnten wir von der behaupteten Vokalähnlichkeit nichts finden. Aber hierbei mag eine subjektive Einstellung Unterschiede machen. Die Analyse durch das bloße Ohr ergab eine Menge von Teiltönen:  $c^2$  stark, aber auch  $g^2$ ,  $b^2$  hervortretend;  $c^3$  bis  $c^4$  schwach. Es war der Eindruck einer harmonischen Obertonreihe, deren Grundton  $c$  oder  $C$  sein würde.  $C$  würde der Schwingungszahl 65 entsprechen, die sich wahrscheinlich mit der Umdrehungszahl der Scheibe deckte. Auch an „Variationstöne“ ließe sich denken, die gleichfalls Obertöne haben (SCHAEFER und ABRAHAM). Jedenfalls waren die Töne vorhanden und war die Klangstruktur eine gänzlich andere als JAENSCH sie voraussetzt.

Vor allem aber fragt es sich, welche Teiltöne schon durch die Zerlegung der „gemischten Sinuskurve“ entstehen müssen, wenn HELMHOLTZENS Voraussetzung richtig ist, daß das Ohr die Fourieranalyse ausführe. Auf meinen Wunsch hat sich Herr stud. math. PHILIPPS in dankenswerter Weise der langwierigen Arbeit unterzogen, rein analytisch (ohne die weniger sichere Ordinatenausmessung) die Amplituden der darin enthaltenen harmonischen Teiltöne zu berechnen. Zur Vereinfachung der Berechnung wurde aber angenommen, daß die oben angegebenen ersten 5 Wellenlängen der Kurve III auf dem Gesamtumfang der Scheibe noch 2mal in der nämlichen Aufeinanderfolge wiederkehren (während in Wirklichkeit ihre Anordnung das 2. und 3. Mal eine andere und nur die Durchschnittslänge dieselbe ist). Dies kann



aber wesentliche Unterschiede nicht zur Folge haben. Es ergaben sich für die ersten 10 Teiltöne folgende Amplitudenverhältnisse in der Kosinus- und Sinusreihe, wenn die Gesamtlänge der Periode = 100 gesetzt wird:

(Kos.)  $-2,7; -0,1; -11,1; -23,4; +67,4; -14,1; +0,4; -7,3; -2,2; -2,3$ .  
 (Sin.)  $+0,6; +4,3; -8,1; +12,5; +60,1; +25,2; +4,2; +3,5; -1,0; -0,3$ .

Es liegt also ein Maximum beim 5. Teilton; aber auch seine Nachbarn nehmen noch etwas daran teil. Da nun die Umdrehung der Scheibe 65 bis 70mal pro Sekunde erfolgte, so muß die durch die 5 verschiedenen Sinuswellen gebildete Gesamtperiode, die 3mal bei jeder Umdrehung auftritt, 195–210mal pro Sekunde wiederkehren. Der Grundton muß also *as* sein. Zu diesem Grundton gehören als 4., 5., 6. Teilton: *as*<sup>2</sup>, *c*<sup>3</sup>, *es*<sup>3</sup>. Dieses Maximum mit seinem Zentrum in *c*<sup>3</sup> ist aber nichts anderes als der Formant eines sehr hellen A. Ja selbst das obere Maximum in der 4-gestr. Oktave, das einem hellen A in der Nähe eigen zu sein pflegt, dürfte in der Kurve stecken, da sich nach Herrn PHILIPPS aus dem Verhalten der Gleichungen auf eine weitere Verstärkung in der Gegend des 15. Teiltons, also des *g*<sup>4</sup>, schließen läßt.

Wenn also ein „gutes A“ von JAENSCH gehört wurde, so läßt sich eine solche Wirkung unter den HELMHOLTZschen Voraussetzungen von dieser „gemischten Sinuskurve“ nur erwarten. Nun ist sie allerdings nicht ganz identisch mit der JAENSCHschen, bei welcher die Periode erst nach einer ganzen Umdrehung genau wiederkehrt und daher der Grundton *C* sein muß. Über diese Kurve kann man nach einer brieflichen Mitteilung Herrn Dr. F. TRENDELENBURGS vom Standpunkte der Theorie auch ohne Berechnung sagen, daß die Amplitude des Grundtons nur schwach ausgeprägt, dagegen der 15. Teilton und die ihm benachbarten Töne sehr kräftig vertreten sein müssen. „Eine solche Annahme stellt die von Ihnen beschriebene Kurvenform schon so gut dar, daß die nächsthöheren Teiltöne sicher nur noch mit kleinen Amplituden vertreten sind. Erst in noch viel höheren Gebieten werden neue Teiltöne auftreten.“ Nun ist der 15. Teilton von *C* *h*<sup>2</sup>, also wieder das Formantzentrum eines hellen A, und die Teiltöne in den viel höheren Gebieten werden wahrscheinlich wieder das obere Maximum darstellen. Wir kommen somit auch für die Kurve III in ihrer Originalgestalt zu derselben Folgerung.

Das einzige, worüber man sich wundern könnte, wäre, daß der Vokal bei unserer Nachprüfung nicht ebensogut zum Vorschein kam. Aber ein ganz natürliches A ist ja auch nicht zu erwarten, da zu einem solchen doch sehr genau abgemessene Stärkeverhältnisse vom Grundton bis zur oberen Formantgrenze gehören, wie sie bei dieser Einrichtung nur durch einen ganz unwahrscheinlichen Zufall entstehen könnten. Im übrigen mag auch eine mehr analysierende psychische Einstellung ungünstig gewesen sein.

c) JAENSCH hat auch die A-Kurve aus HERMANNs photophonographischen Aufnahmen auf einer Scheibe ausgeschnitten. Mit dieser hörten wir als Grundton *es*. Das würde die Oktave des Motor- und Scheibentons sein, wenn der Motor 75,5 Umdrehungen machte, könnte hier aber auch allenfalls mit den 2 Maximis der ebensooft wiederholten A-Kurve zusammenhängen. Daneben war aber eine ganze Reihe von Teiltönen bemerklich, die harmonische Obertöne jenes Grundtones zu sein schienen: *es*<sup>1</sup>, *b*<sup>1</sup>, *g*<sup>2</sup>, *b*<sup>2</sup> und höhere sehr leise Töne. Der Gesamteindruck konnte, wenn nicht analysiert wurde, als A-ähnlich passieren. Die Erklärung ist einfach.

d) Die 3 Kurven *K*<sub>1</sub> *K*<sub>2</sub> *K*<sub>3</sub> waren so ausgeschnitten, daß auf eine periodische Welle, die 9mal bei jeder Umdrehung wiederkehrte, bei *K*<sub>1</sub> eine

periodische, bei  $K_2$  und  $K_3$  eine unperiodische Oberschwingung von der mittleren Periodenlänge  $8^\circ$  aufgesetzt war. Von diesen zackigen Kurven erhielt J. bei  $K_1$  einen Klang ohne Vokalcharakter, bei  $K_2$  und  $K_3$  einen mit Vokalcharakter. Wir hörten bei  $K_1$  einen schönen, sehr naturtreuen Klarinettenklang mit zahlreichen kräftigen Obertönen (besonders bei Tonhöhe  $c^2$ ), konnten aber bei  $K_2$  und  $K_3$  keine wesentlichen Unterschiede bemerken, obgleich wir durch wechselnde Regulierung des Motors Grundtöne verschiedener Höhe ( $c^1, c^2, f^2$ ) erzeugten. Nur etwas Geräusch mischte sich bei, aber kein Vokalcharakter.

Überall also, wo etwas von Vokaleindrücken entstand, waren die dazu nötigen Obertöne vorhanden. Sonach kann ich die Theorie der „gemischten Sinuskurven“ nur als gänzlich mißglückt ansehen. Wie gegenüber HERMANN, bleibt auch gegenüber seinem Nachfolger die OHM-HELMHOLTZsche Betrachtungsweise aufrecht, die alle Klänge als Komplexe gleichzeitiger, übereinandergelagerter Sinusschwingungen auffaßt, welche vom Ohre wieder in ihre Komponenten zerlegt werden.

Aber auch das methodische Prinzip, das JAENSCH dem bisherigen entgegenstellt: daß man von der Veränderung der Reize ausgehen und zusehen müsse, welche Empfindungsveränderungen den verschiedenen Möglichkeiten der Reizveränderung entsprechen, kann ich nicht unterschreiben. Es stellt die Sache nicht erst auf die Beine, sondern vielmehr auf den Kopf. In der gesamten Sinnespsychologie muß als oberster methodischer Grundsatz gelten, von der Analyse der Erscheinungen selbst auszugehen und die Veränderungen der Reize, ohne die es natürlich keine experimentelle Sinnesforschung gibt, prinzipiell nicht auf Grund der bloß physikalischen Möglichkeiten, sondern nach Anleitung der an den Erscheinungen beobachtbaren Eigenschaften und Veränderungen vorzunehmen. So wenigstens prinzipiell. Im einzelnen Fall wird jeder Experimentator nach Bedarf mit Veränderungen auf der einen und anderen Seite vorgehen.

In Erwiderung auf einige 1914 von mir (8) erhobene Einwände hat JAENSCH gewisse für den Vokaleindruck günstige Modifikationen seines Verfahrens empfohlen, auch einige von den früheren abweichende Angaben gemacht. Unsere Schlußfolgerungen werden aber davon nicht berührt. Später hat sein Schüler H. LACHMUND die Methode noch weiter ausgebildet, indem er auf dem Umkreis einer Scheibe 3 unter sich gleiche Serien „gemischter Sinuskurven“ anbrachte (also dieselbe Anordnung wie bei der von Herrn PHILIPPS berechneten Kurve). Dadurch entstand bei der Drehung ein deutlich hörbarer tiefer „Periodenton“, der auch als Stimmton, als Träger der durch den höheren „Frequenzton“ erzeugten Vokalität, aufgefaßt wurde. Ein solcher Stimmton war übrigens jetzt auch bei JAENSCHS A-Kurven für die meisten Versuchspersonen hörbar, während ihn JAENSCH nicht gehört und nicht erwartet hatte. LACHMUND legt nun das größte Gewicht darauf, daß diesem Ton die Aufmerksamkeit zugewandt sein müsse. Erst wenn der „Frequenzton“ in sich unbeachtet bleibe, übertrage er seine Formanteigenschaft auf den Stimmton (S. 15—20, 46, 80ff.).

Mit dieser Wendung, die für die Beobachtungsfähigkeit des Verfassers ein gutes Zeugnis ablegt, nähert er sich ganz wesentlich der richtigen Erkenntnis des Sachverhaltes. Nicht der Formant für sich, sondern sein Zusammenwirken mit tieferen Tönen und einem Grundton schafft die Vokalität. Ihr Träger ist das Klangganze bzw. der Grundton, auf dessen Höhe die des Klangganzen bezogen wird. Wir kommen aber dann eben auf die hier vertretene Auffassung zurück. Die ganze Methode ist nur eine Scheinsynthese, eine sehr künstliche, unter den bisherigen die künstlichste, Art von Nachahmung der Vokale, deren Wesen dadurch mehr verhüllt als aufgehellt wird. Daß bei einer so komplizierten Schallerzeugung auch unharmonische Teiltöne herauskommen, ist gewiß nicht unmöglich. Aber wir haben ja hervorgehoben, daß einzelne Töne dieser Art unter den höheren Klangbestandteilen die Vokalität nicht notwendig beeinträchtigen, unter Umständen sogar dazu mitwirken. Nur die Hauptmasse des Klanges und besonders die tieferen Teiltöne müssen bei einem vollkommenen Vokal harmonisch sein.

### 3. TER KUILES Versuche.

Ogleich die von TER KUILE auf höchst einfache Weise hergestellten künstlichen Vokale ungleich näher als die JAENSCHS an die natürlichen herankommen, können wir darüber nicht anders urteilen. Er erzeugte sie dadurch, daß er an eine Schachtel von der Größe, wie sie zur Aufbewahrung von Visiten- oder Briefkarten dienen, ein Korkstückchen anleimte und dieses durch eine schwingende Stimmgabel berührte. Ich erhielt auf diese Art mit einer elektrisch angetriebenen Gabel von 200 Schw. durch passende Auswahl der Kartons, auch mit den von Herrn TER KUILE bereitwilligst übersandten Kästen, ein deutliches nasaliertes Ä, ein nasaliertes OA und ein ziemlich gutes A. Man kann so, wie TER KUILE selbst bemerkt, auch eine heisere Stimme mit fast komischer Naturtreue nachahmen, wenn die Gabel dem Karton zu nahe gebracht wird<sup>1)</sup>. Die Einrichtung gehört zu den zweckmäßigsten, wenn sie keine theoretische Bedeutung beansprucht. Aber TER KUILE gründet gerade darauf eine Theorie, die wieder von HERMANN'S Grundannahmen ausgeht und die Vokale auch wieder unter die Geräusche einreicht.

In diesem Falle habe ich (mit den Herren ABRAHAM, v. ALLESCH, BAILEY, v. HORNBOSTEL als Mitbeobachtern) das If.-Verfahren angewandt, vor allem für das A auf dem Grundton 200 = *gis*. Der Vokal läßt sich natürlich wie ein gesungener ab- und aufbauen, auch mit Lückenversuchen prüfen. Es fand sich, daß er im wesent-

<sup>1)</sup> Ein ähnlicher Versuch ist schon in m. Tonps. II, 535 beschrieben: „Wenn man eine angeschlagene *a*<sup>1</sup>-Gabel auf ein den Tisch nur lose bedeckendes oder sich wulstförmig darüber erhebendes Blatt Papier aufsetzt, so wird der Ton nâselnd, leicht schnarrend, und nähert sich dem einer Oboe oder eines gewöhnlichen Zungenstimmpfeifchens. Man kann auch einen Anflug von Heiserkeit darin finden.“

lichen aus den 7 ersten harmonischen Teiltönen des Tones 200 bestand. Der Grundton selbst ist äußerst schwach. Doch läßt er sich aus der Leitung bei Abtragung aller höheren Töne noch hören. Er erscheint wieder als schöner einfacher Ton. Mit dem Hinzukommen der Oktave wird er stärker und etwas nach O hin verändert, mit dem 3. Teilton OA-ähnlich, aber etwas instrumental. Mit dem 4., der besonders stark ist ( $gis^2$ ), nähert er sich dem A, noch mehr mit dem 5. ( $his^2 = c^3$ ). Die folgenden machen ihn noch heller und dröhnender, aber nicht mehr wesentlich anders. Bei einem anderen Karton, der ein sehr starkes nasaliertes Ä (franz. in) gab, wieder sehr schwacher Grundton, beim 2. Teilton ein fagottartiger Klang, beim 3. wie eine ferne Trompete oder Posaune, beim 4. A-ähnlich, noch mehr beim 5. und 6., mit dem 7. schon zu scharf, der Oboe angenähert; usw. bis zum 12., ja 20.

Man sieht aus diesen Beispielen, daß es sich wieder um stark zusammengesetzte Klänge handelt, und daß deren Zusammensetzung in demselben Maße, wie sich die Klänge den natürlichen Vokalen nähern, auch den dort gefundenen Strukturen entsprechen. Allerdings sind hier noch Geräuschbeimischungen vorhanden, die bei gut gesungenen Vokalen viel weniger merklich sind. Aber sie haben mit der Vokalcharakteristik nichts zu tun, jedenfalls nichts im günstigen Sinne. Übrigens hat TER KUILE selbst durch resonierende Gabeln bei seinem künstlichen A maximale Resonanz für  $a^2$  (aber auch starke für die ganze Zone  $f^2-cis^3$ ) und bei einem ungewöhnlich hellen, „gleichsam kindlichen“ A Resonanz für  $c^3-f^3$ , bei einem O für  $g^1-c^2$ , bei einem AO für  $es^2$  (Unterformant) festgestellt — alles Ergebnisse, die mit den Formanten der alten Theorie ganz vortrefflich übereinstimmen. So liefern auch diese Versuche nur wieder neue eklatante Bestätigungen dafür.

Es war eben ein verhängnisvoller Fehlgriff HERMANN'S und seiner Nachfolger, die OHM'SCHE Definition des „einfachen Tones“, durch die allein HELMHOLTZ erst Klarheit in die ganze Klanglehre brachte und durch die allein auch die Begriffe „Analyse und Synthese von Klängen“ ihre scharf definierte objektive Bedeutung erhalten, preiszugeben. Es handelt sich dabei aber nicht etwa bloß um eine physikalische Definition. Vielmehr ist gerade das Wesentliche die immer wieder zu bestätigende Tatsache, daß überall da, wo die physikalisch-mathematische Analyse noch eine Mehrheit gleichzeitiger Sinusschwingungen zeigt, auch das geschulte Ohr noch Teiltöne wahrnehmen kann (allzu schwache natürlich ausgenommen), und daß überall, wo reine Sinusschwingungen vorliegen, auch dem Ohr jede weitere Zerlegung unmöglich ist. An dieser Grundtatsache kann keine akustische Theorie mehr rütteln, ohne selbst in Trümmer zu gehen.

## Phonographische, telephonische und ohrenärztliche Beobachtungen.

### I. Die Veränderungen der Vokale bei veränderter Umdrehungsgeschwindigkeit der Phonographenwalze.

Schon bald nach der Erfindung des Phonographen wurde der Apparat zur Prüfung der Vokaltheorien herangezogen. Sind die Vokale durch Formanten von fester absoluter Tonhöhe charakterisiert, so müssen die auf einer Walze aufgenommenen Vokale bei veränderter Umlaufgeschwindigkeit ihren Charakter verändern, mögen übrigens die Formanten harmonisch oder unharmonisch zum Grundton sein. Ist aber die relative Höhe bzw. die Ordnungszahl der charakteristischen Teiltöne entscheidend, so müssen die Vokale im wesentlichen unverändert bleiben. Also nicht zwischen HELMHOLTZ und HERMANN, wohl aber zwischen der Absolut- und der Relativtheorie kann so entschieden werden. In der Zeit der ersten, unvollkommenen Stanniolaufnahmen wurde noch gelegentlich behauptet, daß die Tatsachen im letzteren Sinne sprächen<sup>1)</sup>. Bei geringen Veränderungen der Geschwindigkeit könnte man auch heute noch so denken. Aber bei größeren liegt die Entstellung so klar vor, daß seit HERMANN'S letzter Abhandlung darüber (1911) entgegenstehende Behauptungen nur ganz sporadisch noch aufgetaucht sind (so 1916 bei FLOWERS, den STEFANINI sofort korrigierte). Daß sich sämtliche beobachtete Veränderungen aus einer Absolut-Theorie müssen ableiten oder voraussagen lassen, kann man allerdings nicht erwarten. Vielmehr mögen bei einer stark beschleunigten oder verlangsamten Drehung rein physikalische Faktoren ins Spiel treten, die die Alteration der aus dem Phonographen kommenden Schallwellen mitbedingen. Auch wird in Betracht kommen, daß der Phonograph überhaupt nur auf eine mittlere Zone von Tonhöhen merk-

<sup>1)</sup> So von JENKING und EWING, Nature Bd. 17, 1878; wogegen alsbald CROSS, das. Bd. 18, 1879.

lich reagiert. Voraussagen kann man nur: 1. daß im allgemeinen die Vokale bei verlangsamter Reproduktion in dunklere, bei beschleunigter in hellere übergehen werden; 2. daß die zweiteiligen, aus Formant und Unterformant zusammengesetzten Vokale, eine Verzerrung ihrer Struktur erfahren müssen, die überhaupt keinem der gebräuchlichen Vokale mehr entspricht. Denn wird I verlangsamt, so geht zwar der I-Formant in den E-Formanten über, zugleich sinkt aber der Unterformant mit, während er für E höher gehen müßte. Und wenn wir E beschleunigen, ist in beiden Beziehungen das Umgekehrte der Fall. Auf diesen Umstand hat bereits MILLER mit Recht aufmerksam gemacht. Wieweit und in welcher Art sich diese zu erwartenden Neubildungen in der Beobachtung geltend machen werden, läßt sich wieder nicht voraussagen. Überblicken wir nun die Beobachtungen selbst.

Schon mit dem Stanniophonographen bemerkten GRÜTZNER, LAHR und andere Forscher bedeutende Veränderungen, besonders bei verlangsamter Drehung. Ein eigentümliches Blöken fiel GRÜTZNER dabei auf. Am längsten hielt sich der A-Laut. LAHR'S Angaben stimmen vollkommen mit dem zu Erwartenden: bei Beschleunigung (LAHR gibt nicht an, um wieviel) gingen U in Ü und I, O in Ö und E über, A blieb A. Bei Verlangsamung blieb U U, O ging in Ou über, A in Ao, Ü in U.

Gründlicher hat dann HERMANN die Erscheinungen untersucht (Bd. 139, 1911). Er probierte bestimmte Geschwindigkeitsgrade bis zum doppelten bzw. halben gegenüber der Aufnahmegeschwindigkeit. Auch ihm fiel es auf, daß die Verlangsamung stärker entstellt als die Beschleunigung und daß dabei vielfach eigentümliche Blöklaute auftreten. Im übrigen sind die beobachteten Veränderungen wieder die zu erwartenden. Wenn E, Ü und I durch starke Beschleunigung oft in O und U übergingen, so ist klar, daß hierbei nur die Unterformanten gehört wurden, indem die Hauptformanten die Höhengrenze der Phonographen überschritten. Auch HERMANN bemerkt, daß A sich besonders lange hält. Er bringt dies wohl richtig damit in Zusammenhang, daß die Formantregion des A eine besonders große Ausdehnung hat ( $des^2 - b^2$ , wir würden sogar sagen:  $c^2 - g^3$ , wenn die äußersten Grenzen berücksichtigt werden, die aus dem Auf- und Abbauverfahren zu erschließen sind). Es gibt eben auch besonders viele noch als A geltende Modifikationen vom dunkelsten bis zum hellsten, und in diese wird der Vokal bei beschleunigter oder verlangsamter Reproduktion zunächst übergehen. Aus demselben Umstande versteht man, daß andere alterierte Vokale gern als A bezeichnet werden, indem eben ihre Verstärkungsmaxima in

diese breite Region übertreten. Auch daß eine geringe Verlangsamung im allgemeinen schädlicher ist als eine gleich große Beschleunigung, dürfte HERMANN richtig daraus hergeleitet haben, daß bei der gewöhnlichen Aufnahme in Phonographen der Formant sich mehr in der Nähe seiner unteren Grenze hält, so daß er bei der Verlangsamung leichter aus dem Bereich herausfällt, in welchem der Vokal noch richtig erkannt wird. Es wäre auch darauf hinzuweisen, daß der Aufnahmetrichter nicht resonanzfrei ist. Er dürfte meistens das O vor anderen Lauten verstärken. Im übrigen bleiben einige Resterscheinungen, die wir auch heute noch nicht zu deuten wissen; so der blökende oder nasale Charakter bei Verlangsamung. Doch kommen bei sehr langsamer Drehung zunächst hohe Töne, schließlich stark schnarrende Geräusche hinein. Zuletzt wird der Klang intermittierend.

Ich habe auch selbst 1914 und 1916 mit den Herren ABRAHAM, v. HORNPOSTEL, v. ALLESCH, KREICHGAUER und WERTHEIMER ausführliche Versuche über diese Umwandlungen angestellt. 5 Versuchsreihen mit zunehmender Berücksichtigung der möglichen Variationen wurden gemacht. Von der 2. an wurden außer den Vokalen auf  $c$  und  $c^1$  auch Instrumentalklänge zur Vergleichung herangezogen, und zwar das  $c$  einer gedackten Pfeife,  $c^1, c^2, c^3, c^4, c^5$  von Stimmgabeln,  $f^1, f^2$  von Zungen und eine Vokalröhre nach WILLIS. Dabei zeigte sich, daß der Phonograph Töne von  $c^1$  abwärts überhaupt nicht wiedergibt, ausgenommen, wenn die Reproduktion mit bedeutend größerer Geschwindigkeit erfolgt. In diesem Falle wird z. B. eine  $c^1$ -Gabel bei  $1\frac{1}{2}$ - bzw. 2facher Geschwindigkeit der Reproduktion als  $g^1$  bzw.  $c^2$  gehört. Außerdem hört man nur die den Ton begleitenden Geräusche oder Obertöne. Umgekehrt verschwindet eine  $c^5$ -Gabel in solchen Fällen, wo sie also  $g^5$  oder  $c^6$  geben müßte, weil Töne von der Mitte der 5-gestrichenen Oktave ab gleichfalls nicht wiedergegeben werden. Innerhalb dieser Grenzen werden bei Annäherung daran natürlich auch schon Abschwächungen eintreten, die auf die Vokale nicht ohne Einfluß bleiben können: die hellsten müssen bei gleicher Reproduktionsgeschwindigkeit etwas verdunkelt, die dunkelsten etwas erhellt erscheinen.

Man kann die Geschwindigkeit des von uns benützten Phonographen von einer unteren Grenze aus um das Dreifache steigern. Das Verhältnis zweier Geschwindigkeiten kann jederzeit auf die einfachste und sicherste Weise aus dem Intervall der Töne erschlossen werden, die ein und derselbe aufgenommene objektive Ton dabei ergibt. Da z. B. ein Ton, der bei geringster Wiederabgeschwindigkeit  $c^2$  ist, bei größter zu  $g^3$  wird, so ergibt sich

daraus die ebengenannte Verhältnisbestimmung 1:3. Und so haben wir jederzeit die gewünschten Verhältnisse der Wiedergabegeschwindigkeiten leicht herstellen können.

HERMANN verfuhr so, daß er bei der Aufnahme ein und dieselbe Walze in ihrem 1. Drittel mit der Tourenzahl 75, im 2. mit  $112\frac{1}{2}$ , im 3. mit 150 laufen ließ. Erfolgte nun die Wiedergabe der ganzen Walze einmal mit 75, einmal mit  $112\frac{1}{2}$ , einmal mit 150, so konnten mit derselben Walze 3 Grade der Beschleunigung und 3 der Verlangsamung erzielt werden. Dieses Verfahren ist scharfsinnig und sparsam, aber wohl nicht ganz einwandfrei. Denn da die Aufnahmen bei geringster und bei größter Geschwindigkeit schon nicht genau die gleichen Vokale liefern, so sind die Erscheinungen bei den verschiedenen Beschleunigungen und Verlangsamungen, obschon deren arithmetische Verhältnisse sich einfach berechnen lassen, doch wegen der verschiedenen Ausgangspunkte nicht recht kommensurabel, und man darf nicht von einer Beschleunigung um  $\frac{4}{3}$ ,  $\frac{3}{2}$ , 2 schlechtweg reden, sondern immer nur inbezug auf eine bestimmte Aufnahmegeschwindigkeit.

Ich halte auch das unwissentliche Verfahren, das HERMANN hier bevorzugte, weder für notwendig noch auch für unbedingt nützlich. Es lenkt die Aufmerksamkeit auf das Erkennen oder Erraten eines Vokals; man ist zufrieden, wenn man ihn benennen kann, während man vielleicht nicht so sehr auf die Feinheiten seiner Veränderung achtet wie beim wissentlichen Verfahren. Freilich hat HERMANN recht, wenn er auf die besondere Autosuggestionsgefahr bei Vokalen hinweist. Aber bei geschulten Beobachtern ist diese nicht zu fürchten.

Die Aufnahmen erfolgten bei größter, kleinster und mehreren mittleren Geschwindigkeiten. Die extremen Geschwindigkeiten erweisen sich als ungünstig für die Reproduktion; auch mit den gleichen Geschwindigkeiten reproduziert, liefern sie gewisse Entstellungen, weshalb man ja auch für phonographische Aufnahmen eine mittlere Geschwindigkeit zu benützen pflegt. Hier aber galt es, gerade auch die extremen Fälle zu untersuchen. Die Wiedergabe erfolgte dann jedesmal sowohl in der gleichen als in mehreren anderen Geschwindigkeiten. Die Ergebnisse können hier aus räumlichen Rücksichten nur ganz summarisch mitgeteilt werden.

1. Aufnahme mit größter Geschwindigkeit.

a) Wiedergabe mit gleicher Geschw.: Die Instrumente recht gut, doch die Gabeln  $c^2$ ,  $c^3$  etwas klarinettenartig, auch wohl leicht blökend. Vokale im ganzen gut, doch stark gesungenes A trompetenartig; U merklich erhellt.

b) mit mittlerer ( $\frac{3}{4}$ ) Geschw.: Instrumente noch leidlich treu. Vokale auf  $c^1$  (jetzt  $g$ ) ziemlich gut; bei denen auf  $c$  (jetzt  $G$ ) die äußeren gegen die mittleren verschoben, I nach E, die anderen fast alle nach O hin (A = AO, E und Ä = Ö, Ö = O, U = Uo).

c) mit sehr geringer Geschw.: Bei  $\frac{5}{9}$  starke Veränderungen, Gabeltöne mehr flötenartig, stärkere trompetenartig, Zunge blökend. Vokale auf beiden (jetzt in  $d$  und  $D$  verwandelten) Grundtönen in derselben Richtung



wie vorher bis zur Unkenntlichkeit verändert, zugleich aber rasselnd, schnarrend, die Umlaute nasalisiert, z. B.  $\ddot{U}$  = franz. en; Uhu wie Öhö, Kuckuck ganz seltsam = Tretre.

Bei  $\frac{4}{9}$  klingen Zungen wie Labialpfeifen, Vokale wie dunkles tonloses Rattern. Auf  $c^1$  (jetzt  $B$ ) noch etwas nach Ö oder O klingend, auf  $c$  (jetzt  $B_1$ ) ohne Vokalcharakter.

2. Aufnahme mit  $\frac{8}{9}$  der höchsten Geschwindigkeit (dies ist wohl ungefähr die bei phonographischen Aufnahmen gebräuchliche.)

a) Wiedergabe mit gleicher Geschw.: Gabeln etwas metallisch, sonst alles gut.

b) mit größter Geschw.: Gabeln im ganzen gut, Zungen scharf wie Oboen; Vokale alle etwas heller, doch in der Aufeinanderfolge noch kenntlich.

c) mit nahezu geringster Geschw. (Dezime unter der Aufnahme): Gabeln schwer beschreiblich, unkenntlich, Zungen fagottähnlich, blökend; Vokale nur tiefes Schnarren. (Erfolgt die Wiedergabe nur eine kleine Terz unter der Aufnahme, also mit  $\frac{5}{6}$  Geschw., so sind die Gabeln, auch die Vokale, noch leidlich erkennbar.)

3. Aufnahme mit  $\frac{2}{3}$  der höchsten Geschwindigkeit.

a) Wiedergabe mit gleicher Geschw.: Alles sehr gut. Weitaus die treueste Reproduktion.

b) mit größerer und größter Geschw.: bei  $\frac{6}{5}$  der Aufnahmegeschw. etwas nasalisiert, Vokale teilweise bedeutend erhellt, doch im allgemeinen kenntlich. Bei  $\frac{7}{5}$  in gleicher Richtung weiter geändert.

c) mit kleinerer und kleinster Geschw.: bei  $\frac{2}{3}$  Gabeln gut, Zungen wie Englischhorn, Vokale alle merklich verdunkelt, nur  $U = O$ . Bei  $\frac{1}{2}$  Geschw. Gabeln noch gut, Vokale unkenntlich und sehr schwach, mehr oder weniger starkes tiefes Schnarren.

4. Aufnahme mit  $\frac{1}{2}$  der höchsten Geschwindigkeit.

a) Wiedergabe mit gleicher Geschw.: Gabeln gut, Vokale sehr gut.

b) mit größter Geschw.: Gabel  $c^5$  (angeschlagen) nur trockener Schlag. Bei den Vokalen auf  $c$  im ganzen die dunklen erhellt, die hellen verdunkelt (z. B.  $U =$  franz. un,  $O = A$ ,  $\ddot{O}$  und  $\ddot{A} = A$ ,  $I = E$ ). Das nämliche noch mehr auf  $c^1$ .

5. Aufnahme mit geringster Geschwindigkeit.

a) Wiedergabe mit gleicher Geschw.: Gabeln gut, wenn sie nicht zu stark angegeben waren. Pfeife  $c^1$  wie Trompete. Vokale gut, nur  $U = Uo$ .

b) mit größter Geschw.: Gabel  $c^1$  (jetzt  $g^2$ ) wird hörbar, Gabel  $c^5$  dagegen nur trockenes Klopfen. Vokale wie bei 4 b), vielfach nasalisiert, quäkend.

Außer dieser Versuchsreihe wurde auch eine mit Vokalen auf dem Grundton  $A$  und geringster Aufnahmegeschw. gemacht. Wurde diese Geschw. bei der Wiedergabe allmählich bis zur größten erhöht, so blieben die Vokale bis  $d$  ( $= \frac{4}{3}$  Geschw. so gut wie unverändert. Dann begannen  $U$  heller,  $I$  und  $E$  dunkler zu werden; zuletzt (auf  $d^1 = \frac{8}{3}$  der Aufnahmegeschw.) war  $U = A\ddot{O}$ ,  $O = A\ddot{O}$ ,  $A =$  scharfes  $\ddot{A}$ ,  $E = O$ ,  $I = UO$  (Unterformanten, nachdem die Hauptformanten verschwunden). Also ähnlicher Befund wie bei HERMANN.

Für alle diese Beobachtungen gilt das oben Gesagte: zum größten Teil sind sie aus den Formantverschiebungen verständlich,

indem die ganze Tonmasse eines Klanges mit den relativen Intensitäten ihrer Teiltöne nach oben oder unten verschoben wird, die Maxima daher in verschiedene Formantgegenden kommen. Zu einem anderen Teil beruhen sie aber auf Intensitätsverschiebungen innerhalb des Klanges, indem sich die unteren oder oberen Teile den Grenzen der Wiedergabefähigkeit des Phonographen nähern oder sie überschreiten (vgl. z. B. das Verschwinden des  $c^5$  bei 4b, 5b). Zum 3. Teil endlich sind physikalische Störungen im Spiele, die sich bei sehr geringen oder sehr großen Geschwindigkeiten geltend machen. Diese sind noch nicht näher untersucht, aber ihr Vorhandensein läßt sich deutlich erschließen. Wenn z. B. bei Aufnahmen mit sehr hoher Geschwindigkeit die Klangfarben von Instrumenten schärfer werden, Gabeln trompetenartig klingen usw., auch wenn mit gleicher Geschwindigkeit reproduziert wird (1a, 2a), so ist klar, daß Obertöne hinzugekommen oder verstärkt sein müssen. Bei Reproduktion mit sehr geringer Geschwindigkeit nehmen diese Entstellungen noch zu (1c, 2d). Veränderungen in gleicher Richtung erfolgen aber auch bei mittleren Aufnahmegeschwindigkeiten, wenn mit übertriebener Geschwindigkeit reproduziert wird (3b).

Soweit nur Faktoren der 1. Klasse in Betracht kommen, sind die Veränderungen der Vokale auch in ihren Einzelheiten gut verständlich. So begreift sich leicht, daß bei 1b durch Tieferlegung der ganzen Klangmasse um eine Quarte I nach E, A nach AO, E und Ä nach Ö hin verändert werden: denn die Formantzentren der neuen Vokale liegen ungefähr eine Quarte unter denen der alten. Und so vielfach. Daß Ö zu O wird, beruht auf dem Wegfall des Hauptformanten, der durch die Verlangsamung zu schwach wird. Dagegen ist es wieder nur aus physikalischen Einflüssen zu verstehen, wenn bei Tieferlegung um  $\frac{5}{9}$ , d. h. um eine kleine Septime, aus Uhu Öhö oder gar aus Kuckuck Tretre wird (das r hängt jedenfalls mit den oben erwähnten Intermissionen zusammen).

Zusammenfassend können wir also sagen: Obgleich diese Tatsachen sich nicht durchgängig aus der Lehre von den festen Formanten ableiten lassen, fügen sie sich ihr doch zwanglos ein und tragen zu ihrer Bekräftigung bei, während sie mit einer Relativtheorie in unauflösllichem Widerspruche stehen.

Was die Konsonanten betrifft, so bleiben auch sie durch phonographische Aufnahmen nicht immer ganz unverändert und leiden ebenfalls durch veränderte Reproduktionsgeschwindigkeit (vgl. Kuckuck). Doch ist darauf bisher meines Wissens nicht näher geachtet worden.

## II. Erfahrungen und Versuche in der Telephonie.

Nachdem DU BOIS-REYMOND, H. FR. WEBER und HELMHOLTZ mathematisch gezeigt hatten, daß bei der telephonischen Übertragung nicht bloß Phasen-, sondern auch Amplitudenverschiebungen eintreten müssen, hat HERMANN 1891 (Bd. 48) diese Schlüsse weiter verfolgt und mit Beobachtungen verglichen. Die letzteren faßt er dahin zusammen, daß tatsächlich wesentliche Veränderungen der musikalischen Klangfarbe eintreten, die Sprache aber unverändert bleibe. Also, schloß er, ist das Amplitudenverhältnis der Teiltöne ohne jede Bedeutung für den Vokalcharakter. „Jede Theorie, welche . . . auf Verstärkungsgebiete in den Obertönen des Stimmklanges u. dgl. großen Wert legt, scheint mir mit den Erfahrungen am Telephon und Mikrophon unvereinbar.“

Genauer betrachtet, lassen aber seine eigenen Beobachtungen diese „Unzerstörbarkeit der Sprachlaute“, die auch JAENSCH noch 1913 behauptete, in etwas bedenklichem Licht erscheinen. „Am sichersten unter allen Vokalen erkennbar ist A. Die Vokale E, I, Ä, Ö, Ü werden sehr leicht miteinander verwechselt, dagegen fast niemals mit A, O oder U. O wird mit U leicht verwechselt, fast nie einer dieser Vokale mit A“ (S. 558). HERMANN scheint für diese Verwechslungen nur die geringe absolute Intensität haftbar zu machen („A verschwindet als offenbar lautester Vokal zuletzt“). Aber zunächst wird man doch gerade an die Verschiebung der Amplitudenverhältnisse denken. Jedenfalls sprechen seine eigenen Beobachtungen eher gegen als für seine Schlußfolgerung.

Auch MAX WIEN kam 1905, in Anbetracht der Intensitätsverschiebungen von Teiltönen (Begünstigung der höheren Töne durch die Telephonmembran, auch bei größerer Entfernung vom Ohre oder bei elektrischer Beeinflussung der Telephonströme), die gleichwohl die Verständlichkeit der Sprache nicht aufheben, zu dem Schluß, daß die Teiltontheorie der Vokale durchaus unzulänglich sei. Andere noch nicht hinreichend erforschte Dinge, für die das Ohr eine erstaunliche Empfindlichkeit besitze, müßten wesentlicher sein als die Teiltonintensitäten. Aber hier dürfte die außerordentliche Interpretationskunst des erwachsenen Menschen gegenüber sinnvoll zusammenhängender Rede seiner Muttersprache übersehen sein. Zu Versuchen dieser Art eignen sich daher nur einzelne Sprachlaute oder sinnlose Silben. Aber selbst bei diesen kommt in Betracht, daß ein Sprachlaut nicht ein völlig scharf definiertes Phänomen, sondern einen Typus darstellt, innerhalb dessen gewisse, oft recht beträchtliche Verschiebungen

möglich sind, ehe er in einen benachbarten Typus übergeht (man denke an A oder das so wichtige S). In ähnlicher Weise wird selbst ein musikalisches Intervall, obgleich der Physiker dafür seine mathematische Definition hat, in Wirklichkeit selbst bei erheblicher Verstimmung noch als Quarte, große Terz usw., oder eine schlecht gezeichnete Ellipse innerhalb gewisser Grenzen noch als Ellipse anerkannt. Mit solchen „Verzerrungen“ hat auch die allgemeine Theorie der Vokale zu rechnen.

Untersucht man die Verständlichkeit einzelner Sprachlaute in unserem gegenwärtigen Stadttelephon, so findet sich S nahezu gleich dem F, dieses selbst merklich abgestumpft, auch Ch pal. sehr viel stumpfer, dem Sch ähnlich, geflüstertes E und I nicht unterscheidbar und sehr schwach, Ü etwas blasend, gesprochenes I und Ü außerordentlich schwach, auch E bedeutend schwächer als die dunkleren Vokale. Nach MILLER wird I telephonisch häufig als U gehört, d. h. man hört nur den Unterformanten. Sonst wird alles gut wiedergegeben. Der Fall liegt so, wie wenn die obere Hörgrenze auf etwa  $e^4 = 2600$  Schw. herabgesetzt wäre (vgl. unsere Tabelle S. 94). Aber dies bildet kein Hindernis für das Verstehen der sinnvollen Sprache. Die Amplitudenverschiebungen, die theoretisch abgeleitet wurden, halten sich also praktisch in solchen Grenzen, daß sie innerhalb des für die Sprache entscheidenden Gebietes ohne Wirkung bleiben.

Man kann die Grenzen der im Telephon bzw. Mikrophon vorkommenden Schwingungen auch leicht durch Telephonieren von Stimmgabeltönen nachprüfen. Die Gabeln der EDELMANNschen kontinuierlichen Tonreihe sind bis etwas unter  $c$  (128 Schw.) hörbar, freilich zuletzt sehr schwach<sup>1)</sup>. In der 2- und 3-gestrichenen Oktave sind sie besonders gut, nur in der Klangfarbe gegen die Klarinette hin verändert; es müssen also Obertöne (vermutlich besonders im Mikrophon) hinzukommen. Weiter nach oben ist  $g^4$  noch gut,  $c^5$  (4096) aber schon sehr geschwächt.

Untersucht man obertonreiche Klänge wie die der menschlichen Stimme oder der Streichinstrumente, so kann man allerdings nach der Tiefe zu auch noch Klänge der großen Oktave durchaus gut telephonieren. So kam das von einem Bassisten gesungene  $E = 80$  Schw. noch kräftig heraus; aber die Stimme wurde nach der Tiefe hin immer metallischer, instrumentenähnlicher, offenbar infolge der Intensitätsverschiebung zugunsten der Obertöne. Das  $C$  des Cello kam gleichfalls heraus, sogar noch tiefere, durch

<sup>1)</sup> Hr. KREICHGAUER teilt mir mit, daß er auch oszillographisch die Reaktion eines guten Mikrophons auf die Gabel 100 feststellen konnte.

Herabstimmung der Saite erzeugte Töne, aber mit deutlich stärkerer Oktave und schwachem Grundton. Es ist daher anzunehmen, daß bei so tiefen Klängen der Grundton im Telephon wesentlich nur als Differenzton der Obertöne hinzukommt.

Natürlich hat auch die Telephontechnik sich längst mit der Bestimmung der für das Sprachverständnis unentbehrlichen Frequenzen befaßt. Bereits 1901 fand M. WIEN die größte Empfindlichkeit des Telephons zwischen 500 und 3000 Schw. Versuche im Kaiserlichen Telegraphenversuchsammt zu Berlin 1906 ergaben als wichtigste mittlere Frequenz der telephonischen Wechselströme 800 Schw. und als für die Sprache unentbehrliche Gegend 600—1200. DEVAUX-CHARBONNEL bezeichnet 1000 Schw. als entscheidendste Gegend und 800—1200 als unentbehrliche Frequenzen. Die Grenzbestimmungen wechselten inzwischen einigermaßen, wobei aber auch die Anforderungen an Deutlichkeit verschieden sein mochten. Neuerdings bestimmte der Telegraphendirektor ULFILAS MEYER die Region 500 ( $h^1$ ) bis 2100 ( $c^4$ ) als „notwendig und hinreichend“, vermutete aber, daß jenseits  $c^4$  noch gewisse Frequenzen in der Sprache vorkämen, die für ihre Deutlichkeit nicht mehr wesentlich seien. Als untere Grenze der im Telephon überhaupt vorkommenden Frequenzen bezeichnet er 360 ( $cis^1$ )<sup>1</sup>. Nach BROEMSER (1919) gibt das Telephon alle Frequenzen unterhalb 1200 richtig wieder, schwächer die höheren, das Mikrophon alle unter 800 verstärkt, die über 800 schwächer. Er hält es für genügend, wenn Schwingungen bis 2000 unverzerrt wiedergegeben werden. Unterdrückung der Schwingungen unter 400 veränderte die Klangfarbe der Sprache merklich, beeinträchtigte aber nicht wesentlich ihre Verständlichkeit.

Diese Bestimmungen stehen unter sich und mit unseren Ermittlungen in befriedigendem Einklang<sup>2</sup>). Die zeitweilig angenommene Region 600—1200 war allerdings zu eng gegriffen, doch ist auch daran so viel richtig, daß von 1200 abwärts die empfindlichste Gegend beginnt, bei deren Ausschaltung das Sprachverständnis völlig vernichtet wird. Und wenn sich immer wieder 800 als die überhaupt wirksamste Frequenz erwiesen hat, so ist dies ohne weiteres daraus verständlich, daß hier das Formantzentrum des

<sup>1</sup>) Die in den telephontechnischen Schriften angegebenen Werte  $\omega$  (Kreisfrequenzen) ergeben, durch  $2\pi$  dividiert, die obigen Schwingungszahlen. Für „Schwingungen“ pflegt man dort den Ausdruck „Hertz“ zu gebrauchen.

<sup>2</sup>) Nach Mitteilungen K. W. WAGNERS (2, 3, vgl. „Das Fernsprechen im Weltverkehr“. Aus d. Reichspostministerium Nov. 1923) haben sich die seit 1915 veröffentlichten Arbeiten des Verfassers für das Fernsprechen bereits in erfreulichem Maße nützlich erwiesen.

A liegt, das man als den wichtigsten aller Vokale zu betrachten hat (vgl. S. 97, auch 81, 229, 241, sowie Kap. 10 u. 13). Immerhin ist dies nur theoretisch zu verstehen, und sicher wäre es verkehrt, beim Bau von Mikrofonen auf überwiegende Verstärkung dieser (ohnehin schon durch die beim A erzielbare größere Energie der Lautgebung bevorzugte) Gegend Gewicht zu legen und dagegen die obere und untere Grenzregion der Sprachlaute zu vernachlässigen. Eine Sprache bloß mit U, O, E, I als Vokalen wäre immer noch weit leistungsfähiger als eine bloß mit A-Vokalen. Sie wäre zwar ebenso unschön, aber gerade durch ihre Kontraste zu vielerlei verschiedenen ausdrucksvollen Kombinationen befähigt.

Man kann, wie K. W. WAGNER (1) gezeigt hat, durch Einschaltung von Kondensatoren und Drosselspulen („Siebketten“) in die Telephonleitung auch Ab- und Aufbau- sowie Lückenversuche machen, wie bei unserer Interferenzeinrichtung. Die Drosselspulen dämpfen die Schwingungen proportional, die Kondensatoren umgekehrt proportional der Frequenz. Durch diese kann also die untere, durch jene die obere Region des Klanges abgeschnitten werden. Hierbei ist es ein Vorteil gegenüber der Interferenzmethode, daß man nicht Multipla zugleich mit ausschließt. Doch schienen mir die Abbauprodukte bei Vorführungen zunächst weniger fein differenzierbar und überzeugend als die des Interferenzverfahrens.

In Amerika, wo gleichzeitig ähnliche Einrichtungen aufkamen (CAMPBELL), wurden sie in den letzten Jahren von H. FLETCHER zu systematischen Abbauversuchen benützt, bei denen die Vokale und Konsonanten der englischen Sprache, sowohl für sich allein als innerhalb ganzer Silben, auf ihre Verständlichkeit geprüft wurden, wenn die Frequenzen von der unteren oder oberen Grenze her bis zu beliebigen mittleren Punkten durch elektrische Filter aus Kondensatoren und Spulen ausgeschlossen wurden. Jedesmal wurde die Prozentzahl der Fälle des richtigen Verständnisses festgestellt. Aus solchen Reihen wurden Deutlichkeitskurven abgeleitet, die die Abnahme des Sprachverständnisses mit Vergrößerung des weggenommenen Tonbezirkes veranschaulichen, eine für den Abbau von unten und eine für den von oben her. Auch die Verluste an der physikalischen Gesamtenergie wurden jeweilig ermittelt und besondere Kurven dafür abgeleitet.

Dabei trat u. a. die relativ geringe Bedeutung der tieferen Teiltöne für die Verständlichkeit hervor (woraus man aber nicht schließen darf, daß die Vokale unverändert blieben). Wurden alle Töne unter 500 ausgeschlossen, so betrug die Verständlichkeit immer noch 98%, obgleich dabei die physikalische Gesamtenergie

auf 40% reduziert war. Die Schwingungszahl 1550 ( $g^3$ ) war dadurch ausgezeichnet, daß die Verständlichkeit um den gleichen Betrag, nämlich bis auf 65%, reduziert wurde, wenn die darunter und wenn die darüber liegenden Töne abgeschnitten wurden.

Unter den einzelnen Lauten hebt FLETCHER 3 Typen hervor: der 1., aus langen hellen Vokalen bestehend, darunter I wie in team, liefert immer noch mehr als 97% richtige Fälle, wenn alle Frequenzen unter oder über der Gegend 1000—1500 abgeschnitten werden. Der 2., wozu die kurzen Vokale U, O und E (wie in ten) gehören, hat wichtige Bestandteile unterhalb 1000. Wurden alle Schwingungen unter dieser Grenze ausgeschlossen, so ergaben sich mehr als 20% Irrtümer. Der 3. Typus umfaßt die Reibungskonsonanten S, Z, engl. Th. Sie werden stark mitgenommen, wenn die Schwingungen über 5000 ausgeschlossen werden, nicht aber durch Wegnahme der unter 1500 liegenden Zone. Auf ihre Rechnung kommt es hauptsächlich, wenn ganze Silben durch Ausschluß der über 2500 liegenden Schwingungen geschädigt werden.

Allgemein schließt FLETCHER, daß Schwingungen von 100 bis über 5000 tadellos übertragen werden müssen, wenn die Sprache gut wiedergegeben werden soll, und daß zwar die unter 1000 den größten Teil der Energie beitragen, das Charakteristische aber zumeist durch die über dieser Grenze liegenden Töne gegeben werde. Die Konsonanten seien im allgemeinen schwerer verständlich als die Vokale (mit wenigen Ausnahmen beiderseits), am schwersten aber engl. Th, F und V, welche die Hälfte aller Mißverständnisse verschuldeten.

Die Ergebnisse dieser verdienstvollen Untersuchungen schließen sich gut an die unserigen an. Nur im einzelnen bleibt einiges paradox. So ist unbegreiflich, daß I bei Ausschluß aller Töne über  $ges^3$  oder gar über  $c^3$  noch verständlich sein soll.

Eine merkwürdige Eigentümlichkeit zeigt auch das Verhalten der S-Kurve. Wenn die über 5000 ( $es^5$ ) liegenden Töne ausgeschlossen wurden, war S noch gut verständlich (88%). Wurde die ausgeschlossene Region nach unten erweitert, so war es schwerer zu verstehen und erreichte bei 3000 ( $ges^4$ ) 50%, war also unverständlich. Soweit ist alles nur zu erwarten, denn  $ges^4$  ist unsere Formantmitte für S. Wurde nun aber die ausgeschlossene Region noch weiter bis zu 1800 ( $b^3$ ) ausgedehnt, so hob sich die Verständlichkeit wieder auf 60%, um dann bei weiterer Vergrößerung bis 1200 ( $es^3$ ) wieder auf 50% zurückzusinken. Hier ist die vorübergehende Hebung auffallend. Man könnte daran denken, daß die im S enthaltene Tonhöhe  $a^3$ , die mit zu seinem Charakter beiträgt, durch den Ausschluß der dicht darüberliegenden Geräuschregion deutlicher würde. Aber diese Tonhöhe ist ja keineswegs für S allein charakteristisch. Und so mag diese geringfügige Hebung zufälligen Ursachen entspringen sein.

### III. Ohrenärztliche Erfahrungen.

S. 139ff. besprachen wir die Verschiebung der oberen Hörgrenze mit dem Alter. Bei normalen Ohren bleiben diese bis zum 70. Lebensjahr ohne Schaden für das Sprachverständnis. Die winzige Abstufung des S, die man durch Interferenzversuche mit Abschneidung der Töne oberhalb  $c^6$  nachbilden kann, macht sich im Leben nicht geltend.

Die Grenze des Normalen gegen das Pathologische ist nun aber auch hier eine fließende, und pathologische Beobachtungen selbst bieten ausgiebige Gelegenheiten zur Vergleichung mit den experimentellen Ergebnissen. Bei Labyrinthkrankungen z. B. schreitet meistens die Taubheit von der normalen oberen Hörgrenze nach unten hin fort (was man nach HELMHOLTZ' Hörtheorie so deuten kann, daß die Fasern an der Schneckenbasis, die der Perzeption der höchsten Töne dienen, schädigenden Einflüssen am meisten ausgesetzt sind). Auch bei pathologischen Veränderungen der Gehörknöchelchen scheinen die funktionellen Hemmungen vielfach zuerst die obere Tongegend zu treffen. Es kommt aber auch vor, daß die tiefe Region zuerst geschädigt wird, oder daß Tonlücken in der mittleren Gegend auftreten. Zur Vergleichung des Sprachverständnisses bei den von oben nach unten fortschreitenden Veränderungen bieten sich unsere Tabellen S. 94 und 133, für feinere Differenzen auch die Abbautabellen der stimmhaften und stimmlosen Laute, für die Fälle von Tonlücken die Formantentafeln und das, was über die Erscheinungen bei Lücken- und Stichversuchen oben S. 71ff., 108ff. berichtet wurde.

Nehmen wir nun beispielsweise an, bei einem Patienten mit abwärts fortschreitender Ertaubung sei die obere Hörgrenze für Töne bei  $gis^4$  gefunden worden, so wird er schon mit dem Verständnis des geflüsterten „sieben“ Schwierigkeiten haben. Ist die Hörgrenze für Töne bis  $es^2$  gesunken, so muß nach der Tabelle S. 133 auch „acht“ ihm unverständlich sein. Den umgekehrten Schluß darf man aber natürlich nur dann ziehen, wenn die Voraussetzung einer von oben nach unten fortschreitenden Ertaubung schon nach anderen Kennzeichen gerechtfertigt erscheint. Daß man, wie v. BEZOLD meinte, aus dem Nichtverständnis bestimmter Flüsterlaute sogar auf den Sitz der Krankheit schließen könnte (so sollte z. B. ein nichtverstandenes „5“ charakteristisch sein für akute Mittelohrentzündung, ein „7“, aber auch „6, 2, 20“ für Labyrinthkrankung, „9“ für Tubenverschluß), dürfte sich nicht halten lassen. Aber dafür hat der Ohrenarzt ja andere Kriterien.

Exakte Angaben aus der ohrenärztlichen Literatur, die sich zur Vergleichung eignen, sind leider bisher nur in geringer Zahl



vorhanden<sup>1)</sup>. Es fehlte früher an einwandfreien Apparaten zur Bestimmung der oberen Hörgrenze; auch hat sich erst infolge der neueren phonetischen Untersuchungen das Interesse wissenschaftlich eingestellter Ohrenärzte solchen exakteren Bestimmungen zugewandt. v. BEZOLD war einer der Pioniere, aber seine „Sprachsextete“ hat sich als verfrühte und unhaltbare Verallgemeinerung erwiesen (oben S. 95ff.). Aus neuerer Zeit liegt eine Angabe des holländischen Ohrenarztes und Phonetikers STRUYCKEN vor, dessen Monochord für die Bestimmung der oberen Hörgrenze sich neben der Galtonpfeife bei Ohrenärzten eingebürgert hat; sowie besonders eine Untersuchung von O. CLAUS an 26 Patienten, die durch unsere Tabellen (Abh. 12) veranlaßt ist, und die italienischen Arbeiten von GRADENIGO und PAPALE, von denen das nämliche gilt. Die Übereinstimmung kann als durchaus befriedigend bezeichnet werden. Insbesondere entsprach die Reihenfolge der Schädigungen des Sprachverständnisses beim Herabrücken der Hörgrenze der Reihenfolge in unseren Tabellen. Aber auch die absolute Lage der herabgesetzten oberen Hörgrenze, die bei bestimmten Defekten zu erwarten war, fand sich bestätigt. PAPALE, dessen If.-Studien über Flüstervokale o. S. 207 erwähnt sind, konstruierte daraufhin neue Reihen italienischer Flüsterwörter zu Prüfungszwecken (2). Dasselbe ist für das Deutsche und Englische auf Grund unserer Ergebnisse durch H. LAMPERT geschehen.

STRUYCKEN berichtet über Fälle von Vernichtung der Gehörknöchelchenleitung: „Bei Einschränkung bis auf 7000 Schw. ( $a^5$ ) wird schon die Unterscheidung zwischen manchen Konsonanten schwierig. Sinkt die obere Hörgrenze weiter bis auf 5000 Schw. ( $es^5$ ), dann wird auch das tonlose Ss und Ff nicht mehr erkannt (wie beim Telephon, wo auch diese Schwingungen nicht übertragen werden). Auch die Unterscheidung zwischen I und E fällt dann schwer.“

CLAUS, der ausschließlich Fälle von Labyrinthkrankung zur Untersuchung heranzog, hebt 2 Fälle als Beispiele hervor: „Bei einem Patienten war die obere Tongrenze bis auf  $g^5$  eingeschränkt; dieser hörte alle geflüsterten Vokale sehr gut bis auf das O, das er angeblich nur sehr undeutlich und verschwommen vernahm“ (hier könnte etwa neben der Herabsetzung der oberen Hörgrenze auch noch eine Hörlücke zwischen  $g^1$  und  $c^2$  bestanden haben); „die Konsonanten wurden sämtlich gut perzipiert, mit Ausnahme des Ch palatale. Ein anderer Patient hörte nur noch die Töne von  $d^5$  abwärts; er perzipierte I, Ch pal. und S nicht, alle übrigen Vokale und Konsonanten wurden mühelos verstanden.“

<sup>1)</sup> Was bis 1883 über partielle Gehörausfälle zu finden war, habe ich I, S. 402ff. zusammengestellt. Daraus läßt sich für unseren Zweck trotz bemerkenswerter Einzelheiten im ganzen doch nur entnehmen, daß die „hohe Region“ für das Sprachverständnis besonders wichtig ist. Faßt man diesen Begriff im musikalischen Sinne, wie er von den Autoren damals wohl in der Regel verstanden wurde, so wäre es etwa  $a^2-c^5$ , in Übereinstimmung mit den phonetischen Ergebnissen.

Zunächst schien allerdings sowohl in den STRUYCKENSchen als den CLAUSschen Fällen die obere Hörgrenze regelmäßig um einige Töne höher zu liegen, als nach unserer Tabelle zu erwarten war; aber es standen damals in dieser Tabelle unpraktischerweise nur die unter Berücksichtigung der If.-Breite reduzierten oberen Hörgrenzen für Flüstergeräusche. Achtet man auf die den If.-Einstellungen direkt entsprechenden oberen Grenzen für Töne, die in die gegenwärtige Form der Tabelle mit aufgenommen sind, so verschwinden diese Abweichungen. Ein Patient, der die über einer bestimmten Grenze liegenden Teile der Flüstergeräusche nicht mehr vernimmt, mag doch immerhin auf so starke Einwirkungen, wie es die durchdringenden Longitudinaltöne des STRUYCKENSchen Monochords und die Pfeife der Galtonpfeife zwischen 5000 und 7000 Schw. sind, noch um einige Tonstufen höher hinauf reagieren. Die letztere muß bekanntlich überhaupt stark, rasch und kräftig angeblasen werden, wenn nicht irreführende tiefere „Schneidentöne“ entstehen sollen.

Das Herabrücken der oberen Tongrenze geht nach CLAUS (S. 298) nicht unbedingt und einfach parallel mit dem Verschwinden der Perzeptionsfähigkeit für Flüsterlaute. Dies dürfte, soweit nicht Tonlücken in Betracht kommen, teils mit einer allgemeinen Gehörsschwächung, teils damit zusammenhängen, daß beim Verständnis der Flüstersprache noch so manche psychische Bedingungen und Unterschiede mitspielen, die auf die Tonwahrnehmung nicht oder weniger von Einfluß sind. Doch wurde von den CLAUSschen Patienten die Flüstersprache allgemein nicht mehr verstanden, sobald die Tongrenze unter  $c^4$  herabgesunken war; was mit unserer Tabelle S. 133 vollkommen stimmt.

Über die Prüfung mit stimmhaften Sprachlauten berichtet CLAUS: „War die obere Tongrenze bis auf  $c^4$  eingeschränkt, so wurden alle Vokale mit Ausnahme des I gehört; sank sie jedoch bis auf  $g^3$ , so wurden I und E nicht mehr vernommen [verstanden]. Ging sie noch weiter herunter, etwa bis  $f^2$ , dann war gewöhnliche Konversationssprache nicht mehr verständlich.“ Dies ist nur zu erwarten. Abweichungen dürften hier hauptsächlich dadurch entstehen, daß nicht bloß Hörverlust für die hohe Region, sondern zugleich eine allgemeine Gehörsschwächung vorzuliegen pflegt, wenn zur Prüfung mit stimmhafter Sprache geschritten wird. Darin liegt ein wesentlicher Unterschied gegenüber den Bedingungen unserer Versuche, wo unterhalb der jeweilig ausgeschlossenen höheren Klangteile alles unversehrt blieb oder wenigstens die Schwächung nur minimal war.

Bemerkenswert ist noch, daß sowohl bei der geflüsterten wie der stimmhaften Sprache in den CLAUSschen Fällen das A auch bei weit herabgerückter Tongrenze, wenn selbst O und U nicht mehr verstanden wurden, noch erkennbar blieb. Wieder ein Zeichen jener merkwürdigen Widerstandsfähigkeit dieses Lautes, von der wir schon mehrfach hörten. Hier dürfte besonders die größere Schallkraft<sup>1)</sup> den Ausschlag gegeben haben. Aber es müssen

<sup>1)</sup> A wird nach CLAUS unter den Flüstervokalen am weitesten gehört. Bei meinen If.-Versuchen bemerkte ich vielfach, daß starkes Flüster-A und Sch durch die Wand gehört wurden, weshalb hier leisere Schallgebung oder Verlängerung der Leitung bis ins übernächste Zimmer erforderlich war. Dies entspricht auch den bekannten Beobachtungen von O. WOLF (Sprache und Ohr, 1871) über die Hörweite der Sprachlaute, wonach unter den Vokalen A, unter den Konsonanten Sch am weitesten gehört wird. Damit stimmen auch Versuche von H. RÜEDERER (Üb. d. Wahrnehmung des gespr. Wortes, Diss. München 1916, S. 20) überein.

die unter  $c^3$  liegenden Töne und Geräuschteile hier doch noch einigermaßen erhalten geblieben sein und muß nur eine Schwächung stattgefunden haben, die die an sich schwächeren O und U vernichtete, während A vernehmbar blieb. Bei den If.-Versuchen wäre es ganz unmöglich, daß A noch erhalten bliebe, wenn O und U schon (von oben her) vernichtet sind<sup>1)</sup>.

In einem von CLAUS als besonders instruktiv bezeichneten Fall eines sehr intelligenten Patienten konnte dieser rechts die Monochordtöne um  $g^5$  noch hören und vernahm alle geflüsterten Vokale und Konsonanten gut mit Ausnahme von I, Ch pal. und F. Hier folgt aus unserer Tabelle, daß die Schädigung mindestens bis  $b^4$  hinabreichen mußte. Aber die relativ starken Töne des Monochords hörte er noch bis zu einer Sext höher hinauf. Auf dem linken Ohre war die obere Tongrenze  $g^3$ . Hier vernahm er überhaupt keine Flüsterlaute mehr; dagegen verstand er die in Konversationssprache vorgesagten Vokale gut bis auf I und E. Diese verstand er entweder gar nicht oder hörte I als U, E als O. Hier ist die Übereinstimmung mit der Tabelle S. 94 ohne weiteres klar. (Daß von Schwerhörigen I oft mit U verwechselt wird, ist überhaupt eine alte Erfahrung, die uns nun aber erst durch die Tatsache der Unterformanten verständlich wird.)

Mit einem der CLAUSSchen Patienten machten wir folgenden lehrreichen Versuch an der synthetischen Einrichtung. Der Patient hörte hier aus dem Trichter T (o. S. 44), wo die Töne stärker herauskommen als an dem gewöhnlich benutzten Schlauchende, die einzeln angegebenen einfachen Töne bis zu  $h^4$  als oberer Grenze. Wurde nun aber ein E aus den erforderlichen Teiltönen zusammengesetzt, in welchem also Töne der 4-gestrichenen Oktave als Formanttöne enthalten waren, so hörte er O statt E. Die Formanttöne waren also bei diesem Patienten zwar isoliert vorhanden, wurden aber im E-Klange durch die tieferen Teile des Vokals unterdrückt, und so kam nur dessen Unterformant O zur Geltung. Wenn ich den Fall für meine Ohren nachbilde, indem ich die hohen Töne so schwäche, daß sie zwar für sich allein noch zu hören, aber aus dem Gesamtklang auch bei angestrenzter Aufmerksamkeit nicht mehr herauszuhören sind, dann geht dieser auch für mich in O über — ein interessanter Fall der Unterdrückung höherer Töne durch tiefere.

Bezüglich der höchstliegenden Konsonanten fand CLAUS einen Unterschied in der Hörschärfe erst, wenn die obere Tongrenze bis auf  $c^6$  heruntergerückt war. Dies entspricht in der Tat der höchsten Grenze, bis zu der die Konsonanten hinaufreichen; aber die Schädigung kann hier nur ganz minimal sein<sup>2)</sup>.

<sup>1)</sup> P. v. LIEBERMANN berichtet (Biol. Zentralbl. 1912, S. 746) von einem Arzt, der die Töne oberhalb  $c^2$  nicht mehr hören konnte, aber ein auf  $c$  gesungenes A ausnahmslos richtig erkannte. Richtig kann er es aber nicht gehört haben, sondern nur als ein OA, das ihm das helle A ersetzen mochte, ebenso wie der partiell Farbenblinde im gewöhnlichen Leben mit seinen veränderten Farben zurechtkommt, oft ohne auch nur davon zu wissen. REVEZ meint, es sei von dem vernichteten  $c^3$  dessen abstrakte „A-Vokalität“ doch übriggeblieben. Aber der Mann hörte ja nach dem Bericht die  $c^3$ -Gabel überhaupt nicht, muß also auch für ihre Vokalität taub gewesen sein.

<sup>2)</sup> KOEHLER erwähnt (I, Bd. 64, S. 94), daß in sämtlichen Fällen von Labyrinthkrankungen und Neuritis acustica, die in der Frankfurter Ohrenklinik in dieser Richtung untersucht worden waren, entweder völliger Ausfall des S, F, Ch oder Schwerhörigkeit für diese Laute konstatiert wurde,

GRADENIGO unterschied (1923) 3 Kategorien typischer partieller Hörverluste in den hohen Regionen:

1. Vollständiger Verlust von  $g^2$  aufwärts. Hier sind alle Vokale außer U und die Konsonanten außer R unverständlich; „sette“ und „otto“ ununterscheidbar.

2. Vollständiger Verlust von  $g^5$  aufwärts. Alle Vokale klingen aspiriert, E dem O, I dem U genähert. S ist abgestumpft. Bei Zerstörung schon von  $e^4$  aufwärts sind E und I stimmhaft noch unterscheidbar, geflüstert aber nicht.

3. Nur die 2 Oktaven  $g^3$ — $g^5$  verloren. Auch hier „wunderbare Übereinstimmung“ der klinischen Erfahrung mit If.-Versuchen.

Als Ursache der Erkrankung vermutet er in allen diesen Fällen sklerotische Prozesse im Labyrinth. In einer besonderen Klasse von Fällen, über die er früher berichtete (Su particolarità funzionale di una forma di nevro-labirintite de sifilide congenita, Giorn. d. R. Accad. di Medic. di Torino 1914) ist alles bis  $d^3$  oder  $dis^3$  hinauf unversehrt (nur ein wenig geschwächt), alles Darüberliegende zerstört, nur sehr starke Töne sind noch eine Oktave über  $dis^3$  hinaus hörbar. Das Sprachverständnis ist geschädigt, besonders für S; Näheres leider nicht angegeben.

Man darf im allgemeinen nicht erwarten, daß ohrenärztliche Befunde so einfache und durchsichtige Gesetzmäßigkeiten aufweisen wie die der experimentellen Phonetik, wo wir die Bedingungen selbst setzen und daher relativ gut überschauen können, obgleich auch da oft genug maßgebende Faktoren übersehen werden. Der Krankheitsfall kann Komplikationen darbieten, wie solche im vorangehenden erwähnt sind. Außerdem aber kommen bei den Prüfungen durch das Sprachverständnis die Intelligenz und Beobachtungsfähigkeit des Patienten, die passende Wahl der nachzusprechenden Silben, die Modalitäten der Aussprache, der Stärke und Entfernung, des Prüfungsraumes usw. in Betracht: Schwierigkeiten, die den Ohrenärzten auch sehr wohl bekannt und von CLAUDIUS LAMPERT u. a. nachdrücklich hervorgehoben sind.

Sprachliche Prüfungen haben für den Ohrenarzt den großen Vorteil, daß die richtige Wiederholung von seiten des Patienten ohne weiteres den Beweis des richtigen Hörens liefert, soweit nicht etwa bloßes Erraten oder Erschließen mitwirkt; und die Flüstersprache hat gegenüber der stimmhaften als Prüfungsmittel den dreifachen Vorzug, daß sie eine größere Konstanz der Lautstärke besitzt, daß sie leicht in kleinen Räumen und bei geringer Herabsetzung der Hörschärfe zu verwenden ist, und besonders, daß sich jedes Ohr für sich allein damit prüfen läßt. Dennoch wird man bei gründlicherer Prüfung auch die stimmhafte Sprache

wenn die obere Hörgrenze „schon unter  $c^7$  herabgesetzt war“. Hat man dies dahin zu verstehen, daß sie nur wenig unter  $c^7$  herabgesetzt war, so würde hierin allerdings ein starker Widerspruch gegen unsere Versuchsergebnisse und Tabellen liegen; aber ich würde dann die Zuverlässigkeit der ohrenärztlichen Beobachtungen ganz entschieden bestreiten. Vgl. das im Text Folgende.

heranziehen und selbstverständlich die Tongrenzen und Tonlücken feststellen. Für die obere Tongrenze scheinen die jetzt gebräuchlichen Hilfsmittel genügend, für die untere, sowie für Tonlücken in der mittleren Gegend müßten aber durchaus einfache Töne benutzt werden, also entweder die Pfeiftöne des Mundes oder Klänge, die durch If.-Vorrichtungen von ihren Obertönen befreit sind. Auch bei den Gabeln der BEZOLD-EDELMANNschen „kontinuierlichen Tonreihe“ wäre dies erforderlich, um nicht in Fehlschlüsse zu verfallen, da sie wenigstens die Oktave merklich enthalten.

Es wäre überhaupt sehr zu empfehlen, daß in allen Kliniken, die über größere Räume verfügen, If.-Einrichtungen geschaffen würden, sowohl um einfache Töne herzustellen als um Hördefekte und besonders ihre fortschreitende Ausdehnung von oben nach unten nachzubilden<sup>1)</sup>. Dabei hat auch die If.-Breite sicherlich ihr Analogon in dem jeweiligen Zustande des Organs: wie dort unter der Grenze völliger Vernichtung noch eine kleine Zone der Schwächung liegt, so erstreckt sich ohne Zweifel auch im erkrankten Gehörorgan über die Grenze der funktionell vernichteten Region hinaus noch eine Zone abnehmender Schädigung.

In mancher Beziehung würden allerdings die elektrophonischen Methoden die genannten Zwecke für den Ohrenarzt noch besser erfüllen. In der Neuen Welt haben sich mehrfach Physiker mit Ohrenärzten verbunden, um von deren Patienten mit „Audion-Oszillatoren“ „Audiogramme“ aufzunehmen, d. h. ihre Schallempfindlichkeit von den tiefsten bis zu den höchsten Tönen zu prüfen und in Kurvenform darzustellen, sodann diese Audiogramme mit denen von normalen Ohren und mit den ohrenärztlichen Befunden zu vergleichen<sup>2)</sup>. Der größte Vorzug dieser Methode ist die genaue Meßbarkeit und Regulierbarkeit der objektiven Tonstärken. Sie gestattet aber auch, wie schon erwähnt, Abbau- und Lückenversuche; während Synthesen sich damit bisher nur in unvollkommener Weise erzielen ließen. Ob Ohrenärzte sie in den Kliniken selbst verwendbar finden werden, dürfte besonders vom Kostenpunkt abhängen.

<sup>1)</sup> Für die Prüfung der Hörschärfe hat GRADENIGO schon 1916 (*Arch. ital. di Otol.* Bd. 28, 1918) nach dem Vorgange von WAETZMANN (*Zeitschr. f. Ohrenheilk.* 1911) eine If.-Einrichtung empfohlen. Aber zu den obigen Zwecken wurde sie damals meines Wissens noch nicht gebraucht.

<sup>2)</sup> Außer den Arbeiten von FLETCHER, WEGEL und LANE (*s. Lit.-Verz.*), worin diese Anwendungen eingehend berücksichtigt wurden, vgl. besonders FOWLER und WEGEL, *Audiometric Methods and their Application*, *Western Electric Comp. Eng. Dep. Oct.* 1922. J. P. MINTON, *Proc. Nat. Acad. Sc.* Bd. 7, S. 221ff. (1921). Bd. 8, S. 274ff. (1922), Bd. 9, S. 269ff. (1923, mit J. G. WILSON); *Phys. Rev.* Bd. 19, S. 80ff. (1922), Bd. 22, S. 506ff. (1923).

## 10. Kapitel.

# Systematik der Sprachlaute vom akustischen Standpunkte.

Im 5. Kapitel wurden die Grundeinteilungen der Sprachlaute in Vokale und Konsonanten und in stimmhafte und stimmlose erläutert. Es fragt sich nun, ob auf Grund der inzwischen dargelegten Tatsachen auch über die sachgemäße Untereinteilungen dieser Grundklassen oder über sich damit kreuzende Gesichtspunkte Näheres gesagt werden kann. Es genügt, wenn wir bei den Vokalen nur die stimmhaften, bei den Konsonanten nur die stimmlosen Formen ins Auge fassen, da für die Parallelklassen fast durchweg die gleichen Untereinteilungen gelten. Dabei ist für uns auch hier der akustische Standpunkt maßgebend. Es muß der Klarheit halber versucht werden, ihn so rein und konsequent als möglich durchzuführen, obschon wir nicht leugnen, daß sich für andere Zwecke auch andere Einteilungsgründe empfehlen. Unter den akustischen Tatsachen, die wir der Einteilung zugrunde legen, sollen hier aber nur solche berücksichtigt werden, die auch der unmittelbaren Gehörswahrnehmung nicht ganz verborgen bleiben. Denn nur dann wird sich die Einteilung zu allgemeinerem Gebrauch eignen. Die Ergebnisse der experimentellen Forschung werden zur näheren Erläuterung und tieferen Begründung des unmittelbar Wahrzunehmenden dienen.

Außer den folgenden Einteilungen finden sich bei Sprachforschern noch viele andere, die sich aber zum Teil nur im Ausdruck davon unterscheiden, zu einem anderen Teil genetischer (artikulatorischer) Art sind. So z. B. „schwere—leichte, harte—weiche, weite—enge, kurze—lange, tenues—mediae, vordere—hintere, schlaffe—gespannte, gerundete“ usf. Auf die 3 zuerst genannten wird das Gesetz der „Vokalharmonie“ bezogen, wonach in den türktatarischen, jakutischen, mongolischen, finnisch-ugrischen Sprachen Stammsilbe und Affixum zu derselben Klasse gehören (vgl. BÖHTLINGK, Sprache der Jakuten, 1851; RADLOFF, Phonetik der nördl. Türk Sprachen, 1882; WINKLER, Das Ural-Altäische, 1885). Es handelt sich dabei wesentlich um enge und weite Mundöffnung. So gehören O, A, Ö, Ä zu den weiten oder schweren, Ü, I, russ. Y zu den engen oder leichten Vokalen. Rein akustisch wüßte ich diese Gruppierung kaum zu definieren. Aber organo-genetisch ist ganz wohl zu begreifen, wie sich in der Praxis

des Sprechens eine solche Neigung zur „Angleichung“, nämlich der Mundstellungen, herausbilden kann.

Einen interessanten geschichtlichen Überblick der Vokalsysteme gibt MICHAELIS (2), für das Französische ROUSSELOT (2). Unsere deutschen 8 Vokale sind wohl zuerst so aufgezählt bei REYHER (1679), dann öfters im 18. Jahrh., auch bei A. v. HALLER (1778), der sogar meint, es gäbe wohl überhaupt, selbst in fremden Sprachen, keine anderen. In England herrscht seit WALLIS (1643) der organogenetische Standpunkt, der auch noch von BELL und SWEET ausschließlich durchgeführt wird. Während die akustische Richtung sich immer mehr auf das Vokaldreieck einigt, führte die genetische zu viereckigen, aber immer komplizierteren Konstruktionen.

## I. Vokale.

### 1. Sukzessiv ein- und mehrteilige (Monophthongen, Diphthongen, Triphthongen).

Da der (auch auf Konsonanten anwendbare) Begriff der Diphthongen schon im 4. Kapitel erläutert wurde, ist die Unterscheidung hier nur der Vollständigkeit halber anzuführen. Für Diphthongen ist außer den darin enthaltenen Teillaute das Hinüberschleifen des 1. in den 2. Bestandteil charakteristisch. Auch mögen die Bestandteile selbst gegenüber ihrer isolierten Aussprache oft etwas verändert sein. Außer den offiziellen Diphthongen des Hochdeutschen kommen im wirklichen und besonders im dialektischen Sprechen noch allerlei derartige Gebilde vor (z. B. im Oberbayerischen: „oans, guet“). In den gewöhnlichen vokalischen Diphthongen hat der 1. Teil den Akzent und die längere Dauer, und dieses Merkmal wird auch gelegentlich in die Definition aufgenommen. Doch ist im Grunde das „Ja“ und die 1. Silbe in „Jubel“, wenn J rein gesprochen und nicht mit einem Ch oder gar Teh eingeleitet wird, nichts anderes als ein vokalischer Diphthong aus I und A (U), nur mit dem Akzent auf dem 2. Bestandteil. Ebenso das U $\bar{\cup}$ O im italienischen Uomo. Daher wäre es wohl konsequent, den Begriff der Diphthongie auch auf solche Fälle auszudehnen. Wenn allerdings nach süddeutscher Art in „Theater“ das E in A ohne neue Artikulation hinübergezogen wird, so darf der Begriff nicht auch auf solche Sprachgewohnheiten übertragen werden, wenn er nicht seine feste Begrenzung zuletzt einbüßen soll.

