

(Aus dem Psychologischen Institut der Universität Wien)

Stimme und Persönlichkeit

Von

HERTA HERZOG (Wien)

(Mit 10 Abbildungen im Text)

Inhaltsverzeichnis

	Seite
1. Abschnitt: Methodische Überlegungen zum Problem	300
2. Abschnitt: Das Massenexperiment	305
Kapitel 1: Seine technische Ausgestaltung	305
Kapitel 2: Das eingelaufene Material	310
Kapitel 3: Die Ergebnisse des Experiments	312
1. Die physiologische Bedingtheit der Stimme	314
2. Die Umweltsbedingtheit der Stimme	327
3. Die charakterologische Bedingtheit der Stimme	335
4. Die Stimme der Sprecher	343
Kapitel 4: Der Einsender und seine Antwort	346
3. Abschnitt: Das Stimmerlebnis	350
Kapitel 1: Methodisches	350
Kapitel 2: Die Struktur des Stimmerlebnisses	351
Kapitel 3: Der Inhalt des Stimmerlebnisses	353
Kapitel 4: Das auffällige Moment	355
Kapitel 5: Die direkte Deutung	357
Kapitel 6: Die indirekte Deutung	363
1. Das Auftreten der Einfälle	363
2. Die personale Bestimmtheit der Einfälle	364
3. Die Verwertung der Einfälle	367

1. Abschnitt

Methodische Überlegungen zum Problem¹

In den natürlichen Situationen des commercium individuorum (hier A — dort B) sind alle Sinne damit beschäftigt, Kontaktfaktoren als Steuerungsprinzipien des gegenseitigen Verhaltens

¹ Die erste Anregung zu dieser Untersuchung ging von T. H. PEAR, Voice and Personality, London 1931, aus, der vom englischen Radiosender aus Experimente mit unbekanntem Sprechern machte. Die vorliegende Arbeit ist eine erste Erweiterung des PEARschen Ansatzes, der eine Reihe

wahrzunehmen. Durch die Errungenschaften der Technik wurde eine künstliche Erweiterung des Kontaktbereiches geschaffen. Diese brachte eine Differenzierung des Kontaktes mit sich, die sich psychologisch als eine Differenzierung der Anforderungen an die einzelnen Sinne auswirkt: Telephon und Radio einerseits, der stumme Film andererseits versetzen den B in die Lage eines Blinden oder eines Tauben. Das Problem des Ausdrucks wird durch diese technischen Fortschritte (von B aus) genau so in den Vordergrund des Interesses gerückt wie es auf künstlerischem Gebiet (von A aus) durch den Expressionismus geschah.

Aber nicht nur von anderen Gebieten her, auch von der Psychologie selbst wird das Ausdrucksproblem drängend, indem das, was im Ausdruck zum Vorschein kommt, die Persönlichkeit, ihre Erlebnisse heute im Mittelpunkt der Diskussion stehen. Die Persönlichkeitsforschung einerseits, der Behaviorismus andererseits müssen an der Ausdrucksforschung in hohem Maße interessiert sein. Engt letzterer doch den Ausdruck und seinen Geltungsbereich stark ein, indem er das Ausgedrückte teilweise negiert, die Erlebnisse in Zweifel zieht, erweitert doch erstere den Bereich und sucht den Ausdruck der Persönlichkeit nicht nur in den dauernden Merkmalen eines Individuums, wie z. B. im Körperbau, sondern auch in all den flüchtigen Körperbewegungen, wie z. B. in der Schreibbewegung oder im Gang. Wenn wir im Folgenden den Ausdrucksgehalt der menschlichen Stimme untersuchen, so kommen wir dazu von einer dritten, ganz anders gearteten psychologischen Problemstellung, von der Sprachpsychologie KARL BÜHLERS und einem ihrer Fundamentalsätze: eine Funktion der Sprache ist Ausdruck.

Die Frage, die wir uns nun stellen, lautet: Physiognomik der Stimme oder, exakter formuliert: Inwieweit ist die Stimme eines Sprechers für den Hörer Ausdruck seiner Persönlichkeit? Denn diese Zweiheit von Sprecher — Hörer, das konstituierende Moment jeder Kontaktsituation, ist für den Ausdruck und seine Zeichen-

weiterer experimenteller Untersuchungen folgen werden, deren Anordnung in diesem Text bereits mitgeteilt ist. Die ganzen Arbeiten stehen unter der Leitung von Herrn Professor KARL BÜHLER und fügen sich in den Rahmen seiner Sprachtheorie (siehe unten) ein. Die Verfasserin ist ihm für seine Führung und fortlaufende Beratung außerordentlich zu Dank verpflichtet. Auch Herrn Dr. PAUL LAZARSELD danke ich für seine Hilfe bei der Ausarbeitung der Ergebnisse.

funktion charakteristisch. Ausdruck an sich ist sinnlos, er bekommt Bedeutung erst, wenn jemand da ist, der etwas als Ausdruck erlebt. Und so wollen wir zur Untersuchung unserer Frage auch nicht beim Sprecher einsetzen, sondern bei dem, den der Ausdruck, ob er nun aktiv oder passiv ist, in erster Linie angeht, dem Hörer. Dafür gab es prinzipiell zwei Möglichkeiten der Lösung: die quantitative und die phänomenologische. Die quantitative beantwortet die Frage: ob, wie und inwieweit können der Stimme vom Hörer Daten entnommen werden? Im Vordergrund dieser Fragestellung steht das Problem der Richtigkeit. Die Richtigkeit der vom Hörer angegebenen Daten ist uns ein Maßstab für die tatsächliche Ausdruckshaltigkeit der Stimme des Sprechers. Die phänomenologische dreht sich um das Problem: Wie erlebt der Hörer die Stimme des Sprechers? Nicht die Richtigkeit der aus der Stimme entnommenen Daten, sondern ihre Erlebnisweise steht hier im Mittelpunkt der Betrachtung.

Welche methodischen Konsequenzen ergeben sich aus diesen zwei möglichen Lösungsarten? Die quantitative verlangt, wie wir sagten, eine Methode, mit deren Hilfe es möglich ist, abzuschätzen, welche Daten und inwieweit sie richtig der Stimme entnommen werden können. Dies ermöglicht die Methode des Massenexperiments. Es besteht darin, daß dieselbe Stimme, ohne daß der Sprecher sichtbar ist, von möglichst vielen Hörern gehört und gedeutet wird. Also: ein oder wenige Sprecher — sehr viele Hörer. Es werden besser einige Sprecher sein, um der Bedeutung individueller Modifikationen einigermaßen begegnen zu können. — Der Sprecher muß unsichtbar sein, um der Reinheit der Stimmwirkung willen. Das Zusammenwirken von akustischem und optischem Ausdruck des Sprechers für die Beurteilung des Hörers ist ein weiteres Problem, das einer gesonderten Untersuchung bedarf. Ihm wurde bereits in einer zweiten Arbeit von HELMUTH TURSKY¹ nachgegangen. — Es müssen endlich viele Hörer sein: nur die Verschiedenheit der Hörer nach Geschlecht, Alter, sozialer Stellung, Charakter . . . berechtigt uns in ihrer Gesamtheit zu einigermaßen allgemeinen Aussagen darüber, wie weit die Ausdruckshaltigkeit der menschlichen Stimme reicht.

Dabei ist es offensichtlich notwendig, dieser größtenteils ungeschulten Masse von Laien irgendwelche allgemein geläufige und

¹ HELMUTH TURSKY, Zur Phänomenologie des Zuordnungsaktes zwischen Stimme und Bild des Sprechers. Wiener Dissertation 1932.

verständliche Handhaben zur Formulierung ihrer Ergebnisse zu geben, darauf hinzuweisen, worauf es ankommt und in welchen Terminis man davon sprechen soll. Denn handelt es sich bei den Teilnehmern des Massenexperimentes zwar größtenteils um Laien in bezug auf eine wissenschaftlich exakte Methode der Selbstbeobachtung (die man daher von ihnen nicht verlangen darf), so stellt die Versuchssituation doch andererseits an sie keine Anforderungen, denen sie nicht gewachsen wären. Im Gegenteil: was das Massenexperiment von den Hörern verlangt, ist etwas, was im Leben des Alltags stündlich geübt wird. Das Kind deutet die Stimme der Mutter, der Untergebene die des Chefs und umgekehrt; und nicht nur die Stimme des bekannten Menschen wird ihrem Ausdrucksgehalt nach ausgewertet, sondern auch die des „neuen“ Kollegen, Chefs, . . . Unsere Aufgabe ist also keine solche, die die Kräfte der Befragten übersteigt. Sie bringen dazu genau so eine Schulung mit wie sie der Psychologe für die zweite mögliche Methode mitbringt. *Die von Kasper, Lautschod - 6/2 Jg*

Neben die quantitative stellten wir ja als zweite die phänomenologische Lösungsmöglichkeit. Als geeignete Methode dafür erscheint die geschulte, detaillierte, subjektive Einzelanalyse. Größtmögliche Unvoreingenommenheit, möglichste Universalität der Phänomenologie sind die Hauptbedingungen. Hier handelt es sich nicht um die objektive Feststellung, welche Daten der Stimme und inwieweit sie richtig erkannt werden können, sondern darum, was im Hörer beim Hören der Stimme vorgeht. Dabei tritt der Gesichtspunkt der Richtigkeit naturgemäß zurück und der einer möglichst treuen und vollkommenen Beschreibung des Erlebnisses in den Vordergrund. Dafür genügen eine relativ geringe Zahl von Einzelanalysen, von einem oder einigen wenigen geschulten Selbstbeobachtern an einer Reihe von gehörten Stimmen durchgeführt. Also: Mehrere Sprecher — ein oder einige wenige Hörer.

Natürlich bedingen sich die beiden Methoden, das Massenexperiment und die Einzelanalyse, gegenseitig, eine befruchtet und ergänzt die andere. So sind die Gesichtspunkte für die Anlage des Massenexperimentes nur aus der Selbstbeobachtung zu gewinnen; denn nur aus der detaillierten Analyse des geschulten Beobachters ergibt sich der Maßstab dafür, wo man die undetaillierte Antwort des Laien braucht und wie man es anstellen muß, daß man sie erhält. Umgekehrt ist die Verifikation für das Einzelexperiment natürlich nur durch ein Massenexperiment zu erhalten.

Nur wenn es sich zeigt, daß das Massenexperiment im Groben die Ergebnisse der Selbstbeobachtung bestätigt, erscheint auch die Auswertung der subtilen Einzelanalyse gerechtfertigt.

Die vorliegende Arbeit sucht in einem *dreifachen* Ansatz diesen methodischen Überlegungen gerecht zu werden. Der *erste* Ansatz sind Selbstbeobachtungen, die beim Anhören von 30 Stimmen im Radio angestellt wurden und sich um eine möglichst vollkommene Phänomenologie bemühen. Man kann gegen die Radioübertragung Einwände erheben — ich werde später noch darauf zu sprechen kommen — aber es handelte sich bei diesen Selbstbeobachtungen ja auch darum, Gesichtspunkte für ein Massenexperiment zu gewinnen und dafür erschien die Radioübertragung technisch sehr günstig. Wollte ich die Auswertung nicht unnötig komplizieren, so mußten in beiden die gleichen Bedingungen herrschen. Es war eine der Aufgaben der folgenden Arbeiten¹, diese Bedingungen zu variieren und zu verbessern.

Durch die Auffindung einiger Grunddimensionen nun, die in jedem Stimmerlebnis wiederkehrten und immer wieder für die Charakterisierung der Stimmträger herangezogen wurden, ergab sich die Notwendigkeit eines *zweiten* Ansatzes, der für diese Grunddimensionen eine breitere objektive Basis schaffen sollte. So stellt die zweite Stufe der Untersuchung das Massenexperiment dar. Es sollte zu den qualitativen Ergebnissen der Einzelanalyse die quantitative Antwort geben. Die einzig mögliche Methode war ein Fragebogen, in dem die aufgefundenen Dimensionen so zu formulieren waren, daß der Laie verstand, worauf es ankam, aber in keiner Weise durch den Fragebogen in seiner Unbefangenheit irritiert wurde und die Möglichkeit zu spontanen Erweiterungen hatte.

Die Antworten auf den Bogen aber zeigten, daß man mit einer statistischen Verarbeitung nicht auskommen konnte. So wurde in einem *dritten* und letzten Ansatz eine phänomenologische Klärung der statistischen Ergebnisse durch das eigene Erlebnis vorgenommen. Außer den eigenen Einzelanalysen wurden dazu auch — allerdings mit großer Vorsicht — die ziemlich zahlreichen

¹ H. TURSKY (a. a. O.) und ST. MARIA BONAVENTURA BARTA, Experimentelle Untersuchungen zum Problem der Resonanz (i. Vorb.) haben ihre Vpn. auf Grammophonplatten aufgenommen; Ing. N. THUMB, Stimme und Körperbau (i. Vorb.) arbeitet mit lebenden Stimmen hinter einem Vorhang.

Zusätze der Hörer zum Fragebogen verwendet, soweit sie ernst zu nehmen waren und Phänomenologisches enthielten.

Damit ist der tatsächliche Gang der Untersuchung skizziert. Für die folgende Darstellung aber wollen wir aus praktischen Gründen alles Phänomenologische, das Gegenstand unseres ersten und dritten Ansatzes war, zusammenfassen und in einem ersten Teil die quantitativen, in einem zweiten Teil die phänomenologischen Ergebnisse behandeln.

2. Abschnitt

Das Massenexperiment

Kapitel 1: Seine technische Ausgestaltung

Das Experiment fand am 19., 21. und 23. Mai 1931 im Radio Wien statt. Die *Durchführung im Radio* brachte, wie schon erwähnt, den Vorteil, daß man hier mit der Beteiligung einer Masse rechnen konnte, in der nahezu Alles vertreten war, die also einen durchaus maßgeblichen Ausschnitt aus der Hörerwelt darstellte. Außerdem bot die Radioübertragung eine völlig natürliche Versuchssituation. Die Nachteile einer Radioübertragung, die vor allem in Unreinheiten der Wiedergabe bestehen, brauchen demgegenüber nicht erst hervorgehoben zu werden. Wir haben ihnen übrigens dadurch einigermaßen zu begegnen gesucht, daß wir im Fragebogen die Art des Radiogerätes, mit dem gehört wurde, aufnahmen. Es zeigte sich, wenn ich das gleich vorwegnehmen darf, daß in bezug auf die Richtigkeit zwischen den 62% der Antworten, die von Leuten mit Lautsprechern stammten, und den 38% der Detektorhörer kein Unterschied besteht.

Ein *zweiter* Punkt der Versuchsanordnung, den die folgenden Arbeiten variierten, war die *Textwahl*. Wir haben uns zu einem vorgegebenen Text entschlossen, der von den einzelnen Sprechern einfach vorgelesen wurde, und zwar aus mehreren Gründen. Die Unbefangenheit des Sprechers, besonders der unteren Schichten, wurde durch den Text, für den er nicht verantwortlich war, möglichst gewahrt. Andererseits wurde durch den gleichen unpersönlichen Text der Darstellungsgehalt des Gesprochenen für den Hörer bei seiner Deutung weitgehend ausgeschaltet, sicherlich in einem viel höheren Maß als es bei Freisprechen möglich gewesen wäre.¹

¹ Sr. M. BONAVENTURA BARTA a. a. O. ließ ihre Sprecher in einer natürlichen Sprechsituation (Ansuchen um Gehaltsaufbesserung) auf Grammo-

Diesen Vorteilen gegenüber mußten wir allerdings den Nachteil einer geringeren seelischen Beteiligung des Sprechers in Kauf nehmen. Immerhin wurde der Text, eine Zeitungsnotiz, die einen interessanten Fall enthält, so gewählt, daß innerhalb der Determination noch genug Spielraum für den einzelnen Sprecher blieb:

„Achtung! Achtung! Ich mache eine Mitteilung. Ein wertvoller Hund hat sich verlaufen. Eine große Belohnung ist ausgesetzt! Alle Hundekenner und Hundefreunde werden aufgerufen. Auch die Gendarmen und Wachebeamten bitte ich um ihre Mithilfe. Ich bin erst vor drei Tagen aus Hamburg zurückgekommen und habe von dort den Hund Lux mitgebracht. Lux ist durch seine Leistungen bei der Aufspürung des Mörders Alexanders berühmt geworden. Er sollte nun uns helfen. Ich habe Lux zwei Tage bei mir im Zwinger gehalten und bin gestern mit ihm in den Prater gegangen, um ihn dort arbeiten zu lassen. Er war durch die lange Reise ermüdet. Deshalb gaben wir ihm zuerst eine leichte Aufgabe, ein Versteck aufzuspüren. Wer beschreibt unseren Schreck, als Lux nicht mehr zu uns zurückkehrte. Er hat sich verlaufen und in der fremden Stadt den Weg nach Hause nicht mehr gefunden. Lux ist ein starker Dobermann, vollständig reinrassig, mit einem glänzenden schwarzen Fell, sein Schwanz und die Ohren sind gestutzt. Er ist außerordentlich rasch, kann auf Bäume klettern und über 2 bis 3 m hohe Planken springen. Maulkorb und Halsband hatten wir ihm bei Beginn der Arbeit abgenommen. Es besteht die Gefahr, daß das kluge Tier den Versuch macht, zu seinem früheren Herrn zurückzukehren. Wenn jemand den Hund findet, wird er gebeten, ihn sehr rasch zurückzustellen, da Lux darauf dressiert ist, von Fremden kein Futter anzunehmen und Gefahr besteht, daß der Hund neben gefüllten Schüsseln zugrunde geht. Es hätte keinen Sinn, den Hund zu unterschlagen, denn seine Fähigkeiten können von einem Fremden nicht ausgenützt werden. Wer ihn aber zurückbringt, erhält eine hohe Belohnung.“

Ein *drittes* Moment der Versuchsanordnung, das variierbar ist und in den weiteren Arbeiten konsequent und mit besonderer Sorgfalt variiert wurde, ist die *Auswahl der Sprecher*. In dieser ersten Arbeit erfolgte sie naturgemäß noch ziemlich willkürlich, bloß unter dem Gesichtspunkt möglicher Vielfältigkeit, indem wir uns zwar auf Wiener beschränkten, aber hier Leute verschiedener Herkunft, verschiedenen Alters, verschiedenen Aussehens und natürlich auch verschiedenen Geschlechtes wählten. Die Auswahl erfolgte nicht nach der Stimme, sondern nach dem allgemeinen Typus des Sprechers. Es folgen nun die Daten und ein Bild der Sprecher in der Reihenfolge, in der sie zu Worte kamen; phonplatten aufnehmen. Hier ist es allerdings schwerer, bei der Deutung vom Darstellungsgehalt zu abstrahieren.

wir drucken hier die Bilder mit Genehmigung von „Radio-Wien“ ab.

1. Abend, 19. Mai 1931.

1. Sprecher: Privatdozent, geboren zu Wien 1883 (also 48 Jahre alt). Größe: 1,78 m. Gewicht: 71 kg. Hält sehr viel Vorlesungen und Vorträge; leitet die Arbeiten von Studenten. Hat etwa 6 Jahre im Ausland verbracht.

2. Sprecher: Mittelschülerin, geboren zu Dresden 1917 (also 14 Jahre alt). Größe: 1,63 m. Gewicht: 53 kg. Aus sehr intellektuellem Milieu; sehr gute Schülerin und von ihren Kolleginnen anerkannt.

3. Sprecher: Chauffeur, geboren zu Wien 1901 (also 30 Jahre alt). Größe: 1,75 m. Gewicht: 98 kg. Bekannt guter Fahrer und draufgängerisch.

2. Abend, 21. Mai 1931.

4. Sprecher: Besitzerin einer Bastwerkstätte, geboren zu Wien 1885 (also 46 Jahre alt). Größe: 1,75 m. Gewicht: 72 kg. Tochter und Gattin eines Universitätsprofessors. Ist selbst Doktor der Mathematik, jetzt auf kunstgewerblichem Gebiete tätig.

5. Sprecher: Katholischer Priester, geboren in Deutschböhmen 1873 (also 58 Jahre alt). Größe: 1,76 m. Gewicht: 96 kg. War auch zeitweilig Lehrer und betätigt sich jetzt literarisch. An allen Stellen seines öffentlichen Auftretens vertritt er tolerante Ideen.

6. Sprecher: Kaufmann, geboren in einem ungarischen Ort 1876 (also 53 Jahre alt). Größe: 1,67 m. Gewicht: 74 kg. Seit seinem 14. Lebensjahr in Wien in verschiedenen Branchen tätig. Seit 1914 Kaffeehausbesitzer und von Einfluß in der Genossenschaft.

3. Abend, 23. Mai 1931.

7. Sprecher: Mittelschüler, geb. zu Wien 1919 (also 12 Jahre alt). Größe: 1,57 m. Gewicht: 56 kg. Aus gepflegtem Milieu, eifriger Radfahrer, will Ingenieur werden.

8. Sprecher: Stenotypistin, geboren zu Wien 1904 (also 27 Jahre alt). Größe: 1,72 m. Gewicht: 74 kg. Steht seit ihrem 17. Lebensjahr im Berufsleben. Arbeitet jetzt speziell in der Organisationsabteilung einer großen Bibliothek.

9. Sprecher: Volksschullehrer, geboren zu Wien 1892 (also 39 Jahre alt). Größe: 1,77 m. Gewicht: 72 kg. Gewerkschaftlich tätig, studiert in seiner freien Zeit, beteiligt sich an verschiedenen wissenschaftlichen Versuchen.



Abb. 1. Vp. I.



Abb. 2. Vp. II.



Abb. 3. Vp. III.



Abb. 4. Vp. IV.



Abb. 5. Vp. V.



Abb. 6. Vp. VI.



Abb. 7. Vp. VII.



Abb. 8. Vp. VIII



Abb. 9. Vp. IX

Der eigentliche Versuch wurde von einem Einführungsvortrag Prof. BÜHLERS eingeleitet. Außerdem sprach Dr. LAZARSELD an jedem der drei Tage einleitende Worte, die die Instruktion an die Hörer enthielten. Der Versuch selbst verlief dann so, daß je drei der Sprecher an den drei schon erwähnten Tagen im kleinen Studio der „Ravag“ in Wien ihren maschineschriebenen Text, den sie vorher zum Vermeiden etwaiger Lesefehler einmal durchgelesen hatten, vorlasen. Die Sprecher waren immer allein im Studio, einer konnte den anderen weder sprechen sehen noch hören. Sprechzeit war je zwei Minuten, dann folgte eine zweiminütige Pause, in der Musik gespielt wurde; sie war den Hörern angekündigt worden und sollte zur Niederschrift der Beobachtungen dienen. Die Instruktion an die Hörer bestand nämlich darin, daß sie aufgefordert wurden, auf Grund ihrer Beobachtungen einen Fragebogen auszufüllen. Der nebenstehende Fragebogen wurde zu Beginn jedes Versuchstages diktiert, außerdem war er in der Zeitschrift „Radio Wien“ erschienen.

