

## Singen und Sprechen.

Von  
C. STUMPF.

### Inhalt.

	Seite
§ 1. Unterschied im Tonmaterial . . . . .	38
§ 2. Unterschied in der Einstellung . . . . .	43
§ 3. Übergänge und Mischungen . . . . .	45
a) Sprachgesang . . . . .	46
b) Singendes Sprechen . . . . .	50
§ 4. W. KÖHLERS Anschauung . . . . .	54
§ 5. Zur Phänomenologie der stetigen Tonveränderungen . . . . .	62
§ 6. Rückblick auf KÖHLERS These . . . . .	69
§ 7. Ergebnis . . . . .	73

### § 1. Unterschied im Tonmaterial.

Unser Sprachorgan stellt ein Instrument dar, das sich durch die Vereinigung dreier Eigenschaften auszeichnet: 1. kann man ihm bei gleichbleibender Tonhöhe verschiedene Resonanzeinstellung erteilen und bei gleicher Resonanzeinstellung verschiedene Tonhöhen damit produzieren (Vokale), 2. kann es außer Tönen auch Geräusche hervorbringen (Flüstervokale und Konsonanten), 3. kann es stetige Übergänge ebenso wie sprunghafte erzeugen.

Diese Eigenschaften sind natürlich nicht ausschließlich der menschlichen Stimme eigen. Stetige Übergänge z. B. kann man neben diskreten auch leicht mit manchen anderen Instrumenten herstellen, obgleich sie in unserer Musik bisher nicht offiziell und in größerem Mafsstabe verwendet werden. Aber

---

<sup>1</sup> Aus der im Manuskript überreichten Festschrift zum 70. Geburtstage MAX FRIEDLÄNDERS, hier in stark erweiterter Form.

für die Funktionen des menschlichen Kehlkopfes sind diese drei Eigenschaften in Verbindung miteinander grundwesentlich und charakteristisch. Sie sind sozusagen sein tägliches Brot und kommen in dieser Vereinigung und Vollkommenheit und in der Leichtigkeit der Umstellung aus der einen in die andere Funktion sonst nirgends vor.

Es soll auch nicht gesagt sein, daß unser Sprachorgan nicht noch anderer Künste fähig wäre. So können wir z. B. den gleichen (gleichbezeichneten) Vokal mit Brust- oder mit Kopfstimme singen, wobei phänomenal verschiedene Klangfarben entstehen (Register). Aber wesentlicher und ganz unentbehrlich für ein normales Stimmorgan sind die drei obigen Eigenschaften.

Nur vom stimmhaften Sprechen, nicht vom Flüstern soll im folgendem die Rede sein. Beim Flüstern sind stetige Übergänge, wie überhaupt Tonhöhenveränderungen, nur möglich unter gleichzeitiger qualitativer Veränderung der geflüsterten Laute selbst. Man kann z. B. ganz wohl eine stetige Tonbewegung erzielen, wenn man vom geflüsterten A allmählich zum O oder zum E übergeht, aber nicht auf dem A selbst, solange dessen Nuancierung genau dieselbe bleibt. Selbst bei einer geringen Höhenveränderung des Flüsterlautes wird seine Vokalnuanze mitgeändert.

Von den obigen drei Unterschieden hängt nun der letzte offenbar mit dem des Sprechens und Singens zusammen. Beim Sprechen benützen wir außer sprunghaften fortwährend stetige Tonveränderungen, beim Singen — prinzipiell wenigstens — nur sprunghafte. Beim Sprechen sind gerade die stetigen Veränderungen, auch die ganz geringfügigen, ein wichtiges Ausdrucksmittel, sie beseelen die Sprache. Beim Singen dagegen, wie in der Musik überhaupt (abgesehen von den primitiven Formen), dürfen sie nur ausnahmsweise vorkommen.

Dieses Unterscheidungsmerkmal findet man denn auch bereits im Altertum bei ARISTOXENUS, PTOLEMAEUS, BACCHIUS, GAUDENTIUS, und so im allgemeinen bis in die neuere Zeit.<sup>1</sup>

<sup>1</sup> Eine Ausnahme macht unter den neueren Musikforschern nur HUGO RIEMANN. Er erklärt: „Letzten Endes ist das Prinzip des Melodischen die stetige Veränderung der Tonhöhe, nicht die stufenweise.“ („Wie hören wir Musik?“ S. 7. „Die Elemente der musikalischen Ästhetik.“ S. 40.) RIEMANN meint, daß, wenn auch die Töne äußerlich diskret

Auch Forscher, welche die Musik aus dem erregten Sprechen herleiten wollen, dürften nicht leugnen, daß für die entwickelte Musik, wie sie seit Jahrtausenden in Asien und Europa geübt wird, feste Intonation unentbehrlich ist und sich immer mehr durchgesetzt hat.<sup>1</sup> Sie hat ja auch zur Einführung der Notierung auf Linien geführt. Aufzeichnungen der „Sprachmelodie“ in unserem Notensystem, so oft sie auch versucht wurden, behalten immer etwas Unnatürliches. Die stetigen Tonveränderungen der Sprache lassen sich räumlich nur in Form von Kurven wiedergeben, wie sie auf objektivem Wege z. B. SCRIPTURE durch Vergrößerung grammophonischer Kurven<sup>2</sup> oder EGGERT mit der MARBESCHEN Rufsmethode<sup>3</sup> gewonnen haben.

Vergleicht man die meist sehr kurzen und trotzdem aufeinanderfolgen, doch in der Phantasie ein stetiger Übergang stattfinde. Er geht so weit, es als eine „Manier mancher Virtuosen von großem Ton“ zu bezeichnen, wenn sie „mit Präzision jeden folgenden Ton sofort fest einsetzen“. Das Portament nennt er „die konkreteste Versinnlichung des melodischen Prinzips“. In der zweitgenannten Schrift allerdings wird es bei Streichern und Sängern „eine plumpe Enthüllung der Natur und eine allzu drastische Nachhilfe für die Phantasie“ genannt.

Psychologie war ebensowenig wie Akustik die Stärke des sonst so verdienten Forschers. Hier scheint er mir aber selbst mit der Logik etwas in Konflikt zu kommen.

<sup>1</sup> Vgl. die längeren Ausführungen in meiner Abhandlung „Musikpsychologie in England“, Viertelj.-Schr. f. Musikwissenschaft 1885, S. 264 ff. und in „Anfänge der Musik“ 1911.

J. J. ROUSSEAU, ein Hauptvertreter des sprachlichen Ursprunges der Musik, sagt doch in seinem „Dictionnaire de Musique“, art. Voix: „Je penserais que le vrai caractère distinctif de la Voix de Chant est de former les sons appréciables dont on peut prendre ou sentir l'unisson, et de passer de l'un à l'autre par des Intervalles harmoniques et commensurables; au lieu que dans la voix parlante, ou les sons ne sont pas assez soutenus, et, pour ainsi dire, assez unis pour pouvoir être appréciés, ou les Intervalles qui les séparant ne sont point assez harmoniques, ni leurs rapports assez simples.“

Man kann den Unterschied nicht präziser ausdrücken.

<sup>2</sup> E. W. SCRIPTURE, *Researches in Experimental Phonetics* 1906, S. 63 ff. Abbildungen nach derselben Methode auch in m. „Anf. d. Mus.“, S. 16 ff.

<sup>3</sup> B. EGGERT, *Untersuchungen über Sprachmelodie. Zeitschr. f. Psychol.* 49 (1908), S. 230.

schwankenden, vielfach auch ihrer Höhe nach äußerst un-  
deutlichen Laute des gewöhnlichen, nicht zu stark affektbetonten  
Sprechens mit den langgehaltenen, vollen, kultivierten Tönen einer  
schönen Frauen- oder Männerstimme in einer getragenen Kanti-  
lene, aber auch selbst mit den kurzen in einem Rezitativ oder einer  
gut gesungenen Koloratur, so hat man zwei Welten vor sich,  
die zunächst so verschieden erscheinen können wie Töne und  
Geräusche oder gar Töne und Gerüche. Zwar gibt es unter  
Umständen auch beim Sprechen langgedehnte Silben; aber in  
diesen Fällen stellt sich regelmäÙig zugleich eine vergrößerte  
Schwankung der Tonhöhe ein, und diese ist so wesentlich, daß  
durch sie der Sinn des Ausdruckes vollständig determiniert  
sein kann. Man spricht ein gedehntes „So!“ aufsteigend bis  
zum Umfang etwa einer Duodezime, wenn man die Ver-  
wunderung über etwas ganz Unerwartetes ausdrückt, absteigend  
aber in gleichem Umfang, wenn man etwa jemand auf einer  
Spitzbüberei ertappt hat. In beiden Fällen kann mehr darin  
liegen als in ganzen Sätzen. So ausdrucksvoll sind hier gerade  
die stetigen, in der Musik grundsätzlich unerlaubten Ver-  
änderungen.

Auch ein Unterschied in der Klangfarbe ist nicht zu  
verkennen. Worin er besteht, ist freilich nicht ganz leicht zu  
sagen. HELMHOLTZ meint (Tonempf. <sup>4</sup> S. 183), daß beim  
Sprechen wahrscheinlich durch stärkeren Druck der Stimm-  
bänder gegeneinander, wobei sie aufschlagende Zungen bilden,  
eine knarrende Klangfarbe, d. h. stärkere Obertöne als beim  
Singen gegeben werden. Ist dies richtig (und es hat viel für  
sich), so folgt zugleich, daß die Formanten der Vokale, also  
die unterscheidenden Obertonpartien, kräftiger hervortreten,  
daß aber auch alle Vokale außer U durchschnittlich etwas  
heller werden müssen als beim Singen. Bei diesem kommt  
es in erster Linie auf Deutlichkeit der intendierten Grundtöne  
an; da darf kein Ton, auch nicht der leiseste, so „unter den  
Tisch fallen“, wie es beim Sprechen ganz gewöhnlich geschieht,  
ohne daß darum die Vokalitäten unerkennbar würden.

Zum Teil beruht der Unterschied in der Klangfarbe aber  
auch auf den der Sprechstimme mehr oder weniger beige-  
mischten Geräuschen, mögen sie auf dem von HELMHOLTZ an-  
gedeuteten Wege oder sonstwie zustande kommen.

Durch diesen Unterschied in der Klangfarbe und Geräuschbeimischung ist die geringere Deutlichkeit der Tonhöhen beim Sprechen offenbar wesentlich mitbedingt.

Aber nicht blofs in den einzelnen Klängen, die hier und dort zur Anwendung kommen, liegen wesentliche Unterschiede, sondern auch in ihrer Aufeinanderfolge. Für das Sprechen ist mindestens ebenso wesentlich wie das Schwanken der Tonhöhen der Gebrauch undeutlicher und schwankender, nur etwa in der Richtung bei gleichen Anlässen gleichbleibender Tonabstände, für das Singen aber der Gebrauch fester und von einer Tonlage zur anderen transponierbarer Intervalle. Wer unter Vermeidung stetiger Tonschwankungen jede Silbe auf einem bestimmten Tone spräche, die kurzen, „unbetonten“, tief abgestimmten Silben nur etwas undeutlich in ihrer Intonation, dabei die Intervalle unserer Tonleitern geflissentlich umginge und sich ganz willkürlich bald da bald dort im Tongebiet bei den akzentuierten Silben niederliese, der würde uns immer noch im wesentlichen zu sprechen scheinen. In der Musik dagegen, also auch beim Singen, sind feste transponierbare Intervalle in bestimmter Auswahl (verschieden je nach den Musiksystemen) durchaus notwendig. An sie ist die Gefühlswirkung der Musik in so hohem Mafse geknüpft, dafs diese bei ganz minimalen Abweichungen, die für das Sprechen gar nicht in Betracht kommen, schon ins Gegenteil umschlagen kann. Unreinheit ist ja schlimmer als jede noch so scharfe unaufgelöste Dissonanz, wenigstens für das feiner empfindende Ohr, wie es sich mit der Entwicklung der Musik ausgebildet hat.

Mit diesem Merkmal hängt nun offenbar das der Vermeidung und der Anwendung stetiger Tonveränderungen zusammen, ja das erstere ist sogar das tieferliegende. Denn eben weil feste Intervalle nötig sind, sind auch feste Töne nötig.