Bei den weiteren Einteilungen sind hier natürlich nur die Monophthongen berücksichtigt.

### 2. Simultan ein- und mehrteilige.

Alle stimmhaften Vokale der wirklichen Rede sind Klänge, bestehend aus einer Mehrzahl einfacher Töne. In dieser Hinsicht kann man jedoch eine Unterscheidung machen, die sich ebenso

D. C. MILLER wie dem Verfasser aufgedrängt hat, aber auch schon in HELMHOLTZ' Vokalnotierungen enthalten und von ihm (S. 179ff.) ausdrücklich aufgestellt ist: in einteilige und zweiteilige oder in Vokale ohne einen und mit einem Unterformanten. Die hellen Vokale von Ö bis I bestehen aus 2 durch eine leere oder nur ganz schwach ausgefüllte Strecke getrennten Tonkomplexen, und die Lücke zwischen diesen wächst mit der Helligkeit des Vokals. Die dunkleren hingegen, U, O, A, haben keine solche Lücke, wenigstens nicht unterhalb des Formanten (U nur zwischen Formant und Oberformant).

In unmittelbarer Nähe tritt nach den Resonanzergebnissen auch bei O und A ein sekundäres Maximum auf; aber dieses hat für den Vokalcharakter so gut wie keine Bedeutung.

Mit dem Unterschied der Ein- und Zweiteiligkeit hängen mancherlei Folgeerscheinungen zusammen, wie z. B. der Übergang von E in O, von I in U bei Schwächung oder Wegfall des höheren Bestandteils durch äußere oder innere Hindernisse des Hörens; vielleicht auch Erscheinungen der Sprachgeschichte. Immerhin möchte ich die Bedeutung dieser Disjunktion nicht zu hoch anschlagen. Denn für die unmittelbare Wahrnehmung ist die simultane Zweiteiligkeit kein hervortretendes Merkmal. Man kann die Unterformanten nur bei erhöhter Aufmerksamkeit und Übung direkt wahrnehmen, und genauer sind sie überhaupt nur durch experimentelle Analyse festzustellen.

### 3. Die 5 sog. Hauptvokale.

Vielfach wird ein Rangunterschied gemacht zwischen Haupt- und Neben- oder ursprünglichen und abgeleiteten Vokalen. Vor allem begegnet uns hier die überlieferte Lehre von den 5 Hauptvokalen A E I O U. Diese kommen zwar in der Reinheit, wie sie Phonetikern vorschweben, bei weitem nicht in allen Sprachen vor; in manchen, wie der englischen, sind sie stark in der Minderheit. Aber man denkt sie sich gern als Ideale, von denen die Laute der einzelnen Sprachen nur mehr oder minder unvollkommene irdische Erscheinungsformen seien<sup>1)</sup>. Sie gelten, in

<sup>1)</sup> Selbst mancherlei Metaphysik und Mystik hat sich daran geknüpft. So erschien nach Zeitungsberichten unter den Extravaganzen moderner Kunst ein „Urkunstdrama“ eines jungen Holländers: „Das Weltgericht, eine Tragödie der uralten A E I O U“, wo alles nur in Gebärden und Vokalen vor sich geht. In den alten Zauberpapyri waren es die 7 Vokale des griechischen Alphabets, die solcher Mystik als Unterlage dienten.

Ganz anders steht es mit psychologischen Ausdeutungen der Sprachlaute, wie sie sich auf ihre onomatopoetische Funktion und ihre tatsächliche Verwendung im Sprachgebrauche gründen lassen, obschon auch da seit dem



einem freilich noch näher zu definierendem Sinn, als Urvokale, als Ausgangspunkte und Maßstäbe aller übrigen. Auch die ersten 5 Solmisationssilben enthalten diese Vokale. Im Französischen wird allerdings, gemäß der Aussprache des U, Ü im populären Gebrauche zu den 5 Hauptvokalen gerechnet, da es nun einmal im Alphabet so vorkommt<sup>1)</sup>. Die französischen Phonetiker aber schreiben bei der Aufzählung Ou für U, damit der dunkelste Laut in der Liste nicht fehle.

Worin könnte nun rein empirisch der Vorrang dieser 5 bestehen? Etwa in ihrer Einfachheit für das Gehör? Aber Ö, Ä und Ü sind für das direkte Gehör nicht minder einheitliche Laute wie E und I, wenn auch mit 2 Buchstaben oder mit gleichbedeutenden Strichen oder Punkten auf dem Kopfe geschrieben werden<sup>2)</sup>. Ja, sogar der Blöklaut AOÄÖ ist ein einheitlicher Laut, obschon er mit 4 (oder 6) Buchstaben geschrieben wird. Wenn man Ö und Ü als „Umlaute“ den Hauptvokalen gegenüberstellt und Ä als offenes E definiert, so hat dies sprachgeschichtliche Gründe, die für die rein phänomenologische Betrachtungsweise nicht maßgebend sein können. Denken wir, es wäre statt des mitteldeutschen Dialektes eine Aussprache wie die gegenwärtige braunschweigische zur hochdeutschen geworden, so hätten wir alle Veranlassung, statt des „reinen A“ ein A in dieser Form systematisch zu untersuchen, und man kann es in der Tat ebenso-

---

platonischen Kratylus manches Allzukühne gesagt worden ist (wie wenn W. v. HUMBOLDT in der Lautgruppe St regelmäßig den Eindruck des Beständigen oder J. GRIMM im K etwas Fragendes findet oder A. BÖCKH die Konsonanten als das materielle, tätige, männliche Prinzip den Vokalen als dem formellen, passiven, weiblichen gegenüberstellt). Schöne, anregende Betrachtungen dieser Art, die einem die Muttersprache noch lieber machen können, in GRABOWS „Musik in der deutschen Sprache“.

Abstufungen der „Euphonie“ unter den Vokalen und Konsonanten hat ZIEHEN (Vorlesungen über Ästhetik, Bd. 1, 1923, S. 208ff.) auf statistischem Wege festzustellen versucht, aber auch ihre malende Verwendung in Dichtwerken an interessanten Beispielen erläutert. Daß A und O allen anderen Lauten vorgezogen wurden, läßt sich wohl aus unseren Bemerkungen oben S. 189ff. verstehen. Im übrigen handelt es sich, wie ZIEHEN selbst hervorhebt, meistens um versteckte, dem Urteilenden selbst entgehende Nachwirkungen von Zusammenhängen, in denen die Laute in der individuellen Erfahrung vorkommen.

<sup>1)</sup> So hörte ich in einem Gefangenenlager 1917 ein südfranzösisches Volkslied mit dem humoristischen Refrain A E I O Ü (sic). Auch in einem BALZACschen Roman erinnere ich mich die 5 Vokale mit U = Ü zitiert gefunden zu haben.

<sup>2)</sup> Auch SIEVERS ist geneigt, Ü und Ö als gleichberechtigt neben die übrigen zu stellen (S. 92), obschon er sie mit WINTELER nur als „Vermittlungsvokale“ in das System einreihet.

gut als Gegenstand von Ab- und Aufbauversuchen durch Interferenz wählen, seine Formanten bestimmen und es synthetisch aus einfachen Tönen darstellen.

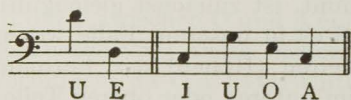
Auch die gewöhnliche Reihenfolge A E I O U läßt sich akustisch nicht rechtfertigen. Sie wurzelt offenbar in der Folge, in der sie im Alphabet auftreten, und hat darum nur historische Berechtigung. Akustisch würden wir sie ihrer Helligkeit nach ordnen und dann die Folge U O A E I erhalten<sup>1)</sup>.

#### 4. Das Helligkeitsprinzip.

Wie nun? Müssen wir etwa diese ganze Lehre ähnlich der von den „5 Sinnen“ zum alten Eisen werfen? Sind alle monophthongischen Vokale untereinander vollkommen gleichwertig und alle Rangordnungen willkürlich?

So weit darf man doch nicht gehen. Vielmehr läßt sich eine Systematik aufstellen, in der diese 5 eine ausgezeichnete Rolle spielen und aus der sie gewissermaßen deduziert werden können; freilich nur als Durchgang von einer noch ursprünglicheren Dreizahl zu noch reicheren Systemen.

Jeder Vokal hat, wie schon mehrfach erwähnt, als Gesamteindruck eine gewisse Helligkeit. Diese Gesamt- oder Komplexhelligkeit, die Klanghelligkeit im prägnanten Sinne, ist von der Tonhelligkeit, durch welche die Höhenordnung der Töne (sowie der Klänge mit vorherrschendem Grundton) bestimmt wird, wohl zu unterscheiden. Auf der nämlichen Tonhöhe bauen sich Vokale von verschiedener Klanghelligkeit auf, und diese Vokalhelligkeiten können sich auch im umgekehrten Sinne wie die Tonhelligkeiten bewegen. Singt man:



so geht die Tonhöhenbewegung immer entgegen der Vokalhelligkeitsbewegung: während jene steigt, sinkt diese und umgekehrt. Herrscht zwischen der Vokalverteilung und den Tonhöhen an akzentuierten Stellen ein allzu starker Gegensatz, so wird er immer-

<sup>1)</sup> Diese Anordnung findet sich schon bei OLEARIUS, Deutsche Sprachkunst 1630 („Diese Ordnung weiset uns die Natur“). Neuerdings ist sie besonders von KOENIG und KOEHLER betont worden, und zwar auf Grund ihrer Lehre von der Vokalität einfacher Töne, durch welche zugleich den 5 Hauptvokalen ihre Stellung als solche aufs neue vindiziert und akustisch begründet werden soll. Auf diese Lehre kommen wir erst im 13. Kap. zu sprechen. Hier beschäftigen uns nur die Vokale im gewöhnlichen Wortsinne, die Gebrauchsvokale, die niemals einfache Töne sind.

hin als ästhetische Unstimmigkeit empfunden, ja, es können Unbequemlichkeiten für den Sänger entstehen.

So z. B. in BEETHOVENS „Bußlied“ bei „Fluch“, in SCHUBERTS „Ode an den Unendlichen“ auf „Jubel“. Auch in „Allmacht“ kommen hervorstechende hohe Töne just auf „Sturm“, „Ruf“, „Flug“. Obgleich in diesem herrlichen Liede kein Ton anders sein dürfte, sind diese Stellen für den Sänger etwas mißlich. Ebenso in BRAHMS' Zigeunerliedern in Nr. 9 das stark und hoch einsetzende „Nur“. In LÖWES Ballade „Prinz Eugen“ gehen bei „Wer-da-Rufe“ die Ton- und Vokalhelligkeiten genau in umgekehrter Richtung. LÖWE benutzt übrigens in demselben Stück diesen Widerspruch mit Bewußtsein zu einer komischen Wirkung, indem er bei „tät den Schnurrbart streichen“ die Stimme auf „Schnurr“ um eine Oktave in die Höhe springen läßt.

Hervorragende Sänger und Chordirigenten ändern öfters den Text mit Rücksicht auf solche Diskrepanzen. Man findet aber auch nicht selten, so z. B. gerade in den Zigeunerliedern, wie an prägnanten Stellen U tief, I und E hoch gelegt sind. Eines der schönsten Beispiele ist das „Und es ward Licht“ in HAYDNS unsterblicher „Schöpfung“; weniger allerdings im englischen Text. In SCHUMANN'S „Paradies und Peri“ schließt die Peri ihren Triumphgesang mit den mehrmals wiederholten Worten: „Wie selig, wie selig bin ich!“, indem sie von  $f^2$  bis  $c^3$  in Terzen hinaufsteigt und auf  $a^2$  und  $c^3$  je 6 Takte liegenbleibt, während zugleich das Orchester mächtig in die Höhe geht. Die Wirkung wird durch die ausschließliche Verwendung der beiden hellsten Vokale im Text bedeutend unterstützt. Auch im Leben übrigens wird einer, der den andern gruseln machen will, sein „Huh!“ nicht in hoher Tonlage, und einer, der „Liebliches“ oder „Winziges“ auch im Tonfall recht anschaulich machen will, nicht in tiefer Stimmlage sprechen. UHLAND'S schalkhaftes „Teelied“, zumal die Strophe: „In Indiens mythischem Gebiete, Wo Frühling ewig sich erneut, O Tee, Du selber eine Mythe, Verlebst Du Deine Blütezeit“, wo fast nur die 3 hellsten Vokale vorkommen, wird jeder unwillkürlich mit hoher Stimme deklamieren.

Die Gesamthelligkeit gründet sich auf die Aussagen des unbefangenen, auch theoretisch unbeeinflussten Gehörs. Wie dessen Urteil zustande kommt, ist zunächst gleichgültig. Doch läßt sich leicht erkennen, daß es in unserem Falle wesentlich abhängig ist von den Tonhelligkeiten (Höhen) der Formanten. Je nachdem der Hauptformant im unteren oder oberen Teile der für die Vokalcharaktere überhaupt maßgebenden Tonregion liegt, sprechen wir von dunklen oder hellen Vokalen (ROUSSELOT nennt sie direkt tiefe und hohe). Allerdings haben auch die Unterformanten ihren Anteil daran. Man kann z. B. im Zweifel sein, ob Ä nicht heller sei als Ü. Dies kommt daher, daß zwar sein Formantzentrum etwas tiefer, aber sein Unterformant erheblich höher ist. Aber bei den 5 „Hauptvokalen“ können Zweifel nicht bestehen. Obschon auch hier E einen höheren Unterformanten als I hat, schlägt der höhere I-Formant in der Gesamthelligkeit durch.

Die Gesamthelligkeit entsteht nicht etwa durch Summation, ebensowenig aber durch Mischung oder Durchschnittsbildung aus

den Teilhelligkeiten. Man kann nicht sagen, daß das I durch seinen Unterformanten in der Helligkeitsskala der einfachen Töne herab-rücke, sowenig als sich aus dem Zusammenklang von *c* und *g* ein *e* bildet. Es handelt sich vielmehr um eine ganz andere Gattung von Helligkeiten, wofür eher der Ausdruck „Helldunkel“ am Platze wäre.

Will man nun ein System von Vokalen nach der Helligkeit aufstellen, so ergeben sich zunächst als ausgezeichnete Punkte der überhaupt dunkelste und der überhaupt hellste, U und I. Fragt man dann nach einem Vokal von mittlerer Helligkeit zwischen diesen Extremen, gleich weit von dem einen und anderen, so ist zweifellos ein helles A ein solcher Vokal. Soll weiter zwischen A und den beiden Grenzpunkten je ein Glied eingeschaltet werden, so liegt O zwischen U und A, E zwischen A und I. Versinnlicht man sich diese Verhältnisse in räumlicher Form, so würden sie sich bis dahin innerhalb einer geraden Linie darstellen lassen:

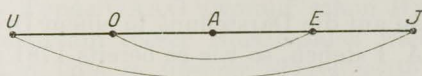


Abb. 3. Schema der 5 „Hauptvokale“.

Die Bogen deuten hier die zu den hellen Vokalen als Unterformanten gehörigen dunklen an, und so wäre dies schon ein übersichtliches und akustisch interessantes Schema und zugleich eine Art Deduktion der 5 „Hauptvokale“.

Die gleichen Abstände in der Zeichnung würden aber der Tatsache nicht gerecht werden, daß der Unterschied von U und O, E und I uns geringer erscheint als der von O und A, A und E<sup>1)</sup>. Dies führt zunächst dazu, zwischen O und A ein OA (AO), zwischen A und E ein Ä einzuschalten. Diese beiden wären auch wieder durch Bogen zu verbinden, da OA (AO) Unterformant für Ä ist. Und wir hätten nun 7 Glieder, also eine besonders heilige Zahl<sup>2)</sup>.

<sup>1)</sup> So sagt auch der Gesanglehrer G. ENGEL (2): „Wir können weit mehr Vokalnancen von O zu A als von U zu O hervorbringen.“

<sup>2)</sup> M. THAUSING legt in seinem viele gute Beobachtungen enthaltenden, aber der akustischen Grundlage entbehrenden „Natürlichen Lautsystem der menschlichen Sprache“ (S. 40) im Ernst Gewicht darauf, daß die 7 Stufen, die er auf jeder seiner 3 Hauptrichtungen unterscheidet, eine heilige Zahl darstellen, die „für die Evidenz des Systems nicht ganz gleichgültig“ sei. Auf diese Art von Evidenz möchten wir lieber verzichten, zumal da bei der Methode der symmetrischen Zwischenglieder innerhalb einer Geraden jedesmal eine ungerade Zahl herauskommen muß und alle ungeraden Zahlen bis 9 (im Kölner Fasching sogar 11) da und dort, dann und wann als heilig gelten.

Aber Ö und Ü, die auf dieser Linie zwischen A und E, auch gelegentlich mit einem davon zusammenfallend (helligkeitsgleich), unterzubringen wären, würden die Symmetrie doch wieder stören; desgleichen würden die Bogen ihrer Unterformanten die der Unterformanten anderer Vokale schneiden.

Außerdem würde es sich hier doch in Wahrheit nicht um eine eigentliche Deduktion handeln, sondern nur um eine möglichst rationelle Anordnung der in der Erfahrung gegebenen, in unserer Sprache nun einmal gebräuchlichen Vokale. Unter diesen hat A eine mittlere Helligkeit. An sich würde man, wäre A nicht tatsächlich gegeben, zwischen U und I vielmehr nur ein mittleres Ü (in österreichischen Landen Ui geschrieben) einfügen. A ist eben nicht eine bloße Zwischenhelligkeit, sondern seinem ganzen Charakter, seiner Qualität nach von den beiden Ausgangspunkten verschieden.

#### 5. Das Vokaldreieck.

So kommt man auf die Darstellung in einem Dreieck, an dessen Eckpunkten U, A, I stehen, so wie es bereits 1781 der junge Mediziner HELLWAG angegeben hat<sup>1)</sup>. Auf diese Konstruktion ist man

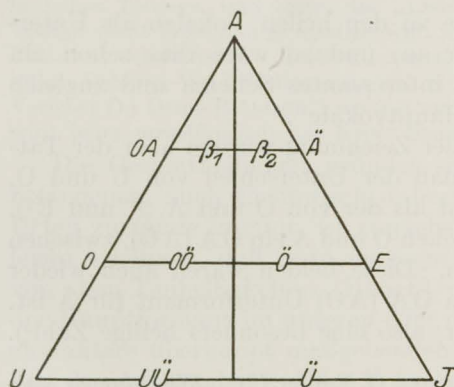


Abb. 4. Das Vokaldreieck.

in Deutschland seitdem zu meist wieder zurückgekommen, auch ohne von HELLWAG zu wissen, und dem Verfasser ist es ebenso gegangen. Aber auch MILLER hat sich darauf geführt gesehen. Wir haben davon bereits bei der Darstellung der Flüsterhöhen (o. S. 145) Gebrauch gemacht, auch dort durch den inneren Zwang der Sache dazu geführt. Dieses Schema gibt die möglichen Nuancen, Übergänge und Verhältnisse

in überraschender Vollständigkeit wieder und ist allen anderen weitaus vorzuziehen.

Wie viele Zwischenstufen man auf jeder Dreieckseite aufträgt, ist zunächst willkürlich, da stetige Übergänge von einem Eck-

<sup>1)</sup> Bekannt wurde das Schema besonders durch CHLADNI, dem es HELLWAG mitteilte (TRAUTMANN, S. 58), der aber seinen Vorgänger nicht nennt.

punkt zum anderen führen. Am zweckmäßigsten scheint es, je 2 Durchgangspunkte auszuzeichnen, wie auf der Abbildung geschehen ist. Auf der U—I-Linie liegt offenbar Ü, näher an I als an U. Setzt man symmetrisch dazu auf die linke Seite ein UÜ, welches sehr wohl herstellbar ist (vgl. oben S. 146), so hat man ein geschlossenes System von 9 Vokalen, unter denen allerdings unser Ö noch fehlt. Will man nur je 1 Zwischenstation auf jeder Linie unterscheiden, so erscheinen 6 Grundlaute. Je nach Bedürfnis kann man aber auch 3 Zwischenstufen (wie BRÜCKE) oder 4 (wie STOCKHAUSEN in seiner „Gesangstechnik“) oder noch mehr, insbesondere auf den seitlichen Dreiecklinien, unterscheiden.

Wir können nun aber auch in das Innere des Dreiecks gehen und finden da auf der Verbindungslinie von O und E unser Ö, dem links ein OÖ entsprechen würde, und auf der Linie von OA nach Ä, die die Linie von A nach den Ö-Lauten hin schneidet, Blöklaute vom Typus OAÄÖ (worauf schon HELLWAG hinwies). Unterscheidet man auf dieser Horizontalen wieder 2 Zwischenvokale,  $\beta_1$  und  $\beta_2$ , die in unserer Schreibweise am besten mit AOäö und Äöao (dunkler und heller Blöklaut) auszudrücken wären, so entstehen 13 Vokale, die sämtlich, ebenso wie alle übrigen hier nicht näher bezeichneten überhaupt möglichen Zwischenstufen, von A ausstrahlen. Will man aber auf der obersten Horizontalen nur 1 Zwischenstufe in der Mitte unterscheiden, so entsteht ein Dreieck im Dreieck (so bei BRÜCKE), und wir erhalten gerade ein Dutzend. Ich würde die erste Form, wie sie in unserer Zeichnung angegeben ist, trotz der Unglückszahl 13 vorziehen. Jedenfalls aber haben wir hier ein geschlossenes und übersichtlich geordnetes System. Suchen wir z. B. das kurze „russische Y“, so liegt es auf der Verbindungslinie von OÖ und UÜ, bald den einen, bald dem anderen näher. Schon HELLWAG bemerkt in den von VIËTOR herausgegebenen Aufzeichnungen vom Jahre 1780: „Zwischen diesen Reihen und Stufen könnte man noch unendlich viele andere einschalten, welche Völker von verschiedenen Sprachen und Mundarten im Sprechen gebrauchen: so ließen sich vielleicht alle Vokale und alle Diphthongen, welche je ein Mensch ausgesprochen hat, gleichsam mathematisch durch Stufen bestimmt angeben.“

Die nasalierten Vokale allerdings lassen sich nur so in dem Dreieck unterbringen, daß man sie durch einen Index ( $O_n$ ,  $\ddot{A}_n$  usw.) den entsprechenden freien Vokalen zuordnet. Diese Unterscheidung folgt eben einem anderen Einteilungsgrund, der sich mit dem gegenwärtigen kreuzt (s. unten).

An dem Vokaldreieck lassen sich nun auch die Grundeigenschaften der Vokale in einfacher und anschaulicher Weise aufzeigen:

a) Man kann sagen, A besitze die größte vokale Sättigung, es entferne sich am weitesten von der Linie der bloßen Helligkeiten. Und man kann überhaupt die Höhe der auf die U—I-Linie von einem Vokal aus gezogenen Senkrechten als Maß seiner vokalen Sättigung betrachten.

b) Die Lage des Endpunktes dieser Senkrechten auf der Grundlinie gibt ferner die Stellung des Vokals in der Helligkeitsreihe. Rechts von der Mittellinie liegen die hellen Vokale, die auch Unterformanten besitzen, links davon die dunklen, die keine Unterformanten (bzw. statt derselben Oberformanten) besitzen.

c) Die Unterformanten selbst findet man überall, wenn man der Horizontallinie in der Richtung nach links bis zur linken Dreieckseite folgt. Die Länge dieser Strecke wächst wie die der „toten Strecken“ zwischen Unterformant und Formant.

d) Selbst in Hinsicht der Klangstärke gibt das Dreieck gewisse Anzeigen: U und I sind die schwächsten, A ist der stärkste Vokal, d. h. man kann ihm die größte Stärke verleihen. So stellt es auch in dieser Hinsicht den Gipfel des Dreiecks dar. Doch möchte ich diesen Gesichtspunkt nicht so wie die übrigen ins einzelne durchführen.)

e) Durch die Senkrechte zur Grundlinie dürfte auch die objektive Stärke des Grundtons angezeigt werden, indem diese der Höhe der Senkrechten reziprok gesetzt werden kann. Wir hörten, daß von U zu A die Stärke des Grundtons ab-, dann nach E und I hin wieder zunimmt. Doch läßt sich auch diese Regel, die sich besonders bei den Resonanzversuchen an den 5 „Hauptvokalen“ deutlich herausstellte, zunächst nur im großen und ganzen als allgemeines Verhalten in Anspruch nehmen. Für die auf dem Dreiecksumfang liegenden Vokale dürfte sie aber ohne weiteres gelten.

f) Daß die beobachteten Tonhöhen der Flüstervokale und deren Übergänge ineinander sich am vollständigsten im Vokaldreieck darstellen lassen, ist bereits im 6. Kapitel gezeigt.

g) Auch in hirnpfysiologischer Richtung wird uns im 13. Kapitel das Vokaldreieck noch als Ausgangspunkt von Hypothesen dienstbar werden.

So möchten wir diese Anordnung als das „natürliche System“ der Vokale in Anspruch nehmen. Denn ein natürliches System ist auf wesentliche Merkmale im Sinne J. ST. MILLS gegründet, d. h. auf solche, an die sich viele andere anknüpfen. Es umschließt

auch die zahllosen historischen Systeme, deren jedes für die betreffende Nation und Zeit besondere Stellen des Vokaldreiecks vor anderen auszeichnet, ähnlich wie bestimmte instrumentale Klangfarben und wie die besonderen Intervalle der verschiedenen gebräuchlichen Leitern, ja auch gewisse absolute Tonhöhen für das Gehör der bezüglichen Nation oder Zeit ausgezeichnet sind, oder auch wie die Stimmen der Haustiere sich für die Landbewohner von allen übrigen Klängen abheben. Aber diese Auszeichnungen sind im allgemeinen akustisch zufällig und bedeutungslos, wenn auch wohl möglicherweise in einer Sprache mehr die dunklen, in einer anderen die hellen Vokale oder eine sonstige besondere Gegend des Vokaldreiecks hervortreten, also ein gewisser akustischer Grundzug herrschend sein kann.

Daß aber auch in sprachgeschichtlicher Hinsicht wenigstens die 3 Eckpunkte unseres Schemas eine Rolle spielen, ist bekannt. Sie sind die indogermanischen Urvokale. Auch auf volkstümliche Zusammenfügungen in der Gegenwart, wie „piff paff puff“, „bim bam bum“, „lirum larum“, kann man hinweisen.

Um die idealen Eckpunkte des Dreiecks, das vollkommenste U, A, I, genau festzulegen, könnte man die Teiltonstrukturen heranziehen. Bezüglich des A hat bereits ROUSSELOT vorgeschlagen, seine besonders vielfachen und ausgeprägten Nuancen durch die Lage oder Schwingungszahl des Formantenzentrums zu definieren, ähnlich wie man eine homogene Farbe durch ihre Wellenlänge definiert. Aber es müßten dann unbedingt auch die übrigen wesentlichen Teiltöne und die relativen Stärken aller angegeben werden, da ja durch eine geringe Verstärkung des unter dem Formanten liegenden  $c^2$  oder  $h^1$  schon eine beträchtliche Verdunkelung entsteht. Zunächst genügen indessen hier wie bei U und I aus direkter Beobachtung geschöpfte Momente zur Bestimmung.

U ist am vollkommensten bei einer Männerstimme in der unteren Hälfte der kleinen Oktave, die ja auch die Grundtöne der gewöhnlichen Männersprache enthält (über die Gründe s. 13. Kap. unter 2). I pflegen Frauen und Kinder am vollkommensten hervorzubringen, weil stärkere Teiltöne ihrer Stimme bis zur 5-gestrichenen Oktave reichen. Über das „an sich beste“ A wird man natürlich streiten. Deduktiv ließe sich sagen, es müsse so hell sein, daß es zwischen den einteiligen und zweiteiligen Lauten gerade in der Mitte steht, d. h., experimentell gesprochen, daß der Geübte bei aufmerksamster direkter Beobachtung zweifelhaft bleibt, ob schon Spaltung in eine obere und untere Abteilung stattfindet oder nicht. Dazu eignet sich vorzüglich das Flüster-A.



Bei dem entsprechenden gesungenen wird dann die Analyse — einerlei auf welchem Wege — bereits eine Ausbuchtung in den Teiltonstärken aufweisen, die aber noch nicht groß genug ist, um zweifelsfrei bemerkt zu werden. Das amerikanische A in „father“ scheint an dieser Grenze zu stehen, da es bei Analysen bald als einteilig, bald als zweiteilig angegeben wird. Darin verraten sich zwar wirkliche kleine Unterschiede der Aussprache; aber diese Schwankungen würden eben nicht so leicht über die Grenze führen, wenn nicht der Laut im allgemeinen an der Grenze läge. Das italienische A scheint ihm gleich zu sein. Dieser Laut wäre also an die Dreieckspitze zu setzen. Unser deutsches A, wie wir es durch die Bedingung möglichst gleicher Entfernung vom O und Ä charakterisierten und den Versuchen zugrunde legten, hat seinen Platz wohl ein wenig links davon.

Hiernach würde man, wenn Sprachforscher es erwünscht finden sollten, auch die entsprechenden idealen Teiltonstrukturen angeben können. Inzwischen sind unsere Strukturtabellen, und zwar für das vollkommenste deutsche A die auf dem Grundton  $c^1$  (wo  $g^2$  am stärksten herauskommt), für das vollkommenste U die auf  $c$ , für das vollkommenste I die auf  $g^1$  und  $c^2$  zu vergleichen.

Eine theoretisch nicht uninteressante Frage hat kürzlich K. HUBER aufgeworfen: ob man recht habe, den Übergang von dunklen zu hellen Vokalen (U—Ü, O—Ö, A—Ä) als einen stetigen anzunehmen. Er verneint diese Frage. Nur die Übergänge zwischen U, O und A, ebenso zwischen Ä, E und I seien streng stetig, die obigen dagegen schlossen rein erscheinungsmäßig jedesmal eine Diskontinuität ein, und zwar lägen die Punkte, wo der Sprung eintrete, in einer Senkrechten, die etwas rechts von A durch das Dreieck zu legen sei.

Mit unserer Anordnung würde dies nicht ganz stimmen, da OÖ, UÜ dort links von der Mittellinie liegen. Aber dies wäre ein sekundärer Unterschied, der sich beseitigen ließe; die Hauptfrage wird dadurch nicht berührt. In bezug auf diese selbst wird man zugeben müssen, daß der Übergang von 1- zu 2-teiligen Lauten, rein abstrakt genommen, einen Sprung bedeutet. Aber schon physikalisch kann sich das Minimum, wie schon erwähnt, stetig vorbereiten, und so kann auch für das Hören, sowohl das analysierende als das nichtanalysierende, der Übergang stetig sein; für das erste insofern, als der Hörende im Zweifel sein kann, an welchem genauen Punkte der helle Bestandteil zuerst auftaucht (s. die If.-Versuche, wo die Aufmerksamkeit speziell auf diese Frage gerichtet war), für das nichtanalysierende Hören insofern, als offenbar eine, sei es auch kleine, Zone innerhalb der Strecke A—Ä existiert,

in der man schwanken kann, wohin man einen gegebenen Laut rechnen soll. Und diese Zone ist nach beiden Seiten auch selbst nicht mathematisch scharf abgegrenzt, sondern ihre Begrenzung kann auch wieder zweifelhaft sein.

Ferner erscheint mir der bergang von A nach  mit dem von A nach U nicht blo kommensurabel, sondern sogar deutlich als kleiner erkennbar. Man mute daher fragen, worin eigentlich noch das Kriterium der Unstetigkeit liegen und durch welche experimentelle Fragestellung darber entschieden werden soll. Sagen wir: „Beim bergang A— wird eben das erste Auftreten einer -Frbung als etwas Neues, Eigenartiges empfunden“, so gilt dies doch ebenso beim bergang O—A. Ich finde daher vorlufig keinen zwingenden Grund zu obiger These.

Manche werden sich versucht fuhlen, das Vokaldreieck mit dem YOUNG-HELMHOLTZschen Farbendreieck oder dem HERINGSchen Farbenviereck zu vergleichen. Das Farbendreieck ist aber aus den Tatsachen der physiologischen Farbmischung abgeleitet, die von denen der Tonverknupfung ganz prinzipiell verschieden sind. Das Farbenviereck ist zwar in 1. Linie auf rein phanomenologische Erwagungen gestutzt; aber schon der Umstand, da man dort nicht ins Innere gehen kann, bedingt (neben anderem) eine grundwesentliche Verschiedenheit. Derselbe Umstand lehrt zugleich, da man mit geometrischen Deduktionen aus solchen Veranschaulichungen vorsichtig sein mu, denn ein Viereck, in welchem gegenuberliegende Seiten nicht durch Linien verbunden werden konnen, das uberhaupt kein Inneres hat, gibt es in der Geometrie nicht. Das Vokaldreieck ist in dieser Hinsicht ein richtiges Dreieck; aber auch daraus darf man nichts Neues mit mathematischer Evidenz ableiten wollen, sondern kann ihm nur etwa Winke fur neue Beobachtungen entnehmen. In der Hauptsache hat es seinen Wert als nachtragliche anschauliche Zusammenfassung der bereits vorliegenden.

## 6. Freie und naselnde Vokale.

Diese Unterscheidung ist fur die normale deutsche Aussprache von geringem Belang. Die naselnde Klangfarbe mancher Stimmen und pathologische Falle haben gleichwohl zu Untersuchungen Anla gegeben. Auch bei normalem Sprechen kommen Annaherungen an das Naseln in bestimmten Zusammenhangen vor, wie bei den Vokalen in „Angst, eng, Enkel“, infolge der vorauswirkenden Einstellung auf den Nasalkonsonanten. Aber zu einer besonderen Einteilung der Vokale wurden uns diese Umstande nicht veranlassen; sie ist hier nur mit Rucksicht auf ihre wesentliche Be-

deutung in anderen Sprachen eingefügt. Experimentell hat sich uns ergeben, daß das Charakteristische der Nasallaute in gewissen Lücken unter ihren tieferen Teiltönen und in der Beimischung von Teiltönen aus der 4-gestrichenen Oktave liegt. Diese Züge bleiben aber auch der unmittelbaren Wahrnehmung bei geschärfter Aufmerksamkeit nicht ganz verborgen. Man braucht nicht lange zu suchen, um in näselnden Klängen etwas Hohes, Dünnes und selbst eine Art Hohlheit zu entdecken, worauf denn auch schon HELMHOLTZ hingewiesen hat.

### 7. Offenes und gedecktes Singen und offene und geschlossene Aussprache.

Diese beiden miteinander verwandten, aber auf ganz verschiedenen Motiven und Vorgängen beruhenden Unterscheidungen<sup>1)</sup> sollen wegen ihrer praktischen Bedeutung hier noch besprochen werden, obschon sie in Wahrheit nicht weitere akustische Unterschiede von Vokalen betreffen, sondern nur die Zuordnung verschiedener Vokale zu gleichen Buchstabensymbolen.

a) Die erste Unterscheidung hängt mit den Erfordernissen der Gesangstechnik und der Klangschönheit zusammen. Wird I von einer Männerstimme im Brustregister auf Tönen höherer Lage, etwa *e*<sup>1</sup>, kräftig gesungen, so nähert es sich, wenn nicht besondere Kunstgriffe angewandt werden, dem E und zuletzt sogar dem Ä. Um dies zu vermeiden, gebraucht der Kunstsänger statt der „offenen“ die „gedeckte“ Stimmgebung (deren Mechanismus nach PIELKES laryngoskopischen, MUSEHOLDS stroboskopischen und SCHILLINGS Röntgenuntersuchungen wesentlich auf Unterschieden in der Stellung des Kehlkopfes und des Kehldeckels beruht, wo durch Verteilung der Arbeit auf die verschiedenen Muskelgruppen, Verlängerung des Ansatzrohres und Verschiebung der Resonanzverhältnisse bewirkt wird). Dann bleibt es in der I-Sphäre, wird aber nicht so kräftig und schmetternd. Bei größerer Kraftentfaltung wird daher doch wieder gelegentlich die offene Stimmgebung auch in der Höhe herangezogen. SCHILLING hebt neuerdings hervor, daß für Kunstsänger immerhin die Möglichkeit bestehe, durch kompensatorische Mundstellungen der Alteration der Vokale auch bei offener Stimmgebung entgegenzuwirken. Aber ohne solche besondere Maßregeln, die wohl nur wenige beherrschen, wird man ihnen unterliegen.

<sup>1)</sup> Zur Vermeidung von Mißverständnissen schlägt SCHILLING vor, die offene Stimmgebung „ungedeckt“ zu nennen, was sich in der Tat namentlich für Gesangspädagogen empfehlen dürfte. Wir bleiben hier zunächst bei dem überlieferten Sprachgebrauch.

Diese spontanen Umwandlungen beim offenen Singen beginnen mit aufsteigender Tonhöhe an verschiedenen Punkten: für U, Ü und I schon bei *fis—g* (wenigstens auf gedehnten Silben), dann folgen die übrigen, zuletzt A, das ebenso wie Ä in beiden Formen wesentlich unverändert in bezug auf seinen Vokalcharakter, wenn auch nicht in bezug auf Klangschönheit, bis in die höchste Lage der Männerstimme gesungen werden kann. In diese beiden Vokale verwandeln sich beim offenen Singen zuletzt alle anderen. Aber die Klangfarbe gleicht dann mehr der eines schreienden Instruments als einer guten Menschenstimme<sup>1)</sup>. Also eine Vernichtung der Vokalunterschiede, wie bei der Frauenstimme von *c*<sup>2</sup> aufwärts, aber aus verschiedenen Ursachen. Denn dort handelt sich um fortschreitende Verarmung an Obertönen der entsprechenden Regionen, hier aber vermutlich umgekehrt um ein Reicher- und Stärkerwerden der Obertöne, besonders in der 2- und 3-gestrichenen Oktave, wodurch die Unterschiede gleichfalls verschwinden müssen. Man kann diese akustischen Umwandlungen auch, wie dort, in die einfache Regel fassen, daß die Vokale sich nach der Mitte des Vokaldreiecks hin verschieben.

Diesen Veränderungen wirkt also die Deckung entgegen und bewirkt, daß die Vokale den vorgestellten, intendierten, nahebleiben. Ob dies völlig gelingt, ist die Frage. SCHILLING ist geneigt, sie zu bejahen: nicht die Vokalisierung, sondern nur die Klangfarbe der Stimme werde durch die Deckung beeinflusst. In Ermangelung eigener Erfahrungen und Beobachtungen nach dieser Richtung glaube ich doch aus seinen Versuchsergebnissen entnehmen zu können, daß kleinere Verschiebungen auch in der Vokalanuance nicht ganz dabei vermieden werden, wenngleich der Vokal innerhalb seiner Sphäre verbleibt. PIELKE hat mit GUTZMANN die Veränderungen bei den 5 Hauptvokalen durch Aufnahme von Schallkurven untersucht, aus denen GUTZMANN die Amplituden der Teiltöne nach dem Fourier-Verfahren berechnete. Die Diagramme zeigen bei den offenen Vokalen Verstärkung des 2., Schwächung des 1. Teiltons. Bei U, O, A ist uns dies ohne weiteres verständlich, bei E und I betrifft die festgestellte Veränderung offenbar die Unterformanten, die sich ja beim Übergang von I zu E von U nach O, bei dem von E zu Ä von O nach AO wandeln, also wieder beidemal in die Höhe gehen.

<sup>1)</sup> PIELKE, der zuerst diesen Vorgang eingehend untersuchte, bringt ergötzliche Beispiele aus der Praxis von Operntenören, wie „Glöck“ statt Glück, „wannevoll“ statt wonnevoll, bemerkt übrigens, daß auch beim scharf akzentuierten Sprechen ähnliche Umwandlungen (Fretz statt Fritz u. dgl.) vorkommen.

Auffallend ist aber in GUTZMANN'S Diagrammen (bei PIELKE 1), daß die Formanten bei A, E, I außerordentlich schwach herauskommen. Im Verhältnis dazu sind die Amplituden der beiden ersten Teiltöne oder des einen von ihnen ganz enorm, was allen sonstigen Kurvenaufnahmen wie auch unseren Befunden widerspricht. Der Gegensatz der hohen und tiefen Teiltöne wächst noch, wenn man statt der Amplituden Intensitäten ( $a^2 n^2$ ) setzt. Bei dem gedeckten A auf  $cs^1$  z. B. verhält sich die Amplitude des Grundtons zu der des Formantenzentrums  $b^2$  wie 59,6:3, die Intensität sogar wie 320:7,3. Dieses verdächtige Ergebnis dürfte in der Membran des Apparates seinen Grund haben. Infolge dieser Erwägungen ist wohl auch der Unterschied zwischen dem Verhältnis des 1. und 2. Teiltons bei offenem und gedecktem Singen nicht so groß anzunehmen, wie er in den Diagrammen erscheint; doch dürfte es im allgemeinen damit seine Richtigkeit haben.

In den Lehrbüchern der Gesangkunst konnte ich, soweit sie mir bekannt wurden, nirgends eine akustische Definition des Unterschiedes finden, auch nicht in J. STOCKHAUSEN'S „Gesangstechnik“, — überall nur Beispiele. Nur der alte M. GARCIA, der die heutigen technischen Ausdrücke dafür noch nicht gebraucht, bezeichnet den Unterschied als den einer hellen und einer dunklen Klangfarbe (timbre clair—sombre). Freilich hebt auch er nicht hervor, daß damit eben die Vokalität selbst sich ändere. Es schweben ihm mehr ästhetische Kategorien vor: die helle Stimmform werde jenseits  $fs^1$  unangenehm kreischend, während die dunkle einen runden, edlen, lieblichen Ton erzeuge.

Übrigens gibt es in der ganzen phonetischen Literatur kein Kapitel, worin die Definitionen und der Sprachgebrauch so durcheinander gehen wie bezüglich der Unterschiede der Stimmgebung und der Stimmregister. Am meisten schwankt der Ausdruck „Kopfstimme“, den manche, wie selbst der verdiente Gesangsprofessor G. ENGEL sehr unsachgemäß mit „Falsett“ zusammenlegen. Vgl. NADOLECZNY'S Übersicht S. 642.

Bei der Frauenstimme kommt eine Art Deckung nur für die tiefere Lage, bis gegen die Mitte der 1-gestrichenen Oktave vor. Für die in hohen Lagen angewandte „Kopfstimme“, die eine gewisse Analogie zur Deckung bietet, fand SOKOLOWSKY schwächere Obertöne als bei der Bruststimme. GARTEN gibt an (3, VIII., S. 20), dabei höhere Formanten gefunden zu haben; was bedeuten würde, daß die Vokale (wahrscheinlich handelte es sich nur um U, O, A) heller gegeben wurden. Aber hier treten ohnedies unvermeidliche Veränderungen der Vokale ein, gegen die auch kein Registerwechsel mehr helfen kann.

Das Ganze ist sonach nur eine Angelegenheit der Stimmetechnik, nicht eigentlich der Vokaltheorie. Den Vokalen wird dadurch keine neue Dimension und keine neue Klasse hinzugefügt, wie es etwa bei den nasalierten Vokalen der Fall ist. Gesangstheoretisch bleibt es bemerkenswert, daß die vom Sänger tatsächlich hervorgebrachten Vokale in der Höhe nur durch besondere Maßregeln mit den vom Text vorgeschriebenen und vom Sänger selbst intendierten in Übereinstimmung gebracht werden und auch mit diesen Maßregeln noch kleineren Abweichungen aus-

gesetzt sind. Aber mit der Klassifikation der Vokale selbst hat dies nichts zu tun.

b) Wenn in der Theorie und Geschichte der Sprache von dem Unterschiede offener und geschlossener Vokale die Rede ist, so handelt sich's nicht um Erfordernisse und Bedürfnisse der Stimmgebung oder Lautschönheit, sondern um gewisse historisch in Gebrauch gekommene Unterschiede in der Aussprache eines Buchstabenzeichens. Linguisten pflegen akustische Erläuterungen darüber ebensowenig wie Gesangslehrer über die vorangehende zu geben; auch sie verweisen nur auf die verschiedene Hervorbringungsweise. Den akustischen Unterschied muß man aus den Beispielen erschließen. Hiernach verstehen auch die Linguisten durchgängig unter einem geschlossenen U, O, A ein in unserem Sinne dunkleres, unter einem offenen ein helleres (nach O, A, Ä neigendes); unter einem gedeckten Ö, Ä, Ü, E, I aber umgekehrt ein helleres, unter einem offenen ein dunkleres, nach den in der Helligkeitsordnung vorausgehenden Vokalen hinneigendes oder ganz in sie übergehendes. Sie pflegen Ä überhaupt nur als eine offene Form des E zu bezeichnen. Dies hat seine Grundlage in der Sprachgeschichte, die zu verschiedenem Aussprechen ein und desselben Buchstabenzeichens geführt hat. Man schreibt „Herr, fett, hell“, spricht aber „Härr, fätt, häll“, wenn auch das Ä etwas heller ist als etwa in „Mähren“. Ähnliches im Italienischen. Solche Mehrdeutigkeiten gehören eben zu dem ererbten Bestand aller historischen Sprachen. Immerhin ist die deutsche Sprache darin ärmer als viele andere; schon das Französische kennt ja außer dem E = E und dem E = Ä noch ein E = Ö (dem sich unser tonloses E nur nähert).

Zwar daß jedes Buchstabensymbol eine gewisse Breite seiner akustischen Bedeutung haben muß (ebenso wie jeder Ton- und Farbename, wenn sie nicht durch Schwingungszahlen oder Wellenlängen definiert werden), ist überhaupt selbstverständlich. Aber wenn man einmal die im Deutschen gebräuchlichen, wohlunterschiedenen und in der Regel auch durch besondere Zeichen ausgedrückten 8 Klassen auseinanderhält, so wäre vom rein logischen Standpunkt doch zu wünschen, daß sie sich wenigstens nicht überschneiden möchten. Das „offene E“ ist eben akustisch kein E mehr; es ist nicht „dieselbe Farbe in anderer couleur“, sondern eine andere Farbe. Akustisch ist Ä sowenig ein E, wie O ein U. Sagen wir: Käse, so sind die beiden Vokale ebenso verschieden wie die in Juno oder Rütli. Vergleiche ich Brumbären mit Brombeeren, so sind die ersten Vokale nicht verschiedener als die zweiten.

Man könnte vielleicht zur rationellen Rechtfertigung der Mehrdeutigkeit auf die verschiedenen Schlüssel hinweisen, durch die in der Notenschrift ein und dieselbe Note grundverschiedene — bis zu mehr als Oktavenweite absteigende — Tonbedeutungen erhält. So seien auch in der Buchstabenschrift den Zeichen je nach dem einzelnen Falle gleichsam unsichtbare Schlüssel vorgesetzt, im Französischen sogar in den verschiedenen Akzenten sichtbare. Die Praxis unserer Umgangssprache zeige aber, daß solche im Deutschen nicht vonnöten seien. Der Vergleich könnte bis zu einem gewissen Grade stimmen, wenn er auch naturgemäß hinkt. Indessen lebt sich auch so ja das heranwachsende, Schreiben und Lesen lernende Geschlecht in den Sprachgebrauch ein, und wir denken nicht an Abschaffung.

Aber die Theorie darf sich nicht dadurch beirren und zu Wendungen verleiten lassen, als läge hier eine doppelte Form des Vokallautes selbst vor, als bedeuteten „offen und geschlossen“ eine besondere Dimension des akustischen Eindruckes, nach der sich ein Vokal bei unveränderter Stellung im Vokaldreieck noch wandeln könnte. Die Einteilung kreuzt sich nicht mit der der Helligkeiten und der mit diesen variierenden Vokalcharaktere, sondern ist darin schon enthalten<sup>1)</sup>. Der offene Vokal liegt stets mehr nach der Dreiecksmitte zu als der geschlossene von gleicher Buchstabenbezeichnung.

Für die lebendige Sprache, ihre Geschichte und Theorie, ist die akustisch verschiedene Wiedergabe ein und desselben Buchstabenzeichens im Sinne offener und geschlossener Aussprache namentlich dadurch von Bedeutung, daß sie eng mit der von kurzen und langen Silben verknüpft ist und damit offenbar auch kausal zusammenhängt<sup>2)</sup>. VIËTOR verbindet beide geradezu in der Definition als „lange und geschlossene, kurze und offene Vokale“. Es scheinen tatsächlich ausnahmslos die offenen Vokale kurz, die geschlossenen gedehnt zu sein (wie „fett — Seele“, „Mord — Wohl“<sup>3)</sup>). Doch ist nicht umgekehrt jedes kurze E zugleich ein offenes, vgl. Bett, Vetter, Rettung, Gesell, Elle. Ebenso kommen lange E vor, die mehr offen gesprochen werden. So scheint mir das E in „Wer?“ auch bei Dehnung näher an Ä als an E zu liegen.

Aber sogar noch ein dritter Unterschied kommt hinzu: die offen und kurz gesprochenen Vokale werden vielfach zugleich

<sup>1)</sup> Selbst ein so vortreffliches Werk wie VIËTORS Phonetik scheint mir von diesem, dem Sprachforscher allerdings naheliegenden, Mißverständnis nicht ganz frei zu sein.

<sup>2)</sup> SIEVERS, der die geschlossenen und offenen Formen mit WINTELER durch  $u^1-u^2$ ,  $o^1-o^2$  usf. bezeichnet, versucht für das Zusammentreffen mit Länge und Kürze eine psychologische Erklärung (S. 230), die auf der richtigen Fährte sein dürfte.

<sup>3)</sup> Auch in den Beispieltabellen der Gesangsschulen, bei STOCKHAUSEN, FRIEDLÄNDER, NOË und MOSER usf. fällt offen — geschlossen durchweg mit kurz — lang zusammen.

stärker hervorgestoßen, was ja auch psychologisch einigermaßen naheliegt. Auch beim Gesange gehört Kultur der Stimme dazu, um beim Forte in der Höhe geschlossen zu vokalisieren<sup>1)</sup>.

Aber um so deutlicher ist auch, daß ein Einteilungsgrund, der sich aus 2 oder 3 verschiedenen Attributen (Helligkeit, Zeitdauer, Stärke) zusammensetzt, deren Veränderungen nicht notwendig und immer parallel laufen, als Fundament einer Klassifikation sein Mißliches hat.

Selbst individuelle und provinzielle Eigenheiten wirken gelegentlich mit, um diese Angelegenheit zu verwirren. Wenn z. B. BRÜCKE das E in „ehrlieh“ als Analogon des französischen è anführt, so kann man sich des Gedankens nicht erwehren, daß die langjährige Wiener Umgebung ihn dabei beeinflußte, obschon er ein geborener Berliner war. Wenn ferner VIËTOR „Pferd, Erz, Wert“ als Beispiele des geschlossenen E benutzt, so scheinen mir diese auch nicht besonders glücklich gewählt. Soviel ich hören kann, ist die hochdeutsche Aussprache dabei in der Regel nicht so Ä-frei, wie etwa in „Rede, See“. Ein nachfolgendes R scheint verdunkelnd (öffnend) auf das E zu wirken. Zugleich zeigen die Beispiele wieder, daß lange Vokale nicht immer zugleich im vollen Sinne geschlossen zu sein brauchen, so wenig wie kurze immer offen sein müssen.

Daß überhaupt Zwischenstufen vorkommen, bei denen sich die individuelle Variabilität besonders geltend machen wird, liegt in der Natur der Sache und gehört mit zu den Unterschieden gegenüber dem Falle der musikalischen Schlüssel. So wird man in „Fehlen, Leben“ die 1. Silbe bald mehr geschlossen, bald mehr offen gesprochen hören; dem Verfasser würde die genaue Mitte zwischen E und Ä am angemessensten erscheinen. „Lerche“ steht schon wieder näher am reinen Ä, fällt aber auch noch keineswegs damit zusammen. Spricht man „Hehr, Heer, Herd, Herr, Hähne“, so werden wohl die meisten diese Reihenfolge der allmählichen Verdunkelung (Öffnung) approbieren.

BRÜCKE hat derartige Übergangsstufen feinhörig in Klassenform festgelegt. Aber sie dürften in der Praxis eine stetige Reihe zahlloser Nuancen bilden. Ein einfaches und sicheres Mittel zur Verständigung über Helligkeitsgrade bildet die Angabe der Tonhöhe beim Flüstern. So gibt der Gesangspädagoge TR. HEINRICH sein Ä in „Väter“ mit einem etwas erhöhten  $b^3$ , sein E in „Erde“ mit  $g^{is^3}$  an; beiden Notierungen könnte ich nur zustimmen. Die Tonhöhen des E in den soeben erwähnten 5 Wörtern sind in meiner Aussprache:  $c^4$ ,  $b^3$ ,  $a^3$ ,  $as^3$ ,  $g^3$ . Um einen halben Ton können sie immerhin je nach der Stärke und dem Affekt, mit denen geflüstert wird, schwanken, aber kaum mehr.

## 8. Multiple Vokale.

Ogleich die Vokale im gewöhnlichen Sinne, als farbige Laute individueller menschlicher Stimmen, ein in sich geschlossenes System bilden, ist es doch von entscheidender Wichtigkeit, daß dieses System selbst nur einen Ausschnitt aus den zahllosen für das menschliche Ohr wahrnehmbaren Klangfarben darstellt und

<sup>1)</sup> „Man wird finden“ — sagt STOCKHAUSEN —, „daß die offenen Vokale den stärkeren, die geschlossenen den schwächeren Ton begünstigen.“



stetig in seine Nachbarn übergeht. Zunächst können schon durch die menschliche Stimme selbst vokalartige Klänge hervorgebracht werden, die im Wortschatz keiner Sprache enthalten sind: indem nämlich mehrere Personen gleichzeitig verschiedene Vokale auf demselben Grundton singen oder sprechen. Diese wollen wir hier multiple Vokale nennen. Sie würden an sich, soweit ihre Teile noch einigermaßen unterscheidbar bleiben, unter die simultan mehrteiligen gehören, von denen unter 2. die Rede war, sind aber dadurch unterschieden, daß sie in der normalen Sprache nicht vorkommen.

Zur genauen Herstellung von Doppelvokalen muß man die Einzelvokale in verschiedene Leitungen sprechen lassen, die sich vor dem Ohre des in einem anderen Raume befindlichen Beobachters vereinigen, und muß durch einen drehbaren Hahn oder eine ähnliche Vorrichtung für genau gleichzeitigen Beginn und Schluß des Gehörseindrucks sorgen. Natürlich kann man in analoger Weise auch 3 und mehr, ja, alle unsere 8 Vokale vereinigen. Ich habe solche Versuche nur probeweise mit je zweien angestellt. Manche Vokale geben ohne weiteres einheitliche Mischprodukte, wie A und E, andere mischen sich nicht, wenigstens nicht für unsere Auffassungsgewohnheiten; doch wäre es wohl möglich, daß bei häufiger Darbietung in entsprechenden Stärkeverhältnissen auch da einheitlichere Eindrücke zustande kämen. A und Ö verschmelzen noch ziemlich gut, schwerer dagegen A und I. Solche Unterschiede gibt es ja aber auch unter unseren gewöhnlichen Vokalen, insofern die Vokale mit leeren Strecken weniger verschmelzen bzw. leichter durch das bloße Gehör zerlegt werden können als die aus einem Stück bestehenden, wie O oder A.

Sehr eingehend hat sich in den letzten Jahren K. HUBER mit solchen „Vokalmischungen“ beschäftigt. Die Versuche (weit über 600) waren unwissentlich, d. h. die Urteilenden hatten keine Kenntnis der jeweiligen Komponenten und konnten sie nur in 7% der Fälle richtig erschließen. Die 8 Vokale wurden in allen möglichen Kombinationen zu je zweien gemischt. Sie erschienen in sehr ungleichem Maße zur Einheit verschmolzen (4 Stufen). Die qualitativen Ergebnisse sind noch nicht ausführlicher veröffentlicht. Doch wird betont, daß die Resultante keineswegs immer der Ähnlichkeit nach zwischen den Komponenten lag, daß A durch keine einzige Kombination zustande kam, dagegen alle übrigen Qualitäten des Vokaldreiecks durch Mischungen von U, A, I in verschiedenen Intensitätsverhältnissen darzustellen waren. Es wurden aber auch Versuchsreihen mit Verteilung der beiden Eindrücke an beide Ohren („dichotischer“ Zuleitung gemacht<sup>1)</sup>). Dieselben Verschmelzungsstufen traten dabei auf, nur modifiziert

<sup>1)</sup> Versuche in dieser Form schon bei v. HORNBOSTEL (1, S. 75ff). Er verknüpfte so auch Vokal- mit Instrumentalklängen (Versuch 13) und hebt nicht mit Unrecht die Bedeutung solcher Beobachtungen für die physiologische Hörtheorie hervor.

durch die räumlichen Eigenheiten des dichotischen Hörens. Qualitativ aber ergaben sich hier immer Zwischenvokale; z. B. mischten sich A und I zu einem in ihrer Verbindungslinie auf dem Vokaldreieck liegenden Vokal, also Ä oder E. Doch gelang die Mischung von O und E zu Ö nur schwer und unvollkommen, die von U und I zu Ü überhaupt nicht mehr.

Es finden sich in HUBERS Ergebnissen, soweit sie mitgeteilt sind, neben völlig durchsichtigen und lehrreichen auch einige Paradoxien, die ich nicht aufzulösen wüßte. Eine Nachprüfung der Einrichtungen, die er freundlichst gestattete, führte nicht zur Einigung, lehrte aber wenigstens, daß es hier ganz besonders auf vollkommen genaue Nuancierung der Vokale in jedem Einzelfalle, sowie auf unveränderte Hinüberleitung in den Beobachtungsraum (bei unwissentlichen Versuchen ist lange Leitung erforderlich) ankommt.

In theoretischer Hinsicht kann ich der Vokalmischung nicht die grundlegende Bedeutung zuerkennen, wie sie der Farbenmischung zukommt. Dort werden Spektralfarben oder wenigstens nahezu einfache Farben von bestimmten Wellenlängen zur Deckung auf der Netzhaut gebracht. Hier hingegen handelt es sich schon um hoch zusammengesetzte Schwingungen. Auch haben die Mischungsvorgänge in der Netzhaut und im Sehzentrum kein Analogon im Ohr und im Hörzentrum, wo niemals aus 2 genügend verschiedenen Tönen ein mittlerer oder ein Geräusch (analog dem durch Rot und Grün erzeugten Weiß) entsteht. Ich glaube daher nicht, daß man über das Wesen der Vokale auf diesem Wege prinzipielle Aufschlüsse erwarten darf. Aber zur Anwendung und Bewährung bereits gewonnener Strukturkenntnisse, als Probe auf das Exempel werden sorgfältige Untersuchungen wie die HUBERSche immer wertvoll bleiben.

Wie die Teiltontabellen der kombinierten Vokale bei monotonischer bzw. amphotonischer Zuleitung aussehen müssen, läßt sich bei gleicher Stärke der Einzelvokale aus unseren Teiltontafeln ungefähr voraussagen, und danach könnte man auch ihre direkte Synthese aus einfachen Tönen bewirken. Zu berücksichtigen ist natürlich, daß die Grundtöne und die zusammenfallenden Teiltöne sich verstärken und infolgedessen auch die Wechselwirkungen unter den Teiltönen hinsichtlich ihrer physiologischen Intensität (12. Kapitel) Verschiebungen erleiden müssen. So kann z. B. bei U + I der I-Formant infolge Verstärkung des Grundtons (evtl. auch seiner Oktave) ganz unterdrückt und einfach ein U gehört werden. Überhaupt wird durchschnittlich der Mischvokal wegen der Verstärkung des Grundtons etwas dunkler ausfallen als jeder der Einzelvokale.

Niemals kann, wenigstens bei monotonischer Zuleitung, aus U + I irgendein anderer Vokal als einer dieser beiden, auch nicht ein Ü, herauskommen; ebensowenig aus O + E ein Ö. Dies ist so unmöglich, wie daß aus dem Zusammenklang c + g ein e entstände, ja, es ist sogar derselbe Fall: denn die Formanten von O und E und der dazwischen liegende Formant des Ö sind ja eben Töne von bestimmter Höhe. Die Wirkung von U + I, O + E kann

nur die sein, daß der Formant des I bzw. E gegenüber seinem Unterformanten geschwächt, unter Umständen unterdrückt wird. Höchstens könnten höhere, an sich unwesentliche Teiltöne des U und O sich mit solchen des I und E zu einer etwas veränderten Nuance der tiefen Vokale verbinden.

Auch A + I kann monotisch nicht einen auf der entsprechenden Dreieckslinie liegenden Vokal, Ä oder E, geben. Hier muß ein neuer Vokal entstehen, ein durch einen viel breiteren oder gespaltenen Unterformanten entstelltes I, wenn anders dessen Formant überhaupt genügend zur Geltung kommt. Wie freilich dichotisch zugeleitete Vokale sich anhören werden, kann man, da es sich dabei nicht um bloß physikalische Addition der Schwingungen, sondern um physiologische Prozesse handelt, nicht so einfach aus den Einzeltabellen voraussagen.

Auch die Mischvokale müssen ex constructione ihre Formanten haben, die fest bleiben, wenn die beiden Einzelvokale zu einer anderen Tonhöhe übergehen. Ihrer Komplexqualität nach würden sie sich, soweit sie einheitlich, also auch in einheitlicher Helligkeit erscheinen, entweder im Vokaldreieck anordnen oder in einer sonstigen geschlossenen Figur mit geradliniger Basis, in welcher unser Vokaldreieck enthalten wäre. An die Stelle der seitlichen, in der A-Spitze zusammentreffenden Geraden würde vielleicht eine einzige gekrümmte Linie treten. Aber es lohnt sich nicht, hierüber zu spekulieren. Das Interesse, das solche Fälle bieten, dürfte sich darauf beschränken, daß sie Beispiele liefern für die selbst innerhalb der menschlichen Stimme noch möglichen Vokalfarben und Material für Studien über einheitliches und mehrheitliches, evtl. monotisches und dichotisches Hören.

Wenn man die Mühe nicht scheut, könnte man auch noch weitere Komplikationen dadurch erzielen, daß die verschiedenen Personen ihre verschiedenen Vokale auf verschiedenen Grundtönen singen. Ja, man kann auch Vokale mit anderen Klängen mischen und dabei durch Intensitätsverschiebungen selbst stetige Übergänge aus dem einen in das andere Gebiet herbeiführen. Tatsächlich hören wir solche Mischungen immer bei begleiteten Gesängen, in denen auch alle möglichen Intensitätsverhältnisse vorkommen. Multiple Vokale im allgemeinen hören wir in der musikalischen Praxis immer, wenn mehrere Sänger oder Chöre gleichzeitig auf verschiedenen Texten singen, sei es in identischen oder in verschiedenen Tonfolgen. In der Opernmusik ist dies ja nicht selten der Fall. Allerdings treffen dabei die Sänger, wenn sie nicht unisono singen, nur gelegentlich und zufällig auf denselben Tönen zusammen. Aber auch bei verschiedener Tonhöhe müssen

Mischvokale herauskommen, wenn jeder seinen eigenen Text singt, nur noch kompliziertere. Wissenschaftliche Beobachtungen lassen sich dabei nicht machen, da es sich fast immer um eine Art *Parlando* in raschem Tempo handelt. Der Zweck dieser Singweise ist aber auch keineswegs, einheitliche neue Vokalfarben zu erzeugen; im Gegenteil soll der Hörer die Texte nach Möglichkeit auseinanderhalten, was ihm freilich auch nur sehr unvollkommen und — wenn er es überhaupt für der Mühe wert hält — nur mit Hilfe des Textbuches oder des Erratens aus der Situation heraus gelingt. Praktisch können daher diese hybriden Vokalbildungen nicht die Bedeutung in Anspruch nehmen, wie sie den durch Schallhindernisse entstehenden und durch künstlichen Abbau nachzuahmenden defekten Bildungen zukommt. In der Instrumentalmusik verhält es sich anders: da sind multiple Klangfarben von größter Wichtigkeit (s. 15. Kap. unter 9).

#### 9. Vokalitäten außerhalb der menschlichen Stimme.

Die multiplen Vokale bilden einen Übergang zu den zahllosen vokalähnlichen Klangfarben, die in der Kunst wie in der Natur vorkommen und für das menschliche Ohr vernehmbar sind. Schon wenn man mit unserer synthetischen Einrichtung den Formanten des I mit dem Unterformanten des E verbindet, entsteht ein Gebilde, das man mit der Stimme selbst kaum wird erzeugen können. Ebenso bei Phonographenversuchen (oben S. 229). Auch bei musikalischen Instrumenten finden sich Annäherungen an Vokalisches, auf die wir noch näher zurückkommen werden. Nimmt man noch Tierstimmen dazu, das Tirili der Lerche (anscheinend ein Ultra-I ohne Unterformanten), das Uhu der Eule, das Mäh des Lammes usf.<sup>1)</sup>, sowie mehr oder minder vokalähnliche Klänge der unbeseelten Natur, so muß man sicher zugeben, daß zwischen den menschlichen Vokalen und den übrigen Klangfarben scharfe Grenzen nicht bestehen. Auch das Merkmal der festen Formanten ist nicht an die Sprachlaute allein gebunden, sondern muß sich überall finden, wo durch eine feste Resonanz bei wechselndem Grundton Teiltöne von gleichbleibender absoluter Höhe verstärkt werden. Auch bei Instrumenten scheint dergleichen bis zu einem gewissen Grade vorzukommen (s. 15. Kapitel). Ein

<sup>1)</sup> SOKOLOWSKY hat sogar bei einem „sprechenden Hunde“, aus dessen Bellen man „Hunger, Kuchen, Ruhe“ heraushören wollte, eine U-Kurve aufgenommen und daraus regelrecht nach FOURIER 10 Teiltöne auf dem Grundton *dis*<sup>1</sup> berechnet, unter denen *dis*<sup>2</sup> der stärkste war (KATZENSTEINS Arch. f. exp. Phonetik B. 1, 1913). Das gebellte U muß doch stark nach O hin geklungen haben.

Instrument freilich, dessen Resonanz man so bequem und fein umstellen kann wie bei der menschlichen Stimme, gibt es bisher nicht.

Aus diesen Betrachtungen geht aufs neue hervor, wie wenig berechtigt es ist, für die Vokale der menschlichen Sprache ein ganz singuläres, unvergleichliches Wesen in Anspruch zu nehmen. Wie die Vokale einer einzelnen historischen Sprache nur ein Ausschnitt aus den für individuelle menschliche Sprachwerkzeuge überhaupt möglichen sind, so sind diese selbst wieder nur ein Ausschnitt aus der Gesamtheit der möglichen Klangfarben und hängen durch stetige Übergänge mit den übrigen zusammen.

Was sie von diesen unterscheidet, ist nur jener Vertrauheitscharakter, wie er im höchsten Grade den Lauten der Muttersprache, in geringerem aber allen durch ein menschliches Sprachorgan erzeugbaren Klängen zukommt. Er verleiht ihnen den Anschein der Unvergleichbarkeit. Aber er hat nichts mit einzigartigen Eigenschaften des Klangmaterials zu tun, sondern wurzelt nur in dem seelischen Verhalten der dem Menschengeschlecht und einer besonderen Nation angehörigen individuellen Hörer. Dadurch sind für jeden von uns Typen, Kategorien entstanden, auf die er alle anderen Laute bezieht. Aber ähnlich verhält es sich doch auch mit den heimischen Instrumenten gegenüber der Menge der möglichen instrumentalen Klangfarben. Es ist nicht einmal wahrscheinlich, daß den Lauten der Muttersprache in bezug auf die Leichtigkeit der Hervorbringung oder der Auffassung und des Wiedererkennens ein besonderer Vorzug für das sprechende Individuum innewohnt.

## II. Konsonanten.

Für die Klassifikation der Konsonanten wird seit alter Zeit von den Linguisten die Erzeugungsweise in den Vordergrund gestellt, von ihrem Standpunkte mit Recht, insofern organogenetische Beschreibungen zur Nachbildung anleiten sollen. Doch hatte man natürlich auch die akustischen Verschiedenheiten mit im Sinne, auf deren Erzeugung es ankommt; und kein Einsichtiger konnte jemals leugnen, daß das Ohr hier letzte richterliche Instanz sein muß. Uns kommt es nach dem Plane unserer Arbeit auf diese Verschiedenheiten allein an. Aber im Endergebnis müssen beide Einteilungen sich decken, da die Intention des Sprechenden bei der Einstellung seiner Sprachwerkzeuge eben auf die Hervorbringung akustisch verschiedener Eindrücke für den Hörer gerichtet ist. Wir werden finden, daß sich für die Grundeinteilung akustische,

für die Untereinteilungen genetische Gesichtspunkte in erster Linie empfehlen.

Wie schon erwähnt, dürfte auch für die Konsonanten die Grundeinteilung in Mono- und Diphthongen (Triphthongen) in Betracht kommen, obschon sie hier bisher nicht gebräuchlich ist. Warum sollte man nicht die schon anerkannten Gruppen der Aspiraten (wie Kh, Th)<sup>1)</sup> und der Affrikaten (wie Pf, Ts = Z, Ks = X) mit anderen, etwa dem slawischen Č, einem T mit auslautendem äußerst weichem, kurzem und hellem sch, unter den gemeinsamen Begriff konsonantischer Diphthongen fassen? Es verhält sich anders mit Zusammenfügungen wie Pt, die nicht mit einheitlicher Artikulation gesprochen werden und nicht durch einen stetigen Übergang akustisch verbunden sind.

Ja, es ließe sich daran denken, auch von Diphthongen zu reden, die aus Vokalen und Konsonanten gemischt wären, überall da nämlich, wo ein gleitender Übergang zwischen ihnen stattfindet, durch den sie zu einer artikulatorischen wie akustischen Einheit verschmelzen, wie bei den aspirierten Vokalen (Ha, Hu) oder den afrikanischen Schnalzlauten, in denen ein Explosivkonsonant in einen hellen oder dunklen Vokal mündet. Freilich würden diese „gemischten Diphthongen“ zuletzt in den lebendigen Strom der Rede überführen und die scharfe Abgrenzung des Begriffes gefährden. Es muß den Sprachforschern überlassen bleiben, bis zu welcher Grenze sie solche Erweiterungen ratsam finden.

Die weitere Einteilung hält sich auch hier an die Monophthongen. Und zwar ist es bei den Konsonanten vom akustischen Standpunkte das Richtige, sogleich die Hauptgruppen der Geräusche überhaupt zu unterscheiden, und die Konsonanten in diese Gruppen einzufügen<sup>2)</sup>. Damit ist von vornherein das erreicht, was bei den Vokalen zuletzt betont wurde, die Eingliederung in die Gesamtheit der verwandten Gehörseindrücke<sup>3)</sup>. Zunächst sei an die allgemeine Phänomenologie der Geräusche (5. Kapitel) erinnert und einiges Erläuternde hinzugefügt. Wir unterscheiden sie spezifisch von den Tönen und Klängen. Die allgemeine Geräuschqualität ist nicht

<sup>1)</sup> Im Norddeutschen werden die Tenues überhaupt mit angehängtem leichten h gesprochen, und es ist auch kaum möglich, sie ohne jede Spur eines nachfolgenden h kräftig auszusprechen. Im Süddeutschen werden sie bekanntlich von den Mediae fast nicht unterschieden.

<sup>2)</sup> So verfährt auch NAGEL und kommt im wesentlichen zu denselben 3 Grundklassen, wie sie im folgenden unterschieden werden.

<sup>3)</sup> Zum Überfluß könnte man auch hier als Übergang multiple Konsonanten auf dieselbe Weise herstellen, wie sie oben für multiple Vokale beschrieben ist, und so die merkwürdigsten Bildungen erzeugen, die zum Teil dem Fauchen wütender Katzen nicht unähnlich sein dürften.

weiter definierbar, weist aber analoge Modifikationen oder Veränderungsrichtungen auf wie die der Töne: Stärke, Helligkeit (Höhe), Volumen, musikalische Qualität, mehr oder minder ausgeprägte Farbigkeit, räumliche und zeitliche Eigenschaften. Die Flüstervokale sind ausgeprägt farbige, die stimmlosen Konsonanten wesentlich farblose Geräusche. Die meisten Gehörserscheinungen sind aber Mischungen von Tönen und Geräuschen. Die stimmhaften Konsonanten sind solche Mischungen.

Manche Erscheinungen werden überhaupt mit Unrecht den Geräuschen zugewiesen. So ist das Donnern eine ausgesprochen tonale Erscheinung („tonat“), wenn auch seine Tonhöhe wegen ihrer Tiefe und Veränderlichkeit schwer erkennbar ist. Es wird durch die Pauke oder eine Blechscheibe von tiefer Abstimmung deutlich nachgeahmt. Solche verkehrte Zuordnungen rühren daher, daß man sich über die wesentlichen Merkmale nicht klar ist. Nennt man Geräusch jede Gehörserscheinung, die hinsichtlich ihrer Höhe oder sonstiger Eigenschaften schwer oder gar nicht mit dem bloßen Ohre bestimmbar ist, dann gehört freilich der Donner in den meisten Fällen darunter. Dann gehören aber auch die Töne der 6- und 7-gestr. Oktave dazu, und sie sind tatsächlich auch mehrfach darunter subsumiert worden. Aber ich würde das wesentliche Unterscheidungsmerkmal zwischen Tönen und Geräuschen eben nicht in diesem Nebenumstande finden, bei dem es ja auch sehr auf die akustische Anlage und Übung des Hörenden ankommt, sondern in der Natur ihrer Grundqualität.

Die Mannigfaltigkeit der Geräusche ist ungeheuer<sup>1)</sup>. Aber 3 Hauptklassen lassen sich akustisch unterscheiden, wobei aller-

<sup>1)</sup> AUERBACH zählt (5, S. 276, vgl. 113) nicht weniger als 92 Geräuschnamen auf, denen sich noch manche (z. B. Krachen) hinzufügen ließen. Auffallend ist dabei die außerordentlich hohe Zahl der sprachlichen Onomatopöen. Freilich würden wir einige der aufgezählten Geräusche nicht als solche gelten lassen, so: Blöken, Krähen, Lachen, Miauen. Dies sind unzweifelhafte Klänge, sogar mit deutlichen Tonhöhen und Vokalcharakteren. Die Melodie eines Hahnenrufs kann häufig exakt in Noten wiedergegeben werden. Musikalische Werturteile müssen hier natürlich ausgeschlossen bleiben. Wenn uns die Klangfarbe der Tierstimmen musikalisch oft mißfallen mag, so gilt dasselbe von vielen exotischen Menschenstimmen und Instrumenten.

Für die Theorie des Hörens ist es nicht ohne Bedeutung, daß auch Geräusche bis zu einem gewissen Grade durch das bloße Gehör analysiert werden können. Vgl. o. S. 135. Ein schönes Beispiel gibt die Analyse des Meeresbrausens in STRINDBERGS Erzählung „Am offenen Meere“ S. 143, wo ein Brausen, Zischen, Schraßen, Rasseln, Sausen, Rollen unterschieden werden, die, verschiedenen Ursachen entstammend, sich bald zeitlich folgen, bald aber auch zugleich auftreten.

dings nicht die qualitative, sondern die zeitliche Beschaffenheit, ein etwas äußerliches akustisches Merkmal, als Einteilungsprinzip dient:

a) Dauergeräusche, d. h. solche, die in gleichbleibender oder stetig veränderlicher Beschaffenheit längere Zeit verharren (bzw. ihrer Natur nach verharren können), wie das des Regnens, des Rauschens überhaupt, auch alle Geräusche durch Streichen über eine nahezu glatte Unterlage. Dies sind die Geräusche par excellence („Geräusch“ von „Rauschen“). Zu ihnen, und näher zur Gruppe der farblosen Dauergeräusche, gehören die meisten Konsonanten; so die Reibelaute Sch, S, F, Ch, H und die Nasalkonsonanten M, N, Ng. L glauben wir akustisch richtiger zu den Vokalen rechnen, und zwar als nasale Modifikationen des U (stimmhaftes L) Ū oder Ö (stimmloses L) fassen zu müssen<sup>1)</sup>. Funktionell mag es dabei immer zu den Konsonanten gezählt werden, da es wie andere einen Vokal einleitet oder abschließt und zwei Vokale trennt oder verbindet.

Die genannten Unterklassen zeigen nicht nur genetisch, sondern auch akustisch verschiedene Charaktere. Aber kaum wird es gelingen, sie in dieser Hinsicht durch bestimmte Merkmale zu definieren. Wir können nur sagen, welchen Kombinationen äußerer Reize die Eindrücke entsprechen und daß dabei die Gehörsnervenendigungen in Zonen von bestimmter Breite, Tonlage usw. erregt werden. Aber das Phänomen selbst ist immer ein nahezu einfaches, nur in sehr geringem Grade durch das Gehör analysierbar.

Die Reibelaute lassen sich noch in die 3 Unterklassen der Zischlaute (Sch, S), Schiebelaute (Ch) und Blaselaute (F, H) zerlegen, die sich zwar hauptsächlich genetisch, außerdem aber doch auch erscheinungsmäßig genügend voneinander abheben.