Kapitel 2: Das eingelaufene Material

An dem Experiment haben sich 2700 Hörer beteiligt, von denen freilich nicht alle die ganze Versuchsreihe mitmachten. Wir erhielten von ihnen im ganzen 5135 Einzelantworten.

Ein paar Daten zu den Einsendern: Unter den 2700 Hörern waren die beiden *Geschlechter* annähernd gleichmäßig vertreten. Es schickten 1383 Männer, 1317 Frauen ein.

Der *Altersaufbau* der Einsender zeigte (absolut und relativ) eine stetige Zunahme der Beteiligung mit zunehmendem Alter, die in der Altersklasse zwischen 50 und 60 Jahren ihren Höhepunkt erreichte, also in der Altersklasse, die unter den aktiv im Leben Stehenden die größte Lebenserfahrung hat.

Der *sozialen Schicht* nach entfiel der Hauptanteil auf die Intelligenzler; die Beteiligung flacht sich in dem Maß ab als der Beruf manueller wird. Es schickten ein unter den *Männern*: 38 % Akademiker gegenüber 7 % Arbeitern. Bei den *weiblichen* Einsendern waren mehr als die Hälfte „Hausfrauen“. Unter den Berufstätigen stehen nicht die Akademikerinnen (12 %), sondern die Lehrerinnen und Beamtinnen (19 %) an erster Stelle, was sich aus dem absoluten Berufsaufbau erklären läßt. Auch hier sinkt das Interesse mit abnehmender Intellektualität: Arbeiterinnen waren nur 4 % am Versuch beteiligt.

FRAGEBOGEN

Wir bitten unsere Hörer, diesen Fragebogen entweder auszuschneiden oder abzuzeichnen und auszufüllen. Die Zusendung des Bogens erbitten wir an die Ravag, Wien, I., Johannesgasse 4b.

Meine Ansichten über die Sprecher bei dieser Umfrage sind die folgenden:

Sprecher	Geschlecht	Alter	Art des Berufes oder der Beschäftigung	Ist der Sprecher gewohnt, Befehle zu geben?	Wie sieht der Sprecher aus?		Ist die Stimme angenehm?
					Größe	Dicke	
19. Mai	1						
	2						
	3						
21. Mai	4						
	5						
	6						
23. Mai	7						
	8						
	9						

* Man verwende für die Größe die Zeichen: g (groß), m (mittel), k (klein).
Für die Dicke die Zeichen: d (dick), m (mittel), h (hager).

Art des benützten Apparats:

Detektor? Kopfhörer?

Lampenapparat?

Lautsprecher?

Beruf oder Beschäftigung des Hörers?

Alter? Geschlecht?

Wohnort?

Dabei sind die Motive der Beteiligung für Männer und Frauen, die dem gleichen Beruf angehören, völlig übereinstimmend. So sind z. B. innerhalb der Intellektuellen der *Akademiker* und der *Künstler* durch eine mehr skeptische Haltung charakterisiert. Ersterer nimmt am Versuch teil im Bewußtsein seiner eigenen wissenschaftlichen Schulung, die sich an einem „Rätselraten“ beteiligt und mit Kritik nicht spart; letzterer in Ablehnung der Möglichkeit einer wissenschaftlichen Erfassung dieser Frage gegenüber künstlerischer Intuition. (In diesen beiden Haltungen offenbaren sich bereits die beiden extremen Deutungswege.) Im Gegensatz zu Akademikern und Künstlern ist die Einstellung bei den *Lehrern* und *mittleren Beamten* eine mehr positive; sie würdigen den Versuch aus ihren eigenen Erfahrungen heraus und kritisieren, wenn überhaupt, aus der Praxis, während der Akademiker theoretisch kritisiert. Und man kann sagen: Je weniger intellektuell der Beruf wird, um so weniger intellektuelle Bedenken, um so unbedenklicher und positiver wird die Einstellung, ein Befund, der durch die phänomenologische Analyse ohne weiteres verständlich werden wird.

Charakteristisch war auch die *äußere Form* der Einsendungen: so zeigte sich, daß die *weiblichen* Einsender, wenn sie nicht den bereits vorgedruckten Fragebogen benutzen konnten, fast nie ihre Einsendungen in Form einer Tabelle schickten: sondern die *jüngeren* meist schlagwortartig, auch untereinander geschrieben für die einzelnen Vpn., die *älteren* immer in Briefform, in fortlaufenden Sätzen, meist auch ohne neuen Absatz für die Beurteilung eines neuen Sprechers. Die *Zuschriften* der *männlichen* Einsender waren hingegen schon rein äußerlich in der Regel übersichtlicher, sachlicher, aber unpersönlicher, ein Umstand, der sich auch inhaltlich nachweisen ließ. Viele der weiblichen Einsender, besonders die älteren und sozial tiefer stehenden, benützten nämlich die Einsendung, um „der lieben Ravag etwas von ihrer dankbaren Hörerin zu erzählen“, während auch die persönlichen Zusätze der männlichen Einsender größere Sachlichkeit zeigten, irgendwie mit dem Thema zusammenhängen.

Kapitel 3: Die Ergebnisse des Experiments

Schon aus den Einzelanalysen ging hervor, daß sich die Ausdruckshaltigkeit der Stimme auf das Ganze der Persönlichkeit des Stimmträgers erstreckte, also ganz grob genommen den Körper-

bau, das Milieu und die Innerlichkeit angehe, was nach unseren alltäglichen Erfahrungen durchaus nicht als ein völlig überraschendes Resultat zu bezeichnen ist. Man denke nur an sprachliche Bezeichnungen wie: „alte Stimme“, „fette Stimme“, „gütige Stimme“, „Leutnantston“, . . . in denen diese umfassende Ausdruckshaltigkeit aufscheint.

Das Ravag-Experiment nun war vor allem darauf angelegt, die Gruppe der *physiologischen Daten* herauszuarbeiten. Die Gründe dafür: Erstens einmal war es das Naheliegendste, die Stimme als physiologischen Faktor zuerst auf ihren Ausdrucksgehalt im physiologischen Bereich (der noch dazu von allen drei Komplexen am anerkanntesten ist) zu untersuchen, zumal da durch die moderne Körperbauforschung so weitgehende Beziehungen zwischen den einzelnen Körperbaumerkmalen aufgefunden wurden. Zweitens sind gerade die Fragen nach körperbaulichen Daten für eine Massenbefragung geeignet, da sie am leichtesten exakt und eindeutig bestimmbar sind. Auch die Verifikation der Ergebnisse fällt hier leicht. — Wir haben uns im Fragebogen auf ein paar ganz drastische, unmißverständliche Fragen wie Geschlecht, Alter, Größe und Dicke des Sprechers beschränkt und detailliertere Angaben völlig dem Belieben des Hörers anheimgestellt.

Auch die Frage der *Umweltsbedingtheit*, des Milieus des Sprechers haben wir nur in einem Punkt des Fragebogens angeschnitten, nämlich in der Frage nach dem Beruf, von dem wir aus Erfahrung wußten, daß hier tatsächlich Affektionen der Stimme nachweisbar wären. Auch hier wurden tiefergreifende Angaben dem spontanen Zutun des Hörers überlassen.

In einem noch viel höherem Maß mußten wir es dem Hörer freistellen, ob und was er an *charakterologischem* Ausdrucksgehalt der Stimme entnahm. Der Charakter, besser gesagt die Innerlichkeit, ist heute ein so umstrittenes Problem, daß wir es nicht in den Vordergrund der Befragung schieben wollten und konnten, wobei die charakterologische Ausdruckshaltigkeit der Stimme keineswegs unterschätzt wurde. Es ist im Gegenteil klar, daß in der natürlichen Kontaktsituation der wahrnehmbare Ausdrucksgehalt vor allem nicht ohne weiteres Wahrnehmbares, also die Innerlichkeit Betreffendes angeht und daß dieses durch die Übung des Alltags geschulte Erfassen der spezifischen charakterologischen Ausdruckshaltigkeit sich auch in der Versuchssituation durchsetzen würde. Der Punkt des Fragebogens: „ist die Stimme des Sprechers

sympathisch“ war nicht nur Anlaß zu näherer Beschreibung der Stimme, sondern er wurde auch durch sehr zahlreiche charakterologische Angaben spontan erweitert, ebenso wie die Frage nach der „Befehlsgewohnheit“, die einerseits erweiternde Angaben betreffs des Milieus, andererseits charakterologische Zusätze brachte.

Die Ergebnisse stützen sich daher erstens auf die Antworten zu den entsprechenden Punkten des Fragebogens, zweitens auf die spontanen Zusätze der Einsender. Durchschnittlich hat jeder der 2700 Einsender 6 Zusätze gemacht (wobei jede Aussage, auch wenn sie denselben Punkt betraf, als ein Zusatz gezählt wurde). Auf die einzelnen Gruppen verteilen sich die Zusätze folgendermaßen:

Tabelle 1

Die prozentuelle Verteilung der zusätzlichen Angaben zum Fragebogen

Zusätze betreffend	in %
Körperbau	25,5
Milieu	11,3
Charakter	46,0
Beschreibung der Stimme	17,2
Summe	100

Mit der Interpretation dieser Tabelle wollen wir uns im Folgenden beschäftigen.

1. Die physiologische Bedingtheit der Stimme

1. Von allen physiologischen Momenten modifiziert das *Geschlecht* die Stimme am eindeutigsten. Das Geschlecht der *erwachsenen* Sprecher wurde eindeutig erkannt. Bei Vp. I und Vp. III kommt überhaupt keine Fehldeutung vor, bei den 5 übrigen je eine. In allen Fällen stammt sie von Hörern, die über starke Störungen beim Apparat klagten, wir dürfen sie also vernachlässigen.

Anders steht es bei den zwei *jugendlichen* Sprechern: die 14jährige weibliche Vp. II wurde in 1,3 % als Bub gedeutet, in 0,15 % wurde ihr Geschlecht als unbestimmbar angegeben, in 98,55 % tatsächlich erkannt. — Der 12jährige männliche Sprecher Vp. VII wurde in 52,5 % als Mädels gedeutet, in 1,3 % konnten die Hörer keine sicheren Angaben machen, in nur 46,2 % wurde er als Bub erkannt.

Fehldeutungen in bezug auf das Geschlecht kommen also nur bei jugendlichen, noch nicht vollentwickelten Vpn. vor. Und zwar ist die geschlechtliche Differenzierung vor der Mutation der Stimme (jedenfalls der männlichen) auffallend gering. Die weibliche 11jährige Vp. in den englischen Experimenten Prof. PEAR'S¹ wurde noch in 8,1 % für einen Buben gehalten, bei unserer 14jährigen ist der Prozentsatz bereits auf 1,3 % gesunken. Bei Knaben sind zwar die Mutationserscheinungen viel krasser, aber auch die Verwechslungsmöglichkeiten viel größer. Unsere Vp. befand sich in dem Vorstadium der eigentümlich hohen, gespannten, daher „kreischenden Stimme“. Wer dieses künstlich in die Höhe Getriebene nicht heraushörte (Gründe dafür: schlechter Apparat, keine Radioerfahrung), deutete die mutationsbedingte kreischende Stimme als milieubedingt und kam zu Urteilen wie: „Naschmarktweib“. Die Deutung erschwerend kam noch hinzu, daß sich die Vp. zu Beginn des Versuches räusperte und dieses Räuspern den reinen, daher hohen Stimmtönen enthielt. Daher stehen hier auch relativ mehr unbestimmte Angaben und Doppeldeutungen.

2. Während also die Stimme durch das Geschlecht sehr stark und beim erwachsenen Sprecher auch völlig eindeutig modifiziert wird, ist das *Alter* aus der Stimme nicht ohne weiteres richtig entnehmbar. Die folgende Tabelle enthält in der ersten Kolonne das tatsächliche Alter unserer Vpn., in der zweiten das durchschnittlich geschätzte Alter, u. zw. in der ersten Abteilung bei richtiger, in der zweiten bei unrichtiger Geschlechtsbestimmung.

Tabelle 2

Die Altersschätzungen zu den Vpn.

Nummer und Beruf der Vpn.	Alter in Jahren	Geschätztes Alter in Jahren	
		bei richtiger Geschlechtsbestimmung	bei falscher
Vp. VII Mittelschüler	12	14; 10	24; 93
Vp. II Mittelschülerin	14	25; 92	16; 21
Vp. VIII Stenotypistin	27	26; 53	
Vp. III Chauffeur	30	44; 87	
Vp. IX Lehrer	39	38; 8	
Vp. IV Akademikerin	46	34; 47	
Vp. I Privatdozent	48	36; 81	
Vp. VI Cafetier	53	34; 74	
Vp. V Priester	58	48; 38	

¹ T. H. PEAR, a. a. O. S. 163 f.

An Allgemeingültigem läßt sich aus diesen Altersschätzungen nur herausholen, daß offensichtlich die Tendenz besteht gegen einen *Mittelwert* hin die jüngeren Sprecher zu über- und die älteren zu unterschätzen. So werden Vp. I, Vp. IV, Vp. V, Vp. VI unter-, Vp. II und Vp. VII überschätzt. Bei den Vpn., deren Alter tatsächlich in diesen mittleren Bereich fiel, erfolgte dann eine überraschend richtige Deutung. Dazu gehören die 27jährige Vp. VII und die 39jährige Vp. IX. Die 30jährige Vp. III, die bedeutend überschätzt wurde, ist hier auszuschalten, weil ihre Stimme durch Aufregung und infolge des erzwungenen Hochdeutsch vollkommen entstellt war.

Auch PEAR¹ konnte übrigens in seinen Ergebnissen diese Tendenz nach dem Mittelwert nachweisen. Bei einem Vergleich der Ergebnisse finden wir auch dort eine Überschätzung der jüngeren und eine Unterschätzung der älteren Sprecher und übereinstimmend mit unseren Ergebnissen eine richtige Schätzung seiner ungefähr 40jährigen Vp. IV. Nur fand PEAR im mittleren Altersbereich eine durchaus ungleichmäßige Schätzung, die er darauf zurückführt, daß hier die Tendenz nach dem Mittelwert zu vernachlässigen sei zugunsten individueller, von der Stimme ausgehender Tendenzen, die er als solche zur Ungenauigkeit der Schätzung definiert. In unseren Ergebnissen wurden, wie schon oben erwähnt, zwei der Sprecher im mittleren Altersbereich richtig getroffen und nur die durch Aufregung völlig entstellte Stimme des Chauffeurs überschätzt. Es scheint also so zu sein, daß die Treffer deshalb erfolgten, weil die allgemeine Tendenz nach dem Mittelwert hier eben eine Tendenz zur individuellen Genauigkeit darstellte, nicht aber, daß sie hier zu vernachlässigen wäre und nur die individuellen Tendenzen wirksam sein sollten. Doch müssen wir uns jeder endgültigen Antwort enthalten, da PEARS Angaben auf einem größeren Material (5 Vpn. dieses Alters gegenüber nur 3 bei uns) fußen.

Im übrigen fanden aber auch wir übereinstimmend mit PEAR neben der Tendenz nach dem Mittelwert *individuelle, von der Stimme ausgehende Tendenzen* für die Schätzung verantwortlich. Die dritte Tendenz, die PEAR anführt, nämlich die zur *Genauigkeit* der Schätzung, konnten wir wohl als ein Bestreben nach möglichst bestimmten Angaben nachweisen, aber sie bestimmt offenbar nicht so sehr das Alter der Schätzung als die Formulierung.

¹ T. H. PEAR, a. a. O. S. 164 ff., S. 174.

Zu der Tendenz nach dem Mittelwert wäre noch zu bemerken, daß ihre Grundlagen genauerer Untersuchung bedürfen. Offenbar liegt sie ja im Wesen jeder Schätzung und wird in unserem Versuch durch die Radioübertragung, die eine dem Hörer vollkommen bewußte Verschlechterung der Schätzungsmöglichkeit bedeutet, besonders nahegelegt. Sie liegt aber möglicherweise auch in dem zu schätzenden Objekt, der Stimme des Sprechers, die tatsächlich sehr bald eine bestimmte Lage erreicht und sich sehr lange in ihr erhält, ein Umstand, der die Annahme eines mittleren Wertes für diese Zeit des Gleichbleibens der Stimme nahelegt; so daß also für diese Tendenz nach dem Mittelwert sowohl der tatsächliche Ausdrucksgehalt der Stimme als auch die Einstellungsart des Hörers, der sich für das Alter eines unbekanntem Menschen zuvörderst nur ganz grob interessiert und daher mit einem Mittelwert begnügt, heranzuziehen wären.

Jedenfalls bleibt innerhalb dieser allgemeinen Tendenz genügend Spielraum für den individuellen Sprecher und den individuellen Hörer. Was die Bedingtheit der Schätzung durch den *individuellen Sprecher* anlangt, so wirkt zwar auch das objektive, aber in noch viel stärkerem Maße das „*subjektive*“ Alter des Sprechers und insofern sind die individuellen Tendenzen, wie PEAR annimmt, auch solche zur Ungenauigkeit der Schätzung. Denn der Hörer antwortet diesen individuellen Faktoren innerhalb seiner allgemeinen Tendenz mit sehr feinen Nuancierungen. So wurde z. B. Vp. VI (53 Jahre) unvergleichlich stärker unterschätzt als Vp. V (58 Jahre); einem objektiven Altersunterschied von 5 Jahren entspricht ein Unterschied in der Schätzung von 13 Jahren. Vp. VI fühlt sich jünger, sieht jünger aus, will nicht alt sein. Besonders am Anfang des Experimentes bemühte er sich recht frisch und lebendig zu sprechen und diesem primären Ausdruckswillen antwortete der Hörer mit dem zu geringen Alter. Bei Vp. V fehlt dieses Bemühen frisch und jung zu erscheinen; er wurde in viel geringerem Maße unterschätzt.

Was die Bedingtheit der Schätzung durch den *individuellen Hörer* anlangt, so handelt es sich hier vor allem um das *Bestreben nach möglichst genauen*, das ist engen *Schätzungen*. Es ist rein äußerlich auch nachweisbar durch Korrektur der einmal gefällten Altersschätzung auf Grund gewisser, während des Sprechens wahrgenommener Veränderungen der Stimme. So heißt es z. B. bei

Vp. VI manchmal: „Ist mir zuerst jünger vorgekommen, dann aber als guter Vierziger, weil die Stimme viel weniger lebendig war.“

Vor allem aber äußert sich dieses Streben nach exakten Angaben darin, daß bei jedem der 9 Sprecher über 60% der Altersangaben auf 1 Jahr genau sind. Wenn aber das Alter in Form eines Intervalles angegeben wird, erweist sich die Tendenz nach dem Mittelwert insofern auch hier als wirksam, als das beliebteste das 5jährige ist (rund 15 % aller Angaben entfallen darauf). Weiter wird bei den älteren Sprechern das 10jährige (bei der 58jährigen Vp. V 10 % gegenüber 4 % bei der 27jährigen Vp. VIII.), bei den jüngeren das 2jährige (8 % bei der 27jährigen Vp. VIII, gegenüber 1,3 % bei der 58jährigen Vp. V) bevorzugt. Auch bei diesem Intervall besteht übrigens die Tendenz mit wenigstens einem der Glieder des Intervalles eine durch 5 teilbare Zahl zu erreichen.¹

Die Verteilungskurve der Altersschätzungen ergibt folgendes Bild (siehe nebenstehende Abb. 10).

Bemerkenswert an diesen Kurven ist einmal die allgemein durchgängige Größe der Streuung: der 53jährige Cafetier (Vp. V) wird ebenso für unter 25 gehalten, wie der 30jährige Chauffeur (Vp. III) für über 60. Ferner, daß die Form der Verteilung eine durchaus ungleiche ist, daß es also neben symmetrischen Verteilungskurven für Vp. I, Vp. II, Vp. VIII auch viel ungleichmäßigere wie z. B. für Vp. V gibt. Die Zweigipfligkeit in der Verteilungskurve von Vp. III ist deutlicher Ausdruck der widerstrebenden Schätzungstendenzen. Schließlich wäre noch hervorzuheben, daß die Kurve der dem Alter nach richtig getroffenen Vp. VIII durchaus vergleichbar ist mit der der überschätzten Vp. III und der stark unterschätzten Vp. I; der Alterstreffer stellt sich also ebenso dar wie die Täuschungen: d. h. offenbar sind beide das Produkt der gleichen oben dargestellten Tendenzen.

Im allgemeinen ist also die Altersbestimmung eine sehr unverläßliche und unsichere, obwohl die Hörer in ihren Angaben relativ sehr bestimmt sind und von Zweifeln über ihre Deutung wenig berichten. Offenbar ist erstens die Stimme für das Alter wenig ausdrucks haltig. Zweitens wird diese geringe Ausdruckshaltigkeit für den unbefangenen Deutenden auch gar nicht subjektiv wichtig und damit auch nicht zum Gegenstand kritischer Über-

¹ Das deckt sich mit Feststellungen von K. MARBE.

legungen. Daß die Möglichkeit zu verschieden starker Determiniertheit der Angaben, (wie sie bei den Geschlechtsangaben, die — objektiv viel richtiger — subjektiv viel mehr in Zweifel gezogen werden, zwar fehlt, bei den Altersangaben aber besteht), kein hinreichender Grund für die subjektive Sicherheit der Angaben des Hörers ist, ergibt sich aus den Antworten zu einer

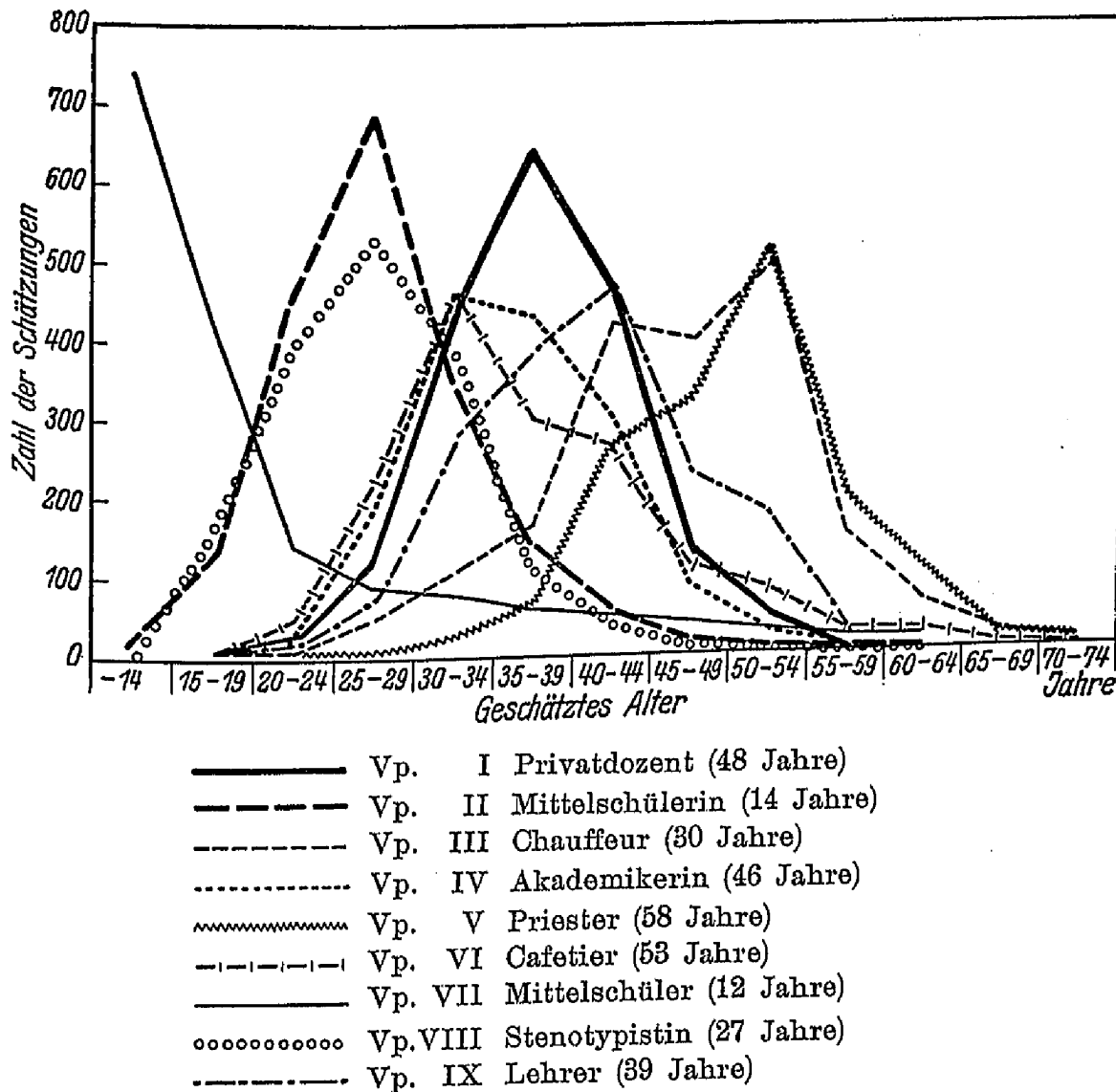


Abb. 10. Die Verteilung der Altersangaben zu den einzelnen Sprechern.

dritten Gruppe physiologischer Daten, die den Körperbau im engeren Sinn angehen.