Wenn wir also sagen: beim Singen produziere der Kehlkopf Tonfolgen aus festen Tönen und Intervallen, beim Sprechen aber neben einzelnen diskreten vielfältig stetige Tonveränderungen, sowie Tonabstände von schwankender Abmessung, so dürfte die Unterscheidung nach der rein lautlichen Seite hin den Tatsachen entsprechen. Man kann unter Voraussetzung dieser Erläuterungen auch kurz sagen: Singen ist gebundene,

Sprechen freie Tongebung. Dafs auch in rhythmischer Hinsicht beim Sprechen, selbst dem in „gebundener Rede“, gröfsere Freiheit als beim Singen besteht, kann darunter mitverstanden werden.

## § 2. Unterschied in der Einstellung.

Mit alledem soll natürlich weder das Singen noch das Sprechen definiert, es sollen nur die beiden Tätigkeiten voneinander unterschieden werden, und dies auch nur nach der lautlichen Seite. Ohne Zweifel ist die ganze Intention oder Einstellung des Singenden eine andere als die des Sprechenden, von dieser so verschieden, wie eben die künstlerische, mehr auf die Gefühlswirkung abzielende, und die theoretische oder praktische, vorwiegend auf Mitteilung von Tatsachen und Gedanken oder auf Handlungen abzielende Geistesverfassung verschieden sind. Aber diese allgemiesten Einstellungen sind hier nicht das Entscheidende. Denn der Evangelist in BACHS Passionen und der Prediger auf der Kanzel berichten beide über die Tatsachen der Leidensgeschichte und wollen beide damit rühren. Aber der eine singt, der andere spricht. Entscheidend sind zuletzt speziellere Einstellungen: dem einen schweben in Verbindung mit den nämlichen Tatsachen und Gefühlswirkungen Tonhöhen, Intervalle und melodische Wendungen vor (guten Sängern oft auch bestimmte Modifikationen der Klangfarbe, des Einsatzes usw.), dem anderen nur etwa die Deutlichkeit und Eindringlichkeit des Vortrags, nicht aber die spezifisch-musikalischen Eigenschaften des Lautmaterials. Geht die Intention darüber hinaus auf gröfsere Zusammenhänge, so sind auch diese entsprechend verschieden: der musikalische Sinn und die Gesamtstruktur der Melodie bestimmt die innere Haltung des Sängers, der logische Sinn und die Gesamtform des Satzes, ja bis zu einem gewissen Grade auch der Zusammenhang, in dem er steht (Gegensatz, Beweis usw.), bilden den Bewußtseinsrahmen des Sprechers.

Am schärfsten ist der Unterschied natürlich, wenn wir die Instrumentalmusik („reine, absolute Musik“) dem gewöhnlichen Prosasprechen gegenüberstellen. Hier die verschiedenen Systeme artikulierter, mit Vorstellungen und Begriffen verknüpfter Worte

und ihre Abfolge in den jeweiligen grammatischen Formen, dort bestimmte Intervallsysteme, Melodie- und Harmonieformen, an die zufolge einer langen Entwicklung mannigfache, fein abgestufte, eigenartige Gefühle geknüpft sind. Diese Unterschiede sind nicht blofs objektiv vorhanden, sondern sind auch im Ausführenden mindestens als unmittelbar auslösungsbereite Dispositionen, zuweilen und teilweise aber auch als aktuelle Vorstellungsmassen wirksam. Es ist kein „psychologischer Fehlschluss“, wenn wir die inneren Einstellungen mit dadurch definiert sein lassen.

Aber es besteht in dieser Hinsicht noch ein Gradunterschied. Bei dem Musizierenden pflegt die Ausführung doch mehr im Vordergrund zu stehen als beim Sprechenden. Dieser widmet dem Technischen nur in besonderen Fällen, wie beim Einüben einer Rede oder Deklamation, nähere Beachtung, und dann sind es andere Eigenschaften, auf die er merkt. Der Musizierende, und speziell der Sänger, muß fortwährend die Reinheit der Intonation, die dynamischen Schattierungen und sonstigen Vortragsbedingungen unter Kontrolle halten. Auch der gemeine Mann, der singt, wie ihm der Schnabel gewachsen ist, kann nicht umhin, dabei mehr als beim Sprechen die Bewegungen seines Stimmorgans und deren Produkte zu beachten, wenn die Sache nicht sehr übel verlaufen soll.

Über den Mechanismus des Nachsingens eines gehörten (oder vorgestellten) Tones hat W. KÖHLER beachtenswerte, von den bisherigen abweichende Gedanken entwickelt (Psycholog. Beiträge zur Phonetik in *Katzensteins Archiv f. experim. u. klin. Phonetik* I S. 19 ff. 1913). Danach fände nicht eine feste Assoziation zwischen der Tonvorstellung und einer bestimmten Muskelinnervation statt, sondern eine direkte Anpassung der Stimmlippenschwingungen an die Frequenz der Erregungen in der Schnecke (bzw. im Gehirn). Dem Bewußtsein bliebe nur die Aufgabe einer möglichst strengen Fixation des gehörten (vorgestellten) Tones. Ich halte dies für sehr wahrscheinlich, würde allerdings den Einfluß der Übung, sowie einer individuellen Anlage stark betonen. Die Kontrolle, von der wir oben im Texte sprechen, betrifft aber nicht blofs das Treffen eines einzelnen Tones, sondern auch das der Intervalle und des gesamten Vortrags. Und hierbei bleibt dem Bewußtsein doch sehr viel zu leisten.

O. ABRAHAM hat beim Singen eines Volksliedes („Deutschland, Deutschland“) bedeutende Abweichungen von den reinen Intervallen selbst von seiten musikalischer und routinierter Sänger festgestellt und

die Richtung dieser Abweichungen in einer gewissen Abhängigkeit von der allgemeinen Melodiebewegung gefunden (Tonometrische Untersuchungen an einem deutschen Volkslied, Psychol. Forschung, Bd. 3, 1923, S. 1—22). Doch ist zu beachten, daß beim Vortrag einer Melodie mit Text und zumal einer so umfangreichen ( $G-c^1$ ) und mit so viel Affekt habituell verknüpften, wie „Deutschland, Deutschland“, die Einstellung nicht dieselbe ist wie bei isolierten Intervallen oder etwa einem aufgelösten Dreiklang auf la—la—la; und besonders wenn noch ausdrücklich die Anweisung zu möglichstem Reinsingen eingeschärft wird. Hier dürften bei kunstmäßig geschulten Sängern, die ihr Stimmorgan so in der Gewalt haben, daß es ihren Intentionen jederzeit genau folgt, doch die Abweichungen von einer festen Stimmung geringer sein (noch geringer wahrscheinlich bei guten Violinisten). Daß diese feste Stimmung mit der „reinen“ zusammenfalle, ist damit noch nicht gesagt, wie sich ja auch bei meinen Versuchen über Beurteilung der Reinheit vorgelegter Intervalle gewisse konstante Abweichungen ergeben haben. Aber diese in Nebeneinflüssen wurzelnden Abweichungen stehen der Überzeugung nicht entgegen, die aus der Entstehung und Struktur unseres Musiksystems fließt, daß die ihm zugrunde liegende und von den Ausführenden grundsätzlich intendierte Stimmung die physikalisch reine ist.

Die Einstellungen der Ausführenden übertragen sich aber *mutatis mutandis* auch auf den Hörenden, soweit dieser durch seine Nationalität und seinen individuellen Bildungsgang auf das Dargebotene vorbereitet ist. Er hört dann heraus, was jener hineinlegt. Wo beim Sprechen zufällig feste Intonationen und reine Intervalle unseres Tonsystems zum Vorschein kommen, bleiben sie als solche unbeachtet, und darum bleibt auch ihre Gefühlswirkung rudimentär oder fällt aus.

### § 3. Übergänge und Mischungen.

Kehren wir nun zu dem hervortretendsten rein akustischen Unterscheidungsmerkmal, dem Gebrauche stetiger gegenüber sprunghaften Tonveränderungen und schwankender gegenüber festen Intervallen, zurück, so liegt es in der Natur der Sache, daß trotz der prinzipiell scharfen Unterscheidung auch Übergänge und Mischungen vorkommen müssen. Das Extrem einer Sprache mit ausschließlich stetigen Tonveränderungen gibt es in der Wirklichkeit nicht, und so können darin beliebige Annäherungen an das Singen Platz greifen. Überdies müssen die nationalen, lokalen und individuellen Gewohnheiten zu gewissen Regelmäßigkeiten des Tonfalles führen, die sich den



festen Intervallen nähern. Umgekehrt sind auch Annäherungen des Singens an das Sprechen denkbar. Wir haben so tatsächlich einerseits den Sprachgesang, namentlich in der aufereuropäischen Musik weit verbreitet, sowie das Parlando und das Portamento unserer Sänger; und wir haben andererseits das singende Sprechen, das uns im Leben so oft begegnet. Von der einen wie der anderen Seite strecken sich Verbindungslinien hinüber.

Fassen wir diese beiden Klassen von Zwischenstufen etwas näher ins Auge.

a) Beim Parlando-Singen läßt sich der Sänger schon rhythmisch freier gehen, so wie es der Prosasprache, dem Sprechen im prägnantesten Sinn, eigen ist. Die Intervalle werden in unseren „Rezitativen“ in der Regel noch festgehalten.<sup>1</sup> Aber bei Naturvölkern finden sich hierin Übergänge, die man fast ebensogut zu der einen wie der anderen Gattung rechnen kann, zumal da auch die Intervalle des Gesanges dort meist nicht so konstant innegehalten werden, wie wir es verlangen. Dennoch wird der Unterschied wohl niemals ganz verwischt und merkt man auch dem Vortragenden eine verschiedene Einstellung an, je nachdem er nur parlando singt oder wirklich spricht.

Ein sehr charakteristisches Parlando wird vielfach auch hervorgebracht durch Festhalten einer einzigen konstanten Tonhöhe. So findet man bei Indianern gelegentlich ganze Gesänge auf einem gleichbleibenden Ton, ohne daß doch — wie in dem bekannten besonders ausdrucksvollen Liede von CORNELIUS<sup>2</sup> —

<sup>1</sup> Bei einer Aufführung von SCHREKERS „Schatzgräber“ in Berlin verfiel der Kanzler im letzten Akte tatsächlich in Sprechen und erzielte damit eine sehr drastische Wirkung. Aber es sind hier in der Partitur bestimmte Noten für jede Silbe vorgeschrieben, die der Komponist, wie er mir sagte, auch innegehalten wissen will, in denen er sich allerdings dem Tonfall der natürlichen Rede möglichst annähert. Der Übergang zum wirklichen Sprechen wird vom Hörer hier doch wie eine *μετάβασις εἰς ἄλλο γένος* empfunden.

<sup>2</sup> Natürlich wäre die Eintönigkeit hier nicht so ausdrucksvoll, wäre sie nicht eben ein Ausnahmefall und nicht durch den Text motiviert. KRETZSCHMAR sagt von diesem Liede (Gesammelte Aufsätze über Musik I, 1910, S. 28): „Es ist Deklamieren, ausdrucksvolles Sprechen und Stammeln,

durch begleitende, geistreich wechselnde Harmonie die Monotonie aufgehoben würde. Auch die erzählenden Lieder der Zentral-Eskimo halten sich nach F. BOAS' Aufschreibungen auf einem Ton, von dem sie nur an akzentuierten Stellen um einen Halbton abweichen.<sup>1</sup> Das Psalmmodieren im Gregorianischen Gesang läßt gleichfalls die Stimme ganze Sätze hindurch auf einem Tone verweilen, aber den Anfang und Schluß machen hier Intervalle der diatonischen Leitern. Ebenso liest RAOUL in den „Hugenotten“ den Brief der Königin auf dem konstanten Ton *c*<sup>1</sup> vor, von dem er nur am Schluß auf die tiefere Oktave herabgeht.<sup>2</sup> Und schon in CIMAROSAS „Matrimonio segreto“ singt der alte Geron in dem Duett mit Geronimo sein in einem halblauten Parlando vorzutragendes Selbstgespräch auf dem konstanten Tone *b*, nur zweimal durch die höhere Oktave unterbrochen.