Bemerkenswert sind vom phänomenologischen Standpunkte die Reihenbildungen unter den Reibelauten. Sie sind der U—I-Reihe unter den Vokalen analog. So bilden die Zischlaute vom dunkelsten und breitesten Sch bis zum hellsten, dünnsten, schärfsten S, ebenso die Schiebelaute vom dunkelsten, breitesten Ch gutt. bis zum feinsten, hellsten Ch pal. oder alv., nicht minder die Blaselaute unter sich eindimensionale Reihen. Diese Reihen lassen sich aber nach beiden Seiten noch verlängert denken, und es kommen in der Natur dunkle und helle, scharfe und weiche Geräusche aller Arten vor, die über die sprachlich erzeugbaren hinausreichen. Jenes ultrahelle subjektive Zischen, von dem oben S. 138 die Rede war, läßt sich ebenso als Endglied einer Ch- wie einer

<sup>1)</sup> Ähnlich BREMER S. 132 und andere; vgl. o. S. 120ff.



Sch—S-Reihe fassen. Wir selbst können aber auch noch Extreme hervorbringen, die nicht mehr sprachlich verwendbar sind: so das Geräusch beim Blasen durch den von den Lippen ohne Beteiligung der Zähne gebildeten Kanal, wie man auf eine heiße Suppe bläst. Dieses so gut wie gänzlich tonfreie, extrem-weiche Geräusch ist wohl zu schwach, um in der Sprache Verwendung zu finden.

Wie das Rauschen der Zentraltypus der Geräusche überhaupt, so sind die Zischlaute der Zentraltypus der Sprachgeräusche. (Das Sch und Z in den Worten „Rauschen“ und „Zischen“ sind ohne Zweifel auch wieder nachahmender Natur.)

b) Unterbrechungsgeräusche<sup>1)</sup>, vom schärfsten Rattern, z. B. eines schlechten Motorrades, bis zu der nur eben merklichen Rauigkeit beim Verschieben des Fingernagels über eine fein gerippte Fläche. Unter den Konsonanten gehört hierher das R in seinen verschiedenen Formen. Nach der Stärke und Schärfe des Intermittierens kann man eine stetige Reihe von der schärfsten bis zur eben merklichen Rauigkeit herstellen, aber es unterliegt noch Unterschieden je nach der Beschaffenheit des unterbrochenen Geräusches, das meist auch mit Tönen vermischt ist. Auch ein sehr intensives, freilich auch sehr unschönes Nasen-R läßt sich erzeugen, sowie ein intermittierendes stimmhaftes Ng, das ebenfalls in der Nase seinen Hauptsitz hat, aber wohl in keiner Sprache benützt wird.

KOEHLER bemerkt mit Recht, daß es verkehrt wäre, R als eine Summe aufeinanderfolgender Knalle oder Stöße zu definieren. Es ist ein einheitlicher Eindruck. Aber ebenso verkehrt wäre es, zu leugnen, daß dieses Ganze zeitliche Teile hat, die immer wieder neu ansetzend rasch aufeinanderfolgen, und daß dieser Charakter bei darauf gerichteter Aufmerksamkeit leicht wahrgenommen werden kann. Komplexqualitäten mit solchen zeitlich und intensiv differenzierten Teilen gibt es auch sonst genug. Wird ein Ton mit abnehmender Stärke angegeben, sei es, daß wir ihn dabei stetig aushalten oder in kurzen Zeitabständen wiederholen, so konstatieren wir auch nicht eine bloße Summe schwächer werdender Töne, sondern ein „Diminuendo“, das als solches, als Ganzes aufgefaßt wird. Erzeugen wir langsame Schwebungen, die durch stetige Erhöhung des einen Tones immer schneller werden, so unterliegen die Intensitätsschwankungen selbst wieder einer Veränderung höherer Ordnung. Dabei erscheinen sie allerdings von einer gewissen Schnelligkeit an nicht mehr als Intensitätsschwankungen.

<sup>1)</sup> Der Logiker stoße sich nicht daran, daß zwischen dauernden und Augenblicksgeräuschen eigentlich kein Drittes denkbar ist und die Unterbrechungsgeräusche vielmehr nur ein besonderer Fall der Dauergeräusche sind. Wer will, kann die Einteilung leicht danach umformen, indem er die Dauergeräusche in zeitlich gleichförmige und intermittierende zerlegt. Denken wir das 1. Merkmal zu „Dauergeräuschen“ immer hinzu, so ist von vornherein auch logisch alles in Ordnung.

sondern als ein Rollen, Schwirren usw., wie es auch das R ist, weil eben der Übergang zwischen den einzelnen Maximis unmerklich wird und diese unverbunden, diskret aufeinanderfolgen. Wahrscheinlich könnten wir uns keine von diesen Erscheinungen vorstellen, wenn wir immer nur einzelne Töne als solche wahrgenommen hätten. Insofern sind sie also neue Qualitäten. Aber nicht in dem Sinne neu, wie die Empfindungen eines neuen Sinnes es wären, sondern nur eben als Komplexqualitäten innerhalb des Tonsinnes, und zwar als solche, die durch bloße zeitliche und Stärke-Veränderungen unterscheidbarer Bestandteile bedingt sind. Wie das Ganze nicht über den Teilen, so dürfen auch die Teile nicht über dem Ganzen in der Beschreibung vergessen werden.

Vom deutschen R in seiner doppelten, lingualen und gutturalen, Form unterscheidet sich stark das englische „umgekehrte R“ (ɹ), wie in „rice“; das Intermittieren tritt darin außerordentlich zurück (noch mehr wohl im amerikanischen), dagegen ein vokalisches, etwa O-ähnliches Element mehr hervor<sup>1</sup>). Das englische Schluß-R, wie in „her“, ist überhaupt fast unhörbar und mehr eine bestimmte vokalische Abschlußform des vorhergehenden Vokales als ein besonderer Laut<sup>2</sup>).

c) Augenblicksgeräusche, die wieder in allen Intensitätsstufen vorkommen, vom furchtbarsten Krach bis zum leisesten Ticken oder dem „Einlochknall“ O. ABRAHAMS (oben S. 136 Anm.). Ein Ton kann durch extreme Zeitverkürzung in ein solches Geräusch übergehen. Richtiger würde man hier allerdings sagen, er gehe in einen seiner Qualität nach undeutlichen Eindruck über, den man ebensowohl den Geräuschen wie den Tönen beizählen kann; ähnlich wie sogar extrem schwache Temperatur- und Berührungsempfindungen nach WUNDERLIS Versuchen verwechselt werden können. Manche kurze Geräusche, bei denen die Zeitverkürzung nicht so weit getrieben ist, sind als eigentliche Töne anzusprechen, wie z. B. ein Paukenschlag. Bei weniger sonorem Material werden Mischerscheinungen auftreten, die man je nach dem vorwiegenden Bestande da- und dorthin rechnen kann. Aber es gibt auch reine und unzweifelhafte Augenblicksgeräusche.

So sind nun auch Augenblickskonsonanten als Hauptklasse zu statuieren. Unter ihnen bilden die Verschuß-, allgemeiner Explosivlaute, die wichtigste Gruppe. Die Unterklassen, Tenues und Mediae, dürften sich akustisch wesentlich nur durch ihre Stärke unterscheiden. Unseren Analysen (6. Kapitel) sind nur die Tenues zugrunde gelegt, weil sie sich für Interferenzversuche besser eignen; aber das Wesentliche wird auch für die Mediae gelten.

<sup>1</sup>) PAGET und CRANDALL haben es mit ihren Methoden akustisch untersucht.

<sup>2</sup>) Nach SWEET (Handb. of Phonetics) wird es nach einem langen A oder Öe, wie in bar, her, heard, überhaupt von diesen absorbiert.

Einigermaßen stark angegeben, sind die Explosivkonsonanten stets mit schwachen (evtl. geflüsterten) Vokalen verquickt, ähnlich wie die durch äußere Mittel erzeugten kräftigeren Augenblicksgeräusche durch die beigemischten Töne zugleich einen gewissen Vokalcharakter tragen (man kann ein A, E, O klatschen). Je reiner aber das Geräusch ist, wie bei den stimmlosen Mediae, um so weniger ist ein solcher erkennbar.

Besondere Bemerkungen erfordern J und W. Das erste fällt, wenn es rein gesprochen wird (nicht Tja für Ja oder Herrchese für Herr Jesus) zumeist mit einem sehr kurzen I zusammen; dieses bildet mit dem darauffolgenden Vokal einen vokalischen Diphthongen mit Betonung des 2. Bestandteils. Nur nach einem T nimmt es unvermeidlich den Charakter eines ganz schwachen Reibelautes an, wie in dem plattdeutschen Lüttj, Lüttjmann.

Das deutsche W ist, als stimmloser Laut isoliert ausgesprochen, von äußerster Schwachheit, wie das H, und von diesem und einem stumpfen F nicht sehr verschieden. Im Zusammenhange des stimmlosen Sprechens ist es nur eine besondere Einsatzform eines Flüstervokals, verschieden von dem H-Einsatz (der Aspiration), aber ebenso unselbständig. Diese Einsatzverschiedenheiten ließen sich nach dem bloßen Gehör, besser noch mit Unterstützung graphischer Methoden näher beschreiben. Die Lippenbewegungen sind jedenfalls deutlich verschieden. Stimmhaft aber erscheint mir das W als ein ebenso einsetzendes U, das aber sehr rasch in einen anderen Vokal übergeht. Folgt ihm ein U selbst, wie in „Wunde“, so geht der Einsatz direkt in dieses über; die ganze Silbe „Wu“ ist dann eben nur ein eigentümlich einsetzendes U<sup>1)</sup>.

Im englischen W tritt die vokalische Seite noch stärker in die Erscheinung. Im Schweizerdeutschen gehört nach WINTELER W neben L und J überhaupt zu den Vokalen, den „rein tönenden, niemals weichen Lauten“.

In dieser Weise also ordnen sich die als Konsonanten bezeichneten Laute in die Fülle der akustischen Erscheinungen. Es bleibt nur hinzuzufügen, daß die durch die menschlichen Sprachwerkzeuge herstellbaren Konsonantengeräusche, und besonders die der Muttersprache, ebenso wie die Vokale jenen Vertrautheits-

<sup>1)</sup> Nach MICHAELIS (1, S. 29) ist die vokalische Auffassung des J und W auch von KRAÜTER (Saargemünd) vertreten worden. MICHAELIS selbst findet einen leisen Reibelaut in beiden, vermutet aber hier einen Unterschied zwischen Süd- und Norddeutschen. Vgl. auch NADOLECZNY S. 670: „Bei den Lippenlauten ist zu bemerken, daß das W im Süddeutschen zwischen den Lippen, im Norddeutschen zwischen Unterlippe und Oberzähnen gebildet wird.“

charakter angenommen haben, der sie psychologisch von allen übrigen Lauten unterscheidet, aber nicht akustischer Art ist.

Ein Unterschied gegenüber den Vokalen dürfte darin liegen, daß die sprachlich herstellbaren Vokale sich in einem geschlossenen System darstellen lassen, und daß wohl alle in diesem System möglichen Nuancen in irgendwelchen Sprachen oder Dialekten Verwendung finden, während wir noch mancherlei farblose Geräusche durch die Sprachwerkzeuge hervorbringen können, die doch in keiner Sprache als Konsonanten gebraucht werden, wie das vorerwähnte Blasen oder das Grunzen mit geschlossenem Munde.

Unter den älteren Schematismen der Sprachlaute findet sich auch eine geistreiche Anordnung des großen Philologen AUGUST BÖCKH, worin die Konsonanten mit den Vokalen zusammen in einem gemeinschaftlichen Bilde vereinigt werden („Vom Übergange der Buchstaben ineinander“ 1808, eine Jugendarbeit, die er aber später in seine „Gesammelten Werke“ einreichte). Ausgehend von dem an die damalige Naturphilosophie anklingenden Gedanken, das Alphabet als Organismus müsse sich in der für alle Organismen gültigen Form einer in sich zurücklaufenden Linie darstellen lassen, ordnet er die Buchstaben in eine Kreislinie. Die Anordnung der 5 „Hauptvokale“ ist dieselbe wie im Vokaldreieck: U und I stehen an den Endpunkten des horizontalen Durchmessers, A am unteren Ende des vertikalen (auch HELLWAG hatte seinem Dreieck diese umgekehrte Form gegeben). An U und I schließen sich dann nach oben hin die nächstverwandten Konsonanten W und J an. Dies läßt sich noch ganz wohl rechtfertigen. Aber gegenüber A steht am obersten Punkte des Kreises — C, ein Buchstabe, dem nachgerühmt wird, daß er in zweifacher Form (als K und Ts) ausgesprochen werde und sich daher als Mitte eigne. Wenn die Buchstaben nach den dadurch bezeichneten Lauten eingeteilt werden sollen, wie es doch hier der Fall ist, dürfte dieser überhaupt nicht besonders aufgeführt werden. Die Konsonanten werden in folgender Anordnung eingefügt: W, P, T, K, C, S, R, L, N, M, J. Man erkennt an diesem Versuch, namentlich an der Einreihung des R, nur wieder, daß es unmöglich ist, die Konsonanten ebenso wie die Vokale auf einen gemeinsamen Nenner (Helligkeit) zu bringen, sondern daß sie nur durch mehrere Einteilungsgründe umspannt werden können.

## 11. Kapitel.

# Einheitliches und mehrheitliches Hören.

### 1. Allgemeinere Fragestellungen.

In den vorangehenden Teilen dieses Buches suchten wir die objektiven Komponenten der Sprachlaute rein empirisch festzustellen, ohne uns darum zu kümmern, wie es kommt, daß diese Reize in ihrem Zusammenwirken gerade diese und keine anderen Laute für unsere Wahrnehmung hervorbringen. Diese prinzipielleren und tiefergreifenden theoretischen Fragen, die den Psychologen und Gehirnphysiologen am meisten an der ganzen Sache interessieren müssen, sollen uns jetzt noch beschäftigen. Aber wir werden sie nicht restlos lösen können, sondern uns zum großen Teil nur auf Anfänge experimenteller Forschung und auf Hypothesen beschränken müssen.

Hauptsächlich 3 Fragen sind zu beantworten:

1. Unter welchen Bedingungen hört man die objektiv gegebene Summe der Komponenten subjektiv als eine Einheit, wie dies bei den Sprachlauten der Fall ist?

2. In welcher Weise beeinflussen sich gleichzeitige Töne gegenseitig in Hinsicht ihrer subjektiven Stärke, und wie verhält sich die Gesamtstärke zur Stärke der einzelnen heraushörbaren Töne?

3. Wie verhält sich die spezifische Qualität eines Vokals zu den Eigenschaften der Teiltöne, aus denen er besteht?

Analoge Fragen tauchen auch für die Konsonanten auf. Doch ist da die Fülle der Spezialprobleme wie der verfügbaren Tatsachen weit geringer.

Zu Versuchen in diesen 3 Richtungen sind einfache Töne wieder das beste, in den meisten Beziehungen sogar das einzig zuverlässige Material; weshalb unser von allen Obertönen gereinigtes Pfeifensystem auch hierfür gute Dienste tut.

Wir beginnen mit der ersten dieser Fragen, für die noch am meisten Erfahrungen vorliegen, schicken aber einige allgemeinere Bemerkungen voraus, um die Grundbegriffe und deren Bezeichnungen festzulegen.

## 2. Komplex- und Gestalteigenschaften.

Wir sprechen in der Beschreibung der Erscheinungen von einem Komplex überall da, wo ein Ganzes von absoluten Inhalten gegeben ist, während von einer Gestalt im weitesten Sinne gesprochen wird, wo ein Ganzes von Verhältnissen vorliegt. Das Gesichtsfeld in einem bestimmten Augenblick, ein individueller Klang oder Mehrklang das Ganze der augenblicklich überhaupt einem Individuum gegebenen Empfindungen aller Sinne (Körpergefühl, Gemeingefühl) sind Beispiele von Komplexen. Melodien, Akkorde, Rhythmen aller Sinne, Raumformen sind Beispiele von Gestalten. Sowohl Komplexe als Gestalten können gegliedert oder ungliedert wahrgenommen werden, d. h. mit oder ohne Unterscheidung der Teile. So wird ein Dreiklang von dem einen Hörer gegliedert, von dem anderen ungliedert erfaßt. Jeder gegliederte Komplex hat eine wahrnehmbare Gestalt, d. h. es bestehen zwischen seinen Teilen bestimmte Verhältnisse, die ein Ganzes bilden. Gestalten haben die Eigentümlichkeit, auf Komplexe aus anderen absoluten Inhalten gleicher Gattung übertragbar zu sein. Verhältnisse sind eben nicht an bestimmte absolute Inhalte gebunden, und so gilt dies auch für Verhältnisganze. So können Melodien, Akkorde transponiert, geometrische Gestalten in beliebigen Größen vorgestellt werden. Dagegen wird ein Komplex als solcher durch jede Veränderung auch nur eines seiner Teile mehr oder minder verändert. Der Klarinettenklang oder der Vokal O ist ein anderer auf *c* und auf *cis*, wenn wir sie auch infolge bestimmter gleichbleibender Eigenschaften beide Male unter dieselbe Klangspezies ordnen. Auch der Mehrklang *ceg* ist als Klang ein anderer als *d fis a*, obgleich sie mit Rücksicht auf bestimmte sinnliche Wirkungen, im Falle des gegliederten Vorstellens auch infolge gleicher Gestalt, als ein identischer Akkord (Durdreiklang) bezeichnet werden.

Jedem unzergliederten Komplex kommen als solchem gewisse Eigenschaften zu: Komplexeigenschaften. So ist ein Klang weich, hart, rau, voll, leer, ein Vokal dunkel, hell. Alle Grundeigenschaften, durch die wir die letzten Elemente der Erscheinungen, in die ein Komplex aufgelöst werden kann, charakterisieren, z. B. Qualität, Stärke, Helligkeit, finden sich auch in den Komplexen wieder, aber sie bedürfen hier besonderer Feststellungen.

Ebenso haben aber auch unzergliedert wahrgenommene Gestalten bestimmte Eigenschaften: Gestalteigenschaften. Ein Rechteck kann schmal oder breit, schlank oder plump, eine Melodie von verschiedenem Tonumfang (Ambitus), raschbewegt oder schwerfällig sein.

Daß niemals zwei Erscheinungen oder Funktionen im normalen Bewußtsein zusammen gegeben sind, ohne durch irgendeine, sei es auch nur sehr lose, Beziehung zu einem Ganzen verknüpft zu sein, und daß dieses Ganze nicht identisch ist mit der bloßen Summe seiner Teile, ist eine alte Erkenntnis. Aber seit den letzten Dezennien sind die Eigenschaften, die einem Ganzen als solchem zukommen, immer mehr beachtet und diskutiert worden. F. KRUEGER hat wohl zuerst den Ausdruck „Komplexqualität“ eingeführt (Wundts Psychol. Studien Bd. 2, S. 221. 1906), v. EHRENFELS den Ausdruck „Gestaltqualität“ (Vierteljahrsschr. f. wiss. Philosophie Bd. 14, S. 249. 1890). Beide Forscher knüpften daran Betrachtungen, von denen namentlich die über Gestaltqualitäten folgenreich geworden sind. Die im Text vorgetragene Abgrenzung zwischen dem Begriff des Komplexes und dem der Gestalt ist in dieser Form bisher meines Wissens nicht vorgenommen worden, scheint mir aber im Interesse der Klarheit notwendig zu sein. Ich würde daher beispielsweise Klangfarben nicht (mit v. EHRENFELS und CORNELIUS) als Gestaltqualitäten, ebensowenig (mit KRUEGER) Konsonanz und Dissonanz als Komplexqualitäten definieren. Wir setzen ferner für „Qualitäten“ hier den allgemeineren und farbloseren Ausdruck „Eigenschaften“, da unter den Eigenschaften eines Klanges, auch denen eines Vokals, doch wieder seine Qualität als eine neben anderen Eigenschaften unterschieden werden muß. In die stark angeschwollene sachliche Diskussion treten wir hier nur insoweit ein, als es die vorliegende Angelegenheit erfordert.

Viele Bemerkungen über Komplexeigenschaften (Gesamthelligkeit, Gesamtstärke) findet man neuerdings bei WERNER.

Vokale, wie überhaupt Klänge mit Obertönen, sind klassische Beispiele ungliederter Komplexe im gewöhnlichen Wahrnehmungsleben, Klangfarben klassische Beispiele von Komplexeigenschaften. Denn wenn durch eine Summe objektiv gegebener Töne ein Vokal für das Gehör entstehen soll, so ist eine Voraussetzung dafür, daß die Komponenten nicht als einzelne Töne für sich bemerkt, sondern nur als ein undifferenziertes Ganzes gehört werden. Sobald ich auch nur einen einzigen Ton heraushöre, ist der Gesamteindruck als solcher alteriert. Man hört das Ganze zwar neben dem Teilton fortklingen, und wenn dieser nur ein schwacher und unwesentlicher Klangteil ist, so wird die Klangfarbe auch nicht erheblich verändert. Ist es aber ein stärkerer, oder sind es gar mehrere starke Töne, so ist die Klangfarbe als solche zerstört. Auch die Formantttöne der Vokale entfalten ihre volle Wirkung nur unter der Bedingung, daß sie nicht für sich wahrgenommen werden, und nur solange dies der Fall ist<sup>1)</sup>. Er-

<sup>1)</sup> KOEHLER sagt allerdings (1, Vorl. Mitt. S. 101): Durch Heraushören der Formantttöne „verändert sich der nebenbei gehörte Gesamtklang des gesungenen Vokals gar nicht sehr deutlich; sondern neben oder über ihm, gewissermaßen her austauchend, erklingt nur auch noch der betreffende Teilton“. Aber dies dürfte nur für den ersten Augenblick zutreffen. Bei längerer Dauer der Beobachtung und konzentriertem Beachten der Teil-töne wird der Vokal für mich alteriert oder ganz zerstört. (Vgl. Tonpsych. II, S. 529 und MACH, daselbst.)

staunlich bleibt es immer, daß die objektive Tonsumme  $c^2 + g^2 + c^3 + e^3$  in einem bestimmten Stärkeverhältnis ihrer Glieder als Vokal A gehört wird und daß objektiv hierbei wirklich nichts weiter als eben diese Summe gegeben ist. Sie wird ja auch tatsächlich nicht immer als A gehört, sondern unter Umständen als ein harmonischer Akkord (Konkord) aus ungleich starken Tönen. Wann also ist das eine, wann das andere der Fall? Mit der Beantwortung dieser Frage ist zwar das Rätsel noch keineswegs gelöst, aber es sind wenigstens die Bedingungen, unter denen erfahrungsgemäß die Erscheinung auftritt, näher beschrieben.

Wir lassen dabei die Frage auf sich beruhen, ob beim einheitlichen (ungegliederten) Hören einer objektiven Tonmehrheit die Sinnesempfindung selbst eine ungeteilte Einheit ist, oder ob sie nur der Auffassung als Einheit erscheint. Die Unterscheidung der Empfindung von ihrer Auffassung, der Perzeption von der Apperzeption, die ich mit HELMHOLTZ teile, gilt manchen Neueren als überflüssig und verkehrt. Wir brauchen hier nicht darauf einzugehen. Wenn wir vom ungegliederten, einheitlichen Hören sprechen, mag es von jedem nach seinem allgemeinen psychologischen Standpunkt ausgelegt werden.

Man könnte vielleicht einwenden, schon unsere Fragestellung sei überflüssig und das ganze Problem nur scheinbar: denn auch objektiv liege keine bloße Summe vor, da die Sinusschwingungen nicht nebeneinander verliefen, sondern in Wirklichkeit eine einheitliche nicht-sinusförmige Welle bildeten, aus der sie nur als Kunstprodukte der mathematischen Analyse erschlossen würden. Aber in unserer synthetischen Einrichtung sind sie doch tatsächlich und real voneinander getrennt. Im Trichter oder im Gehörgang werden sie zwar auch hier zu einer Gesamtwelle vereinigt, in der Schnecke aber nach HELMHOLTZ doch wieder gesondert; ebenso wie die natürlichen Vokale, die vom Munde bis dahin unzerlegt verlaufen, hier zerlegt werden. Und wenn man auch HELMHOLTZ' Theorie nur als Hypothese oder als Bild gelten lassen will: daß irgendwo und irgendwie eine physiologische Zerlegung des akustischen Prozesses stattfinden muß, ist ein nicht zu umgehendes Postulat. Man kann also sagen, daß der letzte Zustand auf der physischen Seite, von dem wir bisher einigermaßen eine erfahrungsmäßige Kenntnis haben, der des getrennten Verlaufes sei.

Wie ist es aber auf dem psychologischen (phänomenalen) bzw. gehirnphysiologischen Gebiete? Hier ist nicht etwa erfahrungsgemäß die mehrheitliche, gegliederte Auffassung des Gehörten das Ursprüngliche in der individuellen Entwicklung, sondern ohne Zweifel die einheitliche. Es wäre eine verkehrte und längst veraltete Ansicht von dem seelischen Entwicklungsgange, wollte man voraussetzen, daß die Teilempfindungen, z. B. die den Sinusschwingungen entsprechenden einfachen Töne, zuerst alle atomistisch, als diskrete Einheiten, nebeneinander dem Bewußtsein vorschwebten und erst allmählich zu Ganzheiten zusammenwüchsen. Das Gegenteil ist offenbar: Erfahrung, Übung, Aufmerksamkeit bewirken in erster Linie zunehmende Gliederung der Eindrücke. Erst später unterliegt auch die Zusammenfassung und Vereinheitlichung psychischen Einflüssen



(vgl. Tonps. II, S. 69ff.). Auch gelangt das mehrheitliche Hören niemals dazu, sämtliche Bestandteile oder auch nur alle stärkeren und unter sich gleich starken Bestandteile eines hochzusammengesetzten Klanges gleich deutlich nebeneinander zu hören. Es treten immer nur einzelne oder wenige Teile in solcher Weise auf einmal hervor. Selbst von extrem musikalischen Menschen werden kaum mehr als 4 Töne, beim Hören von Obertönen nur 1–2, in streng gleichzeitiger Sonderung aufgefaßt, derart also, daß die Aufmerksamkeit allen zugleich zugewandt ist. Aber ein sehr rascher Wechsel der Aufmerksamkeit zwischen den verschiedenen Stimmen und ein leichtes und sicheres Erfassen jeder beliebigen Stimme ist damit verbunden und für das im höheren Sinne musikalische Hören charakteristisch.

Alle diese Einschränkungen und Vorbehalte lassen die Tatsache unberührt, daß der erwachsene Mensch, sei er musikalisch oder nicht, den gleichen objektiv gegebenen Bestand an sinusförmigen Komponenten einmal einheitlich, einmal mehrheitlich hört. Gerade unsere synthetischen Versuche stellten diesen Unterschied immer wieder vor Augen. Jeder Beobachter ist überrascht, wenn er den künstlichen Vokal gehört hat und nun durch Einstecken eines Höröhrchens in die einzelnen Leitungen sich überzeugt, daß eine Anzahl von Tönen daran beteiligt und daß nichts weiter erforderlich ist, als sie wieder zu vereinigen. Aber schon der Umstand, daß langes Ausprobieren der Mischungen und Stärkeverhältnisse meinerseits vorausgehen mußte, bis einheitliche Eindrücke zustande kamen, beweist, daß hier ganz besondere Bedingungen zusammenkommen müssen. Namentlich bei den hellen Vokalen, die aus zwei durch eine Lücke getrennten Klangmassen bestehen, kommt es darauf an, diese Dualität unmerklich zu machen. Sonst hört man z. B. beim I ein U + I oder einen Klang, den man sowohl (mit Ignorierung des I) als U wie auch ebensogut (mit Ignorierung des U) als I deuten kann.

Die obige Frage verlangt also zweifellos eine Beantwortung.

### 3. Günstige Bedingungen für das einheitliche Hören.

In des Verfassers „Tonpsychologie“ sind ausführlich die Bedingungen besprochen, von denen die Wahrnehmung der Teile in einem Tonganzen abhängt<sup>1)</sup>. Hier sollen nur die besonderen Umstände erwähnt werden, die die Analyse beeinträchtigen und das synthetische Hören begünstigen, wenn ein Klang aus har-

<sup>1)</sup> Bd. II, § 21–23. Einzelne der dort erwähnten Bedingungen wurden später ausführlicher untersucht, so der Einfluß der Zeitdauer (Zeitschr. f. Psych. Bd. 17, Beitr. z. Akust. Heft 4) und die Verteilung an beide Ohren (das. Bd. 70, Beitr. z. Ak. H. 8).

Nicht ohne Interesse war es mir bei den synthetischen Vokalversuchen, als ich die höchsten Pfeifen auf ihre Obertöne untersuchte, um diese durch Interferenz auszuschließen, daß ich sogar in diesen Klängen der 4- und 5-gestrichenen Oktave noch Grundton und Oktave deutlich durch das bloße Ohr auseinanderzuhalten lernte. Dies ist vielleicht der schwierigste Fall der Analyse eines Zweiklanges durch das bloße Ohr.

Über den Unterschied des Gehörsinnes vom Gesichtssinn in Hinsicht des analysierenden Verhaltens vgl. meine Darlegungen Zeitschr. f. Psych. Bd. 75, S. 330ff. 1916, Beitr. z. Ak. H. 9, S. 17ff. und „Die Attribute der Gesichtsempfindungen“, Abh. d. Berliner Akademie v. J. 1917, S. 50ff.

monischen Teiltönen (ganzahligen Multiplis des Grundtones) vorliegt, wie es bei den Vokalen und den gewöhnlichen Instrumentalclängen der Fall ist. Hier tragen zur Einheitlichkeit des Gesamtklanges besonders bei:

a) Die Verschmelzung der konsonierenden, besonders der 6 tiefsten, Teiltöne.

Es ist bekannt, daß die Töne einer Oktave im Zusammenklang sich dem Eindruck eines einzigen Tones nähern. Von Unmusikalischen werden sie darum tatsächlich in den meisten Fällen mit einem wirklichen Einklang verwechselt; aber auch in der Musik gilt das Singen und Spielen in Oktavengängen noch als „Unisono“. Diese Erscheinung kann nicht aus dem Zusammenfallen des höheren Tones mit dem 1. Oberton des tieferen erklärt werden, schon darum nicht, weil sie auch bei völlig einfachen Tönen auftritt; sie muß vielmehr auf noch unbekanntem physiologischen Gesetzmäßigkeiten beruhen. Sie findet sich aber, wenngleich in abnehmendem Grade, auch bei Quinten, Quarten und überhaupt allen konsonanten Intervallen. Aber auch Intervalle wie 4:7, 5:7 machen noch einen merklich einheitlichen Eindruck. Bei den Oktavenerweiterungen der einzelnen Intervalle kehren ähnliche Unterschiede wieder: die Duodezime verhält sich ihrer Verschmelzung nach zur Dezime wie die Quinte zur Terz usw.

Ich habe die Tatsachen der Verschmelzung eingehender verfolgt und glaube darin das Wesen der musikalischen Konsonanz zu finden. Es scheint mir jetzt, daß sie nicht das eigentlich definierende Merkmal bilden, sondern nur ein „Proprium“ im Sinne der alten Logik, d. h. ein die Konsonanz regelmäßig begleitendes, auch in den graduellen Abstufungen (wenigstens innerhalb der nämlichen Oktave) damit parallel laufendes Merkmal. Aber diese Frage kann hier auf sich beruhen.

In einem Klang aus harmonischen Teiltönen sind nun bis zum 6. oder 8. die sämtlichen Verschmelzungsstufen gleichzeitig verwirklicht, und auch unter den höheren Teiltönen finden sich vielfach Oktaven, Quinten, Terzen usw. wieder. Sogleich die stärkste Verschmelzung findet zwischen den zwei ersten Teiltönen statt, dann folgen die übrigen Stufen, teils zwischen unmittelbar benachbarten, teils zwischen entfernteren Gliedern. Freilich kommen bei den hellen Vokalen höhere Teiltöne und darunter auch dissonante hinzu, weshalb denn auch namentlich bei E und I Unter- und Oberformant nicht so schwer wie etwa die Bestandteile des A durch das bloße Gehör auseinanderzuhalten sind. Doch kommen bei obigen Vokalen andere Faktoren der Vereinheitlichung zu Hilfe (vgl. unten S. 288), so daß im gewöhnlichen Leben die Zwiespältigkeit nicht auffällt.

So beruht die Einheitlichung eines Vokals in erster Linie auf dieser seiner Zusammensetzung aus harmonischen, insbesondere konsonierenden Teiltönen.

b) Die reine Stimmung aller harmonischen Teiltöne. Verschmelzung tritt auch bei nur annähernd reiner Stimmung der bezüglichen Intervalle auf. Aber die vollkommen reine Stimmung aller harmonischen Teiltöne eines Klanges hat noch besonders vereinheitlichende Wirkungen. Wie schon hervorgehoben wurde, fallen hier die Differenztöne stets mit tieferliegenden Gliedern der Reihe zusammen oder rufen solche, wenn sie objektiv fehlen, subjektiv hervor. Die Folge ist eine stärkere Vereinheitlichung des Ganzen gegenüber einer Reihe mit unharmonischen Gliedern, in welcher nur weitere nicht verschmelzende Glieder hinzukommen<sup>1)</sup>. Zugleich werden die Schwebungen, die etwa ein mittelbares Mehrheitsurteil hervorrufen könnten, auf das geringste mögliche Maß reduziert.

c) Die identische Örtlichkeit, aus der die Teiltöne kommen und in der sie dem Hörenden erscheinen.

Sopran und Alt, 1. und 2. Geige des Orchesters heben sich infolge ihrer räumlichen Verteilung auf dem Podium auch für das Gehör deutlicher voneinander ab und werden auch von Unmusikalischen leichter auseinandergelassen, als wenn die Sänger und Spieler alle bunt durcheinandergemischt wären<sup>2)</sup>. Und so ist überhaupt räumliche Verteilung der Klangquellen dem analytischen, räumliche Identität dem synthetischen Verhalten günstig. Vielleicht hilft schon die Mitwirkung des Gesichtseindrucks; doch liegt ja in solchen Fällen auch immer in gewissem Maße eine ungleiche Verteilung auf beide Ohren vor, wobei physikalisch sowohl Phasen-, als Intensitäts-, als Zeitdifferenzen auftreten. Welche von diesen für die Lokalisation maßgebend ist, war lange strittig, ist aber jetzt meines Erachtens zugunsten der Zeitdifferenz entschieden<sup>3)</sup>. Doch wirken diese minimalen Zeitunterschiede des Eintreffens in beiden Ohren, die bis zu 0,015 Sek.

<sup>1)</sup> Ich machte bereits I, Bd. II, S. 330 auf die anscheinende Paradoxie aufmerksam, daß durch Vermehrung der Töne eines Zusammenklanges unter Umständen der einheitliche Eindruck erhöht werden kann: wenn nämlich stärker verschmelzende Töne, besonders tiefere Oktaven, hinzutreten. Dies ist der Fall bei den Differenztönen harmonischer Teiltöne.

<sup>2)</sup> Extrem Unmusikalische, die sonst überhaupt Mehrklänge fast immer nur einheitlich hören, können durch Verteilung zweier genügend verschiedenen Stimmgabeln an beide Ohren zur Analyse gebracht werden. Tonpsych. II, S. 363.

<sup>3)</sup> v. HORNBOSTEL und WERTHEIMER, Sitzungsber. d. Berliner Akad. d. Wiss. 1920, S. 388. v. HORNBOSTEL (3).

heruntergehen können, nicht als wahrgenommene Bewußtseinsinhalte, sondern nur als physiologische Reize für das Auftreten von Lokalzeichen, auf denen dann die objektive Lokalisation beruht; und wenn genau gleichzeitig einsetzende Tonquellen bei räumlicher Verteilung besser auseinandergehalten werden, so ist es nicht jene minimale Zeitdifferenz im Eintreffen der Schalle an sich, sondern die dadurch hervorgerufene verschiedene Lokalisation, die die Analyse begünstigt.

Das Vorhandensein angeborener Lokalzeichen  $p$ ,  $q$  für beide Ohren geht zwingend daraus hervor, daß man eine von rückwärts an das rechte oder linke Ohr gehaltene Stimmgabel ohne jede Kopfwendung unfehlbar richtig lokalisiert und 2 gleichzeitig vor beide Ohren gehaltene Gabeln von verschiedener Höhe richtig auf beide Ohren verteilt. Im letzteren Fall ist natürlich vorausgesetzt, daß man die Tonhöhen selbst deutlich auseinanderhalten und die höhere als höhere, die tiefere als tiefere sicher beurteilen kann (Tonpsych. II, S. 52ff.). Mehreren Beobachtern und mir gelang es sogar, bis zu 10 genau gleichzeitige einfache Töne, die in den verschiedensten Kombinationen auf beide Ohren verteilt wurden, in wenigen Minuten richtig zu lokalisieren, wobei Drehungen des Kopfes durch einen Kopfhalter ausgeschlossen waren (BALEY, Zschr. f. Psych. Bd. 70; m. Beitr. z. Akust. H. 8). Die Lokalzeichen müssen den Tonempfindungen von Anfang an attributiv eigen sein; nur ihre Beziehung auf beide Ohren, d. h. ihre Assoziation mit den betreffenden Tast- und Gesichtsempfindungen ist individuell erworben, desgleichen ihre Verwertung für die Erkenntnis der Lage äußerer Klangquellen im Gesichtsraum.

Man kann fragen, ob nicht sogar beim einohrigen Hören gröbere lokale Verschiedenheiten in der Provenienz der Klangquellen das mehrheitliche Hören begünstigen. Hält man in unserer synthetischen Einrichtung nach Abnahme des Trichters  $T$  das Ohr direkt vor die nebeneinander in konzentrischen Kreisen angeordneten Öffnungen der Leitungen für sämtliche Teiltöne, so kommt der einheitliche Eindruck schwerer zustande, nicht bloß weil jede kleine Verschiedenheit in der Haltung des Ohres Intensitätsverschiebungen unter den Teiltönen hervorruft und die richtige Stellung schwer zu finden ist, sondern wohl auch darum, weil sich die kleinen Richtungsverschiedenheiten der Schallstrahlen geltend machen. Darum ist es wesentlich, diese wieder in einer Schlauchleitung gesammelt dem Ohre zuzuführen. Wie allerdings Richtungsverschiedenheiten sich monotonisch geltend machen sollen, ist nicht leicht zu sagen. Ich möchte die Tatsache selbst auch nicht als sicher hinstellen.

Je feiner nun verschiedene Lokalisation der Klangkomponenten sich im Bewußtsein geltend macht und je stärker sie das mehrheitliche Hören begünstigt, um so besser ist für das einheitliche Hören gesorgt, wenn, wie bei den gewöhnlichen Instrumentalklängen und den Vokalen, sämtlichen Teiltönen die gleiche Lokalisation zukommt.

Doch ist nicht etwa gleiche Lokalisation in allen Fällen unbedingtes Erfordernis. Wenn man die einen künstlichen Vokal bildenden Töne, auch die des Formanten, auf das rechte und linke

Ohr verteilt (dichotisches Hören), so können unter besonderen Umständen die entsprechenden Vokale gehört werden. Dabei scheint aber vorausgesetzt, daß sie annähernd schon in jedem Ohre für sich allein zustande kommen. SCHOLE bediente sich eines „Doppelresonators“, d. h. zweier verschieden abgestimmter an beide Ohren verteilter Resonatoren, durch die also verschiedene Teiltöne rechts und links verstärkt wurden. Wurde nun U auf  $G = 96$  Schw. gesungen und der eine Resonator auf  $b^1$ , der andere auf  $d^2$  gestimmt, so wurde ein reines O gehört. Wurde auf der einen Seite  $f^2$ , auf der anderen  $h^2$  verstärkt, so hörte man ein Ao. Die Versuche sind aber insofern nicht rein genug, als die übrigen Teiltöne des gesungenen U in nicht genau übersehbarer Weise mitwirkten.

Ich habe solche Versuche durch getrennte Zuleitung der Töne des synthetischen Apparates zu beiden Ohren angestellt. Die 4 Töne  $c^2 g^2 c^3 e^3$  wurden in den für ein A auf  $c^1$  sonst erforderlichen Stärkeabstufungen mittels zweier in die Öffnungen bei  $T$  gesteckten T-Röhrchen auf das rechte und linke Ohr verteilt, z. B.  $c^2 c^3$  rechts,  $g^2 e^3$  links; oder  $c^2 e^3$  rechts,  $g^2 c^3$  links. Hierbei war nun freilich der D. T.  $c^1$  höchstens auf der linken Seite zu erwarten, da bei dichotischer Darbietung Differenztöne im allgemeinen nicht zustande kommen (vgl. Tonps. II, S. 496). Für mich war aber überhaupt ein einheitlicher Eindruck so nicht zu erzielen. Dagegen gelang der Versuch mit 2 unmusikalischen Beobachtern, Dr. v. ALLESCH und O. KAUMANN. v. A. hörte in beiden Fällen dichotisch ein A, er hörte aber ein solches auch auf jedem Ohr allein, wengleich etwas anders, auf dem einen heller, auf dem anderen dunkler. Das beidohrige A war „voller, reicher, aber nicht so glatt und einheitlich“. Ähnlich drückte sich K. aus. Auch aus diesen Versuchen schließe ich, daß bei dichotischer Zuleitung nur dann der nämliche Vokal entsteht wie bei gewöhnlichem beidohrigen oder einohrigen Hören, wenn er schon wenigstens auf einem Ohre annähernd diesen Vokalcharakter besitzt. Daß aber für die eben erwähnten Personen der A-Charakter schon durch je 2 Töne allein zustande kommen konnte, verstehen wir aus den S. 195 erwähnten Tatsachen.

#### d) Die genaue Gleichzeitigkeit.

Alle Teiltöne eines Instrumental- oder Vokalklanges beginnen und schließen im allgemeinen streng gleichzeitig, wenn auch im einzelnen Fall Ausnahmen vorkommen. Ist bei einer Synthese zuerst nur 1 Ton vorhanden, zu dem dann ein 2., 3. usf. hinzutreten, so unterscheidet auch der Ungeübte die neuen weit leichter von dem vorherigen Bestand, zumal wenn er die sukzessive ein-

tretenden Töne kurz vorher auch schon einzeln gehört hat und sie jetzt zu hören erwartet. Wenn man an der synthetischen Einrichtung die einfachen Teiltöne in maximaler Stärke vom tiefsten bis zum höchsten fortschreitend zusammenfügt, so hört man jeden Ton bis zum letzten und höchsten bei seinem Auftreten deutlich heraus; während es sehr schwer oder unmöglich wäre, in dem fertigen Ganzen jeden einzelnen Ton zu erkennen. Bei meinen synthetischen Versuchen war es für die Vereinheitlichung ungünstig, daß vorher schon bekannte Teiltöne zum Ausprobieren sukzessive hinzugefügt werden mußten. Daher war der drehbare Hahn außerordentlich nützlich, durch den der Gesamteindruck zeitlich scharf abgeschnitten zum Gehör kam.

Die Töne  $c^1 c^2 g^2 c^3 e^3 g^3 b^3 c^4$  (harmonische Teiltöne des  $c^1$ -Klages), in voller Stärke zusammen angegeben, liefern einen sehr einheitlichen zungenartigen Klang. Schaltet man nun  $c^1$  abwechselnd aus und wieder ein, so hört man es als einfachen Ton hinzutreten, und es bleibt noch eine kurze Weile als solcher isoliert neben dem Gesamtklange hörbar, bis es gleichsam verschlungen und assimiliert wird. Ebenso  $c^2$ . Der scharf näselnde Klang bleibt dabei immer auf dem Grundtone  $c^1$  liegen, da dieser als Differenzton erhalten bleibt, wenn er auch wohl beim Ausfall eines jener Töne schwächer wird. Wird derselbe Versuch mit  $g^2$  gemacht, so klingt das Ganze beim Wegfall dieses Tones weit weniger scharf, wie wenn es aus der Ferne käme; aber es bleibt (oder wird wieder) einheitlich. Werden  $g^2$  und  $c^3$  zugleich herausgenommen, so ist die Veränderung noch stärker. Aber man entwöhnt sich durch diese Versuche schnell des einheitlichen Hörens. Die wieder eingefügten Töne bleiben immer länger isoliert.

Eine hierhergehörige Erfahrung beschreibt auch schon HELMHOLTZ: als er zu der Stimmgabel  $b$  die Gabel  $b^1$  fügte, blieben die Töne zunächst noch getrennt und erschien die tiefe Gabel als U-ähnlich, dann aber trat eine Vereinheitlichung ein, und der Charakter des Gesamtklages war der eines O (vgl. dazu Tonpsych. II, S. 352).

e) Günstige Intensitätsverhältnisse unter den Teiltönen.

Ein einzelner stärkerer Teilton (besonders ein dissonanter), der nicht von höheren und tieferen in abnehmender Stärke umgeben ist, drängt sich dem Gehör leicht als Einzelton auf. Anders ist es bei den Vokalen. Die Stärken nehmen von Teilton zu Teilton in der Regel langsam ab oder zu, was in der Resonanzbreite der Rachenhöhle und ihrer Abteilungen seinen Grund hat. Diese Art der Stärkeverteilung trägt wesentlich zu dem einheitlichen Ein-

druck bei. Aber auch im übrigen scheinen die Intensitätsverhältnisse bei den Vokalen in einer für das einheitliche Hören besonders günstigen Weise gegeneinander abgestuft.

JAENSCH macht neuerdings gegen die Beweiskraft der HELMHOLTZschen Gabelsynthese, aber auch gegen die unserer Pfeifensynthese und gegen die Definition der Vokale als Klänge geltend, daß sich das Schallphänomen mit der Stellung des Beobachters ändere: „Im allgemeinen muß man eine ganz bestimmte Stellung einnehmen, um das Vokalphänomen in guter Reinheit zu erhalten, vorausgesetzt daß die Gabeln oder Pfeifen frei im Versuchsraum aufgestellt sind und der Schall nicht etwa durch Schläuche zusammengeleitet dem Ohre des Beobachters zugeführt wird“ (3 S. 99). „Hier vollzieht sich nun der Übergang von klangartigem zu vokalartigem Schallphänomen, ohne daß an den Tonquellen selbst und ihrem Funktionszustande das Geringste geändert würde“ (S. 109).

Gewiß, an den Tonquellen braucht dabei nichts geändert zu werden. Aber ihr „Funktionszustand“ ändert sich gar sehr, denn sie kommen bei jeder Veränderung in der Stellung des Beobachters mit anderen Intensitätsverhältnissen in seine Ohren. Auch dem Verfasser ist es schon zu Beginn seiner Synthesen öfters begegnet, daß im Schallzimmer und in unmittelbarer Nähe der Pfeifen durch eine zufällige Drehung oder ein Bücken des Kopfes ganz plötzlich ein einheitlicher (damals ein instrumentaler) Eindruck auftauchte, während vorher nur eine Summe einzelner Töne gehört wurde. Daran ist nicht ein geheimnisvoller Übergang zu einer neuen Schallgattung schuld, sondern einfach der zufällige Eintritt der für einheitliches Hören günstigen Intensitätsverhältnisse, wobei selbst die Reflexionen von den Zimmerwänden und die von der Kopfstellung in hohem Maße abhängigen Stärkeverschiebungen der Differenztöne (vgl. m. Abh. 7, S. 99) von Einfluß sein müssen. Um nicht von solchen Zufälligkeiten abhängig zu sein und die erforderlichen Intensitätsverhältnisse selbst herstellen zu können, wurden die Töne eben in den Regulierapparat geleitet und aus dem zusammenfassenden Trichter und Schlauch abgehört.

f) Eine besondere physiologische Einstellung.

Zu leugnen ist aber nicht, daß auch bei völlig unveränderter Schallgebung und Schallzuleitung sich noch innere, und zwar zunächst rein physiologische Einflüsse geltend machen können. Es ist mir bei den synthetischen Versuchen oft aufgefallen, daß ein bis dahin nur mehrheitlich gehörter Komplex plötzlich zu einem einheitlichen Klangeindruck verschmolz, ohne daß äußere

Veränderungen, ja auch ohne daß Einflüsse der Willens- und Aufmerksamkeitserichtung stattgefunden hatten. Solche Erfahrungen erwecken den Eindruck, daß der Übergang vom mehrheitlichen zum einheitlichen Hören (ebenso natürlich der umgekehrte) durch einen bestimmten zentralphysiologischen Vorgang vermittelt wird, über den sich freilich zur Zeit nichts weiter sagen läßt.

Auch von anderen ist diese synthetische Einstellung schon bemerkt worden; so z. B. nach mündlicher Mitteilung von MAX PLANCK, der den Fall vergleicht mit dem Zusammenschmelzen und Plastischwerden der beiden Eindrücke beim Stereoskopieren. Auch JAENSCH hat bei seinen Versuchen zur Vokalnachbildung solche Beobachtungen gemacht. Neuerdings hat sich H. WERNER bei seinen Untersuchungen zur Intensitätspsychologie vielfach auf diesen Unterschied der analytischen und synthetischen Einstellung geführt gesehen. Er betont auch das Vorhandensein individueller Unterschiede.

In meinen Versuchen aus früherer Zeit über das Erkennen der Komponenten bei sehr kurzdauernder Darbietung von Mehrklängen aus Pfeifentönen, die in verschiedener Anzahl und Auswahl angegeben wurden, aber immer nur in den Verhältnissen der 6 ersten harmonischen Teiltöne zueinander standen und gleiche Stärke untereinander hatten (Zeitschr. f. Psych. Bd. 17), wechselten bei einem sehr intelligenten, aber von nervösen Stimmungen abhängigen Beobachter das synthetische und das analytische Verhalten in höchst merkwürdiger Weise miteinander ab. In manchen Versuchsreihen gab er fast immer nur an, einen Ton zu hören. „Er analysierte nicht das eine Mal besser, das andere Mal schlechter, sondern einmal analysierte er, und zwar besonders sicher und genau, das andere Mal analysierte er nicht, wenigstens nicht bei konsonanten Zusammenklängen in der hier angewandten Lagerung. Es ist, als wenn — um mich in einem physiologischen Bild auszudrücken — das durch Übung erworbene, gesonderte Funktionieren der einzelnen Hörganglien oder Hörprozesse durch einen über die ganze Hörphäre ausgebreiteten Hemmungsvorgang beeinträchtigt wäre.“ Der Bericht fügt noch hinzu, daß durch eine Art Autosuggestion das Beharren bei einem einmal eingeschlagenen analytischen oder synthetischen Verhalten begünstigt wurde. Von Autosuggestion würde ich in diesem Falle heute nicht mehr sprechen, aber als Tatsache sind uns ja solche „Perseverationen“ jetzt vielfach bekannt.

Bei den Aufbauversuchen mit der Interferenzmethode (o. S. 55) waren es besonders die hellsten Vokale Ü, E, I, bei denen mit dem Auftreten der Formantengegend zunächst ein gespaltener Klang gehört wurde (O + E, U + I). Hier liegen eben Unterformant und Formant in der Tonreihe weit auseinander. Aber mit Vervollständigung des Formanten und dem Eintritt der günstigen Intensitätsverhältnisse stellte sich die Einheit wieder her. Es kam jedoch auch vor, daß, als ich selbst den Vokal schon ganz einheitlich hörte, bei einem anderen Beobachter das analytische Verhalten noch fast bis zum vollendeten Aufbau erhalten blieb und dann plötzlich in das synthetische umschlug.

Auch bei Lückenversuchen mit dem A entstand gelegentlich ein ähnlich labiler Zustand, ja sogar ein Zusammen von Analyse und Synthese. So wurden, als beim A auf *as* nur die Teiltöne *as*, *as*<sup>2</sup>, *c*<sup>3</sup> und einige höhere übrigblieben, ein „dunkles A“, daneben aber noch gesondert ein dunkler Untergrund und hohe Bestandteile gehört. Ebenso als nur *as*, *es*<sup>2</sup> und die



hohen Elemente da waren, „OU mit hohen, mehr isolierten Bestandteilen“; als nur *as*, *c<sup>3</sup>*, *es<sup>3</sup>* da waren: „kein A, sondern U mit hohen gellenden Beittönen; nur wenn man sich künstlich vom Analysieren fernhält, kann man auch hier etwa A finden“ (Aussage des Beobachters). Hier gibt es also eigentümliche Zwischen- und Mischzustände in bezug auf das einheitliche Hören und den Vokaleindruck. (Das Vorkommen gradueller Übergänge zwischen ein- und mehrheitlichem Hören hat auch F. KRUEGER einmal im Arch. f. d. ges. Psych. Bd. 1, S. 228 betont und seine Bedeutung für die Lehre von den unbemerkten Teilinhalten hervorgehoben.)

#### g) Psychologische Faktoren.

Das einheitliche Hören ist nicht ganz unabhängig vom Willen. Beim analysierenden ist dies längst bekannt und liegt offen zutage. Aber man kann sich auch auf einheitliches Hören einstellen, kann dem Vereinheitlichungsprozeß innerlich entgegenkommen. Auch dies ist mir am Hörschlauch meiner synthetischen Einrichtung nicht selten aufgefallen. Mit der einheitlichen Einstellung war dann zugleich eine Vokalfarbe gegeben, mit der mehrheitlichen verschwand sie. Freilich muß alles übrige, die Klangzusammensetzung usw., in günstigem Sinne vorbereitet sein, wenn dieser psychische Faktor wirken soll. Es ist nur wie das Schließen einer Klappe.

Aber noch ein anderer, gewissermaßen entgegengesetzter psychischer Faktor dürfte mitwirken: die Gewöhnung. HELMHOLTZ hat das Prinzip aufgestellt, daß Sinnesempfindungen, die regelmäßig zusammen auftreten und gerade in dieser Vereinigung eine bestimmte Funktion für unser Bewußtsein, insbesondere für die Erkenntnis der Außenwelt haben, dadurch zu einem Ganzen werden, das wir nur mit Mühe wieder in seine Teile auflösen können. Die Anwendungen, die er davon auf das Heraushören der Obertöne und auf die Wahrnehmung der Doppelbilder der beiden Augen macht, sind zwar schweren Einwürfen ausgesetzt, und im ersteren Falle hat er sie sogar selbst in den späteren Auflagen der „Lehre v. d. Tonempf.“ stillschweigend wieder gestrichen. Aber das Prinzip als solches läßt sich wohl vertreten. In unserem Falle sind es wieder namentlich die 3 hellsten Vokale Ü, E, I, auf die ich es anwenden möchte. Es ist hier besonders merkwürdig, daß Unterformant und Formant, zwischen denen eine weite Nullstrecke liegt, als ein Klang gehört werden. Man kann auf die Schwäche der Unterformanten hinweisen. Aber schwächere tiefe Töne pflegen sich neben stärkeren hohen Tönen immerhin leicht geltend zu machen. Ich möchte daher glauben, daß hier wirklich die lebenslängliche Gewöhnung mit im Spiele ist<sup>1)</sup>.

<sup>1)</sup> Viele der jetzt in Berlin gebräuchlichen Automobilhupen geben ein Signal, das man als ein tiefes, aber zugleich schneidend scharfes U bezeichnen möchte. Bei einiger Aufmerksamkeit entdeckt man darin einen vom Grundton weit abliegenden, der 4- oder 5-gestr. Oktave angehörigen starken Ton,

Ja, die Vokale überhaupt dürften unter diesen Gesichtspunkt fallen. HELMHOLTZ hebt bereits hervor (S. 170), daß die menschliche Stimme für das unbewaffnete Ohr weniger leicht zu analysieren sei als andere Schallquellen. So allgemein ist dies zwar nicht zuzugeben: denn bei einer guten Männerstimme hört man, wenn der Sänger auf einer gleichbleibenden Tonhöhe langsam von U nach I übergeht, die Obertöne ebenso deutlich in die Höhe steigen, wie bei der Vokalröhre von WILLIS oder einem ähnlichen toten Instrumente. GRASSMANN hatte sich so auf die Vokale eingeübt, daß er beim A 10 Obertöne mit bloßem Ohr heraushören konnte. Aber beim gewöhnlichen Singen und Sprechen kommt dergleichen nicht vor, da der Singende die Vokale nicht so ineinander übergehen läßt und der Sprechende überhaupt nicht lange genug auf einem Tone verweilt, um für das Heraushören von Teiltönen Gelegenheit zu bieten, vor allem aber die Intention des Hörenden auf ganz andere Dinge gerichtet ist. So kann es wohl sein, daß gerade die Sprachlaute für uns in besonderem Maße einheitlich geworden sind, besser gesagt: daß bei ihnen das Analysieren durch das bloße Ohr schwerer fällt, als bei den musikalischen Instrumenten. Bei diesen mag die Neigung des guten abendländischen Musikers, Akkorde und polyphone Gebilde nach Möglichkeit in allen ihren Teilen aufzufassen, die Analyse auch einfacher Klänge begünstigen und eine der einheitlichen Auffassung entgegenstehende Gewöhnung begründen.

Außer dem Willen und der Gewöhnung wirken unter Umständen noch weitere psychische (bzw. psychophysische) Faktoren in einem das einheitliche Hören begünstigenden Sinne. So der augenblickliche Zusammenhang, die Einstellung auf Vokale überhaupt oder auf bestimmte Vokale (vgl. oben S. 50ff.).

Man kann die Gesamtheit dieser psychischen Faktoren auch als „psychische Einstellung“ im weiteren Sinne der rein physiologischen Einstellung an die Seite setzen, wobei es jedem überlassen bleibt, auch hier physiologische Mechanismen im Gehirn wirksam zu denken. Wir haben keine Veranlassung, auf die letzten Fragen der allgemeinen Psychophysik einzugehen.

also eigentlich die Konstitution eines I, aber doch mit einem so viel gewaltigeren und eindringlicheren unteren Bestandteil, daß wohl niemand den Gesamtklang einem I gleichsetzen würde. Hier ist es also umgekehrt wie bei den hellen Vokalen: der zusammengesetzte Laut wird trotz der scharfen hohen Beimischung als U gehört und diese wirkt nur wie ein „Oberformant“. Aber auch hier könnte man mit HELMHOLTZ sagen, daß die Gewöhnung an diese Art von Klang als Zeichen eines äußeren Objektes die einheitliche Auffassung besonders begünstige.

## Die gegenseitige Beeinflussung der Teiltonstärken und ihr Verhältnis zur Gesamtstärke des Klanges.

### I. Gegenseitige Beeinflussung.

Mehrfach ist schon erwähnt, daß die Intensitäten der Teiltöne, wie sie durch die gewöhnlichen objektiven Maßbestimmungen ermittelt werden, selbst wenn sie mit den durch besondere Untersuchungen festgestellten, die ungleiche Empfindlichkeit des Gehörs für die verschiedenen Tonhöhen ausdrückenden Koeffizienten multipliziert werden, nicht den physiologischen Intensitäten der Gehirnprozesse, welche die unmittelbare Grundlage und Vorbedingung der Klangerscheinungen sind, zu entsprechen brauchen, ja in gewissen Punkten sicher nicht entsprechen. Viel eher ließe sich dies von unseren rein phänomenalen Stärkekategorien annehmen, welche die unmittelbaren Wirkungen oder (nach parallelistischer Anschauung) die psychische Innenseite der letzten Gehirnprozesse darstellen. Aber diese Kategorien beziehen sich nicht auf die Teiltöne während ihres Zusammenklingsens, sondern nur auf die vor oder nachher isoliert gehörten Teiltöne. Zweifellos treten während und zufolge des Zusammenklingsens gewisse Wechselwirkungen zwischen den physiologischen Prozessen, teilweise schon im Ohre, teilweise erst im Hörzentrum des Gehirns ins Spiel, wodurch erst die endgültigen physiologischen Stärken festgelegt werden. Auf die subjektive Differenztonbildung, die ihren Sitz im Gehörorgan hat, wiesen wir schon öfters hin. Aber auch abgesehen davon ist die Frage, in welcher Richtung und welchem Ausmaße gleichzeitige Tonprozesse einander innerhalb des Organismus beeinflussen. Erst wenn wir darüber zureichende Kenntnisse besitzen, werden wir in der Lage sein, auf die wirklichen Stärkeverhältnisse der Teiltöne im Endstadium ihres Verlaufes zu schließen.

Bis auf die jüngste Zeit ist diese Frage arg vernachlässigt worden. Doch ist sie durch neuere Untersuchungen der Beantwortung wesentlich näher gerückt. Stellen wir die wichtigsten bisherigen Ergebnisse zusammen!

## 1. Ältere Beobachtungen.

Aus Beobachtungen ALFRED MAYERS, MACHS und des Verfassers<sup>1)</sup> schien bereits hervorzugehen, daß gleichzeitige Töne im allgemeinen einander etwas von ihrer Stärke abziehen, daß aber tiefere weniger durch höhere geschädigt werden als umgekehrt. Nimmt man 2 Gabeln im Intervall einer Oktave oder auch etwa einer Quarte und schwächt die höhere soweit ab, daß sie durch die tiefere eben verdeckt wird, so ist sie für sich allein immer noch gut hörbar. Wird dagegen die tiefere durch die höhere eben verdeckt, so ist sie auch für sich schon nahe an der Schwelle. Aber MAYER ging zu weit, wenn er die Verdrängung tiefer durch hohe Töne überhaupt für ausgeschlossen hielt, wie denn auch schon die gewöhnliche musikalische Erfahrung dagegen spricht. MACHS Beobachtungen waren am Klavier angestellt und lassen sich da leicht nachprüfen; er hatte sie aber nicht ganz richtig gedeutet.

1897 stellte ich mit MAX MEYER weitere Versuche an, bei denen als Klangquelle die elektromagnetischen Stimmgabeln des HELMHOLTZschen Apparates für Vokalsynthesen dienten, die den Vorteil konstanten Tönens und leichter Stärkeregelung boten. Die Versuche wurden gemacht, um MEYERS Vermutung zu prüfen, daß die simultane Stärkebeeinflussung den Grund für die Tonverschmelzung der konsonanten Intervalle enthalte. Hierbei zeigte sich 1. wieder, daß von 2 gleichzeitigen Tönen der tiefere dem höheren mehr abzieht als umgekehrt; 2. aber schien es einerlei, ob es sich um konsonante oder dissonante Intervalle handelte, womit also die obige Vermutung feile; 3. glaubte ich aus den Versuchen entnehmen zu können, daß das Stärkeverhältnis zweier Töne durch Hinzufügung eines dritten nicht geändert wurde (abgesehen von der Verstärkung durch Differenztonbildung); 4. war bei großen Tonabständen kein deutlicher gegenseitiger Einfluß zu bemerken. Endlich schien mir, wenn beide Töne der höheren Lage angehörten, die gegenseitige Beeinflussung geringer als in der Mitte. Aber die Versuche blieben fragmentarisch, die Thesen in so allgemeiner Form nicht hinreichend gesichert, und so wurde nichts darüber veröffentlicht<sup>2)</sup>.

<sup>1)</sup> S. die Berichte Tonpsych. II, S. 225ff., 231, 418—421.

<sup>2)</sup> M. MEYER hat kurz darauf (Zeitschr. f. Psych. Bd. 17. 1898) auf rein rechnerischem Wege eine Tabelle für die Amplitudenverhältnisse entworfen, bei denen unter Voraussetzung bestimmter theoretischer Grundanschauungen der höhere von 2 Tönen verschwinden müsse. Er fand die Werte verschieden je nach dem Phasenverhältnis. Aber bloße Phasenverschiedenheiten haben sicher hierauf keinen Einfluß. Außerdem legte MEYER noch

Im Zusammenhange mit den Vokalstudien nahm ich 1918 auch diese Frage wieder auf und konnte jetzt mit ganz einfachen Tönen arbeiten, wie dies auch hier für die Sauberkeit der Ergebnisse unumgänglich ist. Auf diese Untersuchung wie auf eine neuere aus dem Berliner psychologischen Institut werden wir weiter unten zurückkommen, nachdem die Ergebnisse systematisch durchgeführter amerikanischer Arbeiten dargelegt sind, über die hier ausführlicher berichtet werden soll, da sie, wie die MILLERS, in Deutschland noch wenig bekannt sind.

## 2. Neuere amerikanische Untersuchungen.

In diesen Arbeiten aus dem wissenschaftlichen Laboratorium der Vereinigten amerikanischen Telephon- und Telegraphengesellschaften und der Western Electric Company zeigt sich aufs neue, wie Ziele der Technik auf die entferntesten, scheinbar rein theoretischen Fragen zurückführen, aber auch wie der von praktischen Zielen geleitete Forscheringenieur sich über manche begriffliche Skrupel hinwegsetzt, mit denen sich Theoretiker oft den Fortschritt erschweren, so berechtigt sie sonst sein mögen. Das Recht nachträglicher Kritik muß diesen freilich verbleiben; aber manchmal wird es eben doch heißen dürfen: „Solvitur ambulando“. Die hier hauptsächlich in Betracht kommende Arbeit ist die von WEGEL und LANE. Eine sehr übersichtliche und elegante Darstellung dieser und anderer Untersuchungen desselben Kreises hat aber auch FLETCHER (1) gegeben.

Die Messungen waren rein physikalischer Art, aber die Fragestellung und die gesuchten Antworten lagen auf psychologischem, phänomenalem Gebiete. Die Verfasser benutzten einen besonders empfindlichen „Audiometer“ aus Röhrensender, Widerstand und Thermophon (statt des Telephons). Sie legten eine Untersuchung KNUDSENS über die absoluten und Unterschieds-Schwellen für die Stärke wie für die Höhe der Töne zugrunde. Die phänomenale Stärke (loudness) eines gegebenen Tones definierten sie durch die Zahl der unterscheidbaren Stärkegrade zwischen der Schwelle und der gegebenen Stärke, physikalisch durch den 10fachen Logarithmus des Verhältnisses zwischen dem Druck des gegebenen Tones auf das Trommelfell zu dem Druck bei dem Schwellenwert (die Drucke ließen sich aus der im Thermophon absorbierten elektrischen Energie

andere Voraussetzungen zugrunde, die mit seiner eigenen damaligen Hörtheorie zusammenhängen. Aber diese selbst führt in ihren Konsequenzen zu starken Widersprüchen mit den Beobachtungen. Deshalb müssen wir von diesem an sich interessanten Versuch einer theoretischen Lösung des Verdeckungsproblems hier absehen.

berechnen). Die Stärkeeinheit ist definiert durch den 10fachen Logarithmus der Unterschiedsschwelle.

An diesen Konzeptionen hätte wohl FECHNER seine Freude gehabt. Ich glaube auch nicht, daß man sie bemängeln kann, soweit es sich um reine Definitionen handelt, bei denen doch nur die Zweckmäßighkeitsfrage gestellt werden kann. Daß sie aber zweckmäßig waren, scheinen die schönen Gesetzmäßigkeiten zu zeigen, die sich aus den Versuchen in Verbindung mit diesen Definitionen ergaben.

Erfolgreich waren auch die (mir nicht aus dem Original bekannt gewordenen) Versuche von MC KENZIE, die Lautheit von Tönen verschiedener Höhe, ja von einfachen und zusammengesetzten Klängen, und sogar von Klängen und Geräuschen zu vergleichen, was bis dahin als wenig aussichtsvoll und sehr schwierig gegolten hat. Die zu vergleichenden Erscheinungen wurden in rascher Aufeinanderfolge, als eine Art Triller gegeben und die Fragestellung wurde auf die Gleichmäßigkeit dieses Trillers gerichtet. (Die Methode ist dem für heterochrome Helligkeitsvergleichen öfters angewandten Flimmerversfahren nachgebildet.) Es ergab sich, daß einfache Töne verschiedener Höhe dann als gleich laut erschienen, wenn sie um die gleiche Anzahl ihrer Stärkeeinheiten über ihrer Schwelle lagen. Für Töne und Klänge dagegen, ebenso für Klänge und Geräusche ergaben sich bei gleicher Lautheit ungleiche Anzahlen von Stärkeeinheiten. Klänge nahmen mit der Zahl ihrer Stärkeeinheiten schneller an Lautheit zu als einfache Töne.

Nun wurde ein Ton von bestimmter Höhe und Stärke gegeben und ein anderer, tieferer oder höherer, soweit geschwächt, daß er eben noch gleichzeitig neben jenem wahrzunehmen blieb. Dann wurde der erste in bestimmtem Maße verstärkt und festgestellt, um wie viele Stärkeeinheiten der zweite zunehmen mußte, um wieder merklich zu werden. So wurden von 200 bis zu 5000 Schw. die Schwellen des „verhüllten“ (masked) Tones bei verschiedenen Stärkestufen des „verhüllenden“ ermittelt und in einer Kurve zusammengefaßt. Diese Kurven sind von hohem Interesse und gewähren ein anschauliches Bild der etwas komplizierten Verhältnisse.

Ein Ton übt danach seine größte verhüllende Wirkung auf Töne, die ihm der Höhe nach benachbart sind, jedoch außerhalb der Schwebungszone liegen. Die Wirkung erstreckt sich sowohl nach oben als nach unten. Erweitert man den Höhenabstand nach beiden Seiten, so sinkt die Wirkung nach unten hin ziemlich rasch bis zur Nullgrenze, nach oben hin dagegen nur langsam. Dieser Unterschied ist aber nur bei starken verhüllenden Tönen bedeutend; hier müssen die hohen verhüllten Töne ganz beträchtlich gesteigert werden, um daneben hörbar zu werden, während die tiefen nur einer geringen Erhöhung ihrer Schwelle bedürfen. Bei schwachen, verhüllenden Tönen macht sich dieser Unterschied nur wenig geltend. Die These ALFRED MAYERS wird also von den Verfassern in ähnlichem Sinne berichtet, wie es mir von Anfang an notwendig schien.

Hatte beispielsweise der verhüllende Ton 400 Schw. und wurde dieser um 60 Einheiten über seine Schwelle gesteigert, so mußte ein gleichzeitiger Ton von 1250 Schw. um 46 Einheiten gesteigert werden, ein Ton von 3000 aber nur um 8, um daneben merklich zu sein. War der verhüllende Ton = 200, so konnte er um 55 Einheiten verstärkt werden, ohne daß ein Ton 3000 verdrängt wurde.

Aus dem Verlaufe der Kurven zieht FLETCHER auch Folgerungen für Komplexe aus mehr als 2 Tönen. Es seien z. B. gegeben die 3 Töne 400, 300, 2000 mit den relativen Stärken 50, 10, 10. Dann kann man den Kurven gemäß nur den 1. und 3. hören; 300 liegt noch unter seiner Schwelle und müßte um 16 Einheiten verstärkt werden, um wahrnehmbar zu sein. Haben aber dieselben 3 Töne die Stärken 80, 40, 40, so werden nur die beiden ersten gehört, der ganz hohe unterdrückt.

Es ist nicht ersichtlich, ob hier eine bloße Folgerung vorliegt oder ob sie durch Beobachtung kontrolliert wurde. Bei der Folgerung ist vorausgesetzt, daß das Stärkeverhältnis zweier Töne durch einen dritten nicht geändert wird. Dies schien mir allerdings aus wenigen vorläufigen Beobachtungen hervorzugehen (s. o.), aber es ist nicht selbstverständlich und bedürfte wohl noch einer genauen experimentellen Prüfung.

Allgemein schließen die Verfasser, daß bei Verstärkung eines Mehrklanges in gleichem Verhältnis aller Teile die tiefen Töne sich stärker geltend machen müssen; und finden dies auch in der Erfahrung bestätigt.)

Bei Verteilung zweier Töne an beide Ohren (dichotisch nach unserer Bezeichnungsweise) fanden sich die Verhüllungswirkungen ungeheuer viel geringer. Was sich davon noch zeigte, glauben die Verfasser auf Hinüberdringen der Töne durch die Knochenleitung, also auf monotisches Hören zurückführen zu müssen. Unleugbar geht aus diesem Unterschiede hervor, daß die Verhüllungserscheinungen im wesentlichen nicht zentraler, sondern peripherischer Natur sind<sup>1)</sup>.

Die Verfasser machen endlich ihre Beobachtungen auch zur Hauptstütze einer neuen Hörtheorie, die zwar ebenso wie die HELMHOLTZsche auf Resonanz zurückgeht, aber nicht für jeden Ton besondere Fasern oder Fasergruppen resonieren läßt, sondern immer die ganze Grundmembran der Schnecke, nur mit verschiedener Lage der Resonanzmaxima. Dabei werden aber wie

<sup>1)</sup> MINTON erhebt (Phys. Rev. Bd. 22, 1923) dagegen den Einwand, daß eine schwach tönende Gabel vor dem einen Ohre durch eine gleichgestimmte stärkere vor dem anderen Ohre verdeckt werde, was nur zentralphysiologisch zu deuten sei. Aber diese Erscheinungen beim Unisono gehören überhaupt zu einer anderen Tatsachengruppe, über welche namentlich v. HORNBOSTEL (1, S. 84ff.) zu vergleichen ist.

bei HELMHOLTZ die hohen Töne näher an die Basis, die tiefen näher an die Kuppel der Schnecke verlegt.

In einigen Punkten bleiben dem Leser, dem die benutzten Einrichtungen nicht durch den Augenschein, die benutzten Klangquellen nicht durch eigenes Nachprüfen bekannt sind, noch Fragezeichen, durch welche indessen die Hauptergebnisse kaum berührt werden dürften.

KNUDSENS Untersuchungen über die Unterschiedsschwellen für Tonhöhen lieferten ungewöhnlich hohe Werte, z. B. für einen Ton von 1000 Schw. 3, für einen von 500 Schw.  $1\frac{1}{2}$  Schw., also mindestens das 6fache der früher ermittelten Schwellen. Diese hohen Werte haben ihren Grund vor allem wohl in der ungewöhnlichen Methode: es wurden durch vielfaches Abwechseln der beiden, durch ganz kurze Pausen getrennten Töne Reihen gebildet, die auf ihre Gleichmäßigkeit hin zu beurteilen waren. Wie weit sich diese Methode bewährt, und welchen besonderen Fehlerquellen sie ausgesetzt ist, muß abgewartet werden. Außerdem aber ist KNUDSEN offenbar überhaupt nicht auf die schlechthin niedrigsten Schwellenwerte ausgegangen, wie sie ein besonders fein veranlagtes und hochgeübtes Ohr erzielen kann, sondern auf Durchschnittswerte von vielen Personen. Bedenkt man aber die enorme Unfähigkeit des Tonhöhenurteils bei manchen Unmusikalischen (Tonpsych. I, 313ff. u. ö.), so dürften Durchschnittswerte nur dann einen Sinn haben, wenn eine gewisse Siebung der Versuchspersonen stattgefunden hat und nur Individuen eines gewissen Typus zusammen genommen werden. Ich weiß nicht, in wie weit KNUDSEN so verfahren ist.

Ein weiteres Fragezeichen betrifft die Verhüllungserscheinungen bei Annäherung an den Einklang. Wenn die Schwelle des verhüllten Tones hierbei immer höher wird, und wenn dieses Gesetz keine Ausnahme erlitte, müßte sie beim Einklange selbst unendlich hoch werden; und dies würde auch insofern stimmen, als der verhüllte Ton dann vom verhüllenden unter keiner Bedingung mehr zu unterscheiden ist, er mag verstärkt werden, so viel er will. Aber vor diesem Punkte treten Schwebungen auf, „Störungen des Zusammenklanges“, wie sie HELMHOLTZ nannte, die sich in verschiedenen Modifikationen bis über den Umfang eines oder mehrerer Ganztöne hinaus erstrecken. Zuerst sind, vom Einklange ausgehend, die Töne immer noch nicht zu unterscheiden, aber die wahrgenommene einheitliche Tonhöhe unterliegt bestimmten Veränderungen; dann werden sie allmählich unterscheidbar, lassen aber noch einen 3. Ton zwischen sich hören usw. Diese mehrfach ausführlich untersuchten und beschriebenen Erscheinungen finden in den Kurven der Verfasser ihren Ausdruck durch eigentümliche Knickungen. Aber es ist nicht klar, inwiefern diese Knickungen die Intensitätsverhältnisse der beiden Töne an diesen Stellen wiedergeben sollen, da es sich eben hier teils nur um einen, teils um 3 Töne handelt und eine Schwelle des schwächeren gegen den stärkeren Primärton nur etwa insofern angegeben werden kann, als man diejenigen Intensitäten aufsucht, bei denen der schwächere eben noch merkliche Schwebungen des stärkeren bewirkt. Aber auch in diesem Falle wird er, wie die Verfasser selbst bemerken, nicht neben diesem wahrgenommen, sondern nur aus den Schwebungen erschlossen. Also hat die Schwelle und die Schwellenkurve hier eine gänzlich veränderte phänomenale Bedeutung. Richtiger wäre es wohl gewesen, in diesen Bezirken die Kurven überhaupt zu unterbrechen.

Die erwähnten Knickungen treten in den Kurven aber nicht bloß in der Gegend des Einklanges, sondern auch an Stellen auf, welche den 3 ersten harmonischen Obertönen des verhüllenden Tones entsprechen. Die Ver-



fasser haben denn auch Schwebungen an diesen Stellen beobachtet, führen sie aber nicht auf objektive, sondern auf subjektive Obertöne zurück. Objektiv seien nur ganz verschwindende Obertöne vorhanden (ihre Intensität soll nach einer brieflichen Mitteilung von H. FLETCHER an v. HORNBOSTEL zufolge harmonischer Analyse nur 0,0001 von der des Gesamtklanges betragen). Die subjektiven Obertöne seien die Folge der nichtlinearen Reaktion des Ohres auf starke Töne (so hat sie bereits HELMHOLTZ S. 263 abgeleitet).