3. Wir haben im Fragebogen die Frage nach *Größe und Dicke* des Sprechers gestellt und dazu die meisten Fragezeichen, die meisten Hinweise auf mögliche Fehldeutungen, aber auch die detailliertesten Zusätze bekommen, d. h. mit den Größe-Dicke-Angaben wurde offensichtlich für den Hörer sehr Bedeutsames getroffen. Und deshalb ist ihm auch die Frage der Richtigkeit sehr

wichtig. (Übrigens auch, wie wir später noch sehen werden, die Frage der Erlebnisweise dieser Daten.) Bleiben wir aber vorläufig bei dem statistischen Ergebnis. Gestellt war die Frage so: *ist der Sprecher groß, mittel, klein, ist er dick, mittel, dünn?* Dabei ist die Fragestellung allerdings eine sehr grobe und ungenaue, im einzelnen werden natürlich die Ansichten der verschiedenen Hörer bis zu welcher Anzahl von Zentimetern jemand noch als „mittel“ zu bezeichnen ist, auseinandergehen. Aber uns kam es ja vor allem darauf an, ob die Stimme für die Gestalt des Sprechers ausdrückshaltig sei. Und da zeigte sich Folgendes.

Die *Größe* unserer Sprecher wurde mit Ausnahme von Vp. III (aus den schon erwähnten Gründen) getroffen; und zwar nicht nur absolut, sondern auch relativ: d. h. in ihrer Abstufung zwischen den einzelnen Sprechern. Unser größter Sprecher, der Privatdozent, und unser kleinster, der Cafetier, wurden folgendermaßen geschätzt¹:

Tabelle 3

Die Größeschätzungen zum größten und kleinsten unserer Sprecher in Prozent

Nummer u. Größe der Vpn.	groß	mittel	klein	ohne Antw.	Summe in %
Vp. I (1,78 m)	61,3	37,1	1,3	0,3	100
Vp. VI (1,67 m)	29,1	50,3	19,7	0,9	100

Das Maximum liegt also eindeutig in der „richtigen“ Kolonne. Die Angaben in den anderen Kolonnen und der Grad der Streuung ist maßgebend für die Relation der Größen der einzelnen Sprecher. Zur Darstellung der *Größenreihung* verwendete ich einen Koeffizienten, der die Differenz der Abweichungen vom Mittelwert bezogen auf die Gesamtheit aller Angaben enthält. Bezeichnen wir alle Angaben zu „groß“ mit +, alle zu „mittel“ mit =, alle zu „klein“ mit —, so lautet der Koeffizient:

$$\frac{(+)-(-)}{(+)+(=)+(-)}$$

+1 hätte somit ein einstimmig als „groß“ erkannter, —1 ein einstimmig als „klein“ erkannter Sprecher. Die übrigen Deutungen müssen zwischen +1 und —1 liegen.² Die Größenreihung unserer Vpn. auf Grund dieses Koeffizienten ergibt folgendes Bild:

¹ Leider hatten wir keinen ganz kleinen Sprecher, da die in Aussicht genommene kleine Vp. VI knapp vor Beginn des Versuches erkrankte.

² Siehe LAZARSELD, Statistisches Praktikum. Jena 1929.

Tabelle 4
Die Größenreihung der Sprecher

Geschlecht, Nummer u. Beruf der Vpn.			tatsächliche Größe in m	Schätzungswert
männlich:	{	Vp. I Privatdozent	1,78	0,59
		Vp. IX Lehrer	1,77	0,49
		Vp. V Priester	1,76	0,48
		Vp. III Chauffeur	1,75	0,05
		Vp. VI Cafetier	1,67	0,09
weiblich:	{	Vp. IV Akademikerin	1,75	0,39
		Vp. VIII Stenotypistin	1,72	0,33
jugendlich:	{	Vp. II Mittelschülerin	1,63	0,29
		Vp. VII Mittelschüler	1,57	0,31

Mit einer einzigen Ausnahme entspricht die geschätzte Größenfolge der wirklichen. Nur der Chauffeur, Vp. III, erscheint als „mittel“, während er in Wirklichkeit ausgesprochen groß ist. Wie gesagt, hat die Aufregung des Sprechers Schuld an dieser Einschätzung; man erkannte wohl seine Dicke, unterschätzte aber seine Größe. Außerdem spielt hier möglicherweise die typische Vorstellung des „gemütlichen Wieners“ eine hindernde Rolle.

Im Zusammenhang mit der Größenbeurteilung des Chauffeurs tauchte übrigens die Frage auf, ob nicht eine mögliche Fehlerquelle darin liege, daß es bestimmte bevorzugte Größe-Dicke-Kombinationen gäbe, die vom Hörer bei Auftreten des einen Kombinationsgliedes ganz automatisch, ohne stimmliche Grundlage, gesetzt würden. Wenn dem so wäre, müßten in einer 9 Feldertafel, in der wagrecht die drei möglichen Größe-, senkrecht die drei möglichen Dickeangaben stehen, in der also die Größe-Dicke-Kombination für die einzelnen Sprecher eingestrichelt werden kann, sich erstens bei allen Sprechern auffallende Maxima in der- oder denselben Größe-Dicke-Kombinationen finden, es brauchte sich weiter das Maximum für jeden Sprecher nicht mit den durch den Koeffizienten errechneten, unabhängig voneinander dargestellten Größe- und Dickewerten decken. Wir haben eine solche Tafel angelegt und es hat sich gezeigt, daß dem nicht so ist, daß also die Gestaltangaben, d. h. die Kombination zwischen Größe und Dicke nicht rein automatisch gefunden wird. Die Deutung unterliegt im Gegenteil gerade hier viel charakteristischeren Gesetzmäßigkeiten.

Betrachtet man nämlich die Angaben über die *Dicke* der einzelnen Sprecher, so scheint es auf den ersten Blick, als ob sie weniger richtig wären als die Größenangaben. Eindeutig erkannt wurde nur das positive Extrem, nämlich die ausgesprochen „dicken“ Vp. III (Chauffeur) und Vp. V (Priester), während z. B. das negative Extrem, der magerste unserer Sprecher (Vp. I, der Privatdozent) als „mittel“ erscheint.

Tabelle 5

Die Dickeschätzungen zum dicksten und hagersten unserer Sprecher in Prozent

Nummer u. Dickenmaß ¹ der Vpn.	dick	mittel	hager	ohne Antw.	Summe in %
Vp. III (56,0 kg)	73,1	19,2	6,9	0,8	100
Vp. I (39,9 kg)	10,8	51,1	37,9	0,2	100

Umgekehrt wurde z. B. die ausgesprochen mittelstarke weibliche Vp. VIII (74 kg) als „hager“ bezeichnet.

Stellt man ferner analog der Größenreihung eine *Reihung* nach der tatsächlichen und geschätzten *Dicke* (dargestellt durch den gleichen oben zitierten Koeffizienten) auf, so ergibt sich:

Tabelle 6

Die Dickenreihung der Sprecher

Geschlecht, Nummer u. Beruf der Vpn.			Ihr Dickenmaß ¹ in kg	Schätzungswert
männlich:	Vp. III	Chauffeur	56,0	0,67
	Vp. V	Priester	54,6	0,66
	Vp. VI	Cafetier	44,3	0,30
	Vp. IX	Lehrer	40,7	0,06
	Vp. I	Privatdozent	39,9	0,33
weiblich:	Vp. VIII	Stenotypistin	43,0	0,54
	Vp. IV	Akademikerin	41,1	0,24
jugendlich:	Vp. VII	Mittelschüler	35,6	0,1
	Vp. II	Mittelschülerin	32,5	0,52

Danach zeigt sich ein Bevorzugen der Angabe „mittel“, das offenbar zurückzuführen ist auf ein Versagen der Unter-

¹ Als Dickenmaß verwendete ich den Quotienten aus Gewicht (in kg) und Größe der Vpn. (in m) und bekomme so das auf 1 m Größe bezogene Gewicht jeder Vp.

scheidung zwischen „mittel“ und „hager“ (kraß bei Vp. I und Vp. VIII, weniger stark bei Vp. IX und Vp. VI). Man könnte daraufhin versucht sein zu sagen: Es ist klar, daß die „fette“, die „asthmatische“ Stimme von Vp. III und Vp. V richtig nach der Dicke ausgewertet wurde; für weitere und feinere Unterscheidungen ist aber die Stimme nicht ausdrucks haltig, daher bevorzugen die Hörer, wenn diese eindeutigen krankhaften Nebengeräusche in der Stimme fehlen, die Deutung „mittel“, so wie sie beim Alter sichtlich am stärksten durch die Tendenz nach dem Mittelwert bestimmt werden.

Dieser Auslegung widersprechen aber die *Zusätze* der Einsender, die *Gestaltangaben* enthalten. Da diese Zusätze völlig spontan erfolgten und einen Punkt im Detail behandeln, der ohnedies im Fragebogen relativ ausführlich (mit 2 Fragen) vertreten war, gewinnen sie als Interpretation zu den Größe-Dicke-Angaben an Bedeutung (wenn sie auch ihrem vollen Umfange nach vom Leser vielleicht erst später, nach dem phänomenologischen Bericht, gewürdigt werden können). Eine ganze Gruppe von physiologischen Zusätzen (31 %; siehe Tabelle 7) betrifft nämlich *formale Merkmale* des Sprechers, d. h. also *Gestalt-eigenschaften* oder *spezieller Formen von Körperteilen*. Zu letzteren gehören: Angaben über die Form des Kopfes, der Gesichtszüge und Details daran (Stirn, Nase, Zähne . . .), der Schultern, des Brustkorbes, der Extremitäten. Es zeigte sich dabei, daß diese Gestaltangaben völlig im Sinn der KRETSCHMERSCHEN Typen¹ formuliert waren, d. h. immer wieder die drei typischen Gegensätze: schmal — lang, hoch — derb, breit — rund hervorhoben und danach die einzelnen Sprecher richtig charakterisierten.

So erscheinen der Privatdozent (Vp. I) und die Mittelschülerin (Vp. II) als Leptosomen, als „grazil“, „schmal“, „mit markantem Kopf“, „mit langen, schmalen Händen“, „mit edlen Bewegungen“ . . ., der Chauffeur (Vp. III), der Priester (Vp. V), der Cafetier (Vp. VI), der Mittelschüler (Vp. VII), die Stenotypistin (Vp. VIII) als Pykniker, (natürlich in den verschiedensten Legierungen), also als „breit“, „mit Bauch“, „ungelenk“, „kurzgliedrig“, „mit rundem Kopf“, „mit vollem Gesicht“, „mit runden, breiten Schultern“, „mit breitem Brustkorb“, „mit runden, weichen Händen“, schließlich die Akademikerin (Vp. IV) und der Lehrer (Vp. IX)

¹ E. KRETSCHMER, Körperbau und Charakter, 7. u. 8. Aufl., 1929, S. 17 ff.
21*

als Athletiker mit „eckigen Formen“, „männlichen, derben Zügen“, „starken Zähnen“, „großen, derben Händen“. Dabei herrscht in der Charakterisierung einige Unsicherheit nur bei dem Cafetier (Vp. VI), offenbar zufolge seiner „veränderlichen Stimme“, die auch bei der Altersbeurteilung Schwierigkeiten machte. Im übrigen hatten es ja die Hörer nicht mit reinen Körpertypen zu tun und die Streuungen in den nicht den Haupttypus charakterisierenden Rubriken geben in ihrer Abstufung ein — allerdings nicht sehr genaues — Maß der „Legierung“. Aber die Deutung strebt vor allem, wie wir später noch zeigen werden, nach dem Charakteristischen, dem Grundtypus und der kommt in den Zusätzen völlig richtig heraus.

Interpretieren wir nun mit diesen Zusätzen, die allerdings von einer geringeren Zahl von Hörern stammen, aber offensichtlich von solchen, die den Versuch genau und aufmerksam mitmachten, die statistischen Ergebnisse des Fragebogens, so zeigt sich, daß in der Gruppe der „mittelstarken“ die verschiedensten Typen zusammengefaßt werden, verschieden nicht nur vom objektiven Standpunkt, sondern auch in der Erlebnisweise der Hörer. So wird zwar der Cafetier (Vp. VI) als Pykniker, der Lehrer (Vp. IX) als Athletiker, der Privatdozent (Vp. I) als Leptosomer erfaßt und in Einzeldaten auch so charakterisiert, in den allgemeinen Dickeangaben aber werden alle drei als „mittel“ bezeichnet. Offenbar liegen also bei den Dickeangaben die Verhältnisse viel komplexer als bei der Größe, d. h. die Frage: „ist der Sprecher dick, mittel, hager“ ist zu wenig differenziert als daß darin das Gestalterlebnis des Hörers ausgedrückt werden könnte. So wird der Hörer zur Oberflächlichkeit verführt: einerseits läßt er sich von dem groben Eindruck der „asthmatischen Stimme“ leiten und ist geneigt, jede Stimme ohne diese charakteristischen Nebengeräusche als „mittel“ zu bezeichnen, weil eine Differenzierung der Dickeangaben erst von der Gestalt her für ihn bedeutsam wird und andererseits läßt er Überlegungen zur Geltung kommen, die mit dem Ausdrucksgehalt der Stimme für die Frage, die er beantworten soll, nur mehr in sehr losem Zusammenhang stehen: so heißt es bei Vp. VIII häufig: „ist eine moderne Frau, daher sicher schlank“ oder bei Vp. II: „ist noch jung, daher wohl auch schlank“. Der scheinbare Mißerfolg der Dickeangaben liegt also vom Hörer aus gesehen in der zu groben Fragestellung begründet, darin, daß die Stimme Ausdruck für viel detailliertere Gestaltmomente ist.

In den Gestaltdetails ist aber nur eine Seite an den Zusätzen der Einsender, soweit sie den Körper des Sprechers angehen, getroffen, nämlich das formale Moment. Ein weiteres Moment bringen alle die Zusätze, die sich mit dem *Farbeindruck* beschäftigen. Sie machen prozentuell sogar einen viel größeren Teil der Zusätze aus. Erstens, weil im Fragebogen überhaupt nicht danach gefragt war. Zweitens ergänzt der Hörer, der durch den Ausdrucksgehalt der Stimme den lebendigen Sprecher erlebt, natürlich spontan vor allem das Moment an der Wirklichkeit, das gegenüber einer natürlichen Kontaktsituation am stärksten vernachlässigt wurde, nämlich das Farbmoment. Denn die Gestaltqualität haftet auch dem Akustischen primär an, die Farbqualität nur auf Grund synästhetischer Übertragung. Dazu kommt drittens, daß, wie wir später sehen werden, die stimmliche Grundlage für alle Farbangaben vor allem im Stimmtimbre liegt und dieses unter den Stimmdimensionen die wichtigste Rolle spielt.

Wie sind nun diese optischen Zusätze formuliert? Sie betreffen Angaben über Haare, Augen und Teint der einzelnen Sprecher. Und zwar nehmen sie in der Gesamtheit aller physiologischen Zusätze (3886) folgenden Platz ein:

Tabelle 7

Die prozentuelle Verteilung der physiologischen Zusätze

Zusätze betreffend	in %
Haare, Augen, Teint	58
Formen von Körperteilen	17
Gestalteigenschaften	14
Körperliche Zustände	11
Summe	100

Die Angaben der letzten Gruppe betreffen sowohl den Gesundheitszustand der Stimme, der Atmungsorgane usw. als auch allgemeinere Angaben über den Gesundheitszustand des Sprechers. Natürlich fehlt hier dem Laien für eine eindeutige Zuordnung die Schulung. Aber drastische Veränderungen, das Asthma beim Chauffeur (Vp. III), dem Priester (Vp. V) und dem Cafetier (Vp. VI) kommen doch deutlich heraus. Ebenso wird die mutierende Knabenstimme erkannt, allerdings noch häufiger als „verkühlt“ beurteilt. Im übrigen wirkt sich die Tendenz des Hörers nach möglichst detaillierten Angaben gerade hier, wo es vor allem auf sachliche Schulung

ankommt, am verhängnisvollsten aus und verführt ihn z. B. dazu, die gleichförmige Sprechweise von Vp. II auf „Kränklichkeit“, „Müdigkeit“, die aufregungs- und bildungsbedingte schlechte Sprechweise von Vp. III auf „schlechte Augen“ zurückzuführen.

Mehr als die Hälfte aller Angaben der obigen Tabelle fallen aber in die erste Gruppe. Allerdings sind nur 73 % davon (auf alle physiologischen Zusätze bezogen 42 %) Farbangaben. Daneben gibt es andere Qualitätsbestimmungen, in denen auch das Gestaltmoment eine große Rolle spielt (z. B. Zusätze über die Haarqualität, Haartracht, Form der Augen, Art des Blickes, Qualität der Haut).

Unter den Farbangaben sind die Zusätze über die Haarfarbe (56 %) weitaus am häufigsten. Sie erweisen sich als weitgehend richtig überall dort, wo es sich um die Scheidung blond—dunkel handelt und sind unsicher dort, wo es sich um Altersveränderungen der Haarfarbe handelt, wo also sekundär die Altersbeurteilung in die Deutung hereingenommen wird; in erhöhtem Maß gilt dies auch für Angaben über die Haarquantität, wenn z. B. dem überschätzten 30jährigen Chauffeur eine „Glatze“ zugeschrieben wird.

Die Zusätze über die Augenfarbe sind weniger zahlreich und bestimmt, aber auch sie treffen ziemlich richtig. Es läßt sich allerdings zeigen, daß Haar- und Augenfarbe (mit Ausnahme der älteren „melierten“ Vp. V) in den Angaben völlig übereinstimmen. Und es fragt sich, welche der Angaben die primäre ist. Leider stimmen nämlich bei unseren Vpn. Haar- und Augenfarbe überein. Es scheint allerdings so, wenn wir die Begründungen heranziehen, als ob beide Daten primär und unabhängig voneinander gedeutet würden. Doch müßte man hier jedenfalls Resultate bei blond-dunkeläugigen Sprechern (und umgekehrt) abwarten.

Die Hautfarbe wird getroffen. Es zeigt sich dabei eine Abstufung zwischen den einzelnen Sprechern, die völlig ihrem Körperbautypus entspricht. So erscheint z. B. die zarte Mittelschülerin (Vp. II) als „blaß“, die pyknische Stenotypistin (Vp. VIII) als „gesund gefärbt“, der dicke Chauffeur (Vp. III) als „aufgedunsen“.

Eine ähnliche, den Typus treffende Charakterisierung der einzelnen Vpn. ergibt sich auch in den Angaben über die Haarqualität. So werden z. B. die „glatten“ Haare der leptosomen Vp. II dem „weichen“ Haar der pyknischen Vp. VII gegenübergestellt.

Weitere Angaben über die Haartracht, die die als „intellektuell“ erkannte Vp. I z. B. als „glattrasiert“, den als „Fiaker“ beurteilten Chauffeur (Vp. III) als „schnurrbarttragend“ kennzeichnen, die bei den weiblichen Sprechern viel häufiger sind als bei den männlichen und bei den jüngeren häufiger als bei den älteren, suchen ebenso wie Angaben darüber, ob der Sprecher eine Brille trägt, wie Angaben über milieubedingte Teintveränderungen („gepudert“ bei den Intellektuellen, „tätowiert“ beim Chauffeur), ebenso wie ästhetische Wertungen das Bild, das der Hörer aus der Stimme gewonnen hat, durch typische Details weiter auszuführen. Sie sind, soweit

sie überhaupt ernst zu nehmen sind, weniger wichtig für unsere quantitative Auswertung als für die spätere phänomenologische Untersuchung.

Fassen wir zusammen, so ergibt sich: die Frage nach der Größe und Dicke des Sprechers wurde von den Hörern spontan erweitert durch Details des Aussehens, die oft überraschend richtig waren, daneben allerdings auch arge Fehldeutungen aufwiesen. Als wichtigste Punkte wollen wir zwei herausstellen: 1. die angegebenen Details wiesen in ihrer Spontaneität eine beträchtliche Vielseitigkeit und Lebendigkeit auf. Um ihre Bedeutung wirklich abschätzen zu können, bedarf es offensichtlich einer phänomenologischen Untersuchung über ihr Zustandekommen. 2. Bei allem Nuancenreichtum ergaben sich die Treffer in den Angaben dort, wo Typisches im Sinne der KRETSCHMER'schen Typen herausgehoben wurde. Wir hatten freilich keine reinen Typen, außerdem gehörten die Sprecher verschiedenem sozialen Milieu an. In diesem Punkt können also unsere Ergebnisse nur Anregungen dafür sein, den Zusammenhang zwischen Körperbautypus und Stimme unter verbesserten Bedingungen zu untersuchen. Wenn wir dafür die Forderung aufstellen, daß soziale Unterschiede der Sprecher möglichst ausgeschaltet werden, so erscheint sie zwar a priori einleuchtend, da die Tatsachen des Dialektes (im weitesten Sinne), des Berufstones jedem geläufig sind. Wir können diese Forderung aber auch sehr exakt aus unserem Material belegen, indem wir aufzeigen, inwieweit das Milieu des Sprechers in der Stimme zum Ausdruck kommt.