Diese monotonen Rezitationen bleiben der gewöhnlichen Sprache auch vielfach durch den Verzicht auf feste Metren nahe. Dennoch fallen sie zweifellos unter den Begriff des Singens. Man reduziert nur die Leiterintervalle fast ganz auf die Prime, weil der an die übrigen habituell geknüpfte Gefühlsausdruck vermieden werden soll, wo es sich bloß um einen Bericht u. dgl. handelt. Übrigens läßt sich leicht bestätigen, daß man auch in der Sprache des Lebens beim Vorlesen eines Aktenstücks oder beim Ausrufen einer Meldung unwillkürlich monoton wird, offenbar aus dem Gefühl heraus, daß das Gesagte

---

das einen ganz ergreifenden Eindruck hinterlassen kann.“ Der Subsumtion unter das Sprechen oder gar Stammeln würde ich aber doch widersprechen, und sie ist auch sicherlich von dem geistreichen Musikforscher nur als ein kräftiger Ausdruck zur Beschreibung der künstlerischen Wirkung gemeint.

<sup>1</sup> Notenbeispiele vgl. in „Anf. d. Musik“ S. 93, 157, 170.

<sup>2</sup> Seltsamerweise ist dies gerade eine Frage. Man kann aber aus der interessanten Abhandlung von F. MIES „Über die Behandlung der Frage im 17. und 18. Jahrhundert“ (*Zeitschr. f. Musikwissenschaft*, 4. Jahrg. S. 286 ff.) ersehen, daß zwar die aufsteigende Bewegung, besonders der phrygische Schluß mit Aufstieg zur Dominante, in der Vertonung der Frage vorherrscht, daß aber auch genug Ausnahmen vorkommen. Bei LULLY z. B. ist die Frage meist abwärts geführt. „Die Franzosen fragen großenteils anders als die Welschen“ — schreibt einmal GRAUN an TELEMANN. Vielleicht hat dieser Umstand auch bei MEYERBEER mitgewirkt.

nicht als Ausdruck der eigenen Gedanken oder Gemütsbewegungen erscheinen soll. Beim Psalmodieren ist diese Deutung allerdings nicht anwendbar, hier dürfte es sich vielmehr zunächst darum handeln, daß man eigentlich gehoben sprechen, deklamieren würde, wenn nicht das Zusammensprechen Vieler mit ihren verschiedenen und schwankenden Tonhöhen eine so abscheuliche und zugleich das Verständnis schädigende Wirkung täte. Dazu kommt aber gewiß auch das positive Bedürfnis der Uniformierung und Verstärkung der von vielen gemeinsam vorgetragenen Gebete und der Ausschaltung aller individuellen Unterschiede gegenüber dem Unendlichen. Man kann experimentell beobachten<sup>1</sup>, wie die sonst leicht erkennbare individuelle Eigenart des Sprechens (abgesehen von gröberen Klangfarbenunterschieden) verschwindet, sobald ein bestimmter Ton festgehalten wird. Die Anfangs- und Schlufsintervalle aber und der Tonfall beim Punkt, Komma, Fragezeichen, wie sie im liturgischen Gesange vorgeschrieben waren und noch sind, wurden so gewählt, daß sie den in der Sprache gebräuchlichen möglichst nahe blieben. Gleitende Tonbewegung in größerem Umfange findet sich an bestimmten Stellen in Gesängen der Naturvölker, besonders ein Abwärtsgleiten am Schlusse. Reste solcher Bewegungen kommen auch im europäischen Volksgesange vor.<sup>2</sup> Sie stellen gleichfalls leichte Annäherungen an das Sprechen dar.

In der Kunstmusik gehört das Portamento hierher, wie es von Sängern und Streichern gebraucht wird. Ein stark aufgetragenes Portamento soll nur selten angewandt werden, wenn etwa der Leitton besonders drängend zur Tonika hinübergeführt werden soll oder an Stellen wie „Fülle der Liebe“ in GOETHE-BRAHMS' Rhapsodie, wo HERMINE SPIES mit dem aufsteigenden Portamento von *a*is bis *e*<sup>2</sup> eine ergreifende Wirkung erzielte (jetzt soll auch SIGRID ONEGIN den gewaltigen Schritt

<sup>1</sup> Vgl. PASSOWS u. SCHAEFFERS „Beiträge z. Anatomie . . . des Ohres usw.“ Bd. 17, 1921, S. 184.

<sup>2</sup> S. Anf. d. Musik, S. 82f. Auch in den von BARTÓK herausgegebenen Volksgesängen der ungarischen Rumänen (4. Bd. der „Sammelbände f. vergleichende Musikwissenschaft“ v. STUMPF und v. HORNPOSTEL, vgl. Einführung S. X, XVI) finden sich viele gleitende Bewegungen, besonders am Schlufs ein Hinabgleiten zur Unterquarte.

in dieser Weise, mit vielleicht noch ausgeprägterem Portament, vollziehen). Aber in fast unmerklicher Form, nur angedeutet, ist es in vielfacher und ganz unentbehrlicher Verwendung. In Verbindung mit einem leisen Vibrieren beseelt es den Vortrag. Streicher können Legato-Stellen, wie z. B. die ersten zwei Noten von SCHUMANN'S A-dur-Quartett ( $fis^2-h^1$ ), unmöglich ausdrucksvoll wiedergeben, wenn sie nicht beim Verlassen der ersten (auf der A-Saite gegriffenen) Note den Finger leicht abwärts gleiten lassen. Das Gleiten muß aber sehr rasch, nur eben merklich, erfolgen und die zweite Note doch wieder scharf eingesetzt werden, so daß der Eindruck der festen und reinen Tonhaltung nicht verloren geht. Es wäre interessant, dieses Verfahren einmal graphisch aufzunehmen, da für die direkte Beobachtung die Tonbewegung viel zu undeutlich und flüchtig ist. Bei den modernen Streichern wird freilich das Portament oft ungebührlich übertrieben. Viele können eine Kantilene nicht ohne beständiges Hinübereutschen von einem Ton zum anderen vortragen. In Kaffeehäusern mag dies wirken, in Konzertsälen beleidigt es feinere Ohren. Es verstößt gegen Geist und Wesen der Tonkunst. Unleidlich sentimental wirkt es bei der Zither mit ihren ohnedies schon larmoyant piepsenden Tönen. So darf also auch der Sänger selbst, dem diese Vortragsmanieren nacheifern wollen, nur mit größter Diskretion von gleitender Tongebung Gebrauch machen.

Es gäbe noch andere Möglichkeiten stetiger Tonbewegung in der Musik. Auf eine solche, sogar in Quintenparallelen, habe ich schon einmal (Tonpsychol. II, 180) hingewiesen. Aber man muß fast fürchten, dergleichen zu verraten. Solche Effekte dürften wieder nur äußerst sparsam zu ganz bestimmten Ausdruckszwecken eingestreut werden, wenn nicht Musik aufhören soll, Musik zu sein.

Eine Annäherung an stetige Tonbewegungen bewirken aber auch schon rasche chromatische Gänge von größerem Umfange. Dabei entfällt das eigentümliche Ethos, der spezifische Gefühlsausdruck, der an die Intervalle der diatonischen Leiter, auch an die alterierten, geknüpft ist, es bleibt nur der Eindruck des Dahinrollens über eine gewisse Strecke des Tonreiches; erst der Schlufston bringt wieder seiner Leiterstellung gemäß ein bestimmtes Intervallgefühl mit sich. Eine Kette von Viertel-

tonstufen würde die Annäherung natürlich noch weiter treiben. Im arabisch-türkischen Stil ist in der Tat der Gesang vielfach durchsetzt mit Koloraturen, bei denen mit großer Schnelligkeit eine Reihe kleinster Intervalle (natürlich nicht in scharfer Abmessung) durchlaufen werden, was dem Europäer leicht wie ein stetiges Auf- und Abheulen erscheint. Aber es sind hier nicht wirklich gleitende Bewegungen beabsichtigt; wie denn auch auf Instrumenten mit festen Griffen ähnliche Gänge ausgeführt werden. Ich habe während des Krieges in Gefangenenlagern viel dergleichen gehört, besonders gut in dem griechischen Internierungslager zu Görlitz bei Stücken, die offenbar unter türkischem Einfluß standen.

Es dürfte wohl auch die Ornamentierung in unserer eigenen Musik (Triller, Doppelschlag usw.) nicht außer Zusammenhang mit dem Bedürfnis stehen, die Fesseln der starren Intonation wenigstens scheinbar zu lockern und sich der freibewegten Tongebung des ausdrucksvollen Sprechens zu nähern, ohne doch der an die Intervalle geknüpften Gefühlswirkungen verlustig zu gehen. Doch wirken hier vielerlei Motive zusammen.<sup>1</sup>

b) Singendes Sprechen andererseits ist ein Sprechen, bei dem sich feste und diskrete Tonhöhen dem Hörenden aufdrängen. Spricht man zunächst möglichst affektlos, trocken und nicht zu laut ein einsilbiges Wort, so ist es immer leicht, die Tonhöhe, die unter diesen Umständen nahezu konstant bleibt, auf dem Klavier anzugeben, vorausgesetzt natürlich, daß der Sprechende nicht eine zu rauhe Stimme hat.<sup>2</sup> Hier ist Sprechen und Singen dem Klangbestande nach tatsächlich

<sup>1</sup> Vgl. BARTÓK a. a. O. über die Ausfüllung der Quinte oder Quarte durch rasche Durchgangstöne.

Wie aus dem gehaltenen Ton durch das Streben nach Beseelung das Vibrando, Balancement, Tremolando und daraus weiter die Zerlegung in mehrere gleiche oder benachbarte Töne hervorgehen, hat R. LACH, Studien zur Entwicklung der ornamentalen Melopöie, S. 47f. gut geschildert. Die Hereinziehung des Darwinismus in diese Dinge, sei es auch nur als Analogie, kann ich allerdings nicht glücklich finden.

<sup>2</sup> Der Ton entspricht dem neutralen Zustande der Stimmlippen zwischen Anspannung und Entspannung, dessen phonetische Bedeutung W. KÖHLER in *Katzensteins Archiv f. experim. u. klin. Phonetik* I (1913) S. 14ff. untersucht hat. Für mich liegt dieser „phonische Nullpunkt“ bei A—c.

dasselbe, obgleich die Aufmerksamkeit des Hörers einmal auf den Ton, das andere Mal etwa auf den Dialekt, die Aussprache der Vokale und Konsonanten oder den Sinn des Gesprochenen gerichtet sein kann und er demgemäß das eine Mal Gesang, das andere Mal bloße Sprache zu hören angibt.

Selbst mehrsilbige Interjektionen und Ausrufungsformeln kann man oft genug ebensogut als gesungene auffassen. So das überraschte „Nanu!“ des Berliners, das je nach dem Grade des Erstaunens von der Terz bis zur Septime und weiter geht, wobei zugleich die Betonung der zweiten Silbe wächst; oder ein lebhafteres „Jawohl!“ „Gewifs!“ oder ein resigniertes „Ja, Ja!“ — gar nicht zu reden von der Musik der „warmen Würstchen“ auf thüringischen Bahnstationen. Überall da pflegen die Intervalle sehr deutlich und keine merklichen Tonschwankungen vorhanden zu sein. Ist dies nun Singen oder Sprechen? Ich möchte sagen: es ist Singen, wenn die Intervalle so weit im Bewußtsein zur Geltung kommen, daß das spezifische Intervallgefühl der Quarte usf., wie es sich bei musikalischen Menschen eines bestimmten Kulturkreises ausgebildet hat, dadurch ausgelöst wird. Außerdem ist es Sprechen. Also je nach dem Subjekt und seiner Verfassung. In der Gehörserscheinung als solcher ist kein wesentlicher Unterschied. Etwas lässiger pflegt ja immerhin der Sprechende die Tonbildung zu behandeln, und etwas Nachsicht muß also der Hörer mit dem Vortrag haben, wenn er Gesang hören will.