Das wirkliche Vorhandensein solcher subjektiven Obertöne beim gewöhnlichen Hören habe ich Tonpsych. II, 260ff. bezweifelt und bezweifle es heute noch, obgleich sie rechnermäßig bei fortgesetzter Tonverstärkung zu erwarten sind. Direktes Hören zeigt mir trotz langer Übung nichts davon, auch bemerke ich keine entsprechende Veränderung der Klangfarbe durch bloße Verstärkung eines einfachen Tones meiner synthetischen Einrichtung innerhalb der hier möglichen Stärkegrenzen. Mit Dr. v. ALLESCH versuchte ich eine Prüfung durch schwebende Hilfsgabeln.  $c^1$  der synthetischen Einrichtung gab, möglichst stark ins Ohr geleitet, mit der etwas verstimmtten Hilfsgabel  $c^2$  in der Tat Schwebungen, ebenso mit  $g^2$ , aber nur, wenn das Hörrohrchen an die Wand des Gehörganges angedrückt wurde, nicht aber bei bloßer Luftleitung. Außerdem kommt in Betracht, daß die Hilfsgabel  $c^2$  durch Differenztonbildung mit  $c^1$  Schwebungen auf diesem Tone, die sog. „tiefen Schwebungen“ verstimmtter Oktaven, geben muß. Dasselbe Ergebnis erhielten wir für  $g^1$  als Grundton, wo die Schwebungen mit  $g^2$  auch deutlich auf  $g^1$  lokalisiert schienen, also Differenztonschwebungen waren.  $d^2$  schwebte auch hier nur beim Andrücken des Röhrchens an den Gehörgang.

Übrigens ist mir auch nicht klar geworden, inwiefern subjektive Obertöne als Erklärung dafür dienen sollen, daß ein höherer Ton leichter durch einen tieferen verdrängt wird, als umgekehrt. Die Verfasser scheinen anzunehmen, daß der auf diese Obertöne fallende Teil der Gesamtenergie dem höheren, aber nicht dem tieferen Tone entzogen werde; aber die sehr kurzen Andeutungen lassen nicht erkennen, warum.

Die Verfasser weisen ferner auf subjektive Kombinationstöne hin, die gleichfalls aus der nichtlinearen Reaktion des Gehörorgans entspringen müssen (auch hier merkwürdigerweise, ohne HELMHOLTZ' Vorgang und die vielen seither darüber gepflogenen eingehenden Untersuchungen, insbesondere die WAETZMANN'S seit 1907, zu erwähnen). Ihr wirkliches Vorhandensein bis zu solchen 3. Grades erschließen sie aus Schwebungen mit Hilfsgabeln. In der Behauptung subjektiver Kombinationstöne stimme ich den Verfassern natürlich bei, sind doch die beiden stärksten seit lange bekannt. Aber die Schwebungsmethode, an sich freilich die feinste, führt unter Umständen zu Fehlschlüssen, indem durch das Hinzubringen der Hilfsgabel selbst Kombinationstöne erzeugt werden können. Ich habe die hierbei zu beobachtenden Vorsichtsmaßregeln erörtert und die bei einfachen Primärtönen wahrnehmbaren Kombinationstöne unter Berücksichtigung aller Umstände ihres Auftretens, auch des sehr wichtigen Stärkeverhältnisses der Primärtöne festgestellt. Die Untersuchung ist der Aufmerksamkeit der amerikanischen Physiker entgangen<sup>1)</sup>.

<sup>1)</sup> Vgl. m. Abh. 7. Daß alle Einzelheiten der an K. T. zu beobachtenden Erscheinungen aus HELMHOLTZ'S Grundgedanken herzuleiten wären, soll nicht behauptet werden. Warum bei kleinen Intervallen der 1. Differenzton so stark, bei hohen Primärtönen sogar unter Umständen stärker ist

## 3. Eigene Beobachtungen.

Ich berichte nun über noch unveröffentlichte Beobachtungen, zu denen die völlig einfachen und in ihrer Stärke fein regulierbaren Töne der synthetischen Einrichtung, statt der physikalischen Messung aber wieder die rein phänomenalen Stärkekatégorien benützt wurden. Ich hatte es nicht bloß auf den Fall der Verdrängung eines Tones durch einen anderen abgesehen, sondern notierte auch bloße Abschwächungen. Die Frage war: wie verändert sich durch das Zusammenklingen der schwächere, wie der gleichstarke, wie der stärkere Ton? Doch betrafen die meisten Beobachtungen auch hier die Punkte der Ebenmerklichkeit und des Verschwindens. Es wurden die Tonpaare  $g^1 c^2$  (Quarte),  $c c^1$ ,  $c^1 c^2$  (Oktaven),  $c^1 e^2$  (Dezime) und  $c^1 c^3$  (Doppeloktave), aber auch noch andere größere Intervalle geprüft.

a) Waren die beiden Töne deutlich ebenmerklich ( $= 1/8$ ), so blieben sie beim Zusammenklingen fast ausnahmslos unverändert. Jedenfalls war von einem Verdrängen des einen durch den andern nichts zu bemerken. Oft schien mir sogar die Deutlichkeit, wenn auch nicht die Stärke, der Komponenten wie durch eine Art Kontrastwirkung im Zusammenklänge größer. Bei der Oktave schien mir aber zuweilen der tiefere Ton wirklich etwas verstärkt (s. unter g).

b) Waren beide Töne ungleich stark und lag der stärkere im Stärkebereich 1—2, so war entweder keine oder nur eine schwache schädigende Wirkung des stärkeren auf den schwächeren zu bemerken. Auch trat kein ausgesprochener Unterschied zutage, je nachdem der höhere oder der tiefere stärker war.

Z. B. wenn bei der Quarte oder der Oktave der eine Ton 1, der andere  $1/8$  betrug, blieb der letztere im Zusammenklang unverändert hörbar, auch

als diese selbst, warum auch der D. T.  $2t - h$  vielfach dieselbe, manchmal eine größere Stärke besitzt als  $h - t$ , wogegen wieder  $2h - t$  minimal schwach ist: warum  $3h - 2t$  bei einfachen Primärtönen noch schwach hörbar,  $2h - 2t$  dagegen ganz unhörbar ist: dies und anderes muß wohl mit der besonderen Beschaffenheit der Membranen unseres Gehörorgans (Trommelfell, ovales Fenster), vielleicht auch der Knöchelchenkette, zusammenhängen. Für die Schwäche von  $2h - t$  mag auch in Betracht kommen, daß dieser Ton immer oberhalb der Primärtöne liegt, daher von diesen in höherem Grade als die tiefen Differenztöne subjektiv geschädigt wird.

Die von mir und mehreren Mitbeobachtern bemerkte Erscheinung unharmonischer K. T. bei der verstimmtten Oktave und Duodezime (7, S. 121) scheint sogar auf zentrale Einflüsse in einzelnen Fällen hinzuweisen. Bei der Seltsamkeit dieser Erscheinung sei eine analoge Beobachtung aus der gegenwärtigen Untersuchungsreihe erwähnt. Als nämlich die Oktave  $c^1:c^2$  der synthetischen Einrichtung sich zufällig ein wenig verstimmt hatte, war bei jedem Schwebungsmaximum ein unerklärlicher K. T.  $es^2$  deutlich zu beobachten.

wenn er der höhere war. Also keine merkliche Erhöhung der Schwelle. War  $c^2 = 1$ , so verschwand die tiefere Quarte  $g^1$  erst dann, wenn sie auch isoliert schon untermerklich war. Bei der Dezime  $c^1 = 1$ ,  $e^2 = 1/8$  blieb  $e^2$  ebenmerklich. Fast schien es sogar wieder durch eine Art Kontrast deutlicher zu werden. Ähnliches bei der Doppeloktave.

c) Dagegen trat bei bedeutender Stärke des einen Tones, wenn er gellend ( $= 3$ ) oder wenigstens mittelstark ( $= 2^{1/2}$ ) war, eine erhebliche Schädigung des anderen, schwächeren bei Intervallen bis zur Doppeloktave jederzeit ein; und zwar im allgemeinen eine größere Schädigung, wenn er über als wenn er unter dem schädigenden lag. Ein gellender Ton tötet einen höheren, selbst wenn dieser beträchtlich über seiner Schwelle liegt.

Bei der Quarte z. B. verschwand der höhere Ton bei  $g^1 = 2^{1/2}$ ,  $c^2 = 1$  oder  $g^1 = 3^{1/2}$ ,  $c^2 = 1^{1/2}$ . Der tiefere Ton verschwand bei  $g^1 = 1/8$ ,  $c^2 = 2$ .

Bei der Oktave unterdrückte der tiefere den höheren, wenn jener  $= 3$ , dieser  $= 1^{3/4}$  war. Es war dann noch eine Veränderung der Klangfarbe zu beobachten. War aber der höhere  $= 1$ , so machte sich überhaupt keine Veränderung bemerklich. War  $c^1 = 2^{1/2}$ ,  $c^2 = 1/4$ , so war dieses noch eben merklich, bei  $1/8$  dagegen nicht mehr. Der höhere unterdrückte den tieferen, wenn jener  $= 4$ , dieser  $= 1^{1/2}$  oder  $2$  war. Ebenso bei  $c^2 = 2^{1/2}$ ,  $c^1 = 1/8$ , wogegen bei  $c^1 = 1/4$  dieses noch hörbar blieb. Hier machte also die Umkehrung keinen Unterschied.

Bei der Dezime drückte der tiefere Ton mit  $2^{1/2}$  den höheren von der Ebenmerklichkeit auf Unmerklichkeit herab; um merklich zu bleiben, mußte dieser auf  $1/2$  verstärkt werden. Der höhere unterdrückte den tieferen erst mit Stärke  $3$ .

Bei der Doppeloktave ebenso; nur blieb der tiefere Ton mit  $1/8$  noch neben dem höheren hörbar, wenn dieser selbst auf  $3$  gesteigert wurde.

d) Dafür, daß bei ungleicher Stärke auch der stärkere Ton eine Einbuße erlitte, findet sich unter dieser Beobachtungsreihe kein Anhaltspunkt; ich habe darauf allerdings auch nicht so geachtet wie auf die Schädigung des schwächeren. Sollte es in geringem Maße der Fall sein, so müßten zur Feststellung wohl auch feinere Mittel angewandt werden.

e) Ob der Konsonanz- und Verschmelzungsgrad einen Unterschied macht, erscheint auch nach diesen Versuchen zweifelhaft. Doch möchte ich einen geringen Einfluß wenigstens bei der Oktave nicht ausschließen (vgl. die Schwelle des höheren Tones bei Quarte und Oktave).

f) Daß der schädigende Einfluß des stärkeren Tones auf den schwächeren mit ihrem Abstand in der Tonreihe (von Terzen anfangend) abnimmt, hat sich auch hier im allgemeinen bestätigt. Bei einem Tonabstand von 4—5 Oktaven, z. B. bei  $c$  und  $a^4$ , schien mir überhaupt keine Schädigung, weder eine gegenseitige bei gleicher, noch eine einseitige bei ungleicher Stärke stattzufinden. Der schwächere Ton verschwand im Zusammenklang bei

der nämlichen Stärke wie im isolierten Zustand, und es machte dabei auch keinen Unterschied, ob es der höhere oder tiefere war: immer blieb seine Schwelle unverändert.

g) Bei gleicher Stärke der Töne im isolierten Zustand erschien im Zusammenklange der tiefere merklich stärker als der höhere.

Dies tritt vor allem aus einem besonderen Grunde regelmäßig ein bei der Oktave. Der tiefere ist eben hier zugleich Differenzton. Selbst bei Stärke  $\frac{1}{8}$  für beide Töne scheint der tiefere etwas verstärkt. Erschienen im Zusammenklange  $c$  und  $c^1$  beide = 2, so zeigte sich bei isolierter Angabe  $c^1 = 1$ ,  $c^2 = 2$ . Einigermaßen könnte Ähnliches auch bei der Quinte zutreffen, da der tiefere Ton hier sowohl 1. als 2. Differenzton ( $h-t$  und  $2t-h$ ) ist; namentlich wenn der Ton  $2t$  als Oberton real vorhanden ist. Und so noch in anderen Fällen.

Aber auch wo keine Differenztonbildung dem tieferen zugute kommt, scheint er im Zusammenklange begünstigt zu sein. Dies ist mir besonders aufgefallen bei Versuchen mit Dreiklängen, die mit Hilfe des Regulierungsapparates so gegeben wurden, daß die Töne im Zusammenklange untereinander gleich stark, und zwar jeder mit der Stärke 2 erschienen. Wurden die Töne dann einzeln gegeben, so war der tiefste der schwächste und nahm überhaupt die Stärke im allgemeinen mit der Höhe zu. Beispielsweise waren bei  $c^1 g^1 g^2$  oder  $g^1 e^2 b^2$  oder  $c^1 e^2 d^3$  jedesmal die isolierten Stärken:  $1\frac{1}{2}$ ,  $1\frac{3}{4}$ , 2.

In manchen Versuchen schien aber der mittlere Ton unter dem Zusammenklange besonders zu leiden, namentlich bei Tönen von geringem Abstand. So ergaben  $e^2 g^2 b^2$  die isolierten Stärken  $1\frac{1}{2}$ ,  $2\frac{1}{2}$ , 2. Selbst wenn ich den Zusammenklang so regulierte, daß die simultanen Stärken = 2,  $\frac{1}{2}$ ,  $1\frac{1}{2}$  waren, so daß man  $g^2$  kaum mehr heraushören konnte, erwiesen sich die isolierten Stärken =  $1\frac{3}{4}$ ,  $2\frac{1}{2}$ , 2.

Hiermit dürfte die Tatsache zusammenhängen, daß man Differenztöne, die zwischen 2 Primärtöne zu liegen kommen (wenn nämlich diese weiter als eine Oktave auseinander liegen), außerordentlich schwer heraushören kann<sup>1)</sup>. Auch bei Versuchen über

<sup>1)</sup> Auf diese Erklärung habe ich auf Grund analoger Beobachtungen schon 7, S. 142 hingedeutet. Damals machte ich folgende (dort nicht mitgeteilte) Beobachtungen an der „Flaschenorgel“. Läßt man  $g^2 d^3$  erklingen, so erscheint der D. T.  $g^1$ . Fügt man nun noch den Flaschenton  $c^1$  hinzu, so wird dieses  $g^1$  unterdrückt; fügt man dagegen den Flaschenton  $c^2$  (statt  $c^1$ ) hinzu, so bleibt es hörbar. Ferner:  $e^2 a^2$  zusammen geben die D. T.  $a$  und  $a^1$ . Fügt man nun den Zweiklang  $e^1 + c^2$  hinzu, so wird  $a^1$  unter-

die Analyse eines Tongemisches durch unmusikalische oder exotische Individuen findet sich regelmäßig, daß sie die äußeren Töne leichter als die mittleren heraushören<sup>1)</sup>.

Das Auffallendste bei diesen Dreiklangversuchen war aber, daß der tiefste Ton im Zusammenklang nicht bloß weniger geschwächt als die höheren, sondern gegenüber seinem isolierten Zustande geradezu verstärkt erschien; also ähnlich wie bei der Oktave, aber ohne daß man hier entsprechende Differenztöne dafür haftbar machen kann. Es wäre sehr zu wünschen, daß solche Versuche auch mit objektiven Stärkemessungen der isolierten Töne wiederholt würden; denn das Ergebnis ist so neu und befremdlich, daß ich meinen nicht zahlreichen Beobachtungen hierüber noch kein entscheidendes Gewicht beilegen möchte.

Vergleichungen der Stärke eines isolierten Tones mit der Stärke desselben Tones als Teil eines Zusammenklanges unterliegen ja auch einer eigentümlichen Schwierigkeit. Man hat zuerst den Eindruck, als sei hier eine exakte Vergleichung überhaupt unmöglich, als handle es sich um Stärken in verschiedenem Sinne des Wortes. Es spielt eben die Deutlichkeit des Heraushörens mit: der deutlicher herausgehörte Ton, z. B. ein dissonierender Ton innerhalb eines sonst konsonanten Zusammenklanges, wird leicht als stärker beurteilt, ohne es wirklich (auch nur subjektiv wirklich) zu sein. Im allgemeinen mag die Konzentration der Aufmerksamkeit auf einen einzelnen Ton durch die übrige Klangmasse erschwert sein. Aber unüberwindlich sind diese Schwierigkeiten nicht, und andere Beobachter fanden sie nicht einmal besonders groß (vgl. unten 4).

g) Der tiefste Ton eines Zwei- oder Mehrklanges behält den Charakter als Grundlage des Klangganzen bei abnehmender relativer und absoluter Stärke noch lange bei. Er verliert ihn erst, wenn er sich seinem Schwellenwerte nähert. Bei der Oktave  $c^1 c^2$ , wenn  $c^2 = 2^{1/2}$ , hatte  $c^1$  mit der Stärke 1 entschieden noch die Führung. Sie wurde erst schwankend mit  $c^1 = 1/4$ , wo dieses schon nahe an seiner simultanen Schwelle ist. Es begleitet dann den stärkeren Ton nur noch wie ein leichter Schatten<sup>2)</sup>.

drückt,  $a$  aber bleibt hörbar. Also ein gewöhnlicher D.-T.  $g^1$ , der unterhalb der erzeugenden Primärtöne liegt, wurde durch einen neu hinzugefügten tieferen (nicht zu weit davon entfernten) Primärtönen  $c^1$  unterdrückt, aber nicht durch einen höheren  $c^2$ ; und von 2 D. T. wurde derjenige, der zwischen 2 neu hinzugefügten Primärtönen lag, unterdrückt, aber nicht der darunter liegende.

<sup>1)</sup> Tonpsych. II, S. 362ff. Beitr. z. Akust. u. Musikwiss. Heft 3, S. 106.

<sup>2)</sup> Vgl. Tonpsych. II, 384ff., 421.

Der Grund, warum wir den tiefsten Ton eines Klanges, auch wenn er nicht der stärkste ist, doch als Fundament des Ganzen und als den eigentlichen Träger der Tonhöhe auffassen, liegt wohl hauptsächlich in seinem größeren Volumen und den daran geknüpften räumlichen Assoziationen<sup>1)</sup>. Darauf einzugehen würde hier zu weit führen, nur die Tatsache selbst ist für uns wichtig.

Vergleicht man die unter 3. mitgeteilten Ergebnisse mit den Schwellenbeobachtungen der amerikanischen Physiker unter 2., so herrscht eine befriedigende Übereinstimmung: sie ordnen sich in deren „Verhüllungskurven“, soweit ich sehe, widerspruchlos ein, freilich ohne daß physikalische Maßbestimmungen eine zahlenmäßige Vergleichung ermöglichten. Dagegen gehen sie gemäß der erweiterten Fragestellung in verschiedenen Richtungen darüber hinaus.

4. Nehmen wir nun alle unter 1—3 besprochenen, hinreichend beglaubigten Tatsachen zusammen, so ergibt sich für die Vokalstruktur, daß sich die subjektiven Stärken der Teiltöne, wie sie im isolierten Zustande festgestellt werden, beim Zusammenklang im Gehörorgan und im Gehirn zugunsten der jeweilig tieferen Teiltöne verschieben müssen. Zu der Verstärkung, die der Grundton als Differenzton aller unmittelbar aufeinanderfolgenden harmonischen Obertöne und die auch andere tiefere Teiltöne als Differenztöne höherer Teiltöne erhalten, treten nun noch die Verschiebungen der relativen, ja, wahrscheinlich sogar der absoluten Stärke, wie sie durch die Begünstigung der tieferen Töne im Zusammenklang gegeben sind. In manchen Fällen müssen nach den obigen Gesetzlichkeiten hohe Obertöne durch einen aus mehreren starken, tieferen Tönen bestehenden Formanten so geschwächt werden, daß sie sich sogar in der Klangfarbe kaum mehr geltend machen können. Doch erweisen sich diese Fälle, wenn man unsere synthetischen Tabellen daraufhin durchsieht, als selten; zumeist werden doch die höheren Teiltöne über ihrer Schwelle bleiben.

Aber auch wenn der Grundton des gesungenen oder gesprochenen Vokals trotz dieser verstärkenden Faktoren in der Hirnrinde immer noch schwächer vertreten ist als mancher Oberton, so versteht man aus der zuletzt erwähnten Tatsache, daß er Grundton bleibt und die Rolle der Klanghöhenbestimmung nicht etwa an die Oktave oder Duodezime oder an den Formanten abgibt.

<sup>1)</sup> Vgl. ebenda I, 202 ff.; II, 46ff. u. ö.

## 4. Physikalische Intensität herausgehörter Teiltöne.

Eine bemerkenswerte Ergänzung dieser Ergebnisse bildet die Untersuchung, die 1922 Fr. EBERHARDT im Psychologischen Institut der Berliner Universität auf Veranlassung W. KOEHLERS ausgeführt hat. Mit Hilfe der oben S. 17 erwähnten LEWINSchen Einrichtung wurde die relative physikalische Intensität eines in einem Klangganzen enthaltenen Tones bestimmt. Dann wurde eine Versuchsperson (Vp) veranlaßt, unter einer großen Zahl von objektiv hergestellten Intensitätsstufen eines isoliert dargebotenen Tones von gleicher Höhe wie der vorher untersuchte diejenige Stufe auszusuchen, die der Stärke des in und aus dem Klangganzen gehörten Tones gleichkam. Hierauf wurde diese Intensität wieder physikalisch bestimmt. Sie erwies sich durchgehends als kleiner gegenüber der ersten Bestimmung. Hieraus ließ sich schließen, in welchem Maße der Ton durch den gleichzeitigen Tonkomplex subjektiv an Stärke geschädigt war. Der zu bestimmende Ton war in einer Versuchsreihe als 3. Teilton eines Klanges ( $g^2$  im Klange von  $c$ ), in einer 2. Reihe als Komponent eines Zwei- oder Dreiklanges aus annähernd gleichstarken Tönen ( $g^2$  mit  $es^2$  oder mit  $b^2$  oder mit beiden), in einer 3. Reihe als Teil eines Zweiklanges ( $e^2$   $gis^2$ ) mit veränderlichem Stärkeverhältnis der Komponenten gegeben.

Die 8 Beobachter konnten die Vergleichung mit beträchtlicher subjektiver Sicherheit ausführen und urteilten bei Wiederholung der Versuche in annähernd gleicher Weise, aber — und dies ist besonders bemerkenswert — sie differierten untereinander außerordentlich. In der 1. Serie betrug die Intensität zwischen 1,3% und 29% der vorher gemessenen physikalischen Intensität. In der 2. Serie betrug die Schädigung für alle Personen weniger (Intensität bei  $es^2$   $g^2$  zwischen 6 und 89%, bei  $g^2$   $b^2$  zwischen 4,5 und 72%, bei  $es^2$ ,  $g^2$ ,  $b^2$  zwischen 4 und 68%); die Rangordnung der Vp war aber ziemlich dieselbe wie in der 1. Serie. In den meisten Fällen erschien  $g^2$  etwas mehr geschädigt, wenn es mit  $b^2$  als wenn es mit  $es^2$  zusammen war (eine Abweichung gegenüber den früheren Ermittlungen, wonach ein tieferer Ton stärker schädigt), am meisten aber, wenn es als mittlerer Ton zwischen beide eingefügt war, obgleich es hier an sich sogar etwas stärker gegeben wurde. In der 3. Serie nahm mit wachsender Stärke des  $gis^2$  die Intensität des herausgehörten  $e^2$  ab, z. B. bei Erhöhung durch 4 Intensitätsstufen von 1—140 sank sie für eine Vp von 73 auf 1%, und ähnlich auch für die anderen Vp. Die Verfasserin fügt hier bei, daß nach alltäglichen Erfahrungen zahlreiche leisere Töne einem gegebenen weniger Abbruch tun als ein einziger lauter Ton.

Das bemerkenswerteste Ergebnis dieser Versuche ist neben der Feststellung der großen individuellen Differenzen die große Schwächung der Beitäöne eines Klanges gegenüber der viel geringeren Schwächung annähernd gleichstarker Komponenten eines Mehrklanges. (Dies stimmt auch mit den vorher erwähnten Untersuchungen, wonach der schwächere Ton im allgemeinen mehr leidet als der gleich starke.) In einem Versuche mit dem sehr obertonreichen Zungenklang auf *c* wurde der 5. Teilton  $e^2$  selbst von dem im Heraushören unübertrefflichen Prof. v. HORNBOSTEL nur mit 4% seiner wirklichen Energie vernommen. Dagegen hörte derselbe Beobachter aus dem Dreiklange  $cis^2 e^2 gis^2$  bei annähernd gleicher Stärke aller Töne den mittleren noch mit 71% heraus.

Bereits KOEHLER hatte (1, Vörl. Mitt., S. 100ff.; 1, III, S. 129ff.) nach Beobachtungen an sich selbst von den Obertönen nur einen „schwachen Rest“ finden können und auch auf SEEBECK hingewiesen, der die nämliche Beobachtung gegen OHMS Klanglehre ins Feld geführt hatte. KOEHLER benutzte sie denn auch schon 1912 als Stütze einer Modifikation der HELMHOLTZschen Hörtheorie, wonach, ähnlich wie in der Umgestaltung der amerikanischen Forscher (s. o.), nicht mehr einzelne Fasern oder kleine Fasergruppen, sondern immer nur größere Regionen der Grundmembran Träger der Resonanz sein würden. Wir erwähnen diese Folgerungen und Theorien, ohne hier dazu Stellung zu nehmen.

Die Verfasserin ist nicht, wie WEGEL und LANE, auf Schwellenbeobachtungen ausgegangen. Aber nach gelegentlichen Bemerkungen scheint es, daß die von ihr ermittelten individuellen Unterschiede sich nicht in entsprechenden Verschiedenheiten der Schwelle kundgeben, d. h. daß das Heraushören eines Tons aus einem Komplex gleichzeitiger Töne nicht gleich großen individuellen Unterschieden unterliegt. Verf. erwähnt z. B., daß das Heraushören des mittleren Tones aus dem Dreiklange  $c^2 e^2 g^2$  mehreren, auch weniger dafür begabten, Vp noch möglich war, wenn er objektiv so geschwächt wurde, daß seine Simultan-Stärke auf 0,3% herabging. Selbst bei 0,1% blieb er noch „tropfenweise“ hörbar. Wurde er dann ganz weggelassen, so machte sich regelmäßig wenigstens noch eine Änderung der Klangfarbe bemerklich. Mit der musikalischen Anlage der Vp möchte die Verf. die Ungleichheit der herausgehörten Intensitätsbeträge nicht ohne weiteres in Verbindung bringen, wenn auch hochmusikalische Personen in der Regel größere Beträge heraushörten. Immerhin — musikalische Anlage läßt sich ja verschieden definieren, und die persönliche Aussage darüber ist nicht maßgebend — dürfen wir annehmen, daß die Neigung zum analysierenden Hören, durch welches allein mehrstimmige Musik voll verstanden und genossen werden kann, durch höhere subjektive Intensität der



herausgehörten Töne begünstigt wird und daß auch umgekehrt durch eine angeborene Neigung und häufige Übung des Analysierens die herausgehörten Töne allmählich an Intensität gewinnen.

Untersuchungen dieser Art versprechen überhaupt mancherlei Vergleichungs- und Berührungspunkte mit der musikalischen Praxis. So wird der Komponist einen Ton, der nach seinen Erfahrungen im Zusammenklange besonders leidet, den er aber gleich stark mit anderen hervortreten lassen will, eben durch mehrere Stimmen oder ein stärkeres Instrument zu Gehör bringen, usf. Aber um solche Vergleichen mit der Praxis zu ermöglichen, müßten die Untersuchungen noch bedeutend spezialisiert werden, wozu eine bequemere Versuchstechnik erforderlich wäre<sup>1)</sup>.

Wie sich die Ergebnisse dieser Versuche zu den unter 3 erwähnten verhalten, ist noch nicht abzusehen, da diese eben fast nur an mir selbst gewonnen sind. Ich vermute aber, daß auch für Individuen, bei denen der stärkere Ton den schwächeren bis auf wenige Prozente seiner Stärke herabdrückt, die größeren Regelmäßigkeiten sich ähnlich herausstellen würden.

Jedenfalls ist nichts davon bekannt, daß in dem Charakter der Vokale und sonstigen Klangfarben, wo es sich ja nicht um herausgehörte Teiltöne, sondern um einheitliches Hören handelt, solche individuellen Unterschiede sich geltend machten. Nun wäre es aber eine mehr als paradoxe Annahme, daß durch das Heraus hören selbst, also durch die auf einen Teilton konzentrierte Aufmerksamkeit, dieser Ton bis auf nahezu Null geschwächt würde, da doch im Gegenteil die Aufmerksamkeit, soweit sie überhaupt Einfluß auf die Stärke gewinnt, nur verstärkend wirkt. Man kommt daher, scheint mir, zu einer etwas komplizierten Vorstellung, die sich zunächst etwa in folgendem anatomischen (später vielleicht ins Physiologische, Funktionelle zu übertragenden) Bilde wiedergeben läßt.

Mehrere Schichten im Nervensystem könnten unterschieden werden, in denen sich die Intensitätsverhältnisse gleichzeitiger Töne verschieden gestalten. In den peripheren Endgebilden des Acusticus mögen ungefähr die Verhältnisse stattfinden, wie sie in WEGELS und LANES, sowie in meinen Beobachtungen auf-

<sup>1)</sup> Wenn nach den Regeln der Komposition gerade die Terz eines Dreiklanges im allgemeinen nicht durch Stimmenverdoppelung verstärkt werden soll, so wird dies daran liegen, daß der „charakteristische Ton“, der Dur und Moll unterscheidet, als solcher auch bei geringer subjektiver Stärke sich der habituellen Aufmerksamkeit des musikalischen Hörers besonders aufdrängt, und daß die herrschende Stellung des Grundtons, aber auch der „Dominante“, in keinem Falle gefährdet werden darf. Hier sind also ästhetische Gründe maßgebend. Übrigens ist auch Verdoppelung in der Oktave nicht identisch mit Verstärkung.

treten. Und diese Verhältnisse mögen, zunächst nur wenig modifiziert, auch in einer unteren Schicht des Hörzentrums noch erhalten bleiben. Hier wird also dem schwächeren Tone durch den gleichzeitigen stärkeren etwas abgezogen, dem höheren mehr durch den tieferen als umgekehrt. Aber der Abzug beträgt lange nicht so viel wie in den EBERHARDTSchen Versuchen, höchstens so viel als dort bei den Besthörenden. In dieser 2. Schicht wird die Klangfarbe und der Vokalcharakter bestimmt. Dann folgt aber noch eine obere Schicht, in die nur herausgehörte Töne eintreten und wo sich jene großen individuellen Unterschiede geltend machen. Beim Übergang in diese 3. Schicht erfolgt eine individuell sehr verschiedene Einbuße an Energie, bei manchen Individuen unter günstigen Umständen nur eine geringe, bei anderen und unter ungünstigen Umständen aber ein Herabsinken bis fast zum Nullwerte. Das Heraushören ist für dieses Herabsinken eine *conditio sine qua non*, aber ich würde aus dem obenerwähnten Grunde nicht sagen, daß es die Ursache wäre. Diese muß vielmehr in zentralen Widerständen liegen, die sich der vollen Entwicklung der physiologischen Tonprozesse entgegensetzen. Ich möchte diese Schicht auch nicht geradezu als „Apperzeptionszentrum“ gegenüber dem darunterliegenden „Perzeptionszentrum“ bezeichnen, obgleich man der Kürze halber sich dieser Ausdrücke bedienen mag: denn die in meinen und den amerikanischen Untersuchungen gefundenen Gesetzmäßigkeiten setzen doch auch schon Apperzeption, Heraushören der Komponenten voraus. Die beiden letzten Schichten müssen nur eben unterschieden werden, um die EBERHARDTSchen Ergebnisse und die wohl damit zusammenhängenden außerordentlichen individuellen Unterschiede in der Fähigkeit des Heraushörens, in der einheitlichen und mehrheitlichen Auffassung der Musik einigermaßen verständlich zu machen. Für das Verständnis der Vokalstrukturen bietet die EBERHARDTSche Untersuchung keine neuen wesentlichen Anhaltspunkte; sie sollte hier nur als interessante Fortführung der allgemeinen Fragen erwähnt werden, zu denen man sich durch Vokalstudien zuletzt geführt sieht.

## II. Verhältnis der Teiltonstärken zur Gesamtstärke des Klanges<sup>1)</sup>.

Physikalisch setzt man die Gesamtstärke eines Klanges einfach gleich der Summe seiner Teiltonstärken und drückt die letzteren in Prozenten der ersten aus. Bei der Beschreibung der Klang-

<sup>1)</sup> Hierzu vgl. Tonpsych. II, 423ff. und „Attribute der Gesichtsempfindungen“, Abh. d. Berlin. Akad., Phil.-hist. Kl. 1917, S. 46ff.

erscheinungen als solcher liegt aber die Sache ganz anders. Hier ist der Klang als dieser bestimmte Einzelklang mit seinen Klangcharakteren, auch seiner Klangstärke, gerade dann in ausgeprägteste Weise gegeben, wenn die einzelnen Töne nicht herausgehört und für sich beachtet werden; er ist phänomenologisch keineswegs die Summe seiner Teile, und seine Stärke ist nicht die Summe ihrer Stärken. Zwar schreiben wir einem Dreiklang auch dann, wenn uns die Komponenten nebeneinander, jeder in seiner Höhe und Stärke, gegenwärtig sind, immer noch außerdem eine Gesamtstärke zu, er kann als Ganzes forte oder piano klingen. Aber in solchen Fällen richtet sich unser Urteil nach der Stärke der Komponenten und, wenn diese nicht ganz gleich ist, nach dem stärksten unter ihnen. Bei bedeutender Ungleichheit wird man das Urteil überhaupt ablehnen<sup>1)</sup>.

Ich habe nun früher behauptet und halte dies auch heute fest, obschon es fast überall Widerspruch gefunden hat und von einem Berichterstatter als ein wunderlicher Einfall angesehen wurde, daß ein Klangganzes, mag es aus unter sich gleich oder verschieden starken Teilen bestehen, niemals stärker ist als der stärkste seiner Teile im isolierten Zustande. Nur wenn durch Differenztonbildung einer der Teile, bzw. der stärkste, im Zusammenklang noch verstärkt wird, wird auch das Ganze entsprechend stärker.

Diese Behauptung, die ich früher nur an Kombinationen von Klavier- oder Stimmgabeltönen erprobte, habe ich nunmehr auch mit völlig einfachen Tönen nachgeprüft und richtig befunden. Bei solchen Versuchen muß aber sorgfältig darauf geachtet werden, daß man eine vermehrte Fülle oder Breite des Eindrucks, wie sie jeder Tonmehrheit gegenüber den Komponenten eigen ist, nicht mit größerer Stärke verwechsle.

Die Versuche wurden mit Tönen des Vokalapparates angestellt, deren jeder auf möglichste Schwäche (Ebenmerklichkeit) eingestellt war. Die Frage war: ist der Gesamteindruck deutlich stärker als der des einzelnen Tones? Hierzu mußten Töne gewählt werden, die nicht durch Differenztonbildung einen unter ihnen verstärkten (wie dies z. B. regelmäßig bei einfachen Oktaven der

<sup>1)</sup> Übertragungen vom Teil auf das Ganze finden sich auch bei anderen Sinnen. Fühlen wir uns heiß, so bewirkt schon die Abkühlung eines Gliedes, z. B. der Hände, den Eindruck einer Gesamtkühlung, und umgekehrt bei der Kälte. Besonders unter den „geometrisch-optischen Täuschungen“ kommt Analoges häufig vor. Aber die Analogien sind natürlich überall mit Unterschieden verbunden, da jedes Sinnesgebiet seine Besonderheiten hat.

Fall ist, wo dann natürlich auch das Ganze stärker erscheint als jeder Teil), aber auch nicht durch Schwebungen die Hörbarkeit des Ganzen erhöhten. Ich wählte folgende Kombinationen: 1.  $c^1 c^3 c^5$ ; 2.  $g^1 e^2 c^3 g^3 d^4 c^5$ ; 3.  $fis^4 as^4 c^5$ ; 4. dieselben nebst  $e^5$ . Differenztöne werden natürlich auch hier vielfach erzeugt, aber sie fallen nicht mit einem der Primärtöne zusammen und können auch nicht stärker sein als einer von diesen; sie waren im Gegenteil so schwach, daß ich niemals etwas davon wahrnehmen konnte. In den 3 ersten Fällen blieb der Zusammenklang für mich im wesentlichen analysierbar, in den beiden letzten Fällen aber nicht. Bei 3. hat auch Dr. RIEFFERT mitbeobachtet.

Das Ergebnis war stets dasselbe: keine Verstärkung. Das Ganze steht ebenso dicht vor dem Verschwinden, wie jeder Teil. Ich könnte höchstens sagen, daß das Ganze eindrucksvoller wäre, weil es eben voller ist, aber nicht, daß es stärker wäre. Macht man die Teile etwas übermerklich, so trifft das nämliche auch wieder für das Ganze zu: es ist nur in gleichem Grade übermerklich.

Wie paradox auch diese Tatsache vielen erscheinen mag, ist sie doch nicht so unverständlich, wenn man für die einzelnen Töne eine getrennte Lokalisation in der Hirnrinde annimmt, ähnlich wie sie in der Schnecke nach HELMHOLTZ stattfindet<sup>1)</sup>.

Auch wenn man einohriges Hören mit doppelohrigem vergleicht, erfolgt keine wirkliche Verstärkung bei letzterem. Natürlich dürfen hier nur ganz geringe Stärken benützt werden, da bei größeren die Überleitung von einem Ohr zum anderen, sei es durch die Luft- oder die Knochenleitung, schon für jedes Ohr allein Verstärkung bewirken kann. Der Ton wird nur voller, breiter. Tonreize, die für ein Ohr (bei gleicher Hörschärfe, oder bei ungleicher für das besser hörende) unmerklich sind, können nicht durch zweiohriges Hören merklich werden. Diese von mir 1890 gegen TARCHANOFF und PREYER behauptete Tatsache hat v. HORNPOSTEL neuerlich (1, S. 84ff.) durch sorgfältige Untersuchungen bestätigt: „Ist ein Ton definitiv und unzweifelhaft unter die Schwelle gebracht, dann

<sup>1)</sup> Zu demselben Ergebnis führte eine Untersuchung von H. BARKHAUSEN und G. LEWICKI, Die Empfindlichkeit des Ohres für nichtsinusförmige Töne, Physik. Zschr. Jhrg. 25, S. 537ff. 1924; woraus hervorgeht, daß für die Hörbarkeit im Telephon der subjektiv lauteste Teilton eines Klanges maßgebend ist. Wurden die übrigen abgedrosselt, so wurde zwar die Klangfarbe, aber nicht die Hörbarkeit verändert. Ausgenommen blieben natürlich die Fälle der Amplitudensteigerung benachbarter Töne durch Schwebungen, wobei die Autoren einen solchen gegenseitigen Einfluß in der Gegend um 1000 Schw. noch bei einer Differenz von 200 Schw. merklich fanden.

kommt auch keine zweiohrige Erscheinung mehr zustande.“ Auch H. WERNER hat (S. 23) Versuche mit demselben Ergebnis angestellt. Bei übermerklichen Stärken beurteilten allerdings seine Beobachter das zweiohrig Gehörte als stärker. Aber er selbst führt dies auf eine Verwechslung der Stärke mit der „Eindringlichkeit“ zurück, vor der man sich eben hüten muß. Ein wirklich geübter Beobachter wird solchen Verwechslungen, die WERNER euphemistisch als verschiedene „Intensitätserlebnisse“ oder „Erscheinungsformen komplexer Intensitäten“ bezeichnet, selten unterliegen; aber für Anfänger sind gewiß Warnungen am Platze.

Neuerdings behauptet WITTMANN, eine wirkliche Verstärkung mit 2 Telephonen erzielt zu haben<sup>1)</sup>. Aber seine Versuche, die v. HORNBOSTEL nachprüfte, leiden an dem Fehler, daß er die Wechselwirkung der beiden Telephone nicht beachtete, wodurch freilich schon physikalisch eine Verstärkung entstehen muß. Hielt v. HORNBOSTEL ein Telephon vor das eine Ohr, setzte aber das zweite nur neben das (überdies fest verstopfte) andere Ohr an den Kopf, so erfolgte gleichfalls Verstärkung<sup>2)</sup>. Für solche Versuche dürfen selbstverständlich nur gegenseitig ganz unabhängige Tonquellen benützt werden.

Infolge dieses Sachverhalts kann es auch für einen Mehrklang aus Tönen verschiedener Höhe keinen Unterschied in der Stärke machen, ob alle Komponenten ein und demselben Ohre dargeboten oder ob sie an beide Ohren verteilt werden, soweit im letzteren Fall überhaupt eine einheitliche Klangerscheinung zustande kommt.

Mit Vorstehendem ist nun auch alles, was in meinen Analysen und Synthesen der Vokale in Hinsicht der Klangstärke des Ganzen und der Teile zutage trat, in guter Übereinstimmung. Beim Aufbau durch die Interferenzmethode z. B. war der Grundton eines gesungenen Vokals (ebenso der eines obertonreichen Instruments) immer sehr schwach, bei Vokalen auf *C* sogar unhörbar. Trat dann ein Teilton nach dem anderen hinzu, so wuchs fortwährend die Klangstärke des Ganzen, aber nur weil und wenn der Grundton selbst durch Differenztonbildung immer stärker wurde. Das gleiche galt auch von seinen nächsten Obertönen, da auch diese durch Differenztonbildung verstärkt werden (der Ton 2 z. B. durch die Töne 5 und 3, 6 und 4 usf.). Niemals wächst die Klangstärke bloß durch das Hinzukommen neuer Teiltöne, wenn diese nicht etwa selbst schon stärker sind als alle vorherigen. Sie ist stets

<sup>1)</sup> Arch. f. d. ges. Psychol. Bd. 51, S. 70ff. 1925.

<sup>2)</sup> Über das Nähere wird v. H. demnächst in der Z. „Psychol. Forschung“ berichten.

gleich der Stärke des stärksten Teiltons, soweit hier überhaupt Vergleichung möglich ist.

Ausdrücklich sei aber hinzugefügt, daß die hier aufgestellten Behauptungen sich nicht auf Geräusche und ihre Teile beziehen sollen. Hier scheint vielmehr tatsächlich eine Summierung der Teilintensitäten stattzufinden. Die Stärke eines S z. B. ist bedingt durch eine über einen größeren Tonbezirk verteilte Erregung, ebenso die eines Knalles<sup>1)</sup>. Hier bewirkt nun jedes Herausnehmen einzelner kleiner Schwingungsbezirke (durch Stich- oder Lückenversuche) entschieden eine Schwächung des Ganzen. Ebenso war bei hellen Flüstervokalen in den Aufbauversuchen regelmäßig zu bemerken, daß sie jenseits des Formanten, wenn sie also qualitativ schon tadellos waren, immer noch stärker wurden, oft bis nahe an die obere Tongrenze. Verhielte es sich so wie bei den Tönen, so müßte bei Herausnahme einzelner Abschnitte die Gesamtstärke solange ungemindert bleiben, als nicht der stärkste Geräuschteil, also etwa der Formant oder ein Unterformant oder der die Tonhöhe des Geräusches bedingende Teil, davon betroffen wird. Nun ist dies zwar der Fall, wenn der herausgenommene Teil über diesem stärksten liegt, da durch Interferenz auf höhere Töne auch die tieferen physikalisch in gewissem Maße geschwächt werden. Aber soweit meine Beobachtungen in dieser Hinsicht reichen, vermindert bei Geräuschen eben jede If.-Einstellung, auch in tieferen Teilen, die Gesamtstärke.

<sup>1)</sup> Auch bei Tönen werden allerdings von objektiv ganz einfachen Schwingungen nach der Resonanzhypothese infolge der Resonanzbreite eine Anzahl benachbarter Fasern der Grundmembran miterregt, die erst zusammen mit der stärksterregten dem Tone seine subjektive Stärke verleihen. Bei den Geräuschen hat aber auch schon der Reiz selbst immer eine bedeutende Breite.

## Psychophysik der Sprachlaute.

Versuchen wir nunmehr möglichst allgemeine Erklärungsgründe für die beschriebenen akustischen Tatsachen zu finden, so ist wohl begreiflich, daß es nicht ohne psychophysische Hypothesen abgehen kann, da die phänomenale Analyse zuletzt auf gewisse Grundtatsachen führt, die in der Natur und dem Zusammenwirken der letzten Gehirnprozesse wurzeln, auf denen alle Eigenschaften unserer Sinnesempfindungen beruhen, über die aber zunächst nur hypothetische Vorstellungen entwickelt werden können. Selbstverständlich bevorzugen wir unter diesen solche, aus denen die beobachteten Erscheinungen am einfachsten und leichtesten abgeleitet werden können. Zuerst und hauptsächlich kommen wieder die gesungenen und stimmhaft gesprochenen Vokale in Betracht, da hier das umfassendste Material vorliegt. Wir beginnen mit einer Übersicht der wichtigsten Punkte, die dann durch Ausführungen über die U-Formanten noch ergänzt werden.

### 1. Allgemeinste Ergebnisse betreffs gesungener und stimmhaft gesprochener Vokale.

a) Zwischen der Absolut- und der Relativtheorie fällt die Entscheidung prinzipiell zugunsten der ersteren, wie dies übrigens heute fast alle Phonetiker lehren und wie man es nach der Art der Lautproduktion (gleichbleibende Kiefer- und Mundstellung bei ein und demselben Vokal auf verschiedenen Tonhöhen) von vornherein erwarten mußte. Doch ist an der Relativtheorie richtig, daß jedem Vokal gewisse Struktureigentümlichkeiten zukommen, die in allen Tonlagen, wenn auch nicht überall gleich ausgeprägt, wiederkehren. Namentlich unterscheiden sich hierin die dunklen und die hellen Vokale, wovon die ersteren relativ wenige Teiltöne, und zwar von vorwiegend niedrigen Ordnungszahlen, besitzen, die letzteren aber zahlreiche Teiltöne, unter denen die mit höheren Ordnungszahlen eine mit der Helligkeit zunehmende Rolle spielen. Ferner ist es richtig, daß mit steigender Tonhöhe des Grundtones die Vokalformanten merklich, wenn auch nur langsam und keineswegs in gleichem Schritte mit dem Grundton, in die Höhe rücken.

Endlich ist auch dies an der Relativtheorie richtig, daß U überhaupt keinen festliegenden, sondern einen mit dem Grundton veränderlichen Formanten hat, nämlich den Grundton selbst.

b) Die gewöhnlichen, nicht nasalierten Vokale bestehen in allen irgendwie wesentlichen Teilen ausschließlich aus harmonischen Teiltönen (ganzzahligen Multiplis) des Grundtones. Sollten ganz hohe unharmonische Töne gelegentlich beteiligt sein, so sind sie doch ohne Bedeutung für die unterscheidenden Vokalcharaktere. Nur bei den Nasales könnten unter den hohen Beiltönen, die zu dem näselnden Charakter wesentlich mitwirken, allenfalls auch unharmonische vertreten sein.

c) An dem jeweiligen Vokalcharakter sind nicht alle Teiltöne gleichermaßen beteiligt. Sehr wichtig sind die untersten, am entscheidendsten aber die des „Formanten“, dessen Definition eben darin besteht, daß er die jeweilig wesentlichsten Teiltöne umfaßt. Er wird in der Regel nicht durch einen einzigen Ton gebildet, sondern, mit Ausnahme einiger weniger Fälle, aus einer Mehrzahl von Teiltönen, deren Stärke mit der Höhe zu- und dann wieder abnimmt. Die Lage des Maximums („Formantenzentrums“) verschiebt sich mit dem ganzen Formanten bei aufsteigendem Grundton innerhalb bestimmter enger Grenzen, dem „Formanten im weiteren Sinne“ oder der Formantegend, die ihrerseits eine fest bestimmte Lage im Tonreich besitzt.

Nur dann, wenn schon der 2. Teilton (die Oktave des Grundtones) in den Formanten fällt, genügt dieser einzige hinzukommende Ton zur Entstehung einer deutlichen Vokalität. So entsteht O, wenn zum Grundton  $c^1$  seine Oktave in genügender Stärke hinzutritt; so A, wenn zu  $c^2$   $c^3$  hinzukommt. Dies hängt offenbar mit der besonderen Stärke und dem hohen Verschmelzungsgrad des 2. Teiltones zusammen.

Handelt es sich um Teiltöne sehr hoher Ordnungszahl, so sind sie in einem guten Vokalklang nicht lückenlos vorhanden, sondern nur in geringer Zahl und in Abständen, die allzu starke Rauigkeiten ausschließen.

d) Außer den Hauptformanten gibt es in bestimmten, häufigen Fällen Nebenformanten. U, O, A haben in unmittelbarer Nähe der Klangquelle bei starker Intonation, U auch bei schwacher, obere Maxima, die wir aber nur beim U, wo ein solches immer vorhanden ist, als Oberformanten bezeichnen; die hellen (von Ö an) haben stets Unterformanten. In beiden Fällen sind die Nebenformanten von den Hauptformanten durch relativ oder absolut leere Tonstrecken getrennt. Außerdem ist aber auch die Gesamtheit der übrigen Teiltöne, soweit sie eine gewisse Stärke überschreiten, nicht



unbeteiligt an der jeweiligen Klangfarbe. Ein U auf  $c^1$  kann nicht so klingen wie ein U auf  $c$ . Bei A, Ö, Ä sind besonders auch die unmittelbar unter dem Formanten liegenden Tonstrecken (Brücken) durchaus unentbehrlich; die Kontinuität der Teiltöne darf hier nirgends unterbrochen sein.

Zuweilen, namentlich in der Nähe der Klangquelle, sind noch mehr als 2 relative Maxima unterscheidbar (vgl. S. 21, 25ff. 28, 180). Diese Unterteilungen sind aber nicht von allgemeinerer Bedeutung.

e) Was die relative Stärke der Teiltöne betrifft, so ist der Grundton häufig, besonders beim A, objektiv sehr schwach oder gar nicht vorhanden, wird aber als Differenzton der Obertöne subjektiv verstärkt. Außerdem finden überhaupt im Ohr (und im Gehirn?) Verschiebungen der objektiven Stärkeverhältnisse zugunsten der tieferen Teiltöne statt. Daß der Formant den Hauptteil oder auch nur den Höhepunkt der physikalischen Klangenergie in sich schließt, ist im allgemeinen anzunehmen und durch Messungen vielfach erhärtet; aber es erscheint nicht als unbedingt notwendige Forderung, vielmehr bleibt denkbar, daß in manchen Fällen schon seine relative Stärke gegenüber den umgebenden Tönen in Verbindung mit seiner absoluten Lage im Tongebiet und mit der sonstigen Struktur des ganzen Klanges den Ausschlag gibt.

f) Jede nicht allzu winzige Änderung in der Stärke eines Teiltones innerhalb der Haupt- oder Nebenformanten, aber auch der darunterliegenden Töne, besonders des 1. und 2. Teiltönen, bedingt eine Änderung der Vokalfärbung, und zwar wird diese dunkler, wenn tiefere, heller, wenn höhere Teile verstärkt werden. Insofern läßt sich sagen, daß bei gegebenem Grundton eine bestimmte Vokalnuanze nur durch eine einzige bestimmte Kombination von Teiltönen in bestimmten Stärkeverhältnissen hervorgebracht werden kann. Niemals gibt es hier äquivalente Zusammenstellungen, etwa in der Weise, daß statt eines Tones ein etwas höherer zusammen mit einem etwas tieferen eintreten könnte. Dagegen wird durch Ausfüllung der toten Strecken der hellen Vokale mit schwachen Teiltönen und durch die über dem Formanten liegenden Teiltöne des A die Vokalität nicht merklich verändert und sind Kombinationen mit und ohne diese Bestandteile vokalisch gleichwertig. Nur die Stimmfärbung kann dadurch geändert werden.

g) Jeder Vokal durchläuft, wenn er experimentell ab- oder aufgebaut wird, bestimmte Rückbildungs- und Entwicklungsstufen, entsprechend dem jeweiligen Bestand an Teiltönen. Diese Stadien sind aber in dem unversehrten bzw. voll entwickelten Vokal durch unmittelbare Beobachtung nur sehr bruchstückweise

zu erkennen. Sie sind nicht alle darin „aufgehoben“ in HEGELSCHEM Doppelsinne des Wortes. Man kann z. B. U nicht wohl als einen wahrnehmbaren Teil des Ä bezeichnen, obschon es bei seinem Abbau schließlich übrig bleibt und beim Aufbau die Grundlage bildet. Ebenso wenig ist O eine Teilerscheinung des A, obschon dieses durch Hinzufügung weniger Teiltöne daraus erwächst. Immerhin lassen sich vielfach Entwicklungsstufen auch in dem fertigen Vokal wiedererkennen, namentlich hört man bei den hellen Vokalen bei darauf gerichteter Aufmerksamkeit stets den dunklen O- oder U-artigen Unterformanten (auch bei den Flüster-vokalen, hier z. B. sehr deutlich beim Ä das AO). Aus dem stimmhaften A glaube ich den tiefsten Bestandteil, seine U-Grundlage, jetzt gleichfalls herauszuhören. Aber niemals würde es gelungen sein, auf diesem Wege die Gesamtheit der Auf- und Abbauprodukte festzustellen.

h) Alle diese Rückbildungs- und Entwicklungsstadien sind aber selbst wieder Vokale, wenn sie gesondert dargeboten werden: eine keineswegs selbstverständliche Tatsache. Nur durch Lückenversuche werden zuweilen instrumental gefärbte, insbesondere klarinetten- oder fagottähnliche Klänge erzeugt.

i) Auf Grundtönen über  $c^2$ , teilweise schon von  $g^1$  an, verlieren die Vokale mehr und mehr ihre Charakteristik und sind von  $c^3$  an überhaupt so gut wie ununterscheidbar.

k) Desgleichen ergibt sich und wird durch die Erfahrung bestätigt, daß auf tieferen Grundtönen die Vokale nicht immer gleich deutlich gesungen werden können. Für U und I vgl. o. S. 255, u. S. 315ff. O gerät am vollkommensten auf Tönen zwischen  $g$  und  $c^1$ , weil da die starken ersten Obertöne  $g^1$  bis  $c^2$  in den Formanten fallen. U. s. f.

l) Stimmhaft gesprochene Vokale sind von gesungenen in Hinsicht der Zusammensetzung aus Teiltönen bei gleicher Deutlichkeit der Aussprache nicht verschieden. Sie haben dieselben Formanten und weisen beim Ab- und Aufbau dieselben Stufen auf; aber sie haben, obschon sich im allgemeinen mehr Geräusch beimischt, durchschnittlich größere Deutlichkeit. Hierzu trägt der Umstand bei, daß in der lebendigen Sprache die Tonhöhe oft schon innerhalb einer einzigen akzentuierten Silbe stetigen Schwankungen unterliegt und so die Formantregion gewissermaßen abgestreift wird und die Valenzen eines Vokals vollständiger zur Geltung kommen (ähnlich schon KRUEGER, MARTENS und PIPPING). Dazu kommt aber, daß der Sprechende, der in erster Linie vom Hörer verstanden werden will, automatisch auf deutlichere Aussprache, und gleichzeitig der Hörende auf schärferes Erfassen dieser

Unterschiede eingestellt ist als beim Singen, wo die Intentionen des Sängers wie des Hörers in erster Linie den musikalischen Eigenschaften gelten.

m) Wie sich die Formanten der deutschen Vokale und die Grenzen der stimmhaften und stimmlosen Sprachlaute in die wichtigsten Orts- und Grenzbestimmungen des Tonreiches einfügen, möge schließlich folgendes Schema veranschaulichen. Es bedarf nicht der Bemerkung, daß die Orte der Formanten hier nur ganz im Groben angezeigt werden können.

Orts- und Grenzbestimmungen des Tonreiches.

	n	Formanten	
$fis^7$	23410		Obere Hörgrenze
$c^7$	16554		
$fis^6$	11705		Obere Grenze des normalen Altersgehörs
$c^6$	8277		Obere Grenze der Konsonanten und Flüstervokale
$fis^5$	5853		
$c^5$	4138	) I	Obere Grenze der Musik
$fis^4$	2926	) E	
$c^4$	2069	) Ä, Ö, Ü	Obere Grenze der Singstimme
$fis^{3:}$	1463	) A	
$c^3$	1035		
$fis^2$	732		
$c^2$	517	) O	{ Mitte des Tonreiches Untere Grenze der Flüstervokale
$fis^1$	366		
$c^1$	259	U	
$fis$	183		
$c$	129	↓	Untere Grenze der Konsonanten
$C$	65		Untere Grenze der Singstimme
$C_1$	32		Untere Grenze der Musik
$C_2$	16		
$Fis_3$	11		Untere Hörgrenze.

Die obere Grenze der zur Melodienbildung verwendbaren Töne liegt bei  $c^5$ , da jenseits dieser Grenze die Intervalle rasch undeutlich und von  $g^5$  an überhaupt unerkennbar werden. Bei der Orgel werden allerdings, wie mir Herr Orgelbaumeister KLAIS (Bonn) mitteilt, in größeren Werken jetzt die Pfeifen „gemischter Stimmen“ gelegentlich noch bis  $a^6 = 14\ 000$  Schw. geführt, aber sie wirken da eben nur als Obertöne ihrer tieferen Oktaven und in Verbindung mit diesen. Ähnlich ist es an der unteren Grenze, wo die Töne von 32 Schw. abwärts bis 16 Schw. nurnmehr in Verbindung mit ihren höheren Oktaven gebraucht werden, die sie aber auch schon als Obertöne (oft kräftiger als die Grundtöne selbst) mit sich führen. Die menschliche Stimme geht in der Höhe zuweilen noch bis  $c^4$ , in der Tiefe bis  $A_1 = 54$  Schw. Die untere Hörgrenze ist, wie die obere, nur schwer genau festzustellen. Sie scheint unter günstigsten Umständen noch etwas unter  $C_2$ , bei etwa  $Fis_3 = 11$  Schw. zu liegen. Danach umfaßt das Tonreich in extremen Fällen rund 11 Oktaven, wovon aber nur 7 musikalisch ver-

wertbar sind. Die Mitte liegt in  $c^2$ , wenn anders das logarithmische Gesetz für die Beziehung zwischen Tonhöhen und Schwingungszahlen gültig ist, wie dies E. H. WEBER und FECHNER behaupteten und noch kürzlich ABRAHAM und v. HORNBOSTEL auf einwandfreiem Wege bestätigten (Z. f. Psych. Bd. 89, S. 233ff., im Gegensatz zu WUNDTs ganz unmöglicher Behauptung, daß gleichen Tonabständen gleiche Differenzen der Schwingungszahlen entsprächen).

## 2. Die Formanten des stimmhaften U.

In bezug auf diesen Vokal weichen wir von der gewöhnlichen Auffassung dadurch ab, daß wir ihm einen beweglichen Formanten zuschreiben, insofern also die Relativtheorie vertreten<sup>1)</sup>. Außer diesem beweglichen Hauptformanten enthält er aber einen festen Nebenformanten. Wir wollen diese beiden Thesen, Vorhererwähntes teils zusammenfassend, teils ergänzend, jetzt noch näher begründen.

Alle einfachen Töne unterhalb  $g^1$  klingen für sich allein schon U-artig. Aber es ist nicht zu leugnen, daß die der kleinen Oktave dem natürlichen U am nächsten kommen. Den Grund suche ich darin, daß die Männerstimme, die das U am dunkelsten gibt, sich vorzugsweise auf Grundtönen dieser Lage bewegt. Geht ein besonders tiefer Baß im Gesange bis zu  $C$  herunter, so liegt doch der starke und objektiv beträchtlich stärkere 1. Oberton in der kleinen Oktave und wirkt mit dem Grundton zusammen formierend. Beim Aufbau durch If.-Röhren war  $C$  für sich allein überhaupt nicht zu hören, beim Grundton *Ges* war dieser zwar hörbar, klang aber für sich allein überdampf und erhielt erst durch den Hinzutritt des *ges* einen natürlichen U-Charakter (o. S. 56ff.).

Mit Rücksicht darauf könnte man also immerhin von einem festen Formanten sprechen, der dann etwa von  $c$  bis  $c^1$  reichen müßte, denn alle Töne dieser Zone machen durch das Hinzutreten zu ihren tieferen Oktaven diese zu einem natürlicheren U. Aber der Unterschied von dem Grundton-U ist doch nur graduell: die Dampfhöhe des Eindruckes wird nur gemildert. Es ist nicht eine so spezifische Umwandlung, wie wenn ein U durch den Hinzutritt des O- oder A-Formanten zum O oder A wird. Die Dampfhöhe hat ja auch bei dem empirischen U schon sehr verschiedene Grade, namentlich ist das der Frauen- und Kinderstimme bedeutend heller, weil die Grundtöne sich schon dem  $g^1$  nähern oder es überschreiten. Deswegen scheint es mir die Tatsachen richtiger auszudrücken, wenn wir statt eines festen Formanten in der kleinen Oktave einfach einen von  $g^1$  abwärts beweglichen, mit dem Grund-

<sup>1)</sup> Soviel ich sehe, hat sich nur SCHOLE (S. 13ff.) in demselben Sinn ausgesprochen.

ton zusammenfallenden Formanten statuieren. Wie wäre sonst auch mit der Tatsache fertig zu werden, daß auf  $g^1$  selbst noch ein recht brauchbares U gesungen und gesprochen werden kann? Ließe man aber etwa den festen Formanten selbst nach oben bis  $g^1$  reichen, so müßte man doch anerkennen, daß hier der Formant mit dem Grundton zusammenfällt, ja daß dasselbe auch bei allen Grundtönen bis hinab zu  $g$  der Fall sein muß. Erst hier würde ein Teilton aus der festen Formantregion, nämlich eben  $g^1$ , in der gewöhnlichen Weise als Oberton zum Grundton hinzutreten. Es ist daher nicht einzusehen, wie man um die Theorie des beweglichen Formanten beim U herumkommen soll.

Eine feste Resonanz der Mundhöhle muß freilich auch da vorhanden sein, obgleich sie nicht leicht zu bestimmen ist. Sie dürfte hauptsächlich bei  $e^1-f^1$  liegen. Ich schließe es daraus, daß die U-Resonanz um einige wenige Töne unter der O-Resonanz liegen muß. Nun wird eine schwingende  $b^1$ -Gabel, vor der zum O-Sprechen geöffneten Mundhöhle gehalten, zweifellos verstärkt gehört (HELMHOLTZ). Also wird die U-Resonanz etwa die obige Lage haben, obgleich beim U selbst das analoge Experiment mit einer  $e^1$ - oder  $f^1$ -Gabel kein deutliches Ergebnis liefert. Man kann auch anführen, daß das natürliche Flüster-U die Tonhöhe des  $f^2$  hat und daß dieses wohl nur als Überblasungsgeräusch des  $f^1$  aufgefaßt werden kann (o. S. 164), woraus sich wieder  $f^1$  als Höhe des Grundtones (Grundgeräusches) ergibt.

Also eine feste Resonanzgegend für U geben wir zu, und zwar dieselbe, die schon DONDERS und in der späteren Zeit auch HELMHOLTZ statuierte. Aber sie ist nicht zugleich die Formantgegend. Sie hat nur die Bedeutung, daß die dadurch verstärkten Teiltöne ein tieferes U besonders sonor und kräftig machen. Daher wird eine Männerstimme diesen Vokal besonders gut um  $f$  herum singen, wo  $f^1$  oder ein Nachbarton als Oberton hinzutritt<sup>1)</sup>. Aber es wäre auch ohne diesen Oberton schon ein U. Wäre  $f^1$  fester Formant oder auch nur Formantzentrum, so wäre ein U nur unter der Bedingung überhaupt möglich, daß Teiltöne aus dieser Gegend im Klange vorhanden wären. Man könnte dann kein U auf  $c^1$  oder auch nur auf  $a$  singen oder sprechen.

Wir bleiben also beim beweglichen Hauptformanten. Dagegen ist zweifellos ein fester Oberformant im gewöhnlichen U

<sup>1)</sup> HELMHOLTZ beobachtete beim Singen des U eine zunehmende Erschütterung im Mund und an den Trommelfellen, wenn er sich dem Grundton  $f$  näherte. Dies führte ihn zunächst dazu,  $f$  als charakteristischen Ton des U anzusehen. Aber es wird wohl mit der besonderen Stärke des in die Resonanzgegend fallenden 1. Obertones zusammenhängen.

enthalten, und zwar ist es die Gegend  $d^2-gis^2$ . Sei es, daß ein einziger Oberton oder daß mehrere in diese Zone fallen: sie sind bei Männer- wie Frauenstimmen ganz regelmäßig verstärkt und beeinflussen die Naturtreue des Vokals so merklich, daß wir sie als Oberformant bezeichnen dürfen. Sie machen den Vokal markiger, körperlicher, metallischer. Die Bedeutung dieser Gegend wurde auch schon früher mehrfach hervorgehoben. So erwähnt NAGEL  $f^2-gis^2$ , SCHOLE  $b^1-g^2$ , ABRAHAM  $d^2$  als regelmäßige Bestandteile des U. Es ist zwar wieder nur eine Bereicherung der Klangfarbe damit gegeben. Aber ein Nebenformant hat ex definitione nicht den spezifischen Vokalcharakter in erster Linie hervorzubringen, wie dies bei dem angeblichen Hauptformanten  $f^1$  der Fall sein müßte; er hat nur zur vollen Naturtreue mit-zuwirken.

Man kann diesen Oberformanten sehr leicht direkt beobachten, wenn man U auf Tönen zwischen  $d$  und  $h$  singt. Die Obertöne in der Gegend von  $d^2-gis^2$  treten dabei sehr stark heraus (selbst während einer Bahnfahrt habe ich unter starkem Geräusch diese Beobachtung machen können). Auch hierin liegt ein Grund, warum U in der kleinen Oktave besonders gut gesungen wird. Doch kann  $fis^2$  auch noch als 8. Teilton eines tiefen Fis im U kräftig vertreten sein<sup>1)</sup>. Man kann den Versuch auch so machen, daß man nach dem Singen des U mit möglichst gleichbleibender Mundstellung einen Pfeifton erzeugt. Wenn ich die 4 Töne  $B, es, g, b$  singe, so erhalte ich immer den Pfeifton  $es^2$  oder  $f^2$ .<sup>2)</sup>

### 3. Die U-I-Linie des Vokaldreiecks.

Nehmen wir nun zuerst an, es gäbe nur solche Vokale, die auf der U-I-Linie des Vokaldreiecks liegen. Dann würden wir im wesentlichen ohne weitere Erklärungsmittel mit der bekannten Eigenschaft der „Tonhöhe“ auskommen. Diese ist das Analogon der Helligkeit bei den Farben, und die Reihe der Tonhöhen von den tiefsten bis zu den höchsten ist eine Anordnung nach ihren Helligkeitsgraden. Diese Helligkeitsgrade sind aber zugleich in der Hauptsache die Vokalcharaktere der U-I-Reihe. Es gilt hier ohne Zweifel, daß schon die einfachen Töne einen

<sup>1)</sup> Es sind übrigens in allen diesen Fällen auch sehr hohe Teiltöne aus der 4-gestrichenen Oktave nicht allzu schwer herauszuhören. Sie machen den Eindruck, als kämen sie aus der Nase.

<sup>2)</sup> Dieser Pfeifton braucht natürlich nicht ein harmonischer Teilton des gesungenen zu sein, da die Resonanz des Lippenraumes, wenn er für sich angeblasen wird, nicht genau dieselbe ist, als wenn er in Verbindung mit der ganzen Mundhöhle steht.

Vokalcharakter besitzen. Die tiefen bis  $g^1$  klingen wie U, die hohen, etwa von  $c^4$  an, wie I, die dazwischen liegenden gleichen den Übergängen auf der Vokallinie, UÜ und Ü, von denen der erste in der deutschen Sprache nicht vorkommt, aber an sich möglich und anderwärts auch vorhanden ist. Allerdings decken sich die einfachen Töne nicht völlig mit den betreffenden Vokalen. Das U der Töne unterhalb  $c$  klingt unnatürlich dumpf und zerflossen, das I oberhalb  $c^5$  unnatürlich hell und spitz, auch die dazwischen liegenden Töne sind nicht geradezu identisch mit empirischen Vokalen. Aber es bedarf in der mittleren Region, wie sie dem durchschnittlichen Umfang der menschlichen Stimme entspricht, nur relativ geringer Zusätze, um die Töne diesen Vokalen gleichzumachen.

Über diesen Punkt herrscht unter den Phonetikern eine seltene Übereinstimmung, wenn sie auch meistens keine ganz einfachen Töne benützten. An den Pfeiftönen des Mundes, die keine Obertöne enthalten, kann man sich leicht überzeugen, daß die allertiefsten (unter  $c^2$ ) U-, die um  $b^3$  Ü-, die höchsten I-Charakter tragen. Beobachter, in deren Muttersprache UÜ-Laute vorkommen, würden wahrscheinlich in den Pfeiftönen zwischen  $g^2$  und  $c^3$  einen solchen Vokalcharakter wiedererkennen. Mir will dies allerdings nicht gelingen.

Sehr deutlich sind U, Ü, I auch in den subjektiven Tönen des „Ohrenklingens“, die bei manchen Personen sehr häufig auftreten und offenbar einfache Töne sind. Ich habe früher selbst zahlreiche (580) Fälle beobachtet und ihrer absoluten Höhe nach bestimmt (m. Abh. 5), aber auch inzwischen immer wieder solche Beobachtungen machen können und dabei meine Aufmerksamkeit speziell auf den Vokalcharakter gerichtet. Er entsprach genau dem obigen Bilde. Diese Töne reichen nach der Tiefe nur bis etwa  $c^1$ , selten noch ein wenig tiefer, und tragen hier U-Charakter, sind aber äußerst schwach und erscheinen, vielleicht aus eben diesem Grunde, noch weit dumpfer und zerflossener als die einfachen objektiven Töne gleicher Höhe. Dagegen sind die häufigen subjektiven Töne der 3- bis 5-gestrichenen Oktave scharf ausgeprägt und hell und die von  $f^4$  bis  $h^4$  geben ein sehr reines I<sup>1)</sup>. Da mein

<sup>1)</sup> Ein Ultra-I kann ich gegenwärtig sogar willkürlich jederzeit durch Verschieben des Unterkiefers subjektiv erzeugen. Es entsteht dann ein scharfer, im rechten Ohre lokalisierter Ton (vermutlich durch Zerrung in der Kette der Gehörknöchelchen), dessen Höhe ich durch Vergleichung mit der Galtonpfeife = 9000 Schw. ( $d^6$ ) bestimmte. Er gleicht diesem objektiven Ton auch vollkommen in seinem Vokalcharakter: ein sozusagen transzendentes I. Wie ich bei K. L. SCHAEFFER (Verh. d. dtsh. otolog. Ges. 1909, S. 15) lese, ist dieselbe Erscheinung auch schon von anderen bemerkt worden.

absolutes Tonbewußtsein im letzten Jahrzehnt infolge abnehmender Übung im Musizieren (wovon es zweifellos im höchsten Grade abhängig ist) an Sicherheit abgenommen hat, diente mir bei subjektiven Tönen der Vokalcharakter oft geradezu als Hilfe zur Erkenntnis der Tonlage, die dann durch die genauere Bestimmung am Klavier bestätigt wurde. Man kann die Höhe eines subjektiven Tones aus der I-Gegend nach seinem Vokalcharakter (reines, überhelles, oder mehr Ü-ähnliches I) mindestens innerhalb der Fehlergrenzen einer Quarte bestimmen; eine noch genauere Bezeichnung wird wohl kaum gelingen. Beim Ü ist die Zone enger. Hört man einen subjektiven Ton mit ausgesprochener Ü-Qualität, so kann man sicher sein, daß er bei  $a^3-h^3$  liegt, wie denn auch beim Pfeifen eines Ü fast stets  $b^3$  genau getroffen wird.

Stellt man durch das Interferenzverfahren einfache Töne her, wobei man weiter in die Tiefe gehen kann, so findet man U am natürlichsten durch die Gegend von  $c-c^1$  repräsentiert. War bei unseren If.-Versuchen der isolierte Grundton eines gesungenen Vokals (einerlei welches) tiefer als  $c$ , so wurde er zwar immer als U bezeichnet, erlangte aber seine Vollkommenheit als U erst durch den Hinzutritt eines Obertones aus der kleinen oder I-gestrichenen Oktave. Noch bis  $ges^1$  tragen die einfachen Töne guten U-Charakter. Dagegen lautete das Urteil bei dem isolierten Grundton  $c^2$  zumeist: „eigentlich gar kein Vokal“, oder nur: „U-ähnlich“.

Man kann auch genetisch und entwicklungsgeschichtlich verstehen, was zur Erzeugung und Ausbildung der Lautreihe U-I führte. Die Sprech- und Singwerkzeuge des Menschen gestatten eine doppelte Veränderung in Hinsicht der erzeugbaren Tonhöhen: man kann die Stimmlippen des Kehlkopfes zur Erzeugung verschiedener Grundtöne innervieren, und man kann bei unveränderter Grundtonhöhe die Kiefer-, Zungen- und Lippenstellung verändern, wodurch verschiedene Teiltonbezirke verstärkt und die Klangfarben verändert werden. Diese doppelte Möglichkeit mußte zu den ersten Erfahrungen gehören, sobald die Stimme zum Gefühlsausdruck oder zu Mitteilungszwecken in Funktion trat. Bei Klangveränderung durch Resonanzverstellung richtete sich die Aufmerksamkeit natürlicherweise auf die beiden Extreme der dunkelsten und hellsten Lautgebung, und so entstanden U und I. Da wir aber einfache Töne nicht durch den Kehlkopf hervorbringen können und die Lippenöffnung beim U notwendig den Oberformanten, die Mundhöhlenresonanz beim I den Unterformanten aus dem durch den Kehlkopf erzeugten Klange verstärken, so mußten die wirklichen Laute eben diese Zusammensetzung erhalten. Aber diese Zusatzteile modifizieren die Charak-



tere der einfachen Töne, die hier die Formanten ausmachen, nur wenig. Das U wird voller, markiger durch seine Obertöne, das I gleichfalls voller, körperhafter, gesättigter durch seine Untertöne: Unterschiede, wie sie allen Klängen gegenüber einfachen Tönen zukommen. Es sind nur leichte Über- und Untermalungen, die den Eindruck extremer Dunkelheit und Helligkeit unter den uns nun einmal zu Gebote stehenden Stimmklängen nicht verwischen.

Gleichwohl sind diese leichten Modifikationen phänomenologisch nicht selbstverständlich, sondern vom allgemeinen Standpunkt einer Theorie der Komplexeigenschaften durchaus zu beachten. Man könnte zunächst erwarten, daß die Helligkeiten der Unter- und Oberformanten sich mit denen der Formanten mischten, daß also U durch seinen Oberformanten einfach und schlechtweg heller, I durch seinen Unterformanten dunkler würde. Dem ist aber nicht so. U rückt nicht einfach in der Helligkeitslinie hinauf, I herunter, sie werden nicht zu einem mittleren  $\ddot{U}$  oder  $U\ddot{U}$ , ebensowenig wie ein Ton durch Hinzufügung eines höheren oder tieferen zu einem mittleren Tone wird. Die Veränderung ist vielmehr, so unbedeutend sie sein mag, doch etwas prinzipiell Neues, Eigenartiges. Wir haben sie mit metaphonischen Ausdrücken aus dem Raum- und Farbengebiete geschildert. Vielleicht wäre es am zutreffendsten, von einem Helldunkel zu sprechen, da es eben eine Komplexeigenschaft aus der Gattung der Helligkeiten ist. Als Komplexeigenschaft ist sie nicht streng kommensurabel mit den Helligkeiten der einfachen Töne, ohne einen Ort auf ihrer Linie. Wir legen darauf Gewicht, weil schon hier der psychophysische Tatbestand auftaucht, der uns im folgenden noch unverkennbarer entgegentritt: daß die Komplexeigenschaften, obschon sie mit denen der Elemente zusammenhängen, doch nicht restlos aus ihnen begriffen werden können, sondern als direkte Wirkungen besonderer zentral-physiologischer Prozesse angesehen werden müssen, die durch das gleichzeitige Auftreten mehrerer einfacher Prozesse ausgelöst werden.

#### 4. Die Vokale außerhalb der U-I-Linie.

##### KOEHLERS Vokalitäten.

Es ist offenbar unmöglich, Laute wie A, Ä, Ö als bloße Abstufungen oder Verknüpfungen von Gliedern der U-I-Reihe abzuleiten. Sie sind nicht bloße Abwandlungen oder Verhüllungen des U, Übergänge zwischen U und I, auch nicht bloß reichere, vollere Glieder dieser Reihe, sondern etwas qualitativ anderes. Hier verspricht nun zunächst die KOEHLERSche Lehre (1) weiterzuführen. Danach besitzen die einfachen Töne, zunächst die von

$c$  bis  $c^6$ , von vornherein außer ihrer Höhe noch eine Vokalqualität (Vokalität) als immanente Eigenschaft, und zwar ist diese die eigentliche Prinzipalqualität des Tonreiches, da sie wirklich qualitative Unterschiede aufweist, ähnlich dem System der getönten („bunten“) Farben, während die Höhenunterschiede nur eine stetig in gleicher Richtung veränderliche Abstufung besitzen. Zwischen  $c$  und  $c^6$  liegen die Vokalitäten U O A E I in dieser Reihenfolge. Jede erstreckt sich über 2 Oktaven, hat aber ihr Zentrum und ihre schärfste Ausprägung in der Mitte dieses Spielraumes: U bei  $c^1$ , O bei  $c^2$ , A bei  $c^3$ , E bei  $c^4$ , I bei  $c^5$ <sup>1)</sup>. Die zwischen diesen „ausgezeichneten Punkten“ liegenden Töne tragen Übergangscharakter (UO, OA usw.), der je nach der Lage dem einen oder anderen Grenzvokal ähnlicher ist. Das folgende Schema mag dies veranschaulichen:

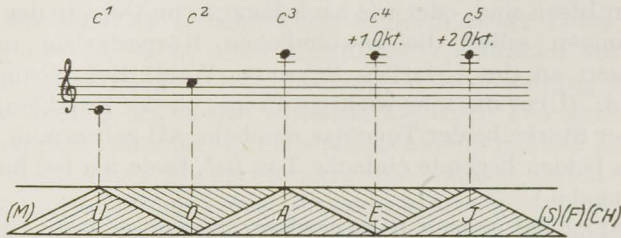


Abb. 5. Schema der Vokalitäten nach KOEHLER.