2. Die Umweltsbedingtheit der Stimme

Der Stimme als dem Vollzugsorgan der Sprache kommt die Funktion zu, Verständigungsmittel im sozialen Kontakt zu sein. Dabei handelt es sich nicht um eine einseitige, sondern um eine wechselseitige Bindung und Beeinflussung, die offenbar auch im Kontaktmittel irgendwie angelegt und nachweisbar sein muß. Unsere Frage geht nun dahin, ob und wie diese zweiseitige Beeinflussung auch in der Stimme nachweisbar ist. Wir wissen, daß wir in der jeweiligen Sprechsituation unsere Stimme nach dem jeweiligen Adressaten verändern oder zumindest verändern können. Inwieweit wird die Stimme von der Gesamtheit der Adressaten, an die sie sich gewöhnlich richtet, beeinflusst, inwieweit ist die Herkunft des Sprechers aus der Stimme entnehmbar? So lautet unser Problem.

1. Wenn wir aus dem ganzen Fragenkomplex den *Beruf* herausgriffen und ihm einen Punkt unseres Fragebogens widmeten, so hatte dies seine guten Gründe. Gegenüber Abweichungen ins Soziologische war der Beruf der Punkt, der die größtmögliche psychologische Reinlichkeit verbürgte. Ferner konnten wir erwarten, daß die Geläufigkeit des Zusammenhanges zwischen Beruf und Stimme die Deutung unterstützen würde und daß die Frage eindeutig und speziell genug gestellt war, um darauf brauchbare Antworten zu erhalten. Die Ergebnisse haben uns recht gegeben.

Tabelle 8

Die Berufsschätzung der erwachsenen männlichen Sprecher in Prozent

Nummer und Beruf der Vpn.	Akad. Stud. fr. Beruf	Beamt. Lehr.	Kaufm. Angest.	Arbeiter Gewerbetreib.	Militär Private ?	Summe in %
Vp. I Privatdozent	44,8	30,4	10,5	0,3	14,0	100
Vp. V Priester	35,2	29,7	18,1	4,9	12,1	100
Vp. IX Lehrer	13,9	35,1	25,6	19,6	5,8	100
Vp. VI Cafetier	33,6	27,3	30,5	2,9	5,7	100
Vp. III Chauffeur	1,3	1,3	17,7	77,7	2,0	100

Tabelle 9

Die Berufsschätzung der erwachsenen weiblichen und der jugendlichen Sprecher in Prozent

Nummer und Beruf der Vpn.	Akad. Stud. fr. Beruf	Beamt. Lehr.	Gesch.-Frau, Di- rektre	Angest. Arbeiter	Hausfr. Priv.	Versch. ?	Summe in %
Vp. IV Akademikerin	26,8	47,9	7,7	2,9	12,9	1,8	100
Vp. VIII Stenotypistin	31,8	44,5	1,9	7,5	11,6	2,7	100
Vp. II Schülerin	30,4	46,9	0,8	9,9	8,6	3,4	100
Vp. VII Schüler	60,3	5,8	2,0	26,6	3,7	1,6	100

Die Scheidung *intellektuell-manuell* ist aus der Stimme eindeutig durchzuführen. Der Chauffeur (Vp. III) wurde von mehr als $\frac{3}{4}$ der Hörer als dem Arbeiterstand zugehörig erkannt, umgekehrt der Privatdozent (Vp. I) als Angehöriger eines Intelligenzberufes. Gewöhnlich geht ja die Treffsicherheit noch weiter: Der Privatdozent (Vp. I), der Priester (Vp. V), der Mittelschüler (Vp. VII), der Lehrer (Vp. IX) und die Stenotypistin (Vp. VIII) wurden wirklich mit der größten Prozentzahl (zwischen 35 % und 45 %) als Akademiker, Mittelschüler, Lehrer und Beamtin erkannt. Man

muß hier allerdings vorsichtig sein; denn es wäre denkbar, daß ein Intellektueller, der stark Dialekt spricht, von vielen Hörern als Angehöriger eines manuellen Berufes bezeichnet würde. In unserem Material liegt in der Überschätzung des Cafetiers (Vp. VI) der entgegengesetzte Fall vor. Da er als Jude ein ziemlich dialektfreies Deutsch sprach, wurde er in der Wertung bedeutend hinaufverschoben. Er erscheint mit 33,6 % als Akademiker, mit 60 % als Intellektueller und als Kaufmann nur in 30,5 %. Die Resultate beim intellektuellen weiblichen Sprecher sind weniger gut als beim männlichen. Sie sind schablonenmäßiger, mit deutlicher Bevorzugung der Berufskategorie „Beamtin—Lehrerin“. Offenbar ist für die als intellektuell erkannte Frauenstimme diese Kategorie am geläufigsten. Es läge nahe, eine Begründung dafür darin zu suchen, daß die Frau auf vielen Berufsgebieten erst relativ kurze Zeit tätig ist und den Hörern einfach die Erfahrung, die, wie wir noch zeigen werden, gerade hier sehr wichtig ist, fehlt.

Im übrigen haben in obiger Tabelle nicht nur die Prozentzahlen in der „richtigen“ Berufsklasse Bedeutung: ihre Nuancen liegen gerade in den „falschen“ Angaben und ihrem gegenseitigen Vergleich. So fällt beim Lehrer (Vp. IX) die relativ sehr starke Besetzung der Rubrik „Arbeiter“ auf: darin findet ein starker Einschlag, den der Lehrer tatsächlich hat und der teilweise für sein gezwungenes Hochdeutsch verantwortlich ist, seine Deutung.

Noch stärker erscheint ein proletarischer Einschlag beim Mittelschüler (Vp. VII), der auf ein Versagen der Deutung zurückzuführen ist, da das geschlechts- und altersbedingte Kreischen der Stimme fälschlich als milieubedingt ausgelegt und zu Angaben wie „Köchin“, „Hausbesorgerin“, „Naschmarktweib“ verwertet wurde. Umgekehrt bedeutet, wie wir schon erwähnten, die starke Besetzung der Rubrik „Akademiker“ beim Cafetier (Vp. VI) eine starke soziale Überschätzung, trägt aber auch dem Typus des intelligenten, geistig regsamen Geschäftsmannes Rechnung. Charakteristisch ist auch das Abnehmen der Intellektualität vom Privatdozenten (Vp. I) zum Priester (Vp. V), negativ nachweisbar im Absteigen der Prozente bei „Akademiker“, positiv im Ansteigen in der Rubrik „Arbeiter“.

Die Feinheiten der Deutung gehen aber über solche soziale Differenzierungen noch hinaus; nur mußte dieser qualitative Nuancenreichtum in der quantitativen Betrachtungsweise einer stark zusammengefaßten Tabelle, auch wenn man sie ihren Relationen nach interpretieren kann, verloren gehen. Es würde zu weit führen, hier materialiter auf alle Einzelheiten einzugehen. Zur Illustration des Tatbestandes folgen nur ein paar Hinweise:

Beim Privatdozenten (Vp. I) wird z. B., sofern er als Akademiker erkannt wurde, immer mehr die naturwissenschaftlich-praktische Seite betont, es heißt also „Chirurg“, „Chemiker“, „Ingenieur“; beim Priester (Vp. V) wird dagegen das spezifisch menschliche Interesse betont: er ist der „Priester“, der „heilende Arzt“, und diese Unterscheidung wird auch in den „falschen“ Angaben durchgehalten. Vp. I ist der „Schauspieler“, der „Journalist“, Vp. V der „Musiker“, der „Dichter“, Vp. I ist der „Großindustrielle“, der „Rentier“, Vp. V der „Kaufmann“, der „Pensionist“. Von diesen beiden Sprechern wird der Cafetier abgehoben als der liebenswürdige, aber geschäftstüchtige Vertreter des Intellektualismus, er erscheint als „Advokat“, als „Redakteur“, als „Humorist“, wenn er nicht als „Kaufmann“ erkannt wurde. Ebenso werden die weiblichen Sprecher voneinander deutlich unterschieden: Die Akademikerin und Bastwerkstättenbesitzerin (Vp. IV) ist die praktisch tätige, die „Ärztin“, die produktive Künstlerin, also die „Schriftstellerin“, die „Kunstgewerblerin“. Die Stenotypistin (Vp. VIII) erscheint als „Mittelschullehrerin“, reproduzierende „Sängerin“. Bei annähernd gleichmäßiger prozentueller Verteilung erscheint Vp. IV als die „Dame“, Vp. VIII als die „Hausfrau“.

Interessant ist schließlich die Charakterisierung des Chauffeurs (Vp. III). Hier liegt die Nuance darin, daß der Beruf des Chauffeurs zwar von vielen Hörern völlig richtig erkannt wurde, daß aber in sehr vielen Fällen auch eine Rückversetzung des Berufstypus um rund 30 Jahre erfolgte, die mit der durch die Befangenheit der Stimme verursachten Altersüberschätzung völlig parallel läuft: der Chauffeur erscheint als „Fiaker“. Dabei ersetzt die Geläufigkeit des Berufstypus offensichtlich völlig seine mangelnde Aktualität. In den Zuschriften stehen höchst lebendige, detaillierte und — völlig übereinstimmende Schilderungen über den Fiaker, der beim Zeugler steht und zum Fahren einlädt, die offenbar als illustrierendes Vorstellungsbild auftauchten und in ihrer Lebendigkeit die durch die Befangenheit der Stimme hervorgerufene Unsicherheit des Hörers wettmachten.

Einen statistischen Beleg für diese Unsicherheit haben wir in der *Weite* der Berufsangaben, d. i. in der Zahl der Glieder, die die Berufsangabe eines Hörers zu einem Sprecher enthält. Obwohl sich die Einsender bei keinem Sprecher über die soziale Schicht so im klaren sind wie beim Chauffeur, sind hier die speziellen Berufsangaben am wenigsten bestimmt, werden die meisten Möglichkeiten angeführt: auf 1908 Einsender kommen 2193 Berufsangaben, d. i. auf 100 Einsender 115 Angaben. Nun nimmt die *Weite* der Berufsangaben mit fortschreitendem Versuch (und zunehmender Schulung) überhaupt ab. (Am zweiten Versuchstag kommen auf 100 Einsender nur mehr 109, am 3. 106 Angaben). Aber bei den übrigen Sprechern des 1. Versuchstages stehen immerhin nur 112 Angaben von 100 Hörern. Die *Weite* der Angaben hängt also sichtlich mit der Stimme des Sprechers III zusammen,

mit seiner Befangenheit, aber auch mit der Tatsache, daß er der einzige wirkliche Dialektsprecher war. Es ist für den Hörer viel schwerer innerhalb des Tatbestandes „Dialekt“ nach der Stimme feinere Nuancierungen zu treffen als beim Intellektuellen, weil hier die einzelnen Berufstypen sprachlich viel weniger unterschieden sind: mit Dienstmann, Fiaker, Wirt verbinden wir viel weniger differenzierte Berufsstimmen wie mit Arzt, Schauspieler, Offizier. Wir wollen der Phänomenologie nicht vorgreifen, aber so viel scheint sich schon aus der Statistik zu ergeben: es gibt 3 Formen, in denen die Berufsangaben gemacht werden:

1. die generelle Angabe einer Berufsklasse („Arbeiter“),
2. die illustrierende Angabe einer Berufsklasse („Arbeiter, vielleicht Chauffeur“),
3. die spezialisierende Angabe eines oder mehrerer Spezialberufe („Arzt“, „Arzt oder Schauspieler“).

Diesen 3 Formen des Deutungsergebnisses entsprechen sichtlich 3 Phasen des Deutungsablaufes, in deren *erster* der Umstand, ob der Sprecher Dialekt spricht oder nicht, für die Scheidung intellektuell oder manuell ausgewertet wird (generelle Berufsangabe), in deren *zweiter* die Art des Dialektes eine Illustration der generellen Berufsklassenangabe nahelegt (illustrierende Angabe), in deren *letster* Vorstellungen von Berufstypen mitwirken, die besonders durch die Klangfarbe ausgelöst werden und zur Angabe von Spezialberufen führen. Inwieweit diese Berufstypen nicht nur ihrem Stimm-, sondern auch ihrem visuellen Typus nach für die Deutung wichtig sind, bedarf (wie dieser ganze Absatz) der phänomenologischen Klärung.

Im Gegensatz zur Frage nach dem Körperbau des Sprechers, die in der Formulierung: „ist der Sprecher groß, mittel, klein“, „dick, mittel, dünn“ nur sehr allgemein gestellt war, haben wir nach der Herkunft des Sprechers mit der Frage nach dem Beruf sehr speziell gefragt. Mit der ersten Fragegruppe begaben wir uns in ein noch völlig neues Gebiet, wir konnten im Vorhinein seine Möglichkeiten durchaus nicht abschätzen, wir mußten die Fragestellung daher möglichst allgemein halten. Mit der zweiten Problemgruppe trafen wir einen Zusammenhang, der uns aus den Erfahrungen des Alltags als bestehend völlig geläufig war und nur exakt herausgearbeitet werden mußte. Wir konnten, ja wir mußten speziell fragen. Demgemäß sind zur ersten Gruppe die Zusätze der Hörer, die ja den Fragebogen ergänzten, spezieller, zur zweiten allgemeiner: sie sind ferner zur ersten viel zahlreicher (25,5 %) als zur zweiten (11,3 %). Und während die Erweiterung der einen in der Differenzierung des im Fragebogen angeschnittenen Problems liegt, liegt die der anderen in einer Ausdehnung des Problemkreises.

Die Herkunft der Sprecher wurde nicht nur in *sozialer*, sondern auch in *lokal-nationaler* und *konfessioneller* Hinsicht charakterisiert. Prozentuell sind die Zusätze zur sozialen Herkunft, weil sie mit dem Fragebogen am nächsten zusammenhängen und weil in dieser Hinsicht jeder Sprecher deutlich zu charakterisieren ist, am häufigsten (69% aller Zusätze über die Herkunft); dann folgen mit 23% die Angaben über die lokal-nationale und schließlich mit 8% die über die konfessionelle Herkunft.

Die beiden letzten Gruppen sind, so sehr sie das Richtige treffen, deshalb zahlenmäßig viel kleiner, weil sie sich nur auf die in dieser Beziehung deutlich modifizierten Sprecher beziehen. Es wurde die *Konfession* nur bei den jüdischen Sprechern, die *Herkunft* nur bei solchen, die mit einem ausländischen oder mit dem typisch wienerischen Akzent sprachen, hervorgehoben. Dabei wurde nicht nur die Art der lokal-national-konfessionellen Herkunft richtig getroffen, sondern es kamen auch die Intensitätsabstufungen heraus.

Der Cafetier (Vp. VI), der tatsächlich ungarisch-tschechischer Herkunft ist, aber einen viel stärker tschechischen als ungarischen Akzent hat, erscheint $2\frac{1}{2}$ mal so oft als „Slawe“ wie als „Ungar“. Die leise reichsdeutsche Sprechweise des Privatdozenten, die von einem längeren Aufenthalt in Deutschland herrührt, fällt zwar auf, aber in einem prozentuell viel geringeren Maß als der Dialekt des Wiener Chauffeurs oder der Akzent des Cafetiers (4% gegen 27% gegen 40%).

Bezeichnenderweise ist für den österreichischen Hörer, mit dem wir es ja in erster Linie zu tun haben, der slawisch-ungarische Akzent des Cafetiers um so viel auffälliger als der Wiener Dialekt des Chauffeurs. Konfrontiert mit der Berufsdeutung scheinen die besseren Ergebnisse beim Chauffeur dafür zu sprechen, daß die Erkenntnis des Wiener Dialektes dieser Vp. zwar von sehr vielen Hörern gemacht, aber offenbar infolge der Geläufigkeit des Dialektes gar nicht als Deutungsergebnis gewertet wurde, sondern als unmittelbares Stimmerkmal für die Berufsdeutung sofort ausgewertet wurde. Andererseits wurde der ausländische Akzent des Cafetiers in seiner Modifikation als selbständiges Deutungsergebnis erlebt und ging nicht so ohne weiteres, vom Hörer unbeachtet, in die Berufsdeutung über. Es erweist sich jedenfalls auch dafür, wann ein Moment an der Stimme „auffällt“ und als Deutungsergebnis abgehoben wird, die Notwendigkeit einer phänomenologischen Betrachtungsweise.

Während also diese lokal-national-konfessionellen Zusätze zwar in die Berufsdeutung mehr oder minder bewußt eingingen, aber doch die Frage der Herkunft des Sprechers nach zwei neuen Gesichtspunkten hin ausdehnen, stellen die Zusätze zur *sozialen Herkunft* der Sprecher einen bloßen Kommentar der Berufsangaben dar. Der Wert dieser Angaben über Unterschiede der Geburt und Erziehung zwischen den einzelnen Sprechern, über ihre pekuniären Verhältnisse, über die Eleganz ihrer Kleidung, über ihre allgemeinen Familien- und Lebensverhältnisse, über den Grad ihrer Bildung liegt einerseits darin, daß sie durch die Spezialisierung einer bereits sehr speziell gestellten Frage die Durchgängigkeit der Tendenz nach einer möglichst speziellen Erfassung des Sprechers erweisen; andererseits sind sie eine willkommene Kontrolle für die Richtigkeit unserer Interpretation der statistischen Ergebnisse der Berufsdeutung, die sie vollauf bestätigen.

2. Bisher wurde der Beruf des Sprechers rein behavioristisch betrachtet. Das lag im Sinne unserer Fragestellung, die die Modifikation der Stimme von außen her auf ihre Deutbarkeit hin untersuchen wollte. Gerade die Frage des Berufes gibt aber die Möglichkeit an ein ganz anderes Problem anzuknüpfen, nämlich daran, *ob und wie die Innerlichkeit in der Stimme zum Ausdruck kommt*. Denn die Berufsbestimmung bedeutet nicht nur eine Milieubestimmung. Sie enthält — implizite zumindest — eine Charakteristik der Persönlichkeit nach ihren *Interessen* hin (wobei wir Interessen durchaus in dem Sinn verstehen, wie sie von unseren Hörern, dem Sprachgebrauch des Alltags folgend, angewendet wurden). Wir fanden ja, wie Sprecher, die in die gleichen Berufskategorien fielen, tatsächlich nach ihren Interessen voneinander abgehoben wurden, wie z. B. der Privatdozent als der sachlich, der Priester als der menschlich Interessierte charakterisiert wurde. Nun haben wir in unserem Fragebogen gewissermaßen eine charakterologische Klärung der Berufsangaben ermöglicht durch die Frage: „*ist der Sprecher befehlsgewohnt oder nicht?*“ Wir hatten es aus Gründen methodischer Reinlichkeit unterlassen, charakterologische Fragen in den Fragebogen aufzunehmen. Mit der Frage nach der Befehlsgewohnheit aber wurde ein Punkt der Innerlichkeit getroffen, der auch seine objektive Geltung hat und sich in seiner Mittlerstellung zwischen Innen und Außen als Ansatzpunkt eignen mußte.

Die Antworten bestätigen diese Vermutung. Das statistische Ergebnis leidet nämlich sichtlich unter einer gewissen Unstimmig-

keit der Einsendungen: in einigen wird die Frage, ob „befehls-
gewohnt“ rein äußerlich, man könnte sagen automatisch nach dem
Berufe beantwortet (also für „Intellektueller“ die Antwort „befehls-
gewohnt“, für „Nicht-Intellektueller“ die Antwort „nicht befehls-
gewohnt“ gegeben), in anderen — und das ist die überwiegende
Mehrzahl — wird das „befehls-
gewohnt“ beantwortet im Sinn der
auch im Beruf sich äußernden Haltung des Sprechers der Außen-
welt gegenüber, wie sie durch seine ganze Persönlichkeit bedingt
wird. Rein äußerlich schon ist der Unterschied oft dadurch nach-
weisbar, daß sich viele dieser Einsender nicht mit der bloßen
Antwort „ja“ oder „nein“ begnügen, sondern dazufügen, *wie* der
Sprecher befiehlt, *wie* er gehorcht; und das, worüber sie dabei
spontane Aussagen machen, sind solche Daten der Innerlichkeit.

Tabelle 10

Die Verteilung der Antworten auf die Frage nach der Befehls-
gewohnheit der Sprecher in absoluten Zahlen und in Prozent

Nummer u. Beruf der Vpn.	befehls- gewohnt		nicht bef.- gewohnt		unent- schieden		Summe	
	absol. Zahlen	%	absol. Zahlen	%	absol. Zahlen	%	absol. Zahlen	%
Vp. I Privatdozent	1699	89,1	203	10,6	5	0,3	1907	100
Vp. IV Akademikerin	1361	88,2	177	11,5	5	0,3	1543	100
Vp. V Priester	873	56,6	666	43,1	5	0,3	1544	100
Vp. IX Lehrer	758	48,1	812	51,5	7	0,4	1577	100
Vp. VIII Stenotypistin	798	47,5	876	52,2	6	0,3	1680	100
Vp. VI Cafetier	574	37,1	967	62,5	6	0,4	1547	100
Vp. II Mittelschülerin	414	22,0	1473	77,0	20	1,0	1907	100
Vp. VII Mittelschüler	251	15,0	1382	82,4	44	2,6	1677	100
Vp. III Chauffeur	236	12,4	1666	87,3	6	0,3	1908	100

Wir greifen nur ein paar Zahlen heraus: die hohe Besetzung
von „befehls-
gewohnt“ bei Vp. IV gegenüber Vp. VIII ergibt sich
nicht als eine automatische Folgerung der Berufsangaben, die ja
die Akademikerin verglichen mit der Stenotypistin stark unter-
schätzen, sondern aus der „Energie“ und „Herrschaft“ einer-
seits, der „Liebenswürdigkeit“ und „Freundlichkeit“ andererseits,
also aus Momenten der Innerlichkeit, die durchaus primär der
Stimme entnommen werden. Sie sind in der Mehrzahl keine
bloß charakterologische Formulierung der Berufsangaben. Im
Gegenteil! In vielen Fällen verhelfen sie offenbar innerhalb der
unabhängig von ihnen festgestellten Berufs-
klasse zu speziellen

Berufsangaben, indem sie die Erinnerung an beruflich fixierte Stimmtypen auslösen.

So fanden wir z. B. in der Kategorie „Lehrerin, Beamtin“ bei der Akademikerin und Stenotypistin annähernd gleich viel Angaben. Die Abhebung der beiden Vp. geschieht durch charakterologische Momente, indem die „energische“ Akademikerin doppelt so oft als Lehrerin wie als Beamtin, die „liebenswürdige“ Stenotypistin doppelt so oft als Beamtin wie als Lehrerin erscheint.