Spricht man in größerem Zusammenhang, so kann eine Annäherung an das Singen schon eintreten, wenn der Sprechende auch nur auf einer einzelnen Silbe ungewöhnlich lang eine bestimmte Tonhöhe festhält, sei's auch unwillkürlich, oder wenn er oft und nachdrücklich auf diese Tonhöhe zurückkommt. Ebenso erweckt eine ungewöhnlich hohe Stimme leicht den Eindruck des Singens. Besonders aber der stereotype Gebrauch bestimmter, hervorstechender Intervalle an akzentuierten Stellen. Man wird dann geneigt, die stetigen Teile der Stimmbewegung außer acht zu lassen und sich an diese ausgezeichneten Punkte der Sprachkurve zu halten.

Durch die Anwendung eines solchen auswählenden und summarischen Verfahrens kann man auch schon in der gewöhnlichen, nicht „singenden“, Rede leicht Melodien finden.

So sind die vielen Versuche entstanden, die Sprachmelodie in unseren Noten zu fixieren. Es kommt ihnen nur ein sehr relativer Wert zu, da gerade das Wesentlichste der Sprache dabei außer acht bleibt. Immerhin können sie gewissen Zwecken dienen, namentlich der Vergleichung der einzelnen Sprachen und Dialekte, ja auch der Individuen in bezug auf den Tonfall bei bestimmten Gelegenheiten (SIEVERS) oder der Festsetzung der verschiedenen auf- oder absteigenden Intervalle, durch die ein und dasselbe Wort im Chinesischen und in afrikanischen Sprachen mehrere verschiedene Bedeutungen erhält.<sup>1</sup> Dafs dabei gleichwohl, wenn nicht objektive Methoden benützt werden, ein weniger geübtes Gehör starken Täuschungen ausgesetzt ist, scheint mir unter anderem F. SARANS durchgeführte Notierung von GOETHES „Zueignung“ zu lehren<sup>2</sup>, wonach sich das ganze Gedicht in seiner eigenen Deklamation und der mehrerer anderer Sprecher fast nur innerhalb eines Ganztons bewegen würde. Ich kann dies nur für eine starke Gehörstäuschung halten, auch bei norddeutschem Tonfall, es sei denn, dafs eine ganz abnorme Sprechweise vorläge, deren gleichzeitiges Vorkommen bei den untersuchten Individuen doch äufserst unwahrscheinlich wäre. Das Auf- und Absteigen im allgemeinen mag dabei aber richtig wiedergegeben sein, nur nicht die Intervalle.<sup>3</sup>

<sup>1</sup> Vgl. u. a. HEEPE, Experimentalphonetische Untersuchungen über die Tonhöhen im Jaunde. Abh. des Hamburgischen Kolonialinstituts, Bd. 24. Die hier wiedergegebenen Kurven sind auch für die Stetigkeit der Tonbewegung auf einer Silbe lehrreich. Eigentümlich ist das von H. hervorgehobene Vorherrschen der Quarte, Quinte und Oktave, also der ersten Konsonanzen, als Zielpunkten der Bewegung, natürlich nicht in exakter Abstimmung. Auch dafs die Stimme am Schlufs oft bis zu einer ganzen Oktave herabsinkt, hat sein Analogon in primitiven Gesängen.

<sup>2</sup> F. SARAN, Melodik und Rhythmik der Zueignung Goethes. 1903.

<sup>3</sup> W. KÖHLER, der in seinen „Akustischen Untersuchungen“ III (*Zeitschr. f. Psychol.* 72, S. 106 ff.) den Sprachnotierungen doch wohl allzu skeptisch gegenübersteht, hebt mit Recht hervor, dafs man den Tonumfang des Sprechens vielfach unterschätze. Kritisches über frühere Notierungsversuche, namentlich die L. MERKELS, in meiner oben S. 40 zitierten Abhandlung in der Viertelj.-Schr. f. Musikwiss., S. 278 ff., wo auch gewisse Einwirkungen der Sprachmelodie auf den Gesang und selbst auf die Instrumentalmusik besprochen werden.

Beim „singenden Sprechen“ ist nun der stets vorhandene Tonfall der Rede ausgeprägter, auffallender, es treten gewisse Intervalle mit besonderer Regelmäßigkeit auf, zumal solche, die ihrer Gröfse oder Richtung nach dem Hörer ungewohnt sind und darum seine Aufmerksamkeit auf sich ziehen.<sup>1</sup> So kann der Eindruck graduell in den des Singens übergehen oder sich ihm wenigstens nähern.

Beim Deklamieren zur Musik (Melodram) wird von guten Sprechern die Annäherung künstlich vollzogen, um den Gegensatz zu mildern. So pflegt sich L. WÜLLNER (und ähnlich tat es früher POSSART) in BYRON-SCHUMANN'S „Manfred“ vielfach auf akzentuierten Stellen derart der Begleitung zu akkommodieren, daß wirklich ein Zwischending von Singen und Sprechen entsteht. Er spricht gleichsam in einer bestimmten Tonart, wenn er auch nirgends die genauen Intervalle gebraucht. Die Wiederkehr bestimmter, mit den Instrumentaltönen zusammenfallender absoluter Tonhöhen an ausgezeichneten Punkten bewirkt hier diesen Eindruck. Auch sonst gibt es in der künstlerisch gehobenen Sprache manche Annäherungen an das Singen. So war es eine Art musikalischen Genusses, wenn CHARLOTTE WOLTER etwa „Eilende Wolken, Segler der Lüfte“ rezitierte. Auch ihre subjektive Einstellung dürfte hier zwischen Singen und Sprechen in der Mitte gelegen haben. Schon ein besonders schöner Stimmklang, wie sie ihn besafs, kann durch seinen physischen Reiz zu der musikalischen Wirkung beitragen. In solchen Fällen ist „singendes Sprechen“ auch nichts weniger als ein Tadel.

Da es nützlich ist, über eine der alltäglichen Beobachtung so offen vorliegende Sache auch die Meinungen geistreicher Nichtfachmänner zu hören, so sei hier noch eine Äufserung CARL SPITTELE'S<sup>2</sup> angeführt. Er dreht gewissermaßen die alten Definitionen um. Die sogenannte singende Aussprache habe nichts mit dem musikalischen Singen zu tun, nichts mit festen Tonhöhen, Dreiklangfolgen, jähem Wechsel von hohen und tiefen Tönen. Der Italiener, der am besten singe, spreche nicht singend. „Der Slawe, der bald in den höchsten Fisteltönen, bald im Brummbafs redet, und ganze Sätze auf eine einzige Note stimmt, singt nicht. Der

<sup>1</sup> Darum hört man, wie H. LACHMUND (Vokal u. Ton, *Zeitschr. f. Psychol.* 88, S. 33) richtig bemerkt, bei dem eigenen Dialekt die Tonhöhen oft nicht, während Fremde ihn für singend erklären.

<sup>2</sup> Lachende Wahrheiten, 1920, S. 168 ff.: „Singende Aussprache“.



Süddeutsche und Schweizer dagegen, wenn er sich noch so monoton zwischen zwei naheliegenden Intervallen (sc. Tönen) bewegt, singt. Nämlich unter Singen versteht man das Schwanken der Tonhöhe innerhalb eines einzigen Vokals, und wäre die Schwankung noch so gering. Solch eine Schwankung entsteht aber, wenn der Sprechende einen Gefühlsausdruck in den Vokal legen will. Ein freudiges Ja, ein ärgerliches Nein, durchläuft beim Süddeutschen eine ganze Tonskala, was einem Romanen, einem Slawen, niemals widerfährt. Was tun denn diese in vorliegendem Fall? Sie verwenden das Tempo, die Tonstärke und die Tonhöhe, sie fügen allerlei Adjektiva hinzu, aber unter keinen Umständen werden sie von der einmal angeschlagenen Tonhöhe innerhalb des nämlichen Vokals hinauf- oder hinabgleiten. Nie werden Sie aus italienischem Munde ein No in so ausdrucksvoller Weise vernehmen, wie unser „Nein“; es bleibt bei dem festen No. Sie können den Unterschied innerhalb des Deutschen selbst wahrnehmen: schildert einer seine überstandenen furchtbaren Zahnschmerzen derart, dafs er durch die Töne, die er in das a und u des Wortes „furchtbar“ legt, Eindruck zu machen versucht, dann singt er. Der andere, Nichtsingende, verlegt zu demselben Zweck das ganze Wort „furchtbar“ in die höchste Fistelregion.“

Diese und noch weitere Bemerkungen SPITTELEERS haben unstreitig viel Wahres in sich, und man mufs zugeben, dafs mit dem „singenden Sprechen“ verschiedene Eigenschaften gemeint sein können. Die Annäherung kann eben auf verschiedenen Wegen erfolgen; wir brauchen das wesentlichste Unterscheidungsmerkmal darum nicht aufzugeben. Gewifs werden ungewöhnlich grofse Schwankungen der Tonhöhe auf einer einzigen Silbe als singendes Sprechen empfunden, weil sie die Aufmerksamkeit auf die Tonbewegung als solche hinlenken. Aber kleinere Schwankungen benützt doch auch das nicht-singende Sprechen beständig, und zwar auch im Italienischen und Französischen. Das „No“ des Italieners freilich wird so kurz gesprochen, dafs zu Tonveränderungen kaum Zeit ist, wie sie dem Engländer auf seinem „Yes“, „No“ und dem Deutschen bei Ja und Nein möglich sind. „Furchtbar“ würde ich unter den angegebenen Umständen durch eine sehr hohe erste und tiefe zweite Silbe ausdrücken; wenn es arg war, vielleicht durch eine Duodezime und mehr, selbst mit Hilfe der Kopfstimme bei der ersten Silbe, aber nicht durch eine stetige Tonveränderung. Ich sehe also nicht ein inwiefern dieses Beispiel SPITTELEERS These erhärten soll.

#### § 4. W. Köhlers Anschauung.

Das alte Merkmal: feste Töne und Intervalle gegenüber gleitenden und schwankenden, vermag also nicht blofs dem prinzipiellen Unterschied, sondern auch den in Wirklichkeit

vorkommenden Annäherungen, Mischungen und Übergängen am besten gerecht zu werden.

Gleichwohl ist auf experimental-psychologischer Seite eine tiefergreifende ganz radikale Trennung mit Nachdruck vertreten worden. Nach KÖHLER würde den Sprachlauten gegenüber musikalischen Klängen ein Attribut, eine sinnliche Grundeigenschaft fehlen, die er als „Tonhöhe“ bezeichnet. Das Tonmaterial wäre qualitativ verschieden: Töne mit und ohne Tonhöhe.<sup>1</sup> Unter Tonhöhe versteht KÖHLER (S. 2) „die Eigenschaft, auf die man gerichtet ist, wenn man einen Ton oder Klang nachsingend treffen will oder wenn man einen gegebenen Ton als Quinte eines anderen bezeichnet“.

Wir nennen jetzt zumeist „Tonhöhe“ die Eigenschaft, die sich parallel mit den Schwingungszahlen verändert und wonach eben tiefe und hohe, tiefere und höhere Töne unterschieden werden. Prägnanter nennt man sie wohl auch „Helligkeit“. Dagegen unterscheiden wir davon die Eigenschaft, die Tönen von gleicher Buchstabenbezeichnung, z. B. allen *D*'s, abgesehen von den Zahlenindices, gemeinsam ist. Diese werden seit LOTZE und BRENTANO als „musikalische Qualität“, auch kurz als „Qualität“ bezeichnet. Also *c* und *c*<sup>1</sup> haben dieselbe musikalische Qualität, die sie eben zu *C*'s macht, aber ungleiche Höhe oder Helligkeit. Die musikalische Qualität ist auch für die Unterscheidung der Intervalle innerhalb der Oktave wesentlich (wenngleich nicht allein ausschlaggebend), und sie ist es, auf der das „absolute Tonbewußtsein“ in erster Linie beruht, bei dem ja häufig der Tonname richtig, die Höhenlage falsch angegeben wird.