Auch diese Lehre ist in mehr oder weniger ähnlicher Form schon früher aufgetaucht; und daß sie mehrfach von gegenseitig unabhängigen Forschern aufgestellt wurde, läßt darauf schließen, daß etwas Richtiges daran sein muß. Zuerst findet sie sich bei WILLIS. Seine Pfeifentöne waren freilich nicht in unserem Sinne einfach. Dann hat 1870 R. KOENIG<sup>2)</sup> für die fünf „Hauptvokale“ bereits ein Oktavengesetz aufgestellt, veranlaßt durch den Umstand, daß 3 unter HELMHOLTZ' charakteristischen Tönen ( $b^1$ ,  $b^2$ ,  $b^3$ ) im Oktavenverhältnis standen. Er setzte alle Zentren infolgedessen um einen Ton tiefer als später KOEHLER. Stimmgabeln von entsprechender Höhe, die er mit den Vokalbuchstaben bezeichnete, finden sich noch in Laboratorien. KOEHLER ist aber ganz unabhängig von KOENIG durch psychologische Versuche zu seiner Aufstellung gekommen. Aus bloß statistischen Zusammenstellungen glaubte auch TREBS (Arch. f. d. ges.

<sup>1)</sup> Wir setzen diese musikalischen Notenbuchstaben für die von KOEHLER angegebenen Schwingungszahlen (I, II, S. 130). Diese stammen von 4 Beobachtern und liegen mit einer einzigen Ausnahme sämtlich zwischen 261 und 264 bzw. den entsprechenden Multiplis. 261 ist aber =  $c$  für  $a^1 = 435$ , 264 ist =  $c$  für die ältere Stimmung  $a^1 = 440$ . Daher ist die obige einfache Bezeichnungsweise berechtigt.

<sup>2)</sup> Quelques expériences etc. S. 42ff, 64. Die Originalabhandlung in den C. R. der Pariser Akademie 1870.

Psychol. Bd. 14, 1909) ein Oktavengesetz erschließen zu können. Auch v. WESENDONK fand Vokalähnlichkeiten bei seinen Flaschentönen. U reicht ihm bis  $b^2$ , A findet er etwa bei  $d^3-g^3$ , E bei  $as^3$ , I bei  $b^3-f^4$ . ROUSSELOT gab ebenfalls schwache Vokalähnlichkeiten einzelner Gabeltöne zu. Ebenso HENSEN. GRASSMANN hatte dagegen nur U-, Ü- und I-Charaktere in der Tonreihe gefunden. Auch LAHR hörte sogar noch eine Gabel von 1000 Schw. ( $c^3$ ) als helles U; erst mit tieferen zusammen wurde sie zu A. Eine Zwischenstellung nimmt JAENSCH ein. Er erklärt: „Niemals habe ich reine Töne gehört, die in wirklich überzeugender Weise wie A oder E geklungen hätten. Damit verträgt sich sehr wohl die Tatsache, daß es immer möglich ist, einen reinen Ton herauszufinden, der dem reinen A oder dem reinen E ähnlicher ist als alle anderen Töne.“

Die empirischen Vokale mit ihrer komplizierten Zusammensetzung aus Teiltönen wären hiernach nur Annäherungen an die reinen Urvokale, Übergänge oder Mischungen, in denen nur etwa ein Element besonders überwiegt, so wie nach PLATON die Dinge der Erscheinungswelt nur unvollkommene Abbilder der an sich einfachen Ideen sind, oder wie nach HERING im Gebiete der Sinneserscheinungen selbst die gewöhnlichen Körperfarben nur Annäherungen an die Urfarben Schwarz, Weiß, Rot, Grün, Gelb, Blau sind. (Über die sehr wichtige Frage, ob der Zweiklang  $c^2+c^3$  in gleicher Stärke beider Töne das nämliche AO geben muß wie der zwischen beiden liegende einfache Ton  $fis^2$ , finde ich bei KOEHLER keine Angabe.)

KOEHLER hat jedoch bald gewisse Zusätze gemacht, wodurch sich die Lehre der älteren Auffassung nähert, indem er z. B. darauf hinweist, daß der volle Charakter des E und I nur durch Beifügung tieferer Töne zu erreichen sei, und daß zu einem guten A alle Teiltöne, die überhaupt eine A-Valenz besitzen (also in dem Raume  $c^2-c^4$  liegen), zusammenwirken müssen (1, III, S. 96). Ja, er dürfte gegenwärtig nichts gegen die Behauptung einwenden, daß gerade die empirischen Vokale, gut ausgesprochen, die reinen und urbildlichen seien. In der neuesten Veröffentlichung hierüber (3, S. 449) schreibt er den einfachen Tönen, ähnlich wie JAENSCH, eine geringere „vokale Sättigung“ zu als den empirischen Vokalen und vergleicht den Fall geradezu mit dem von Papieren, in denen ein bedeutender Grauehalt überall die spezifische bunte Nuance in hohem Grade verhüllt. Was er dagegen festhält, ist die These, daß in der Reihe der einfachen Töne, und zwar relativ am stärksten an den bezeichneten Punkten, sich weitgehende Ähnlichkeiten mit den Vokalen finden und daß diese Vokalitäten in der Theorie der empirischen Vokale eine entscheidende Rolle spielen müssen.

In der Tat: wären die Vokalitäten der einfachen Töne zugleich die idealen Urvokale, vergleichbar den HERINGSchen Urfarben, wie man nach der ersten Fassung der Lehre annehmen müßte,

so sollte man erwarten, daß die richtige Einordnung in das System der Grundqualitäten bei Normalhörenden und Geübten keine Schwierigkeiten machte. Versuche mit Farben, bei denen die analoge Aufgabe gestellt war, eine gegebene Farbe den HERINGSCHEN Urfarben zuzuordnen, bzw. als Übergangsfarbe (Blaugrün usf.) in ihre Reihe einzuordnen, ergaben mir bei 9 normalsichtigen Personen so gut wie volle Übereinstimmung (unter 230 Urteilen nur 5 abweichende). Anders aber Versuche, die ich 1913 mit zwei der besten unter KOEHLERS Beobachtern über die Vokalität beliebig herausgegriffener, ganz oder nahezu einfacher Töne machte (8, S. 331ff. bzw. 43ff.). Sie ergaben arge Divergenzen, derart, daß beispielsweise  $c^3$  mehrfach als U bezeichnet und auch sonst oft Vokale genannt wurden, die in KOEHLERS Schema bis zu 2 Oktaven tiefer stehen. Die einfachen Pfeiftöne des Mundes wurden von der Vp. v. M. unter 12 Fällen, die sich von  $c^2-d^4$  erstreckten, 6mal als  $\ddot{U}$  bezeichnet, obgleich dieser Vokal, wie ihr bekannt war, gar nicht unter KOEHLERS Vokalitäten vorkommt: zwar ein gutes Zeugnis für unbefangene Beobachtung und uns wohlverständlich, da es sich jedesmal um einen Ton zwischen  $fi^3$  und  $d^4$  handelte, aber nicht in Übereinstimmung mit der KOEHLERSCHEN These. Auch Nachprüfungen von anderen Seiten lieferten bisher keine Bestätigung des Oktavengesetzes<sup>1</sup>).

Die überraschende Exaktheit, mit der die „ausgezeichneten Punkte“ in KOEHLERS eigenen Versuchen von den Vp. getroffen wurden und die selbst JAENSCH das Oktavengesetz als unbezweifelbar richtig erscheinen ließen, hängt wohl mit der besonderen, von ihm gewählten Methode zusammen, insofern er nicht die Aufgabe stellte, die richtige Vokalbezeichnung beliebig herausgegriffener Töne zu finden, sondern die Vp. die Umgebung eines seiner Zentren in auf- und absteigendem Verfahren absuchen und den optimalen Punkt eines vorgegebenen Vokals bestimmen ließ. Es bleibt gleichwohl zu erklären, worauf die fast bis auf die Schwingungszahl übereinstimmende Fixierung der in Oktaven übereinander liegenden Punkte durch die verschiedenen Vp. beruhte. Wir müssen uns des Urteils darüber enthalten; nur eine Wiederholung der Versuche mit entsprechenden Variationen der Umstände könnte Aufklärung schaffen<sup>2</sup>). Ich wagte einstweilen die Vermutung (8),

<sup>1</sup>) RICH S. 121ff., 136, 142. Nach privater Mitteilung auch K. HUBER (München).

<sup>2</sup>) Besonders dürfte von Einfluß sein, innerhalb welcher Tongrenzen man Töne vorlegt; z. B. wären, wenn das beste A bestimmt werden soll, etwa folgende Strecken zu versuchen:  $g^1-c^3$ ,  $c^2-d^3$ ,  $f^2-g^3$ ,  $b^2-g^3$ ,  $b^2-c^4$ . Vermutlich würde sich das Optimum mehr oder weniger mit der Strecke

daß die musikalische Auszeichnung des Tones  $\mathcal{C}$  in allen Oktaven für unser heutiges Bewußtsein nicht ohne Einfluß auf das Urteil der Beobachter gewesen sein könnte. Obgleich diese kein absolutes Tonbewußtsein besaßen und zum Teil sogar sehr unmusikalisch waren, scheint mir ein solcher Einfluß nicht ganz undenkbar. Gerade bei Unmusikalischen könnten absolute Höhen maßgebender sein, während für Musikalische die bloßen Tonverhältnisse wichtiger sind. Die Ordnung der Vokale nach ihrer Komplexhelligkeit würde dann ihre Verteilung auf die verschiedenen  $\mathcal{C}$ 's bewirken. Jedenfalls ist es sehr unwahrscheinlich, daß das Oktavenverhältnis der KOEHLERSchen Vokalitäten ganz unabhängig wäre von dem doch unzweifelhaft bestehenden Oktavengesetz der musikalischen Qualitäten, und ist auch die genaue Koinzidenz der Schwingungszahlen mit denen der  $\mathcal{C}$ 's auffallend. Aber hier mögen diese Vermutungen auf sich beruhen.

Nach KOEHLER und v. MALTZEW soll allerdings gerade das musikalische Oktavengesetz in der Höhe eine Ausnahme erleiden: von  $fis^4$  bis  $c^5$  wurden nämlich die Töne von Stimmgabeln fast immer zu tief nachgesungen (natürlich in der eigenen Stimmlage) und die Intervalle mit kleineren verwechselt. Daraus wird geschlossen, daß der seiner Schwingungszahl nach als  $c^5$  bezeichnete Ton vom Gehör als  $h^4$  oder  $b^4$  vernommen werde. Dieses „normale Falschhören“ zeigte sich aber nach KOEHLER nicht bei den Vokalitäten, indem das beste I nicht bei  $b^4$  oder  $h^4$ , sondern genau bei  $c^5$  gefunden wurde.

Indessen ist die Unterschätzung von Intervallen aus Stimmgabeltönen auch schon in der mittleren Tonlage ein allgemeinerer Zug. Die Oktave erscheint auch hier fast wie eine Septime. Auch der Verfasser, v. HORNBOSTEL und andere musikalische Beobachter finden sich dadurch immer aufs neue überrascht. Sogar die Mitglieder des JOACHIM-Quartetts erklärten bei solchen Versuchen die physikalisch reine Oktave mit großer Regelmäßigkeit für zu klein, obschon sie natürlich auf ihren Instrumenten die Oktaven in Doppelgriffen mit tadelloser Reinheit erzeugten<sup>1)</sup>. Dieses Verhalten nimmt nach ABRAHAM und v. HORNBOSTEL bei Doppel- und Tripeloktaven noch zu; und um solche multiple Oktaven handelt es sich ja bei der Transposition in die eigene Stimmlage, auch wenn sie unbewußt erfolgt. Die Erscheinung wurzelt wahrscheinlich darin, daß die Komplexhelligkeit gegebener Klänge nicht einfach proportional der Schwingungszahl des Grundtones zunimmt, weil die Obertöne immer schwächer werden gegen den Grundton und zuletzt überhaupt wegfallen. Auch bei Stimmgabelklängen handelt es sich wegen der darin enthaltenen Oktave um Komplexhelligkeiten. Dadurch wird nun aber das Urteil über das Intervall irreführt. Bei hohen Pfeifen (die man noch in der 5-gestrichenen Oktave mit Differenztönen gut abstimmen kann) muß sich dasselbe geltend machen,

selbst verschieben. Es wären dann aber auch die besten  $\ddot{U}$ ,  $\ddot{A}$ ,  $\ddot{O}$  zu suchen. Die Töne jeder Strecke wären in einer Reihe in auf- und absteigender Folge, in einer anderen aber auch ganz bunt durcheinander vorzulegen.

<sup>1)</sup> STUMPF und MEYER, Maßbestimmungen über die Reinheit konsonanter Intervalle. Zeitschr. f. Psych. Bd. 18, S. 364ff; m. Beitr. z. Akustik H. 2, S. 127 ff.

und so war es auch in v. MALTZEWS Versuchen. Gegen die untere Tongrenze hin kehrt sich der Sachverhalt und damit auch das Urteilsergebnis um: die Obertöne nehmen an Stärke gegen den Grundton immer mehr zu, und die Höhe des Klangganzen wird überschätzt. Von einem Falschhören würde ich aber in allen diesen Fällen nicht sprechen, zumal bei längerer Fortsetzung der Versuche und gesteigerter Konzentration auf den Fragepunkt das Urteil gegen den physikalisch richtigen Ton zu konvergieren scheint. Vgl. auch m. Abh. 8, S. 319ff. (31ff.).

Die Formantzentren, wie sie nach unseren und vielen anderen Untersuchungen nunmehr als festgestellt gelten dürfen, liegen für O, A und I um einige Töne tiefer als KOEHLERS Zentren. Nur das für E fällt ziemlich genau mit  $c^4$  zusammen (ganz wenig darüber). Beim U, dem wir einen beweglichen Formanten zuschreiben, liegen die optimalen Grundtöne auch mehr nach unten als nach oben von  $c^1$ .

Auch viele Details unserer Beobachtungen stimmen nicht mit KOEHLERS These überein. So Resonanzversuche mit Knabenstimmen, in denen A  $c^3$  nur äußerst schwach oder gar nicht vertreten war (o. S. 30, 200). So die dominierende Stellung des  $g^2$  bei Lückenversuchen (o. S. 73ff.). Unter den Synthesen weisen nur die auf  $c^2$  ein A mit einem leichten Maximum bei  $c^3$  auf: aber hier steht eben, wenn nur harmonische Obertöne zugelassen werden, überhaupt kein tieferer zur Verfügung.

Wenn gleichwohl wiederholte Versuche unter entsprechenden Variationen der Umstände KOEHLERS ausgezeichnete Punkte bei einfachen Tönen bestätigen sollten, so müßte man eben annehmen, daß bei isolierten einfachen Tönen die Vokalqualitäten sich aus irgendeinem Grunde gegen die Lage der Formantzentren innerhalb der empirischen Vokalklänge verschieben (oder umgekehrt). Dies wäre an sich gewiß denkbar. Aber die Ursache dieser Verschiebung bliebe zunächst dunkel.

Große Schwierigkeiten machen der Lehre aber auch die „Umlaute“. Ä (offenes E) ließe sich zwar als Übergang zwischen den benachbarten Urvokalen A und E, als einfacher Ton von der Höhe etwa des  $g^3$ , verstehen. Ü müßte allenfalls als Mischung von U und E (I), Ö als Mischung von O und E definiert werden. Aber schon diese doppelte Erklärungsweise (teils Übergangs- teils Mischform) erregt Bedenken. Außerdem kann ich überhaupt keinen prinzipiellen akustischen Unterschied zwischen den 5 sog. Hauptvokalen und den Umlauten finden. Auch würde sich fragen: wenn U und E (I) sich zu Ü mischen, warum mischen sich nicht auch A und I oder O und I, also die Urvokalitäten von  $c^3$  und  $c^5$  oder von  $c^2$  und  $c^5$ ?

Bei den Konsonanten, auf die KOEHLER das Oktavengesetz ausdehnte (nach unten auf M, nach oben auf S, F, Ch pal.), stellten

sich uns noch viel stärkere Abweichungen heraus. Daß F und Ch sogar über die gewöhnlich angenommene Hörgrenze bis zu zwei Oktaven hinaus liegen sollen, haben wir unbedingt ablehnen müssen.

Ist es sonach unmöglich, diese Lehre in der Form des Oktavengesetzes unverändert zu übernehmen und die Tatsachen der empirischen Vokalitäten daraus zu erklären, so läßt sich doch nicht leugnen, daß in den einfachen Tönen gewisse Anklänge an Vokalcharaktere vorhanden sind. Betreffs der U—I-Linie ist dies bereits hervorgehoben. Aber auch für A sind in meinen Versuchen einzelne Äußerungen vorgekommen, die darauf hindeuten, obgleich ich keine Versuchsreihen mit dieser Fragestellung (außer den vorhin erwähnten) angestellt habe. So antwortete ein Mitbeobachter, dem ich bei den Synthesenversuchen zwischen Kombinationen einmal den einfachen Ton  $c^3$  allein mit der Frage nach dem Vokaleindruck vorlegte, zunächst zwar: „kein Vokal“, setzte dann aber hinzu: „höchstens allenfalls dem A ähnlich“. Auch mir scheint ein gewisser A-Schimmer über dieser Gegend zu liegen. Als einmal alle Teiltöne eines auf  $c^1$  gesungenen A außer  $c^3$  ausgeschaltet waren, schien mir dieses auch für sich allein einen Anflug von A-Charakter zu haben. Kurz vorher, als noch der Grundton ganz schwach mitklang, notierte ich: „Ao, wie aus der Ferne“. Doch möchte ich als wissenschaftlicher Beobachter meinem Urteil hier nicht unbedingt trauen, besonders wenn es sich um  $a^2$  handelt, wo selbst der Notenbuchstabe von Einfluß sein könnte. KOEHLER, der kein absolutes Tonbewußtsein besitzt, versichert, daß er gerade durch das Beachten der Vokalcharaktere in den Stand gesetzt wurde, über die absolute Höhe mit einiger Sicherheit zu urteilen. Es mögen also hier individuelle Unterschiede in der Deutlichkeit der Ausprägung obwalten, ebenso wie solche bezüglich der „musikalischen Qualität“ stattfinden, die dem absoluten Tonbewußtsein im gewöhnlichen Sinne zugrunde liegt.

Bei ganz Unmusikalischen (zu denen aber KOEHLER nicht gehört) dürfte der Vokalcharakter auffälliger in die Erscheinung treten, während die „musikalische Qualität“ da bis zum Nullwert herabsinken kann. Einen ganz extremen Fall hat KOEHLER selbst (I, III, S. 54ff.) an einem Arzte beobachtet, für den alles Musikalische im Tonbereiche, also Intervalle, Melodien usw., „böhmische Dörfer“ waren. Die Vokale konnte er aber tadellos unterscheiden, ordnete sie ihrer Helligkeit nach richtig und gab bei Flaschen- und Gabeltönen die Höhenlage der 5 Vokale in sehr guter Übereinstimmung mit den KOEHLERSCHEN Aufstellungen an. Auch das A wurde in ziemlich engen Grenzen um  $c^3$  herum lokalisiert.

Dem Institutsgehilfen KAUMANN, der von diesem Typus auch nicht weit entfernt ist, legte ich am Vokalapparat die Töne von  $c^4$  abwärts mit der Frage vor, welcher davon dem A am ähnlichsten sei. Er fand  $d^3$  als solchen. Weiter hinunter schien ihm  $h^2$  schon zu O-ähnlich, noch mehr  $b^2$ ,  $g^2$ . Wurde nun von tiefen Tönen begonnen, so erschien  $g^1$  noch zu O-ähnlich, ebenso  $c^2$ ; dagegen wurden jetzt schon  $e^2$  und noch lebhafter  $g^2$  als gutes A bezeichnet.  $b^2$  erschien jetzt zu hell, nach I hin,  $e^3$  viel zu hell. Es war Sorge getragen, daß der Beobachter nichts von meinen Manipulationen zur Herstellung des Tonwechsels sehen konnte. Also beim absteigenden Verfahren wurde  $d^3$ , bei aufsteigendem  $g^2$  als A-Zentrum angegeben. Der Unterschied ist aus den allgemeinen Erfahrungen bei psychologischen Versuchen begreiflich. Die Mitte  $b^2$  ( $h^2$ ) würde mit unserem Formantzentrum nahe übereinstimmen. Jedenfalls waren für diesen Beobachter, der sich durch jahrelange Mitwirkung bei meinen Versuchen ein sehr gutes Gehör für Vokalfarben angeeignet hat, unstreitig A-Ähnlichkeiten in den einfachen Tönen zwischen  $e^2$  und  $e^3$  vorhanden. Als die Töne  $c^1$   $g^1$   $c^2$   $g^2$   $c^3$  in unregelmäßiger Reihenfolge gegeben wurden, bezeichnete er  $c^1$  und  $g^1$  als U,  $c^2$  mehr als UO,  $g^2$  und  $c^3$  als dunkleres und helleres A, beide aber als gleich gut.

An solchen unmusikalischen Personen mit ausgeprägtem Vokalitätsurteil läßt sich auch der interessante Versuch anstellen, den G. E. MÜLLER einmal in Analogie zu den bekannten Farbmischungsversuchen vorschlug: die stetige Verschiebung des Intensitätsverhältnisses zweier Oktaventöne bis zum völligen Verschwinden des einen. Mischt man Rot und Blau auf dem Farbkreis unter Verschiebung ihrer Sektorenbreiten, so kann man die eine Farbe in die andere durch Violett hindurch stetig überführen. Wie verhalten sich die Tonfarben und überhaupt die Eigenschaften des Tongemisches in solchem Falle? Nur mit völlig einfachen Tönen und stetiger Stärkeregulierung, wie sie unsere Einrichtung ermöglicht, läßt sich dieser Versuch rein ausführen. In Hinsicht der Stärke wurde im vorigen Kapitel sowohl die gegenseitige Beeinflussung der Komponenten beim analysierenden Hören, als die Modifikationen der Gesamtstärke beim nichtanalysierenden schon besprochen. Über die Höhenveränderung des gehörten Klanges kann ein Unmusikalischer und nicht Analysierender sich nur in der Weise äußern, daß er ihn mit bestimmten anderen, z. B. denen des Tonmessers, vergleicht und identifiziert. Hierüber Versuche zu machen ist aber nach sonstigen Erfahrungen kaum nötig: er wird, wenn es sich z. B. um  $c^2 + c^3$  handelt, solange einer der beiden Töne weit überwiegt, dessen Tonhöhe angeben, außerdem aber in seinem Urteil mehr oder weniger schwanken. Es wird ihm ein Klang erscheinen, den er sowohl mit  $c^2$  als mit  $c^3$  identifizieren kann.  $c^2$  wird aber viel länger bevorzugt werden, weil das  $c^2$  des Tonmessers auch schon  $c^3$  in beträchtlicher Stärke enthält. Aber niemals dürfte ein Zwischenton als Tonhöhe des Komplexes angegeben werden, es sei denn zufällig von solchen, die überhaupt jedes akustischen Urteils bar sind und wie Taube ihre Aussagen aufs Geratewohl abgeben. Hier liegt der unüberbrückbare Gegensatz des Tonsinnes zum Farbensinn: es gibt keine Tonhöhenmischungen<sup>1)</sup>.

Wie aber verhalten sich in solchem Falle die Vokalfarben? Darüber wurden mit Hrn. KAUMANN folgende Versuche gemacht. Nachdem er  $c^3$  als ein

<sup>1)</sup> Auch mit der „binauralen Mischung“ von RÉVÉSZ steht es nicht anders. Vgl. Zeitschr. f. Psych. Bd. 75, S. 330ff. 1916; auch in m. „Beitr. z. Akust. u. Musikwiss.“ H. 9.



helles A,  $c^2$  als ein etwas dunkles O beurteilt hatte, wurde  $c^2$  zuerst allein kräftig angegeben, dann  $c^3$  mit ganz langsam von Null an wachsender Stärke beigefügt, dann, nachdem es gleichfalls eine bedeutende Stärke erreicht hatte,  $c^2$  langsam bis zum völligen Verschwinden geschwächt, so daß jetzt  $c^3$  allein übrigblieb. Der Hörschlauch wurde in 2 Arme geteilt, so daß ich gleichzeitig mit dem Beobachter, der auf die Vokalitätsänderungen achten sollte, die durch Drehen der Regulierschrauben bewirkten Stärkeveränderungen kontrollieren konnte.

Das Ergebnis war, daß K., als für mein Ohr  $c^3$  eben merklich wurde (bei  $c^2 = 2$ ,  $c^3 = 1/4$ ), eine Veränderung des Ou nach A hin bemerkte, welche dann immer weiter schritt. Bei  $c^2 = 2$ ,  $c^3 = 1$  war es schon „mehr A als O“. Als dann  $c^2$  geschwächt wurde, blieb das Urteil „A nach O hin“, bis für mich  $c^2$  im Zusammenklange ganz verschwunden war. An diesem Punkte urteilte K.: „jetzt helles A“.

Dieses Ergebnis ließ sich im allgemeinen wohl erwarten. Es entstand nun die weitere Frage: wie, wenn statt  $c^2$   $c^1$  mit  $c^3$  mit solcher stetigen Verschiebung der Intensitäten vorgelegt wird?  $c^1$  wird für sich allein von K. entschieden als U-ähnlich bezeichnet. Sollte nun, wenn auch hier keine Analyse stattfand, ein U direkt, ohne O zu durchlaufen, in A übergehen und rückwärts? Diese Erwartung hegte ich nicht, sondern vermutete, daß im Fall der Nichtanalyse ein Laut entstehen würde, der in den mittleren Stadien gleichzeitig als U- und A-ähnlich beurteilt würde, aber eine gewisse Uneinheitlichkeit besäße, entsprechend den Erfahrungen mit U und I bei den Interferenzversuchen (S. 56ff) und den Synthesen (S. 181 Anm.). Aber der Erfolg war nur die sofortige Analyse. Sobald für mich  $c^3$  neben dem starken  $c^1$  merklich wurde, konstatierte auch K. das Auftreten eines hohen Tones. Hier war also keine Veränderung des Vokalurteils zu erzielen. Man müßte zu diesem Versuche wohl noch unmusikalischere Individuen, ähnlich etwa dem KOEHLERSchen Arzte, heranziehen.

In dieser reduzierten Form also, die auch der gegenwärtigen Fassung von KOEHLERS Lehre nahesteht, erkennen wir die Vokalitäten der einfachen Töne an. Sie sind nur eben „in hohem Grade verhüllt“. Und zwar würde ich die U-I-Charaktere der einfachen Töne als das Verhüllende bezeichnen, analog wie bei den Farben das Grau verhüllend, die Sättigung mindernd wirkt.

Nun könnte man sich zunächst vorstellen, daß diese Vokalähnlichkeiten einfacher Töne nichts weiter wären als Erinnerungen an die durch die betreffenden Töne hauptsächlich bedingten empirischen Vokalcharaktere zusammengesetzter Klänge. Wie das Blatt an den Baum, der Teil an das Ganze, so erinnere auch ein dem Formantzentrum benachbarter einzelner Ton an den vollen Vokal. Man würde dann der KOEHLERSchen Lehre ähnlich gegenüberstehen wie ARISTOTELES der platonischen Ideenlehre: nicht die empirischen Vokale wären mehr oder minder vollkommene Abschattungen der idealen Urvokale, sondern diese wären nur mehr oder minder blasse Erinnerungsbilder der allein ursprünglichen und wirklichen empirischen Vokale.

Indessen, so steht es doch auch nicht. Der Teil erinnert nur dann an das Ganze, wenn er, sei es auch nur sehr nebenher, bemerkt worden ist. Obertöne z. B. erinnern keineswegs an die Klänge, als deren Teile sie gegeben wurden, wenn sie ganz und gar unbemerkt geblieben waren. Es wäre also zwar möglich, daß solche Erinnerungen sich z. B. bei mir selbst im Zusammenhange meiner synthetischen Versuche an die einfachen Töne knüpften. Aber für Beobachter ohne solche habituelle Aufmerksamkeitsrichtung würde die Erklärung nicht zutreffen.

Somit kommen wir zu dem Schlusse, daß tatsächlich in dem Tonbezirk zwischen etwa  $g^1$  und  $g^4$  ein im allgemeinen zwar sehr schwaches, bei manchen Menschen aber (und wohl besonders bei Unmusikalischen) besser ausgeprägtes Spektrum von Vokalfarben sich über die einfachen Töne lege. Diesen kommt also außer ihrer Höhe, Stärke usf. auch noch eine Vokalität oder Farbe als immanente Eigenschaft zu. Jede Tonhöhenveränderung innerhalb dieser Zone ist, wenn sie eine gewisse Schwelle überschreitet, zugleich eine Vokalitätsänderung. Im weiteren Sinne kann man auch die U-I-Charaktere der einfachen Töne, die sich nach unten und oben anschließen, als Vokalfarben bezeichnen. Hier fallen sie eben mit den Helligkeiten zusammen. Im engeren Sinne farbig nennen wir nur die davon abweichenden, oberhalb der Grundlinie des Vokaldreiecks liegenden Vokalitäten. Unter ihnen scheint die A-Vokalität die relativ deutlichste zu sein. So machen sich die Eckfarben des Vokaldreiecks auch in dieser Hinsicht bemerklich.

Die U- und I-Charaktere schließen sich aber nicht etwa nur von außen an diese Zone nach beiden Seiten an, sondern sind als entsprechende Helligkeiten der einfachen Töne auch innerhalb dieser Zone vorhanden. Sie werden nur hier im allgemeinen nicht mit Vokalen identifiziert; nur bei  $b^3$  ist die Ü-Ähnlichkeit frappant, d. h. der Vokal Ü wird auch schon durch einen einfachen Ton mit starker Annäherung wiedergegeben. Doch bleibt eine gewisse Ähnlichkeit mit U den Tönen oberhalb  $g^1$  noch eine gute Strecke weit erhalten (vgl. o. S. 322 v. WESENDONK und LAHR), und eine gewisse I-Ähnlichkeit beginnt auch schon unterhalb  $g^4$ . Die Töne dieser mittleren Zone lassen eben Ähnlichkeiten nach 2 Seiten hin erkennen, z. B.  $c^3$  eine U- und eine A-Ähnlichkeit, die erste nach dem Helligkeitsgrad, die zweite nach der Farbigkeit im prägnanten Sinne. Daher tritt denn auch für verschiedene Beobachter, ja für denselben Beobachter, bei verschiedenen Gelegenheiten bald die eine, bald die andere Ähnlichkeit in den Vordergrund und wird für ihre Aussagen bestimmend.

Bereits in meiner „Tonpsychologie“ hielt ich es für notwendig, die Klangfarben auf Tonfarben zurückzuführen, also auch den völlig einfachen Tönen Farbenunterschiede zuzuerkennen. Aber es schien mir damals noch möglich, diese Unterschiede aus denen der Höhe, des Volumens und der Stärke herzuleiten. Wir sehen jetzt, daß die Anerkennung der Farbigkeit als eines Attributes neben diesen anderen nicht zu umgehen ist, und es bleibt KOEHLERS Verdienst, nachdrücklich darauf hingewiesen zu haben.

Ob dieses neue Attribut der einfachen Töne entwicklungsgeschichtlich ebenso ursprünglich ist wie Höhe und Stärke, kann gleichwohl gefragt werden. Es könnte in einer anderen Weise, als es nach der Erinnerungshypothese der Fall wäre, die Folge der empirischen, an zusammengesetzte Klänge gebundenen Vokalcharaktere sein. Darauf werden wir zurückkommen, nachdem erst für diese selbst eine genügende Erklärung gefunden ist.

Denn aus dem Vokalitätenspektrum der einfachen Töne allein lassen sich die empirischen Vokalcharaktere noch keineswegs verstehen. Wie sollen diese so ausgeprägten, aufdringlichen, für jeden Normalhörenden ohne weiteres erkennbaren und unterscheidbaren Charaktere auf einem so schwachen, für viele Menschen nur mit Mühe oder gar nicht erkennbaren Vokalitätenschimmer beruhen? Vor allem müßten doch, wenn das Formantzentrum, z. B. für A  $g^2$  oder  $a^2$ , seine Urvokalität dem Klangganzen aufdrängte, die Vokalitäten der übrigen, keineswegs verschwindend schwachen, manchmal wahrscheinlich sogar physiologisch stärkeren Teiltöne gleichfalls in den Gesamteindruck des Lautes eingehen. Wo bliebe dann die Ähnlichkeit der empirischen mit den Urvokalen? Ich kann daher auch nicht zustimmen, wenn KOEHLER (3, S. 444) das Vorhandensein der Vokalitäten bei einfachen Tönen als eine notwendige Konsequenz der HELMHOLTZschen Klanglehre bezeichnet.

Hier müssen in allen Fällen nicht so ganz einfache Prozesse und Gesetze walten. Diese Prozesse können aber nicht psychologische sein. Ich wüßte nicht, welche bekannten Gesetzmäßigkeiten aus dem der psychologischen Erfahrung zugänglichen Gebiete sich heranziehen ließen, um aus der Kombination oder Wechselwirkung der Eigenschaften der Teiltöne eines Klanges die empirischen Vokalcharaktere O, A, Ä, Ö, E herzuleiten. Wir müssen daher den physiologischen Weg, und zwar nach Lage der Dinge in diesem Gebiete den Weg physiologischer Hypothesen beschreiten.

## 5. Zwei zentralphysiologische Prozesse.

Der Hypothese, die nun entwickelt werden soll, legen wir wieder das Vokaldreieck (S. 252) zugrunde. Darin kann der Ort jedes Vokals durch 2 Bestimmungsstücke eindeutig definiert werden: durch den Punkt auf der U-I-Linie, zu welchem eine Vertikale von ihm aus hinführt, und durch die Höhe dieser Vertikalen selbst. Betrachten wir daher die Sache psychophysisch, so kann man als physiologische Grundlage des Vokals in der Hirnrinde das Zusammenwirken zweier Prozesse für jeden der dabei beteiligten Töne ansehen. Beide Prozesse sind parallel den Schwingungszahlen veränderlich. Dem einen entspricht die mit der Schwingungszahl zunehmende Helligkeit (Höhe) der Tonempfindung. Er ist seiner Quantität nach unveränderlich (kann es wenigstens sein), seiner Intensität nach durch die physiologische Intensität der augenblicklichen zentralen Erregung definiert. Wir können ihn als U-I-Prozeß oder Grundprozeß bezeichnen. Zu diesem tritt aber innerhalb der Grenzen von etwa  $g^1$  bis  $g^4$  ein zweiter, entsprechend der Vertikalen im Vokaldreieck, der im Zusammenwirken mit jenem eine von der U-I-Reihe abweichende Vokalfärbung des Lautes bedingt. Dieser akzessorische Prozeß möge nach der stärksten Abweichung auch als A-Prozeß bezeichnet werden. Wird ein Ton für sich allein erregt, so kommt dieser Prozeß nur in ganz geringem Maße oder gar nicht zustande. Wirkt er aber mit dem Grundton bzw. noch anderen tieferen Teiltönen zusammen, so macht er die entstehende Vokalität um so unähnlicher der U-I-Linie, je höher die Vertikale des Vokaldreiecks für den fraglichen Ton ist<sup>1)</sup>.

Wie zwingend man zu dieser Vorstellung des Sachverhalts gedrängt wird, lehrt allein schon die sonst ganz paradoxe Tatsache, daß jeder der einfachen Töne  $g$  und  $g^1$

für sich deutlichen U-Charakter hat, während ihr Zusammenklang im richtigen Stärkeverhältnis ein deutliches O ist. Also durch das Hinzutreten eines tieferen, dunkleren Tones ( $g$ ) kann aus einem U ( $g^1$ ) ein O, aus dem dunkleren der hellere Vokal

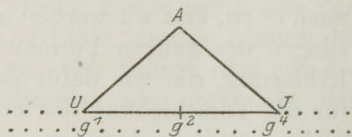


Abb. 6. Grundprozeß und akzessorischer Prozeß.

<sup>1)</sup> Statt des geradlinigen Dreiecks kann man, in Anlehnung an die naturwissenschaftlichen Gewohnheiten, ebensowohl eine Kurve denken, die dann auch allenfalls nach beiden Seiten asymptotisch zur Abszisse verlaufen mag.

werden. Es muß dadurch eine im  $g^1$  latente O-Qualität ausgelöst werden. Man möchte an eine Art katalytischer Wirksamkeit denken. So ist es aber bei allen Teiltönen eines Vokals außerhalb der U-I-Reihe. Bereits KOEHLER hat deshalb mit Rücksicht auf ähnliche Erfahrungen von einer „Chemie der Vokale“ gesprochen.

$g^3$  klingt für sich allein gewiß nicht Ö-artig, am ehesten noch einem dunklen Ü ähnlich, entsprechend der Lage des Tones auf der U-I-Linie. Also der Vokalcharakter des einfachen  $g^3$  ist fast nur durch den Grundprozeß bestimmt. Aber als Oberton in einen passenden Komplex eingefügt, erteilt es diesem die überzeugende Ö-Färbung. Ja, wir hörten, daß selbst ein I-Klang, dem durch ein zufälliges Versehen  $g^3$  beigemischt war, dadurch einen auffallenden Stich nach Ö hin bekam (S. 181 Anm.). Viele andere hierhergehörige Tatsachen liefern die Lückenversuche (2. Kap.).

Der akzessorische Prozeß entfaltet also erst bei Verknüpfung eines Tones mit tieferen und mit einem Grundton, in welchem der Grundprozeß ausschließlich oder fast ausschließlich vertreten ist, der also unterhalb  $g^1$  oder wenigstens unterhalb  $c^2$  liegt, seine volle Wirksamkeit.

Die Einwirkung tiefer Töne auf die Vokalvalenzen höherer scheint innerhalb des Bereiches der menschlichen Stimmtöne geradezu mit ihrer Tiefe zu wachsen. Darauf weist die Tatsache hin, daß nach den Interferenzversuchen der Beginn des A-Formanten für den Grundton  $C$  bei  $c^2$  liegt, für den Grundton  $c$  aber bei  $e^2$ , für  $c^1$  bei  $g^2$ , obgleich das A auf  $c$  und auf  $c^1$  den Teilton  $c^2$  gleichfalls in nicht unbeträchtlicher Stärke enthält. Allerdings ist das A auf  $C$  dunkler und nimmt an Helligkeit über den Grundton  $c$  nach  $c^1$  zu, und wir werden unten hören, daß das langsame Emporsteigen des ganzen Formanten damit zusammenhängt. Aber die Erklärung, die wir dafür geben werden, dürfte nicht ausreichen für die Verschiebung des Formantbeginnes. Diese scheint vielmehr darauf zu beruhen, daß der absolut tiefere Grundton auch aus einem absolut tieferen Oberton bereits einen A-Schimmer hervorzulocken vermag.

Bei den Vokalen außerhalb der U-I-Linie darf überhaupt niemals ein Formantton selbst Grundton sein, wenn der Vokal gut ausfallen soll. Zwar gibt *fis*<sup>1</sup> mit *fis*<sup>2</sup> noch ein leidliches O,  $c^2$  mit  $c^3$  ein leidliches A, indem der Grundton, der am Anfangspunkte der weiteren Formantzone liegt, mit einem kräftigen 1. Oberton, der gegen ihr Ende hin liegt, zusammenwirkt. Aber auf  $b^1$  als Grundton, also auf dem Formantzentrum des O, wird dieser Vokal schon mangelhaft, ebenso A auf  $g^2$ . Ö, Ä, E sind auf ihren Formanttönen als Grundtönen überhaupt unkenntlich.

Übrigens gilt es auch für die U-I-Linie, daß bei den höheren Tönen, die für sich allein Ü- oder I-Charakter tragen, dieser Charakter durch das Hinzutreten tieferer, nämlich ihrer Unterformanten, erst zur vollen Ausprägung kommt. Dies hat bereits KOEHLER bemerkt. Doch bedürfen die Vokale dieser Reihe eines solchen Zusammenwirkens viel weniger, da eben schon die einfachen Töne hier einen viel deutlicheren Vokalcharakter tragen.

Man kann nun in folgender Weise die Qualität der durch den akzessorischen Prozeß bedingten Vokalfärbungen hypothetisch auf eine bestimmte Eigenschaft dieses Prozesses zurückführen. Man kann annehmen, daß er seiner Quantität nach von  $g^1$  bis zur Mitte des A-Formanten,  $g^2$  oder  $a^2$ , zunimmt, dann wieder abnimmt, und daß dadurch die Abweichung der Vokalfärbung von der U-I-Linie, also die spezielle Vokalqualität, die er im Zusammenwirken mit dem Grundprozeß erhält, bedingt ist. Dann ist die jeweilige Vertikale innerhalb des Vokaldreiecks, die wir schon in mehrfacher Beziehung bedeutungsvoll fanden, zugleich das Maß für die Quantität des akzessorischen Prozesses an dem bezüglichen Punkte der Tonreihe.

Dagegen ist der akzessorische Prozeß seiner Intensität nach, und damit auch nach der Stärke der Vokalfärbung, von den Faktoren abhängig, die im vorigen Kapitel untersucht wurden. So bewirkt der Ton  $g^2$ , wenn er als 3. Teilton zu  $c^1$  und  $c^2$  hinzukommt, eine A-Färbung des Klanges in einem seiner Intensität entsprechenden Grade. Tritt auch noch der 4. Teilton  $c^3$  hinzu, so wird das A heller und wieder in einem seiner Intensität entsprechenden Grade. Ebenso beim 5.,  $e^3$ . Aber nun nähert sich die Vokalfärbung schon wieder, zumal wenn der Ton etwas stärker genommen wird, der U-I-Linie, und so nimmt die Ähnlichkeit mit dieser immer mehr zu. Wir führen dies auf die jenseits der Dreiecksspitze wieder abnehmende Quantität des akzessorischen Prozesses zurück.

Es handelt sich hier um spezifische Synergien, wie sie sich auch in anderen Gebieten der Sinneswahrnehmung finden (binokulare Plastik, binaurale Lokalisation, Verschmelzung konsonanter Töne, Zusammenwirken gleichzeitiger Empfindungsprozesse beim Tast- und Temperatursinn usw.).

Man kann den Tatbestand auch so ausdrücken, daß man den zerebralen Tonprozessen zwischen  $g^1$  und  $g^4$  bestimmte physiologische Vokalvalenzen (nach dem Ausdrucke KOEHLERS) zuerkennt, die unter den angegebenen Bedingungen wirksam werden. Diese Valenzen und die ihnen entsprechenden Quantitäten des akzessorischen Prozesses kann man aus dem Vokaldreieck

ablesen, nicht aber irgendwelche relative Intensitäten der Teiltöne. Wir weisen darauf ausdrücklich hin, weil man von den Schwingungskurven her gewohnt ist, die Ordinaten als Amplituden zu deuten. Über die Empfindungsstärken der zu einem guten Vokal gehörigen Teiltöne orientieren unsere synthetischen Tabellen in Verbindung mit den im vorigen Kapitel behandelten Gesetzmäßigkeiten.

Die Bedeutung, die wir der Quantität des akzessorischen Prozesses für die Vokalqualität hypothetisch zugeschrieben haben, könnte man in weiterer Ausgestaltung der Hypothese auch für den Grundprozeß selbst in Anspruch nehmen. Hier wäre die gesetzliche Beziehung sogar eine einfachere: die Quantität des Grundprozesses würde sich ohne Wendepunkt genau parallel mit der Helligkeit des Vokales, die hier zugleich seine Vokalität ist, verändern. Und zwar wäre es am ansprechendsten, zu denken, daß mit zunehmender Helligkeit des Lautes (also abnehmender Wellenlänge des Reizes) die Quantität des Prozesses abnähme. Aber hier würde vorläufig jede mit den Schwingungszahlen parallel laufende Modifikation des zentralen Tonprozesses die gleichen Dienste tun, während bei den im engeren Sinne farbigen Vokalen für das Auf- und Absteigen der Dreiecksvertikalen kaum eine andere Deutung als die gegebene möglich scheint.

Unsere Hypothese darf aber nicht so verstanden werden, als sollten daraus die spezifischen Qualitäten der Vokale in dem Sinne abgeleitet werden, daß man einem, der niemals ein A gehört hätte, deduktiv ein Bewußtsein davon beibringen könnte. Sie soll nur für das Abweichen der Vokale von der U-I-Linie eine physiologisch plausible Unterlage angeben. Warum diese Abweichungen sich uns in der Erscheinung als O, A, E darstellen, wird natürlich überhaupt niemals erklärt werden; es gehört zu den letzten psychophysischen Grundtatsachen, die wir einfach als gegeben hinzunehmen haben. Doch dürfen wir wohl behaupten, daß innerhalb dieser Begrenzung der Aufgabe die hier entwickelten Vorstellungen die Tatsachen richtiger und vollständiger zusammenfassen als etwa das sonst auch so ansprechende KOENIG-KOEHLERsche Oktavengesetz, und daß sie so wenig als möglich über die Beobachtungen selbst hinausgehen. Auch erfüllen sie eine Forderung, die bei der Aufsuchung physiologischer Unterlagen der Phänomene stets beachtet werden muß: daß man Qualitatives so weit als es nur möglich ist auf Quantitatives gründe.

Nur doktrinäre Anhänger eines unklaren „psychophysischen Parallelismus“ werden sich daran stoßen, daß hier einem psychologisch einfachen, für das Bewußtsein in keiner Weise weiter auf-

lösbaeren Tone 2 Gehirnprozesse zugrunde gelegt werden. Dieser vulgäre Parallelismus hat ja überall Löcher und bedarf vor allem einer viel präziseren und spezialisierteren Fassung. Ist doch auch Lila oder Braun phänomenal eine einfache Farbe, welcher physiologisch mindestens 2 Prozesse zugrunde liegen müssen<sup>1)</sup>. Immerhin äußert sich die Doppelheit des physiologischen Vorganges psychologisch darin, daß die Vokalität der Tonempfindungen zwischen  $g^1$  und  $g^4$  in der Tat nicht so eindeutig ist wie außerhalb dieser Grenzen, und daß man, je nach der Richtung der Aufmerksamkeit und individuellen Gradunterschieden in der Stärke der beiden Prozesse, aus demselben Ton, z. B.  $c^3$ , einmal einen Anklang an A, ein anderes Mal nur eine auf der Linie zwischen U und I liegende Vokalität heraushört.

### 6. Besondere Gesetzlichkeiten

hinsichtlich des Zusammenwirkens der Teiltöne.

a) Muß der Formant stets den schlechthin stärksten Teilton des Vokalklanges enthalten?

Diese Frage berührten wir schon o. S. 214. Wenn man die Tabellen der objektiven Messungen durchsieht, wie sie durch Analyse der Schwingungen in ihre Komponenten gewonnen wurden, so kommt man zwar für das A wohl ausnahmslos zur Bejahung der Frage, nicht aber für die helleren Vokale. Hier liegt z. B. in PIPPINGS Tabellen die weitaus größte Intensität für E und I im Unterformanten. Und so auch sonst öfters. Auch in unseren Resonanztabellen und in den synthetischen Tafeln, welche die subjektiven Stärken der isolierten Teiltöne wiedergeben, verhält es sich nicht anders. Der Unterschied zwischen Formant und Unterformant ist dort zwar nicht so groß, aber das Maximum liegt doch ebenfalls für E und I regelmäßig im Unterformanten.

Nun interessieren uns aber hier weder die objektiven noch die subjektiven Stärken der isolierten Teiltöne, sondern die aus ihrem Zusammenwirken im Ohr und Gehirn resultierenden. Hier finden nach dem vorigen Kapitel Verschiebungen der Stärkeverhältnisse zugunsten der tieferen Töne statt. Danach ist anzunehmen, daß bei den hellen Vokalen der Unterformant in den Gehirnvorgängen noch mehr gegen den Formanten hervortritt, daß also die obige Frage für diesen Fall wenigstens verneint werden muß.

<sup>1)</sup> Vgl. u. a. G. E. MÜLLER, Zeitschr. f. Psych. Bd. 10, S. 1ff., ferner meine Arbeit über die Attribute der Gesichtsempfindungen, Abh. d. Preuß. Akad. d. Wissensch. 1917, und EWALDS postume Abwehr willkürlicher erkenntnistheoretischer Voraussetzungen in der Hörtheorie, Zeitschr. f. Sinnesphysiologie Bd. 53, S. 213ff.



Man kann sich dies in folgender Weise zurechtlegen. Wir bezeichnen Rosa in der Regel nicht als rötliches Weiß, sondern als weißliches Rot, auch wenn das Weiß einigermaßen überwiegt. Nur wenn es so sehr überwiegt, daß der rote Anteil fast auf Null herabsinkt, pflegen wir uns umgekehrt auszudrücken. Der Farbenton ist eben für unsere Auffassung das Auffallendere gegenüber der tonfreien Seite. So ist es nun auch recht wohl möglich, daß in einem guten I die dunkle Unterlage, namentlich der Grundton, zentral stärker wäre als der stärkste Ton des Formanten, oder daß der ganze Unterformant stärker wäre als der ganze Formant: trotzdem würden wir es als I in Anspruch nehmen, weil eben schon eine schwache Beimischung aus der Formantengegend, wenn sie nur eine gewisse Schwelle überschreitet, die Auffassung nach dieser Seite determiniert. Wahrscheinlich walten hier aber auch individuelle Unterschiede und gibt es Personen, deren E und I die Unterformanten besonders stark enthalten, vielleicht auch Personen, die sie bei objektiv gleicher Stärke subjektiv stärker als andere hören.

b) Unter den Formanttönen bestimmt der stärkste durch seine spezifische Vokalvalenz die Nuance des Vokals. Die übrigen unterstützen ihn, bringen aber ihre davon abweichenden Valenzen, solange sie die richtige Stärke nicht überschreiten, nicht zur Geltung. Sie passen sich irgendwie dem herrschenden Formantenzentrum an, analog etwa, wie es bei „erzwungenem Mitschwingen“ der Fall ist<sup>1)</sup>. Läßt man aber ihre Stärke über den für eine gute Synthese erforderlichen Grad hinaus wachsen, so werden sie selbständig und verändern die Vokalnumance in ihrem Sinne. Das gleiche ist der Fall, wenn sich das Maximum selbst innerhalb des

<sup>1)</sup> Vgl. Tonpsych. II, 111ff. über die Akkommodation spezifischer Energien, die benachbarten Fasern der Grundmembran zugeordnet sind. Unmittelbare Nachbarschaft in der Schnecke liegt hier freilich nicht vor.

KOEHLER, der die obige Frage schon ins Auge faßt, spricht von einer Addition der gleichen Valenzen und einem Unwirksamwerden der abweichenden. Dies beruht darauf, daß er für jeden Ton 2 farbige Valenzen annimmt, die sich in graduell verschiedener Mischung verbinden, z. B. in der 2-gestrichenen Oktave eine mit der Tonhöhe abnehmende O- und eine gleichzeitig zunehmende A-Valenz (vgl. o. S. 321). Summiert man nun — so scheint er anzunehmen — die in allen Teiltönen des A-Formanten steckenden A-Valenzen (wobei ihre Prozentzahlen doch wohl auch mit einem Stärkekoeffizienten zu multiplizieren sind) und ist die Summe größer als die der O-Valenzen, so werden die letzteren unwirksam. Dies könnte man sich etwa analog denken wie beim Fall auf der schiefen Ebene, wo die zur Ebene senkrechte Komponente infolge der Maschinenbedingungen ausfällt. Mir scheint aber der Tatbestand sowohl in bezug auf die Valenzen als auf ihr Zusammenwirken richtiger in der obigen Weise ausgedrückt zu werden.

Formanten verschiebt, z. B. beim A von  $g^2$  auf  $c^3$  übergeht: das A wird dann eben heller.

Von Interesse sind hier auch die Lückenversuche mit Herausnahme von Formantttönen aus dem A (S. 73 ff., 109 unten). Es geht dadurch in eine Art Ö oder Ä über. Offenbar indem durch Wegfall des zusammenhaltenden und maßgebenden Zentrums eine Disgregation eintritt und die in der 3-gestrichenen Oktave liegenden Valenzen ein ihnen im A nicht zustehendes Übergewicht erhalten, zugleich aber die im unteren Formantteile liegenden O- oder AO-haltigen Töne frei werden und den Unterformanten des Ö oder Ä herstellen.

c) Die Wirkung der unter dem Formanten liegenden Teiltöne.

Sind zwischen dem Grundton und dem Formanten noch andere Teiltöne vorhanden, und zwar in lückenloser Folge, wie namentlich beim A, so kommen die Färbungen dieser Teiltöne gleichfalls nicht in ihrer Eigenart zur Geltung. Sie müssen unterdrückt oder assimiliert werden. Die Töne selbst sind durchaus notwendig, wie uns besonders die „Brückenversuche“ drastisch zeigten und jede Synthese bestätigt. Aber ihre spezifischen Vokalqualitäten sind in dem Vokal nicht erkennbar, und daran kann nichts anderes schuld sein, als die größere Intensität des darüberliegenden Formanten. Man kann in einem guten A nichts von O, in einem Ä nichts von A hören, obgleich Töne mit entsprechenden Valenzen und nicht unerheblicher Stärke darin sind. Nur tiefere, dem Formanten ferner liegende Teile des Klanges können, wenn man es darauf anlegt, schwach herausgehört werden und zeigen dann ihre eigentümlichen Vokalcharaktere, U beim A, UO beim Ö, O oder AO beim Ä. Im unanalysierten Ganzen aber wirken sie nur wie eine Untermalung.

d) Liegen größere Lücken oder nur ganz schwach besetzte Strecken zwischen einer tiefen und einer hohen Abteilung, einem Unterformanten und einem Formanten, wie bei den hellsten Vokalen, so ist das Zusammenwirken überhaupt ein weniger integrierendes. Die hohen Formantttöne sind in der Wirksamkeit ihrer Valenzen lange nicht so vom Unterformanten abhängig, wie die Formantttöne des A von seinen tieferen Komponenten. Doch ist der Unterschied nur graduell.

Bemerkenswert ist hier aber wieder, daß, wenn die Lücke durch schwache Teiltöne ausgefüllt ist (wie in der Nähe eines laut Sprechenden), gleichwohl nichts von deren Vokalitäten in den Klang kommt. Entweder sind die Töne selbst nur objektiv vorhanden, subjektiv aber durch die stärkeren auf beiden Seiten

unterdrückt (o. S. 299), oder es werden wieder nur ihre Vokalvalenzen durch die stärkeren des Formanten verdrängt.

e) Der regelmäßige Oberformant des U in der 2-gestrichenen Oktave und die in der Schallquelle vorhandenen oberen Maxima des U, O und A in der 4-gestrichenen Oktave beeinflussen den Vokalcharakter als solchen gleichfalls nicht. U müßte ja sonst nach A hin verändert erscheinen. Aber es wird nur voller und kräftiger. A wird durch die schwachen, hohen Beitöne in seinem Vokalcharakter überhaupt nicht merklich verändert. Wahrscheinlich werden sie subjektiv durch die tieferen unterdrückt.

f) Der gesamte so resultierende Vokalcharakter erscheint beim normalen Hören stets als Eigenschaft des Grundtones. Wird z. B. auf  $c^1$  ein Ä gesungen oder gesprochen, so schwebt der Ä-Charakter nicht irgendwie über dem  $c^1$ , sondern ist auf das innigste mit dieser Tonhöhe verschmolzen, der er doch an sich ganz fremd ist. Der Klang wird eben als einheitlicher gehört, und alle seine Eigenschaften werden auf den Grundton als Träger bezogen. Es ist nur schwer überhaupt möglich, diesen als einfachen Ton mit seinem weichen, dunklen Charakter herauszuhören. Gelingt es aber, dann allerdings schwebt der Ä-Charakter sozusagen haltlos darüber oder ruht nunmehr auf  $c^2$ . Aber es ist dann aus dem Ganzen ein anderes und im gewöhnlichen Hören nicht vorkommendes Phänomen geworden.

Bei genügender Durchbildung dieser zentralphysiologischen, sozusagen metaphonetischen Untersuchungen, von der wir freilich noch weit entfernt sind, müßte man zu einem deduktiv begründbaren Verständnis dafür gelangen, warum in einem bestimmten Vokal auf bestimmtem Grundton gerade diese und keine anderen Teiltöne in diesen und keinen anderen Stärkeverhältnissen enthalten sein müssen. Vielleicht wird eine spätere Zeit sogar mathematische Formeln ersinnen, die es gestatten, mit wenigen Zeichen die ganze Struktur eines Vokals auszudrücken. Etwas Derartiges hat offenbar bereits GRASSMANN vorgeschwebt.

#### 7. Ursprung des akzessorischen Prozesses.

Man kann die Frage aufwerfen, ob der Prozeß, den wir als akzessorischen bezeichneten, insofern er nur bei einer mittleren Tonstrecke zu dem U-I-Prozeß hinzutritt und sich auf diesem weiter ausgedehnten Grundprozeß aufbaut, nicht auch entwicklungsgeschichtlich als der spätere anzusehen sei; analog wie die getönten Farben nach glaubwürdigen Vorstellungen in der tierischen Entwicklungsgeschichte erst später zu den tonfreien, der Schwarz-Weiß-Reihe, hinzugetreten sind. Aber da es

sich hier zunächst nur um die Vokale der menschlichen Stimme handelt und es höchst unwahrscheinlich wäre, daß sich die Menschen ursprünglich nur mit U-I-Lauten verständigt hätten, wird man die Frage zu verneinen haben. Liegt doch auch ein mindestens ebenso kräftiges allgemeines Motiv für die Erzeugung der A-Vokale (so wollen wir kurz die von der U-I-Linie abweichenden nennen) vor, wie für die der U-I-Reihe. Waren es dort die Extreme der Helligkeit und Dunkelheit, die sich zu Ausdruckszwecken darboten, so mußte hier das Bedürfnis vermehrter Klangstärke zur Produktion führen. U-I-Laute können wir nur mit relativ geringer Stärke erzeugen, A-Laute dagegen gestatten volle Verwendung der Lungenkraft. Man wird niemals einem entfernten Menschen mit U oder I zurufen. Sobald man der vollen Expiration halber die Kiefer- und Mundöffnung vergrößert, entstehen eben prägnant-farbige Laute. A selbst kann man mit stärkster Schallkraft geben<sup>1)</sup>. Und so wird sich auch in dem die Schalleindrücke aufnehmenden Gehirn der akzessorische Prozeß von Anfang an mit dem Grundprozeß verknüpft haben. Dies stimmt auch mit der vergleichenden Sprachwissenschaft, die die 3 Vokale U, A, I als die primitivsten betrachtet. Wir kennen keine Sprache und keine Sprachwurzeln, in denen nur U-I-Laute vorkämen.

Anders liegt die Sache, wenn von Vokalen und Vokalitäten in einem weiteren Sinne die Rede ist, von Lauterscheinungen nur ähnlicher akustischer Art, die irgendwo in der Natur vorkommen. Da ist es wohl denkbar, daß das Attribut der Vokalität ebenso wie das der musikalischen Qualität ursprünglich für den Hörenden fehlte und das Ohr der Urmenschen oder ihrer Vorgänger nur Höhen- oder Helligkeitsunterschiede kannte. Aber so weit erstrecken sich unsere Betrachtungen hier nicht.

Dagegen könnte die unter 4 (S. 328) besprochene Frage, ob nicht die A-Vokalitäten einfacher Töne nur eine Folge der empirischen Vokalcharaktere seien, von hier aus noch einmal aufgeworfen werden. Wir lehnten dort die Meinung ab, daß es sich nur um Erinnerungen an die empirischen Vokale handle. Aber könnten nicht in anderer Weise Residualwirkungen, auch rein physiologischer Art, vorliegen? Könnte nicht der kräftige Vokalcharakter, der für A-Vokale nur durch Verbindung mehrerer Teiltöne zustande kommt, auf die daran beteiligten Glieder ge-

<sup>1)</sup> U ist für den Sänger nicht bloß in hoher Lage, sondern auch bei starker Tongebung ein unbequemer Vokal. In HÄNDELS „Israel in Ägypten“ ersetzen manche Dirigenten bei der oft wiederholten Stelle „Er hat geholfen wunderbar“, wo auf dem U immer eine lange Tonfigur einsetzt, das „wunderbar“ durch „gnadenvoll“.

wissermaßen abgefärbt haben? Wenn man die ausgeprägten Vokalcharaktere nun doch einmal nicht aus den blassen Vokalitäten der einfachen Töne begreifen kann und eine Art zentralphysiologischer Chemie in jedem Falle zu Hilfe nehmen muß, wäre da nicht geradezu der umgekehrte Weg zu versuchen?

Der Verfasser glaubte sich in der Tat den Fall zuerst so zurechtlegen zu sollen. Das Kausalverhältnis wäre danach so zu denken, daß der empirische Vokalcharakter durch die Verbindung an sich völlig vokalitätsloser einfacher Töne entstände und diesen dann erst durch einen rein physiologischen (individuell ungleich starken) Übertragungsprozeß ein Schimmer von Vokalität zuwüchse.

Aber auch diese Deutung läßt sich kaum durchführen. Man wird schwerlich Analogien aus der Erfahrung dafür anführen können, daß Komplexqualitäten auf die Komplextteile übergangen, wenn diese nachher in isoliertem Zustande wahrgenommen werden. Nur das Umgekehrte läßt sich in gewissem Umfange durch naheliegende Fälle belegen (die Suppe wird salzig durch beigemischtes Salz usw.). Wenn dagegen beispielsweise noch so oft bei dem Zusammenklange von *c* und *d* eine unangenehme Gefühlsempfindung erlebt wurde, wird man sie doch niemals bei einem dieser Töne für sich allein erleben. Und wenn die Zungenklänge *c* und *fis* zusammenklingend durch ihre Obertöne  $g^1$  und  $fis^1$  Schwebungen geben, wird man doch niemals eine Rauigkeit zwischen den zusammenklingenden einfachen Tönen *c* und *fis* hören oder auch nur zu hören glauben (darum ist es ja eine vergebliche Ausflucht, wenn die Anhänger der HELMHOLTZschen Konsonanztheorie die Dissonanz in diesem Fall aus den Erfahrungen an den gewöhnlichen obertonhaltigen Klängen herleiten).

Man muß daher zugeben, daß die Vokalitäten der einfachen Töne ihnen unabhängig von denen der empirischen Vokale zukommen und daß sie die direkten Folgen physiologischer A-Prozesse sind. Aber es bleibt dabei, daß sie, abgesehen von der U-I-Reihe, im allgemeinen nur schattenhaft in die Erscheinung treten, am deutlichsten noch die A-Vokalität selbst, bei der in der Gegend um  $g^2$  der akzessorische Prozeß nach dem Obigen seine größte Quantität erreicht.

8. Die Formantverschiebung mit der Höhe des Grundtones und das Unkenntlichwerden der Vokale jenseits  $c^2$ .

Diese beiden Erscheinungen sind augenscheinlich aus einer gemeinsamen Ursache herzuleiten, nämlich aus dem Erfordernis einer Mehrzahl harmomischer Teiltöne innerhalb des Formantgebietes.

Von vornherein ist klar, daß mit steigendem Grundton auf einen bestimmten absoluten Tonbezirk immer weniger harmonische Teiltöne kommen. Denn diese sind ein Gerüst, in dem den einzelnen Stufen mit steigenden Ordnungszahlen immer kleinere Intervalle entsprechen. Man braucht nun dieses sich verjüngende Gerüst nur in Gedanken von unten nach oben zu verschieben, um die Folgerung einzusehen. In den Bezirk  $g^2-c^3$  z. B., den Hauptteil des A-Formanten, fallen für den Grundton  $C$  die Teiltöne vom 12. bis 16., für  $c$  nur die vom 6. bis 8., für  $c^1$  nur der 3. und 4., für  $c^2$  der einzige Teilton 2. Man wird daher unwillkürlich den Vokal mit steigender Tonhöhe immer heller, an hohen Obertönen reicher intonieren, soweit es mit seiner Charakteristik noch irgend verträglich ist, d. h. der Formant wird höher, nähert sich dem des Ä. Man erhält so für A bis zum Grundton  $c^2$  immer mindestens 3 Teiltöne innerhalb des weiteren Formantbereiches ( $c^2-g^3$ ), wobei aber A auf  $c^2$  selbst schon ungewöhnlich hell wird. Bei dem Grundton  $e^2$  fällt außer diesem selbst nur noch der Oberton  $e^3$ , bei  $a^2$  überhaupt kein Oberton mehr in den weiteren A-Formanten; der Grundton für sich allein aber wirkt nicht.

Für O erhält man innerhalb des weiteren Formantbereiches, wenn wir diesen mit  $c^1-des^2$  ansetzen, durch solche kleine Verschiebungen der Hauptresonanz bei Grundtönen bis  $g^1$  immer mindestens 2 Teiltöne mit O-Valenzen; weiter hinauf allerdings nur einen, aber eben damit auch nur ein schlechteres O.

Die ausgenützten Teile der zur Verfügung stehenden weiteren Formantregion rücken so immer höher hinauf, aber natürlich bei weitem nicht in gleichem Maße wie der Grundton, da ihrer Verschiebung durch den beschränkten Umfang dieser Region feste Grenzen gezogen sind. Das Prinzip der Absolut-Theorie, die im allgemeinen feste Tonlage der Formanten, bleibt daher gewahrt, und doch muß man der Relativ-Theorie auch ein Stück Wahrheit zugestehen.

Kurz gesagt: der Formant im weiteren Sinne, d. h. die Zone des Tonbereiches, innerhalb deren die Valenzen eines Vokals überhaupt liegen können, ist unverrückbar. Aber innerhalb dieser Zone verschiebt sich der Formant im engeren Sinne, d. h. der für einen bestimmten Grundton in Betracht kommende Ausschnitt, mit steigendem Grundton von den unteren zu den oberen Teilen der Zone.

Zugleich erhellt aber, daß und warum von der Mitte der 1-gestrichenen Oktave an Alterationen der Vokale eintreten müssen, die schließlich nicht mehr mit ihrem Charakter verträglich sind. Notwendig muß sich U dem O, O dem A nähern, da schon

für den Grundton  $c^2$  nur Obertöne von  $c^3$  an zur Verfügung stehen. Die hellsten Vokale aber verlieren ihre richtigen Unterformanten, die ja bei O und U liegen, und nähern sich einander, teils weil die Unterformanten wesentlich zur Unterscheidung beitragen, teils weil in der für diese Vokale entscheidenden Gegend von  $b^3$  bis  $b^4$  schon für den Grundton  $c^2$  nur noch 4 Teiltöne liegen, deren Stärkeverhältnisse durch die Mundresonanz nicht mehr genügend für die Unterscheidung dieser beiden, ohnehin nicht sehr verschiedenen Vokale abgestuft werden können. So gleichen sich die Extreme den mittleren Vokalen an. Diese selbst aber werden mit dem Hinaufrücken des Grundtones über  $c^2$  immer unkenntlicher. Denn immer mehr schieben sich die großen Abstände zwischen den ersten Teiltönen (Oktave, Quinte, Quarte) über die Formantgegenden, so daß in diese zuletzt fast nur noch Lücken fallen. Für den Grundton  $c^3$  sind nur noch  $c^4$ ,  $g^4$  und allenfalls  $c^5$  disponibel. Damit lassen sich aber die Ausgaben des Vokalbudgets nicht bestreiten.

Da die beiden hier besprochenen Tatsachengruppen notwendige Folgen der Zusammensetzung der Vokale aus harmonischen Teiltönen sind, dagegen für die Theorie der unharmonischen Formanten unverständlich bleiben müssen, so ist damit wieder ein Beweis für die erste Lehre geliefert.

#### 9. Der Vokalcharakter der Abbauprodukte.

Diese bei den Interferenzversuchen zu beobachtende Tatsache und die analogen Erscheinungen bei wachsender Entfernung der Schallquelle oder Störungen der äußeren Schalleitung oder der Gehörperzeption<sup>1)</sup> kann man nicht etwa einfach aus der allgemeinen Einstellung des Beobachters auf Vokale herleiten. Denn bei Lückenversuchen treten doch gelegentlich instrumentale Färbungen auf, obschon die Einstellung auf Vokale gerichtet ist. Sie sind nur um so auffälliger.

Die Erklärung ergibt sich zunächst aus den synthetischen Strukturtafeln. Man sieht daraus nicht bloß, daß beim Abtragen der Teiltöne von oben herab immer wieder Vokale, sondern auch welche Vokale entstehen müssen. Immer muß natürlich der defekte Vokal dunkler sein als vorher, also im Vokaldreieck weiter nach links liegen. Aber die hellsten Vokale müssen wegen der leeren Strecken direkt in ihre Unterformanten U und O übergehen, die mittleren, wie Ä und A, durch Zwischenglieder. Zuletzt muß

<sup>1)</sup> Bei Entfernung der Schallquellen verschwinden die höheren Töne früher als die tieferen (vgl. Tonpsych. I, S. 208 ff.), ebenso bei fortschreitenden Labyrinthkrankungen (s. o.). Dabei kommen analoge Abbauprodukte zum Vorschein, wie in unseren Interferenztabellen.

überall (soweit der Grundton überhaupt hörbar ist) U übrigbleiben, da dieses der Vokalcharakter aller einfachen Töne bis gegen  $c^2$  hin ist.

Aber mit diesem Hinweis auf die Strukturtafeln ist die Frage noch nicht voll erledigt, sondern eigentlich nur zurückgeschoben. Man erkennt die Lücke und zugleich die Richtung, in der ihre Ausfüllung gesucht werden muß, deutlicher in dieser Fassung: „Wie kommt es, daß wir jene künstlichen und natürlichen Abbauprodukte immer auch direkt sprachlich erzeugen können, daß sie mit solchen zusammenfallen, die in individuellen, dialektischen, nationalen Vokalbildungen wirklich vorkommen?“ Es wäre immer denkbar, daß sie uns zwar als Verstümmelungen menschlicher Laute durch höhere Gewalten erfahrungsmäßig bekannt, aber nicht durch darauf gerichtete Intention erzeugbar wären.

Die Antwort muß sein, daß die Einstellung der Sprachorgane, die für einen Vokal, abgesehen vom U, erforderlich ist, immer auch Einstellungen auf die daraus durch Abbau resultierenden Vokale in sich schließen muß; die auf E z. B. eine Teileinstellung auf O (für die hinteren Teile der Mundhöhle), die auf I eine auf U usf. Es muß physiologisch unmöglich sein, die für E oder I erforderliche Mundresonanz anders als unter dieser Bedingung herzustellen. Warum dies aber so ist, muß und wird die Physiologie der Sprachwerkzeuge auseinandersetzen, die hier nicht unseres Amtes ist. Sie ist meines Wissens bisher dieser Spezialfrage noch nicht nähergetreten.

Die hier besprochene Tatsache ist für die menschliche Stimme wohl ebenso bedeutungsvoll wie die Zusammensetzung der Vokale aus harmonischen Teiltönen. Wie durch diese ihr Wohlklang, so ist durch jene ihre Widerstandsfähigkeit gegen ungünstige Verhältnisse der Schallübermittlung bedingt. Man wird bei solchen Hindernissen der Leitung oder der Gehörperzeption zwar die Vokale mehr oder weniger mißverstehen, aber nicht so leicht die menschliche Stimme mit einem Instrument oder einer sonstigen Klangquelle verwechseln, obschon natürlich auch dies vorkommen kann.

#### 10. Die Vokalität der Flüstervokale.

Die Flüstervokale lassen sich nicht etwa auf leise Töne zurückführen, die den Flüstergeräuschen nur beigemischt wären. Denn die darin erkennbaren „Tonhöhen“ liegen beim U, O und A um eine Oktave höher als ihre Formanten (o. S. 164). Sollten sie von beigemischten Tönen herrühren, so müßten diese auch ihre Vokalitäten den Flüsterlauten aufprägen, was mit den wahrge-



nommenen Vokalcharakteren in Widerspruch steht. Dies führte uns dazu, sowohl die Höhen als die Vokalitäten zu den immanenten Eigenschaften der Flüstergeräusche selbst zu rechnen. Übrigens tragen auch laute Geräusche, selbst knallartige, wie z. B. die durch Händeklatschen entstehenden, einen gewissen Vokalcharakter. Man kann ein A, U, E klatschen.

Bei den Flüstervokalen liegt die Sache sogar insofern einfacher als bei den stimmhaften Vokalen, als keine variablen Grundtöne vorhanden sind, sondern ein für allemal durch jedes Vokalgeräusch ein bestimmter Tonbezirk ausgefüllt wird, innerhalb dessen die Formantengegend besonders intensiv oder dicht sein muß. Durch diese stetige Ausfüllung erlangen die Flüstervokale trotz ihrer Schwäche eine genügende Deutlichkeit. Es ist dadurch auch ein Ersatz gegeben für den Wechsel des Grundtones bei den gesungenen und für seine stetige Veränderung bei den gesprochenen Vokalen; dort wird die ganze (weitere) Formantzone durchwandert oder abgestreift, hier ist sie auf einmal ausgefüllt.

Zentralphysiologisch können wir auch hier einen Grundprozeß für die U-I-Linie und daneben einen akzessorischen (A-) Prozeß für die übrigen Vokale voraussetzen. Man könnte sogar auch die Frage nach der Vokalität einfacher Töne hier in gewissem Sinne wiederholen, obschon es einfache Geräusche im Sinn eines einzigen sinusförmigen Wellenzuges nicht gibt. Es wäre zu versuchen, künstliche Geräusche zu erzeugen, die wenigstens nur eine ganz kleine Zone der Region  $g^1-g^4$  umfaßten. An diesen wäre zu prüfen, ob ihnen für sich allein ein O- oder A- oder E-Charakter zukäme. Ich vermute, daß es der Fall sein würde. Mit der Interferenzmethode sind solche Geräusche leicht herzustellen, werden aber außerordentlich schwach, daher ist die Beobachtung ihrer Vokalität noch schwieriger, als sie es schon bei den einfachen Tönen ist.

Auffallend ist die Lage der Formanten des geflüsterten U und O: der des U hat nach den If.-Versuchen ziemlich genau die Lage des O-Formanten gesungener Vokale, der des geflüsterten O die des A-Formanten, obgleich die Kiefer- und Mundstellungen beim Flüstern dieselben sind wie beim Singen und lauten Sprechen. Auch in MILLERS graphischen Aufnahmen sind die charakteristischen Frequenzen bei den dunklen (einteiligen) Flüstervokalen durchweg merklich tiefer als bei den gesungenen (Tabelle S. 237)<sup>1)</sup>. GARTEN, dem diese Diskrepanz gleichfalls auffiel, erklärte sie daraus (3, VIII, S. 19ff., IX, S. 6, Anm.), daß beim Flüstern die

<sup>1)</sup> MILLER selbst hebt vielmehr die etwas höhere Lage der oberen Frequenzen bei den hellen Flüstervokalen hervor; aber dieser Unterschied ist in seiner Tabelle sehr unbedeutend und kehrt sich beim Ä sogar um.

Stimmritze weiter sei als beim Singen und Sprechen. Dadurch werde trotz gleicher Kiefer- und Mundöffnung die Resonanz des ganzen Raumes besonders beim U und O verschoben, und zwar in dem Sinn und der Größe der Formantverschiebung. Aber die Paradoxie scheint mir damit noch nicht ganz geklärt. Es bleibt doch immer noch unverständlich, warum der Hörende nicht infolge dieser Resonanzverschiebung statt des U eben ein O, statt des O ein A hört. Man würde allenfalls verstehen, daß der Flüsternde während er eine dem gesungenen U entsprechende Kiefer- und Mundstellung erzeugt, doch einen Resonanzraum herstellt, der der Tonhöhe des O-Formanten entspricht. Aber warum hört man dann nicht eben ein O? Die Vokalitäten selbst als Eigenschaften der einfachen Töne, bzw. die Vokalvalenzen, müssen sich bei den Geräuschen verschoben haben. Warum? Darüber wüßte ich nur etwa die Hypothese vorzubringen, daß der auf die U-Vokalität gerichtete Impuls des Flüsternden schon in den Anfängen dieser Sprechweise eine zwangsläufige Umbildung der Vokaleigenschaften der Flüstergeräusche innerhalb dieser Grenzen im sensorischen Gehörszentrum bewirkt habe. Die dem gesprochenen oder gesungenen U entsprechende Kiefer- und Mundstellung ist in der motorischen Sphäre des Großhirns durch bestimmte Prozesse vertreten. Diese finden sich beim stimmhaften Sprechen und Singen regelmäßig in Begleitung derjenigen Prozesse in der sensorischen Sphäre, die der U-Vokalität entsprechen. Beim Flüstern hat nun — so können wir denken — diese regelmäßige Begleitung nachgewirkt und die den Geräuschen dieser Höhenlage an sich zukommenden Vokalitäten verdrängt bzw. umgebildet. Ob man sich diese Umbildung phylogenetisch oder ontogenetisch (auch heute noch vor sich gehend) vorzustellen hat, mag bei dem hypothetischen Charakter der (den Verfasser selbst wenig befriedigenden) Erklärung dahingestellt bleiben. Bei den HERINGschen „Gedächtnisfarben“, die eine gewisse Analogie bilden, scheint die phylogenetische Deutung die Oberhand zu gewinnen.

## 11. Vokale und Geräusche.

Nicht wenige Forscher haben geradezu nur den Geräuschen Vokalität zuerkannt und die der laut gesprochenen oder gesungenen Vokale aus den beigemischten Geräuschen abgeleitet oder gar die Vokale selbst allgemein als Geräusche definiert (DONDERS, TRAUTMANN, LLOYD, HERMANN, WUNDT, JAENSCH, ABRAHAM<sup>1)</sup>).