Vergleichen wir die Beurteilung des Privatdozenten und der Akademikerin hier und in den Berufsangaben, so rückt sie — gegenüber einer Unterschätzung der Akademikerin dort — die beiden Sprecher nahe zusammen; der Cafetier, in der Berufsschätzung dem Priester fast gleichgestellt, erscheint hier in einem ganz beträchtlichen Abstand von ihm. Die Hörer erklären diese Deutung mit der „liebenswertig entgegenkommenden“ gegenüber der „gütigsten“ Stimme des Priesters. In bezug auf die *Richtigkeit* bedeuten also die Antworten auf die Frage nach der Befehlsgehnheit der Sprecher eine deutliche *Korrektur der Berufsangaben*. und zwar offensichtlich gerade deshalb, weil sie das charakterologische Moment in den Vordergrund der Betrachtung rücken. Das würde aber ferner besagen, daß die Stimme tatsächlich für gewisse Momente der Innerlichkeit ausdrückshaltig ist, wofür auch der Umstand spricht, daß die automatische Beantwortung der Frage nach der Befehlsgehnheit eine relativ so seltene war. Allerdings konnten wir hierbei graduelle Unterschiede zwischen den einzelnen Sprechern nachweisen: d. h. die Anzahl der rein automatischen Beantwortungen ist dort am größten, wo die charakterologische Ausdruckshaltigkeit am kleinsten ist. (Beim Lehrer wird immer wieder darüber geklagt, daß seine Stimme „so wenig besagt“, er erscheint beruflich mit rund 50% als Intellektueller und die Antwort auf die Frage der Befehlsgehnheit ergibt dasselbe Bild.)

Dieser Zugang zur Innerlichkeit, den wir hier aufdecken konnten, ist allerdings einigermaßen erschwert und isoliert erfaßbar nur dort, wo die Antworten mit charakterologischen Merkmalen begründet waren. Wir haben aber in unserem Material auch Zusätze, die völlig rein die Innerlichkeit der Sprecher betreffen.

3. *Die charakterologische Bedingtheit der Stimme*

Die meisten Angaben (40%) unter den spontanen Zusätzen gehen die *Innerlichkeit* der Sprecher an. Diese hohe Prozentzahl erklärt sich nicht nur daraus, daß im Fragebogen keine rein

charakterologische Frage enthalten war, die Hörer also alles Diesbezügliche in Form von Zusätzen schreiben mußten, sondern auch aus der immer wieder hervorgehobenen Stärke der charakterologischen Ausdruckshaltigkeit der Stimme. Sie ist verständlich aus dem Wesen des Ausdrucks: in ihm offenbart sich vor allem das wahrnehmungsmäßig nicht unmittelbar Faßbare, die große Unbekannte, die als Motor oder als Steuerungsprinzip einer jeden menschlichen Handlung anzusetzen ist. Ein Zugang zu dieser Unbekannten ist die Ausdeutung dieser Handlungen oder der durch sie zustande gekommenen Werke auf ihren „Ausdrucksgehalt“ hin, ist die Ausdeutung der Schreibbewegung, wie sie KLAGES in seiner Graphologie vornimmt, ist die Ausdeutung der Sprechbewegung, wie wir sie vornehmen. Wenn wir diese Unbekannte mit dem von KARL BÜHLER geprägten Terminus „Innerlichkeit“ bezeichnen, so ist dies vorläufig vor allem eben ein Terminus: letztes Ziel, vielleicht die Hauptaufgabe der theoretischen Psychologie ist die Bestimmung dieses Terminus. Was können wir aus unserem Material dazu beitragen?

An den Antworten der RAVAG-Hörer interessieren uns also die quantitativen Ergebnisse, aber mindestens ebenso stark ihre Formulierung. Sie darf allerdings nicht absolut betrachtet werden, sondern nur in bezug auf die Situation, aus der heraus sie entstanden ist. Dort, wo der Hörer ohne jeden Zwang, aber auch ohne jede Anleitung zu Worte kommt, wird er offenbar in den ihm geläufigen Terminis des Alltags sprechen. Letztere sind ihrem Bedeutungsgehalt nach offensichtlich weder exakt noch eindeutig bestimmt, noch werden sie, gerade wegen ihrer Abgeschliffenheit, in vollem Verantwortungsbewußtsein gesetzt. Der Wert der Zusätze ist daher vor allem ein lexikalischer: d. h. wir bekommen in ihnen ein Vokabular der zur Beschreibung der Innerlichkeit geläufigen Termini.

Bei einem Versuch, diese ungefähr 6000 Zusätze einigermaßen zusammenzufassen, erweist sich die Einteilung von KLAGES¹ als sehr praktisch, d. h. man kann sämtliche Angaben in den fünf Gruppen unterbringen, mit denen KLAGES die Persönlichkeit beschreibt, wenn man sie im behavioristischen Sinn erweitert und sein erlebnismäßig fundiertes Persönlichkeitsschema durch eine

¹ L. KLAGES, Persönlichkeit, in: das Weltbild, Bücher des lebendigen Wissens, hrsg. von Hans Prinzhorn II. Bd., S. 30 ff.

hinzugefügte Rubrik „aktuelles Erlebnis“ vervollständigt. Wir brauchen aus dieser Übereinstimmung mit KLAGES durchaus keine theoretischen Konsequenzen zu ziehen. Sie ergibt sich aus dem Ausgangspunkt der KLAGESSchen Einteilung, der Sprache und den in ihr vorkommenden Anregungen. Die Zusätze gliedern sich also in Angaben über:

- 1. Die Fähigkeiten der Sprecher (23,1%)
- 2. das Temperament „ „ (14,3%)
- 3. die innere Art „ „ (37,2%)
- 4. die innere Harmonie „ „ (2,0%)
- 5. die Haltung zur Außenwelt und die Auseinandersetzung mit ihr (16,1%)
- 6. die aktuellen Erlebnisse der Sprecher (7,3%)

Die Tabelle dient vor allem der Übersichtlichkeit, zahlenmäßig ist sie nur so zu interpretieren, daß z. B. das, was wir unter „innerer Art“ zusammenfassen, von den Hörern am häufigsten hervorgehoben wurde, also in der Stimme am stärksten zum Ausdruck kommt. Ein detaillierter Bericht über die Art, wie die Sprecher in diesen einzelnen Gruppen (oft außerordentlich treffend) charakterisiert wurden, müßte für den Leser, der sie nicht kennt, langweilig sein. Wir können ihm und uns die Situation erleichtern durch die Tatsache, daß für den einzelnen Sprecher die Angaben in den einzelnen Gruppen auch quantitativ nicht konstant waren. Es gibt immer eine Merkmalsgruppe, die ihn besonders charakterisiert.¹ Wir wollen diesen Umstand als das Ausleseprinzip für die folgende Darstellung verwenden, geben aber vorher noch die Tabelle dafür.

Tabelle 11

Die prozentuelle Verteilung der charakterologischen Angaben zu den einzelnen Sprechern

Angaben betreffend die charakterologischen Gruppen	Nummer und Beruf der Vpn.									Summe in %
	Vp. I Privatdozent	Vp. II Mittelschülerin	Vp. III Chauffeur	Vp. IV Akademikerin	Vp. V Priester	Vp. VI Cafetier	Vp. VII Mittelschüler	Vp. VIII Stenotypistin	Vp. IX Lehrer	
Fähigkeiten	13	11	11	18	11	8	11	9	8	100
Temperament	5	25	13	7	14	10	11	8	7	100
Innere Art	13	12	15	11	12	9	8	10	10	100
Innere Harm.	12	14	9	3	23	10	6	9	14	100
Haltung	19	12	11	9	8	19	11	7	4	100
Akt. Erleb.	4	12	60	4	4	3	8	2	3	100

¹ Ich folge hier auch dem Gedankengang von H. TURSKY a. a. O. 22

1. Zu den „Fähigkeiten“ rechneten wir Fähigkeiten des Denkens, Wollens und Fühlens. Es gehören im wesentlichen hierher: Angaben über die Intelligenz (28 % dieser Gruppe), die Lebensklugheit (25 %), die künstlerischen Fähigkeiten (5 %), die allgemeine Begabung (3 %), die Energie (37 %) und die Feinfühligkeit (2 %). Innerhalb dieser Gruppe ist also sichtlich die Energie aus der Stimme besonders deutlich faßbar.

„Energisch“, „willensstark“, „setzt durch, was sie sich vorgenommen hat“ sind auch die Formulierungen, in denen die Akademikerin (Vp. IV) gegenüber den anderen Sprechern charakterisiert wird. Auch die künstlerischen Fähigkeiten der Akademikerin, die tatsächlich eine Bastwerkstätte besitzt, kommen in der Deutung gut heraus, ebenso wie die des Priesters, der auch Dichter ist.

2. In dem, was wir als „Temperament“ zusammenfaßten, finden sich vor allem Angaben über die innere Beweglichkeit und Heiterkeit (56 % d. Gr.) und über die Empfindlichkeit der Sprecher (44 %).

So wird die Mittelschülerin in dieser Gruppe besonders charakterisiert und als „nervös“, „empfindlich“ gekennzeichnet, während der Lehrer dagegen z. B. als „gleichgültig“ erscheint.

Die beiden Hauptpaare dieser Gruppe beweglich, heiter—traurig, still, empfindlich—gleichgültig decken sich in ihrer Formulierung mit dem, was KRETSCHMER¹ als die Charakteristika der zyklithymen und schizothymen Temperamente angibt. KRETSCHMER wird aber noch viel weiter bestätigt, insofern, als tatsächlich die Sprecher, die in unserem Experiment mit den körperbaulichen Merkmalen der Leptosomen und Athletiker gekennzeichnet wurden, auch in der Temperamentsbeurteilung deutlich zusammengestellt wurden und zwar in der Kategorie, die KRETSCHMER als die für die Schizothymen charakteristische angibt. Umgekehrt wurden die als die Pykniker Beurteilten von ihnen mit den Merkmalen der Zyklithymen abgehoben. Von der Stimmdeutung her werden also die von KRETSCHMER aufgestellten Temperamentsunterschiede der verschiedenen Körperbautypen bestätigt und als relevant, weil in den Laienzusätzen spontan auftauchend, dargetan. Die Feinheiten der Deutung gehen aber weiter: auch innerhalb einer Temperamentsgruppe werden die Sprecher noch deutlich voneinander abgehoben. Und zwar geschieht dies, schon rein quantitativ erkennbar, einerseits dadurch, daß durch die Verteilung der Angaben auf die beiden Temperamentspole der Platz bestimmt wird,

¹ E. KRETSCHMER, a. a. O. S. 172 ff.

den der Sprecher in der Temperamentsskala seines Typus einnimmt, andererseits durch die Verteilung der Angaben auf das für die andere Temperamentsgruppe charakteristische Gegensatzpaar: denn so wenig wir es mit reinen Körperbautypen zu tun hatten, so wenig sind unsere Sprecher auch reine Temperamentstypen. Zur Erläuterung diene die Gegenüberstellung der Angaben zweier als Pykniker erkannter Sprecher.

Tabelle 12

Die prozentuelle Verteilung der Temperamentsangaben zu zwei pyknischen Sprechern

Temperamentsangaben	Priester (Vp. V)	Cafetier (Vp. VI)
beweglich, frisch	9	72
ruhig, schwernehmend	68	16
empfindlich	5	8
gleichgültig	18	4
Summe:	100	100

Der Priester erscheint 1. als der ruhige, melancholische, der Cafetier als der bewegliche, mehr sanguinische Typ des Sanguinikers: 2. hat der Priester einen stärkeren und auch ausgeprägteren schizothymen Einschlag als der Cafetier (auch in der körperbaulichen Beurteilung).

KRETSCHMER¹ faßt seinen Temperamentsbegriff allerdings sehr weit. Er versteht darunter auch Merkmale, die, zumindest dem üblichen Sprachgebrauch nach, von dem wir in unserer Einteilung ausgingen, ganz offensichtlich eher als „Fähigkeiten“ zu bezeichnen wären, wie z. B. „Dummheit“. Dieses Betonen des Temperaments, des affektiv-willensmäßigen Faktors überhaupt, bedeutet eine theoretische Einseitigkeit, die hier nicht zur Diskussion steht. In unseren Antworten kommen jedenfalls die 3 Bereiche des Denkens, Wollens und Fühlens in gleichem Maße zum Ausdruck. Lassen wir dagegen KRETSCHMERS Theorie undiskutiert und kümmern wir uns, dem Charakter und den Forderungen unseres Materials folgend, um die praktischen Beispiele, mit denen er seine Typen charakterisiert, so enthalten seine Angaben durchaus dasselbe und sind in gleicher Weise formuliert wie die Angaben, mit denen die Ravag-Einsender die Vertreter der beiden Formkreise charakterisierten.

3. Wir haben unter „innerer Art“ zusammengefaßt, was irgendwie den Interessenkreis und die ethische Haltung der Sprecher kennzeichnete. Dazu gehören also Angaben wie: „Ist gern gut“, „Trinker“ (8 % d. Gr.; die Angaben beziehen sich immer auf das Gegensatzpaar, also auch auf die negative Formulierung),

¹ E. KRETSCHMER, a. z. O. S. 172 ff., 225 ff.

„Sportler“ (6 %), „Natur-, Kinder-, Tier-, Menschenfreund“ (4 %), „Frauenfreund“ (2 %), „geistige Interessen“ (3 %), „modern“ (2 %), „sachlich“ (10 %), „aufrichtig“ (10 %), „gütig“ (16 %), „gerecht“ (4 %), „gründlich“ (7 %), „großzügig“ (3 %), „nachgiebig“ (11 %), „eingebildet“ (14 %). Die ethischen Wertungen des Gütig- und Eingebildetseins stehen im Mittelpunkt, sie kommen also in der Stimme für den Hörer am deutlichsten zum Ausdruck. Dabei bedeutet dieses „zum Ausdruckkommen“ nicht nur einen objektiven Tatbestand, sondern auch eine gewisse Intention vom Hörer aus gerade auf diese Momente der Innerlichkeit. Übrigens wurden Zusätze zur inneren Art überhaupt besonders gern gemacht (sie sind zahlenmäßig am häufigsten). Und charakteristischerweise ist auch die Verteilung der Angaben dieser Gruppe auf die einzelnen Sprecher die regelmäßigste: hier wollen und können die Hörer zu allen Sprechern etwas sagen. Die individuellen Unterschiede sind rein quantitativ am wenigsten deutlich (relativ betrachtet!), sie liegen vor allem in den qualitativen Abstufungen, die, wie aus der vorangegangenen Aufzählung ersichtlich ist, nirgends so nuancenreich sind wie hier. Daß das Maximum der Angaben dieser Gruppe beim Chauffeur liegt, das ist bei dem Sprecher, bei dem durch die Aufgeregtheit die Deutung so außerordentlich erschwert wurde, bestätigt diese Feststellung.

Und zwar wird er bezeichnenderweise nicht so sehr seiner ethischen Haltung als seinen Vorlieben nach, in denen mehr sein typischer Charakter getroffen wird, von den anderen Sprechern abgehoben. 64 % aller Angaben in der Rubrik „ißt gern gut, Raucher, Trinker“ entfallen auf diese Vp. Ähnlich verhält es sich übrigens auch beim Mittelschüler, dessen mutierende Stimme die Deutung auf allen Gebieten erschwerte. Bei ihm ist die Rubrik „Sportler“ am häufigsten von allen Sprechern besetzt.

Bei den anderen Sprechern aber steht die Beurteilung der ethischen Haltung im Vordergrund und bestimmt in ihrer Nuancierung die Unterschiede zwischen den einzelnen Vpn. So wird z. B. der Privatdozent als der „etwas herrische, eingebildete Verstandesmensch“ abgehoben von dem „idealistischen Kinder- und Naturfreund“, dem „gütigen“ Priester. Der Cafetier erscheint ihm gegenüber als der „Realist“, der „sich durchsetzen will und kann“ und wird deutlich unterschieden von dem Lehrer als dem „Gewissenhaften“, „Pedantischen“, „Bescheidenen“. In ähnlicher Weise werden auch die weiblichen Vpn. gegeneinander abgehoben.

Erinnern wir uns von hier aus noch einmal an die auf den ersten Blick vielleicht überraschenden Treffer in den speziellen Berufsangaben, so finden wir aus der Übereinstimmung damit, daß wir in der charakterologischen Erfassung der Sprecher in bezug auf die „innere Art“ die Grundlagen für die Berufsangaben zu suchen haben.

4. Wenn wir dazu, was wir unter „*innerer Harmonie*“ zusammenfaßten, so wenig Angaben erhielten, so bedarf dieser Umstand offensichtlich einer phänomenologischen Klärung; obzwar zu berücksichtigen ist, daß eine Gruppe, für die es rein quantitativ weniger Vokabeln gibt — wir zählten dazu vor allem zweierlei Angaben, solche wie „harmonisch, ausgeglichen — unharmonisch, zwiespältig“ (60 % d. Gr.) und solche wie „ausgeprägt — unausgeprägt“ (40 %) — auch weniger häufig vorkommen kann. Hätten wir unserer Gruppeneinteilung einen theoretischen Wert zugeschrieben, so müßte uns die geringe Besetzungszahl aber doch ernstliche Sorgen machen. Von unserem praktischen Einteilungsgesichtspunkt aus sind wir jedoch vollauf berechtigt, diesen Umstand mit dem Hinweis auf die Phänomenologie zu vernachlässigen. Praktisch brauchen wir jedenfalls diese Kategorie und sie bewährt sich auch, weil sie zeigt, daß sie für zwei Vpn., den Priester (Vp. V) und den Lehrer (Vp. IX), besonders charakteristisch ist. Und in gewissem Sinn hat sie auch für den Mittelschüler (Vp. VII), für den es, wie die Tabelle 11 zeigt, keine eindeutige, durch ein Maximum bestimmte charakteristische Gruppe gibt, als Repräsentant des Minimums ihre Bedeutung.

Dieser quantitative Befund enthält eine sehr hübsche qualitative Pointe. Dort, wo es sich im Quantitativen um das absolute (d. h. auf diese Gruppe bezogene) Maximum oder Minimum der Angaben handelt, ist der Befund qualitativ ein völlig eindeutiger. Der Priester erscheint in sämtlichen Fällen als „harmonische“, „ausgeprägte“ Persönlichkeit, der Mittelschüler als „unausgeprägt“. Beim Lehrer aber, wo es sich um ein relatives Maximum handelt, wo also das Maximum auf den Sprecher bezogen ist, herrscht qualitativ eine völlige Unentschiedenheit: er erscheint ebenso oft als „harmonisch, ausgeprägt“, wie als „unharmonisch, unausgeprägt“. Die Vp. also, deren Stimme eine so geringe Ausdruckshaltigkeit besitzt, wie die Hörer immer wieder klagen, wird in der charakterologischen Gruppe charakterisiert, in der diese mangelnde Ausgeprägtheit ihren vollendeten qualitativen und quantitativen Beleg findet. Im übrigen ist diese Gruppe diejenige, bei der auch in bezug auf alle anderen Sprecher die größte Unstimmigkeit in den Antworten zu verzeichnen ist.

5. In der Gruppe „*Auseinandersetzung mit der Außenwelt und Haltung zu ihr*“ faßten wir folgende Angaben zusammen: „sicheres Benehmen“ (27 % d. Gr.), „aufdringlich“ (8 %), „liebenswert“ (37 %), „geschmeidig“ (18 %), „beliebt“ (10 %). Wiederum ist die prozentuelle Verteilung aus dem Ausdrucksgehalt der Stimmen, aber auch vom Hörer her, von dem, woran ihm am meisten liegt, zu verstehen.

Wie wichtig diese Gruppe ist, dafür haben wir einen statistischen Nachweis auch darin, daß nicht nur ein, sondern zwei Vpn. mit dem Maximum dieser Gruppe bedacht werden. Trotz dieser quantitativen Übereinstimmung erscheinen sie aber qualitativ deutlich voneinander unterschieden, indem der Privatdozent (Vp. I) vor allem als „hochmütig“, der Cafetier (Vp. VI) als „geschmeidig“ gekennzeichnet wird.

6. Während die 5 bisher besprochenen Gruppen solche Angaben zusammenfaßten, die irgendwelche Aussagen über die Persönlichkeit, über das bleibende Gepräge der Individualität enthielten, faßten wir in dieser Gruppe alle Zusätze zusammen, die irgendwie die „*aktuellen Erlebnisse*“ der Persönlichkeit angingen. Der Kreis der Erlebnisse war durch die Versuchssituation allerdings ziemlich eindeutig determiniert: 89 % der Angaben betreffen die Aufgeregtheit der Sprecher, 11 % die Beteiligung am Text („hat sich bemüht“, „war aufmerksam“, „war nicht bei der Sache“, „hat ungern gesprochen“). Die Ausdruckshaltigkeit der Stimme gerade für diese aktuellen Momente an der Persönlichkeit, aus den Erfahrungen des Alltags in unzähligen Fällen nachweisbar, wird illustriert durch die unverhältnismäßig starke Prozentzahl, mit der in dieser Gruppe der Chauffeur (Vp. III) bedacht ist. Während bei den übrigen Sprechern das Maximum der Angaben einer Gruppe durchschnittlich bei 14 % liegt, beträgt es hier 60 %. Im übrigen sind die Angaben über die Aufgeregtheit der Sprecher bei den beiden jüngsten Vpn., der Mittelschülerin (Vp. II) und dem Buben (Vp. VII), am häufigsten.

Fassen wir zusammen: Schon in der quantitativen Reichhaltigkeit der zusätzlichen Angaben über die Innerlichkeit scheint gegen eine erkenntnistheoretische Leugnung derselben eine Widerlegung, zumindest eine Mahnung zur Vorsicht in bezug auf apodiktisch negative Urteile vorzuliegen. Wenn wir die Frage der Möglichkeit von Selbstbeobachtungen und eine Erfassung der Erlebnisse durch sie hier völlig undiskutiert lassen, so ergibt sich doch aus unserem Material ein Zugang zum Erlebnis vom Ausdruck her. Andererseits müssen wir das Ausgedrückte über die Erlebnisse hinaus von der Stimmdeutung her erweitern und feststellen, daß die Stimme mehr enthält als die „*aktuellen Erlebnisse*“, daß sie auch Ausdruck der „*Persönlichkeit*“ des Sprechers ist. In der Formulierung der qualitativen Ergebnisse wollen wir zurückhaltend sein. Wir haben wohl eine Anzahl von Gruppen gefunden, in denen die „*Persönlichkeit*“ charakterisiert wurde, aber sie sind vorläufig nicht viel mehr als ein Inventar der gebräuchlichsten und offen-

bar tragfähigsten Vokabeln. Wir können mit unserem Material nicht erfassen, was sie eigentlich bedeuten. Es bleibt also z. B. ungeklärt, ob die mit diesen Vokabeln gemeinten Eigenschaften verschiedenen Schichten der Persönlichkeit angehören, bis zu welcher Schichttiefe sie also reichen, usw. Wir können nur feststellen: 1. solche Vokabeln werden zu jedem Sprecher angeführt. 2. Die Prozentzahl der Angaben ist eine verschieden große, es gibt also eine Intensität der charakterologischen Ausdruckshaltigkeit der Stimme. 3. Die Vokabeln ermöglichen eine sehr feine und treffende Abstufung der Sprecher gegeneinander und zwar gibt es für jeden Sprecher eine Vokabelgruppe, die ihn besonders charakterisiert und nach der diese Abhebung vor allem erfolgt. 4. Jede der angeführten Gruppen zerfällt in eine Anzahl von Gegensatzpaaren. Eine charakterologische Typenlehre, die in charakteristischen Gruppen eine solche Polarität feststellt, wie es KRETSCHMER bei seinen Schizothymen und Zykllothymen tut, erscheint von hier aus einigermaßen gerechtfertigt. 5. Wir konnten überhaupt eine weitgehende Bestätigung der KRETSCHMERSCHEN Ausführungen feststellen, insofern als die in seinem Sinn körperbaulich verschiedenen Typen auch charakterologisch mit den Vokabeln beschrieben wurden, in denen KRETSCHMER ihre Typen gekennzeichnet hätte.