Was nun KÖHLER mit der „Tonhöhe“ meint, die den Sprachlauten fehle, ist offenbar die musikalische Qualität. Denn wenn ein Bassist einen vorgegebenen Ton aus der 2-gestrichenen Oktave mit Bruststimme nachsingt und trifft, so trifft er eben nur die gleiche musikalische Qualität, aber nicht die Höhe. Und wenn KÖHLER unter den Tonerscheinungen,

<sup>1</sup> Bericht über den 5. Kongress für experimentelle Psychologie 1912, S. 152. Ausführlich in „Akustische Untersuchungen III“, *Zeitschr. f. Psychol.* 72 (1915) S. 103 ff.

denen die Tonhöhe fehle, neben der Sprache auch die Töne der 6-gestrichenen Oktave anführt, an denen man zwar noch Helligkeits- oder Spitzigkeitsunterschiede aber keine musikalischen Intervalle mehr wahrnehme, so geht daraus das nämliche hervor. Es ist mir dies auch mündlich bestätigt worden.  $c$ ,  $c^1$ ,  $c^2$  haben also nach K. die nämliche „Tonhöhe“. <sup>1</sup> Der Unterschied zwischen Sprechen und Singen ist KÖHLER zufolge von ähnlicher Art wie der zwischen Geräusch und Ton. Geräusche haben in der Tat, wie es scheint, keine musikalischen Qualitäten. Wenn man Intervalle, Melodien darin erkennt, so mag es an beigemischten Tönen liegen. Nur der Helligkeit nach sind sie abgestuft. So sieht sich KÖHLER dazu geführt, die Sprachlaute tatsächlich den bloßen Geräuschen beizuzählen. Er übersieht nicht, daß man aus der Sprache gewisse Intervalle mehr oder weniger deutlich heraushören kann, will dies aber lieber als ein Hineinhören bezeichnen, und findet, daß in demselben Moment auch der phänomenale Charakter des Sprechens in den des Singens übergehe.

Ähnlich, aber noch weitergehend, urteilt E. R. JAENSCH (Die Natur der menschl. Sprachlaute, *Zeitschr. f. Sinnesphysiologie* 47, 1913, S. 256f.): „Die wirklich ausgeprägten Vokale, die gesprochenen, zeigen das charakteristische Merkmal der Töne, die Höhe, entweder gar nicht oder nur in ähnlich undeutlicher Weise wie die Geräusche. . . . Das andere Merkmal der Tonreihe, die . . . musikalische Qualität, fehlt natürlich den Vokalen erst recht.“

JAENSCH' Schüler H. LACHMUND findet (Vokale und Töne, *Zeitschr. f. Psychol.* 88, S. 33), daß KÖHLER mit Unrecht JAENSCH' Äußerung als Zustimmung zu seiner Ansicht auffasse. „Bei JAENSCH steht, wie sich der Leser unschwer überzeugen wird, etwas ganz anderes. Es heißt daselbst nicht, daß die Höhe den gesprochenen Vokalen fehlt, sondern nur daß sie in ähnlicher Weise undeutlich [nur hier gesperrt], wie bei den Geräuschen, ist und gelegentlich — nämlich in einem sogleich zu erwähnenden Grenzfall [den geflüsterten Vokalen] — auch fehlen kann.“

Der übereifrige Schüler hat hier doch des Guten zu viel getan. Er

<sup>1</sup> Eine harte Terminologie, die sich schwerlich einbürgern wird. S. 112 spricht er denn auch spezieller von der „musikalischen Tonhöhe“, womit ich aber lieber die durch die Notenbuchstaben einschließlichs der Indices angezeigte Stellung in der Tonreihe ausdrücken würde. An dieser Stelle setzt K. übrigens auch ausdrücklich diese „musikalische Tonhöhe“ und die Helligkeit (Höhe im gewöhnlichen Sinn) einander gegenüber.

hat den Wortlaut der Behauptung, von der sich der Leser „unschwer überzeugen“ soll, seinerseits ganz wesentlich verändert. Er ersetzt nämlich „entweder fehlt oder undeutlich ist“ durch „nicht fehlt sondern undeutlich ist“. Das ist sehr zweierlei! In Wahrheit liegt die Sache so: JAENSCH behält sich den Rückzug vor; bei LACHMUND wird er tatsächlich angetreten.

In einer Hinsicht hat aber JAENSCH den Unterschied zwischen gesprochenen und gesungenen Lauten sogar noch mehr als KÖHLER verschärft. KÖHLER will (was er allerdings deutlicher, ausdrücklicher hätte sagen müssen) nur die Tonhöhe im Sinne der „musikalischen Qualität“ den Sprachlauten absprechen, nicht die mit der Schwingungszahl parallel veränderliche Tonhöhe (Helligkeit). JAENSCH dagegen spricht ihnen beides ab, das letztere bedingt („entweder gar nicht oder undeutlich“), das erstere aber unbedingt und ebenso nachdrücklich wie KÖHLER selbst.

Unsere Kritik im Text richtet sich nur gegen KÖHLER. Sie würde aber gegen JAENSCH' weitergehende Behauptung ähnlich lauten.

Scheinbar, nämlich den Worten nach, lehren übrigens auch schon die altgriechischen Theoretiker einen Unterschied in den Attributen: wenn sie nämlich gesungenen Tönen die „Breite“ absprechen, die gesprochene besäßen. Aber sie meinen damit nur jene gleitende Bewegung über eine gewisse Zone des Tonreiches, jene sukzessive Breite, von der wir schon gesprochen haben.

Diese Formulierung der Tatsachen kann ich nun doch nicht unterschreiben. Ich würde sagen, daß nicht das Lautmaterial selbst sich dabei ändere, sondern nur die Richtung und Verteilung der Aufmerksamkeit. Hätte die Aufmerksamkeit die Kraft, das gegebene Lautmaterial selbst in einem so wesentlichen Punkte umzugestalten, daß ihm eine Grundeigenschaft zuwüchse oder geraubt würde, so dürfte man auch Helligkeits- und Stärkeunterschiede nur dann in den Sprachlauten (ebenso in den Geräuschen) anerkennen, wenn zufällig die Aufmerksamkeit des Hörenden auf solche Unterschiede gerichtet ist. Im übrigen wären die Sprachlaute weder hoch noch tief, weder dunkel noch hell, weder stark noch schwach. Wie aber, wenn der Hörende überhaupt nicht auf das Tonmaterial, sondern auf den Sinn des Gesprochenen achtet, wie dies doch eigentlich beim Sprechenhören die Regel ist? Dann wäre eben damit die ganze sinnliche Grundlage, die Tonerscheinung samt all ihren Attributen, verschwunden und nur der abstrakte Gedanke oder allenfalls die visuelle Vorstellung der besprochenen Gegenstände noch vorhanden. Wir würden verstehen, ohne überhaupt etwas zu hören. Auch bei der

Musik selbst kann es vorkommen, daß wir in eine Melodie als solche, in diese bestimmte Tongestalt und ihren Reiz versunken sind, ohne daran zu denken, ob sie laut oder leise, hoch oder tief, rechts oder links von uns gesungen wird. Sollen wir da sagen, daß all diese konkreten Bestimmungen, die Attribute der Stärke, Höhe, Örtlichkeit in keiner Weise im Bewußtsein repräsentiert seien? Und wenn wir ein Buch lesen und nachher nicht anzugeben wissen, ob es in Antiqua oder Fraktur gedruckt war (was leicht vorkommen kann): sollen wir sagen, daß wir es tatsächlich in keiner von beiden Drucksorten oder in einer unbestimmten Zwischensorte gelesen haben? Gewiß kann man die Beschreibung der Erscheinungen in diesem Sinne unter Schaffung einer neuen Terminologie und Kategorientafel umzugestalten versuchen. Aber ich kann vorläufig weder die Notwendigkeit noch die Fruchtbarkeit dieser mühsamen Operation erkennen.

Ohne Zweifel findet eine Änderung in unserem Bewußtsein in dem Moment statt, wo wir eine Quarte, eine Oktave in der gesprochenen Rede bemerken. Aber sie besteht nicht darin, daß sozusagen ein neues Register der Stimme aufgezogen würde, welches eine Grundeigenschaft mehr in seinen Klängen enthielte<sup>1</sup>, sondern darin, daß unsere gewohnten Intervalle, die in der gesprochenen Rede nur sporadisch, zufällig, auf Augenblicke auftauchen, ohne ein Interesse oder eine Gefühlsbetonung mit sich zu führen, jetzt plötzlich von Bedeutung werden. Es ist, kurz gesagt, ein Wechsel der Auffassung, nicht der Empfindung, der Apperzeption, nicht der Perzeption, des psychischen Verhaltens, nicht der sinnlichen Erscheinung.

An der Anerkennung dieser Sachlage ist KÖHLER dadurch verhindert, daß es für ihn unbemerkte Empfindungsinhalte

---

<sup>1</sup> Der Mechanismus dieser automatischen Registrierung müßte nicht im Sprechenden, sondern im Hörenden, und da natürlich im Gehirn gesucht werden, wo allerdings für Hypothesen noch Spielraum genug ist. Geht der Sprecher selbst zum Singen über, so müßte das Auftreten des neuen Attributs auch in den Sprachwerkzeugen selbst, in ihrer Funktionsweise, aufzeigbar sein, was auf Schwierigkeiten stoßen dürfte. Aber wir lassen physiologisch-genetische Fragen hier überhaupt beiseite.

überhaupt nicht gibt.<sup>1</sup> Was wir augenblicklich nicht beachten, ist danach tatsächlich nicht und in keiner Weise im Empfindungsinhalt vorhanden. Wir können auf diese allgemeinere Frage hier nicht eingehen. Aber gerade Konsequenzen, wie die obigen stehen eben auch der allgemeinen These entgegen. Unbemerkte Teilinhalte (übrigens nur ein Grenzfall, da von der vollen Beachtung bis zum vollen Nichtbemerken viele Stufen führen) bieten keine Schwierigkeit für den, der zwischen einer psychischen Funktion und ihrem Inhalt, beispielsweise zwischen dem Bemerkten eines Tones (oder einer Toneigenschaft oder eines Tonverhältnisses) und dem bemerkten Tone selbst (der Eigenschaft, dem Verhältnis) als einer gegebenen Erscheinung unterscheidet. Die Welt der Sinneserscheinungen, die unserer psychischen Verarbeitung unterliegt, birgt eine Fülle von Eigenschaften, Beziehungen, Gesetzmäßigkeiten, die wir erst nach und nach, vollständig wohl niemals, uns zum Bewußtsein bringen. Diese Unterscheidung aber zwischen den Funktionen und den Erscheinungen, ihrem Material und ihrer Unterlage, kehrt in mehr oder weniger versteckter Form doch überall wieder. Und auch die Lehre von der schöpferischen Kraft der Aufmerksamkeit, die neue primäre Bewußtseinsinhalte, neue Dimensionen des Sehraumes, neue Grundeigenschaften des Tongebietes mit einem Schlage produzieren soll, statuiert nicht bloß eine von den Inhalten verschiedene Funktion, sondern mutet dieser sogar Leistungen zu, die die Grenze des Wunderbaren schon überschreiten. In allen Fällen bleibt es von Wichtigkeit, sich auch an unserem konkreten Falle zu vergegenwärtigen, wie alle diese Fragen und Stellungnahmen unter sich zusammenhängen.

Bei dieser Gelegenheit möchte ich gegenüber Mißverständnissen feststellen, daß ich niemals behauptet habe, durch verändertes funktionelles Verhalten, z. B. durch das Bemerkten einer vorher nicht bemerkten Teilerscheinung, könne die gegebene Erscheinung nicht geändert werden, sondern immer nur: sie müsse nicht unter allen Umständen geändert werden, obgleich es unter den gewöhnlichen komplizierten Umständen des psychischen Verlaufes tatsächlich der Fall sein möge. Und selbst in dieser starken Einschränkung war es ausdrücklich „nicht als sicher

<sup>1</sup> S. seine Abhandlung „Über unbemerkte Empfindungen und Urteils-täuschungen“ in der *Zeitschr. f. Psychol.* 66 (1918), S. 51 ff.

beweisbarer Satz, sondern als These oder Hypothese“ bezeichnet. Die Beispiele, die ich anführte, mögen daher der Selbstbeobachtung anderer sich anders darstellen. Aber selbst dann, wenn niemals funktionelle Veränderungen ohne Erscheinungsveränderungen möglich wären, wäre dies kein Gegenbeweis gegen die Unterscheidung beider. Diese ruht noch auf anderen Fundamenten. So habe ich die Sachlage ausdrücklich bereits in meiner Akademie-Abhandlung über Erscheinungen und psychische Funktionen 1907, S. 15 auseinandergesetzt.