<sup>1)</sup> Nach PAGET wäre wenigstens historisch das Flüstern dem lauten Sprechen vorangegangen und hätte seine Vokalcharaktere darauf übertragen. Die englische Sprache speziell sei praktisch auf das Flüstern

Nun hat es zwar mit der Geräuschbeimischung vielfach seine Richtigkeit, namentlich beim A und den helleren Vokalen, aber sie ist auch da äußerst verschieden, je nach der individuellen Beschaffenheit und augenblicklichen Disposition der Stimme<sup>1)</sup>, der Stärke usf., während der Vokalcharakter unabhängig davon, ja sogar deutlicher bei den nichtgeräuschigen Lauten und Stimmen hervortritt. Man kann Vokale auch in einer Entfernung, in der die Geräuschbeimischung ganz verschwunden ist, noch unterscheiden (vgl. HELMHOLTZ S. 117ff.). Vollends die Subsumtion stimmhafter Vokale unter den Begriff der Geräusche kann nur als eine unbegreifliche Verkennung der elementarsten phänomenologischen Tatsachen bezeichnet werden. Gesungene und stimmhaft gesprochene Vokale stehen auch nicht als besondere Schallgattung in der Mitte zwischen Klängen und Geräuschen, sondern fallen bei normalen Stimmen, von unwesentlichen Beimischungen abgesehen, schlechtweg und ausschließlich unter den Begriff der Klänge.

O. ABRAHAM stützt die Annahme, daß nicht die Obertöne, sondern begleitende Geräusche das Wesen der Vokalität ausmachen, auf folgende Erwägung. Wenn man auf einem gegebenen Tone die U-I-Reihe in stetigem Übergange singt, so steigen die Obertöne unstetig in die Höhe. Bei kräftigen Stimmen ist dies besonders deutlich. (Das gleiche ist bei der Vokalröhre von WILLIS zu beobachten, einer Zunge mit verschiebbarem Ansatzrohr, von der mir 2 Exemplare mit verschiedenen Grundtönen zur Verfügung standen.) Man hört aber die Vokale selbst nicht sprunghaft sondern stetig ineinander übergehen. Daraus schließt ABRAHAM, daß die Vokalität nicht von den Obertönen, sondern von den stetig aufsteigenden Geräuschen herkomme.

Hierauf ist aber zu antworten: Wenn man mit der Resonanz-einstellung von einem Teilton der harmonischen Reihe zum nächsten übergeht, so nehmen in der Zwischenzeit beide an der Verstärkung teil, zuerst der tiefere mit überwiegender Stärke, dann beide gleichmäßig, dann der höhere stärker. Dadurch muß aber tatsächlich eine stetige Verschiebung der Klangfarbe für den Nicht-

gegründet. Dessen vollkommenste Verknüpfung mit dem stimmhaften Sprechen finde sich aber nur im Dialekt seiner Heimat „Zumerzet“. Die Beurteilung dieser merkwürdigen Aufstellungen muß Anglisten überlassen bleiben.

<sup>1)</sup> Unter Umständen kann eine ganz leichte Geräuschbeimischung der Stimme auch einen eigenen Reiz verleihen, wie es nach G. ENGEL bei der „etwas verschleierten“ Stimme der JENNY LIND der Fall gewesen sein soll (m. Tonps. II, S. 539).

analysierenden bewirkt werden. Man kann diese Art der Resonanzverschiebung besonders an der Vokalröhre gut beobachten, deren Einstellung man beliebig lange an einem Punkt festhalten kann, um den jeweiligen Sachverhalt festzustellen.

Wenn man den umgekehrten Versuch macht: einen bestimmten Vokal festzuhalten, aber mit dem Grundton stetig in die Höhe zu gehen, so zeigt sich dasselbe. Singe ich ein O zunächst auf den Leitertönen *B, c, d, es, f* mit Aufmerksamkeit auf die hervortretendsten Obertöne, so höre ich immer die jeweilige Duodezime, aber sie beginnt gegen *f* schwächer zu werden, dafür wird hier die Oktave stärker. Die Stärke verteilt sich also so zwischen diesen benachbarten Teiltönen, wie es die Formantgrenzen vorschreiben: während die Duodezime die obere Formantgrenze ( $c^2$ ) überschreitet, tritt gleichzeitig die Oktave ( $f^1$ ) von unten her in die Formantregion ein und übernimmt nun die Führung in der Vokalfärbung. Ganz ebenso ist es natürlich bei stetiger Erhöhung des Grundtones; die Stärkeverteilung ändert sich dann eben auch stetig<sup>1</sup>).

## 12. Zentralphysiologisches über die Konsonanten.

Den Konsonanten als farblosen Sprachgeräuschen müssen, wie den Geräuschen überhaupt, im Gehirn bestimmte, von den Tonprozessen unterschiedene, aber in analogen Veränderungsweisen verlaufende Vorgänge zugrunde liegen. Man kann hier zur Zeit eigentlich nur die beobachteten Erscheinungen in die abstraktere Sprache der Gehirnphysiologie — zuweilen heißt dies auch nur: ins Lateinische — übersetzen. So werden wir für die 3 Hauptklassen der Dauer-, Unterbrechungs- und Augenblicksgeräusche stationäre, intermittierende und momentane Geräuschprozesse in der akustischen Sphäre des Großhirns unterscheiden. Wir werden außer diesen zeitlichen auch Intensitätsunterschiede, z. B. zwischen *Mediae* und *Tenues*, annehmen. Wir werden beim Aussprechen des R durch einen einheitlichen Impuls in der motorischen Sphäre des Sprechers eine zusammenhängende Kette von Teilprozessen in der sensiblen Sphäre des Hörenden (auch des Sprechers selbst) ausgelöst denken, ähnlich wie der Trommelwirbel oder die Schußfolge eines Maschinengewehrs, deren akustische Wirkungen in dieselbe Geräuschklasse gehören, durch einheitliche Akte ausgelöst werden.

<sup>1</sup>) In ähnlichem Sinne haben bereits H. J. MOSER (Katzensteins Arch. f. exp. u. klin. Phonetik Bd. 1, S. 120ff. 1914) und SCHOLE (S. 21) solche Erscheinungen besprochen.

Während aber hier, bei den Zeit- und Stärkeunterschieden, die erscheinungsmäßigen Merkmale sich in analogen Zeit- und Stärkeunterschieden der Gehirnprozesse abbilden, wird man für das eigentlich Qualitative, z. B. für die spezifische Qualität der Nasalkonsonanten, der Zischlaute usw., vergeblich nach gleichartigen Eigenschaften der Gehirnprozesse suchen, sondern sich auf die Feststellung einer regelmäßigen Korrespondenz beschränkt sehen.

Die Helligkeitsunterschiede, die Reihenbildungen in dieser Hinsicht (z. B. vom dunkelsten bis zum hellsten Ch) und das Emporsteigen der Formanten, in dem sie wurzeln, ließen sich vielleicht mit einer abnehmenden Quantität oder, wenn der Begriff anwendbar ist, einer zunehmenden Verdünnung des chemischen Vorganges in der Gehirnrinde, in Zusammenhang bringen. Dieser Zusammenhang wäre verständlich, aber auch er nicht selbstverständlich.

## 14. Kapitel.

### Zur Physik und Physiologie der Sprachlaute.

Fragen dieser Art haben wir bisher nur insoweit berührt, als der Fortgang unserer hauptsächlich beschreibend-phänomenologisch orientierten Untersuchung es erforderte. Aber zum Schluß mögen einige Bemerkungen nicht unerwähnt bleiben, die aus den Beobachtungen folgen oder damit zusammenhängen. Hierbei sollen auch einige die allgemeine Akustik betreffende Fragen zur Sprache kommen.

#### I. Physikalisches.

Daß die Lehre von der multiplen Resonanz im Sinne WHEATSTONES unhaltbar ist, wurde bereits S. 10 erwähnt. Anders verhält es sich mit der multiplen Resonanz in dem Sinne, wie sie später von RAYLEIGH theoretisch und von ZWAARDEMAKER experimentell untersucht wurde<sup>1)</sup>. Darunter ist verstanden die Resonanz eines aus mehreren Teilräumen zusammengesetzten Hohlraumes, wie es der menschliche Rachenraum ist. Die Resonanz (Eigenschwingung) einer Luftmasse wird durch ihre Verbindung mit einer anderen verändert. Wird z. B. ein größerer mit einem kleineren Raum durch einen Kanal verbunden (gekoppelt), so vertieft sich nach ZWAARDEMAKERS Messungen mit dem RAYLEIGHschen Scheibchen die Resonanz des größeren, während sich die des kleineren erhöht. Dabei ist aber auch die Weite des Verbindungskanals, der Grad der Koppelung, von Einfluß. Außerdem ist noch ein 3., tiefstes Resonanzmaximum vorhanden. Diese 3 Maxima sind gleichzeitig in allen Teilen des Gesamtraumes konstatierbar. Umfaßt der multiple Resonator mehr als 2 Teilräume, so wächst natürlich auch die Anzahl der Maxima (Eigenschwingungen). Diese Tatsachen sind von hoher phonetischer Bedeutung, haben aber mit der multiplen Resonanz im früheren Sinne nichts zu tun.

<sup>1)</sup> RAYLEIGH, Theory of Sound 1878, II, Ch. XVI, p. 175. Deutsche Übers. 1880, II, 230ff. AUERBACH 5, S. 459. ZWAARDEMAKER 3.

Man könnte fragen, ob Resonanzräume auch imstande seien, vorhandene Obertöne ganz oder nahezu zu vernichten, wie es doch beim U und in den leeren Strecken der hellen Vokale der Fall sein muß, wenn der aus dem Kehlkopf kommende Klang alle diese Teiltöne kräftig enthält. Diese Möglichkeit ist in der Tat gegeben, da Resonatoren zugleich in gewissem Grade als Interferenzröhren wirken, ebenso wie umgekehrt. Dazu kommt die besondere Beschaffenheit der Mund- und Rachenräume<sup>1)</sup>. An einem mit fleischigen Wänden ausgekleideten Hohlraum habe ich solche Auslöschung auch direkt beobachtet.

Weiter ist es wohl von physikalischem Interesse, daß durch einen Klang von stetig veränderlicher Tonhöhe gleichwohl ein Resonator von fester Abstimmung erregt werden kann; wie auch umgekehrt bei dem GARTENSCHEN Resonanzverfahren der Resonator selbst eine gleitende Tonhöhe besaß. Bei einer durch einen Stempel von  $d^2$  bis  $c^3$  verstellbaren Pfeife konnte ich den Stempel beliebig rasch verschieben: immer wurden doch die zwischenliegenden Resonanzgabeln  $a^2$  und  $b^2$  in Mitschwingung versetzt. Wahrscheinlich nimmt aber bei vergrößerter Geschwindigkeit der Tonbewegung diese Wirkung ab und hört zuletzt auf.

Ferner ist bemerkenswert, daß auch durch Flüstervokale, also leise Geräusche ohne Tonbeimischung, die aus höchst zahlreichen, aber schwachen und in den Wellenlängen durchweg etwas verschiedenen Schwingungen bestehen, Resonatoren erregt werden (o. S. 165). Vermutlich summieren sich hier die Impulse der innerhalb der Resonanzbreite liegenden schwachen Teilschwingungen.

Über Interferenz, ihre physikalischen Einzelheiten und ihre zweckmäßigste Herstellung ist im 2. Kapitel gesprochen. Von besonderer Wichtigkeit in physikalischer Richtung sind wohl die Interferenzerscheinungen bei Geräuschen. Nach den Ergebnissen unserer Versuche nehmen wir an, daß prinzipiell sämtliche Geräusche durch eine hinreichende Anzahl fein abgestufter Interferenzeinstellungen auszulöschen sind. Für die noch nicht ganz geklärte Frage nach der physikalischen Natur der Geräusche dürfte diese Tatsache wesentliche Bedeutung haben. Es scheint sich doch immer um eine große Anzahl objektiver Teilschwingungen zu handeln, die wie ein Tonstaub über eine gewisse, meist sehr weite Tonregion mit ganz geringen Abständen

<sup>1)</sup> HELMHOLTZ S. 182: „Die Luftmasse der Mundhöhle hat nun im Gegenteil (gegenüber Stimmgabeln) geringe Dichtigkeit und Masse, ihre Wände sind, soweit sie von Weichteilen gebildet sind, nicht sehr widerstandsfähig und unvollkommen elastisch, haben bei Erschütterungen viel innere Reibung, wodurch sie Bewegung vernichten.“ Vgl. auch S. 170—171.

ihrer Wellenlängen, aber verschiedenen Amplituden und vielleicht auch verschiedener Dichtigkeit der einzelnen Teilstrecken verteilt sind.

Daß auch Augenblicksgeräusche konsonantischer und sonstiger Art durch Interferenz stufenweise zu vernichten sind und demnach physikalisch unter den aufgestellten Begriff fallen, lehrten unsere Versuche. Ja auch bei lauten Knallen im geschlossenen Raume fand sich das nämliche. Im Freien erhält man bei oszillographischen Aufnahmen von Geschoßknallen gleichfalls kleine Wellenzüge, die nicht bloß auf eine Mehrzahl sukzessiver Teilschwingungen, sondern wohl auch auf höhere, der Grundwelle aufgesetzte Komponenten deuten. Im Grunde tragen aber diese dumpfen, tiefen Knalle mehr Ton- als Geräuschcharakter.

So bleiben unter den Augenblicksgeräuschen nur jene sehr schwachen Knalle übrig, die aus einer einzigen Schwingung, ja aus einem Bruchteil einer solchen bestehen, die man aber richtiger gleichfalls nicht als Geräusche, sondern als extrem verkürzte Klänge bezeichnet. Gehört das Kurvenfragment, das die graphische Darstellung zeigt, nicht einer Sinuskurve an, so kann es gleichwohl der Anfang einer Kurve sein, die in ihrer Fortsetzung periodisch würde, kann also physikalisch als Bild eines extrem verkürzten Klanges aufgefaßt werden, dessen sinusförmige Bestandteile freilich unbestimmt bleiben. Ob übrigens der „Einlochknall“ ABRAHAM'S (o. S. 136, Anm. 1) nicht doch auch durch Interferenz zum Verschwinden gebracht werden kann, wäre noch genauer zu prüfen.

Bereits KOEHLER hat aus Interferenzversuchen mit Geräuschen, bei denen sich allerdings der Einsatz der Klopfgeräusche nicht vollkommen ausschließen ließ, auf die Periodizität der Geräusche geschlossen. Seine Überlegungen (I, III, S. 80ff.) sind überhaupt hier zu vergleichen. Übrigens hat schon 1877 GRASSMANN im obigen Sinne von einer „Geräuschbreite“ gesprochen und ist in m. Tonpsych. II, S. 506ff. der physikalische Geräuschbegriff in genau derselben Weise wie hier formuliert. Auch BRÜCKE, HELMHOLTZ, BARTH u. a. haben die gleiche Ansicht für alle oder viele Geräusche vertreten. Aber es fehlte noch der experimentelle Nachweis.

NAGEL, der die alte Definition der Geräusche als unperiodischer Schwingungen auf Grund der HERMANN'Schen Konsonantenkurven verwirft (sie wird ja auch schon unhaltbar bei einem mit stetiger Tonerhöhung gesungenen Vokal, der dadurch zu einem bloßen Geräusch würde), gibt eine Definition, die heute ziemlich verbreitet ist: „Zurücktreten des Grundtons, überhaupt Fehlen einer leicht bestimmbareren Tonhöhe.“ Wenn dies eine physikalische Definition sein soll, so gilt dagegen, daß auch bei Vokalen, wie beim A, und bei Instrumentalklängen der Grundton sehr schwach und unter Umständen = 0 sein kann. Meint man aber nur die leichte Erkennbarkeit der Tonhöhe des Grundtons oder des ganzen Klanges, so ist dies überhaupt kein physikalisches Merkmal. Überdies fehlt sie auch den Tönen oberhalb  $c^5$  und unterhalb  $C_1$ , deren Einordnung unter die Geräusche mir wenigstens ganz ungerechtfertigt erscheint.



## II. Physiologisches.

Über verschiedene, untrennbar mit unserem Thema zusammenhängende physiologische Fragen ist im vorangehenden mehrfach, besonders im 12. und 13. Kapitel, gesprochen. Hier sollen noch die peripher-physiologischen Fragen über die Bildung der Laute durch die Sprachwerkzeuge und die Übermittlung des akustischen Prozesses an den Hörnerven berührt werden, aber nur so weit, als es der Ertrag unserer Untersuchungen nahelegt.

### 1. Zur Erzeugung der Sprachlaute<sup>1)</sup>.

Kein musikalisches Instrument ist so vielseitigen Gebrauches fähig, wie das menschliche Sprachorgan. Man kann ihm bei gleichbleibender Tonhöhe verschiedene Resonanzeinstellung erteilen und bei gleicher Resonanzeinstellung verschiedene Tonhöhen damit produzieren (stimmhafte Vokale). Ferner kann es außer Tönen Geräusche hervorbringen (Flüstervokale und Konsonanten). Es kann stetige Tonübergänge ebenso wie sprunghafte vollziehen (Grundverschiedenheit des Singens und Sprechens). Es kann sich endlich sogar auf verschiedene Klangfarben einstellen (Register). Diese Eigenschaften sind im einzelnen natürlich auch sonst vorhanden. Stetige Übergänge z. B. kann man neben sprunghaften auch leicht auf manchen Instrumenten herstellen, obgleich sie in unserer Musik nicht offiziell und in größerem Maßstabe verwendet werden. In Registern übertrifft uns weit die Orgel. Aber in Verbindung miteinander sind diese Eigenschaften für das menschliche Sprachorgan grundwesentlich und charakteristisch. Sie sind sozusagen sein tägliches Brot und kommen in dieser Vereinigung sowie in der Leichtigkeit der Umstellung aus der einen in die andere Funktion sonst nirgends vor.

<sup>1)</sup> Der Leser erwarte sich hier nicht ein Eingehen auf die RUTZsche Lehre von den „klanglichen Konstanten“ d. h. der Bedeutung der Körperhaltung und Rumpfmuskulatur, die die richtige Klanggebung und Vortragsweise eines Liedes je nach den verschiedenen Komponisten begünstigen sollen. Es ist freilich wahr, daß das ganze Muskelsystem, ja der ganze Körper vom Kopf bis zur Zehe gewissermaßen mitsingt. Aber von da bis zu solchen „Typen“ ist noch weit; und die Beweisführung erscheint, an den Anforderungen einer exakten Experimentalpsychologie gemessen, ganz und gar dilettantisch. Vgl. dazu A. GUTTMANN (Lit.-Verz.), LEYHAUSEN (Arch. f. d. ges. Psych. Bd. 30. 1913), GEBHARDT (daselbst Bd. 50. 1925). Auch die damit in Verbindung stehenden SIEVERSschen Drahtfiguren, deren Anblick den Sänger zum richtigen, stilgemäßen Vortrag eines Liedes anleiten soll, kann ich leider nur als eine unbegreifliche Selbsttäuschung des berühmten Germanisten und Phonetikers einschätzen.

Die brennendste Streitfrage in genetischer Richtung betraf lange Zeit die Erzeugung der Vokale. Sie wird jetzt vorwiegend zugunsten der HELMHOLTZschen und zuungunsten der HERMANNschen Lehre entschieden. Auch die Gesamtheit unserer Ergebnisse über die stimmhaften Laute läßt sich unmöglich mit der Lehre von der Erzeugung der Sprachlaute durch ein in der Periode der Stimmlippenschwingungen intermittierendes Anblasen des Rachenraumes vereinigen. Nur die Flüstervokale und die stimmlosen Konsonanten werden natürlich durch bloßes Anblasen der Mundhöhle erzeugt, da hier die Stimmlippen überhaupt nicht oder nicht in maßgebender Stärke schwingen. Bei den stimmhaften Vokalen aber erzeugen diese einen sehr zusammengesetzten Klang, der durch die jeweilige Einstellung des Resonanzraumes modifiziert wird, indem gewisse Klangteile verstärkt, andere unterdrückt werden. Dabei schwingen nach EWALDS, MUSEHOLDS und NAGELS Untersuchungen die Stimmlippen im Brustregister vorwiegend in der Weise von „Polster- oder Gegenschlagpfeifen“, sich horizontal gegeneinander bewegend. Der Klang solcher Pfeifen ist, wie ich an einem von Herrn WETHLO gebauten Modell wahrnehmen konnte, außerordentlich schreiend, also äußerst obertonhaltig. Mögen auch zwischen dem Modell und dem Original im lebenden Organismus bedeutende Unterschiede bestehen: in diesem Punkte dürfte es nicht zu weit von der Natur abweichen. Wenn KATZENSTEIN angab, daß er durch Einführung einer Röhre bis in die unmittelbare Nähe der Stimmlippen und durch objektive Aufnahme ihrer Schwingungen auf einer Membran nahezu sinusförmige Bewegungen erhalten habe, so wird dabei durch die Einrichtung selbst irgendwie der Klang vereinfacht worden sein<sup>1)</sup>.

Die Aufgabe, für alle Sprachlaute die in den Strukturtafeln verzeichneten Resonanzmaxima aus entsprechenden Einstellungen der Sprachorgane herzuleiten, steht noch in den Anfängen ihrer Ausführung, zumal da über die Resonanzmaxima selbst bisher noch manche Zweifel bestanden. Aber eine prinzipielle Schwierigkeit dürfte hier nicht mehr vorliegen. Der Oberformant des U wurzelt offenbar in der durch die Lippen gebildeten vordersten Abteilung, die Unterformanten der hellen Vokale bilden sich in dem größeren hinteren Teile des gesamten multiplen Resonanzraumes, usf.

<sup>1)</sup> Ganz ist das Verhalten des Kehlkopfes bei seinen verschiedenen Leistungen noch nicht geklärt. O. WEISS (4, II) schließt aus Versuchen an einem ausgeschnittenen Kalbskehlkopf, daß der schwingende Körper hier stets die Luftmasse über den Stimmbändern sei, ebenso wie bei einer aufschlagenden Zunge. Immerhin könnte die Sache beim menschlichen Gesang etwas anders liegen als bei dem des Kalbes.

ZWAARDEMAKER folgert aus der oben erwähnten Untersuchung der multiplen Resonanz, daß in dem stark zusammengesetzten Resonanzraum des menschlichen Stimmorgans im allgemeinen unharmonische Formanten entstehen müssen, wenn sie auch vielleicht durch kleine unbewußte Umstellungen der Muskulatur, namentlich bei feinhörigen Sprechern und Sängern, zu harmonischen umgebildet würden. Hier können wir uns dem hervorragenden Physiologen nicht anschließen. Wenn anders der Mundraum die Funktion hat, aus den Teiltönen des durch die Stimmlippen erzeugten obertonreichen Klanges durch Resonanz eine Auswahl zu treffen, so können auf diesem Wege doch keine anderen als die in dem Klang enthaltenen Teiltöne zustande kommen, und diese können bei Polsterpfeifen nicht wohl andere als harmonische sein. So hat es sich uns ja auch allgemein bestätigt. Nur insoweit noch sonstige schwingende Gebilde außer den Stimmlippen in dem ganzen System der Sprachwerkzeuge (einschließlich der unterhalb der Stimmlippen liegenden Teile und der nasalen Hohlräume) vorhanden sind und durch den Luftstrom angeregt werden, können sich unharmonische Töne dem Klange beimischen. Aber sie werden bei den nicht nasalierten Vokälen so gut wie keine Rolle spielen.

Man hat es wunderbar gefunden — und VIËTOR stützte darauf (1) seine anfängliche Ablehnung der Resonanztheorie —, daß Kinder mit ihren so viel kleineren Resonanzräumen die gleichen Vokale wie Erwachsene hervorbringen können. Aber erstlich sind die Unterschiede in den Dimensionen nicht so groß wie man erwarten möchte (JESPERSEN hat darüber Messungen angestellt). Zweitens sind tatsächlich gewisse Unterschiede vorhanden; das kindliche U ist nicht so dunkel, das I heller als bei Erwachsenen. Endlich und hauptsächlich stellen Kinder in dem Bestreben, die Laute der Erwachsenen nachzubilden, ihre Sprachwerkzeuge eben so ein, daß die erstrebten Klangfarben nach Möglichkeit herauskommen. Die durch das Knochengerüst des Kiefers bedingten Verschiedenheiten werden durch die Öffnung der Mundhöhle, die Stellung des weichen Gaumens und der Zunge ausgeglichen. Man darf auch nicht übersehen, daß die Kieferstellung auch bei ein und demselben Individuum bedeutende Veränderungen verträgt, ohne daß die Vokale unkenntlich werden. Wenn man einen Finger zwischen die Zähne klemmt und so die Kieferstellung festlegt, lassen sich alle Vokale und sogar die feineren Nuancen, wie ein helleres und dunkleres A, noch deutlich erzeugen. Auch wenn der Kunstsänger den ganzen Stimmapparat so einstellt, daß der Kehlkopf entlastet wird und mehr die peripherischen Teile in Anspruch genommen werden (HABER, GUTZMANN, KATZENSTEIN), bleiben die Vokalcharaktere im wesentlichen erhalten, freilich mit nicht unbedeutenden Modifikationen im einzelnen.

Die Schwingungen der Stimmlippen und der im Ansatzrohr eingeschlossenen Luftmasse teilen sich auch den einschließenden festen Wänden mit, ebenso wie bei den Blasinstrumenten. Wenn

man sich die Ohren mit den Fingern zuhält oder mit Ohropax (sit venia verbo!) verstopft und nun die stimmhaften Vokale spricht, so sind sie durch den behinderten Schallabfluß subjektiv bedeutend verstärkt, und zwar am meisten bei U und I, weniger bei O und E, am wenigsten bei A<sup>1)</sup>. Offenbar nehmen die im hinteren Teil des Ansatzrohres resonierenden tieferen Teiltöne der Vokale, bei E und I die Unterformanten, mehr als die höheren an der Verstärkung teil. Wahrscheinlich hört darum der Sprechende auch bei offenen Ohren seine eigene Stimme beim U und I etwas dunkler, beim A etwas heller als die Nebenmenschen.

Aber auch bei offenen Gehörgängen verraten sich diese der Umgebung von den Stimmlippen mitgeteilten Vibrationen durch das Vibrationsgefühl, dessen zuerst wohl HELMHOLTZ gedenkt, um daraus den charakteristischen Ton des U zu erschließen. Es wurde von HOPMANN nach seiner Intensität und Ausdehnung an Schädel, Hals und Wangen bei den verschiedenen Vokalen untersucht<sup>2)</sup>. Wieder zeigte es sich am deutlichsten und stärksten bei U und I, weniger bei O und E, am wenigsten bei A. Parallel damit ging auch die Ausdehnung der Vibrationsbezirke. Diese Reihenfolge entspricht genau den Stärken der tiefsten Teiltöne, wie sie sich bei verschiedenen Gelegenheiten fanden (o. S. 27, 75, 212), und ist umgekehrt proportional den Höhenlagen dieser Vokale im Vokaldreieck.

Stimmhafte Vokale im vollen Sinne können von den Sprachorganen natürlich nur durch die Stimmlippen erzeugt werden. Aber Patienten, denen die Glottis zugewachsen oder der ganze Kehlkopf exstirpiert ist, können sich doch noch eine „Pseudostimme“ und eine ziemlich verständliche Sprache mit Vokalen und Konsonanten erwerben.

Vgl. hierüber CZERMAK, Wiener med. Wochenschr. 1859. BOSE, Diss. Gießen 1865. STRÜBING, Virchows Arch. f. pathol. Anat. u. Physiol. 1890. GOTTSTEIN, Arch. f. klin. Chir. Bd. 62. GUTZMANN, Zeitschr. f. Laryngol. 1909. H. STERN, das. 1923. GUTZMANN, Sprachheilkunde 3. Aufl., hersg. v. ZUMSTEG 1924, S. 571ff. Die Operierten, die nur durch eine Kanüle atmen können, müssen die zur Stimmbildung nötige Luft aus der Magenhöhle beziehen. Manche bringen es nur zu einer wesentlich stimmlosen Sprache, in der Vokale nur als Auslaute von Konsonanten auftreten, andere aber erwerben unter ärztlicher Leitung eine, wenn auch heisere und tiefe,

<sup>1)</sup> Wie ich nachträglich sehe (aus GRÜTZNER 1), haben dieselbe Beobachtung auch schon GRÜTZNER und der Taubstummenlehrer DEUTSCH gemacht.

<sup>2)</sup> E. HOPMANN, Untersuchungen über die bei gesungenen Vokalen an Kopf und Hals auftretenden Vibrationen. Monatsschr. f. d. ges. Sprachheilk. 19. Jahrg. 1909. Vgl. auch GIESSWEIN 1, S. 36ff. des Sonderabdruckes (Diss.).

doch nicht ganz tonlose Stimme, die bis zu einer oder zwei Oktaven umfassen kann. GOTTSTEIN stellte auf einem Chirurgenkongreß sogar einen solchen Zögling vor, der das Loreley-Lied sang — wobei freilich die bekannte Weise einem weitherzigen Entgegenkommen der Hörer begegnen mochte. Ich hatte auch selbst Gelegenheit, einen von Prof. GLUCK operierten Patienten dieser Art auf seine Vokalisation zu prüfen. Es war kein eigentliches Flüstern (so will es auch STRÜBING nicht nennen), sondern ein leises, heiseres Sprechen, das doch immer etwas Ton hatte. Versuchsreihen mit verschiedenen Vokalen ergaben, daß dieselben von 6 Beobachtern fast immer richtig erkannt wurden. Eine eigentliche Sprachmelodie war nicht vorhanden, die Höhenunterschiede wurden im wesentlichen durch Akzentunterschiede wiedergegeben. Nur geringe Höhenunterschiede waren hervorzubringen, wobei Verlagerungen der Zunge die Hauptrolle spielten, aber immer blieb die Stimme rauh und tief. Pfeifen und Singen waren ausgeschlossen.

An einem von Dr. H. ZUMSTEEG operierten Patienten, der sogar mit Erfolg telephonieren konnte, hat F. TRENDELENBURG mit dem o. S. 213 erwähnten Verfahren trotz der außerordentlich geringen Schallintensität Vokalkurven von U, A und I aufnehmen und an den Bildern durch die Lupe folgendes ersehen können: beim A war die Grundperiode 400, auf sie aufgesetzt der Oberton 800 (Formant) und besonders stark ein hoher Teilton von 2800 bis 3200 Schw., welcher den mehrfach erwähnten hohen Teiltönen des A entspricht. Ein anderes A zeigte als Grundtonperiode 500 und als ersten Oberton 1000. Die Feinstruktur dieser Klänge zeigte aber im Gegensatz zu den gesungenen Vokalen Unregelmäßigkeiten, und diese sowie das abnorm kräftige Auftreten des Gebietes um 3000 verliehen nach Tr. dem Klange seinen rauhen und harten Charakter. „Wir haben hier eine Stoßerregung im Sinne HERMANN'S ohne die zur Erklärung der Bilder der gesungenen Vokale notwendige Einschränkung, daß ein Stoß dem anderen identisch gleicht“ (Briefl. Mitt. an Dr. ZUMSTEEG).

Es ist, soviel ich sehe, noch eine offene Frage, ob die Verschiedenheiten der stimmhaften Vokale nicht doch auch schon im Kehlkopf durch die Schwingungsform der Stimmlippen in gewissem Grade präformiert werden. SCRIPTURE hielt diesen Einfluß sogar für ausschlaggebend. Daß indessen hauptsächlich die Einstellung der Resonanzräume entscheidet, kann nicht bezweifelt werden. Nur eine Mitwirkung der Stimmlippen könnte in Frage kommen. Wie diese beim Falsett in anderer Weise als bei der Bruststimme (nicht in ihrer ganzen Breite, vielleicht nicht einmal gegensschlagend) funktionieren, so wären auch Unterschiede feinerer Art, je nach den Vokalen, denkbar. Die auf einen Vokal gerichtete Innervation des gesamten Sprachapparates würde dann zugleich mit einer bestimmten Einstellung der Resonanzräume auch eine bestimmte Schwingungsweise der Stimmlippen bewirken, und es wären infolgedessen schon hier die Teiltöne des betreffenden Formanten gegenüber anderen begünstigt. AUERBACH, PIPPING, MUSEHOLD u. a. sind denn auch für diese Annahme eingetreten. Aber meines Wissens liegt eine ins einzelne gehende und beweiskräftige Untersuchung noch nicht vor.

Die Resonanzeinstellung des Ansatzrohres ist auch bei den stimmlosen Sprachlauten wesentlich. Hier hat das Ansatzrohr, wie TRENDELENBURG (2) anlässlich der Konsonanten richtig bemerkt, eine doppelte Bedeutung: als selbständige Schallquelle, die vom Luftstrom des Kehlkopfes angeblasen wird, und als Resonator, der gewisse Geräuschteile verstärkt. Die Resonanzeinstellung bedingt die Vokalität der Flüstervokale in derselben Weise wie die der gesungenen. Aber auch die darin wahrzunehmenden Tonhöhen sind von der Resonanzeinstellung abhängig und fallen deshalb im allgemeinen mit den Formantzentren zusammen. Nur bei U, O, A fanden wir es nötig, eine Überblasung des Geräusches in die Oktave anzunehmen. Die in den Konsonanten beobachtbaren Tonhöhen aber, die in keiner näheren Beziehung zu deren Formanten stehen, müssen als Nebenwirkungen der Einstellungen auf diese Geräusche gedeutet werden; d. h. es müssen durch den Luftstrom gewisse Teile des gesamten Raumes zu Geräuschschwingungen von dieser bestimmten Höhe angeblasen werden. Die Aufzeichnung dieser Vorgänge liegt, ebenso wie die bei der Erzeugungsweise der einzelnen Konsonanten, außerhalb unserer Aufgabe.

Die Frage, ob bei den stimmlosen Sprachlauten die Stimmlippen nicht doch auch beteiligt sind, obschon sie keinen Ton hervorbringen, und ob eventuell diese Beteiligung von Vokal zu Vokal variiert, ist gleichfalls noch unentschieden. Unentbehrlich ist diese Mitwirkung zwar nicht; denn bläst man die Mundhöhle unter Ausschluß des Kehlkopfes von außen oder mit einem durch die Nase geführten Schlauch an, so kommen den Flüstervokalen ähnliche Geräusche doch heraus. Aber es wäre immer möglich, daß beim wirklichen Flüstern ganz schwache Mitbewegungen der Stimmlippen erfolgen, wie dies z. B. von STEFANINI behauptet wird.

## 2. Zur physiologischen Hörtheorie.

Den Schluß der mit unserer Untersuchung zusammenhängenden peripher-physiologischen Fragen bildet die physiologische Hörtheorie, d. h. die Bestimmung der Vorgänge bei der Übertragung der physikalischen Reizung auf den Hörnerven. Die bisher erfolgreichste Darstellung dieses Vorganges, die von HELMHOLTZ, fanden wir überall, wo die Überlegungen darauf führten, hervorragend nützlich. Damit ist aber nicht gesagt, daß sie allein, zumal in ihrer konkreten ursprünglichen Gestalt, sich als Grundlage eigne. Bekanntlich ist sie von Anfang an auch mannigfach bekämpft worden. Zuletzt erachtete auch KOEHLER an einflußreicher Stelle (3, S. 451ff.) die Tatsachen des Vokalhörens als unvereinbar mit der

OHM-HELMHOLTZschen Lehre, wonach Klänge nur Summen von Teiltönen seien, die er deshalb als „summative Theorie“ bezeichnet. Denn die spezifischen Klangeigenschaften, besonders aber die Vokalqualitäten, seien daraus nicht zu verstehen.

Wir müssen hier etwas weiter ausgreifen. Es handelt sich um 5 Thesen, die sich übereinander aufbauen:

1. das rein mathematische Theorem FOURIERS, nach welchem jede periodische Schwingung sich algebraisch als eine Summe von Sinusschwingungen ausdrücken läßt;

2. die von OHM verteidigte rein physikalische Definition des einfachen Tones als einer Sinusschwingung der Luft und jeder anderen periodischen Schwingungsform als einer Summe superponierter Sinusschwingungen;

3. HELMHOLTZ' psychophysische Behauptung, daß die Empfindungen bei physikalisch einfachen Tönen im Sinne von 2. auch durch das Gehör nicht weiter zerlegt werden können<sup>1)</sup>, jeder musikalische Klang aber einer Mehrzahl solcher einfachen Töne entspricht, die unter günstigen Umständen auch herausgehört werden können, und daß die Art dieser Zusammensetzung nach Zahl und Stärkeverhältnissen der Teile die Unterschiede der Klangfarbe bedingt;

4. HELMHOLTZ' rein physiologische Hypothese, wonach zahlreiche, verschieden abgestimmte Teilchen innerhalb der Schnecke des Ohres durch Resonanz die eintreffenden zusammengesetzten Schwingungen in Sinusschwingungen zerlegen;

5. die rein psychologische These, daß auch die dem Bewußtsein gegebene Gehörsempfindung bei musikalischen Klängen in analoger Weise wie die akustischen Prozesse aus Teilempfindungen zusammengesetzt sei, obgleich diese nur unter besonderen Voraussetzungen als solche bemerkt werden.

Das Historische betreffend, mögen hier zwei Bemerkungen eingeschaltet sein.

Für OHM handelte es sich nach dem Wortlaut seiner Ausführungen nur um diejenige mathematisch-physikalische Definition eines Tones von der Schwingungszahl  $m$  („Schwingungsmenge“), „welche zugrunde gelegt werden muß, um daraus jede beliebige  $m$  mal in der Sekunde wiederkehrende Bewegungsform abzuleiten“. Als solche bezeichnet er die Formel:  $a \cdot \sin 2\pi (mt + p)$ . Diese Definition hatte er in einer vorausgehenden Untersuchung über Kombinationstöne und Schwebungen als „eine ausgemachte Sache“, „eine von den Vorfahren überkommene“, „aus alter Zeit herstammende“ vorausgesetzt, und verteidigte sie nun gegenüber Experimenten von SAVART und SEEBECK, aus denen hervorzugehen schien, daß

<sup>1)</sup> Diesen Satz hat HELMHOLTZ zwar nicht wörtlich ausgesprochen, aber ohne Zweifel, wenigstens bei schwachen Tönen, stillschweigend vorausgesetzt. Der Verfasser hat ihn I, II, S. 258ff., so ausgesprochen und begründet. A priori selbstverständlich ist er nicht.

jeder regelmäßige oder auch nur annähernd regelmäßige Impuls als Ton empfunden werde. Hier kommt er nun freilich mit phänomenal-akustischen Beobachtungen in Berührung und scheint dabei die 3. These auch schon vorauszusetzen, wagt aber auf diesem ihm unzugänglichen Gebiete keine positive Formulierung.

Daß HELMHOLTZ in seiner Hypothese der Schneckenklaviatur Vorläufer hatte, ist in Deutschland nicht so unbekannt, wie GRADENIGO in einer historischen Darlegung (*Arch. ital. di Otol.* Bd. 27—29, 1916—1918), deren zeitbedingte Animositäten hier auf sich beruhen mögen, voraussetzt. So sind in des Verfassers *Tonpsych.* Bd. 2, S. 98 ff. DU VERNEY, LE CAT, SCARPA, A. v. HALLER, HERDER, TREVIRANUS u. a. erwähnt. COTUGNO allerdings war auch mir entgangen, und es scheint, daß seine Abhandlung (*COTTUMNIUS, De aquaeductibus auris humanae diss. 1775*), in welcher im Zusammenhang mit anatomischen Beschreibungen der Labyrinthgebilde in der Tat der Gedanke der Resonanz der einzelnen Schneckenfasern auf einzelne Töne klar und bestimmt ausgesprochen ist, in Deutschland wenig bekannt wurde. Ob HELMHOLTZ sie kannte, weiß ich nicht. Daß er nicht gerade die Neigung hatte, sich mit fremden Federn zu schmücken, zeigt die „*Physiologische Optik*“, die reichlich mit historischen Nachweisungen bis zu den alten Arabern zurück ausgestattet ist. Wenn wir nun trotz COTUGNO und anderen, teilweise noch bedeutend älteren Vorgängern gleichwohl fortfahren, die Hypothese als die HELMHOLTZsche zu bezeichnen, so geschieht es aus demselben Grunde, aus dem man die Theorie der natürlichen Zuchtwahl nach DARWIN benennt, obgleich schon EMPEDOKLES den Grundgedanken ausgesprochen hat: weil nicht die Idee als solche, sondern ihre Verknüpfung mit zahlreichen Erfahrungstatsachen eine Hypothese erst zum brauchbaren Werkzeug der Forschung macht, wie dies auch im vorliegenden Falle der Verlauf der wissenschaftlichen Entwicklung seit HELMHOLTZ gezeigt hat. Vgl. über diese Prioritätsfrage auch A. KREIDL, *Z. Geschichte d. Hörtheorien*, *Arch. néerland. de physiol. de l'homme et des anim.* Bd. 7, S. 502 ff. 1922, wo noch andere, ältere und neuere, Autoren vor HELMHOLTZ (auch CH. BELL wäre noch zu nennen) für diese Vorstellungsweise zitiert werden. Es scheint, daß sie überhaupt seit der Mitte des 18. Jahrhunderts fast die überwiegende war und erst infolge des Einspruches von MAGENDIE, JOH. MÜLLER und E. H. WEBER so zurückgedrängt wurde, daß sie zu HELMHOLTZENS Zeit als etwas Neues erscheinen konnte.

Übrigens weist KREIDL auf eine noch auffallendere Antizipation unter den Hörtheorien hin: daß nämlich EWALDS „*Schallbildertheorie*“ nicht nur in der Sache, sondern auch im Namen bereits von ARNOLD und anderen ausgesprochen worden war. Man wird sie trotzdem immer EWALD zuerkennen, da er sie durch lehrreiche Versuche gestützt hat.

Die 1. These ist unbestreitbar. Die 2. ist zunächst eine bloße Definition, bei der nur nach ihrer Zweckmäßigkeit gefragt werden kann. Daß es aber zweckmäßig ist, periodische Luftschwingungen mathematisch als Summe von Sinusschwingungen darzustellen, kann nicht bestritten werden. Denn nur dadurch ist es möglich, ihre gesamte Form algebraisch festzulegen und die unendliche Mannigfaltigkeit der Formen auf ein einheitliches Prinzip zurückzuführen, während außerdem nur für die Längen- und Amplitudenunterschiede mathematische Ausdrücke zur Verfügung stehen.



Etwas mehr ist schon zur 3. These zu sagen. Fest steht, daß die durch eine nicht zu starke Sinusschwingung erzeugte Tonempfindung auch subjektiv auf keine Weise in eine Mehrzahl von Tönen zerlegt werden kann, und daß umgekehrt jede physikalisch und mathematisch in Sinusschwingungen auflösbare Welle bei genügender Stärke und Verschiedenheit der Komponenten von Geübten subjektiv in entsprechende Teiltöne zerlegt werden kann. In dem Streite zwischen OHM und SEEBECK hatte dieser zwar darin recht, daß die Obertöne im allgemeinen nicht in der geforderten Stärke gehört werden (o. S. 303). Aber er hatte unrecht, wenn er die Klangfarbenunterschiede nicht ausschließlich auf Unterschiede in der Zusammensetzung der Gesamtwelle, sondern auch auf ihre bloßen Formverschiedenheiten als solche gründen wollte, wie dies vor OHM allgemein geschah (vgl. z. B. CHLADNIS Akustik).

Daß die Formverschiedenheiten der objektiven Tonwellen sich nicht als solche subjektiv geltend machen, ist vor allem bewiesen durch die Einflußlosigkeit der Phasenverschiebungen. Denn obgleich diese die objektive Wellenform gewaltig verändern, bleiben sie für das Gehör nach HELMHOLTZENS immer wieder bestätigt (Nachweise wirkungslos<sup>1</sup>).

Da gleichwohl immer wieder Versuche auftauchen, ohne speziellen Zerlegungsmechanismus auszukommen und das tatsächlich Gehörte oder Hörbare aus der Form der zusammengesetzten Schwingungen selbst herzuleiten, so wolle man nur beispielsweise die beiden folgenden Wellenformen betrachten:

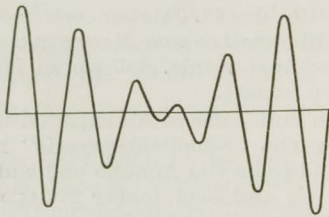


Abb. 7. 5:6.

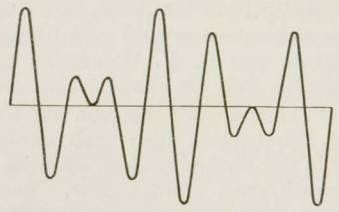


Abb. 8. 5:7.

<sup>1</sup>) Den Fall der Schwebungen hat HELMHOLTZ selbst schon angenommen. KOENIGS Versuch mit seiner „Wellensirene“ scheiterte daran, daß er aus der Sinusform der aufgesetzten Blechscheiben fälschlich auf sinusförmige Luftschwingungen schloß (m. Abh. 3, S. 677 ff.). Aber auch seine sonst ausgezeichnet durchgeführte Versuchsreihe über Kombinationstöne („Stoßtöne“) mit den daraus gegen HELMHOLTZ gezogenen Folgerungen verliert ihre Beweiskraft durch die Obertonhaltigkeit elektromagnetisch erregter Gabeln (daselbst). Endlich haben sich auch HERMANN'S, in seiner Vokaltheorie gründenden, Einwendungen als hinfällig erwiesen.

Die 1. entsteht durch Superposition zweier Sinusschwingungen vom Verhältnis 5 : 6, die 2. durch solche vom Verhältnis 5 : 7 bei gleichen Amplituden und ohne Phasendifferenz zu Beginn der Schwingungen. Im 1. Falle hört man, wenn die Periode sich 100mal in der Sekunde wiederholt, die Töne 500 und 600, außerdem auch ziemlich stark die Differenztöne 100 und 400 und äußerst schwach den Summationston 1100. Im 2. Falle die Primärtöne 500 und 700, die D.T. 200 und 300 und den S.T. 1200. Wie soll man nun diesen Befund aus den Wellenformen verstehen? Die 1. Welle hat 6 Gipfel von ungleicher Größe, die dem Ton 600 entsprechen mögen. Aber wenn man nun den Ton 5 (500) etwa auf die 5 Gipfel beziehen wollte, die nach Abschneidung des höchsten übrigbleiben, so käme man folgerichtig noch zu 4, 3, 2, 1, wovon 4 und 1 auch gegeben sind, nicht aber 3 und 2. Ebenso findet man bei 5 : 7 7 Gipfel von ungleicher Größe, aber auf den Ton 5 kommt man nur durch spezielle Festsetzungen, die nichts weniger als selbstverständlich sind (m. Abh. 6, S. 259ff. bzw. 80ff.). Der Ton 2 ließe sich allenfalls aus den 2 überragenden Elongationen verstehen, wenn die Plus- und Minusseite jedesmal zusammengekommen werden. Aber für den Ton 3, der durchaus kräftig ist, und für das Nichtvorhandensein von 6, 4, 1 ist wieder guter Rat teuer.

Will man etwa statt der Anzahl der Wellengipfel die Längen der innerhalb einer Periode unterscheidbaren Teilwellen als maßgebend für die hörbaren Töne ansehen, so bedarf es zunächst einer Definition dieses Begriffes der Teilwellen, die nach verschiedenen Gesichtspunkten erfolgen kann. Das natürlichste ist, sie durch die Schnittpunkte der Kurve mit der Mittellinie abzugrenzen. In diesem Fall erhält man bei solchen binären Kurven bis zur Duodezime zweierlei Teilwellen in jeder Periode. Die Länge der einen liegt zwischen den Längen der beiden Sinuswellen, die der anderen eine Oktave höher<sup>1)</sup>. Man müßte also an Stelle der beiden Töne einen Oktaventriller aus 2 ganz anderen Tönen hören. Tatsächlich hört man aber die beiden Primärtöne selbst ruhig fortklingen.

Nun kann man freilich die Einzelheiten der Beobachtungen in Hinsicht der Differenztöne auch unter Voraussetzung der Zerlegung in Sinusschwingungen bisher noch keineswegs vollständig erklären<sup>2)</sup>. Aber aus der zusammengesetzten Schwingungsform als

<sup>1)</sup> S. m. Abh. 6, S. 647ff. (68ff.).

<sup>2)</sup> Die größten Fortschritte in dieser Richtung, experimentelle wie theoretische, verdanken wir WAETZMANN (zuletzt in 2). Er bestätigt HELMHOLTZENS Voraussetzung einer Asymmetrie der Schwingungen für die Entstehung von Kombinationstönen, zeigt aber, daß solche Asymme-

solcher sind ja nicht einmal die Primärtöne verständlich, ergeben sich vielmehr direkte flagrante Widersprüche mit den Beobachtungen. Auch ist der Summationston (der nicht etwa auf Obertönen beruht) bisher nur aus HELMHOLTZENS Voraussetzungen abzuleiten.

Bedingen die zusammengesetzten Formen als solche unmittelbar den akustischen Eindruck, so müßte man auch erwarten, daß die wichtigsten typischen Unterschiede in diesen Formen sich als typische Unterschiede in der Gesamterscheinung des Klanges irgendwie geltend machten. Dies ist aber nirgends der Fall, wie mir eine ausführliche Vergleichung und Diskussion aller durch die Zahlen 1—12 ausdrückbaren Kombinationen zweier Sinusschwingungen gezeigt hat<sup>1)</sup>. So gehören z. B. rein der Form nach in gewisser Hinsicht (man kann da natürlich sehr verschiedene Eigenschaften zugrunde legen) 2 : 3 und 4 : 5 zu derselben Klasse wie 4 : 9 und 5 : 11. In anderer Hinsicht gehört wieder 1 : 5 mit 3 : 11, 1 : 3 mit 5 : 11 zusammen usf., während man rein phänomenal keinen Anlaß hätte, die bezüglichen Klangerscheinungen unter einen gemeinschaftlichen Begriff zu fassen. Bei allen Kurven mit den Schwingungsverhältnissen  $n : n + 1$  senken sich die Gipfel in ähnlicher Weise wie bei 5 : 6 gegen die Mitte, um dann wieder zu steigen, und alle Kurven mit den Verhältniszahlen  $n : n + 2$  sind in ähnlicher Weise symmetrisch wie 5 : 7, während doch die zugehörigen Erscheinungen mit wachsenden Verhältniszahlen von den vollkommenen Konsonanzen zu den ausgesprochensten Dissonanzen übergehen.

Dieser letzte, für die Entwicklung der Musik grundlegende Unterschied der Gehörseindrücke ist freilich auch wieder nicht ohne weiteres mit der Zerlegung gegeben, aber es ist dann wenigstens die Bahn frei, indem das Vorhandensein der Tonempfindungen selbst begreiflich wird, zwischen denen Konsonanz oder Dissonanz stattfinden soll. Wollte man etwa nur die Einfachheit und Regelmäßigkeit der zusammengesetzten Wellenformen dafür verantwortlich machen, so würden z. B. die entschiedenen Dissonanzen 5 : 9 und 5 : 11 den Hauptkonsonanzen näherstehen als das Intervall 7 : 10 (beispielsweise  $c : \bar{g}es$ ), welches doch schon ebenso

trien physikalisch auf sehr verschiedene Weise zustande kommen, und daß starke Differenztöne nur durch starke Asymmetrien, wie bei einer einseitig belasteten Gummimembran, erzeugt werden. Aus der so „gleichgerichteten“ zusammengesetzten Schwingung lassen sich nach W. durch Analyse (häufig auch schon durch bloßen Augenschein) die starken D.T. ableiten. Ob die oben S. 296 Anm. angedeuteten Fragen dadurch schon restlos zu lösen sind, scheint mir allerdings, namentlich bezüglich der D.T. höherer Ordnung, noch zweifelhaft.

<sup>1)</sup> Abh. 6, bes. S. 265ff. (86ff.).

wie 5 : 7 eine halbe Konsonanz ist. Keine Gestaltpsychologie, dünkt mich, kann über solche Diskrepanzen hinweghelfen, solange sie sich nur an die Wellenformen hält.

Auch das Auftreten von Schwebungen ist schwerlich ohne vorherige subjektive Zerlegung der zusammengesetzten Welle zu verstehen. Denn fände eine solche nicht statt, so müßten nicht bloß bei benachbarten, sondern auch bei weiter abstehenden Tönen Rauigkeiten auftreten, da sich in einer zusammengesetzten Welle immer höhere und niedrigere Gipfel, also in diesem Sinne Maxima und Minima, befinden. Und wenn durch die wachsende Zahl der Stöße Grenzen der Wahrnehmung gesetzt werden, so wäre doch kein Grund einzusehen, warum die Höchstzahl, bei der die Rauigkeit verschwindet, sich mit der absoluten Höhe verschieben müßte, wie dies tatsächlich der Fall ist. Die Entstehung der Schwebungen muß vielmehr darauf beruhen, daß 1. durch jeden physikalisch einfachen Ton eine gewisse Zone des Perzeptionsorgans im inneren Ohre erregt wird, daß 2. bei Einwirkung zweier nicht zu weit voneinander entfernten Töne diese Gebiete ineinander übergreifen, wobei daher innerhalb der gemeinschaftlichen Zone die Form der zusammengesetzten Schwingung zur Geltung kommt, daß endlich 3. ein solches Übergreifen in den höheren Lagen noch bei größeren Differenzen (aber abnehmenden Verhältnissen) der Schwingungszahlen stattfindet<sup>1)</sup>.

Demnach wird man auch die 3. These nicht wieder aufgeben können. Eine physiologische Zerlegung in Komponenten, wie sie das geübte Ohr wahrnimmt, ist nicht bloß als Hypothese, sondern als unerläßliche Forderung festzuhalten. Irgendwo hinter dem Trommelfell muß eine klangzerlegende Einrichtung existieren. Was dem Bewußtsein unmittelbar an Klangmaterial gegeben wird, kann nicht absolut einfach sein, sondern muß zum mindesten schon gewisse Bruchlinien aufweisen, nach denen es, wenn sich die analysierende Aufmerksamkeit darauf richtet, zerfällt. Die Möglichkeit des Heraushörens gerade dieser und keiner anderen Klangteile muß schon vor der letzten Schicht, an die unmittelbar Bewußtsein geknüpft ist, geschaffen werden, wenn auch für das

<sup>1)</sup> Vgl. Tonps. II, S. 455ff. AUERBACH findet es (5, S. 621ff.) nicht verständlich, warum mir Schwebungen bei großen Intervallen, wie sie namentlich von KOENIG behauptet wurden, bei einfachen Tönen ausgeschlossen erscheinen. Der Grund liegt in den obigen Erwägungen. Leitet man Schwebungen direkt aus den Wellenformen her, was dem Physiker zunächst natürlich erscheinen mag, dann sind allerdings Schwebungen auch bei beliebig großen Intervallen, z. B. 100 : 1200, zu erwarten, und zwar müßten in diesem Falle 100 (nicht 1100) resultieren, die man sehr gut müßte hören können. Aber die Beobachtung zeigt davon bei einfachen Tönen nichts.

wirkliche Heraushören und die relativen Intensitäten der herausgehörten Bestandteile immer noch besondere Bedingungen und Gesetzmäßigkeiten maßgebend sein werden.

Nur die 4. These ist noch nicht über das Stadium der Hypothese, wenn auch einer der brauchbarsten im ganzen Gebiete der Sinnesphysiologie, hinausgewachsen. Es fragt sich noch immer, ob die Zerlegung durch Resonanz im eigentlichsten Sinne erfolgt, ob sämtliche sukzessiv unterscheidbaren Tonhöhen durch besondere Gebilde in der Schnecke vertreten sein müssen, und welches diese Gebilde sind. Die meisten der gegen die Hypothese gerichteten Einwände konnten jedoch entkräftet werden (vgl. WAETZMANN 1 und 3). Verschiedentlich wurden Modelle gebaut, um durch möglichst getreue Nachbildung der Verhältnisse in der Schnecke ihre analysierenden Leistungen zu veranschaulichen. Am schwersten dürfte noch immer der Nachweis fallen, wie so winzige Teilchen wie die Fasern der Grundmembran, auch wenn ihre starke Belastung durch den anatomischen Zusammenhang mit in Rechnung gezogen wird, auf die für uns hörbaren Schallwellen, besonders die der tieferen und tiefsten Töne, ansprechen können<sup>1)</sup>. Von besonderer Wichtigkeit wären daher Beobachtungen, wie sie ALFR. MAYER und HENSEN anstellten, dieangaben, direkt an den Hör- oder Fühlhärchen niederer Tiere (Krebse, Insekten) Reaktionen auf bestimmte Töne beobachtet zu haben. Die Wiederholung solcher Beobachtungen wäre sehr erwünscht, um die tatsächliche Resonanz so winziger Gebilde auf Schallwellen, seien es auch nur die höchsten des Galtonpfeifchens, sicherzustellen. Vielleicht handelt es sich doch in der Schnecke nur um ein Analogon der physikalischen Resonanz, eine „physiologische Resonanz“ (HERMANN), für welche die Pendelschwingungen der Teilchen im Labyrinth nur ein vorläufiges anschauliches Bild abgeben.

Außer der Grundmembran und den CORTISCHEN Pfeilern, oder beiden in Verbindung miteinander, sind auch die Deckmembran die Hörhaare, die Fasern des akustischen Nerven, endlich die Gehirnganglien, für die Zerlegungsfunktion herangezogen worden. Die meisten Fürsprecher findet aber immer wieder die Grundmembran. Über die Eigenschaften, die sie zu dieser Aufgabe besonders zu befähigen scheinen, vgl. HARTRIDGE, Brit. Journ. of Psych. Bd. 11—13 (1921—1922). STEFANINI, Arch. ital. di otol., rinol. e laringol. Bd. 34. 1923. G. WILKINSON und A. A. GRAY, The Mechanism of the Cochlea. 1924 (wo auch ein ausgearbeitetes Modell beschrieben ist). M. GIESSWEIN, Die mechanischen Verhältnisse der Basilmembran usw. in Passow-Schaeffers Beitr. z. Anatomie usw. Bd. 22. 1925. Derselbe, Die

<sup>1)</sup> Daß bei Kindern schon im 2. Lebensjahre Vokalverständnis, ja sogar manchmal Leistungen eines guten musikalischen Gehörs auftreten, bedeutet keine weitere Erschwerung der Theorie, da die Fasern der Grundmembran nach HENSEN schon ganz früh ihre definitive Länge erreichen.

statische Struktur des Cortischen Organs. Zeitschr. f. Hals-, Nasen- u. Ohrenheilk. Bd. 12. 1925. K. L. SCHAEFER und M. GIESSWEIN im Handbuch der Hals-, Nasen- u. Ohrenheilk. v. DENKER und KAHLER Bd. 6, S. 462 ff. 1926. H. HELD, Die Cochlea usw. im Handbuch d. Physiol. v. BETHE u. a. Bd. 11, S. 467 ff., bes. 526 ff. 1926. E. WAETZMANN, Hörtheorien, das. S. 667 ff.

Während alle diese Autoren entschieden für die Resonanzfunktion der Basilarmembran eintreten, versucht F. SPECHT (Physiol. des Hörens, Arch. f. Ohren-, Nasen- u. Kehlkopfheilk. Bd. 114. 1925) eine radikale Umbildung der Vorstellungen sowohl über die Leitung im Mittelohr als die Vorgänge in der Schnecke. Er läßt einfach die Druckschwankungen der Luft, in denen der Ton besteht, sich in das Labyrinthwasser bis zu den Hörzellen fortpflanzen. Aber das Problem des Nebeneinanderhörens und alle Modalitäten dieses Nebeneinander scheinen mir dabei wieder zu kurz gekommen zu sein. Die Form der zusammengesetzten Schwingung, auf die er zurückgreift (S. 64), reicht dazu nicht aus.

Die vereinzelt behauptungen über Erhaltung der Hörfähigkeit nach Exstirpation der ganzen Gehörknöchelchenkette (worauf SPECHT seine Schlüsse gründet), ja sogar nach Zerstörung oder Exstirpation des ganzen Labyrinths (EWALD, WUNDT, neuerdings auch eine ganz nebenbei erwähnte Beobachtung bei MINTON, Proc. Nat. Acad. of Sciences U. S. A. Bd. 8, S. 278) können vorläufig bei weitem nicht unter die gesicherten Tatsachen gerechnet werden. Bei der labyrinthlosen Taube — wo es WUNDT für zweifellos erklärte, daß sie „trotz der bereits teilweise eingetretenen Atrophie des Hörnerven genau ebenso (!) auf Schallreize reagierte wie ein im Vollbesitz seiner peripherischen Sinnesapparate befindliches Tier“ — wird diese Meisterleistung doch wohl auf Tastempfindungen beruht haben.

Die wunderbare Widerstandsfähigkeit des Gehörorgans gegen funktionelle wie pathologische Attacken ist dem Verfasser am eigenen Leibe offenbar geworden. Weder die durch Wochen fortgesetzte stundenlange Beobachtung schriller Schwebungen (die nach HELMHOLTZ dem Hörnerven schädlich wären), noch ebenso ausgedehnte Untersuchungen schneidender höchster Töne der Galtonpfeife („wie wenn ein Draht durch den Kopf gezogen würde“ — sagt PREYER), noch endlich eine dreimalige lebensgefährliche Erkrankung des rechten Ohres, die erste mit Splitterungen im Felsenbein und peinigenden Absonderungen nach dem Mittelohr, die zweite und dritte mit eiteriger Mittelohrentzündung, beide Male mit Perforation des Warzenfortsatzes nach vielen Parazentesen des Trommelfelles, vermochten das Gehör zu schädigen. Die letzte Erkrankung fiel in die Zeit des Beginnes vorliegender Untersuchungen. Dem leider inzwischen verstorbenen Kollegen PASSOW werde ich stets Dank wissen, daß er in den ersten Kriegswochen, als es ihn drängte, draußen zu helfen, die Behandlung leitete, die zur vollen Genesung führte. Das rechte Ohr hört sogar noch etwas feiner als das linke, was mir früher nicht aufgefallen ist. Aber hätte er die Gehörknöchelchen oder gar die Schnecke ausräumen müssen, so wäre dieses Buch doch wohl ungeschrieben geblieben.

In diese nun schon über 60 Jahre währende Diskussion treten wir hier nicht näher ein, da die Sprachlaute nicht mehr Material dafür bieten als beliebige andere Klänge. Es sei jedoch bemerkt, daß die Verschiedenheiten der Lösungsversuche zum großen Teil auch mit verschiedener Problemstellung zusammenhängen, indem man verschiedene akustische Tatsachen als vorzugsweise er-

klärungsbedürftig betrachtet. Viele verkennen doch die Fülle der zu erklärenden Einzelheiten und vor allem das Hauptproblem, wie es HELMHOLTZ selbst vorschwebte.

Dieses liegt nicht etwa in der Unterscheidung isoliert oder in bloßer Aufeinanderfolge gegebener Töne, sondern durchaus in der Zerlegung zusammengesetzter Schwingungen. Wäre unser Gehör dazu nicht fähig, so bedürften wir überhaupt keiner weiteren Einrichtungen im Ohre, als solche zur Weiterleitung des Reizes und zu seiner Umwandlung in den Nervenvorgang erforderlich sind. Daß verschiedene Wellenlängen verschiedene Empfindungen erregen, ist ebenso leicht begreiflich, wie daß die Unterschiedsempfindlichkeit ihre Grenze hat. Man hat die Anzahl der Fasern in der Grundmembran der Schnecke, die WALDEYER auf 18 000 schätzte, verglichen mit der Anzahl der sukzessiv unterscheidbaren Töne: aber für diese Leistung brauchte die Grundmembran überhaupt keine Faserung zu besitzen. Es wäre ferner auch für HELMHOLTZ selbst nicht einmal nötig gewesen, seine ursprüngliche Annahme, daß die CORTISchen Pfeiler die analysierenden Teile seien, bloß um deswillen aufzugeben, weil diese Gebilde den Vögeln und Amphibien fehlen: denn niemand weiß, ob Vögel und Amphibien zur Klangzerlegung befähigt sind.

Nachdem aber nun einmal Klangzerlegung beim Menschen vorkommt und das gesuchte Organ diese Leistung vollbringen muß, so muß es allerdings auch noch zu zahlreichen anderen Leistungen fähig sein. Und hier ist die Stelle, wo die rein phänomenologische Forschung wieder mitspricht, wo sie die Liste der Forderungen, d. h. der Tatsachen aufstellen muß, denen eine gute Hypothese zu genügen hat. Wir wollen im folgenden wenigstens die wesentlichsten darunter zusammenstellen:

a) Die genaue Übereinstimmung der bei der subjektiven Klangzerlegung auftretenden Teiltöne mit den objektiven Komponenten, wie sie die FOURIER-Analyse ergibt. Die musikalische Erfahrung würde in dieser Beziehung viele Abweichungen nahelegen, da wir an die temperierte Stimmung gewöhnt sind, außerdem auch die Teiltöne 7, 11 u. a. in dieser sog. natürlichen Stimmung nicht mit musikalisch gebräuchlichen Intervallen zusammenfallen. Sie werden gleichwohl alle in der „natürlichen Stimmung“ gehört.

b) Das Heraushören auch unharmonischer Teiltöne, wenn entsprechende Komponenten objektiv gegeben sind. In dieser Beziehung weicht das Ohr von der FOURIER-Analyse ab oder geht über sie hinaus, denn dort gibt es nur harmonische Komponenten, die Amplituden der unharmonischen verteilen sich auf die benachbarten harmonischen. (Solche Bestandteile heben allerdings, genau

gesprochen, die strenge Periodizität des Schwingungsvorganges und damit die Voraussetzung der FOURIER-Analyse im Prinzip auf; aber wenn es sich nur um höhere und schwächere unharmonische Teiltöne handelt, wird der periodische Charakter doch genügend gewahrt bleiben, um die Rechnung anzuwenden. Für die Resonanztheorie besteht hier überhaupt keine Schwierigkeit).

c) Die Schwebungen in allen ihren Details, auch den Schwebungstönen (bzw. Zwischentönen).

d) Die Differenz- und Summationstöne, auch diese reich an Einzelheiten.

e) Das mit der Höhe stetig abnehmende Volumen der einfachen Töne.

f) Die Gesamtheit der Schwellentatsachen in bezug auf gleichzeitige und aufeinanderfolgende Töne.

g) Die auffallende Vertiefung eines Stimmgabeltones durch objektive Verstärkung (z. B. bei Annäherung).

h) Die gegenseitige subjektive Beeinflussung gleichzeitiger Töne in ihrer Stärke.

i) Die Unterschiede des zweiöhrigen (amphotischen und dichotischen) Hörens vom einöhrigen.

k) Die subjektiven Töne, von etwa  $c^1$  aufwärts in allen Tonhöhen vorkommend<sup>1</sup>).

l) Die pathologischen Verstimmungen (Doppelthören usw.).

m) Die normalen und pathologischen Ausfallserscheinungen, speziell die von der oberen Tongrenze abwärts schreitenden und die Tonlücken.

n) Die bei Tieren durch experimentelle Zerstörung einzelner Teile der Schnecke (durch übermäßige akustische Reizung oder anatomischen Eingriff) entstehenden Tonlücken<sup>2</sup>).

<sup>1</sup>) Bei vielen Personen scheinen solche Erscheinungen fast gar nicht vorzukommen, bei manchen dagegen, zumal wohl Neurasthenikern, sehr häufig. Ausführliches darüber in m. Abh. 5. Die Theorie hat sie bisher fast unberücksichtigt gelassen, obgleich K. L. SCHAEFER 1909 (Verhandl. d. dtsh. otol. Ges. S. 22 ff.) die ärztlichen Fachgenossen darauf hingewiesen hat. Erst bei J. P. MINTON (Tinnitus etc., Phys. Review Bd. 22. 1923) finde ich eine Studie, die ihrer theoretischen Bedeutung Rechnung trägt. Die Verdeckung objektiver durch subjektive Töne ist da auch experimentell untersucht. Die subj. Geräusche, die immer ausgiebig behandelt wurden (auch neuestens von GRÜNBERG), sind für den Arzt als Symptome ja wohl bedeutungsvoller und für den Patienten zumeist lästiger. Aber theoretisch wären die scharf definierbaren subj. Töne wichtiger. Manches leitet darauf hin, ihren Entstehungsort in den Hörhärchen zu suchen.

<sup>2</sup>) H. HELD und F. KLEINKNECHT erzeugten zuletzt durch eine scharf begrenzte Operation an der Basilarmembran eine Lücke von nur 2 Tönen. Handb. d. Physiol. v. BETHE usf. Bd. 11, S. 532. 1926.



o) Der physiologische Unterschied von Tönen und Geräuschen.

p) Die spezifischen oder zum mindesten graduell außerordentlich großen Unterschiede des Gehörs vom Gesicht in bezug auf Abklingen, Nachbilder, Ermüdung, Induktion und Kontrast (der im eigentlichen Sinne, als gegenseitige qualitative Veränderung, bei Tönen fehlt).

Einzelne dieser Tatsachengruppen, besonders die letzte, mögen allerdings erst im Gehirn, speziell in der Hörsphäre, ihre Wurzeln haben. Sicher gilt dies aber von den musikalisch wichtigsten Erscheinungen, dem Auftreten der „musikalischen Qualitäten“ in den mittleren Lagen des Tonreiches, sowie dem Unterschiede von Konsonanz und Dissonanz und ihren Graden, für den die Schwebungen in keinem Fall ausreichen<sup>1)</sup>, und den damit parallellgehenden Verschmelzungsstufen der Intervalle innerhalb der Oktave. Auch der so auffallende Gehörs- und Gefühlsunterschied zwischen musikalischen und unmusikalischen Personen hat mit dem peripherischen Organ gewiß nur wenig zu tun (man ist nicht auf dem einen Ohr musikalisch, auf dem anderen bei übrigen normaler Hörfähigkeit unmusikalisch). Und so muß gewiß noch manches auf Rechnung der zentralen Organisation geschrieben werden. Aber zunächst wird man zweckmäßigerweise immer im peripherischen Organ nachsuchen, um nicht die Erklärung in ein noch dunkleres Gebiet zurückzuschieben.

Es bleibt immer erstaunlich, in welchem Umfange die HELMHOLTZsche Hypothese den obigen zahlreichen Forderungen entgegenkommt, wenn auch öfters nur mit Hilfe von Zusatzhypothesen. Alle anderen Vorstellungsweisen, auch die ansprechende EWALDSche Schallbilderttheorie, bleiben weit davon entfernt. In einem Punkte hat man aber die Anforderungen an sie meistens sogar unnötig überspannt. HELMHOLTZ selbst war keineswegs der Meinung, daß auf einen einfachen Ton nur eine einzige Faser resoniere; er wußte natürlich, daß infolge der Resonanzbreite eine größere Anzahl auf einmal, und zwar bei Tönen von nicht verschwindender Stärke in gleicher Periodik wie der erregende Reiz, schwingen muß, also immer eine ganze Zone der Basilmembran

<sup>1)</sup> S. m. Beitr. z. Akust. u. Musikwiss. H. 1. Über speziellere Gründe für den zentralen Ursprung der Konsonanzerscheinungen vgl. auch das. H. 6, S. 120ff. (Zeitschr. f. Psych. Bd. 58, S. 325ff.).

In positiver Hinsicht scheint es mir jetzt richtiger, für die Definition statt der Verschmelzungstatsachen die der Verwandtschaft einfacher Töne zugrunde zu legen, die aber ebenso wie die Verschmelzung nur zentral-physiologisch, also hypothetisch, zu erklären ist. Eine solche Hypothese hat kürzlich v. HORNBOSTEL (2) sehr scharfsinnig und eingehend entwickelt. Doch bleiben daneben zunächst noch andere Vorstellungen denkbar.

auf den Reiz antwortet. Diese muß in gleichem Rhythmus schwingen wie der Reiz selbst (erzwungene Schwingungen). Und nicht minder ist anzunehmen, daß die den einzelnen Fasern entsprechenden spezifischen Energien im Gehirn, von denen die Höhe der gehörten Töne zuletzt abhängt, sich zu einem einzigen Erregungszustande vereinigen, da wir tatsächlich (abgesehen von der tiefsten Region, wo ein objektiv einfacher Ton eine kleine Tonhöhenstrecke simultan auszufüllen scheint) nur eine einzige, genau abgegrenzte Tonhöhe vernehmen („Akkommodation der spezifischen Energien“<sup>1)</sup>).

Nun aber ist die Frage, wie viele solcher hinreichend abgegrenzten Zonen sich gleichzeitig in der Grundmembran zu bilden vermögen. Offenbar nur so viele, als ohne weitgehendes gegenseitiges Übergreifen, also ohne Erzeugung stärkerer Schwebungen, gleichzeitig erregt werden können. Diese Anzahl wird verschieden sein, je nach der Stärke der Reize, bei schwachen größer als bei starken. Bei einem mäßig starken Klange mit nach der Höhe immer schwächer werdenden Obertönen ist die Anzahl dieser Zonen immerhin auf etwa 20 zu schätzen<sup>2)</sup>. Die Teiltöne von sehr hohen Ordnungszahlen werden aber auch bei schwacher Erregung wegen ihres dichten Beisammenliegens nicht leicht ganz schwebungsfrei bleiben. Man könnte auch künstlich eine Maximalzahl

<sup>1)</sup> S. Tonps, II, 111 ff. Die neuerdings mehrfach vertretene Anschauung, daß die stärksterregte Faser die Qualität der Empfindung bestimme (WILKINSON und GRAY, BARKHAUSEN und LEWICKI, BUDDE), scheint mir auf dasselbe hinauszukommen.

Manche halten die Annahme spezifischer Energien für die einzelnen Töne überhaupt für überflüssig, gerade weil nach der Resonanzhypothese schon der Prozeß im peripherischen Organ für verschiedene Töne verschieden sei (WEINMANN unter Zustimmung NAGELS). Aber die pathologischen Verstimmungen scheinen darauf hinzuweisen, daß auch eine verstimmte, mit veränderter Schwingungszahl schwingende Faser infolge der festen spezifischen Energien des Hörzentrums die gleiche Tonempfindung wie in ihrem normalen Zustande zur Folge hat. Richtete sich die empfundene Tonhöhe bedingungslos nach dem Erregungszustande im Organ, der seinerseits vom objektiven Reiz durch Resonanz bestimmt wird, so wäre ein Grund für Falschhören nicht einzusehen (I, I, 274 ff., II, 108 ff.).

<sup>2)</sup> Hiermit hängt es offenbar zusammen, daß man bei der Synthese der Vokale die hohen Teiltöne nicht lückenlos einfügen darf, wenn nicht unnatürlich rauhe Klänge herauskommen sollen (s. o. S. 178). Es bildet sich eben im Ohre nur eine geringere Zahl von Schwingungszonen, als für eine lückenlose Reihe erforderlich wäre. Ja auch die Tatsache, daß man aus musikalischen Klängen unter den höchsten Obertönen durch willkürliche Aufmerksamkeit immer nur einzelne, im Oktavenverhältnis zu tieferen stehende Teiltöne heraushört (s. o. S. 13, Anm.), könnte damit zusammenhängen. Die analysierende Aufmerksamkeit ist in ihrer Richtung und Wirkung durch die musikalischen Gewohnheiten des Hörenden mitbestimmt,

gleich starker einfacher Töne, von den tiefsten bis zu den höchsten, herstellen, deren Abstände so abgestuft würden, daß die obige Bedingung bei gleichzeitigem Erklingen erfüllt wäre. Auch so dürften kaum mehr als etwa 25 herauskommen.

Diese Erwägungen und Kriterien scheinen mir für die weitere Ausgestaltung der HELMHOLTZschen Grundvorstellungen nicht ohne Belang, sofern sie die an die anatomische und funktionelle Selbständigkeit der Membranteile zu stellenden Anforderungen präzisieren.

Eine andere Frage ist es, wie viele von diesen selbständig schwingungsfähigen Fasergruppen sich nun auch in der Sinneserscheinung selbst als gleichzeitige und gleichzeitig unterschiedene Töne geltend machen. Tatsächlich können auch Gutmusikalische kaum mehr als 4—5 Töne streng gleichzeitig durch die Aufmerksamkeit deutlich nebeneinander wahrnehmen; der übrige Bestand eines länger dauernden Klanges oder Zusammenklanges wird nur sukzessive erfaßt. Aber daraus darf man natürlich nicht schließen, daß die Sonderung der Erregungen im Gehörorgan sich nur auf 4—5 Zonen der Basilarmembran zu erstrecken brauche. Denn diese analysierende Bewußtseinsleistung vollzieht sich zweifellos erst im Gehirn, das Ohr hat nur die Vorarbeit zu leisten. Es hat die durch Aufmerksamkeitswanderung unterscheidbaren Teile zur jederzeitigen Auswahl richtig gestimmt bereitzuhalten. Die physiologische Sonderung der Komponenten kann nicht erst die Folge oder Begleiterscheinung des Heraushörens sein: die Klänge und Zusammenklänge müssen bereits ihre ganz bestimmte Struktur besitzen, ehe die aussondernde Tätigkeit des Bewußtseins beginnt, und es müssen so viele selbständig erregte Teile nebeneinander existieren, als Komponenten bei genügender Dauer des Klanges herausgehört werden können.

Für Geräusche wäre gemäß ihrer physikalischen Beschaffenheit im Rahmen der Resonanzhypothese ein über ein breites Gebiet der Membran stetig ausgedehnter, in den einzelnen Punkten aber schwacher und von Punkt zu Punkt verschiedener Erregungszustand anzunehmen. In der letzteren Eigenschaft würde der wesentlichste physiologische Unterschied gegenüber den Tönen bestehen: bei diesen, selbst sehr schwachen noch hörbaren Tönen, versetzt ein Sinuswellenzug von bestimmter Wellenlänge

---

wenn es sich auch keineswegs um ein bloßes Hineinhören handelt, sondern nur das herausgehört werden kann und mit sinnlicher Lebhaftigkeit herausgehört wird, was tatsächlich darin enthalten ist. Ein derart begrenzter selektiver Einfluß der Aufmerksamkeitsrichtung auf das Hervortreten bestimmter Zonen der Basilarmembran während des Schwingungsprozesses erscheint mir wohl denkbar.

eine größere Anzahl benachbarter Fasern in einen identischen Schwingungszustand, während bei den reinen tonlosen Dauergeräuschen wahrscheinlich jede Faser durch den ihr adäquaten schwachen Reiz zu ihrer Eigenschwingung angeregt wird. Im Gehirn aber summieren sich die Intensitäten der einzelnen Geräuschteile, während bei den Tönen im wesentlichen vielmehr eine Subtraktion, und zwar allem Anscheine nach schon in der Schnecke selbst, stattfindet (12. Kap.). Jene Summation verhindert aber nicht, daß in einem beschränkten Maße auch Geräusche subjektiv in mehrere (2—3) gleichzeitige gröber unterschiedene Gruppen zerlegt werden können (o. S. 128, 135).