4. Die Stimme der Sprecher

Auch die Grundlage all dieser Deutungen, die *Stimme* selbst, wurde in einer Reihe von Zusätzen (17,2 %) näher beschrieben und zwar erfolgte die Charakteristik der einzelnen Stimmen nach

6 *Stimmdimensionen*:

1. Stimmhöhe (2 %): „hohe“, „tiefe“ Stimme, „großer Stimmumfang“.
2. Stimmstärke (11 %): „leise“, „starke“ Stimme, „wechselnde Stärke“, „gute“, „schlechte“ Betonung, „leiernd“.
3. Rhythmus (1 %): „veränderlicher“, „gleichmäßiger“.
4. Tempo (3 %): „schnelles“, „langsames“, „ruhiges“, „müdes“.
5. Timbre (35 %): „melodisches“, „schrilles“, „ausdrucksvolles“, „helles“, „trübes“, „warmes“, „volles“, „weiches“, „fettes“, charakterologische Wendungen wie: „sanftes“ Timbre.
6. Artikulation (37 %): „stotternde“, „flüssige“, „dialektfreie“, „verstellte“, „affektiertes R“.

In einer siebenten Gruppe wurden alle die Angaben zusammengefaßt, in denen irgendwelche Identifizierungen nach der Stimme (4 %) erfolgten oder Ähnlichkeiten mit der Stimme bekannter Menschen (7 %) konstatiert wurden.

Mit den 6 angeführten Dimensionen ist natürlich keine phonetisch erschöpfende Analyse der menschlichen Stimme gegeben; wohl aber kann man sagen, daß sie anscheinend die für den Laien wesentlichen Punkte bei der Deutung darstellen. Unter ihnen erscheinen wiederum *Artikulation* und *Timbre* als die wichtigsten. Bezeichnenderweise enthalten die Beschreibungen der Stimmen in diesen 2 Dimensionen schon alle die Momente, für die wir einen Ausdrucksgehalt der Stimme feststellen konnten. In der „dialektfreien Artikulation“ ist bereits ein Hinweis auf die Milieubedingtheit des Sprechers enthalten; in der synästhetischen Auffassung des Timbres bei Formulierungen wie „fette“, „klotzige“ Stimme sind bereits irgendwelche Hinweise auf den Körperbau enthalten und in der charakterologischen Auffassung des Timbres wird die Innerlichkeit der Sprecher beurteilt. Die Untersuchung, wie die Deutung tatsächlich erfolgt, ist, wie gesagt, Sache exakter Einzelanalysen. Hier haben wir jedenfalls bereits Anhaltspunkte, ebenso wie in den Identifizierungen und Ähnlichkeiten.

Die *Verteilung* der Stimmbeschreibungen auf die einzelnen Sprecher ist eine ziemlich gleichmäßige. Herausfallend ist nur die Beschreibung des Dialektsprechers, auf den 23 % der Angaben kommen gegenüber durchschnittlich 9,6 % bei den anderen Vpn. Auch die *Art* der Beschreibung ist eine verhältnismäßig einheitliche; die Übereinstimmung zwischen den Hörern ist hier, wo es sich um die Beschreibung des unmittelbar Wahrgenommenen handelt, auch wenn sie schon Deutungen enthält, jedenfalls eine viel größere als bei den früheren Ergebnissen, die reine Deutung sind. Und unter diesem Gesichtspunkt lassen sich auch die Ergebnisse auf den 6. Punkt unseres Fragebogens verstehen, nämlich auf die Frage: „*ist die Stimme des Sprechers sympathisch oder nicht?*“ (s. Tab. 13).

Mit Ausnahme der durch Aufregung und Dialekt entstellten Stimme des Chauffeurs und der mutierenden Stimme des Schülers werden mindestens von rund $\frac{3}{4}$ aller Hörer alle Sprecher als „sympathisch“ beurteilt; ein Hauptkriterium für die Beurteilung ist offensichtlich die *akustisch-musikalische Qualität* (die ja, wie wir oben sahen, ziemlich einheitlich aufgefaßt wird). Ein weiteres Kriterium, wodurch offenbar die Nuancierungen der Antworten zu verstehen sind, liegt in der *Ausdruckshaltigkeit* der Stimme nach *Intensität* und *Qualität*; d. h. die ausdrucksvolle Stimme hat den Vorrang vor der ausdruckslosen (s. die Beurteilung der „ausdruckslosen“

Stimme des Lehrers, der „gleichförmigen“ Stimme der Mittelschülerin). Die qualitative, d. i. hier die charakterologische Ausdruckshaltigkeit wird auf Grund allgemeiner ethischer Wertungen zur Beurteilung, ob die Stimme sympathisch sei, herangezogen. Dabei schneidet unter den männlichen Sprechern die „gütige“ Stimme des Priesters, unter den weiblichen die „liebenswürdige“ der Stenotypistin am besten ab.

Tabelle 13

Die Beurteilungen, ob die Stimme der einzelnen Sprecher sympathisch ist, in ihrer prozentuellen Verteilung

Nummer und Beruf der Vpn.	sympathisch	nicht sympathisch	indifferent unentschied.	Summe in %
Vp. VIII Stenotypistin	92,2	6,8	1,0	100
Vp. V Priester	81,5	16,9	1,6	100
Vp. I Privatdozent	80,8	16,2	3,0	100
Vp. VI Cafetier	79,1	18,5	2,4	100
Vp. IV Akademikerin	78,2	20,2	1,6	100
Vp. IX Lehrer	74,8	22,9	2,3	100
Vp. II Mittelschülerin	73,8	21,8	4,4	100
Vp. VII Mittelschüler	42,4	55,3	2,3	100
Vp. III Chauffeur	35,7	59,3	5,0	100

Wir können als *Gesamtergebnis* unseres quantitativen Lösungsversuches zum Problem „Stimme und Persönlichkeit“ sagen: aus den Zuschriften zum Massenexperiment ergibt sich, daß in der Stimme die physiologischen Daten des Sprechers (als da sind Geschlecht, Alter, Größe, Dicke, Körperformen, Körperfarben, körperliche Zustände), sein Milieu (lokale, nationale, religiöse, soziale Herkunft) und seine Innerlichkeit (Persönlichkeit, aktuelles Erlebnis) in einer für den Hörer weitgehend richtig erfaßbaren Art zum Ausdruck kommen. Zur Klärung der Ergebnisse aber brauchen wir eine zweite Methode, die der phänomenologischen Beschreibung des Stimmerlebnisses. Aus der quantitativen Untersuchung ergaben sich dafür nicht nur Fragen, sondern auch bereits gewisse Hinweise. Ein quantitativ erfaßbares Gebiet, das solche Hinweise in reichem Maße enthält, wurde bis jetzt von uns allerdings vernachlässigt: d. i. die statistische Untersuchung des Zusammenhanges zwischen dem Einsender und der Art seiner Antwort. Damit wollen wir uns in einem vierten und letzten Kapitel befassen.

Kapitel 4. Der Einsender und seine Antwort

Der erste Eindruck, den man von den 2700 Zuschriften hat, ist der einer überwältigenden Mannigfaltigkeit. Beim Durcharbeiten des Materials stellt sich jedoch bald heraus, daß dieser Eindruck vor allem durch die Verschiedenheit der äußeren Form, der Schrift usw. ausgelöst wurde; es springen inhaltlich bereits gewisse Allgemeinheiten der Stellungnahme zum Experiment, auf die schon hingewiesen wurde, hervor. Die statistische Untersuchung der Frage: Einsender und Antwort, die letzten Endes schon eine phänomenologische ist, hatte also als Kontrolle dieses ersten zwiespältigen Eindrucks ihren besonderen Reiz. Wir untersuchten natürlich nur die Antworten auf den Fragebogen, d. i. die Geschlechts-, die Alters-, die Größe-Dicke-, die Berufsdeutung und die Beurteilung, ob die Stimme sympathisch sei und zwar danach, ob die Antworten darauf nach Geschlecht, Alter oder Beruf des Hörers statistisch nachweisbare Zusammenhänge aufwiesen. Von einer Anführung der Tabellen soll hier der Kürze halber abgesehen werden; wir begnügen uns damit, ihre Ergebnisse in ein paar Sätzen darzustellen.

1. Ist die Deutung geschlechtsbedingt?

Da die erwachsenen Sprecher ihrem *Geschlecht* nach sämtlich getroffen wurden, kommen nur die beiden jugendlichen Vpn. in Frage. Wir fanden ein allerdings nicht sehr beträchtlich besseres Abschneiden der weiblichen Hörer, u. zw. in geringerem Maße bei der Mittelschülerin, relativ stark beim Mittelschüler. (Die falschen Schätzungen, in Prozent der Gesamtschätzung aller gleichgeschlechtlichen Hörer dargestellt, ergeben ein besseres Treffen der weiblichen Hörer um 3,8 %.)

Ebenso treffen die Frauen in schwierigeren Fällen die *Altersdeutung* besser. Kraß kommt dies heraus beim mutierenden Mittelschüler und dem aufgeregten Chauffeur. (Drückt man die „Richtigkeit“ der Altersschätzung durch den Quotienten zwischen den richtigen Angaben und den Fehldeutungen aus, so ist er beim Chauffeur z. B. für die männlichen Einsender 1:14, für die weiblichen 1:9.) Eine Ausnahme liegt nur in der Beurteilung der 48jährigen Akademikerin (Vp. IV) vor, die von den Frauen stärker unterschätzt wurde als von den Männern (1:7 gegenüber 1:6).

In der *Größe-Dicke-Deutung* lassen sich weder nach Geschlecht noch nach Alter und Beruf des Einsenders maßgebliche, d. i. über

den Schwankungsbereich des Unabhängigkeitswertes hinausgehende Unterschiede nachweisen.

Anders steht es mit der *Berufsdeutung*. Während Geschlecht und Alter der Sprecher von den weiblichen Einsendern einseitig, d. i. unabhängig vom Geschlecht des Sprechers besser getroffen wurden als von den männlichen, erweist sich die Berufsdeutung als geschlechtsbedingt vom Sprecher aus, d. h. von männlichen und weiblichen Einsendern wird der Beruf des gleichgeschlechtlichen Sprechers in gleichem Maße besser getroffen.

Für die *Beurteilung*, ob die Stimme *sympathisch* sei oder nicht, fanden wir zwischen männlichen und weiblichen Sprechern keinen maßgebenden Unterschied.

2. Ist die Deutung altersbedingt?

Ein altersbedingter Unterschied in der Deutungsrichtigkeit und auch in der Beurteilung der Stimme läßt sich statistisch nicht feststellen.

3. Ist die Deutung berufsbedingt?

In bezug auf *Geschlecht, Alter, Größe, Dicke, Beruf*, soweit es sich also um die Richtigkeit der Deutung handelt, lassen sich auch hier keine Zusammenhänge nachweisen. Auch für die *Beurteilung*, ob die Stimme *sympathisch* sei, fehlen sie bei der Mehrzahl der Vpn. mit Ausnahme des Chauffeurs (Vp. III) und des Mittelschülers (Vp. VII). Beim Chauffeur sind die Angehörigen sozial tieferstehender Berufe geneigt, die Unsicherheit des Sprechers als „unsympathisch“ zu werten, während die sozial und intellektuell Hochstehenden darauf hinweisen, daß die Sprechweise zwar nicht fließend, die Stimme aber wohlklingend und daher sympathisch sei. Der Arbeiter fühlt sich und seinen Stand durch das Stottern des Chauffeurs blamiert, er beurteilt ihn aus diesem verletzten Standesbewußtsein heraus. Ähnlich liegen die Dinge auch beim Mittelschüler; auch hier empört sich gewissermaßen das Standesbewußtsein des Schülers. Die allgemeine geschlechts- und altersbedingte Erscheinung der Mutation wird dann immer in irgendeiner persönlichen Form, als vom Sprecher verschuldet, umschrieben: „hat sich verkühlt“, „schreit zu viel“. Die intellektuellen Unterschiede äußern sich also nicht in der objektiven Richtigkeit der Deutung, sondern in der größeren oder geringeren Subjektivität der Beurteilung.

Umgekehrt gehen die geschlechtlich differenzierten Ergebnisse in bezug auf die objektive Richtigkeit der Deutung, wie wir sie

unter 1. darstellten, offenbar auf emotionale Einstellungsunterschiede zurück. Wir konnten ja schon in der Formulierung der Sätze im allgemeinen eine mehr sachlich-unpersönliche Einstellung der männlichen, eine mehr emotional-persönliche der weiblichen Einsender nachweisen. Sie verhilft letzteren zu einem Bessertreffen in der Erkenntnis des Geschlechts der beiden jugendlichen Sprecher und zu einer feineren Erfassung des Alters. Gerade für Altersnachweise aus der Stimme einer hochentwickelten Pflege des Aussehens gegenüber, die Altersspuren verwischt, sind offenbar die weiblichen Einsender besonders feinhörig. Gelegentlich führt allerdings dieses stark persönliche Interesse auch zu schlechteren Ergebnissen (größere Unterschätzung der ältesten weiblichen Vp.), da die größere Feinhörigkeit auch eine größere Empfindlichkeit dem subjektiven Alter gegenüber bedeuten kann.

Übrigens wird diese vorläufig vielleicht etwas willkürlich scheinende Interpretation der statistischen Ergebnisse im nächsten Kapitel vom Zustandekommen der Deutung her, die gerade in diesen zwei Punkten des Geschlechts und Alters stark resonanzmäßig verläuft, belegt werden.

Daß bei der Berufsdeutung eine Reihe rationaler Überlegungen, ein Aufsuchen verschiedener Momente in der Stimme, die als Deutungshilfen verwendet werden könnten, eine Rolle spielt, haben wir bereits gesehen: hier schneidet daher der gleichgeschlechtliche Hörer rein sachlich, durch seine größere akustische Erfahrung, besser ab. Wenn der gleichberufliche dagegen nicht besser trifft wie der Angehörige eines fremden Berufes, so kann dies nur dahin interpretiert werden, daß für ihn die allzu spezielle Erfahrung gegenüber der einheitlichen, das Typische betreffenden Erfahrung des fremdberuflichen keinen Vorteil bedeutet. So hat der bäuerliche Einsender zwar keine so große subjektive Sicherheit bei der Beurteilung des Privatdozenten, aber andererseits kennt er den intellektuellen Städter nur in einer beschränkten Anzahl ihm zu Typen gewordener Vertreter und hat damit eine einfachere, weil unbeschwertere Deutungsbasis.

Eine volle Ergänzung und Bestätigung des bisher Gesagten kann erst das Stimmerlebnis bringen. Bis zu einem gewissen Grad liegt sie auch in den Angaben einer Gruppe von Hörern vor, bei denen die Deutungsbedingungen objektiv einigermaßen rein und besser überschaubar sind: an unserem Versuch haben sich auch eine Anzahl *Blinder* beteiligt. Wir können ihre Deutungsergebnisse denen der Sehenden gegenüberstellen. Was ihre Stellungnahme

zum Experiment anlangt, so betonen sie ihre größere akustische Feinhörigkeit, erklären aber, daß damit natürlich der Entfall der optischen Wirklichkeit nicht wettgemacht werden könne, daß ferner das Radio die Stimme sehr vergrößere. Diese Äußerungen werden durch die Deutungsergebnisse drastisch belegt. Mit Ausnahme des Berufes treffen sie alle Punkte, an die ein quantitativer Maßstab überhaupt angelegt werden kann, besser!

Das *Geschlecht* der Mittelschülerin wurde von ihnen eindeutig erkannt, das des Mittelschülers um 26,8 % öfter getroffen als von den Sehenden. Die *Altersdeutung* ist, wenn auch nicht richtig, doch viel besser, d. h. die Über- bzw. Unterschätzungen sind bedeutend geringer. In der *Gestalterfassung* finden wir nicht nur eine eindeutigere Charakteristik in den zusätzlich gemachten Detailangaben, sondern auch ein Bessertreffen der allgemeineren Frage nach Größe und Dicke; und zwar wird die Größeschätzung der Sehenden, wo diese der Wirklichkeit widerspricht (beim unterschätzten Chauffeur, Vp. III) korrigiert und die Dicke-Schätzung verbessert.

In all diesen Belangen ist also der Blinde zufolge seiner größeren akustischen Geschultheit dem Sehenden gegenüber im Vorteil. Die mangelnde optische Erfahrung wird durch sie nicht nur kompensiert, sondern sogar übertroffen. Bei der Berufsdeutung aber schneidet der Blinde trotz seiner akustischen Feinhörigkeit nicht besser ab als der Sehende. Hier wird nämlich für die Deutung die völlige Erfahrung der Wirklichkeit relevant, die Stimme wird nicht nur direkt, sondern auch indirekt, mit Hilfe von Indizien, beurteilt und der Ausfall eines Zuganges zu solchen Indizien wird durch die größere Empfindlichkeit in anderen Gebieten nur aufgewogen.

Fassen wir zusammen: Betrachtet man das Problem „der Einsender und seine Antwort“ unter einem quantitativen, das ist statistischen Gesichtspunkt, so heißt das, daß man sich darüber klar wird, ob es allgemeine Zusammenhänge zwischen dem Einsender und der Richtigkeit der Deutung gibt. Wir konnten solche allgemeine Zusammenhänge in einem verhältnismäßig geringen Maß nachweisen. Wir fanden ferner, daß sie offenbar auf allgemeine einstellungs- oder schulungsbedingte Unterschiede der Hörer zurückzuführen seien. Mit der Feststellung der Bedeutung dieser beiden Faktoren ist aber auch die Unproduktivität einer quantitativen Untersuchung darüber erklärt. Wir brauchen zur Klärung

eine phänomenologische Betrachtungsweise, die auf das Einzelindividuum zurückgeht und sich nicht um die Richtigkeit, sondern um das erlebnismäßige Zustandekommen der Deutung kümmert.

3. Abschnitt

Das Stimmerlebnis

Kapitel 1: Methodisches

Die Phänomenologie war Ausgangspunkt unserer Überlegungen, Grundlage des Massenexperimentes: sie muß auch Endpunkt, Erläuterung der Resultate des Experiments sein. Allgemein gültige, letzte Sätze über „das“ Stimmerlebnis aufzustellen, geht daher über den Rahmen dieser Arbeit. Sie hätten die sorgfältige Analyse der Stimmerlebnisse einiger Vpn. gebraucht. (In den folgenden phänomenologischen Arbeiten dieser Serie wurde auch dieser Weg der Untersuchung beschritten.) Wir konnten uns hier mit weniger begnügen: d. h. wir konnten aus der Analyse einer Anzahl von Stimmerlebnissen eines Hörers Gesichtspunkte für den Fragebogen gewinnen, wir konnten sie, ergänzt durch die Zusätze der Hörer über das Zustandekommen der Deutung, für die phänomenologische Auswertung ihrer Ergebnisse heranziehen. Das Material dieses Kapitels sind also:

1. 30 von mir verfaßte Protokolle, die folgendermaßen angelegt wurden: während einer Woche habe ich bei fast allen Sprechern, die in Radio Wien zu irgendwelchen Zeiten über irgendwelche Themen sprachen, je 5 Minuten zugehört und alles, was ich beim Hören ihrer Stimme erlebte, möglichst unbefangen und genau protokolliert. Die Versuchssituation war also die des Massenexperimentes, nur der Akzent lag wo anders. Die längere Versuchsdauer (5 Minuten gegenüber 2 Minuten) war technisch notwendig. Es erwies sich als günstiger nicht nachträglich, sondern während des Erlebnisses zu protokollieren und dabei durfte die Unbefangenheit des Erlebens nicht durch Zeitmangel für das Stenogramm beeinträchtigt werden. (Ein lautes Diktieren während des Erlebens war wegen der akustischen Art der Darbietung störend.)

2. Die zweite Materialgrundlage sind die phänomenologischen Zusätze der Ravaghörer, soweit sie verläßlich schienen. Sie sind sehr zahlreich und oft sehr fein, aber sehr verstreut. Es mangelt ihnen im wesentlichen nur die Geschlossenheit. Sie dienten vor allem als Kontrolle meiner eigenen Stimmanalysen. Ergebnis:

Sie brachten eine volle Bestätigung der schon aus ihnen ableitbaren Gesetzmäßigkeiten, die im Folgenden an Hand dreier Protokolle dargestellt werden sollen.

Kapitel 2: Die Struktur des Stimmerlebnisses

Protokoll 1:

„Herzliche, warme Stimme — männlich tief. — Ist mittelgroß und mittelstark, weil er eine zwar kräftige, aber nicht sehr langatmige Stimme hat. — „Reiche Ernte“: das sagt er mit steigender Tonhöhe: freundlich. — „Ein“ wieder so die Stimme hinaufziehend: erinnert mich an R (ein Bekannter): hat auch so ein rötliches Gesicht, ist ebenfalls blond. — Schnelles Sprechtempo: der Sprecher ist lebhaft und jung. — Er artikuliert und betont die Endsilben so genau: ist offenbar aus niedrigem Milieu, hinaufgekommen, aber tüchtig. — Diese Sprechweise erinnert mich auch noch an jemand anderen; ich hab so das Gefühl, ich weiß die Richtung, wo ich suchen muß: „rund“, rollt dabei das „r“ so — wie A. (auch ein Bekannter): ist auch jung. Das ist auch der, den ich früher suchte. — Ich warte, ob die Ähnlichkeit mit der Stimme des A stark ist. Denn ich spüre noch einen ungeklärten Rest in der Stimme. Er drängt sich immer mehr vor, will gelöst werden: aber ich weiß nicht, was ich damit anfangen soll. Es fällt mir immer nur auf, was zu den schon erwähnten ähnlichen Stimmen paßt. Der Sprecher ist blond, freundlich, jung, mittelgroß, mittelstark, bürgerlich in Allem.“

Der Anteil für die Individualität des Hörers liegt, wie schon aus diesem einen Protokoll hervorgeht, vor allem in den „einfallenden“ Personen, d. h. den Bekannten, die erinnert werden, weil ihnen die Stimme des Sprechers in irgendeiner Beziehung ähnlich ist. Interindividuell konstant dagegen ist ihr struktureller Wert, d. i. der Platz, der ihnen in der Struktur des Stimmerlebnisses zukommt. Das im vorangegangenen Protokoll geschilderte Stimmerlebnis zerfällt nämlich (ebenso wie alle anderen Protokolle) ganz deutlich in 3 Phasen:

Die *erste* Phase enthält einen *resonanzmäßigen*¹ Eindruck, eine unmittelbare Reaktion auf die Stimme. Einerseits nimmt er in der Heraushebung eines Momentes an der Stimme schon das später

¹ Siehe KARL BÜHLER, Die Axiomatik der Sprachwissenschaften. *Kantstudien* 38 (1933).

durch Einzelbelege aus der Stimme gewonnene Urteil vorweg, das zwar differenzierter ist, aber eindeutig die Richtung des ersten Eindruckes festhält; oder anders formuliert: der erste resonanzmäßige Eindruck enthält bereits die Tendenz auf die besonders charakteristischen Momente. Andererseits wird durch ihn die affektive Seite des Stimmerlebnisses bestimmt, also die gefühlsmäßige Einstellung des Hörers zum Sprecher veranlaßt. Man könnte daher geradezu von einer „antizipierenden Funktion“ des ersten Eindruckes sprechen.