Wesentlich anders liegt nun freilich die Sache, wenn der Sprechende selbst während des Sprechens seine Aufmerksamkeit dem Tonfall seiner Rede zuwendet. Dann wird das gesamte Tonmaterial sich ändern, auch dann freilich nicht durch einfaches Hinzutreten eines Attributs und nicht durch die bloße Zuwendung der Aufmerksamkeit an sich, sondern infolge durchaus angebbarer und verständlicher Zwischenglieder. Wer etwa die beiden Sätze: „Bist du spazieren gegangen? Ich bin spazieren gegangen“ im Angesichte der HELMHOLTZschen Notierung der Sprachmelodie oder in Erinnerung daran spricht, um die Richtigkeit der Notierung zu prüfen, der ist in einer ganz anderen Bewußtseinsverfassung, sowohl den Inhalten als den Funktionen nach. Die Noten schweben ihm vor und er hat sich eine Aufgabe gesetzt, zwar nicht die des Treffens, sondern nur des Vergleichens. Aber kann er sich dem Einflusse der Noten ganz entziehen? Schwerlich. Unwillkürlich wird sich doch sein Tonfall den Noten mehr oder weniger anpassen. Die Sprache wandelt sich dabei allmählich in Gesang, wie Daphne in einen Lorbeerbaum. Aber worin besteht die Verwandlung? In erster Linie im besseren Hervortreten der schon vorhandenen, wenigstens in den betonten Silben vorhandenen, Tonhöhen. Dieses Hervortreten selbst wird vermittelt durch eine gedehntere und zugleich präzisere Wiedergabe der Töne auf den Vokalen und durch Zurücktretenlassen der Konsonanten, wie dies beim Singen Vorschrift ist, vor allem aber durch Vermeidung der kleinen Tonschwankungen zu Beginn und Schluß von Silben, durch strenge Beschränkung auf die notierten Töne und Intervalle in möglichst genauer reiner Stimmung. In demselben Maße, wie diese Veränderung fortschreitet, wird das Sprechen zum Singen.

Auch in der Subsumtion der Sprachlaute unter die Ge-

räusche können wir sonach KÖHLER nicht folgen. Zu verkennen ist freilich nicht, daß die Sprache in weit näherer Beziehung zu den Geräuschen steht als der Gesang. Die Flüstersprache besteht im wesentlichen aus Geräuschen, denen nur schwache Tonelemente beigemischt sind. Bei der stimmhaften Sprache spielen die Konsonanten, die größtenteils Geräuschformen sind, eine ganz entscheidende Rolle. In slawischen Sprachen gibt es sogar Wörter, die nur aus Konsonanten bestehen. Beim Gesang müssen dagegen, wie erwähnt, die Konsonanten gegen die Vokale zurücktreten und dürften wohl Wörter wie Krě auch von tschechischen Sängern vermieden werden. Selbst die Vokale sind beim Sprechen in der Regel geräuschiger als beim Singen (s. o. S. 41). Viele Sprechstimmen haben etwas Belegtes, Pelziges, ohne daß es uns irgendwie störte. Bei totaler Heiserkeit kann man noch sprechen und verstanden werden, während der Sänger dann zum Schweigen verurteilt ist. Dies alles kann zugegeben werden. Aber daraus folgt noch lange nicht, daß man darum die Sprachlaute einfach den Geräuschen beizuzählen oder auch nur irgendwie zu koordinieren hätte.

Schließlich ist auch nicht einzusehen, wie sich die oben skizzierten objektiv vorhandenen mannigfachen Mittelstufen und Übergänge zwischen den extremen Formen des Singens und Sprechens sollten verstehen lassen, wenn ein so fundamentaler Unterschied unter den Klangerscheinungen selbst bestände. Wo sollte die Sprache anfangen? Etwa wenn jede Spur einer erkennbaren Tonhöhe getilgt wäre? Dann gibt es überhaupt kein Sprechen. Oder wo sollte, wenn man vom Sprechen ausgeht, das Singen anfangen? Etwa erst wenn zwischen sämtlichen Tönen feste und genau wiederkehrende Intervalle beständen? Dann sind die meisten exotischen Gesänge und alle unrein gesungenen unserer Musik noch ein bloßes Sprechen.

Wir können also die hier vorgeschlagene Grenzregulierung nicht als eine glückliche Lösung betrachten.



## § 5. Zur Phänomenologie der stetigen Tonveränderungen.

Eine Seite der Sache bleibt indessen noch zu erwägen, die zu tieferem Eindringen auffordert und uns aus einer gewissen Richtung der KÖHLERSchen Anschauung näher führt.

In allen Sinnesgebieten sind stetige und sprunghafte Übergänge begrifflich scharf zu unterscheiden. Ein stetiger Übergang oder eine Ton-Bewegung im prägnanten Wortsinn ist nicht eine Summe aufeinanderfolgender Töne, mag auch physiologisch eine diskrete Reihe wenig verschiedener Fasern erregt werden, sondern eine einheitliche, qualitativ eigenartige Tonerscheinung, so einheitlich wie ein festliegender Ton.<sup>1</sup> Ob die Stetigkeit dabei wirklich eine ursprüngliche immanente Eigenschaft der gegebenen Sinneserscheinung ist (wie ich glaube), oder ob nur unmerklich verschiedene diskrete Empfindungen rasch aufeinanderfolgen, die aber durch ein apriorisches Verfahren des Verstandes oder einer „produktiven

<sup>1</sup> Vgl. Tonpsychol. I, 183 ff. Diese Ausführungen über die Erscheinungsweise des Kontinuums wurzeln in den gründlichen Untersuchungen F. BRENTANOS in seiner Metaphysikvorlesung v. J. 1868/9, wo er anlässlich der eleatischen Argumente den Begriff der Bewegung als der „fließenden Wirklichkeit“ von dem der ruhenden unterschied. Weiter zurück mag wohl die aristotelische Definition der Bewegung als „ἐνέοργεα ἀτελής“ die Vorläuferin dieser Unterscheidung gewesen sein. Dafs BRENTANO von der objektiven Wirklichkeit sprach, während hier von den Erscheinungen die Rede ist, macht nichts aus, da der Begriff der objektiven Bewegung natürlich aus dem Anschauungsbilde der erscheinenden gewonnen werden mufs.

Bewegungen in diesem allgemeinen Sinne lassen sich, scheint es, bei den Qualitäten aller Sinne neben diskreten Veränderungen herstellen. In neuerer Zeit hat man bei den Bewegungen im engeren Sinne, den durch die Gesichts- und Tastempfindungen wahrgenommenen räumlichen Bewegungen, diese Eigenart der sinnlichen Erscheinung näher studiert und die „optischen Bewegungsempfindungen“ sogar als eine selbständige Qualitätenklasse, einen besonderen Sinn neben den Wahrnehmungen ruhender Gröfsen und Gestalten in Anspruch genommen. Dies geht aber zu weit. Man müfste sonst auch die bewegten Tonhöhen einem anderen Sinn als dem Tonsinn zuweisen.

Es ist möglich und liegt im Sinne der neueren „Gestaltpsychologie“, den Begriff der Bewegung unter den der Gestalt zu subsumieren, indem man sie als eine stetige Sukzessivgestalt definiert. Doch kann diese weitere Verallgemeinerung hier beiseite bleiben.

Einbildungskraft“ zu einer stetigen Linie verschmolzen werden, braucht uns hier nicht zu beschäftigen.<sup>1</sup> Jedenfalls glauben wir stetige Übergänge ebenso wie sprunghafte sinnlich wahrzunehmen und ist zwischen beiden ein prinzipieller begrifflicher Unterschied. Nur insofern könnte man beide unter einen Hut bringen, als man den ruhenden Ton als einen Grenzfall ansähe, dem man sich bei Verkleinerung des Bewegungsraumes immer mehr nähert, gewissermaßen als einen unendlich wenig bewegten Ton. Das wären aber begriffliche Operationen, die die Verschiedenheit der sinnlichen Eindrücke in sich selbst nicht aufheben.<sup>2</sup>

Wie verhält es sich nun mit den beiderseitigen Attributen? — Die Frage ist bisher noch kaum besprochen. Die ganze Phänomenologie der bewegten Erscheinungen liegt ja noch in den Windeln, wie auch die experimentelle Untersuchung erst begonnen hat. Folgendes dürfte sich vorläufig sagen lassen.

Ohne Zweifel hat auch der bewegte Ton eine Höhe und eine Stärke. Seine Höhe erstreckt sich nur eben in fließender Form vom Anfangs- bis zum Endpunkte der Bewegung, z. B. von *c* bis *g*, wobei diese Strecke in gleichbleibender oder in wechselnder Richtung und dann mit den verschiedensten Umkehrpunkten (auch über die Grenzpunkte hinaus) durchlaufen werden kann. Die Stärke kann währenddessen konstant bleiben oder sich gleichfalls zwischen bestimmten Grenzen in verschiedener Weise ändern. Immer wird es bei stetiger Veränderung nicht ein Aggregat aufeinanderfolgender einzelner Höhen oder Stärken, sondern eine einheitliche Höhe oder Stärke sein, nur eben in anderem, fließendem, Sinne. Übrigens kann

<sup>1</sup> Vgl. dazu Tonpsychol. a. a. O. und II, 116 (Akkommodation der spezifischen Energien).

<sup>2</sup> Über diese Schlussweise bzw. ihren Mißbrauch sagte kürzlich F. HILLEBRAND (Zur Theorie der stroboskopischen Bewegungen, *Zeitschr. f. Psychol.* 89, 1922, S. 213) ebenso richtig wie schalkhaft: „Betrachtungen, die sich auf dem Übergang einer endlichen Größe zum Nullwert aufbauen, sind nie ganz ungefährlich. Man kann leicht dahin kommen, einem Menschen, der nichts besitzt, noch immer ein Vermögen zuzuschreiben . . . Aber die Philosophen manövrieren mit 0 und  $\infty$  von jeher mit einer Sorglosigkeit, als wenn es sich um echte Zahlen handelte, deren Handhabung ihnen sichtlich viel schwerer fällt.“

sich die Stärke ja auch schon bei einer festen Tonhöhe während deren Dauer stetig verändern und kann umgekehrt während einer Tonbewegung eine Stärke die andere unstetig ablösen.

Man müßte für die bewegten Tonhöhen besondere Bezeichnungen und Symbole erfinden und hat es zuweilen versucht. Verläuft die Bewegung in gleichbleibender Richtung und Stärke, so ist sie zwar durch Angabe ihrer Grenzpunkte im allgemeinen charakterisiert, aber selbst dann gibt es noch zahllose Modifikationen je nach dem Tempo; und dieses selbst kann für die einzelnen Teilstücke verschieden sein, wofür besondere Epitheta und Zeichen erforderlich wären. Noch mehr bedürfte es solcher für die Unterschiede des Höhenverlaufes selbst zwischen den beiden Endpunkten. Bleibt die Stärke konstant, so bedarf es für sie keiner anderen Maßbezeichnungen als für die eines ruhenden Tones; verändert sie sich aber ebenfalls, so sind wieder besondere Kategorien einzuführen, wie sie die Musik für ihre diskreten Tonschritte und die festen Einzeltöne im „Crescendo, Diminuendo“ usw. und den entsprechenden anschaulichen Zeichen besitzt.

Lassen wir die „Vokalität“ als eine selbständige Grundeigenschaft neben Höhe und Stärke bei ruhenden Tönen gelten so kommt auch dieses Attribut bewegten Tönen in analoger Weise wie Höhe und Stärke zu und kann auch seinerseits in Bewegungsform auftreten. Wenn z. B. bei einem freudig überraschten „Ah!“ die Stimme stetig in die Höhe und wieder herabsteigt, kann der Vokalcharakter gleichbleiben oder sich mitverändern.