Wenn nun KOEHLER die OHM-HELMHOLTZsche Theorie eine „bloß summative“ nennt, so trifft eine solche Benennung auch auf eine in dieser Weise ausgestaltete Vorstellung insofern zu, als sie den Klang in der Schnecke aus einer Summe relativ selbständiger, räumlich gesonderter Teilschwingungen bestehen läßt. Wenn er aber eine solche bloß summative Theorie für ungenügend erklärt, um den einheitlichen Eindruck und die qualitativen Verschiedenheiten der Vokale, der Klänge überhaupt, daraus zu verstehen, so möchte ich glauben, daß die beiden großen Forscher ihr Ziel so hoch überhaupt nicht gesteckt hatten. OHM, der sich selbst für ganz und gar unmusikalisch erklärte und der einen gegensepielenden Freund zu Hilfe rufen mußte, um an seiner Stelle eine Beobachtung über die Oktave zu machen, wollte sicherlich nicht behaupten, daß ihm beim Hören eines Trompetenklanges eine Anzahl verschiedener Teiltöne nebeneinander, sozusagen nur durch ein Pluszeichen verbunden, deutlich unterschieden im Bewußtsein gegeben wären. Aber auch HELMHOLTZ hätte dieser Formulierung schwerlich zugestimmt. Er ließ die Empfindung mit allen ihren Eigenschaften erst in der Hirnrinde zustande kommen und verlangte nur (unseres Erachtens mit Recht), daß schon vorher eine rein physiologische Zerlegung stattfinden müsse, durch welche die Möglichkeit des mehrheitlichen Hörens für das Bewußtsein gegeben und die herauszuhörenden Tonhöhen vorausbestimmt würden. Aber sicher hat er nicht geglaubt, mit dieser Zerlegung im Ohre schon alles erklärt zu haben. Die Regeln, die er für den Zusammenhang der Klangfarbenunterschiede mit den Verschiedenheiten der Teiltonstruktur gibt, sind rein empirisch gefaßt und umgehen die Frage nach dem Warum. Er dürfte erst in der Hirnrinde die Faktoren gesucht haben, die aus der bloßen Summe ein Ganzes machen und dessen sämtliche Eigenschaften endgültig bestimmen. Jedenfalls kann man seine Lehre in dieser Richtung zu ergänzen versuchen, ohne sie um dessentwillen im Prinzip umzustößen.

Nun bleibt nur noch der 5., rein psychologische Streitpunkt. HELMHOLTZ ließ die Mehrheit der Teiltöne, wie er sie im Ohr statuierte, auch in dem Gehörseindruck selbst noch bestehen. Erscheint uns der Klang einheitlich, so werden sie zwar gehört, aber nicht als Teile des Gehörten bemerkt. Er unterschied daher mit LEIBNIZ Perzeption und Apperzeption. Ich halte auch diese Seite seiner Lehre für unbedenklich, ja aus allgemeineren psychologischen Gründen für unumgänglich. Aber sie ist mit seiner physiologischen Hörtheorie nicht untrennbar verwachsen. Wer unbemerkte Teilinhalte unserer Empfindungen leugnet, braucht nur an Stelle jener „Perzeption“ einen reinen Gehirnzustand zu setzen, der keine psychologische Seite hätte, gewissermaßen ein Vorzimmer des Bewußtseins, in dem die Teile als solche noch vorhanden sind, aus welchen dann aber unter bestimmten Umständen erst jener letzte Rindenprozeß hervorgeht, dessen psychisches Äquivalent der bewußte Klangeindruck mit allen seinen Klangeigenschaften ist. Solange nur die hier besprochenen Tatsachen in Frage kommen, hat es keinen Zweck, darüber zu disputieren. Doch verhehle ich nicht, daß mir dem gesamten Tatsachenkreise der Psychologie nur die Anerkennung unbemerkter Teilinhalte (die aber mit der Behauptung unbewußter psychischer Funktionen keineswegs zusammenfällt) zu genügen scheint.

KOEHLER bekämpft die „summative Theorie“ aber nicht bloß wegen der unbemerkten Teilerscheinungen, sondern auch wegen der Unmöglichkeit, daraus Tatsachen wie die Intervall- und Akkordfärbungen und vor allem die Vokalqualitäten abzuleiten. Diese Eigenschaften seien nun und nimmer aus bloßer Summierung der elementaren Toneigenschaften zu verstehen. In dieser Behauptung hat er vollkommen recht. Aber man muß hinzufügen, daß auch die Anerkennung der von ihm selbst beobachteten Vokalitäten der einfachen Töne nicht zu diesem Ziele führt. Denn wie sollten Ö und Ü durch bloße Summierung jener Urvokalitäten zustande kommen? Wie sollten Ä, aber auch schon A und O durch Summierung der Vokaleigenschaften der darin enthaltenen Teiltöne entstehen, bzw. als deren Summe begriffen werden? In der Tat könnten diese ohnehin so schwachen Vokalfärbungen einfacher Töne ebensogut ganz fehlen: das Problem wäre das nämliche. Es müssen eben, worauf K. wieder mit Recht hinweist (3, S. 456, 462), nach der Zerfällung einer Schalleinwirkung in Sinusschwingungen noch ganz spezifische Prozesse hinzukommen, um diese Neubildungen zu erzeugen. HELMHOLTZ hat darüber keine Vermutungen ausgesprochen; wir glaubten, einige Gesetzmäßigkeiten über diese Prozesse formulieren

zu können. Aber wenn seine Voraussetzungen, wie wir sie unter 1—5 aufzählten, keine Bestimmungen darüber einschließen, so hindert dies doch nicht, daß sie innerhalb der Sphäre und für die Tatsachen, auf die sie zugeschnitten sind, gültig bleiben. In allen Fällen dürfen wir auf diese höchst wertvollen sogenannten „summativen“ Vorstellungsweisen nicht eher verzichten, als bis eine neue, noch leistungsfähigere geschaffen ist.

Durchaus zustimmen muß ich aber KOEHLER darin, daß die von mir selbst früher versuchte Erweiterung und Vertiefung der HELMHOLTZschen Klangfarbenlehre durch die Lehre von den Tonfarben, so wie ich sie damals definierte, noch nicht ausreichte, um den Erscheinungen speziell im Gebiete der Vokalfärbungen voll gerecht zu werden. Die Erkenntnis dieser Unzulänglichkeit ist es wohl auch gewesen, von der seine kritische Stellungnahme sich auf die HELMHOLTZsche Grundlegung ausdehnte. Ich hoffe aber, die Lücke durch die voranstehenden Untersuchungen nach Kräften ausgefüllt zu haben; auch wird uns das folgende Kapitel noch einmal auf diese allgemeinere Frage zurückführen.

## 15. Kapitel (Anhang).

### Über Instrumentalklänge.

Obschon die meisten Methoden der Vokalforschung sich ohne weiteres auf Instrumentalklänge übertragen lassen, liegen doch seit HELMHOLTZ, der auch hier den Hauptfortschritt vollzog, abgesehen von rein physikalischen Studien über Saitenschwingungen oder die Luftbewegung in Pfeifen, nur wenige Experimentaluntersuchungen darüber vor; ausführlichere nur von RITZ, MEISSNER, HERRMANN-GOLDAP, KOEHLER (1, I), MILLER. Nach HELMHOLTZ sind bekanntlich für instrumentale Klangfarben nicht die absoluten Höhen der Teiltöne maßgebend, sondern ihre Ordnungszahlen, d. h. bestimmte Intensitätsverhältnisse innerhalb der Reihe, also nach unserem Ausdrucke „bewegliche Formanten“. Es sind aber auch hier gelegentlich feste und sogar unharmonische Formanten behauptet worden (MEISSNER, HERRMANN-GOLDAP). Nach der psychologischen Seite hin versuchten der Verfasser und KOEHLER HELMHOLTZENS Lehre weiter auszubauen.

#### 1. Erkennen instrumentaler Klangfarben.

Es ist sehr gebräuchlich, Klangfarbe als dasjenige zu definieren, was Klänge verschiedener Instrumente bei gleicher Tonhöhe noch voneinander unterscheidet. Aber schon HELMHOLTZ, dann auch der Verfasser, wiesen darauf hin, daß hierbei noch viele andere Kennzeichen mitwirken: begleitende Geräusche infolge der Art der Tonerzeugung, Ansatz und Verlauf des Klanges, durchschnittliche oder maximale Stärke der Tongebung, charakteristische Rhythmen und Figuren der Tonbewegung usf. Diese Momente wollen wir äußere nennen, die in der Teiltonstruktur liegenden innere.

Um nun zunächst einen genaueren Begriff zu bekommen, wieviel von der Urteilssicherheit, die sich bei guten Kennern findet, auf innere, wieviel auf äußere Merkmale zurückzuführen ist, stellte ich (1910) Versuche an, bei denen die vom Ansatz und Verlauf des Klanges abhängigen Kennzeichen dadurch ausgeschlossen wurden, daß nur ein zeitliches Mittelstück des Klanges

ausgeschnitten und 2 Sekunden lang dargeboten wurde. In den gegenüberliegenden Wänden zweier durch einen Korridor getrennten Zimmer (I und II im Schema oben S. 44) wurden Öffnungen von je 20 qcm angebracht, deren Weite die Gefahr einer Veränderung der Klangfarbe, wie sie bei Röhrenleitungen gerade besonders für Instrumente besteht, ausschloß. Diese Öffnungen waren mit gefütterten Klappen versehen, die bei der Intonation des Instrumentes geschlossen waren, dann für die angegebene kurze Zeit geöffnet wurden. Im 1. Zimmer wurden in bunter Reihenfolge die Klänge folgender Instrumente angegeben: Flöte, Fagott, Posaune, Oboe, Klarinette, Tenorhorn, Waldhorn, Trompete, Kornett, Violine, Cello, Resonanzstimmgabel. Unter diesen stehen die der Blechinstrumente einander dem Klange nach relativ nahe, sonst aber sind die Verschiedenheiten groß genug. Daß diese Instrumente vorkommen konnten, war den Urteilenden bekannt, aber nicht, welche davon in einer bestimmten Versuchsreihe vorkamen und in welcher Ordnung. Die Tonhöhen lagen in der kleinen bis 2-gestr. Oktave, je nachdem sie für die jeweiligen Instrumente am meisten gebräuchlich und charakteristisch waren<sup>1)</sup>. Die Urteilenden waren: H., vorzüglicher Akustiker und Musikforscher; F., gleichfalls akustisch gut vorgebildet und zugleich Komponist; B., Klavierlehrer an der Hochschule für Musik und Dirigent; Ha., Militärmusikdirektor; M., Instrumentenmacher und Lieferant von Blasinstrumenten für Militärmusik.

2 Vorreihen, bei denen nur M. urteilte und nur Blasinstrumente vorkamen, die er selbst geliefert hatte, ergaben unter 22 Fällen nur 7 Fehlerurteile (Flöte verwechselt mit Oboe, Waldhorn mit Posaune, Trompete mit Kornett und Klarinette, sowie Kornett mit Trompete und Oboe). Eine 3. Reihe lieferte unter 27 Fällen bei F. nur 1 Fehlerurteil, bei H. dagegen 12, bei M. 13. Eine 4. Reihe unter 37 Fällen bei Ha. 15, bei B. 17, bei H. 24 Fehlerurteile; darunter nicht bloß naheliegende Verwechslungen, wie Flöte mit Stimmgabel, Kornett mit Trompete, Posaune mit Horn, Klarinette mit Oboe, Fagott mit Cello, sondern auch solche zwischen Stimmgabel und Trompete bzw. Kornett, Violine und Waldhorn bzw. Fagott, Flöte und Fagott,

<sup>1)</sup> Man könnte hier einen Versuchsfehler finden, sofern immer dieselbe Tonhöhe hätte benützt werden müssen. Dies wäre auch richtig, wenn es sich um eine vergleichende Analyse der Klangverschiedenheiten gehandelt hätte. Aber hier kam doch auch in Betracht, daß die verschiedenen Instrumente ihre bezeichnendsten Farben in verschiedener Tonlage haben. Andererseits mußte auch vermieden werden, daß aus der Lage selbst Schlüsse gezogen wurden. Es wurde daher ein Mittelweg eingeschlagen, indem die disponiblen Töne in bestimmte mittlere Grenzen eingeschränkt wurden.

Auch die Stärke der Klanggebung wurde in mittleren Grenzen gehalten.



Violine und Oboe. Einer hörte mit Vorliebe ein Englisch-Horn, obgleich dieses niemals vorkam, ein anderer wieder elektrische Gabeln, je nach der subjektiven Einstellung. Am besten wurden Flöte, Klarinette, Fagott und Waldhorn erkannt.

Aus diesen etwas summarischen Versuchen, die nur eine erste Vorstellung von dem Sachverhalt geben sollten, geht doch soviel hervor, daß die individuellen Unterschiede im Erkennen von Instrumenten aus der reinen Klangstruktur auch unter sehr Musikalischen und Geübten noch bedeutend sind, daß manche selbst unter so schwierigen Umständen große Sicherheit, andere aber unerwartete Ratlosigkeit zeigen, daß also nur aus dem Zusammenwirken der äußeren Klangmerkmale mit den inneren bei solchen, die zeitlebens ihre Aufmerksamkeit den Eigentümlichkeiten der einzelnen Instrumente zugewandt und sich etwa auch mit der Handhabung der wichtigsten vertraut gemacht haben, eine hervorragende Urteilssicherheit entsteht. Man darf daher auch nicht erwarten, daß die akustische Analyse hier so einfache und übersichtliche Verhältnisse zutage fördere, wie bei den Vokalen, deren Unterscheidung auch bei kurzer Darbietung den meisten Erwachsenen keine Schwierigkeiten macht.

## 2. Hauptergebnisse bezüglich der Klangstrukturen.

Es wurden nun, nachdem die Vokaluntersuchungen zu einem gewissen Abschlusse gediehen waren (Anfang 1916), die dort angewandten 3 Methoden: Resonanzversuche mit nachklingenden Stimmgabeln, Interferenzversuche, Synthesen aus einfachen Tönen, auch auf Instrumentalklänge ausgedehnt. Aber die Mannigfaltigkeit der Instrumente ist gewaltig, auch wenn man sich auf die der gegenwärtigen europäischen Musik beschränkt — denken wir nur an die noch immer wachsende Menge der Orgelregister mit den wunderbarsten Klangfarben, deren jedes eigentlich ein besonderes Instrument darstellt —, und man hat sie nicht so leicht wie die Vokale zur jederzeitigen Verfügung im Laboratorium. Vom Klavier mußte abgesehen werden. Darum blieb die Untersuchung hier fragmentarisch<sup>1)</sup>. Aber auch die Anwendung unserer Methoden machte hier öfters Schwierigkeiten, von denen noch zu reden sein wird. Gleichwohl scheinen gewisse allgemeine Züge daraus hervorzugehen, die eine Stellungnahme zu den oben erwähnten Fragen gestatten und zugleich aufs neue die gemeinschaftlichen Grundlagen der vokalen und instrumentalen Klangfarben erkennen lassen.

<sup>1)</sup> Sie wird gegenwärtig von Dr. ERICH SCHUMANN, der mir auch bei den letzten Kontrollversuchen 1926 geholfen hat, fortgesetzt.

Es wurden untersucht: A- und B-Klarinette, Fagott und Kontrafagott, Tenorposaune, B-Trompete, Waldhorn in F, Flöte und Pikkolo, Oboe, Viola (als ein mittleres Streichinstrument) und der Klang eines akustischen Zungenapparates von der Art der bekannten APPUNNSchen Tonmesser. Wir stellen sogleich die wichtigsten Ergebnisse voran:

a) Musikalische Instrumente, besonders Blasinstrumente, haben oft eine unerwartet große Zahl von Teiltönen. In der tiefen Lage können es über 30 sein.

b) Auch hier fanden sich bei den untersuchten Instrumenten nur harmonische Teiltöne.

c) Auch hier waren mehrfach relativ feste, von der Höhe des Grundtones unabhängige Stärkemaxima bemerkbar.

Die festen Maxima freilich, welche HERRMANN-GOLDAP angibt, stimmen mit den hier gefundenen (s. u.) zumeist wenig überein. Aber seine Arbeit ist bereits von KOEHLER einer scharfen Kritik unterzogen worden und läßt in der Tat trotz guter Bemerkungen im einzelnen an Beweiskraft vieles zu wünschen, wengleich man zugeben muß, daß die meisten Tabellen, wenn die Werte richtig bestimmt sind, mit einem festen Intensitätsverhältnis, wie es HELMHOLTZ verlangt, unvereinbar sind. Für das Waldhorn scheint geradezu ein festes Maximum bei  $es^1-es^2$  und für die Oboe ein solches bei  $g^3-c^4$  aus seinen Tabellen hervorzugehen. Bei der Klarinette hingegen sind auf den Grundtönen  $a^1$  bis  $e^2$  regelmäßig Grundton und Duodezime verstärkt, ganz im Sinne von HELMHOLTZ (der  $g^1$ -Klang fällt aus der Reihe, er hat in H.-G.s Tabellen überhaupt keine Duodezime). Bei den übrigen Instrumenten ist das einzige, was klar aus den Tabellen hervorgeht, die Verschiebung des Amplitudenmaximums gegen den Grundton hin mit wachsender Höhe des Klanges; eine bedeutsame Tatsache, die aber mit den festen Maximis an sich nichts zu tun hat und von H.-G. nicht erkannt wurde. Für die schwächeren Klänge der Flöte und Violine benützte er das Phonoskop von WEISS, aber offenbar in noch unvollkommener Methodik, da er bei der Violine selbst in der tiefsten Oktave  $g-g^1$  fast überall nur 2-3 Teiltöne fand, wobei die Aufnahmen auch noch als „sehr gut“ bezeichnet werden. Aus solchen Aufnahmen läßt sich überhaupt nichts schließen.

d) Diese Stärkemaxima sind aber hier nicht das hauptsächlich Ausschlaggebende, sondern noch wesentlicher ist eine von der Höhe des jeweiligen Grundtones abhängige, nur relativ zu diesem gleichbleibende Gruppe von Teiltönen. Diese möge hier als Hauptformant bezeichnet werden, die festen Stärkemaxima dagegen, wo sie vorhanden sind, als Nebenformanten. Zur vollen Charakteristik des Klanges sind aber auch sie erforderlich, ebenso wie die Nebenformanten bei den Vokalen.

Die Hauptformanten sind also hier beweglich, die Nebenformanten fest. Für die ersten war HELMHOLTZ, für die letzten waren seine Gegner im Rechte.

Die Hauptformanten erstrecken sich in der Regel vom 4. bis zum 6. Teilton, d. h. also: die spezifische Klangfarbe (im engeren Wortsinn) ist mit 6 Teiltönen in den richtigen Stärkeverhältnissen für Urteilsfähige erkennbar. Doch ist hier die obere Formantengrenze weniger genau anzugeben als die untere.

e) Die Klangfarbe besitzt jedoch damit in der Regel noch nicht ihre volle Ausprägung. Dazu müssen zumeist noch höhere Tongruppen hinzukommen, und unter diesen findet sich nach Umständen auch wieder ein fester, an absolute Höhen gebundener Nebenformant: so bei der Klarinette und dem Fagott der „Näselformant“, der vom Ende der 3-gestrichenen Oktave beginnt und sein Zentrum in  $d^4$  hat.

An diesen Aufstellungen ist besonders wichtig, daß der Hauptformant nicht immer zugleich den schlechthin stärksten Teilton einschließen muß. Er muß nur diejenigen Teiltöne einschließen, die (in ihrem Zusammenwirken mit den tieferen Teiltönen, besonders dem Grundton) für den Klangcharakter des Instrumentes den Ausschlag geben, also sozusagen die stärkste färbende Kraft unter diesen Umständen haben. Ein relatives Maximum wird freilich die Töne des Hauptformanten vor ihren Nachbarn auszeichnen, aber es braucht nicht das absolute in der ganzen Klangmasse zu enthalten. Wir haben schon bei den Vokalen vermutet, daß die ausschlaggebenden Töne nicht immer und notwendig zugleich die energetisch stärksten seien (o. S. 214). Noch deutlicher fand sich dies bei den Konsonanten, wo die Geräuschmaxima wahrscheinlich in der Gegend der direkt wahrnehmbaren Tonhöhen liegen, die Formanten aber sich höher hinauf erstrecken (o. S. 166).

Das Gesagte werde nun an konkreten, nach unseren 3 Methoden geordneten Beispielen erläutert.

### 3. Resonanzversuche.

Hier ist, wie bei den Vokalen, zu beachten, daß die Angaben sich nur auf Klänge in unmittelbarer Nähe und großer Stärke beziehen, da die Resonanzgabeln dicht an die Klangöffnung des Instrumentes gehalten wurden, um alle in dem Klang enthaltenen Teiltöne festzustellen. Auch die technischen Maßnahmen waren die nämlichen. Eine Schwierigkeit lag darin, daß die genaue Abstimmung des Grundtones zu den verfügbaren Gabeln als Teiltönen bei den Blasinstrumenten nicht so leicht wie bei der menschlichen Stimme zu erzielen war. Wir können daher für die gewonnenen Stärkezahlen im einzelnen nicht immer dieselbe Garantie übernehmen wie dort und beschränken uns auf die Mitteilung der gesichert erscheinenden Befunde.

a) Die Tenorposaune hat sowohl auf  $c$  als auf  $c^1$  eine ununterbrochene Reihe harmonischer Teiltöne bis  $c^5$ , darunter am stärksten und überhaupt von außerordentlicher Stärke (bis zu 5) beide Male die Töne von  $c^1$  bis  $c^3$ . Auf  $c$  übertrafen  $c^2$  und  $c^2$ , auf  $c^1$   $c^2$  und  $g^2$  die übrigen noch etwas an Stärke. Für den Grundton  $c^1$  lag auch noch ein 2. Maximum bei  $d^4$ .

Hier wurde auch der Versuch gemacht, alle Gabeln auf einmal zur Resonanz zu bringen (vgl. o. S. 32): wirklich erklangen beim Grundton  $c$  alle Gabeln von  $c^1$  bis  $g^3$  in einem vollen Akkord. Aber bei der räumlichen Trennung der Gabeln war der Eindruck nicht einheitlich genug, um den Posaunenklang synthetisch naturgetreu wiederzugeben.

b) Die B-Trompete hat auf  $c^1$  und  $c^2$  gleichfalls eine ununterbrochene Teiltonreihe bis  $c^5$ , und selbst die höchsten sind noch bedeutend übermerklich (etwa  $1\frac{1}{4}$ ). Die Hauptmaxima liegen für  $c^1$  bei  $c^2$  und  $g^2$ , für  $c^2$  bei  $c^2$  und  $c^3$ . Im letzteren Falle folgt aber noch ein Maximum bei  $c^4$  und  $e^4$ . Auf dem selten (z. B. im Parsifal) beanspruchten  $c^3$  hat  $c^5$  immer noch die Stärke  $1\frac{1}{2}$ .

c) Die B-Klarinette hat auf  $c^1$  und  $c^2$  wiederum eine fast ununterbrochene Reihe von Teiltönen; wir sagen hier „fast“, weil auf  $c^1$  der 2. Teilton  $c^2$  in Übereinstimmung mit der physikalischen Theorie ausfällt. Auf  $c^2$  ist aber der 2. Teilton da ( $c^2 = 4$ ,  $c^3 = 2$ ,  $g^3 = 3$ ), und die höheren geradzahigen Teiltöne sind in beiden Fällen nur wenig schwächer als die sie umgebenden ungeradzahigen.

Dieser Befund steht im Widerspruch mit der gewöhnlichen Meinung, daß der Klarinette schlechtweg die geradzahigen Teiltöne fehlen, weil in zylindrischen Röhren solche nicht zustande kommen können. Meines Wissens hat nur MEISSNER in seiner hinterlassenen Abhandlung (S. 595) eine entgegengesetzte Angabe gemacht. Aber die Klarinetten sind eben, wie C. SACHS bei Erwähnung unserer Ergebnisse bemerkt (S. 332), nicht rein zylindrisch. SACHS fügt bei: „Je mehr man überbläst, desto stärker werden die geradzahigen, daher der Glanz höherer Lagen.“ Bei einem  $c^3$  resonierte in der Tat  $c^4$  sehr stark, bei einem  $f^3$  wurde  $f^4$  auch mit Interferenz isoliert und erwies sich so stark wie der Grundton selbst. Im übrigen darf man nicht vergessen, daß wenigstens die ersten geradzahigen Teiltöne, der 2. und 4., auch als subjektive Differenztöne der ungeradzahigen hervorgebracht oder, wenn sie schon objektiv vorhanden sind, verstärkt werden.

Der Grundton ist hier immer stark, auch die Duodezime hat höhere Stärkegrade. Sehr stark ferner die 4-gestr. Oktave, namentlich  $g^4$ , welches auf allen 3 Grundtönen ( $c^1$ ,  $c^2$ ,  $c^3$ ) ein relatives

Maximum darstellt. Auch mit bloßem Ohre sind solche Obertöne leicht herauszuhören. Bei einem  $as^1$  traten mir sogleich  $as^4$  und  $c^5$  entgegen, die auch entsprechende Gabeln kräftig zum Erklingen brachten. In MILLERS Klangspektrum der Klarinette treten hohe Teiltöne gleichfalls stark hervor<sup>1)</sup>, ebenso bei FLETCHER<sup>2)</sup>  $c^4$ ,  $d^4$ ,  $e^4$  (neben dem Grundton  $c^1$  und der Duodezime  $g^2$ ).

d) Das Fagott (mit dem Kontrafagott in 8 Beobachtungsreihen untersucht) hat auf den Grundtönen  $C$ ,  $c$ ,  $c^1$  gleichfalls viele Teiltöne. Aber die tieferen scheinen objektiv beträchtlich schwächer zu sein als bei der Klarinette. Mit den Gabeln ließen sie sich erst von  $c^1$  an feststellen, weil tiefere Gabeln überhaupt nicht resonieren. Aber dieser Ton selbst gab keine oder nur eine äußerst schwache Reaktion. Auch  $g^1$  (bei  $C$  und  $c$ ) war noch schwach. Die tiefen Teiltöne scheinen also hier, wie in manchen Fällen der Vokale, wesentlich als Differenztöne hinzuzukommen. Dagegen liegt ein bedeutendes Maximum für alle Grundtöne bei  $c^2-g^2$  (Stärken bis 4). Dann folgt wieder ein jäher Abfall. Ein kleines Nebenmaximum liegt vielleicht noch bei  $d^3-e^3$ . Aber die Teiltöne sind bis  $as^4$  noch festzustellen.

Auf  $C$  wurde auch das Kontrafagott untersucht. Es gab wesentlich andere Werte. So waren  $c^2$ ,  $g^2$ ,  $d^3$ ,  $g^3 = 0$ . Der Klang ist auch für das direkte Hören bedeutend verschieden, voluminöser, runder, aber nicht so metallreich wie der des Fagotts. Mit dem bloßen Ohr waren noch  $a^3$ ,  $b^3$ ,  $d^4$  heraushörbar (man kann überhaupt bei den Instrumenten in der Nähe die Obertöne oft leicht mit dem bloßen Ohr heraushören).

e) Von der Flöte wurden mehrere Typen untersucht, weil hier auch die Unterschiede des Materials und der Bauart von Interesse sind. Die Vergleichung einer guten, modernen Ebenholzflöte mit der goldenen des Virtuosen Prof. PRILL, wobei auch die erste von einem ausgezeichneten Bläser, Prof. SCHÜNEMANN, angegeben wurde, ergab folgende Stärkereihen (die Stärkezziffern hier wieder mit 4 multipliziert):

	$c^1$	$c^2$	$g^2$	$c^3$	$e^3$	$g^3$	$\bar{b}^3$	$c^4$	$d^4$
Goldene Fl. . .	10	8	6	4	1	1	1	$\frac{1}{2}$	$\frac{1}{2}$
Holzflöte . . .	8	8	6	3	1	$\frac{1}{2}$	0	$\frac{1}{2}$	0

<sup>1)</sup> S. 201. Leider ist die Grundtonhöhe nicht angegeben. Wahrscheinlich war sie  $c^1$ ; dann fallen die Maxima mit denen FLETCHERS zusammen. Auffallend ist aber die geringe Stärke des 1. und 3. Teiltöns.

<sup>2)</sup> FLETCHER 3 S. 428. Der Grundton ist hier als  $c$  bezeichnet, muß aber nach den auf der Abszissenachse angegebenen Schwingungszahlen (Doppelschwingungen nach S. 429) in unserer Bezeichnung  $c^1$  gewesen sein.

Ein sicheres Urteil über die alte Frage nach dem Einfluß des Materials würde ich darauf allein nicht gründen, da mehr Exemplare verglichen werden müßten. Doch hat auch MILLER (S. 183) bei einer goldenen Flöte mehr und stärkere Teiltöne gefunden. Und schon HELMHOLTZ beobachtete (S. 157ff.) bei Holzpfeifen eine weichere Klangfarbe und weniger Obertöne als bei Metallpfeifen.

Außer diesen beiden modernen Flöten wurden aber auch 2 alte untersucht, eine konische aus Buxbaumholz, die etwa 100 Jahre zählen mag und eine noch ältere aus Ebenholz mit einer einzigen Klappe, wohl aus der Zeit Friedrich des Großen. Analyse wie Synthese bestätigten den direkten Eindruck, daß der moderne Flötenklang metallischer, reicher an Obertönen ist. Er ist dadurch für Virtuosenzwecke brauchbarer geworden, hat aber sein eigentümliches Ethos einigermaßen verloren.

Nach ERFF (Grundfragen der neueren Instrumentation, Zeitschr. Melos Jg. 4, S. 528. 1925) unterscheiden sich heute die Farben von Flöte, Oboe, Klarinette in der Gegend der 2-gestr. Oktave weniger voneinander als zur Zeit BEETHOVENS. Auch unter den Blechinstrumenten finde eine Annäherung zwischen Trompete, Horn, Posaune statt, die den homogenen Zusammenklang und den Zusammenschluß der Gruppen fördere, aber andererseits Farbverlust bedeute. Sie erfolge durch Angleichung der Mensuren. Solche säkulare Wandlungen verdienen großes Interesse und bilden eine gewisse Analogie zu den Erscheinungen des Lautwandels in der Sprachgeschichte.

Bei der Pickelflöte waren auf  $c^3$  die ersten 3 Teiltöne  $c^3$ ,  $c^4$ ,  $g^4$  kräftig, etwa  $= 2$ , höhere nicht zu finden.

f) Endlich wurden mit der Resonanzmethode Waldhorn und Viola geprüft, doch mußten die Resonanzversuche zu früh abgebrochen werden. Es sei nur erwähnt, daß sich beim Horn die unteren Teiltöne besonders stark gegenüber den folgenden fanden und daß bei Grundtönen von  $c-c^2$  ein festes Maximum in der Gegend von  $c^2$  zu liegen schien. Auch bei der Viola waren auf den tieferen Saiten sowohl der Grundton als die nächsten Teiltöne ziemlich stark.

Für das Horn vgl. auch HERRMANN-GOLDAP oben S. 377. In KOEHLERS Trommelfellkurven, deren Einzelheiten bei den Instrumenten schön und deutlich hervortreten, liefert das Horn bei den Grundtönen von  $c^1-g^1$  mit der FOURIER-Analyse immer ein Maximum auf dem 2. Teilton; bei  $as^1$  aber sinkt es auf den 1. (S. 36). Dies könnte in gleichem Sinne gedeutet werden: daß nämlich das Maximum mit dem Grundton nur innerhalb einer bestimmten Zone in die Höhe geht, und daß diese Verstärkungssphäre, das „feste“ Maximum, von etwa  $as^1-g^2$  anzusetzen wäre (auch die Vokalformanten sind ja nicht absolut fest, sondern als harmonische Teiltöne vom Grundton mitbedingt). Aber gäbe es keine weiteren Tatsachen, die für feste Maxima sprechen, so würde man dieses Verhalten auch auf die

allgemeine Verschiebung des Intensitätsmaximums gegen den Grundton mit dessen Emporsteigen deuten können. MILLER gibt bei den Diagrammen für das Horn (S. 204) leider die benutzten Tonhöhen nicht an, nur die Stärkeunterschiede; aber bei den höheren Stärkegraden liegt gleichfalls beide Male das Maximum auf dem 2. Teilton. Wahrscheinlich lagen also die Grundtöne auch in der Gegend von  $c^1$ .

#### 4. Interferenzversuche.

Das Technische betreffend, macht sich bei den Instrumentalklängen die Gefahr der Veränderung durch die Hauptleitung mehr als bei den Vokalen geltend, weil sie vielfach bedeutend reicher an Obertönen sind und die Leitung zugleich bei besonders starken Klängen verlängert werden muß, wobei nun die tieferen Teiltöne leicht mehr als die hohen geschwächt werden und der Klang darum schärfer, schreiender als beim freien Hören ans Ohr gelangt. Es muß daher in solchen Fällen durch die Stellung des Bläasers vor der Leitung, passende Trichter, sehr weite Röhren u. a. für möglichste Unversehrtheit des Klanges gesorgt werden. Jedenfalls muß er so herüberkommen, daß er als Klang dieses bestimmten Instrumentes zweifelsfrei wieder erkannt wird. Ist dies auf keine Weise zu erreichen, dann muß auf diesen Forschungsweg eben verzichtet werden.

Die folgende Tabelle stellt die Formanten (nach der obigen Definition) für verschiedene Instrumente und Grundtöne zusammen. In der ersten Rubrik stehen die bei den Resonanzversuchen gefundenen festen Maxima.

Formanten und Maxima bei Instrumentalklängen.

Instrument	Grundton	Annähernd feste Maxima	Beweglicher Hauptformant		
Tenorposaune	$c$ $c^1$	} $c^1 - e^3$	$c^2 - c^3$ $c^3 - c^4$		
B-Trompete	$c^1$ $c^2$		} $c^2 - c^3$	$g^2 - g^3$ $g^3 - c^4$	
A-Klarinette	$cis$ $cis^1$ $fis^1$ $fis^2$	} $cis^4 - gis^4$		$cis^2 - cis^3$ $gis^2(cis^3) - gis^3(cis^4)$ $fis^3 - cis^4$ $fis^4 - cis^5$	
B-Klarinette	$c^1$ $c^2$ $c^3$		} $d^4 - b^4$	$e^3 - b^3$ $c^4 - d^5$ $c^5$	
Kontrafagott	$C$ $G$			} $c^2 - g^2$	$g^1 - c^2$ $d^2 - g^2$
Fagott	$c$ $c^1$				$c^2 - g^2$ $c^3 - g^4$

Aus dieser Übersicht dürfte hervorgehen, daß der Formant im oben erwähnten Sinne — als derjenige Klangteil, durch dessen

Hinzukommen zu den tieferen der Klang als der des betreffenden Instrumentes kenntlich wird — mit der Tonhöhe, und zwar im allgemeinen im gleichen Verhältnis wie diese, emporsteigt, also nicht an die absolute Höhe, sondern an die Ordnungszahlen bestimmter Teiltöne gebunden ist und seine Lage nur relativ zum Grundton beibehält. Er fällt aber nur ausnahmsweise einmal (beim Fagott auf  $c$ ) zusammen mit dem festen Stärkemaximum. Die beiden Bezirke können sich auch ergänzen oder überschneiden, da eben der eine fest, der andere beweglich ist.

Die Interferenzversuche, bei denen ich mich vielfach der Mitarbeit des hervorragenden Instrumentenkenners Prof. SACHS erfreute, lieferten aber auch, abgesehen von den Formanten, noch andere Ausbeute: die Veränderungen der Klänge mit jedem neu hinzutretenden Teilton, das allmähliche Heranwachsen und die Vervollkommnung der Klangeigentümlichkeiten der Instrumente bis zu den höchsten Teiltönen, waren außerordentlich instruktiv. Einiges Besondere sei hier angeführt:

a) Der durch Ausschaltung aller Obertöne isolierte Grundton ist bei der Klarinette von der Stärke  $2-2\frac{1}{2}$  (8—10), beim Fagott aber, wenigstens in den tiefen Lagen, äußerst schwach. Erst wenn weitere Teiltöne hinzutreten und die Höhe von etwa  $c^1$  erreichen, wird auch der Grundton durch Differenztonbildung kräftig. Bei der Tenorposaune ist es ebenso. In beiden Fällen wird durch diese hinzukommenden Teiltöne, zumal den 2. und 3., der Gesamtklang außerordentlich verändert, mächtiger, dröhnender. Der Grundton für sich allein ist bei exakter Isolierung überall von demselben, in der Tiefe höchst weichen Charakter und bei gleicher Höhe in allen Instrumenten qualitativ durchaus ununterscheidbar.

b) Bei der Klarinette wird, namentlich auf tieferen Grundtönen, das eigentümlich Hohle des Klanges bemerkbar, sobald der 3. Teilton auftritt. Daß das Fehlen oder die Schwäche des 2. Teiltones sich auch in der Erscheinung, und zwar in dieser Weise, geltend machen müsse, läßt sich, wie auch hier betont werden soll, keineswegs a priori deduzieren, sondern ist als eine bemerkenswerte psycho-physiologische Tatsache zu verzeichnen.

Die Klarinette zeigt beim weiteren Aufbau des Klanges besonders starke Wandlungen. Zum Eindruck des Hohlen kommt mit den Teiltönen der 4-gestr. Oktave der des Näseldnen (der sich aber in der höheren Lage des Instrumentes, etwa von  $g^2$  an, verliert), endlich mit Hinzutritt der 5-gestr. Oktave der des leicht Schmetternden, der allerdings zum Teil ein Kunstprodukt der Interferenzleitung ist. Unter gewöhnlichen Umständen tritt diese Eigenschaft an tiefen Klarinettenklängen nicht hervor, und auch



noch in der 2-gestr. Oktave ist der Klang namentlich beim piano entzückend weich und sanft. Aber in höheren Lagen und bei stärkerer Tongebung tritt die Schärfe doch sehr in die Erscheinung. Auch sind die einzelnen Klarinettenarten darin noch verschieden, am schärfsten nach SACHS (S. 335) die nur ganz ausnahmsweise verwendete C-Klarinette.

Sehr lehrreich ist folgender Versuch. Läßt man mit einer A-Klarinette ihren tiefsten Ton *cis* blasen, stellt aber die Interferenz so ein, als gälte es, den gesamten Klang auf dem Grundton *cis*<sup>1</sup> (einschließlich dieses Tones selbst) auszulöschen, so kommen nur der Grundton *cis* und seine ungeradzahigen Teiltöne zum Vorschein. Man hat dann tatsächlich diejenige Klangstruktur, die theoretisch dem Klarinettenklang eigen sein soll. Ergebnis: der Klang hat noch unzweifelhaften Klarinettencharakter, aber er ist weicher und dunkler als vorher. Die ungeraden Teiltöne als solche können also nicht an der Schärfe oder dem Näseln schuld sein.

Denselben Versuch machte ich mit einer Zunge des Tonmessers von 200 Schw. (*gis*), die sehr zahlreiche Obertöne enthält. Wurden alle geraden Teiltöne ausgeschlossen, so entstand aus dem scharfen metallischen Zungenklang ein eigentümlich angenehmer weicher Klang, der keinem Instrument besonders ähnlich war, allenfalls zwischen Waldhorn und Englisch-Horn in der Mitte stand. Die Duodezime war gut herauszuhören.

Einen verwandten Versuch hat schon RITZ auf der Geige gemacht. Streicht oder zupft man eine Saite in der Mitte, so gibt sie einen Klang, in dem die geraden Teiltöne fehlen. RITZ nennt ihn „eigentümlich näselnd“ (S. 72). Dies kann ich aber nicht bestätigen, wüßte auch nicht, woher das Näseln kommen sollte; er ist nur hohl und weich.

Vergleicht man die Fälle: 1, 2, 4 . . . und 1, 3, 5 . . ., die beide in keinem Instrument verwirklicht sind, in bezug auf die Bildung von Differenztönen, so würde der Unterschied sein, daß im 1. Falle nur Glieder der Reihe selbst dadurch verstärkt werden (denn die D.T. 4-1, 6-1 usf. sind so gut wie unhörbar; vgl. m. Abh. 7, S. 126ff.), im 2. Falle dagegen die objektiv nicht vorhandenen geraden Glieder, wenigstens 2 und 4, in merklicher Stärke subjektiv hinzutreten müssen.

c) Das Fagott wurde, da es die auffallendsten Entwicklungsstufen und die komplizierteste Klangstruktur aufwies, auf den 3 Grundtönen *C*, *c*, *c*<sup>1</sup> mit der Interferenzmethode untersucht. Man kann hier beim Aufbau 4 Zonen unterscheiden. Ist man beim Aufbau von *C* oder *c* mit dem Einschieben der If.-Röhren bis zum oberen Ende der Formanten gelangt, so ist das Instrument erkennbar, aber es hat noch etwas Posaunenartiges. Dann tritt

mit den Teiltönen der 3-gestr. Oktave, besonders ihrer oberen Hälfte, ein Schwirren hinzu, das jedenfalls durch die Schwebungen der nahe beisammenliegenden Töne  $d^3$  bis  $g^3$  bedingt ist. Hierauf folgt mit den Teiltönen von  $as^3$  bis  $c^5$ , besonders denen aus der unteren Hälfte der 4-gestr. Oktave, das Näseln. Endlich von  $ci^5$  bis mindestens  $g^5$  das eigentümlich Pelzige dieses Instrumentes, das durch die Menge der ganz hohen feinen Teiltöne hinzugebracht wird.

Diese 4 Entwicklungsstadien kann man auch am vollen Klang noch erkennen; aber es verhält sich damit wie mit dem Herausheören der einzelnen Teiltöne: man hört sie in allen Graden der Deutlichkeit, je nachdem man gerade eingestellt ist. Ist man aber ganz auf eine Zone konzentriert, so fällt natürlich der Eindruck der Klangfarbe selbst hinweg. Ist man weniger ausschließlich darauf konzentriert, so bleibt dieser neben dem Teileindruck der Zone noch einigermaßen bestehen.

Beim Aufbau des Kontrafagott-*C* durch Hineinschieben der *Hf.*-Röhren war der Grundton selbst für sich allein nicht zu hören, erschien aber auch nicht als Differenzton mit dem 2. Teilton *c*, der in der Stärke 1 auftrat. Erst als auch der 3. *g* (mit der Stärke  $1\frac{1}{2}$ ) hinzukam, erschien zugleich ein schönes markiges *C*. Der Gesamtklang hatte hier schon das Imposante dieser tiefen Baßlage. (Vgl. das gesungene *C* oben S. 56 in demselben Aufbaustadium.) Beim Hinzukommen des 4. Teiltones  $c^1$ , der mit Isolierversuchen =  $1\frac{1}{2}$  gefunden wurde, wurde aber *C* selbst leiser, und der Gesamtklang schien jetzt vorübergehend die Höhe von  $c^1$  zu haben, der Grundton nur wie ein akustischer Hintergrund; aber dies änderte sich wieder mit dem Hinzutreten höherer und stärkerer Teiltöne: die Gesamtmasse schien dann doch wieder auf *C* lokalisiert.

Auch Bestandteile aus der höchsten Region (dem Näselformanten) wurden beim *C* des Kontrafagotts dadurch hörbar gemacht, daß auf alle Teiltöne vom Grundton bis zu  $c^4$  eingestellt wurde. Denn dabei werden die Multipla nicht völlig mit ausgeschlossen (oben S. 48). Man hörte ein feines Schwirren, aus welchem durch weitere Einstellungen  $d^4$  als besonders maßgebender Ton herauspräpariert wurde. Durch Einstellung 3,5 wurde auch er vernichtet. Weitere Versuche lehrten, daß außer  $d^4$  auch  $c^3$  und  $fis^3$  stark in dem Klange vertreten waren, dagegen die zwischen ihnen liegenden Töne nicht oder nur schwächer. Unter günstigen Umständen hörte man einen dissonanten Akkord aus diesen 3 Tönen (Teil eines umgelagerten Dominant-Septimenakkordes). Man kann — wie bei den näselsnden Vokalen, nur noch deutlicher — beobachten, daß  $c^3$  und  $fis^3$  einen leichten Ä-Charakter,  $d^4$  aber

das Näselerde in den Klang bringen, das durch die nächstfolgenden Teiltöne noch verstärkt wird.

d) Beim Horn, das auf den 4 Höhen  $C, c, c^1, c^2$  mit Interferenz untersucht wurde, wies der tiefe  $C$ -Klang beim Aufbau einen langen Entwicklungsgang auf. Aber er trägt nicht eigentlich die populäre, typische Hornfarbe (wird auch nur ganz selten verlangt), sondern ist fast schmetternder und rauher als der der Posaune. Der Grundton für sich allein ist hier ein schönes leises Brummen, der 2. Teilton merkwürdigerweise bei Isolierversuchen überhaupt nicht zu hören, der 3. schwach; aber dessen Hinzukommen macht den Grundton rauher. Der 4., der isoliert mit der Stärke 2 gehört wird, bringt eine bedeutende Veränderung: von da ab wird der Klang immer mächtiger und zugleich einheitlicher, von  $c^2$  an metallisch; mit  $as^2$  ist er fertig.

Zur Feststellung der If.-Formanten auf den verschiedenen Grundtönen bin ich beim Horn nicht gekommen. Doch wurden Isolierversuche, um die Stärken der Teiltöne zu finden, auch hier gemacht. Bei  $C$  erwies sich  $c^1$  besonders stark, bei  $c, c^1, c^2$  dagegen gemeinschaftlich  $c^2$ . Dies letztere steht in Einklang mit den obigen Resonanzbefunden. Wenn das Maximum beim Grundton  $C$  auf  $c^1$  herabgeht, so wird dies mit dem veränderten Klangcharakter zusammenhängen.

##### 5. Synthesen.

Zur Synthese wurden, wie bei den Vokalen, nur Teiltöne herangezogen, die das Charakteristische des Klanges, wie er in mäßiger Entfernung erscheint, erkennen lassen, wenn auch in unmittelbarer Nähe nach dem Zeugnis der Resonanzversuche noch mehr Teiltöne vorhanden sind. Die Vergleichung mit dem natürlichen Klang erfolgte hier nur teilweise durch seine Herüberleitung aus dem Zimmer III (s. oben S. 44); meistens wurde er im Korridor bei geöffneter Türe angegeben. Die Herren Prof. SACHS und SCHÜNEMANN wirkten bei diesen Versuchen vielfach mit, und sowohl die Urteile dieser beiden Kenner als das der Bläser selbst bestätigte die Naturtreue des künstlichen Klanges. Vielfach haben wir uns auch durch Stichversuche (Herausnehmen eines Tones aus dem schon fertigen künstlichen Klange) über die mehr oder minder wesentliche Bedeutung und die Wirkungsart einzelner Teiltöne vergewissert. Dennoch könnte ich für die Genauigkeit der Synthesen hier nicht so zuversichtlich wie bei den Vokalen eintreten, da nicht Gelegenheit war, sie so oft zu wiederholen und nachzuprüfen. Zunächst wurde, mehr der Übung halber, der Klang der Zunge  $c^1$  unseres Tonmessers nachgebildet,

was mit 10 Teiltönen gut gelang; weitere bis zum 24. konnten schwach beigelegt werden, erwiesen sich aber als entbehrlich. Die folgenden Tabellen geben nun die Teiltöne der dargestellten Klänge auf den 4 Grundtönen  $c$ ,  $c^1$ ,  $c^2$ ,  $c^3$  bei mäßig starker Tongebung. Die Klangstärke bedingt bei Instrumenten, besonders Blasinstrumenten, außerordentliche Unterschiede in der Klangfarbe, also der Stärkeverteilung unter den Teiltönen. Jede Tabelle dieser Art kann daher nur für eine bestimmte Klangstärke Gültigkeit beanspruchen. Dies geht auch aus MILLERS Aufnahmen hervor und ist von ihm stets hervorgehoben. Beim Horn ist hier nur geringe Stärke (Piano) vorausgesetzt, weil nur dann der spezifisch-mollige Hornklang erscheint. Für das Forte wären weit mehr und nach der Höhe relativ stärkere Teiltöne erforderlich.

Die Stärkezeiffern sind wieder mit 4 multipliziert. Sie geben hier aber nur die Stärkeverhältnisse innerhalb jeder einzelnen Kolumne an. Es kann dadurch z. B. nicht die Stärke des 6. Teiltones der Trompete mit der des 6. bei der Klarinette verglichen werden, sondern nur die der Teiltöne innerhalb des nämlichen Instruments. Bei den synthetischen Vokaltabellen ist dies anders, weil die nachzubildenden Vokale unter gleichen Umständen, besonders mit möglichst gleicher Expirationsstärke und aus gleicher Entfernung angegeben und die künstlichen an der gleichen Stelle (bei  $B_2$ , oben S. 44) abgehört wurden. Hier aber mußte man sich darauf beschränken, bei den nachzubildenden Klängen nur eben eine mittlere Tonstärke vorzuschreiben und den künstlichen Klang, wenn er zu schwach aus der Schlauchöffnung kam, am Trichter T abzuhören, wobei andere absolute Stärken als bei  $B_2$  herauskamen. Unbedingt nötig war aber auch hier eine kurze Dauer und ein Abschneiden des Einsatzes bei dem natürlichen Klange. Wurde dieser durch die Leitung  $S_2$  erheblich verändert, so mußte er ohne Röhrenleitung in einiger Entfernung angegeben und das Abschneiden des Einsatzes dadurch bewirkt werden, daß man die beiden Ohren mit den Fingern verschloß und nur auf einen Augenblick während der Klangdauer öffnete. Eine gleich kurze Dauer des künstlichen Klanges wurde dann durch Zudrücken und Öffnen eines an den Trichter T gesetzten kleinen Höserschlauches erreicht.

Der Klarinettenklang erscheint auch hier besonders kompliziert. Der Ton  $c$  ist auf der gewöhnlichen (Sopran-) Klarinette nicht vorhanden; die am tiefsten hinabreichende  $A$ -Klarinette geht nur bis  $cis$ . Aber der  $c$ -Klang konnte so dargestellt werden, daß er sich an das natürliche  $cis$  als farben-gleiche Fortsetzung anschloß und wurde in dieser Form von den HH. SACHS und SCHÜNEMANN „sehr gut“ gefunden. Da aber, worauf mich kürzlich Hr. Dr. SCHUMANN hinwies, der Instrumentenmacher MORITZ (Berlin) durch Einfügung einer verlängerten „Birne“ unterhalb des Mundstückes die  $A$ -Klarinette versuchsweise bis  $c$  geführt hat, wobei dieser Ton zwar etwas schwierig, aber mit tadelloser Klarinettenfarbe herauskommt, so benutzten wir diesen Klang als Vorbild zu einer neuen Synthese, deren Stärkezeiffern in der obigen Tabelle stehen. Sie wurden von Dr. SCHUMANN und mir unabhängig festgestellt, mit nur geringen Differenzen. Oberhalb  $g^4$  enthielt die Mischung noch die Töne  $a^4$ ,  $c^5$ ,  $e^5$ , die von Dr. SCHUMANN mit den minimalen Stärken  $\frac{1}{4}$ ,  $\frac{1}{2}$ ,  $\frac{1}{4}$  (1, 2, 1) isoliert aus der Leitung gehört wurden, aber im

## Synthesen von Instrumentalklängen.

Grundton  $c = 128$  Schw.

$g^4$	3			
$\overline{fis^4}$	1			
$e^4$	5			
$d^4$	6			
$c^4$	6			
$h^3$	4			
$\overline{b^3}$	5			
$\overline{a^3}$	6			6
$g^3$	8			4
$\overline{fis^3}$	9			4
$e^3$	8	6	2	8
$d^3$	6	6	4	6
$c^3$	5	12	6	10
$\overline{b^2}$	4	16	0	16
$g^2$	5	20	10	20
$e^2$	6	16	12	18
$c^2$	1	13	12	14
$g^1$	10	10	10	10
$c^1$	1	8	12	10
$c$	7	0	10	2

Klarinette  
Tenorposaune  
Horn  
Viola

Grundton  $c^1 = 256$  Schw.

$g^5$	2							
$e^5$	4							
$c^5$	4							
$\overline{b^4}$	4	4						
$\overline{a^4}$	4	6						
$g^4$	5	8						
$\overline{fis^4}$	4	6						
$e^4$	6	6						
$d^4$	6	12	4					
$c^4$	6	8	5				2	
$\overline{b^3}$	6	12	6				2	
$g^3$	8	11	8			4	4	
$e^3$	8	10	6			6	4	
$c^3$	10	12	12	6		10	6	5
$g^2$	12	20	20	10	12	10	4	
$e^2$	10	6	12	16	12	8	6	
$c^2$	8	14	10	12	10	6	5	

Trompete  
Klarinette  
Tenorposaune  
Horn  
Viola  
Neue Flöte  
Alte Flöte

Grundton  $c^2 = 512$  Schw.

$g^5$	4					
$e^5$	6					
$c^5$	8					
$\overline{b^4}$	8	5				
$g^4$	10	6		6	6	
$e^4$	10	6		10	6	
$c^4$	12	10		8	8	
$g^3$	12	16	5	8	12	8
$c^3$	20	8	10	12	16	12
$c^2$	18	16	20	8	16	16

Trompete  
Klarinette  
Horn  
Viola  
Neue Flöte  
Alte Flöte

Grundton  $c^3 = 1024$  Schw.

$g^5$			5			
$e^5$	4	4				
$c^5$	8	6	8			
$g^4$	8	9	8	8		
$c^4$	16	14	15	10	10	
$c^3$	14	12	12	16	16	

Trompete  
Klarinette  
Viola  
Neue Flöte  
Alte Flöte

Gesamtklänge wahrscheinlich einflußlos, von den stärkeren darunterliegenden unterdrückt waren. Der künstliche Klang erschien uns allen im wesentlichen durchaus charakteristisch; nur eine gewisse leise Härte fehlte ihm noch, die vom Geräusch der aufschlagenden Zunge herrühren dürfte und sich darum mit den Pfeifen nicht nachbilden läßt. Der 2. Teilton ist hier, ob schon er nach den Resonanzversuchen in den tieferen Lagen objektiv fehlt, doch auch schwach eingefügt, da der Klang so einheitlicher wurde, ohne seine charakteristische Hohlheit einzubüßen. Der Ton kommt ja aber auch bei dem natürlichen Klang als Differenzton im Ohre schwach hinzu.

Zu bemerken ist noch, daß die Töne der 4-gestr. Oktave nicht so stark genommen werden durften, wie man nach den Resonanzversuchen und dem direkten Heraushören (besonders für  $g^4$ ) hätte erwarten müssen, wahrscheinlich darum, weil die tieferen Teiltöne von der Einrichtung nicht mit solcher Stärke geliefert werden, um das Gleichgewicht, d. h. die Einheitlichkeit des Klanges, so starken Obertönen gegenüber aufrechtzuhalten. Analoges gilt auch für die anderen Instrumente.

Gelegentlich habe ich auch eine Oboensynthese auf  $g^1$  ausgeführt, bei der der Grundton schwach, dagegen Obertöne von  $g^2$  bis  $fis^4$  stark waren<sup>1)</sup>.

Bei der Viola war die Naturtreue noch nicht völlig erreicht; es fehlt eben das Streichgeräusch und das eigentümlich Belegte der beiden umspannenen Saiten. Durch Schwebungen ganz hoher, dicht beisammenliegender leiser Töne schien es einigermaßen nachgebildet zu werden.

Die Flötensynthesen stimmen nicht ganz mit den Resonanzergebnissen, in denen auf  $c^1$  der Grundton am stärksten ist, wurden aber von Prof. PRILL als „sehr gut“ qualifiziert; vielleicht ist durch die stärkere Hervorhebung des  $g^2$  die A-Ähnlichkeit des natürlichen Klanges noch mehr akzentuiert, der Klang in gewissem Sinne vervollkommenet.

Beim Horn (piano) wurde auf  $c^1$  der interessante Stichversuch gemacht, den 2. Teilton  $c^2$ , der besonders stark sein muß, nach gelungener Synthese wieder herauszunehmen. Der Klang wurde dadurch vollkommen unkenntlich und näherte sich stark dem der Klarinette.

Man kann mit der synthetischen Einrichtung natürlich auch Ab- und Aufbaureihen analog denen der Interferenzmethode ausführen, indem man, nachdem eine befriedigende Synthese erzielt ist, die Pfeifen sukzessive von oben her bis zur tiefsten abstellt und dann wieder einstellt. Auch diese Versuche sind lehrreich, wurden aber nicht systematisch durchgeführt.

## 6. Allgemeines über das Wesen und die Unterschiede der Klangfarbe.

Die Ergebnisse dieser experimentellen Studien geben nun Veranlassung, auf die in des Verfassers „Tonpsychologie“ besprochenen allgemeineren Fragen in Hinsicht der Klangfarbe zurückzukommen, die Betrachtungen in gewisser Richtung zu ergänzen und mit den jetzt gewonnenen Anschauungen über die Vokalfarben zu verknüpfen.

Man muß, wie schon dort ausgeführt wurde, unterscheiden zwischen dem Charakter eines Instrumentes im Sinne des musi-

<sup>1)</sup> Auch bei MILLER S. 201 hat die Oboe einen schwachen Grundton, aber starke Obertöne in der 3-gestr. Oktave (vorausgesetzt, daß es sich um eine Grundtonhöhe von etwa  $c^1$  handelte).

kalischen Ausdrucks, wie er sich in der Praxis herausgebildet hat, seiner Klangfarbe im weiteren und seiner Klangfarbe im engeren Sinne. Der Charakter oder das Ethos eines Instrumentes für unser gegenwärtiges Musikgefühl, z. B. das Romantische des Horns, das Feierliche der Posaunen, das Festlich-Frohe oder auch Kriegerische der Trompeten, das Dämonische und auch wieder Grotesk-Komische des Fagotts, das Schäferliche der Flöte und Oboe usf. sind in hohem Grade historisch und assoziativ bedingt. Darauf soll hier nicht eingegangen werden. Zur Klangfarbe im weiteren Sinne gehören alle Eigenschaften, die dem Empfindungsmaterial als solchem zukommen, also neben der Struktur aus Teiltönen auch die schon oben erwähnten äußeren, aber darum doch keineswegs unwesentlichen Merkmale, vor allem die Art des Einsatzes und des Aushaltens der Töne (man denke an das Klavier, an das eigentümlich schwerfällige Einsetzen des Horns u. a.), dann die zahlreichen Verschiedenheiten nach der Art des Streichens, Blasens, Zupfens (wobei auch die Stelle einen wesentlichen Unterschied macht) und die begleitenden Blase-, Streich- und Anschlaggeräusche. Die Klangfarbe im engeren Sinne endlich ist objektiv identisch mit der Zusammensetzung aus Teiltönen, subjektiv mit der Gesamtheit der daraus entspringenden Komplexeigenschaften.

Diese Gleichsetzung kann trotz vereinzelter Widersprüche nach unzulänglichen Experimenten nicht mehr bestritten werden. Völlig einfache Töne gleicher Höhe unterscheiden sich, wenn auch etwaige räumliche Unterschiede getilgt sind, nur noch durch ihre Stärke<sup>1)</sup>. Es ist ferner unbestreitbar, daß bei den europäischen Orchesterinstrumenten, abgesehen etwa von den Schlagwerkzeugen und gelegentlich verwendeten Glocken, so gut wie ausschließlich harmonische Teiltöne die Klangfarbe ausmachen. Über die Bedeutung dieses Umstandes gilt das schon bei den Vokalen Gesagte (o. S. 188ff.). Klänge mit nur ungeradzahligem Teiltönen

<sup>1)</sup> Wenn F. VOLBACH (Das moderne Orchester, 2. A. 1921, S. 16) sich für die gegenteilige Behauptung auf die Klänge der „Stimmgabeln und gewisser Orgelpfeifen“ beruft, so sind dies eben keine ganz einfachen Töne. Darum sind auch seine ausgedehnten synthetischen Versuche mit Orgelpfeifen, mit denen er niemals neue Klangfarben erzielt zu haben angibt, unrein und beweisunkräftig. Sie müssen aber noch andere Mängel gehabt haben, denn auch mit Orgelpfeifen kann man, wie MILLER gezeigt hat, bei sorgfältiger Versuchseinrichtung Vokal- und Instrumentalfarben mannigfacher Art erzielen. Wenn die sämtlichen Pfeifen meiner synthetischen Einrichtung im Schallzimmer, wo sie noch nicht von Obertönen gereinigt sind, zusammenklingen, so hört man in einiger Entfernung doch schon einen recht einheitlichen, höchst metallreichen Instrumentalklang.

gibt es dagegen nicht. Es gibt nur Klänge, denen in tiefer Lage der 2. Teilton objektiv fehlt, aber selbst da tritt er als subjektiver Differenzton schwach hinzu.

Die Klangfarbenunterschiede im engeren Sinne treten am ausgeprägtesten in der tiefen Region hervor, weil hier der größte Reichtum an Obertönen vorhanden ist, müssen sich aber gegen die Höhe hin immer mehr verringern. Schon beim  $c^3$  sind sie, wenn man von der Stärke und sonstigen äußeren Merkmalen absieht, recht gering. Ein Blick auf die obigen synthetischen Tabellen zeigt die Ursache. Es ist also ähnlich wie bei den Vokalen von  $c^2$  an. Doch bleiben die Instrumente länger unterscheidbar, da sie nicht in gleichem Maße wie jene auf feste Formanten angewiesen sind. In die 4- und 5-gestr. Oktave reichen nur noch wenige, wie Orgel, Klavier, Violine, Harfe, Piccolo, die in der Praxis durch äußere Merkmale, Ansatz und Haltung des Tones, leicht unterschieden werden.

Der Grundton, dessen Stärke bei den einzelnen Instrumenten sehr verschieden ist, nimmt im allgemeinen an Stärke gegenüber den Obertönen zu, je höher die Tonlage und je leiser die Tongebung genommen wird. Diese beiden Gesetzlichkeiten treten besonders in den graphischen Aufnahmen mit seltener Übereinstimmung hervor, wenn sie auch nicht immer richtig erkannt oder gedeutet werden<sup>1)</sup>. Sie sind mir aber auch bei den eigenen Untersuchungen entgegengetreten, sowohl bei Musikinstrumenten als bei Zungen des Tonmessers. Auch darin verhalten sich die Instrumentalklänge ähnlich den Vokalen, wo sie in den Ergebnissen und Tabellen nach allen 3 Methoden ihre Bestätigung finden. Beide Gesetzlichkeiten sind nur Teilerscheinungen jener Verschiebung der physikalischen Energie nach dem Grundton zu, von der das Verschwinden der Klangfarbenunterschiede bedingt ist.

In manchen Fällen kann man zweifelhaft sein, ob nicht die vorgefundenen „festen Formanten“ bloß infolge dieser Verschiebung der Energie gegen den Grundton hin in die Tabellen kommen. Vgl. die Bemerkungen zu HERRMANN-GOLDAP oben S. 377. Es scheint mir indessen nicht möglich, sie nur auf diese Tatsache zurückzuführen.

Beobachtungen über den Einfluß der Tonlage auf Blasinstrumente kann man in der praktischen Musik besonders gut anstellen, wenn mehrere davon in konzertierender Weise zusammenwirken. So ist mir bei einer

<sup>1)</sup> Vgl. zum Einfluß der Höhe u. a. HERRMANN-GOLDAP S. 85, KOEHLER I, I, S. 36 (Waldhorn), MILLER S. 193 (Flöte), GARTEN 3, IX, S. 38 (Vokal A). Zum Einfluß der Intensität u. a. WEISS 2, MILLER S. 193 (Flöte), 204 (Horn).



Aufführung von MOZARTS Serenade für 13 Blasinstrumente durch Mitglieder der Kapelle der Berliner Staatsoper viel Derartiges aufgefallen, z. B. daß das Horn im *po* und *mf* infolge der Weichheit und zerflossenen Breite seines Klanges öfters, ähnlich wie die einfachen Töne, um eine Oktave vertieft erschien; daß die Oboe in tiefer Lage ( $c^1-g^1$ ) pistonartig, sehr schön, mehr dem Vokal A als dem E verwandt, klang, was auf stärkere Teiltöne der 2-gestrichenen Oktave hinweist; u. dgl.

Vom Gesichtspunkt der reinen Klangfarbe sind nach den Analysen wie den Synthesen die Verschiedenheiten zwischen den Tonlagen eines Instrumentes zuweilen größer als die der Instrumente selbst. Die Klarinette ist in dieser Hinsicht kaum noch als ein Instrument zu bezeichnen; sie hat in jeder ihrer 3 Oktaven ein anderes Register, natürlich unter stetigen Übergängen, so daß man allenfalls auch nur 2 unterscheiden kann, das des hohlen und das des ausgefüllten Klanges, die um  $c^2$  ineinander übergehen. Vom Fagott gilt Ähnliches. Bei der Viola hat besonders die oberste Saite eine durchaus verschiedene (nicht immer angenehme, zu dünne und näselnde) Klangfarbe, deren unvermittelter Eintritt zu Verbesserungsversuchen (*Viola alta*) geführt hat. Beim Waldhorn sind die tiefsten Klänge, wie *C*, mehr posaunenartig usw.

Man kann geradezu fragen, was in solchen Fällen noch den verschiedenen Regionen eines Instrumentes gemeinschaftlich ist. Wir werden der Frage unter 8. nähertreten.

Daß aber auch dasselbe Instrument in derselben Region durch bestimmte Kunstgriffe eine veränderte Klangfarbe erhält, ist bekannt: Blasinstrumente durch „Stopfen“, Streichinstrumente durch Sordinen und Flageolettspiel, andere durch Modifikationen des Blasens, des Zupfens an verschiedenen Stellen der Saiten usw. Hierbei ändert sich eben immer die Klangstruktur, besonders im Sinne der Annäherung an einfache Töne. Aber wirklich einfach sind die Töne des Flageoletts ebensowenig wie die der Fistelstimme (RITZ). Bei der Violine sind es gerade die feinen Unterschiede je nach der Art der Bogenführung, die die Ausdrucksfähigkeit dieses seelenvollsten Instrumentes noch erhöhen. Ähnlich bei den übrigen Streichinstrumenten<sup>1)</sup>. Auch hat jede der Saiten wieder ihre eigene Klangfarbe; der nämliche Ton klingt anders auf der *G*-, anders auf der *D*-Saite der Geige oder des Cello, anders auch auf der freien Saite und anders auf der durch den Finger verkürzten. Man

<sup>1)</sup> Vgl. hierüber das gründliche Werk WILH. TRENDELENBURGS: Die natürlichen Grundlagen der Kunst des Streichinstrumentspiels, 1925. Bei den ungewöhnlichen Stricharten des „Flautato“ (wo der 2. mehr als der 1. Teilton hörbar wird) und des „Sul ponticello“ bemerkt Tr., daß auch hohe unharmonische Teiltöne entstehen können (S. 25 ff.).

spielt darum eine Melodie soweit als möglich auf derselben Saite, und oft ist dies auch vorgeschrieben. Die freie Saite wird von hervorragenden Spielern — bei JOACHIM war dies regelmäßig zu beobachten — sowenig wie möglich gebraucht, was allerdings nicht bloß im Streben nach gleichmäßiger Klangfarbe seinen Grund hat, sondern auch in den durch den musikalischen Zusammenhang gebotenen Modifikationen der Tonhöhe, die nur durch Greifen erzielt werden können. Endlich ist sogar auf der nämlichen Saite ein Unterschied, je nachdem ein, sei es freier sei es gegriffener, Ton auf einer oder mehreren der übrigen Saiten gleiche mitschwingende Töne erregt. So klingt ein auf der *D*-Saite der Geige gegriffenes  $a^1$  durch das Mitklingen der freien Saite kräftiger und voller als ein  $b^1$ , ebenso ein  $g^1$  auf derselben Saite durch das Mitklingen desselben Tones als Teiltones der *G*-Saite etwas kräftiger als ein  $f^1$ .

Da dem Musiker alle diese Modifikationen aus der Praxis geläufig sind, auch wenn er sie sich nicht theoretisch zum Bewußtsein gebracht hat, so gehören sie für ihn eben mit zum Begriffe des jeweiligen Instrumentes, ebenso wie die Modifikationen eines Vokals in den verschiedenen Stimmregistern in dem Begriff des Vokals eingeschlossen sind. Aber sie zeigen auch wieder, wie wenig sich die Begriffe „Klangfarbe“ und „Instrument“ gegenseitig decken.

Man kann fragen, welches unserer Instrumente die mildeste, also obertonärmste Klangfarbe, welches die schärfste, also obertonreichste besitze, und wie man etwa ein klangschönstes, ideales Musikinstrument definieren könnte.

Halten wir uns nur an die bei uns gebräuchlichen Instrumente, so sind wohl die beiden Extreme auf der Orgel vertreten. Am mildesten sind bekanntlich die weitmensurierten „gedackten“ Register. Aber den 3. Teilton habe ich hier immer, auch bei schwachem Anblasen, mit bloßem Ohre relativ sehr stark hören können. Mit einer Hilfsgabel von annähernd gleicher Höhe gibt er kräftige Schwebungen. Beim Register „Hohlflöte“ konnte ich mit bloßem Ohr sogar den 5. Teilton, wenn auch schwach, noch hören (I, II, S. 161). In dieser Hinsicht muß ich MILLER widersprechen; er müßte denn bei seinen Synthesen Pfeifen benützt haben, bei denen durch irgendwelche besondere Kunstgriffe die Duodezime auf ein Minimum reduziert war. Die einfachsten Klangquellen sind doch Resonanzgabeln, die bei schwachem Anschlagen mit dem Finger keine Obertöne, bei stärkerem Streichen mit dem Bogen nur die Oktave, bei sehr starkem noch eine äußerst schwache Duodezime hören lassen (der störende unharmonische

Beiton, eine verstimmte Duodezime, kann durch Anbringung eines Gummiringes in der Gegend des unteren Drittels der Zinken beseitigt werden). Aber sie sind als musikalische Instrumente nur vorübergehend gebraucht worden; die vor 100 Jahren gelegentlich gebauten Stimmgabelklaviere sind Kuriositäten geblieben.

Die schärfsten unter den gebräuchlichen Klangfarben finden sich wohl gleichfalls auf der Orgel. Betrachtet man sie als ein einziges Instrument, so ist natürlich das volle Werk, auch wenn nur eine Note oder ein Oktavenkomplex angegeben wird, von der allergrößten Zusammengesetztheit und gleichwohl einheitlich. Aber auch unter den einzelnen Orgelregistern kommen äußerst scharfe Farben vor. Ein mir von Herrn Orgelbauer KLAIS (Bonn) übersandter Pfeifensatz gibt Klänge, die man nur im Zusammenklang mit anderen Registern musikalisch erträglich machen kann.

Ein ideales Instrument wäre nach MILLER ein solches, dessen Teiltöne mit zunehmenden Ordnungszahlen gleichmäßig an Intensität abnehmen (S. 211ff.). Nach seinen Tabellen würde der durchschnittliche Hornklang diesem Ideale nahekommen (S. 202ff.). Freilich würden die charakteristischen Eigenheiten bei einer solchen Struktur verlorengehen. Auch der Hornklang ist nach MILLERS eigener Analyse niemals in Wirklichkeit so gebaut, weder bei starkem noch bei schwachem Blasen. Der berechnete Durchschnitt ist ja ein niemals realisiertes Abstraktum. MILLER hat dann aus 10 Resonanzgabeln, die in den Verhältnissen der harmonischen Teiltöne zueinander standen, einen Klang künstlich zusammengesetzt, den er als vollkommen einheitlich und doch so voll und weich schildert, wie er selten von einem Musikinstrument gehört werde.

Einen Versuch ähnlicher Art habe ich mit noch mehr Gabeln 3 Jahrzehnte lang in meiner Psychologievorlesung am Schlusse des akustischen Kapitels vorgeführt. Es waren 5 gleichzeitige reine Dur-Dreiklänge auf den Grundtönen 100, 200, 400, 800, 1600. Die Verhältnisse dieser 15 Töne sind nicht identisch mit denen der harmonischen Teiltonreihe, doch ist diese vom 1. bis zum 6. Teilton lückenlos und sind weitere 6 in Abständen darin vertreten. Der Zusammenklang ist auch so sehr einheitlich, aber noch reicher und von wunderbarer Schönheit. In seinem langsamen Verklingen wirkt er wie eine Sphärenharmonie; und da die höheren Gabeln rascher verklingen als die tiefen, so kommen auch Intensitätsverhältnisse von der Gleichheit bis zu jedem Grade der Abnahme von unten nach oben dabei zustande. In gewissem Sinne mag dies nun wohl ein idealer Klang genannt werden. Man könnte auch recht wohl einen Mechanismus ersinnen, der das

gleichzeitige Anstreichen oder Anschlagen der Gabeln besorgte, die dann an bestimmten Stellen einer musikalischen Komposition, etwa am Schluß eines feierlichen Hymnus, erklingen könnten. Aber ein ganzes „Mixtur-Register“ dieser Art ließe sich natürlich nicht ohne die größten Umständlichkeiten herstellen, und geschähe es, so wäre noch die Frage, ob die fortgesetzten fünffachen Dreiklangsparellen nicht selbst das moderne Ohr unangenehm berühren würden.

Auf den nämlichen Grundtönen ist in der Sammlung unseres psychologischen Instituts auch der reine Moll-Dreiklang vertreten. Der Unterschied ist sehr lehrreich: der Wohlklang bedeutend geringer.

Kirchenglocken sollten nicht im Dreiklang, nicht einmal im Durdreiklang gestimmt werden, da ihre Töne doch hauptsächlich in der Aufeinanderfolge zum Vorschein kommen, wo es mehr auf melodische als auf harmonische Intervalle ankommt und Halbtonstufen stimmungsvoll wirken können. Vollends entsetzlich wäre es, wenn die Automobilhupen (täuscht mich nicht die Erinnerung, so hat ein großer Naturforscher einmal in Zeitungen diesen Wunsch ausgesprochen) in Dreiklängen abgestimmt würden, die natürlich auch noch meist unrein herauskommen würden.

#### 7. Tiefere Erklärungsgründe.

Wie hängen nun die Komplexeigenschaften, die wir der Klangfarbe im engeren Sinne zurechnen, das Weiche, Volle, Hohle, Scharfe, Schmetternde usw., mit ihrer Teiltonstruktur zusammen?

Schon die bloße Anzahl der Teiltöne bewirkt einen ersten Unterschied: der einfache Ton klingt gegenüber dem zusammengesetzten unanalysierten Klang dürrig, wesenlos; dieser erscheint voller, reicher. „A pure tone is a poor tone“, sagt MILLER treffend. Man hat, wie es scheint, durch die Beschaffenheit der Komplexempfindung eine unmittelbare Kunde davon, ob eine Sinnessphäre des Gehirns mehr oder weniger ausgefüllt ist. Auch beim mehrstimmigen Gesang wird man einen 6-stimmigen von einem 3-stimmigen Chor bei gleicher Gesamtzahl der Sänger durch die größere Fülle unterscheiden, auch ohne die einzelnen Stimmgattungen auseinanderzuhalten und zu zählen (wozu übrigens bei Sechsstimmigkeit, zumal im Palestrinastil, selbst Geübte nicht so leicht imstande sind).

Den Farbenunterschied des einfachen „Tones“ vom „Klange“ hat auch HELMHOLTZ, der diesen Bezeichnungen zuerst feste Begriffe unterlegte, hervorgehoben. Im übrigen begnügte er sich hier mit einigen rein empirischen Regeln, die einer Zurückführung auf tiefere Gründe bedürftig und fähig waren. Daß er dazu nicht kam, hatte seinen Grund in der unvollständigen Beschreibung der

Eigenschaften einfacher Töne. Er schreibt ihnen außer den Höhenunterschieden nur noch solche der Stärke zu, „da die Form einfacher Wellen vollständig gegeben ist, wenn ihre Schwingungswerte gegeben ist“ (S. 120). Physikalisch ist dies unbezweifelbar richtig, denn Sinusschwingungen können sich außer in der Länge, von der die Höhe des Tones abhängt, nur noch in der Amplitude unterscheiden. Aber damit ist nicht gesagt, daß die unter Vermittlung verwickelter Gehirnprozesse daran geknüpften sinnlichen Erscheinungen nicht noch andere Unterschiede aufweisen, die sich mit der Höhe zwar parallel, aber nicht notwendig in überall gleichem Grade verändern. Dies ist in der Tat der Fall. HELMHOLTZ selbst sagt von den einfachen Tönen: „sie klingen sehr weich und angenehm, ohne alle Rauigkeit, aber unkräftig und in der Tiefe dumpf“. Damit sind doch schon mehrere Unterschiede, wie Weichheit und Schärfe, Dumpfheit und Helligkeit, Kräftigkeit und Unkräftigkeit, zugegeben. Aber diese Beschreibung muß erläutert, ergänzt und teilweise auch berichtigt werden.

Die Höhe, die wir mit Helligkeit identisch setzen, involviert die gewaltigen Unterschiede von den tiefsten, dunkelsten bis zu den höchsten, hellsten Tönen. Damit gehen parallel Unterschiede des Volumens, das mit zunehmender Höhe abnimmt, aber sich in der mittleren Lage weniger stark verändert. Tiefen Tönen ist eine gewisse Breite und Verschwommenheit eigen, mit der sie uns wie ein Medium umfassen, während die hohen unstreitig immer spitzer werden, und beides nicht etwa bloß im figurlichen Sinne, auf Grund bloß assoziierter räumlicher Vorstellungen, sondern zufolge einer immanenten räumlichen oder raumähnlichen Beschaffenheit. Wenn HELMHOLTZ die tiefen Töne „dumpf“ nennt, so ist damit vielleicht ihre Dunkelheit und ihr Volumen gleichzeitig gemeint (wie wir von einem dumpfen Schmerz reden). Genau genommen ist dies aber schon eine Mehrheit von Eigenschaften.