Im vorliegenden Protokoll wird der erste Eindruck, „herzliche, warme Stimme“ zwar im Verlauf des Erlebnisses durch weitere Momente an der Stimme erweitert, aber er stellt doch gewissermaßen das Grundgerüst für die weitere Deutung, d. i. für die einfallenden Personen und die auffälligen Momente an der Stimme dar.

Die *zweite* Phase des Stimmerlebnisses kann kurzerhand als ein Prozeß des „*Auffallens* — *Einfallens*“ bezeichnet werden. An der Stimme „fallen“ einzelne Momente „auf“, die für die Deutung verwertet werden. Diese Verwertung geschieht auf zweierlei Weise, direkt oder indirekt: *direkt*, indem aus diesen speziellen auffälligen Stimmqualitäten auf irgendwelche Persönlichkeitsmerkmale geschlossen wird (z. B. führt die Tonhöheschwankung in Protokoll 1, wie sie bei Aussprache von „reiche Ernte“ bemerkt wurde, zu dem Schluß: „freundlich“); *indirekt*, indem auf Grund dieser Stimmerkmale ähnliche Stimmen „einfallen“, d. h. erinnert werden, die für die Deutung verwertet werden (z. B. erinnert in Protokoll 1 die Aussprache des Wortes „ein“ an R, dessen Aussehen für die Deutung herangezogen wird). Diese beiden Deutungsmethoden, die indirekte und die direkte, unterstützen sich in der Regel, indem damit die auffälligen Momente an der Stimme möglichst erledigt werden. Im Gegensatz zu der Resonanz der ersten Phase erfolgt die Deutung hier auf Grund von Schlüssen. Der Hörer benimmt sich wie ein Detektiv, der gewisse Momente an der Stimme als Indizien in einem Schlußverfahren verwertet. „Indizienverfahren“ nennt denn auch KARL BÜHLER¹ diese zweite Deutungsmethode.“

Die *dritte* Phase wird am besten als „*Lösungsbewußtsein*“ gekennzeichnet. (In dem angeführten Protokoll 1 kommt dies sehr deutlich heraus.) Im Lösungsbewußtsein werden die Leistungen der direkten und indirekten Deutung der vorhergehenden Phase

¹ KARL BÜHLER, a. a. O.

registriert. Es werden also 1. die Einfälle auf ihre Berechtigung und ihren Geltungsbereich untersucht (das geschieht durch ein Abwarten, wie stark die Ähnlichkeit ist; oder durch Aufsuchen von weiteren stimmlichen Belegen für den Einfall). Und 2. wird festgestellt, was an der Stimme noch ungeklärt ist. Es gehen zwar von hier noch Tendenzen nach der weiteren Bewältigung der Stimme aus, aber es ist das Charakteristikum dieses Lösungsbewußtseins, daß die „Generalinventur“ eben erst am Schluß des Erlebnisses gemacht wird, d. h. dann, wenn man aus der Stimme nichts mehr herausholen kann. Dieser tote Punkt findet sich in jedem meiner Stimmerlebnisse, er erscheint auch in den Radio-Zuschriften häufig, wenn es z. B. heißt: „ich bin nur bis zu einem gewissen Punkt mit der Stimme fertig geworden“; oder „weiter sagt mir die Stimme nichts“. (In Protokoll 1 ist dieses Nichtweiterkönnen verschleiert durch eine Zusammenfassung der Deutung, indem gewissermaßen der ungeklärten Ähnlichkeit die Fülle von Ergebnissen gegenübergestellt wird.) Die Erklärung dafür ergibt sich aus der Art der Deutung: an der Stimme eines Sprechers, die ja an sich ein einmaliges individuelles Phänomen darstellt, fallen einzelne Momente auf, die die Stimme in irgendeiner Weise charakterisieren, d. h. aber ihr ein typisches Gepräge geben. Und über diese eklektische Erfassung der auffälligen Momente, die zu einer typischen Deutung führt, kommt der Hörer eigentlich nicht hinaus; früher oder später ergibt sich ein ungeklärter Rest, etwas, das akustisch wohl irgendwie wahrgenommen wird, aber nicht als auffälliges Moment klassifiziert werden kann, d. h. nicht in die Reihe der Stimmerfahrungen des Hörers einzuordnen und daher auch nicht zu deuten ist. Wir kommen darauf noch zurück.

Jedes Stimmerlebnis weist diese eben skizzierte Struktur auf. Natürlich sind aber Deutlichkeit und Bedeutung der einzelnen Phasen sehr verschieden stark betont, wie aus der folgenden, mehr inhaltlichen Betrachtungsweise noch hervorgehen wird.

Kapitel 3: Der Inhalt des Stimmerlebnisses

1. Der resonanzmäßige Eindruck ist einmal mehr *inhaltlich*, d. h. auf den Sprecher bezogen und einmal mehr *affektiv* gefärbt, d. h. das Ich des Hörers tritt stärker in den Vordergrund (in den Protokollen heißt es dann: „netter Mensch“, „sympathischer Mensch“). Diese stärkere affektive Färbung ergibt sich dann, wenn die Stimme keine so auffallenden Besonderheiten, vor allem

in der Klangfarbe, enthält, während deutlichere Modifikationen der Klangfarbe sofort mehr qualitativ erfaßt werden und die affektive Seite in den Hintergrund tritt. Damit habe ich schon die stimmliche Grundlage des ersten resonanzmäßigen Eindruckes erwähnt: sie liegt vor allem im Timbre. Diese Grundlage wird allerdings in den Protokollen nur manchmal und nachträglich, gewissermaßen als Beleg für die ungeheuer starke subjektive Sicherheit der resonanzmäßigen Deutung angeführt. Charakteristisch ist jedenfalls, daß hier für den Erlebenden das Primäre die Deutung ist, während in der 2. Phase die Stimmqualitäten bewußt für die Deutung verwertet werden.

Ob im weiteren Stimmverlauf nicht manchmal ein Urteil resonanzmäßig erfolgt, will ich nicht entscheiden. In meinen Protokollen kommt es nicht vor. Ich glaube auch, daß man das Resonanzprinzip in seiner Wirksamkeit nicht überschätzen darf und daß es mit dem 1. Eindruck ohnehin die wesentlichste, d. h. für den Deutungsfortgang wichtigste Leistung ausmacht.

2. Die zweite Phase, die inhaltlich weitaus am reichsten ist, soll an Hand von zwei weiteren Protokollen unter Berücksichtigung folgender Punkte etwas ausführlicher besprochen werden:

- a) Das auffällige Moment an der Stimme.
- b) Seine direkte Verwertung,
- c) Seine indirekte Verwertung: der Einfall.

Vorerst die 2 Protokolle.

Protokoll 2.

„Sympathischer Mensch. — Mitteltiefe, männliche Stimme; spricht gleichmäßig, aber lebendig schnelles Tempo, mittellaut: ist mittelgroß, mittelalt, nicht aufgereggt, schlank. — Spricht das Ende eines jeden Wortes etwas aufgerauht: raucht viel; gehört zu dem Typ der Menschen mit borstigen Haaren. — Schaut so ähnlich aus wie B.: hat auch so ein markantes Gesicht und blonde Haare. — Eigentlich habe ich eine solche Stimme noch nie gehört, aber ich werde dadurch weniger gestört. — Eben bemerke ich wiederum seine lebendige Sprechweise: er ist offenbar ein sehr optimistisches Temperament. — Scheint aber etwas hochmütig zu sein, weil er das Wortende immer tiefer und etwas schwingend spricht: ist sicher auch ein abgeschlossener Mensch. — Spricht ganz dialektfrei und ist sichtlich redegewandt: betont vernünftig, aber mit Nachdruck: ist offenbar ein Intelligenzler, hat Erfolg gehabt, ist energisch. — Hier, wo keine Ähnlichkeiten auftauchen,

bin ich in meinem Urteil viel sicherer. — Ich hab das Gefühl aus der Stimme alles Wichtige herausgeholt zu haben. — Es sind allerdings noch Feinheiten da, aber mit denen weiß ich nichts anzufangen“.

Protokoll 3.

„Sachlicher Mensch. — Zurückhaltend tiefe Stimme. — Spricht die Worte ganz gleichmäßig (~): ist eine abgeschlossene Persönlichkeit. — Merkwürdig ist die Aussprache des „en“. — Komisch, wie er bei „Kobolde“ das „ld“ sagt: so was hab ich schon gehört. — Auf der Universität. — „Bevor“ sagt er wie B., aber alles andere ganz anders. Ich weiß bestimmt, er ähnelt jemand anderem. Es liegt nicht nur an der Aussprache, sondern auch am Tonfall. — Plötzlich bei „ihnen“: redet so wie der H. — Ja, das stimmt! Aber dieser Sprecher ist jünger, weil er lauter spricht, schaut aber so aus wie H.: Gelehrtentyp; ist Gelehrter, sicher Intellektueller nach der Dialektfreiheit und Sachlichkeit. — Ist aber lebhafter, weil er schneller spricht; auch größer und stärker, weil er eine tiefere, langatmigere und lautere Stimme hat. — Im weiteren Sprechen bestätigt sich die Ähnlichkeit mit H. der Stimme nach. Aber es ist noch etwas anderes da. Ist es eine verschwommene Ähnlichkeit? Es läßt mich nicht mehr los, aber ich kann nicht herausfinden, was es ist.“

Ich komme nun zur Besprechung der einzelnen Punkte.

Kapitel 4: Das auffällige Moment

Die Ravag-Einsender beschrieben die Stimmen der Sprecher in ihren Zusätzen nach 6 Stimmdimensionen. In völliger Übereinstimmung damit können wir aus den 3 zitierten (und auch anderen) Protokollen konstatieren, daß die „auffälligen“ Momente nichts anderes sind als die individuellen Modifikationen dieser 6 Stimmdimensionen. Wir stellen im Folgenden ein paar Beispiele aus den 3 Protokollen zusammen:

1. Stimmhöhe und ihre Schwankungen (z. B. Prot. 1: „männlich tiefe Stimme; . . . sagt er mit steigender Tonhöhe: freundlich.“ Prot. 2: „mitteltiefe Stimme: mittelgroß . . . etwas hochmütig, weil er das Wortende tiefer spricht.“ Prot. 3: „zurückhaltend tiefe Stimme; . . . auch größer, weil tiefere . . . Stimme“).
2. Stimmstärke und ihre Schwankungen (z. B. Prot. 1: „ist mittelgroß und mittelstark, weil er eine zwar kräftige . . .“

- Stimme hat; . . . betont die Endsilben so genau: tüchtig.“
 Prot. 2: „mittellaute Stimme: schlank; . . . betont vernünftig, aber mit Nachdruck: Intelligenzler, hat Erfolg gehabt, ist energisch.“ Prot. 3: „ist jünger, weil er lauter spricht“).
3. Rhythmus (z. B. Prot. 2: „Spricht gleichmäßig, aber lebendiges Tempo: nicht aufgeregt; . . . spricht die Wortenden schwingend: sicher ein abgeschlossener Mensch.“ Prot. 3: „Spricht die Worte ganz gleichmäßig: ist eine abgeschlossene Persönlichkeit“).
4. Tempo (z. B. Prot. 1: „schnelles Sprechtempo: lebhaft, jung.“ Prot. 3: „ist lebhafter, weil er schneller spricht“).
5. Stimmtimbre (z. B. Prot. 1: „herzliche, warme Stimme.“ Prot. 2: „sympathischer Mensch“. Prot. 3: „sachlicher Mensch“).
6. Artikulation (z. B. Prot. 1: „artikulierte die Endsilben so genau: stammt aus niedrigem Milieu.“ Prot. 2: „dialektfrei, redegeübt: Intelligenzler.“ Prot. 3: „ist Gelehrter, sicher Intellektueller nach der Dialektfreiheit und Sachlichkeit“).

Vom phonetischen Standpunkt ist natürlich die menschliche Stimme in diesen 6 Stimmdimensionen nicht erschöpfend zu beschreiben. Aber sie drängen sich dem die Stimme erlebenden und sie deutenden (Laien-)Hörer auf, er hebt sie als „auffällig“ heraus — wobei wiederum über ihre etwaige phonetische Zerlegbarkeit nichts ausgesagt werden soll — und macht sie für seine Deutung, soweit er dafür überhaupt Kriterien angeben kann, oder besser gesagt, spontan angibt, verantwortlich. Phänomenologisch erweisen sich also diese 6 Stimmdimensionen als relevant. Ihre Wertigkeit ist dabei eine stark variable. Es kommen weder alle diese Dimensionen in jedem Protokoll vor, noch kommt ihnen, wenn sie vorkommen, überall die gleiche (konstante) Gewichtigkeit zu. An der individuellen Stimme werden jeweils verschiedene Momente auffällig und ihre Deutungsergiebigkeit ist eine verschiedenen starke. (Oder vom Resultat her: es gibt für jeden Sprecher besondere, charakteristische Momente.)

Abgesehen von der Frage nach der Art der Verwertung und etwaigen dabei nachweisbaren Gesetzmäßigkeiten, die wir im nächsten Kapitel behandeln wollen, erhebt sich die Frage: inwieweit an diesem Auffälligwerden die individuelle Persönlichkeit des Hörers mitbeteiligt ist und wieso es überhaupt dazukommt, daß dem Hörer ein Moment an der Stimme „auffällt“. Man muß hier scheiden nach der Verwendung des auffälligen Momentes:

wenn es auf Grund einer Ähnlichkeit auffällt, d. h. zu einem Einfall und zu indirekter Verwertung führt, ist es natürlich stärker durch den Hörer, durch seine individuelle Stimmerfahrung bedingt. Wenn es an sich auffällt und direkt gedeutet wird, ist die Abhängigkeit von der Individualität des Hörers sichtlich geringer. Das ergibt sich auch ganz deutlich aus den Zusätzen der Ravag-Einsender über die Stimme. Denn wir fanden dort, daß die Hörer in der Auswahl der Qualitäten, die sie für die einzelnen Sprecher besonders hervorheben, und auch in der inhaltlichen Bestimmung dieser Qualitäten weitgehend übereinstimmten. Man kann also sagen, daß das *Auffälligwerden* stimmlicher Momente, soweit es nicht durch ganz individuelle Ähnlichkeit bedingt wird, ein sichtlich *allgemeines* ist, daß es offenbar allgemeinen Stimmgesetzen unterliegt.

Damit kommen wir zu der zweiten Frage, zu dem *Warum* des Auffälligwerdens. Für den Erlebenden ist es gegeben durch das Wie des auffälligen Momentes: eben weil die Stimme in dieser oder jener Beziehung so oder so geartet ist, wird dies bemerkt. Nun kann aber etwas auffallen nur in Hinblick auf etwas anderes, auf Grund einer Ähnlichkeit oder eines Kontrastes. Die Veränderungen der Stimme eines *bekannt*en Menschen, hervorgerufen durch seine aktuellen Erlebnisse, bemerken wir in Hinsicht auf seine uns bekannte „Normalstimme“. An der Stimme eines *unbekannt*en Sprechers fällt mir etwas auf in Abhebung von meinen gesamten Stimmerfahrungen. Jeder Mensch hat solche Stimmerfahrungen. Daraus erklärt sich die Übereinstimmung unter den Hörern. Manchmal trifft nun die spezifisch modifizierte Stimmqualität des unbekannten Sprechers eine spezielle Stimmerfahrung des Hörers; dann reagiert er speziell mit dem, was er aus dieser Stimmerfahrung weiß (durch Verwertung des Einfalles). In anderen Fällen fehlt diese spezielle Beziehung, der Hörer reagiert allgemein und direkt aus der Gesamtheit seiner Stimmerfahrungen. Die individuellen Unterschiede zwischen den Hörern sind im letzteren Fall gering, im ersteren offensichtlich größer. Das wird noch verständlicher, wenn man die Art der Deutung berücksichtigt.

Kapitel 5: Die direkte Deutung

Vergleicht man die angeführten Protokolle, so ergibt sich ein viel geschlosseneres Resultat in Protokoll 2, bei weitgehend direkter Deutung. Die auftauchenden Ähnlichkeiten, ihre Identifizie-

rung nimmt die Aufmerksamkeit des Hörers übermäßig in Anspruch und lenkt ihn von dem unbefangenen Erleben und Beurteilen der Stimme, wie es ohne spezielle Ähnlichkeiten möglich ist, ab. Das Urteil ist bei direkter Deutung sicherer, nicht nur weil bei der indirekten Deutung die Einfälle verwirren und ablenken, weil im ersten Fall die Stimme an und für sich beurteilt wird, während sie im zweiten Fall nur vergleichsweise beurteilt wird, sondern auch deshalb, weil bei der direkten Deutung die aus der gesamten Stimmerfahrung gewonnene Allgemeinkenntnis spricht, während sie bei Verwertung des Einfalles zugunsten eines Spezialwissens — das allerdings auch dieser allgemeinen Gesetzmäßigkeit unterliegt — zurücktritt.

Trotz der breiten allgemeinen Basis, auf die sie sich stützt, erfolgt die direkte Deutung durchaus *individuell*, von der gegebenen Stimme ausgehend. Genau so, wie sich der unbefangene Erlebende nicht nach dem eigentlichen Grund für das Auffälligerwerden der einzelnen Momente an der Stimme des Sprechers fragt, sondern ihr Auftreten als ganz natürlich hinimmt, kümmert er sich bei der Deutung nicht um ihre eigentlichen Grundlagen. Er schafft die Zuordnung scheinbar *ad hoc* und folgt dabei durchaus dem individuellen Stimmcharakter. Darin äußert sich das Erlebnismäßige der Stimmdeutung, das Einwirken eines individuellen Sprechers auf einen individuellen Hörer.

Betrachtet man aber die Leichtigkeit und Sicherheit der Zuordnung im einzelnen Erlebnis, die Konsequenz in der Art der Zuordnung mehrerer Erlebnisse und die Allgemeinheit der Zuordnung, wie sie sich aus dem Vergleich mit den Begründungen der Angaben der Ravag-Einsender ergibt, so muß man annehmen, daß es zwischen den angegebenen Merkmalen des Sprechers und bestimmten Stimmerkmalen bestimmte gesetzmäßige Zusammenhänge gibt und daß sie, aus den Erfahrungen des täglichen Lebens geläufig, wenn auch vielleicht nicht *explicite* gewußt, das Fundament jeder Stimmdeutung darstellen. Sehen wir uns daraufhin die in Kapitel 4 gegebene kurze Zusammenstellung der Verwertung der auffälligen Momente an, so können wir tatsächlich eine solche Spezifität der Ausdruckshaltigkeit der einzelnen Stimmdimensionen feststellen:

Die *Stimmhöhe* erscheint in meinen Protokollen als ein Kriterium für die Geschlechts-, Alters- und Größenbestimmung.

Diese Art der Auswertung hat offensichtlich ihre physiologische Berechtigung: bei Männern ist tatsächlich der Durchmesser des Kehlkopfes wesentlich größer, dadurch die Stimmbänder länger und die Stimme tiefer. Die Altersveränderung in der mutierenden Knabenstimme ist bedingt durch physiologische Veränderungen des Kehlkopfes. Bei Erwachsenen hat SEEMANN¹ Stimmhöheveränderungen im Sinn einer Abnahme der Höhe mit zunehmendem Alter feststellen können.

Für den Zusammenhang zwischen Stimmhöhe und Größe konnte SCHILLING² ein Zunehmen der Bässe mit Zunehmen der Größe gegenüber einem Abnehmen der Tenöre nachweisen; ebenso einen Zusammenhang zwischen Langköpfigkeit und Stimmhöhe, der mit den Resultaten DALMAS³ übereinstimmt, der ein frappantes Überwiegen der Hochstimmigen unter den (langköpfigen) Schizophrenen nachwies.

Außer dieser (objektiv auf ihre Berechtigung überprüfbaren) physiologischen Ausdrucksfunktion hat die Stimmhöhe, wie jede andere Stimmdimension, auch eine *psychologische* Ausdrucksfunktion. Denn es scheint so, als ob bestimmte charakterologische Formulierungen vor allem in Hinblick auf eine bestimmte Stimmdimension erfolgten. So wird die Stimmhöhe, besonders aber die Stimmhöheschwankungen, für Angaben verwertet, die wir in der Gruppe der „Auseinandersetzung mit der Außenwelt“ zusammenfaßen; obgleich gerade hier der Eindruck der ganzen Stimme, der Kombination, in der die Dimension auftritt (die in jeder Deutung eine Rolle spielt), besonders wichtig ist, und auch für die Nuancen der Formulierungen im Verein mit den verschiedenen Abstufungen der Stimmhöheschwankungen ausgewertet wird. Die Innerlichkeit ist etwas viel zu Komplexes und wird auch als solches erlebt, als daß sie vom Hörer so schematisch, wie wir es aus rein technischen Gründen bei der Darstellung der Ravag-Ergebnisse taten, in Kategorien eingeteilt würde. Der Deutungsvorgang ist immer der: hier die ganze Stimme — dort die Persönlichkeit, die charakterisiert wird. Die Spezifität der psychologischen Ausdrucksfunktion will nur besagen, daß in der Deutung die Tendenz besteht, eine spezielle Stimmdimension herauszuheben, weil sie für den Sprecher besonders charakteristisch ist und sie in einem ganz speziellen Sinn zu formulieren. Bezeichnenderweise erfolgte die

¹ N. SEEMANN, Les fonctions sexuelles et la croix, *Otolaryngologica slavica* 1930.

² N. SCHILLING, Stimmuntersuchungen an Studenten an der Universität Freiburg. Ber. 2. Vers. d. dtsh. Ges. f. Sprach- u. Stimmheilkunde, 1928.