Neuerdings hat ISSERLIN auf objektivem Wege, durch Auszählung der Zacken von Vokalkurven, gezeigt, daß in der lebendigen Sprache mit der Tonhöhe tatsächlich auch die Klangfarbe der Vokale stetig wechselt, indem der Formant einunddesselben Lautes während des Aussprechens beim Frage-satz steigt, bei der Affirmation sinkt.<sup>1</sup> Mit dem Formanten ändert sich dabei natürlich der Vokal selbst, es ist eben, genau gesprochen, nicht mehr dasselbe A oder O, sondern ein sich stetig erhellendes oder verdunkelndes. Es wäre dann also

<sup>1</sup> ISSERLIN, Psychologisch-phonetische Untersuchungen (Sitzungsber.), *Zentralbl. f. d. gesamte Neurologie u. Psychiatrie* 26, S. 77.

ähnlich wie beim Flüstern (s. o. S. 39). Dagegen fanden sich solche gleitende Veränderungen der Klangfarbe bei gesungenen Vokalen nicht. Also auch in dieser Beziehung der Unterschied der schwankenden und festen Tongebung. Immerhin wird man die Festigkeit auch nur *cum grano salis* verstehen dürfen.

Wie steht es endlich mit der „musikalischen Qualität“ bewegter Töne? Gibt es da eine Identität der Oktaven, gibt es musikalische Intervalle, gibt es absolutes Tonbewußtsein?

Setzen wir zuerst den Fall, daß es sich nur um kleine Höhenschwankungen, etwa im Rahmen eines Halbtons, handle und daß eine zweite Tonbewegung um eine Oktave höher als die erste, sonst aber in genau analoger Weise nach Richtung und Verlauf erfolge, so wird man hier wohl noch einen Oktaven-eindruck haben, indem man in Gedanken einen festen für den schwankenden Ton einsetzt, wie man es bei schlechter Musik, bei einer tremolierenden Stimme ja auch wohl oder übel tun muß. Aber das ist nicht eigentlich das verlangte Urteil auf Grund des bewegten Tones als solchen. Die Frage ist also, ob auch größere Bewegungen, etwa von *c* bis *g* oder *gis*, als Träger einer einheitlichen musikalischen Bewegungsqualität aufgefaßt werden, derart, daß man ihre Wiederholung in der Oktave als Wiederauftreten einer gleichen musikalischen Qualität, ihre Wiederholung in der Quarte als Quartintervall zwischen zwei bewegten Tönen auffaßt, und daß Geübte imstande wären, eine gegebene Bewegung als eine bestimmte, in Bewegungsform auftretende musikalische Qualität (mit Absehen von der Höhenlage) zu bezeichnen.

Systematische Versuche darüber liegen nicht vor. Es muß dabei in erster Linie experimentell dafür gesorgt werden, daß der Anfangs- und der Schlufspunkt der Bewegung in keiner Weise objektiv ausgezeichnet werden, daß namentlich kein Verweilen darauf stattfindet. Dies ist durch Abschlussvorrichtungen an einer durch eine Wand geführten Röhrenleitung, durch Ausschneiden eines zeitlichen Stückes aus einer schon im Gange befindlichen Tonbewegung zu erzielen. Meines Wissens haben nur O. ABRAHAM und E. v. HORNBOSTEL im Verlauf ihrer (bisher nicht veröffentlichten) Untersuchung über Vergleichung von Tondistanzen Versuche mit einer solchen

Einrichtung für Tonbewegungen gemacht, weil sie hofften, auf diesem Wege reine Distanzurteile ohne Einmischung von Intervallurteilen zu bekommen. Sie fanden aber, daß tatsächlich auch hier Intervallurteile abgegeben wurden. Man erkennt den Anfangs- und den Schlusspunkt der Bewegung, obgleich sie nur punktuell, streng momentan, gegeben werden. Psychologisch hätte man dies nicht wohl voraussagen können, da eine Erscheinung, die nicht eine gewisse endliche Zeit andauert, sonst nicht erkannt werden kann. Es muß eben doch eine innere Markierung der objektiven Grenzpunkte stattfinden. In der Empfindung oder in der Vorstellung (und wohl schon zentral-physiologisch) müssen die Grenzpunkte sich in Dauertöne verwandeln, an die das Intervallurteil sich halten kann.

So interessant dies Ergebnis ist, liefert es doch nicht etwa den Beweis, daß den Tonbewegungen als solchen musikalische Qualitäten zukommen. Denn es wird hier durch die innere Markierung doch wieder eine Reduktion auf feste Grenztöne erzielt, auf die sich dann das musikalische Urteil bezieht. Werden zwei Tonbewegungen unter solchen Umständen verglichen und als im Oktaven- und Quartenvverhältnis zueinander stehend beurteilt, so sind es in Wahrheit zwei Schritte zwischen je zwei festen Tönen, die verglichen werden (z. B.  $c-es$  und  $c^1-es^1$  oder  $f-as$ ).

Aber, könnte man fragen, ist es denn eigentlich bei den bewegten Tonhöhen anders? Nennen wir z. B. eine Tonbewegung höher oder umfangreicher (distanzgrößer) als eine andere, tun wir's dann nicht auch nur, weil die Grenzpunkte der einen höher bzw. weiter auseinanderliegen als die der anderen? Wir haben eben kein anderes Mittel, Tonbewegungen in Hinsicht ihrer Höhe wie ihrer musikalischen Qualität zu charakterisieren als durch die Eigenschaften der festen Töne, die an ausgezeichneten Punkten auftreten oder durchlaufen werden.

Dennoch ist hier ein Unterschied zwischen Höhe und musikalischer Qualität. Bei einer aufsteigenden Tonbewegung haben wir, auch abgesehen vom Anfangs- und Endpunkt, auch während des Verlaufes selbst, den ganz zwingenden Eindruck der Höhenveränderung in bestimmter Richtung, der Erhellung oder Verdunkelung, Erhöhung oder Vertiefung. Haben wir

ebenso den Eindruck einer stetigen Umwandlung der musikalischen Qualitäten? Erscheint uns die gesamte so durchlaufene Tonlinie als eine Spirale mit so vielen Windungen als sie Oktaven aufweist?

Ich habe dies früher (Tonpsychologie II, 196 ff.) rundweg verneint und daraus geschlossen, daß die Eigenschaft, die wir jetzt nach LOTZE und BRENTANO als musikalische Qualität bezeichnen, nicht zu den Grundeigenschaften der Töne gehöre. Ich glaubte sie damals empiristisch, als Produkt der individuellen musikalischen Entwicklung, insbesondere der Verschmelzungserfahrungen, erklären zu sollen. Die damals angeführten Beobachtungen erscheinen mir auch jetzt noch unter der dort erwähnten Bedingung richtig, daß man sich eine gleitende Tonbewegung von tiefen zu immer höheren Tönen in raschem Tempo vorführt, ja auch bei einer rasch ausgeführten chromatischen Tonfolge, selbst einem rapiden Glissando über die weißen Tasten. Dabei wird man kaum etwas von Oktavenwiederkehr, von spiraler Bewegungsform bemerken. Die geradlinige Höhenveränderung dominiert. Wenn aber die Bewegung langsam genug erfolgt und die Aufmerksamkeit scharf auf diesen Fragepunkt konzentriert bleibt, dürfte es für einen gut Musikalischen und Geübten nicht unmöglich sein, das Wiederauftreten gleicher musikalischer Qualitäten an bestimmten Punkten (bzw. innerhalb kleiner Teilstrecken) wahrzunehmen. Das Urteil tut sich nur schwerer, diese Punkte (Strecken) herauszuheben, als bei den Anfangs- und Schlusspunkten.

Treffen wir Einrichtungen, wodurch die Bewegung jedesmal, wenn übereinanderliegende Oktaven eines gewissen Tones, z. B. des Ausgangstones, passiert werden, ein wenig verlangsamt oder der Ton ein wenig verstärkt wird (beides auch wieder in stetigen Übergängen), so wird diese Erkenntnis begünstigt, es werden die musikalischen Qualitäten doch auch innerhalb einer stetigen Tonveränderung, und zwar selbst als stetig veränderte „fließende“ Qualitäten bemerkbar werden.

Es scheint mir hiernach bei bewegten Tönen zwischen der musikalischen Qualität und den übrigen Attributen nur ein gradueller Unterschied zu bestehen. An sich ist auch eine stetige Qualitätsveränderung vorhanden und mit der Höhen-

veränderung ebenso integrierend verknüpft, wie es bei sprunghaften Tonveränderungen der Fall ist. Aber sie ist nur für Individuen mit sehr geübtem musikalischen Gehör unter besonders günstigen Umständen bemerkbar. (Wer unbemerkte Eigenschaften der Sinneserscheinungen nicht zugibt, muß eben sagen, daß sie nur unter diesen Umständen überhaupt vorhanden ist. Aber auch dann ist gegenüber den sonstigen Attributen kein prinzipieller, sondern nur ein gradueller Unterschied.)

Der graduelle Unterschied ist immerhin bedeutend genug. Hat man eine in jeder Hinsicht (nach Stärke, Höhe usw.) unregelmäßig schwankende Tonbewegung, wie das Heulen des Sturmes, gehört, so bleibt ein relativ deutliches Bild des Stärkeverlaufes im Bewußtsein, ein vielleicht etwas undeutlicheres der Höhen-, ein noch undeutlicheres der Klangfarbenveränderungen, das undeutlichste aber wohl von den musikalischen Intervallen zwischen den einzelnen Höhenwendepunkten.

Es ist schon früher die Vermutung ausgesprochen worden, daß die Höhenunterschiede phylogenetisch früher aufgetaucht seien als die der musikalischen Qualitäten. Diese dürften sich erst im Verlaufe der Musikentwicklung, im Zusammenhang mit der Auswahl fester Tonstufen ausgebildet haben; wie sie denn auch heute noch in sehr ungleichem Maße bei den Tonempfindungen verschiedener Individuen (bei manchen vielleicht gar nicht) vorhanden sind. Auf die stetigen Tonveränderungen sind sie dann erst von den sprunghaften aus hinüber gewandert. Daraus würde sich verstehen lassen, warum sie bei den stetigen nur rudimentär und schwach vertreten sind. Sie sind da ja auch nur zwecklose Nebenerscheinungen. Entwickelte sich etwa in Zukunft eine Musik, die auf den kunstvollen Gebrauch stetiger Tonveränderungen gegründet wäre und nur aus solchen bestände, so würden die Intervallbeziehungen mit allen daran geknüpften Emotionen auch wieder verschwinden; es müßten dann andere wirksame Reize gefunden werden, um Ersatz zu schaffen und den Jammer wieder in Genuß zu verwandeln.<sup>1</sup>

<sup>1</sup> Ich kann nicht leugnen, daß das polyphone Geheul, welches die Wölfe und Schakale des Zoologischen Gartens öfters im Tage anstimmen (einer fängt an und sofort folgen die anderen) meinem durch futuristisch-atonale Musik schon einigermassen abgehärteten Ohr ein gewisses Ver-

Kehren wir nun zu dem Unterschied des Singens und Sprechens zurück, so ergibt sich, daß in den bewegten Teilen der Rede, bei den größeren stetigen Höhenschwankungen der Stimme, die musikalischen Qualitäten gegen die übrigen Grundeigenschaften stark zurücktreten müssen, wenn sie auch nicht vollständig fehlen. Wenn wir daher die Noten, durch welche die erkennbaren relativ festen Tonhöhen der Rede angegeben werden, durch gerade Striche oder andere Kurven, die die stetigen Übergänge andeuten, miteinander verbinden, so ist dabei zu beachten, daß durch diese das lebendige Sprechen vielfältig durchsetzenden stetigen Tonveränderungen der musikalische Eindruck der Sprachmelodie ganz wesentlich beeinträchtigt werden muß. Aber es darf nicht übersehen werden, daß hieran nur die bewegten Teile der Rede Schuld sind und daß bei diesen eben in der Stetigkeit der Tonveränderung selbst die letzte Ursache dieses Unterschiedes liegt. Das grundlegende unterscheidende Merkmal liegt also tatsächlich in dem Gebrauche stetiger neben sprunghaften Tonveränderungen durch die Sprache gegenüber dem prinzipiellen Ausschlusse stetiger Höhenveränderungen in der Musik. Die griechischen Musiktheoretiker haben dies vollkommen richtig erfasst, wenn gleich nicht allseitig genug erläutert.