Über das Volumen vgl. Tonps. II S. 57 u. ö. RICH S. 121, v. HORNBOSTEL 2, S. 708. WAETZMANN spricht (3, S. 684) die glaubwürdige Vermutung aus, daß die Volumenunterschiede physiologisch bedingt seien durch die nach den höheren Tönen hin abnehmende Breite der erregten Fasergruppe in der Basilarmembran. Dann muß bei stärkerer Erregung auch das Volumen eines Tones wachsen, was den Tatsachen wohl entsprechen könnte.

H. J. WATT hat in seinen beiden interessanten Schriften „The Psychology of Sound“ 1917, und „The Foundation of Music“ 1919 dieses Attribut geradezu als das primäre bezeichnet und die wichtigsten musikalischen Phänomene darauf zurückgeführt. Einer solchen Ausdehnung und Verwertung des Begriffes könnte ich mich aber nicht anschließen.

Die auffallende Breite mancher Klänge, besonders des Waldhornklanges, auch mancher Baß- und Altstimmen, ebenso das eigentümlich Zerflossene

der tieferen subjektiven Töne gegenüber gleich hohen einfachen objektiven Tönen, legen immer wieder die Frage nahe, ob nicht 2 Töne auch bei gleicher Höhe und Stärke sich noch durch ihr Volumen unterscheiden können. Aber bei objektiv erzeugten Tönen müßten doch solche Unterschiede durch objektive Verschiedenheiten des Reizes hervorgerufen werden. Es ist nicht abzusehen, welche Eigenschaften der Tonschwingungen bei gegebener Höhe, Stärke und Richtung des Eintreffens noch als variable Faktoren wirken könnten. Bei den subjektiven Tönen liegt es anders; da könnten vielleicht Ausdehnungsunterschiede in der Erregung der Nervenendigungen oder der zentralen Gebilde noch unabhängig von Höhe und Stärke auftreten. Aber über die Entstehung solcher Töne wissen wir eben überhaupt noch so gut wie nichts.

v. HORNBOSTEL unterscheidet (2, S. 708) neben dem Volumen noch andere ähnliche Eigenschaften, wie Dichtigkeit, Gewicht, bei denen es mir aber zweifelhaft erscheint, inwieweit sie dem ursprünglichen Empfindungsmaterial als solchem zukommen und nicht etwa bloß auf mehr oder weniger zufälligen Vorstellungsassoziationen beruhen, die immerhin große sinnliche Lebhaftigkeit erreichen können. „Behäbigkeit“ gehört sogar sicher in diese Klasse. „Dichtigkeits“-Unterschiede würde ich noch am ehesten zugeben, sofern bei tiefen Tönen mit ihrer Breite auch eine eigentümliche Zerflossenheit auffällt, während hohe kompakter klingen. Aber sind dies nicht doch nur andere Bezeichnungen für die Volumenunterschiede? Mir scheint nur dann eine nachweisbare Berechtigung zur Unterscheidung von Attributen vorzuliegen, wenn sich zeigen läßt, daß sie in gewissem Maße unabhängig variieren. Auch Eigenschaften, die erst in der Verbindung oder Aufeinanderfolge mehrerer Empfindungen auftreten, wie die Unterschiedsempfindlichkeit oder die Beweglichkeit, würde ich nicht zu den Attributen im engeren Sinne zählen, obschon sie für die Charakteristik höherer gegenüber tieferen Tonlagen sehr wesentlich sind.

Man kann ferner sagen, daß die Stärke einfacher Töne bei gleicher Reizstärke mit der Höhe zunehme. Bekanntlich wächst die Empfindlichkeit des Ohres bis in die Gegend der 4-gestr. Oktave. Bei noch höheren Tönen nimmt sie zwar wieder ab, aber dafür tritt eine vermehrte Reizbarkeit des ganzen Nervensystems infolge von Gehörseindrücken (größere „Eindringlichkeit“) an die Stelle, und die höchsten Töne sind, wenn sie die Reizschwelle erheblich übersteigen, von äußerst durchdringender, fast unerträglicher Wirkung.

Diese 3 Eigenschaften: Helligkeit, Breite und Stärke, faßte ich (1, II, S. 524ff.) unter dem Namen der „Tonfarbe“ zusammen und versuchte aus der Tonfarbe der in einem Klang enthaltenen Teiltöne die Klangfarbe des Ganzen zu verstehen. Begreiflicherweise wird ein Klang um so heller, schärfer, durchdringender, je mehr hohe und je höhere Töne zum Grundton hinzukommen, wie die Suppe um so salziger schmeckt, je mehr Salz hinzukommt, indem die Eigenschaften der Teile in gewissem Grade auf das unanalyisierte Ganze übergehen. Zugleich folgerte ich, daß die Klangfarbe keineswegs nur von der relativen, sondern

auch von der absoluten Höhe der Teiltöne (einschließlich des Grundtones) abhängt. 2 Klänge von gleich vielen, in gleichen Intensitätsverhältnissen stehenden harmonischen Teiltönen haben völlig verschiedene Klangfarbe, wenn einer der tiefen, einer der hohen Region angehört; und schon bei geringen Höhenunterschieden müssen notwendig auch entsprechende Klangfarbenunterschiede auftreten. Diese Folgerungen erscheinen mir auch heute unbestreitbar.

Daß auch die Schwebungen der hohen Teiltöne untereinander zu der markigen, unter Umständen schmetternden Beschaffenheit eines Instrumentklanges beitragen, hat schon HELMHOLTZ bemerkt. Namentlich bei tiefen, sehr obertonreichen Klängen, wie denen tiefer Metallzungen (mir standen solche von 20, 30, 50 Schw. zur Verfügung) oder der freien *C*-Saite des Cello (66 Schw.), wo sämtliche benachbarten Teiltöne ebenso viele Schwebungen untereinander machen, muß die starke Rauigkeit entstehen, wie sie tatsächlich beobachtet wird. Auch gewisse näselnde oder prickelnde Beimischungen hängen mit Schwebungen sehr hoher Beitöne zusammen<sup>1)</sup>.

Die Differenztöne, die natürlich für den Klang ebenfalls von Bedeutung sind, indem sie ihn dunkler färben als er ohne sie erscheinen würde, rechnen wir hier von vornherein bei der Bestandsaufnahme dem Klangkomplex zu. In Klängen z. B., die den Durdreiklang (4 : 5 : 6) stark enthalten, sind durch kräftige subjektive Differenztonbildung nicht nur der Grundton, sondern auch die Töne 2 und 3, also die ganze tiefere Teiltonreihe, verstärkt, es kommen also deren Eigenschaften mit zur Geltung. Dies wird den Klang besonders voll und reich machen. Kommt aber der Grundton eines tiefen Klanges ausschließlich oder hauptsächlich als Differenzton der Obertöne zustande (wie dies z. B. bei der Übertragung von Klängen der großen Oktave durch den Rundfunk

<sup>1)</sup> MILLER weist darauf hin (S. 138ff.), daß man das Vorhandensein von Schwebungen in bestimmten Klängen oft schon aus der Kurvenform ablesen könne, und führt als Beispiel einen Klarinettenklang an, in dessen Kurve er Schwebungen hoher Teiltöne erkennt. Die Analyse dieser Kurve ergab ihm die folgenden relativen Amplituden der ersten 12 Teiltöne: 29, 7, 20, 1, 2, 6, 6, 8, 16, 9, 30, 35. Hier sind der 11. und 12. Teilton außerordentlich stark und müssen merkliche Schwebungen in der Frequenz der Grundtonschwingung geben, vorausgesetzt allerdings, daß diese Frequenz nicht die Grenze der Merklichkeit für Schwebungen übersteigt. Leider ist die Höhe des Grundtones nicht angegeben. War es aber ungefähr *a* oder *c*<sup>1</sup>, dann lagen die beiden Teiltöne in der 4-gestrichenen Oktave innerhalb des Näselformanten und gaben miteinander 200–250 Schwebungen pro Sekunde, die sich in dieser Höhe noch als ganz leichtes Prickeln geltend machen.

der Fall sein dürfte), so muß der Klang noch sehr viel rauher und geräuschiger werden, als wenn der Grundton objektiv kräftig vorhanden ist.

Es wurde schon früher (I, II, 525, 531) bemerkt und soll auch hier noch besonders betont werden, daß von einer Mischung der Eigenschaften hier doch nicht ganz in demselben Sinne gesprochen werden kann wie bei den Speisen oder Gerüchen, wo schon die objektiven Bestandteile sich miteinander chemisch oder physikalisch zu einem Gesamtreiz verbinden. Niemals entsteht aus einem hohen und einem tiefen Ton ein mittlerer. Der Grundton rückt nicht infolge der beigemischten Obertöne in der Tonlinie höher hinauf, sondern es erhält nur das Ganze eine größere Komplexhelligkeit. Diese kann uns freilich verleiten, auch den Grundton, durch den wir die Höhe des Klanges definieren, höher anzusetzen als denselben Grundton ohne Obertöne. Aber sobald man aufmerksam beide Erscheinungen vergleicht, erkennt man das Urteil als irrig; und gibt man sie gleichzeitig an, so ist der einfache Ton, wenn er aus gleicher Richtung kommt, nicht mehr gesondert hörbar. Bei Klängen hat es daher einige Berechtigung, ihre Höhe noch von ihrer Helligkeit zu unterscheiden: die Höhe ist immer gleich der Höhe des Grundtones, die Klanghelligkeit dagegen ist größer als die des tiefsten, kleiner als die des höchsten seiner Teiltöne, sie kann der Helligkeit eines zwischenliegenden einfachen Tones einigermaßen zugeordnet werden, liegt aber nicht in der Linie der einfachen Helligkeiten selbst, sondern sozusagen in einer anderen Dimension.

Umgekehrt ist das Klangvolumen kleiner als das des tiefsten, größer als das des höchsten Teiltones, gleichwohl nicht schlechthin identisch mit einem mittleren, sondern einem solchen nur zuzuordnen. In beiden Fällen ist daher von einer Misch- oder Durchschnittsbildung nicht im eigentlichen wörtlichen Sinne zu reden.

Bei den Stärken der Teiltöne wird man ohnedies nicht an eine Mischung denken. Daß aber auch nicht eine Summierung im eigentlichen Sinne stattfindet, wurde schon oben (II. Kap.) erwähnt. Das Klangganze ist nicht stärker als der stärkste Teilton. Es ist nur reicher, dichter, massiver.

Wenn a. a. O. S. 540 gesagt wurde, die Eigenschaften der Teile gingen scheinbar auf das unanalyisierte Ganze über, so ist dieses „scheinbar“ nicht so zu verstehen, als besäßen nur die Teile selbst wirkliche Eigenschaften und seien die Klanghelligkeiten usw. nur Urteilstäuschungen, sondern so, daß eben keine Verstärkung, Erhellung, Zuspitzung in demselben Sinne stattfindet, wie wenn man mit einfachen Tönen von der Tiefe zur Höhe übergeht. Immerhin wird der Ausdruck „scheinbar“ des naheliegenden Mißverständnisses wegen besser vermieden.



Nun muß aber auf Grund der gegenwärtigen Untersuchungen noch eine weitere, damals nur kurz (S. 543) gestreifte Eigenschaft der einfachen Töne zur Erklärung der Klangfarbe herangezogen werden: die Vokalität oder die Tonfarbe im engsten und eigentlichsten Sinne. Freilich sind es hier, wie bei den Vokalen selbst, nicht die Vokalitäten als wahrnehmbare Eigenschaften der einfachen Töne, sondern die physiologischen Vokalvalenzen, die als Erklärungsmittel dienen müssen und können. Denn während die Unterschiede der Helligkeit, Breite, Stärke für jedermann leicht erkennbar sind, bleiben die Vokalähnlichkeiten einfacher Töne, abgesehen von ihrem U- und I-Charakter, den meisten nur schwer erkennbar. Es handelt sich also nur um jene Gehirnprozesse, denen zufolge Töne zwischen  $g^1$  und  $g^4$  in Verbindung miteinander und mit tieferen Tönen die gesättigteren Vokalqualitäten hervorbringen.

Es ist nun nicht anders möglich, als daß unter den zahllosen Kombinationen von Teiltönen in bestimmten Stärkeverhältnissen, wie sie sich in den Instrumentalklängen von verschiedener Tonhöhe finden, beständig auch solche vorkommen, die zufolge der Höhen- und Stärkeverhältnisse der Teiltöne mehr oder weniger deutliche Vokalähnlichkeiten in die Klänge hineinbringen. Wir haben schon auf solche Fälle hingewiesen, und je mehr man darauf merkt, um so klarer erkennt man den Einfluß dieses Faktors auf das gesamte Klangfarbengebiet. Die Extreme U und I wird man nur selten vertreten finden (wie in den Klängen tiefer gedackter Pfeifen und den höchsten Tönen des Pikkolo, der Harfe und der hohen Orgelregister), dagegen fast immer mehr oder weniger Anklänge an UO, O, A, Ä, E. So z. B. O in den Klängen des Horns und der Viola, soweit sie sich in den mittleren Regionen halten und daher unter ihren Teiltönen besonders der UO- und der O-Formant, Töne zwischen etwa  $e^1$  und  $e^2$ , stark vertreten sind<sup>1)</sup>. AÄ in Trompeten- und Klarinettenklängen mittlerer Höhen, Ä und E im Oboenklang, leisestes E auch im Näseln der Klarinette und besonders des Fagotts, wo es durch die Teiltöne zwischen  $c^4$  und  $g^4$  (den E-Formanten) vertreten ist; usf.<sup>2)</sup>. Die festen

<sup>1)</sup> Auch unter den KOEHLERSchen Trommelfellkurven sind die des Waldhorns denen des O zum Verwechseln ähnlich.

<sup>2)</sup> Man könnte fragen, warum Teiltöne der 4-gestrichenen Oktave nur bei bestimmten Instrumenten Näseln bewirken, nicht aber z. B. bei der Tenorposaune, wo für den Grundton  $c^1$  gleichfalls auf  $d^4$  ein 2. Maximum liegt. Aber hier kommt der Näselformant infolge der Stärke aller darunter- und darüberliegenden Teiltöne nicht zu seiner spezifischen Wirkung. Auch im gesungenen A der Männerstimme sind ja in der Nähe meistens Teiltöne dieser Tonlage enthalten, ohne daß sie ein Näseln zur Folge haben müssen.

Nebenformanten der Instrumente, einschließlich des Näselsegmenten, sind nichts anderes als die Vokalformanten, eingefügt in Instrumentalklänge. Wie schon bemerkt, sind mir diese Vokalähnlichkeiten auch bei der Synthese der Instrumentalklänge als Fingerzeige vielfach nützlich gewesen. Übrigens sind sie teilweise auch schon HERRMANN-GOLDAP, KOEHLER und MILLER (A-Ähnlichkeit der Flöte auf  $b^1$  infolge des starken Teiltönen  $b^2$ ) aufgefallen<sup>1)</sup>.

Es ist so zwischen den Vokalen der menschlichen Stimme und den instrumentalen Klangfarben nicht bloß kein prinzipieller Gegensatz, wie er von HERMANN u. a. behauptet wurde, sondern ein sehr naher und enger Wesenszusammenhang. Noch deutlicher als schon im 10. und 13. Kap. erkennen wir nunmehr, wie wenig berechtigt es wäre, den Vokalen eine unbegreifliche Sonderstellung unter den Gehörseindrücken einzuräumen. Sie unterscheiden sich von den Instrumentalfarben nur dadurch, daß wegen der durch die Mundhöhle hergestellten kräftigen Resonanz die Farbvalenzen des „akzessorischen Prozesses“ zwischen  $g^1$  und  $g^4$  am reinsten zur Geltung kommen. Man kann sie geradezu dadurch definieren. Die nämlichen Tongruppen sind bei den Vokalen Hauptformanten, bei den Instrumenten Nebenformanten.

Nichts ist denn auch leichter, als einen synthetisch dargestellten Instrumentalklang in einen Vokal zu verwandeln: man braucht nur einige Drehungen an den Wirbeln des Regulierungsapparates vorzunehmen. So z. B. geht der Hornklang auf  $c^1$  in den Vokal A über, wenn man den Ton  $c^1$  auf Null zurückschraubt und  $e^3$  schwach hinzufügt. Ebenso leicht sind die Flötenklänge auf  $c^1$  und  $c^2$  in A überzuführen durch Schwächung der 2 tiefsten Teiltöne.

Zwischen der Vokalfarbigkeit und den 3 zuerst herangezogenen Eigenschaften einfacher Töne besteht jedoch der Unterschied,

Übrigens scheint auch die Eigenschaft der „Hohlheit“, die sich bei der Klarinette, in geringerem Grade auch beim Fagott, mit dem Näseln verbunden findet, nicht ohne Bedeutung für dieses, obschon beide Eigenschaften begrifflich zu unterscheiden sind und auch getrennt vorkommen. Durch die Hohlheit scheint das Näseln merklicher zu werden; die hohen Teiltöne machen sich durch den Kontrast mit der hohlen Grundlage — gewissermaßen einem negativen Unterformanten — stärker im Klange geltend.

<sup>1)</sup> KOEHLER hat darauf mit besonderem Nachdruck gegenüber HERMANN hingewiesen. Er erinnert an das „Tarantara“, mit dem der römische Dichter den Klang der Tuba wiedergibt. Dieses hat allerdings auch in der Stärke ein Tertium comparationis mit A, da dieser Vokal am stärksten hervorgebracht werden kann, ebenso wie vor der Trompete „Viola, Baß und Geigen schweigen“. Aber unstreitig liegt auch etwas Verwandtes in der Klangfarbe.

daß diese in den Teilen ausgeprägter sind als im Ganzen, jene aber in erheblichem Maße überhaupt erst durch die Verknüpfung mehrerer Teiltöne zustande kommt. Darüber sei auf die Betrachtungen des 13. Kap. zurückverwiesen.

Man könnte schließlich die Frage aufwerfen, ob nicht auch die Eigenschaft einfacher Töne, die wir „musikalische Qualität“ nennen (vgl. S. 91), einen Einfluß auf die Klangfarbe habe. Aber wir wüßten keine Klangfarbenunterschiede zu nennen, die sich darauf mit Bestimmtheit zurückführen ließen und nicht schon aus den oben erwähnten Eigenschaften herzuleiten wären. Nur ein Fall käme vielleicht in Betracht: die Auszeichnung des Grundtones  $\text{C}$  in der Musik. Durch viele Beispiele läßt sich belegen, daß C-dur für besonders kraftvolle und glänzende Wirkungen mit Vorliebe benützt wird, daß Sieg und Triumph, strahlende Helle darin einen oft überwältigenden Ausdruck finden. Daß die mit  $\text{C}$  bezeichneten Schwingungszahlen sich im Laufe der Jahrhunderte einigermaßen verschoben haben, würde keinen durchschlagenden Einwand bedeuten, da die musikalische Qualität, oder wenigstens ihre emotionelle Seite, sich mit verschoben haben kann. Aber die Frage wäre, ob diese glänzende Färbung aus einer schon ursprünglich vorhandenen Sonderstellung der  $\text{C}$ -Empfindungen oder ob nicht umgekehrt die bevorzugte Stellung im Musiksystem und auf der Palette der musikalischen Maler aus rein historischen, mit unserem Notierungssystem zusammenhängenden Ursachen und Gewöhnungen herzuleiten sei. Darum möge auch die obige Frage hier als solche stehenbleiben.

### 8. Klangmerkmale der Instrumente.

Was ist den Klangfarben der verschiedenen Lagen eines Instrumentes phänomenal gemeinsam?

Auf diese, mit dem Wiedererkennen des Instrumentes zusammenhängende Frage können wir jetzt zurückkommen. Es genügt nicht, physikalische Merkmale anzuführen, sondern es müssen die erscheinungsmäßigen Züge aufgewiesen werden, die durch solche Merkmale bedingt sind.

a) Als Folge einer relativ großen oder geringen Anzahl der Teiltöne müssen phänomenale Unterschiede der Fülle und Körperlichkeit (Dichtigkeit), aber auch der Schärfe und Helligkeit des Klanges resultieren. Nun kann es allerdings geschehen, daß ein Instrument X, das auf der Tonhöhe  $c$  mehr Teiltöne als Y aufweist, auf  $c^1$  oder  $c^2$  ebensowenig Teiltöne wie Y auf  $c$  enthält. Aber es wird dadurch nicht ebenso milde und dunkel

werden wie *Y* auf *c*, weil bei Erhöhung des Grundtones die gesamte Klangmasse in eine hellere Region gelangt. Die Helligkeit hängt eben nicht bloß von der Zahl, sondern auch von der absoluten Lage der Teiltöne ab<sup>1)</sup>.

Jedenfalls wird, wenn beide Instrumente wieder eine identische höhere Note angeben, etwa  $c^1$  oder  $c^2$ , ein analoger, wenn auch nicht mehr ebenso großer Unterschied zwischen ihren Klängen vorhanden sein wie vorher. Der Kenner ist aber gewohnt, bei der Beurteilung der Klangfarbe die jeweiligen Tonlagen selbst mit in Rechnung zu ziehen. Auf jeder Tonhöhe, abgesehen von den höchsten Lagen (etwa von  $c^3$  ab), gibt es eben relativ scharfe, mittelscharfe, weiche und sehr weiche Klänge, und jedes Instrument gehört, soweit es überhaupt sich selbst ähnlich bleibt, auf jeder Höhe zu der nämlichen Klasse. Auf diese Art ist es möglich, die Instrumente, trotz der Klangverschiedenheit ihrer Regionen, schon durch ihre relativen Schärfe- und Helligkeitsgrade einigermaßen auseinanderzuhalten und wiederzuerkennen, auch abgesehen von äußeren Kriterien.

b) Ein 2. Moment liefern die beweglichen Formanten. In bezug auf diese hat KOEHLER den Schluß gezogen, ihre Bedeutung für die spezifische Klangfarbe eines Instrumentes müsse auf den darin besonders stark vertretenen Intervallen als solchen beruhen. Denn wenn auf einem Instrument eine Tonleiter gespielt

<sup>1)</sup> Wie feinsinnige Komponisten die durch die absolute Höhenlage mitbedingte Komplexhelligkeit berücksichtigen, möge nur ein Beispiel zeigen. MENDELSSOHN schreibt im 5. und 9. Takt des Allegro maestoso assai, das den Abschluß seiner A-moll-Symphonie bildet, den Klarinetten einen Sprung in die tiefere Oktave vor (nur für diese 2 Takte), der die Melodie vollkommen entstellen würde, wenn man darauf achtete. Man überhört ihn, weil die Klarinetten hier nur zur Füllung und Färbung der melodieführenden Viola dienen. Aber warum diese Sprünge? Offenbar, weil die Gesamtfärbung zunächst eine geheimnisvoll dunkle sein soll, um sich dann immer mehr bis zu vollem Glanze zu steigern. Diese dunkle Färbung würden aber die Klarinetten beeinträchtigen, wenn sie, wie es die Melodie verlangt, an akzentuierten Stellen das höhere  $c^1$  intonierten, das schon in ihr 2. Klangfarbenregister hinübergreift. Vom 13. Takt an durchbrechen sie diese Grenze, dann beginnt aber auch schon die große Steigerung.

In der neuesten Musik sollen allerdings in melodieführenden Stimmen gelegentlich Oktavversetzungen einzelner Töne angewandt werden, die nicht im Klange untergehen, sondern sich dem Hörer aufdrängen und die Melodiegestalt selbst „färben“ sollen (ERPF a. a. O.). Melodien im früheren Wortsinne vertragen dergleichen nicht, aber „atonale“ mögen dadurch vielleicht nicht weiter geschädigt werden. In China macht man's schon lange ebenso mit Quinten (ERICH FISCHER).

wird, so verändern mit dem Grundton auch alle Teiltöne ihre absolute Höhe, nur ihre Schwingungsverhältnisse bleiben dieselben. Nun sind es zwar bei allen Instrumenten die gleichen Schwingungsverhältnisse, nämlich die der aufeinanderfolgenden ganzen Zahlen; aber bestimmte Verhältnisglieder sind durch ihre Stärke ausgezeichnet, z. B. 3 : 4 oder 4 : 5 : 6, und diese stellen dann eben den beweglichen Formanten dar. Also, schließt KOEHLER, muß eine den Intervallen als solchen zukommende phänomenale Eigenschaft, ihre „Intervallfarbe“, die gleichbleibende Klangfarbe eines Instrumentes bedingen. „Aus den Intervallfarben setzt sich demnach die Klangfarbe zusammen.“

Den letzten Satz dürfte er heute nicht mehr billigen, sofern auf keinen Fall eine bloße Addition stattfinden kann. Die „Farbe“ eines Akkords ist ja nicht die bloße Summe seiner Intervallfarben. Immerhin bleibt denkbar, daß das Vorhandensein eines oder mehrerer besonders starker Intervalle innerhalb der Reihe der Teiltöne dem Gesamtklang eine bestimmte Farbe aufprägte, die dann bei Instrumenten mit beweglichen Formanten in verschiedenen Höhenlagen ungeändert bliebe.

Dies ist aber zunächst eine theoretische Konstruktion, und der Schluß ist nicht einmal ganz zwingend, weil es noch andere Erkennungsmerkmale für die Instrumente gibt. Es fragt sich also, was von einem solchen Einfluß empirisch nachzuweisen ist. Unstreitig hat heutzutage jedes der musikalischen Intervalle und jeder Akkord, zunächst für musikalische Personen, im isolierten Zustand einen besonderen Charakter, wenn auch in der Beschreibung Einstimmigkeit schwer zu erzielen ist (vgl. m. Abh. 2, S. 339ff.). Man mag ihn als Farbe oder, wie es bisher meistens geschah, als ein daran geknüpftes Gefühl oder als eine Gefühls-empfindung bezeichnen. Die Frage ist aber: wieweit dieser Charakter an ein analysierendes Hören geknüpft oder wenigstens daraus entstanden ist, wieweit er auch für unmusikalische Personen vorhanden ist, und wieweit er sich innerhalb eines Einzelklanges, der sich von dem gewöhnlichen Akkord durch die ungleiche Stärke der Komponenten unterscheidet, vorfindet. Die heutigen Intervall- und Akkordfarben sind in der Hauptsache Entwicklungsprodukte unserer harmonischen Musik, also recht jungen Datums. Es ist zwar anzunehmen, daß der reine Dur-Dreiklang, wo er etwa zufällig früher vorkam, immer schon etwas angenehmer oder weniger unangenehm als andere Zusammenklänge empfunden wurde; aber wir haben keinerlei Anzeichen dafür, daß man ihn besonders hochgeschätzt oder geflissentlich aufgesucht oder gar Einzelklänge nach den verschiedenen darin verstärkt vorkommenden Dreiklangslagen

unterschieden und wiedererkannt hätte. Wenn exotische Völker, wie Chinesen, Araber, Instrumente und Stimmen von einer so schreienden Klangfarbe ertragen, ja lieben, wie wir sie nicht in unserem Orchester, jedenfalls nicht bei Soloinstrumenten, dulden würden (es sei denn in der exotisch bedingten „Jazz“-Musik), so könnte man daraus schließen, daß sie in der Tat kein Gewicht auf die in unseren Instrumentenklängen vertretenen Akkordwirkungen legen. Freilich gäbe es dafür auch noch andere Erklärungsgründe.

Und so müssen wir uns darauf beschränken, die wenigen, aber nicht unwichtigen Eigenschaften von Einzelklängen namhaft zu machen, die man psychologisch glaubwürdig aus Intervall- oder Akkordfarben herleiten kann und die auch für Unmusikalische noch in Betracht kommen. Vor allem gehört dahin der volle, runde Klang der Instrumente, in denen der 1. und 2. oder auch noch 3. harmonische Teilton in allen Regionen besonders stark vertreten sind. Denn auch Zusammenklänge aus 2 oder 3 selbständigen Klängen in diesen Verhältnissen (Oktaven, evtl. mit Duodezimen) wirken in gleichem Sinne, weil sie sich infolge der hohen Verschmelzung dieser Intervalle am meisten dem Eindruck eines einzigen Tones nähern und doch zugleich infolge ihrer Erstreckung auf der Tonlinie voller klingen als jeder allein. Dergleichen der Eindruck des Hohlen, Leeren, wenn die Teiltöne 1 und 3 in dem Klange hervortreten. Auch dieser Eindruck ist einer Duodezime aus zwei gleich starken selbständigen einfachen Tönen eigen. Ferner werden Klänge, in denen der Dur-Dreiklang in einer seiner 3 Lagen, besonders in der Primärlage 4 : 5 : 6, stark vertreten ist, sich, für uns Heutige wenigstens, durch besonderen Glanz und sinnliche Annehmlichkeit auszeichnen. Umgekehrt werden Klänge, in denen dissonante Intervalle, wie 8 : 9, oder Akkorde wie 7 : 8 : 9 oder auch nur 5 : 6 : 7 (etwa gleich einem verminderten Dreiklang) besonders stark vertreten sind, auch schon abgesehen von den Schwebungen, einen etwas weniger angenehmen Beigeschmack haben<sup>1)</sup>.

Dies etwa wären die Eigenschaften von Klängen, die aus „Intervall- und Akkordfarben“ plausibel hergeleitet werden können

<sup>1)</sup> In einem von MILLER (S. 201) analysierten Klarinettenklang ragen die Töne 8, 9, 10 gleichmäßig durch ihre Stärke über ihre Umgebung hervor. Aber da auch der 1. und 3. Teilton ziemlich stark sind, so erteilt dieser dissonante Mehrklang in Verbindung mit der dunklen Unterlage dem Ganzen doch nur die eigentümlich düstere Färbung, die die tiefere Klarinettenlage zum Ausdrucke des Elegischen, Romantischen, Dämonischen in entsprechendem Zusammenhange besonders befähigt.

und die, soweit überhaupt die Klangzusammensetzung auf verschiedenen Tonhöhen identisch bleibt, zur Wiedererkennung eines Instrumentes beitragen können. Aber so ausschließlich wie KOEHLER würde ich diesen Faktor nicht betonen, einmal, weil es mit der Wiedererkennung überhaupt nicht so gut steht wie er wohl voraussetzte, dann weil die Klangzusammensetzung nicht so identisch bleibt wie es abstrakt nach HELMHOLTZ zu erwarten wäre, endlich und besonders weil andere Kriterien entscheidender mitwirken. Damit kommen wir auf ein weiteres Moment:

c) Auch die festen Formanten, besonders die Vokalformanten, deren Vorkommen in Instrumentalklängen KOEHLER selbst schon betont hat, tragen sicherlich dazu bei, daß der Klang eines Instrumentes in verschiedenen Höhenlagen sich selbst ähnlich bleibt, solange eine gewisse Höhengrenze nicht überschritten wird, jenseits deren sich der Vokalformant unter den Obertönen nicht mehr oder nicht hinreichend vertreten findet, bzw. einem anderen Platz macht (wie bei der Viola auf der *A*-Saite, wo der *O*-Formant durch einen schwachen *Ä*-Formanten ersetzt wird): Wandlungen, mit denen der Kenner eben auch rechnet. Der Näselformant bleibt länger erhalten, da er der 4-gestr. Oktave angehört; er dürfte für die Instrumente, die ihn enthalten, ein Hauptmerkmal bilden.

Zu diesen inneren Merkmalen kommen in der Praxis des musikalischen Hörens die äußeren, die in den verschiedenen Lagen des Instruments der Regel nach gleichbleiben und daher ganz besonders zur Identifikation verhelfen.

Wir schilderten unter 1. die starken Irrtümer, die bei künstlichem Ausschluß solcher Kriterien entstehen können, erkennen aber nunmehr, warum und wodurch auch dann noch richtige Urteile überhaupt möglich waren. Sie wären sicher sogar weit zahlreicher ausgefallen, wären nicht unter den Instrumenten mehrere einander sehr nahestehende gewesen (es sollte damals gerade auch der Unterschied der einander ähnlichen und der unähnlichen Klangfarben in seinem Einfluß auf die Urteilszahlen geprüft werden). Bei den Vokalen unterscheiden wir im Deutschen nur 8 Typen. Würde man nur 8 instrumentale Klangfarben aussuchen, die in gleicher Weise die wesentlichsten möglichen Unterschiede von den mildesten bis zu den schärfsten Farben repräsentierten, jede von ihren Nachbarn ungefähr gleich weit abstehend, so würden zweifellos auch hier bessere Ergebnisse erzielt werden. So gute freilich wie bei den Vokalen, die schon das dreijährige Kind sicher wiedererkennt, würde man aus naheliegenden Gründen auch dann kaum erwarten können.

## 9. Gruppierungen und Mischungen.

a) Für die Klassifikation der Instrumente hat man bisher immer den genetischen Gesichtspunkt, die Art der Klang-erzeugung, zugrunde gelegt<sup>1)</sup>. In der Kompositions- und Dirigier-praxis unterscheidet man die Gruppen der Streich-, Holz-, Blech- und Schlaginstrumente. Von HAYDN bis MAHLER und STRAUSS ist die Orchestrierung damit gut gefahren. In neuester Zeit scheint man diese Gruppen aufzulösen zugunsten mannigfaltigerer „opalisierender“ Farbenspiele und „Klangfarbenmelodien“ (A. SCHÖNBERG). Dies würde zu akustischen Einteilungsgründen führen<sup>2)</sup>.

Während aber die Vokale infolge der Grenzen der Mund-stellungen ein begrenztes und in sich stetig zusammenhängendes System bilden, würde eine rein akustische Klassifikation der unendlich vielen möglichen Instrumentalfarben ihre Schwierigkeiten haben. Unmöglich wäre es nicht, gewisse Grundeinteilungen aufzustellen, in welche aber nicht bloß die Klangfarben- sondern auch die sonstigen akustischen Unterschiede aufzunehmen wären. So z. B. Instrumente mit stetig veränderlichen und mit festge-gebenen Tönen; mit konstanter und mit abklingender Tonstärke; mit wesentlich gleicher und mit stark veränderlicher Klangfarbe je nach der Tonlage, dem Stärkegrad und der Erzeugungsart; mit festen Maximis und ohne solche; mit verschwindender, schwacher, starker, überwiegender Geräuschbeimischung und rein geräuschige. Doch mag abgewartet werden, ob praktische Bedürfnisse zur Durch-führung solcher Gesichtspunkte drängen.

b) Weit mehr als bei den Vokalen interessieren bei instru-mentalenen Klangfarben multiple Formen, d. h. durch Ver-bindung mehrerer Instrumente entstehende Klangmischungen. Denn sie spielen in der Orchestermusik die allergrößte Rolle. Die Entwicklung derselben bringt in beschleunigtem Tempo immer neue Kombinationen, zum Teil allerdings auch mit neuen Instru-menten. Auch die Mischungen der mannigfachen Orgelregister, die teilweise selbst schon „Mixturen“ darstellen, und die von Orchester- und Gesangstimmen gehören hierher. Doch ist Prin-zipiellles nicht weiter zu besprechen. Immer bestimmt die Gesamt-summe der Teiltöne durch ihre eigenen Beschaffenheiten nach Hel-

<sup>1)</sup> Die logisch korrekteste und zugleich umfassendste, auch die exotischen und primitiven Instrumente einschließende genetische Klassifikation bei v. HORNBOSTEL und SACHS, Systematik der Musikinstrumente, Zeitschr. f. Ethnol. 1914.

<sup>2)</sup> Vgl. H. ERPF, Zeitschr. Melos 1905, S. 523ff., 532.



ligkeit, Stärke, Volumen, Vokalvalenz, Intervallen, resultierenden Schwebungen und Kombinationstönen, sowie begleitenden Geräuschen die Farbe der Mischung.

Auf die früher (I, II, S. 545ff.) behandelte Frage, wodurch wir Instrumente ungleicher Klangfarbe innerhalb einer solchen Mischung überhaupt noch unterscheiden können, während sie doch ihre sämtlichen Teiltöne in einer Art Gütergemeinschaft zusammenlegen, soll hier nicht noch einmal eingegangen werden<sup>1)</sup>. Hervorgehoben sei nur, daß eines der wichtigsten unter den damals aufgeführten Momenten, nämlich die ungleiche Lokalisation, nicht bloß für die Unterscheidung der Instrumente, sondern auch für die Klangfarbe der Mischung selbst von Bedeutung scheint: der Eindruck der räumlichen Breite, der Raumfüllung, und zwar im eigentlichen, optisch-räumlichen Sinne, wird durch die Verteilung der Instrumente oder Stimmen auf dem Podium (in manchen Fällen sogar auf verschiedene Teile des Konzertraumes, wie bei G. GABRIELI, in BERLIOZ' Requiem, auch gelegentlich bei den Knabenstimmen im I. Chor der Matthäuspassion) dem Klangeindruck hinzugefügt. Wie dies psycho-physiologisch wirkt, mag hier dahingestellt bleiben; natürlich ist dabei die Verteilung auf beide Ohren wichtig, aber mit ihr wirkt die visuelle Lokalisation und Raumvorstellung zusammen. Es ist mir immer merkwürdig erschienen, wie eigentümlich breit ein großer Chor oder eine Instrumentenmenge auch im Pianissimo wirken, ebenso eine leise murmelnde oder murrende Volksmenge auf dem Theater: es ist ein Eindruck, der durch keine einzelne Klangquelle von einem bestimmten Ort aus nachgeahmt werden kann, da er eben gerade mit der räumlichen Verteilung selbst zusammenhängt. Die kleinen Ungenauigkeiten der Intonation bei den Einzelnen bringen allerdings auch eine gewisse qualitative Breite in die Tonmasse, aber damit ist die Eigentümlichkeit des Eindruckes noch nicht vollständig beschrieben. Es scheint der Klangfarbe ein neues Moment zuzuwachsen, das nur Klangmischungen, nicht Einzelklängen, eigen ist. Beim Forte ist es vielleicht weniger auffallend, weil eben die Stärke selbst die Aufmerksamkeit auf sich zieht. Aber auch hier bringt es einen neuen Zug in die Erscheinung.

Vor kurzem hat der Radiobetrieb eine Einrichtung eingeführt, bei der eine große Anzahl von Mikrofonen an die Instrumente des Orchesters und auf der Bühne so verteilt werden, daß die Leitungen umschichtig zum rechten und linken Ohr des Teilnehmers führen. Durch dieses „stereophone“ Hören soll eine größere Plastik des Eindruckes erzielt werden, die unter

<sup>1)</sup> Vgl. darüber neuerdings v. HORNBOSTEL I, S. 78ff.

den obigen Gesichtspunkt fiele und als eine Verräumlichung des Klanges zu bezeichnen wäre. Es liegen aber darüber meines Wissens noch keine genügenden wissenschaftlichen Versuche vor.

So könnte man nun versuchen, die Wirkungen neuer orchestraler Mischungen sogar vorauszusagen. Wenn die Komplikationen immer größer werden, ist freilich, wie selbst bei der Berechnung der Gravitation von Körpern gegeneinander, eine exakte Voraussagung nicht mehr möglich und hätte in unserem Falle auch keinen praktischen Zweck, weil man die Wirkungen einfacher ausprobiert als berechnet. Der Theorie aber kann nur an den allgemeinsten Prinzipien gelegen sein, wie sie schon bei der Kombination weniger Elemente zutage treten.

#### 10. Blick auf frühere und künftige Untersuchungen.

Punkte der Übereinstimmung und der Abweichung unserer Ergebnisse gegenüber früheren Arbeiten über Instrumentalklänge sind im einzelnen bereits erwähnt. Nur in methodischer Beziehung sei einiges hinzugefügt. HELMHOLTZ' größtes Verdienst in diesem Gebiete liegt in der organischen Verknüpfung der Klanganalyse mit Betrachtungen über die Bauart und Tonerzeugungsweise der Hauptinstrumentenklassen. Wir beschränkten uns auch hier auf den rein phänomenalen Standpunkt, der allein schon genug neue Arbeit erfordert. Aber unstreitig muß eine künftige umfassende Behandlung der instrumentalen Klangfarben auch wieder zugleich dem genetischen Gesichtspunkt Rechnung tragen, z. B. der Frage nach dem Ursprung der festen Maxima. MEISSNER, der zuerst solche bei den Holzblasinstrumenten behauptete, suchte die Quelle hauptsächlich im Schalltrichter. Man wird der Frage nachgehen müssen, nachdem zuerst die Existenz und Lage dieser Maxima noch besser sichergestellt ist.

RITZ untersuchte bei Streichinstrumenten mit gutem Erfolge ihren Bestand an Obertönen mit Hilfe der auf jeder Saite zu erzeugenden Flageolettöne. Er fand bis zu 10 Teiltönen (gelegentlich, wie auf der tiefsten Saite eines Amati-Cello, noch sehr schwache höhere). Die Anzahl nahm ab mit dem Aufsetzen des Fingers und der Verkürzung der Saite. Das Hauptmerkmal der Streichinstrumente sucht er gleichwohl nicht im Klange selbst, sondern im Reibegeräusch des Streichens, das er den Konsonanten der Sprache vergleicht. Die fleißige Untersuchung ist heute noch lehrreich.

MEISSNER bediente sich (1881—1901) einer freilich noch unvollkommenen graphischen Methode, HERRMANN-GOLDAP der

durch HERMANN und WEISS eingeführten graphischen Vorrichtungen. Solche sind auch von den späteren ausschließlich benutzt worden, am erfolgreichsten von MILLER. Wenn wir mit den drei obigen Methoden ein Stück weitergekommen zu sein glauben, so möchten wir doch auch hier nicht den hohen Wert der Kurvenuntersuchungen, insbesondere mit den Mitteln der heutigen Elektrotechnik, in Abrede stellen. Manches einzelne kann überhaupt nur so untersucht werden, z. B. die feinen, flüchtigen Übergangserscheinungen, wie sie während des Strichwechsels des Bogens bei den Streichinstrumenten auftreten, oder der Klangverlauf bei den Klaviertönen<sup>1)</sup>. So wird denn auch in Zukunft nur der Wettbewerb der Methoden weiterführen. Um die Methoden selbst auszuprobieren und zu vergleichen, würden sich zunächst Orgelpfeifen der verschiedenen Register am meisten empfehlen, da sie vermittels eines konstanten Gebläses beliebig lang gleichmäßig auszuhaltende Töne geben<sup>2)</sup>.

Die Erforschung der Instrumentalfarben ist gegen die der Vokale stark zurückgeblieben. Aber nicht nur für die Theorie, auch für die Praxis wären hier noch lohnende Früchte zu ernten, und gerade diese praktischen Bedürfnisse, die vom Instrumentenbau und von den verschiedenen Formen mechanischer Reproduktion der Musik herkommenden Antriebe, werden am wirksamsten dafür sorgen, daß das Versäumte nachgeholt wird. So zweifle ich beispielsweise nicht daran, daß die akustische Definition eines hervorragenden Violinklanges durch Aufzeigung seiner Teiltonstruktur gegeben werden kann (vermutlich sind die tieferen Teiltöne absolut und relativ stärker als bei schlechten Geigen) und daß diese Erkenntnis in Verbindung mit physikalischen Erwägungen den Geigenmachern eine nachträgliche Rechtfertigung für manche rein empirisch gefundene Regel und einen Kompaß für ihre Entdeckungsfahrten zu neuen Konstruktionen liefern kann.

<sup>1)</sup> Über beides vgl. die Kurven MILLERS S. 194ff. und 207ff.

<sup>2)</sup> Herr Orgelbaumeister JOH. KLAIS in Bonn hat in lebhaftem Interesse für solche Untersuchungen dem Berliner Psychologischen Institut eine geeignete Auswahl solcher Pfeifen zur Verfügung gestellt, wofür ich ihm hier Dank sagen möchte. Ebenso danke ich Hrn. Orgelbauer FRANZ (Berlin) für vielfache Mithilfe bei der Instandhaltung des synthetischen Pfeifensystems.

## Nachträge.

Zu S. 84—86: Über den Einfluß der Konsonanten auf die Vokale vgl. auch JULIUS STOCKHAUSEN, Das Sanger-Alphabet. Signale, 1872, Nr. 30—39.

Zu S. 223 Z. 25 v. u.: Durch Ubertragung der Vibrationen des Motors auf die Scheibe konnten diese Tone ins Telephon gelangen.

Zu S. 224 vor c:

HR. PHILIPPS wird uber seine Berechnungsweise in verallgemeinerter Form demnachst in der Z. f. Sinnesphysiologie berichten.

Wie mir nachtraglich Herr Kollege KOEHLER mitteilt, hat er schon fruher gleichfalls solche Rechnungen uber JAENSCH' Kurven mit analogen Ergebnissen rein analytisch durchgefuhrt. Wir geben hier nach seinen Tabellen die Amplitudenverhaltnisse fur die Kurven II und III bei JAENSCH, wovon II bei 65—70 Umdrehungen pro Sekunde einen „A-artigen Klang“, III aber ein „gutes A“ gaben. Kurve II besteht aus 15 aufeinanderfolgenden Sinuswellen, denen auf der Scheibe Kreisbogen von 22—26° entsprechen, Kurve III aus ebenso vielen von den Bogenlangen 20—28°, also starkerer Streuung.

Teiltone	1	2	3	4	5	6	7	8	9	10
Amplituden fur II	0,1	0,2	0,2	0,3	0,2	0,2	0,9	1,0	0,7	1,1
Amplituden fur III	0,1	0,3	0,5	0,6	0,2	0,4	2,5	1,8	1,7	2,2

Teiltone	11	12	13	14	15	16	17	18	19	20	21
Amplituden fur II	1,5	1,3	4,9	7,3	28,7	6,6	3,9	3,8	1,0	0,3	—
Amplituden fur III	3,1	2,5	9,1	3,2	21,8	10,4	4,9	8,7	1,7	2,4	1,2

„Fur beide Kurven“ — schreibt KOEHLER — „ist das wesentliche Ergebnis, da sie Klange mit deutlichen Verstarkungsgebieten sind. Da beide Male das Zentrum des Verstarkungsgebietes bei 65—70 Umdrehungen der Scheibe auf etwa 1000 Schw. (= nahe  $c^3$ ) fallt, so mu ein ‚A‘ zustande kommen, und schon HELMHOLTZ wurde dies aus den gleichen Kurven haben ableiten mussen. II wirkt weniger gut gesattigt als ‚A‘, weil der eine Ton 15 zu stark hervorragt; III ist besser, weil mehrere Komponenten im Resonanzgebiet sich dem 15. an Intensitat nahern. JAENSCH erwahnt (S. 230) bei II eine gewisse Rauigkeit, bei III eine Gerauschbeimischung — kein Wunder, wenn so nahe benachbarte starke Tone miteinander in der Frequenz von 65—70 pro Sekunde schweben.“

Ubrigens hebt KOEHLER hervor, da man auch ohne diese hochst muhsamen Rechnungen schon aus den Konstruktionsbedingungen dieser „gemischten Sinuskurven“ das Vorherrschen des 15. Teiltons erschlieen konne, wie dies auch F. TRENDELENBURG (in diesem Buche S. 224) getan hat, mit dessen Voraussage KOEHLERS Rechnung vollkommen ubereinstimmt.

In einer weiteren Kurve (IV) hat dann JAENSCH eine noch größere Streuung der Wellenlängen vorgenommen, von 18–30°. Diese Kurve gab mit derselben Umdrehungszahl nur ein „Geräusch mit vokalisiertem Einschlag“. Hierüber schreibt KOEHLER (und man kann ihm nur zustimmen): „Wenn man sieht, wie sich die Analyse von II zu III verschiebt, so ist von vornherein klar, daß der Übergang von III zu IV dieselbe Verschiebung noch fortsetzen muß, und daß bei IV eine sehr breite Zone von starken Teiltönen um den 15. herum entsetzlich schweben und rasseln muß, wie das zu JAENSCH' Beobachtung der Klänge paßt.“

Bei alledem ist aber auch nicht zu vergessen, daß selbst ungemischte Sinuskurven, solche von gleichbleibender Länge und Höhe, wie JAENSCH' Kurve I, schon wesentlich verändert aus dem Telephon kommen (oben S. 223), daß daher auch bei den gemischten Kurven der Bestand an scharf miteinander kollidierenden Obertönen noch über die aus der FOURIER-Analyse ersichtlichen hinaus vermehrt werden muß, ja daß auch starke Differenz-töne im Telephon hinzukommen (SCHAEFER, WAETZMANN), deren Kollision untereinander das Rasseln noch vermehren muß.

### Druckberichtigungen.

- S. 47 ist in der Tabelle nach  $\frac{\lambda}{4}$  einzufügen: „in cm“.
- S. 147 ist im 1. Notenschema vor der 1. Note das b-Zeichen zu streichen.
- S. 236 Z. 4 v. u. lies „3, 4“ statt „2, 3“.
- S. 239 Z. 4 v. o. lies „Abstumpfung“ statt „Abstufung“.
- S. 282 Z. 1 v. o. lies „Einheitlichkeit“ statt „Einheitlichung“.
- S. 302 Z. 1 v. o. lies „5“ statt „4“.
- S. 354—368: In den Kolummentiteln der linken Seite lies „Physiologie“ statt „Psychologie“.

## Literaturverzeichnis.

### Abkürzungen:

- Z. = Zeitschrift, B. = Band, J. = Jahrgang, H. = Heft.  
Ann. d. Ph. = Annalen der Physik und Chemie.  
Beitr. = Beiträge zur Akustik und Musikwissenschaft, her. v. C. STUMPF.  
P.-Sch. = Beiträge zur Anatomie, Physiologie, Pathologie und Therapie des  
Ohres, der Nase und des Halses, her. v. A. PASSOW u. K. L. SCHAEFER.  
Pflüg. Arch. = Pflügers Archiv für die gesamte Physiologie.  
Z. Ps. = Z. f. Psychologie und Physiologie der Sinnesorgane (von B. 41,  
1906 an, in 2 Abteilungen: I. Z. f. Psychologie, II. Z. f. Sinnesphysiologie).

- ABRAHAM, O. Töne und Vokale der Mundhöhle. Z. f. Ps. I, B. 74. 1906.  
AUERBACH, F. (1) Untersuchungen über die Natur des Vokalklanges. Ann.  
d. Ph. (2), Ergänzungs-B. 8, 2. Stück 1876 (Bandjahr 1878).  
— (2) Bestimmung der Resonanztöne des Mundes durch Perkussion.  
Daselbst (3) B. 3. 1878.  
— (3) Über die GRASSMANNSCHE Vokaltheorie. Daselbst (3) B. 4. 1878.  
— (4) Die physikalischen Grundlagen der Phonetik. Z. f. franz. Sprache  
u. Literatur B. 16. 1894.  
— (5) Akustik, in WINKELMANN'S Handb. d. Physik 2. Aufl. B. 2. 1909.  
BENJAMINS, C. E. Über den Hauptton des gesungenen oder laut gesproche-  
nen Vokalklanges. Pflüg. Arch. B. 154. 1913. B. 155. 1914.  
BEVIER, L. (New-Brunswick) (1) Phonet. Abhandlungen, Physical Review  
B. 10, 14, 15 (1900—1902).  
— (2) Die akust. Analyse d. Vokale durch phonograph. Aufzeichnung.  
Physikal. Z. B. 1. 1900.  
— (3) The acoustic Analysis of the Vowel A. Die neueren Sprachen B. 8. 1900.  
BEZOLD, F. v. (1) Die Stellung d. Konsonanten in d. Tonreihe. Z. f. Ohren-  
heilk. B. 30. 1897.  
— (2) Das Hörvermögen der Taubstummen. 1896.  
— (3) Über d. funktionelle Prüfung d. menschl. Gehörorganes. 1897.  
BOEKE, J. D. Mikroskopische Phonogrammstudien. Pflüg. Arch. B. 50. 1891.  
BREMER, O. Deutsche Phonetik B. 1. 1893.  
BROCH, O. Slavische Phonetik. 1911.  
BROEMSER, PH. Theoret. Grundlagen der Konstruktion zweckmäßiger  
Mikrophone. Telegr.- u. Fernsprechtechnik 1919, Nr. 1.  
BROSS, KAZ. Über d. BEZOLDSche Sprachsekte. P.-Sch. B. 9. 1916.  
BRÜCKE, E. Grundzüge d. Physiologie u. Systematik d. Sprachlaute. 1856.  
CLAUS, G. Veränderungen des Hörvermögens bei abwärts fortschreitender  
Einengung d. oberen Tongrenze. P.-Sch. B. 19. 1923.  
CRANDALL, J. B. (1) Preliminary Analysis of 4 Semi-vocal Sounds (L, M,  
N, Ng). Physical Rev. (2) B. 23. 1924.  
— (2) and SÁCLA, C. F. Dynamical Study of the Vowel Sounds. Western  
Electr. Comp. Engin. Dep. Bell System. Techn. Journ. B. 3, Nr. 2. 1924.  
— (3) The Sounds of Speech. Daselbst B. 4, Nr. 4. 1925.

- CROSS, CH. R. HELMHOLTZ' Vowel Theory and the Phonograph. *Nature* B. 18. 1879.
- DONDERS, F. C. (1) Über die Natur der Vokale. *Arch. f. Holländ. Beitr. z. Natur- u. Heilkunde* B. 1. 1857. (Brief an E. BRÜCKE.)  
 — (2) Zur Klangfarbe d. Vokale. *Daselbst* B. 3. 1864; *Ann. d. Ph.* (2) B. 123. 1864.
- DWELSHAUVERS, M. F. V. Expériences sur l'intensité relatif des harmoniques dans les timbres de la voix. *Mém. de la Soc. Roy. des Sciences de Liège* (2) B. 16. 1880.
- EBERHARDT, MARG. Über die phänomenale Höhe u. Stärke von Teiltönen. *Psychol. Forschung (Z.)* B. 2. 1922.
- ENGEL, G. (1) Die Vokaltheorie von HELMHOLTZ u. die Kopfstimme. *Progr. d. Neuen Akad. d. Tonkunst.* 1867.  
 — (2) Studien zur Theorie des Gesanges. REICHERTS u. DU BOIS REYMONDS *Arch. f. Anat., Physiol. etc.* 1869.  
 — (3) Die Konsonanten d. deutschen Sprache. 1874.  
 — (4) Über d. Begriff d. Klangfarbe. *Vorträge d. Berliner philos. Gesellsch. N. F.* 12. H. 1887.
- EWALD, J. R. Zur Konstruktion von Polsterpfeifen. *Pflüg. Arch.* B. 152. 1913.
- FLETCHER, H. (1) The Nature of Speech and its Interpretation. *Journ. of the Franklin Institute.* B. 193, No. 6. 1922.  
 — (2) Physical Measurements of Audition and their Bearing on the Theory of HEARING. *Daselbst.* B. 196, No. 3. 1923.  
 — (3) The Physical Criterion for Determining the Pitch of a Musical Sound. *Physical Review* (2) B. 23. 1924.
- FLOWERS, J. B. The true Nature of Speech. *Proc. Americ. Inst. of Electr. Engineers* B. 35. 1916.
- FRANKFURTH, W., u. THIELE, R. Experimentelle Untersuchungen zur BEZOLDSchen Sprachsext. *Z. f. Ps.* II, B. 47. 1912.
- GARTEN, S. (1) Über die Verwendung der Seifenmembran zur Schallregistrierung. *Z. f. Biol.* B. 56. 1911.  
 — (2) Ein Schallschreiber mit sehr kleiner Seifenmembran. *Ann. d. Ph.* (4) B. 48. 1915.  
 — (3) Beiträge zur Vokallehre. *Abh. d. Sächs. Akad. d. Wiss., math.-phys. Kl.* B. 38, Nr. VII—IX (IX mit F. KLEINKNECHT). 1921.
- GISSWEIN, M. (1) Über die Resonanz der Mundhöhle und der Nasenräume. *P.-Sch.* B. 4. 1911.  
 — (2) Die Resonanzbeziehungen zwischen Stimme u. Brustorganen. *P.-Sch.* B. 22. 1925.
- GILDEMEISTER, M. (1) Untersuchungen über die obere Hörgrenze. *Z. f. Ps.* II, B. 50. 1918.  
 — (2) Bemerkungen zur Theorie des Hörens. *Daselbst.*
- GRABOW, A. Die Musik in der deutschen Sprache. 1879.
- GRADENIGO, G. (1) Il limite dell' udito nelle varie età. *Arch. Ital. di Otol.* B. 28. 1918.  
 — (2) Lacune e difetti tonali di carattere degenerativo familiare. *Daselbst. Suppl.* 1923, B. 3.
- GRASSMANN, H. (1) Übersicht d. Akustik etc. *Gymn.-Programm. Stettin* 1854.  
 — (2) Über die physikalische Natur der Sprachlaute. *Ann. d. Ph.* (3) B. 1. 1877.

- GRÜTZNER, P. (1) Physiologie der Stimme u. Sprache. Hermanns Handb. d. Physiol. I, B. 2. 1879.  
 — (2) Verhandl. d. Gesellsch. deutscher Naturforscher u. Ärzte. B. 2. 1891.  
 — (3) Stimme u. Sprache. Ergebnisse d. Physiol. v. ASCHER u. SPIRO I, B. 2. 1902.
- GUTTMANN, A. Die neuen Entdeckungen der sog. klanglichen Konstanten in der Musik. Katzensteins Arch. f. exp. u. klin. Phonetik B. 1. 1914.
- GUTZMANN, H. (1) Untersuchungen über das Wesen der Nasalität. Arch. f. Laryngol. u. Rhinol. B. 27. 1913.  
 — (2) Die Tonhöhe der Sprechstimme. Monatsschr. f. Sprachheilk. J. 16. 1906.  
 — (3) Untersuchungen über die Grenzen d. sprachlichen Perzeptionen. Z. f. klin. Medizin B. 60. 1907.  
 — (4) Stimme u. Sprache ohne Kehlkopf. Z. f. Laryngol. B. 1. 1909.  
 — (5) Physiologie d. Stimme u. Sprache. 1909. (2. A. wird demnächst erscheinen.)
- HEINRICH, TR. Studien über die deutsche Gesangssprache. 1905.
- HELLWAG, C. De formatione loquela. Diss. Tübingen 1781. (Dazu vgl. VIËTOR: „Aus HELLWAGS Nachlass“. Phon. Stud. B. 2. 1889.)
- HELMHOLTZ, H. Die Lehre von den Tonempfindungen. 4. Aufl. 1877 (1. Aufl. 1863). Kurze Mitteilungen bereits 1857, 1859, 1860; abgedr. in „Gesammelte wissenschaftl. Abhandlungen“ B. 1.
- HENSEN, V. (1) Über d. Schrift von Schallbewegungen. Z. f. Biol. B. 23. 1887.  
 — (2) Die Empfindungsarten des Schalles. Pflüg. Arch. B. 119. 1907.  
 — (3) Die Harmonie in den Vokalen. Z. f. Biol. B. 28. 1891.
- HERMANN, L. „Phonographische Untersuchungen“ u. weitere Abhandlungen in Pflüg. Arch. von B. 45 (1889) bis B. 150 (1913). Hier kommen besonders in Betracht B. 47, 48, 53, 58, 61, 83, 91, 139, 141, 146. Auf diese Bändenummern wird in den Zitaten verwiesen. Vgl. auch HERMANN'S Berichte in d. „Naturwissenschaftl. Rundschau“ J. 4 u. 5.
- HERRMANN-GOLDAP, E. Über die Klangfarbe einiger Orchesterinstrumente. Diss. Königsberg 1908. Auch in d. Festschrift für L. HERMANN 1908. Auszugsweise in Ann. d. Ph. (4) B. 23. 1907.
- HORNPOSTEL, E. M. v. (1) Beobachtungen über ein- und zweiohriges Hören. Psychol. Forschung B. 4. 1923.  
 — (2) Psychologie d. Gehörserscheinungen. Handb. d. normalen u. pathol. Physiologie, her. v. BETHE u. a. B. 11. 1926.  
 — (3) Räumliches Hören. Dasselbst.
- HUBER, KURT. Vokalmischungen und Qualitätensystem d. Vokale. Ber. üb. d. 9. Kongreß d. Ges. f. exp. Psychol. 1926.
- ISSERLIN, M. (1) Psychologisch-phonet. Untersuchungen. Zentralbl. f. d. ges. Neurol. u. Psychiatr. B. 26. 1921.  
 — (2) Psychol.-phonet. Untersuchungen I. Z. f. allg. Psychiatrie B. 75. 1919; II. Z. f. d. ges. Neurologie u. Psychiatrie B. 94. 1924.
- JAENSCH, E. R. (1) Die Natur der menschlichen Sprachlaute. Z. f. Ps. II, B. 47. 1913.  
 — (2) Untersuchungen zur Tonpsychologie. 6. Kongr. f. exp. Psychol. 1914.  
 — (3) u. ROTHE, G. Die psycholog. Akustik d. Sprachlaute. Z. f. Ps. I, B. 97. 1925. (Zuerst in der Festschrift für PIPPING 1924.)
- JENKING, FL. and EWING, A. On the harmonic Analysis of Certain Vowel Sounds. Proc. R. Soc. of Edinburgh B. 28. 1879.
- JESPERSEN, O. Phonetische Grundfragen. 1904.
- KATZENSTEIN, J. (1) Über Probleme u. Fortschritte etc. P.-Sch. B. 3. 1909.  
 — (2) Über Brust-, Mittel- und Falsettstimme. P.-Sch. B. 4. 1911.



- KEMPELEN, W. v. Mechanismus d. menschlichen Sprache nebst d. Beschreibung seiner sprechenden Maschine. 1791.
- KOEHLER, W. (1) Akustische Untersuchungen. I. Z. f. Ps. I, B. 54 (1910), II. Z. f. Ps. B. 58 (1910), III. u. IV. (Vorläufige Mitt.) Z. f. Ps. B. 64 (1912), III. (ausführl.) Z. f. Ps. B. 72 (1915). I. u. II. auch in Beitr. H. 4 u. 6.
- (2) Psychologische Beiträge zur Phonetik. KATZENSTEINS Arch. f. exp. u. klin. Phonetik B 1. 1913.
- (3) „Tonpsychologie“ im Handb. d. Neurologie d. Ohres v. ALEXANDER u. MARBURG. 1923.
- KOENIG, R. Quelques expériences d'acoustique. 1882.
- KRÖNIG, A. Notiz über [geflüsterte] Vokallaute. Ann. d. Ph. (2) B. 157. 1876.
- KRUEGER, F. Beziehungen d. experimentellen Phonetik zur Psychologie. 2. Kongr. f. exp. Psych. 1907.
- TER KUILE, E. Neues zur Vokal- und zur Registerfrage. Pflüg. Arch. B. 153. 1913.
- LACHMUND, H. Vokal und Ton. Z. f. Ps. I B. 88. 1921.
- LAHR, J. Die GRASSMANNsche Vokaltheorie im Lichte des Experiments. Diss. Jena 1885.
- LAMPERT, H. Die Mängel d. bisherigen Hörprüfung mit d. Sprache u. ein Versuch zur Besserung. Z. f. Hals-, Nasen- u. Ohrenheilk. B. 5. 1923.
- LEWIN, KURT. (1) Über einen Apparat zur Messung von Tonintensitäten. Psychol. Forsch. B. 2. 1922.
- (2) Über den Einfluß von Interferenzröhren auf die Intensität obertonfreier Töne. Dasselbst.
- LLOYD, R. J. (1) „Speech Sounds“ u. weitere Abhandlungen in VIÉTORS Z. „Phonetische Studien“ B. 3, 4, 5, 11, 13. 1890—1899. (Von Bd. 7 ab erschienen die „Phonet. St.“ als Beiblatt d. Z. „Die neueren Sprachen“, aber mit fortlaufender Nummerierung.)
- (2) The Genesis of Vowels. J. of Anat. and Physiol. Bd. 31. 1897.
- (3) Consonant Sounds. Proc. Roy. Soc. Edinburgh. B. 22. 1898.
- MARTENS, W. Über d. Verhalten von Vokalen u. Diphthongen in gesprochenen Worten. Unters. m. d. Sprachzeichner. Z. f. Biol. B. 25. 1889.
- MEISSNER, G. Klangaufnahmen an Blasinstrumenten, eine Grundlage für das Verständnis d. menschl. Stimme. Her. v. R. WACHSMUTH. Pflüg. Arch. B. 116. 1907.
- MERKEL, C. L. (1) Anatomie u. Physiol. d. menschl. Stimme (Anthropophonik). 1857.
- (2) Physiologie d. menschl. Sprache (Laetik). 1866.
- MEYER, ULFILAS. Über die Frequenz d. Fernsprechströme. Mitteil. aus d. telegraphentechn. Reichsamt. B. 9. Zuerst in „Telegr.- u. Fernsprechtechnik“ J. 10. 1921.
- MICHAELIS, G. (1) Zur Lehre von d. Klängen d. Konsonanten. 1879.
- (2) Über die Anordnung der Vokale. Herrigs Archiv B. 64 u. 65. 1881.
- MILLER, D. C. The Science of Musical Sounds. 1916. (Die 2. A. 1922 ist fast unverändert; die hier zitierten Seitenzahlen gelten auch für diese.)
- MUSEHOLD, A. Allgemeine Akustik u. Mechanik d. menschl. Stimmorgans. 1913.
- NADOLECZNY, M. Physiologie d. Stimme u. Sprache. Im Handb. d. Hals-, Nasen-, Ohrenheilk. v. DENKER u. KAHLER. 1925.
- NAGEL, W. Physiologie d. Stimmwerkzeuge. Nagels Handb. d. Physiol. d. Menschen B. 4. 1906.

- OHM, G. S. Über die Definition des Tones. *Ann. d. Ph.* (2) 1843. Noch ein paar Worte über die Definition des Tones. Das. 1844. Abgedr. in „Gesammelte Abhandlungen“, her. v. LOMMEL. 1892.
- PAGET, R. A. S. (1) Nature of Vowel Sounds. *Nature B.* 109, Nr. 2733. 1922. Weiteres B. 111, Nr. 2775. 1923. *Proc. Roy. Soc. London B.* 36, Nr. 1. 1923. *Journal Inst. Electr. Eng. B.* 62, Nr. 335. 1924.
- (2) Nature and Artificial Production of Consonant Sounds. *Proc. Roy. Soc. London B.* 106. 1924.
- PAPALE, R. (1) Il metodo interferenziale acustico per l'analisi della voce umana. *Arch. Ital. di Otologia* 1922. Suppl. 3.
- (2) Studi per la determinazione di una scala ototipica italiana. Dasselbst 1923. Suppl. 3.
- PIELKE, W. (1) Über „offen“ und „gedeckt“ gesungene Vokale (nebst Bemerkungen von H. GUTZMANN). *P.-Sch. B.* 5. 1911.
- (2) Über d. Ausgleich d. Stimmbruchs durch die sog. Deckung. 3. Intern. Laryngo-Rhinologen-Kongr. 1911, B. 2.
- (3) Über die Register d. menschl. Stimme. *Z. d. Internat. Musikgesellsch.* J. 12. 1910—1911.
- PIPPING, H. (1) Zur Klangfarbe der gesungenen Vokale. *Z. f. Biologie B.* 27. 1890; B. 31. 1895.
- (2) Besprechung von LLOYDS Arbeiten. *Die neueren Sprachen.* B. 15. 1893.
- (3) Über die Theorie der Vokale. *Acta soc. scient. Fennicae B.* 20, Nr. 11. 1895.
- (4) Zur Phonetik d. finnischen Sprache. *Mém. de la Soc. finno-ougrienne B.* 14. 1899.
- RÉTHI, L. Phonographische Untersuchungen d. Konsonanten. *Sitz.-Ber. d. Wiener Akad. d. Wiss. Math.-naturwiss. Kl. B.* 121. 1912.
- REYHER, S. *Mathesis Mosaica.* 1679. (S. 432ff.)
- RICH, G. J. A Study of Tonal Attributes. *Americ. Journ. of Psych.* B. 30. 1919.
- RITZ, J. Unters. üb. d. Zusammensetzung d. Klänge d. Streichinstrumente. *Diss. Erlangen.* 1883.
- ROUSSELOT, P.-J. (1) Principes de phonétique expérimentale. I. 1897—1901, II. 1901—1908.
- (2) Classification des voyelles orales. *Revue de Phonétique B.* 1. 1911.
- RUBENS, H. u. KRIGAR-MENZEL, O. Flammenröhre f. akustische Beobachtungen. *Ann. d. Ph.* (4) B. 17. 1905.
- SACHS, C. *Handbuch d. Musikinstrumente.* 1920.
- SAMOJLOFF, A. Zur Vokalfrage. *Pflüg. Arch. B.* 78. 1899.
- SAUBERSCHWARZ, E. Interferenzversuche mit Vokalklängen. *Pflüg. Arch. B.* 61. 1895.
- SCHAEFER, K. L. (1) Zur Lehre von den sog. Unterbrechungstönen. *Ann. d. Ph.* (4) B. 13. 1904.
- (2) Der Gehörsinn. In Nagels *Handb. d. Physiol.* III, 2. 1905.
- (3) Tabellen d. Schallgeschwindigkeit u. Tonwellenlängen in Luft bei versch. Temperaturen. *P.-Sch. B.* 1. 1908.
- (4) Untersuchungsmethodik d. akustischen Funktionen d. Ohres. *Tigerstedts Handb. d. physiol. Methodik B.* III, 1. Hälfte. 1914.
- u. ABRAHAM, O. Studien über Unterbrechungstöne. *Pflüg. Arch. B.* 83, 85, 88. 1901. *Ann. d. Ph.* (4) B. 13. SCHAEFER auch in *Beitr. H.* 6.
- SCHILLING, R. Die Deckung des gesungenen Tones im Röntgenbilde. *KATZENSTEINS Arch. f. exp. u. klin. Phonetik B.* 1. 1913.
- Stumpf, Sprachlaute.

- SCHOLE, H. Über d. Zusammensetzung d. Vokale U, O, A. Diss. Kiel. Arch. f. d. ges. Psychol. B. 37. 1917.
- SCHULZE, F. A. (1) Einige neue Methoden etc. (obere Hörgrenze). Ann. d. Ph. (4) B. 24. 1907.
- (2) Die obere Hörgrenze u. ihre exakte Bestimmung. P.-Sch. B. 1. 1908.
- SCRIPTURE, E. W. (1) Researches in Experimental Phonetics. 1906.
- (2) Untersuchungen über die Vokale. Diss. München. 1906.
- SIEVERS, E. Grundzüge der Phonetik. 5. Aufl. 1901.
- STEFANINI, A. (1) Nuove studi sulle vocali. Arch. Ital. di Otologia B. 25. 1914.
- (2) Nuove studi sulle vocali. Dasselbst B. 27. 1916; N. Cimento (6) B. 11. 1916.
- STEVANI, R. Le note armoniche caratteristiche delle vocali italiani. Arch. Ital. di Otologia B. 15. 1904.
- STEWART, J. A. An Electrical Analogue of the Vowel Organs. Nature B. 110. 1922.
- STORM, J. Englische Philologie 2. A. I, 1: Phonetik u. Aussprache. 1892.
- STRUYCKEN, H. J. L. (1) 3 Abhandlungen über die obere Hörgrenze in P.-Sch. B. 3, 5, 6 (1910, 1911, 1913).
- (2) Das Verhältnis der Klangstärke von reinen Stimmgabeln u. von Pfeifentönen. P.-Sch. B. 10. 1918.
- (3) Die Vokalsynthese. In d. Festschrift für PIPPING. 1924.
- STUMPF, C. (1) Tonpsychologie I. 1883; II. 1890.
- (2) Musikpsychologie in England. Vierteljschr. f. Musikwissensch. B. 2. 1886.
- (3) Über Ermittlung von Obertönen. Ann. d. Ph. (3) B. 57. 1896.
- (4) Über die Bestimmung hoher Schwingungszahlen durch Differenztöne. Dasselbst B. 68. 1899.
- (5) Beobachtungen über subjektive Töne u. üb. Doppelthören. Z. f. Ps. B. 21. 1899; Beitr. 3. H.
- (6) Über zusammengesetzte Wellenformen. Mit 2 Fig.-Tafeln v. K. L. und M. SCHAEFER. Z. f. Ps. B. 39. 1905; Beitr. 4. H.
- (7) Beobachtungen über Kombinationstöne. Z. f. Ps. B. 55. 1910; Beitr. 5. H.
- (8) Über neuere Untersuchungen zur Tonlehre. 6. Kongr. d. Ges. f. exp. Psych. 1914; Beitr. 8. H.
- (9) Die Struktur der Vokale. Sitz.-Ber. d. Berliner Akad. d. Wiss. 1918.
- (10) Trompete u. Flöte. In d. Festschrift f. H. KRETZSCHMAR. 1918.
- (11) Zur Analyse geflüsterter Vokale. P.-Sch. B. 12. 1919.
- (12) Zur Analyse der Konsonanten. Dasselbst B. 17. 1921.
- (13) Veränderungen des Sprachverständnisses bei abwärts fortschreitender Vernichtung d. Gehörsempfindungen. Dasselbst.
- (14) Über die Tonlage der Konsonanten u. die f. d. Sprachverständnis entscheidende Gegend des Tonreiches. Sitz.-Ber. d. Berliner Akad. d. Wiss. 1921.
- (15) Singen u. Sprechen. Z. f. Ps. I, B. 94. 1923; Beitr. 9. H.
- (16) Phonetik u. Ohrenheilkunde. P.-Sch. B. 22. 1925.
- (17) u. v. ALLESCH, G. Über d. Einfluß d. Röhrenweite auf d. Auslöschung hoher Töne durch Interferenzröhren. P.-Sch. B. 17. 1921.
- TECHMER, F. Phonetik. 1880.
- THAUSING, M. Das natürliche Lautsystem d. menschl. Sprache. 1863.
- THOMSON, AL. (Odessa) (1) Die Eigentöne der Sprachlaute u. ihre prakt. Verwendung. Indogerman. Forsch. B. 24. 1909.

- THOMSON, AL. (Odessa) (2) Besprechung von ŠČERBA im Arch. f. slav. Philologie B. 34. 1913.  
 — (3) Phonet. Betrachtungen zur russischen Aussprache. Z. f. slav. Philol. B. 2. 1925.
- TRAUTMANN, M. Die Sprachlaute im Allgemeinen u. d. Laute des Engl., Französ. u. Deutschen im Besonderen. 1884.
- TRENDELENBURG, FERD. Objektive Klangaufzeichnung mittels d. Kondensatormikrophons. I und II. Wiss. Veröff. aus d. Siemens-Konzern B. 3. 1924 B. 4. 1925.
- VIETOR, W. (1) Haben die Vokale feste Resonanzhöhen? Z. Phonet. Studien B. 2. 1889.  
 — (2) Elemente d. Phonetik. 6. Aufl. 1915.
- WAETZMANN, E. (1) Resonanztheorie d. Hörens. 1912.  
 — (2) Ton, Klang u. sekundäre Klangerscheinungen. Handb. d. norm. u. path. Physiol., her. v. BETHE u. a. B. 11. 1926.  
 — (3) Hörtheorien. Dasselbst.
- WAGNER, K. W. (1) Über d. Frequenz d. Fernsprechströme. Physikal. Z. J. 11. 1910.  
 — (2) Spulen- und Kondensatorleitungen. Arch. f. Elektrotechnik Bd. 8. 1919 (eingereicht 1915).  
 — (3) Das Fernsprechen auf weite Entfernung. Elektrotechn. Z. 1924, H. 1.  
 — (4) Der Frequenzbereich von Sprache u. Musik. Dasselbst H. 19.
- WEGEL, R. L. and LANE, C. E. The Auditory Masking of one Pure Tone by another and its Probable Relation to the Dynamics of the Inner Ear. Physic. Rev. (2) B. 23. 1924.
- WEISS, O. (1) Die photographische Registrierung d. geflüsterten Vokale u. d. Konsonanten Sch u. S. Zentralbl. f. Physiol. B. 21. 1907.  
 — (2) Die Kurven d. geflüsterten u. leise gesungenen Vokale u. d. Konson. Sch u. S. Pflüg. Arch. B. 142. 1911.  
 — (3) u. SOKOLOWSKY, R. Die physikalischen Grundlagen d. Geräuschwahrnehmung. Pflüg. Arch. B. 180. 1920.  
 — (4) Über d. Entstehung d. Vokale. I und II. Katzensteins Arch. f. exp. u. klin. Phonet. Bd. 1. 1913.
- WESENDONK, K. v. (1) Über d. Synthese d. Vokale aus einfachen Tönen. Physikal. Z. J. 10. 1909.  
 — (2) Über Vokalklänge. Ber. d. deutschen physikal. Ges. J. 19. 1917.
- WETHLO, F. Versuche mit Polsterpfeifen. P.-Sch. B. 6. 1913.
- WHEATSTONE, CH. Besprechung von WILLIS (unterzeichnet C. W.), London and Westminster Review (6) Vol. 28, Oct. 1837. Bandjahr 1838.
- WIEN, M. (1) Die akustischen u. elektrischen Konstanten d. Telephons. Ann. d. Ph. (4) B. 4. 1901.  
 — (2) Über Telephonplatten mit hohen Eigentönen. Dasselbst B. 18. 1905.
- WILLIS, R. Über Vokaltöne u. Zungenpfeifen. Ann. d. Ph. B. 24. 1832 (übers. aus: Transact. Cambridge Philos. Society. Vol. 3. 1829).
- WINTELER, J. Die Kerenzer Mundart. 1876.
- ZAHN, v. Über die akustische Analyse von Vokalklängen. Progr. d. Thomaschule in Leipzig. 1871.
- ZWAARDEMAKER, H. (1) Das presbyakusische Gesetz. Z. f. Ohrenheilk. B. 24. 1893.  
 — (2) D. Einfluß d. Schallintensität auf d. Lage d. oberen Tongrenze. Das.  
 — (3) Multiple Resonantie. Tijdschr. v. Geneeskunde. 1913.