### § 6. Rückblick auf Köhlers These.

Wir verstehen nun aber auch besser das Zustandekommen und die Wurzeln der KÖHLERSchen These. KÖHLER selbst geht von der Stetigkeit der Tonveränderungen in der Sprache aus, ja er scheint geneigt, diskrete Stufen, Schritte in festen Tönen, hier überhaupt für ausgeschlossen zu halten (a. a. O. 105, 108). Gerade in diesen beständigen Schwankungen erblickt er den Grund für das Fehlen der musikalischen Qualität. „Phänomenal gesprochen, ist eben mit dem ‚festen Ton‘ die Tonhöhe [musi-

---

gnügen bereitet. Wie jeder seine Weise durchführt, jede klar und deutlich unterscheidbar, mit verschiedenen Wende- und Höhepunkten, mit voller Unabhängigkeit der Stimmen — eine Art linearer Kontrapunkt —, wie Klage, Wut, Resignation darin ausgedrückt scheinen, das kann fesseln und interessieren, freilich nicht auf längere Zeit. Das Konzert ist denn auch immer nach wenigen Minuten zu Ende.



kalische Qualität] entstanden, die in der gleitenden Sprachbewegung nicht möglich war“ (S. 115).

Soviel ist klar: wenn man die beiden Prämissen zugibt, 1. dafs nur stetige Veränderungen im Sprechen vorkommen, 2. dafs solche mit musikalischer Qualität unverträglich sind, dann folgt KÖHLERS These mit Notwendigkeit, auch wenn man von allen prinzipiellen Fragen über die Existenz unbemerkter Teilinhalte u. dgl. absieht.

Aber die eine wie die andere Prämisse scheint mir eben zu weit zu gehen. Die erste könnte nur etwa vom rein mathematisch-physikalischen oder metaphysischen Gesichtspunkt aufrecht erhalten werden. Wenn man's absolut genau nimmt, behält objektiv wohl jede noch so kurze Silbe ihre Tonhöhe nicht bei. „Alles fließt.“ Ich möchte dies sogar auch subjektiv, für die Phänomene, behaupten. Aber es würde doch genau so auch von gesungenen Tönen gelten. Und da hat man sogar recht erhebliche Schwankungen auf graphischem Wege bei den besten Sängern nachgewiesen. Hierin kann also kein Unterschied liegen. Es könnte sich nur um solche Höhenschwankungen handeln, die sich wenigstens für geübte Ohren als solche bemerklich machen, ja aufdrängen. Für solche Schwankungen kann man aber die obige Prämisse nicht festhalten, d. h. man kann nicht behaupten, dafs beim Sprechen immerfort Höhenschwankungen stattfänden, die für geübte Ohren als solche bemerkbar wären oder gar sich aufdrängten, und dafs infolgedessen auch niemals musikalische Intervalle bemerkbar wären. In den Beispielen zweisilbiger Wörter oder Redensarten, die wir oben S. 51 erwähnten, haben die beiden Silben so gut ihre festen Tonhöhen und Intervalle wie beim Singen. Und so tausendfach.

Dafs aber die zweite Prämisse von KÖHLER tatsächlich zugrunde gelegt würde, könnte zunächst um so mehr wundernehmen, als er nicht blofs für die Höhen im Sinne der Helligkeiten, sondern auch für die Vokalitäten der Sprache stetige Veränderlichkeit in Anspruch nimmt (S. 117). Für die Intensitäten gilt natürlich dasselbe. Wenn nun alle übrigen Attribute stetige Veränderungen zulassen, warum sollten sie bei der musikalischen Qualität ausgeschlossen sein? Die begrifflichen Schwierigkeiten sind keine anderen und gröfseren hier wie dort.

In Wirklichkeit werden denn auch von KÖHLER, wie ich durch mündliche Aufklärung und Hinweis auf eine etwas versteckte, vorher übersehene Stelle seiner Abhandlung erfahre, fließende musikalische Qualitäten der Sprache durchaus nicht aberkannt, sondern nur feste. Nur solche hält er für unverträglich mit der gleitenden Bewegung. Nachdem er sehr richtig den fundamentalen Unterschied diskreter und bewegter Erscheinungen betont hat, fügt er in einer Anmerkung hinzu (S. 116): „In welchem Material die Sprachmelodie im engeren Sinn verläuft, bleibt nach den obigen Ausführungen noch unentschieden; nur die Tonhöhen werden ausgeschlossen. Und man kann nun zwei Möglichkeiten prüfen: entweder ist Sprachmelodie Helligkeitsbewegung, oder aber es gibt ein ‚Hinauf, Hinunter‘ und mehr dynamische Erscheinungen verwandter Art, die sich zu Tonhöhen ähnlich verhalten, wie gesehene Bewegung zu gesehenen Punkten (hier gesperrt), und die ebensowenig in ein Nacheinander von Tonhöhen aufzulösen sind wie Bewegung in sukzessiv aneinandergereihte ‚Raumwerte‘, — daß solche Erscheinungen die Sprachmelodie ausmachen, ist die zweite Möglichkeit. Die Realität solcher Erscheinungen an und für sich müßte sich auch sonst nachweisen lassen.“

Hieraus geht klar hervor, daß KÖHLER unter den „Tonhöhen“, die der Sprache fehlen sollen, nur die festen musikalischen Qualitäten versteht. Er gibt dagegen zu, daß fließende begrifflich möglich sind, und hält es auch nicht für unmöglich, daß solche der Sprache zukommen. Dies bedeutet eine sehr wesentliche Milderung seiner These zum Unterschiede von der Fassung, die sich nach dem Wortlaute zunächst wohl jedem aufdrängt, wonach musikalische Qualitäten überhaupt der Sprache fehlen sollten.

Es dürfte gleichwohl nicht überflüssig sein, über die obige zweite Prämisse noch einiges zu bemerken. Denn die Anschauung liegt ja nahe und wird von Kennern der Akustik vertreten, daß musikalische Qualitäten ihrer ganzen Natur und Entstehungsweise nach nur als diskrete ausgezeichnete Punkte der Tonlinie möglich seien. Jede Tonleiter besteht eben aus einer bestimmten Anzahl fester Stufen, und es ist ja auch richtig und von mir selbst stets hervorgehoben, daß der Ur-

sprung der Musik im eigentlichen Sinne mit der Einführung fester Stufen zusammenfällt, und dafs die wesentlichsten Wirkungen der Musik auch heute noch durchaus an feste Stufen gebunden sind.

Gleichwohl mufs ich, wie vor Jahren<sup>1</sup>, behaupten, dafs die diskreten „historischen Qualitäten“, wie sie in den Leitern vorliegen, einer an sich stetigen Reihe musikalischer Qualitäten entnommen sind. Gäbe es nur eine einzige Leiterform auf der Welt und bliebe sich auch die absolute Höhe der Leitertöne, die Stimmung, stets genau gleich, dann freilich läge die Annahme nahe, dafs die musikalischen Qualitäten nur an ganz bestimmte einzelne Tonhöhen gebunden wären. Aber verschiedene Zeiten und Völker nehmen sich verschiedene Qualitätensysteme heraus, unterscheiden „ausgezeichnete Punkte“ nach sehr verschiedenen Gesichtspunkten; und wir selbst erheben, wenn wir die Stimmung um einen Viertelton höher nehmen, durchweg andere Punkte zu ausgezeichneten als kurz vorher. Mit den Tonhöhen sind meiner Meinung nach auch die benützten Qualitäten durchweg etwas verschoben. Man nennt das neue  $a^1$  freilich auch jetzt noch  $a^1$ , ebenso wie die Anwendung des Wortes und Begriffes „Rot“ auf eine vorliegende Farbe eine gewisse Nuancenbreite einschließt, sich über einen gewissen Abschnitt des Spektrums erstreckt. Aber eine Veränderung der Empfindung liegt doch vor, und zwar ist die Veränderung in Hinsicht der musikalischen Qualität meinem Urteil nach sogar gröfser als die nach der blofsen Helligkeit. Mit der neuen Qualität ist dann durch die Verhältnisse der Oktavenidentität, der Verschmelzungsgrade, der Intervalle ein ganzes System neuer Qualitäten wieder fest gegeben.

Wollte man daran festhalten, dafs die musikalischen Qualitäten, wie sie jetzt vorliegen, ihrer Natur nach diskret sind, so müfste man bei Transposition um einen Viertelton annehmen, dafs das System dieser Qualitäten sich innerhalb einer gewissen Breite der Schwingungsfrequenzen verschieben, d. h. dafs mit etwas verschiedenen Höhen sich die gleichen Qualitäten verbinden können. Das würde zu eigentümlichen Konsequenzen

<sup>1</sup> Bericht über den 6. Kongrefs f. exp. Psychol. 1914, S. 324, 348.

führen, die hier auf sich beruhen mögen. Jedenfalls kann nicht davon die Rede sein, daß ausschliesslich bei bestimmten unveränderlichen Schwingungszahlen sich musikalische Qualitäten fänden, bei den dazwischenliegenden aber überhaupt keine.

Es könnte immerhin daran gedacht werden, daß phylogenetisch der Prozeß umgekehrt verlaufen wäre, daß zunächst nur einige Punkte der stetigen Helligkeitslinie sozusagen einen qualitativen Schimmer annahmen, und daß dieser sich erst allmählich infolge des Wechsels der absoluten Tonhöhen beim Singen und Spielen auf die ganze musikalisch gebrauchte Region ausdehnte, aus deren nun stetig gewordenen Qualitätenschatz dann weiterhin die jeweiligen Gebrauchsleitern entnommen würden. Aber das sind alles Spekulationen, die hier, wo es nur die Beschreibung des vorliegenden Tatbestandes gilt, besser beiseite bleiben.

So scheint mir die Annahme einer an sich stetigen Qualitätenslinie die einzige zu sein, die ungezwungen und ohne Hilfsannahmen mit den elementaren Tatsachen der Musik zurecht kommt.

## § 7. Ergebnis.

In summa: Sprache und Gesang sind in phänomenal-akustischer Hinsicht nicht spezifisch, d. h. nicht durch irgend eine Grundeigenschaft, sondern nur gradweise verschieden. Aber der Gradunterschied ist im allgemeinen doch so einschneidend, daß die beiden Klassen menschlicher Lautprodukte in gewöhnlichen und ausgeprägten Fällen leicht unterschieden werden.<sup>1</sup> Der einschneidende Gradunterschied beruht hauptsächlich darauf, daß die Sprache in vielfältigster Weise und großem Umfang stetige Tonveränderungen benützt, der Gesang dagegen, wie die Musik überhaupt, prinzipiell nur diskrete Tonstufen, während stetige Übergänge hier nur als Vortragmodifikationen in Anwendung kommen, die in der entwickelten Kunst das deutliche Hervortreten der festen Tonstufen in

<sup>1</sup> Es ist also logisch-formell ein ähnliches Verhältnis, wie ich es zwischen „Empfindungen“ und „Vorstellungen“ eingehend nachzuweisen versuchte.

keiner Weise beeinträchtigen dürfen. Die Gefühlswirkungen der Sprache knüpfen sich wesentlich gerade an stetige Tonveränderungen und wo sprunghafte gebraucht werden, kommt es auf einen Viertelton höher oder tiefer nicht im mindesten an, während die musikalischen Gefühlswirkungen in erster Linie an feste Intervalle fester Töne geknüpft sind und kleine Abweichungen von diesen gerade am unangenehmsten empfunden werden.

Mit diesen primären Unterschieden hängen weitere, wie der gröfsere Anteil an Geräuschen, die geringere Deutlichkeit der Tonhöhen bei unakzentuierten Silben, die gröfsere Freiheit des zeitlichen Verlaufes in der gewöhnlichen Sprache zusammen.

Noch verschiedener als die Lauterscheinungen im Hörenden sind die subjektiven Einstellungen des Sprechenden und des Singenden. Sie sind verschieden nach den allgemeinsten Intentionen, wie nach den speziellen Bewusstseinsunterlagen. Aber auch hier können gegenseitige Annäherungen Platz greifen.

---