³ G. DALMA, Körperbau und Psychose, mit Berücksichtigung der konstitutionellen Bedeutung der Stimme. *Z. f. Neurol. u. Psychiat.* 97, 1925.

psychologische Auswertung der Stimmhöhe vor allem auf Grund der Stimmhöheschwankungen, die dem Belieben des Sprechers in höherem Maße überlassen sind, während die absolute Stimmhöhe vor allem physiologisch ausgewertet wurde. Dabei wurden für die Deutung erst die Schwankungen (am Silben-, Wort- oder Satzende) wichtig, die über die sprachlichen Bindungen hinausgingen.

Die Berechtigung dieser psychologischen Deutung können wir nur so einigermaßen überprüfen, daß wir vom primären Ausdruckswillen her sehen, welche Veränderungen wir mit der Stimme vor allem vornehmen, wenn wir bestimmte psychische Daten ausdrücken wollen. Man erhebt tatsächlich seine Stimme, wenn man an die anderen appellieren will. Die Auswertung erscheint also nicht ganz ungerechtfertigt.

In ähnlicher Weise wie die „Ausdrucksfunktion“ der Stimmhöhe lassen sich auch die der anderen Stimmdimensionen im physiologischen und psychologischen Bereich verstehen. So wird die *Stimmstärke* (physiologisch bedingt, unter anderem, durch die Stärke des Atemdruckes, also abhängig vom Bau der Brustmuskulatur und der Intaktheit der Atmungsorgane) für die Alters-, besonders aber für die Größe-Dickeangaben ausgewertet. Von diesem Zusammenhang mit der körperlichen Leistungsfähigkeit erscheint auch die *psychologische* Auswertung für Angaben, die wir als „Fähigkeiten“ zusammenfaßten, ganz plausibel. Laut spricht, wer sich durchsetzen will, wer was zu sagen hat, leise spricht man z. B. im Trauerhause, wenn man damit seine Rücksicht dokumentieren will.

Wichtig ist bei der Deutung der Stimmstärkeschwankungen die starke sprachliche Bindung im Deutschen (expiratorischer Akzent). Die Hörer sind deshalb auch zurückhaltender, ihre Beurteilung erfolgt häufig als eine Beurteilung der „Betonung“.

Das Moment der Schwankung, in den Stimmhöhe- und Stimmstärkeschwankungen bereits enthalten, wird im *Rhythmus* und im *Tempo* als auffällig erlebt und ausgewertet. Beide, als zwei eigentlich abgeleitete Stimmdimensionen, werden primär für psychologische Daten, erst sekundär (und zwar durch die Übersetzung psychologischer Daten in Physiologisches, z. B. der „Lebendigkeit“ in „Jugend“) für die Altersbeurteilung herangezogen. Der Rhythmus als das Gestaltmoment (erlebt als Regelmäßigkeit, Kontinuität, Form und Ausgeprägtheit der Schwankungen) wird ausgewertet für Formulierungen, die wir unter „innerer Harmonie“ zusammenfaßten. Das Tempo als Geschwindigkeit der Sprechbewegung wurde ausgewertet für das „Temperament“, das gewissermaßen die Geschwindigkeit der inneren Abläufe bestimmt.

Wir verstehen von hier, warum in den Zusätzen der Ravag-Hörer die Angaben zur „inneren Harmonie“ so gering sind. Denn der Rhythmus wirkt der Tendenz nach sinnvoller Gliederung ausgesprochen entgegen und konnte deshalb in der Versuchssituation nicht wirklich zur Geltung kommen.

Die formalen Dimensionen, Rhythmus und Tempo, sind ferner durch ihre große Bedeutung für die Erkenntnis der erlebnisbedingten aktuellen Veränderungen der Persönlichkeit charakterisiert. So wurde die „Aufgeregtheit“ des Chauffeurs immer wieder aus der „ungleichmäßigen“ und „stammelnden“ Sprechweise entnommen. Im übrigen wollen wir uns in betreff der Deutungsgrundlagen für die Erkenntnis der Erlebnisse eines Sprechers jeder Behauptung entschlagen, die über unser Material hinausginge.

Unter allen Stimmdimensionen ist der Hörer zur Erfassung des *Timbres* als des Spezifikums jeder individuellen Stimme am geschultesten¹, weil er hier die Erfahrungen aus allen Sinnesgebieten heranzieht, weil er es synästhetisch erfaßt und dabei nicht nur die Grenzen zwischen den einzelnen Sinnesgebieten („warme“, „spitze“, „klangvolle“ Stimme), sondern auch die zwischen der Sinnenwelt und Psychischem aufhebt („gütiges“ Timbre). Das unbefangene Hereinziehen der anderen Sinnesgebiete, vor allem im Sinn von Farbe und Form, bestimmt auch die Deutung. Die Urteile über Haar- und Augenfarbe, soweit sie nicht durch Einfälle zustandekamen, gehen darauf zurück; ebenso kommt es auf diese Weise zu Gestaltübertragungen („scharfe Stimme — scharfes Profil“). Gewisse physiologische Grundlagen sind aber auch hier nachzuweisen: das scharfe Timbre entsteht z. B. durch strenge Abschließung der Resonanzräume und feste Artikulation, die beide die ganze Muskulatur, also auch die Gesichtsmuskeln bewegen und markant machen, da sie Fettansatz verhindern. Aber es braucht jedenfalls zu einer definitiven Bestimmung gerade hier, ebenso wie bei den Farbübertragungen, noch ausführlicher Untersuchungen.

Die synästhetische Auffassung des *Timbres* äußert sich auch in seiner *charakterologischen* Auswertung. Es wird für Angaben betreffs der „inneren Art“ verwertet, die etwas charakterologisch offenbar sehr Wichtiges (s. d. große Prozentzahl der Zusätze des Ravag-Experimentes), aber auch sehr schwer Faßbares darstellen.

¹ Vgl. dazu die Bedeutung, die TENNER, „Über Versmelodie“, *Zschr. f. allg. Kunstsch.* 8, 1913, der Klangfarbe für die Sprache zuschreibt.

Wir bemühen uns darum auch sonst mit synästhetischen Umschreibungen, wie z. B. „innere Kälte“, . . .

Die *Artikulation*, d. i. die Art des Zustandekommens der einzelnen Laute, physiologisch bedingt durch Bau und Stellung der Artikulationsorgane, wird für Angaben über die Art und Gesundheit der Zähne ausgewertet. Scheinen sie noch einigermaßen berechtigt, so sind Altersbeurteilungen nach der Artikulation bereits viel gewagter und ebenso auch Gestaltübertragungen. So heißt es z. B. in Prot. 14: „sagt das „i“ so hell, hat ein breites freundliches Gesicht.“ Auch hier fehlen uns wie beim Timbre für Aussagen über die Berechtigung solcher Übertragungen vorläufig noch die Grundlagen.

Die physiologische Auswertung der Artikulation tritt aber ganz zurück gegenüber ihrer *sozialpsychologischen* Bedeutung. Denn die Artikulation ist keine sechste völlig ebenbürtige Stimmdimension, sie zeigt nur, wie die eigentlichen Stimmdimensionen in der sprachlich bedingten Stellung der Artikulationsorgane wirksam werden. Die Artikulation ist lernbar, sie ist milieubedingt und wird vor allem als Art und Stärke des „Dialektes“, des „Akzents“ erlebt und beurteilt. Im Verein mit dem Timbre, durch das der Interessenkreis des Sprechers erkannt wird, bietet sie die Indizien für die Berufsdeutung.

Die stark sozialpsychologische Ausdruckshaltigkeit der Artikulation drängt auch die rein psychologische Auswertung ganz in den Hintergrund. Wenn sie erfolgt, hat sie einen stark sozialen Beigeschmack. Man denke nur an das „Näseln“ des Hochnäsigen (Aristokraten), die „affektierte Artikulation“ (des Schauspielers).

Damit wäre in Kürze skizziert, wie nach unserem Material die direkte Deutung vor sich geht, welches ihre Grundlagen sind. Dabei ist im Rahmen dieser Fragestellung das Zuordnungsschema nicht als letztes Ergebnis, sondern nur als ein erster Ansatz zu werten, der versucht, die Gesetzmäßigkeiten herauszuarbeiten, die in den zahlreichen Begründungen unseres Materials enthalten schienen. Wir möchten ferner betonen, daß die „Ausdrucksfunktion“ der einzelnen Stimmdimensionen nicht als eine so strenge aufzufassen ist, wie es vielleicht bei der Darstellung herauskam: es wirkt immer, gewissermaßen als Hintergrund der Deutung, die ganze Stimme mit. Das gilt für die direkte Deutung, das gilt in erhöhtem Maß für die indirekte Deutung.

Kapitel 6: Die indirekte Deutung

Unter der indirekten Deutung verstehe ich die Auswertung der Einfälle. Sie spielen im individuellen Stimmerlebnis eine große Rolle; und zwar ergibt sich dies nicht nur aus meinen Protokollen, sondern wir können dafür auch auf die Zusätze der Ravag-Hörer hinweisen, die in 11% von Identifizierungen und Ähnlichkeiten der Stimmen berichteten. Wenn ich im Folgenden auf die drei wichtigsten Punkte, das Auftreten, die personale Bestimmtheit und die Verwertung der Einfälle näher eingehe, so stütze ich mich dabei natürlich vor allem auf die Einzelanalysen, die einen durchaus individuellen Beitrag darstellen, kann aber die einzelnen Sätze auch durch gelegentliche Angaben der Ravag-Hörer, allerdings in geringerem Maß als für die direkte Deutung, stützen.

1. Das Auftreten der Einfälle

In Protokoll 1 werden beide Einfälle durch ein stimmlich-sprachliches Moment ausgelöst: es fällt überhaupt auf und erweist sich außerdem als ähnlich der Sprechweise eines Bekannten. Protokoll 1: „Ein“, das sagt er so hinaufziehend: erinnert mich an R. . . . „rund“, rollt dabei das „r“ so — wie A. In Protokoll 3 fällt die Artikulation überhaupt auf; bei einem besonderen Beispiel wird B. erinnert und schließlich H., der der Stimme des Sprechers nicht nur nach der Artikulation, sondern auch nach dem Tonfall ähnlich ist. Diese *eigentlichen* Einfälle kommen also dadurch zustande, daß ein Moment, das überhaupt auffällig ist in seiner ganz speziellen Färbung sich als ähnlich erweist mit der Stimme eines bekannten Menschen. Die Stimme, die dem Hörer einfällt, ist also immer in irgendeiner Weise auffällig und wird auf Grund ihres speziellen So-seins erinnert.

In Protokoll 2 haben wir es diesen *eigentlichen* Einfällen gegenüber mit einem *uneigentlichen* Einfall zu tun, der nicht durch die Stimme des Sprechers ausgelöst wurde. Es heißt hier: „schaut so ähnlich aus wie B.: hat auch so ein markantes Gesicht und blonde Haare“. Nachträglich wird die stimmliche Kompetenz dieses Einfalles abgelehnt dadurch, daß erklärt wird, die Stimme sei überhaupt in keiner Weise ähnlich. Ich hatte B. am Vortag, zum erstenmal gesehen und überhaupt nicht sprechen gehört, aber er war mir durch sein markantes Gesicht aufgefallen. Die Funktion dieses

Einfalles besteht offenbar nur darin, für die aus der Stimme gewonnene Erkenntnis des Aussehens beispielsweise einen Repräsentanten anzugeben: es fällt daher jemand ein, dessen optische Vorstellung sehr frisch und stark ist.

Eine zweite Einteilung der Einfälle ergibt sich nach der Schnelligkeit ihres Auftretens: der erste Einfall in Protokoll 1 („Ein“, das sagt er so hinaufziehend: erinnert mich an R.“) ist ein *spezifisch-plötzlicher*, d. h. es fällt sofort die Person ein, deren Stimme ähnlich ist (dasselbe gilt für den ersten Einfall in Protokoll 3: „bevor“ sagt er wie B.) oder der Einfall ist ein *allgemein-allmählicher*, d. h. es tritt zuerst nur ein allgemeiner Bekanntheitseindruck auf, der die bewußte Mitarbeit des Erlebenden zu seiner Bewältigung erfordert (Beispiel dafür ist der zweite Einfall in Protokoll 1 und Protokoll 3). Die Mitarbeit des Erlebenden besteht darin, daß der Stimme eine erhöhte Aufmerksamkeit zugewandt wird, die darauf aus ist möglichst viele Ähnlichkeiten festzustellen, bis durch die Häufung der Beobachtungen und durch die inhaltlich reichere Stimmerinnerung (da ja immer nur so viel erinnert wird, als durch die sprechende Stimme ausgelöst wird) die Spezialisierung gelingt. Es wird also versucht: 1. die Stimmähnlichkeiten möglichst exakt festzustellen und damit die Stimme des Bekannten möglichst weit zu erinnern; 2. durch Überlegung, eventuell auch durch Abzählen der in Betracht kommenden Möglichkeiten eine irgendwie örtlich orientierte Fixierung zu treffen (Protokoll 1: „ich weiß die Richtung“; Protokoll 3: „auf der Universität“), die angibt, wo die ähnlichen Stimmen zu suchen sind.

Die Parallele auf optischem Gebiet: Wenn ich z. B. in der Straßenbahn jemanden sehe, der mir irgendwie bekannt vorkommt, werde ich versuchen, festzustellen, was alles mir an diesem Gesicht bekannt vorkommt und mich fragen, wo ich dieses Gesicht schon gesehen habe. — Möglichkeit und Nutzen der örtlichen Fixierung erklärt das Folgende:

2. Die personale Bestimmtheit der Einfälle

a) Ihre Beziehung zum Hörer: Ich habe eine Liste der in allen 30 Protokollen vorkommenden Einfälle (58) angelegt und daraus ergibt sich eindeutig, daß auf Grund von Stimmähnlichkeit nur Leute einfielen, die ich ihrem *Wesen* nach *nicht* sehr gut kenne, von denen mir aber die *Stimme* verhältnismäßig *gut* bekannt ist, die ich oft sprechen höre. Es sind vor allem Professoren, Studenten, Geschäftsleute, von denen ich alles, was ich über sie

weiß, durch ihre Stimme weiß, bei denen mir also die Stimme als Ausdruck für ihre Persönlichkeit ganz geläufig ist.

Offenbar ist die Stimme der Leute, die man nur nach der Stimme kennt, einem relativ wichtiger und daher geläufiger als die Stimme von Menschen, die man überhaupt sehr gut kennt. Ganz parallel dazu zeigt sich — ich habe darüber 10 Leute befragt —, daß manche Menschen sich die Stimmen der nächsten Angehörigen, also der Eltern und Geschwister, relativ schwerer vorstellen können als die Stimmen von Leuten, die sie nur der Stimme nach kennen. Außerdem ergibt sich, daß ich alle diese Leute, die mir einfielen, innerhalb des letzten Jahres gesehen hatte. Die örtliche Fixierung beim Auftreten des Bekanntheitseindrucks erweist sich also als möglich und berechtigt, weil tatsächlich vor allem Leute einfallen, die man nur an bestimmten Orten trifft, die man vor nicht allzulanger Zeit dort gesehen hat.

Ebenso läßt sich von hier aus verstehen, daß (z. B. in Protokoll 3) so viel Zeit verstreicht, bis der allgemeine Bekanntheitseindruck tatsächlich gelöst ist, bis die Person mit der ähnlichen Stimme erinnert wird. Es kommt sichtlich darauf an, wie oft man die Stimmen der einfallenden Leute hörte. H., der mir in Protokoll 3 so lange nicht einfallen wollte, ist ein Professor, dessen Kolleg ich nur sehr wenig gehört hatte. Ob die ähnliche Stimme sofort oder erst später einfällt, hängt natürlich außer der Stärke und Frische der Erinnerung auch davon ab, ob die Stimme der einfallenden Person sehr stark und eindeutig modifiziert ist und ob die Ähnlichkeit mit der Stimme des Sprechers eine sehr große ist.

Mit der Feststellung, daß in meinen Protokollen nur weniger bekannte Personen einfielen, ist natürlich nicht gesagt, daß bei Ähnlichkeit eine sehr gut bekannte Person nicht einfallen würde, sondern nur, daß ich die Stimmen weniger gut bekannter Menschen relativ besser kenne und daß sie mir daher eher einfallen. Aber was die Einfälle für den Fortgang der Deutung bringen, ist gering; meist sind es nicht-charakterologische Daten, weil die Kenntnis der Innerlichkeit durch das oberflächliche Bekanntsein eine sehr unvollständige ist. Die *Stimmerfahrungen* mit Menschen, die wir *sehr gut kennen*, bilden die Grundlage des allgemeinen Wissens über den Zusammenhang zwischen Stimme und Persönlichkeit, das in der direkten Deutung verwertet wird und leisten hier die viel werktätigere Hilfe für den Deutungsfortschritt. Die *Stimmerfahrungen* mit *weniger gut bekannten* Menschen sind stimmlich

lebhafter, sie äußern sich in individuellen Einfällen, aber ihre Funktion liegt nicht in erster Linie darin, den Deutungsfortschritt zu fördern, sondern vor allem in einer Verlebendigung des Stimmerlebnisses, in der Einreihung der unbekanntem Sprecher in die Welt der bekannten Menschen, die einem persönlich nicht allzu nahe stehen, so wie der Sprecher, den man zum erstenmal hört.

b) Die Beziehung zum Sprecher. Die eigentlichen Einfälle werden, wie wir zeigten, ausgelöst durch eine ganz spezifische ähnliche Modifikation der Stimme des Sprechers innerhalb einer allgemeinen Auffälligkeit; und zwar wird immer nur soviel von der Stimme der bekannten Person erinnert als durch die momentan sprechende Stimme ausgelöst wird. Die Einfälle sind also, soweit sie sprachlich bedingt sind, durchaus individuell, der speziellen Qualität folgend. Allerdings mit einer gewissen Ausnahme! Wenn das auffällige Moment ein sehr allgemeines und *typisches* ist, z. B. „Dialekt“ oder ein „fremder Akzent“, so fand ich in meinen Protokollen (das müßte allerdings weiter ausgeführt werden), daß zwar auch hier die individuelle Modifikation den Vorrang hat und den speziellen Einfall hervorruft, daß aber immer noch, gewissermaßen als allgemeiner Vertreter des Dialektes, jemand einfällt, der für mich das Prototyp des leichten Dialektsprechers überhaupt ist, der in allen Protokollen, wo es sich um etwas dialektisierende Leute handelt, wiederkehrt, allerdings ohne für die Deutung, die dem speziellen Einfall folgt, viel verwertet zu werden.

Ich bringe als Beleg dafür die entsprechenden Teile aus 4 solchen Protokollen:

Protokoll 4: „Gutmütig. — Leichter Dialekt: merkwürdig reine „a“: so wie der R. — „den Kindern“ sagt er genau so wie B.: ist auch mittelalt, behäbig . . .“

Protokoll 5: „Nicht sehr sympathisch. — Mittelgebildet, Klasse des R. — Nach Aussprache von „letzten“, „auch“ erinnert er mich an S.: ist auch etwas derb, blond . . .“

Protokoll 6: „Etwas unsympathischer Mensch. — Merkwürdige Aussprache von „Architektur“, „Perspektive“: ungarischer Akzent so wie W. — Sagt gerade „Kitsch“: so ähnlich wie F.: hat auch so ein markantes Gesicht . . .“

Protokoll 7: „Gütiger Mensch. — Gebildet, weil er sinnvoll betont. — Gepflegte Stimme, aber leicht ungarischer Akzent: schwächer wie bei W. — „Kraft“ sagt er genau so wie U.: rötliches Gesicht, Brille so wie U. . .“

Beim *eigentlichen* Einfall besteht die Beziehung zwischen Sprecher und eingefallener Person in einer spezifisch ähnlichen Modifikation der beiden Stimmen; der *uneigentliche* Einfall (siehe

Protokoll 2) wird durch eine allgemeine, aus der Stimme gewonnene optische Deutung angeregt, die er weiter spezialisiert. Ein Mittelglied zwischen beiden stellt der Einfall dar, der durch eine *allgemeine akustische* Deutung angeregt wird (siehe Protokoll 4 bis 7), die er ebenfalls spezialisiert. Der uneigentliche Einfall, der durch eine optische Deutung angeregt wird, ist ihm gegenüber für die Verwertung wichtiger, weil hier die Stimme keine so direkten Anhaltspunkte enthält, während der akustisch angeregte typische Einfall keine Bedeutung für den Deutungsfortschritt hat, weil hier die Stimme selbst weitere Modifikationen enthält und ihnen entsprechend durch spezielle Einfälle deutbar ist. Darüber im Folgenden mehr.

3. Die Verwertung der Einfälle

Es hängt offenbar sehr stark von der Persönlichkeit des Hörers ab, inwieweit er die Daten von eingefallenen Personen auf den Sprecher überträgt. Lassen wir die Frage der Berechtigung der Verwertung vorerst ganz beiseite und kümmern wir uns nur um die *tatsächliche Verwertung*, so ist allgemein Folgendes festzustellen: der Unterschied gegenüber der direkten Deutung liegt darin, daß dort der Hörer an die Deutung mit einem ganz bestimmten Zuordnungsschema herangeht. Das fehlt hier, der speziellen Natur des Einfalles gemäß. Die Zuordnung geschieht tatsächlich ad hoc, wobei zweierlei festzustellen ist:

1. Die Deutung erfolgt in meinen Protokollen und auch in den Ravag-Zuschriften immer *in Hinblick auf die sprechende Stimme*, darauf, was der Stimme „zuzutrauen ist“, wie es in einer Zuschrift heißt. Daraus erklärt sich die Leichtigkeit und Unbefangenheit der Zuordnung. Die Einfälle sind eigentlich nur Anregungen für etwas, wozu die Stimme selbst Belege enthält.

2. Die Verwertung der Einfälle erfolgt, wie wiederum in meinen Protokollen und in den Angaben der Radiohörer übereinstimmend herauskommt, in erster Linie für *nicht-charakterologische* Angaben, vor allem für Daten des *Aussehens*. Für die charakterologische Deutung ist die Stimme selbst so ausdrucksstark, bringt der Hörer eine solche Allgemeinschulung durch den Alltag mit, daß die direkte Deutung bei nicht allzu zwingender Ähnlichkeit der indirekten vorgezogen wird. Die Verwertung der Einfälle für das Aussehen erscheint gerechtfertigt damit, daß hier die Schulung des Hörers durch den Alltag für eine direkte Erkenntnis

eine viel geringere ist: in der Regel sieht man die Leute, mit denen man spricht, aber ihr Inneres muß auf verschiedensten Wegen, und auch durch die Stimme, erschlossen werden. Außerdem liegt natürlich die optische Verwertung eines optischen Vorstellungsbildes sehr nahe, zumal der Hörer gemäß der personalen Bestimmtheit der Einfälle das Aussehen der eingefallenen Personen am besten kennt.

Ad 1. Zur *Kontrolle* durch die Stimme: Die Kontrolle durch die sprechende Stimme kommt z. B. in Protokoll 3 deutlich heraus, wenn es heißt: „Redet so wie der H., aber dieser Sprecher ist jünger, weil er lauter spricht . . . lebhafter, weil er schneller spricht . . . auch größer und stärker, weil er eine tiefere, langatmigere und lautere Stimme hat.“ Nur aus dieser Vormachtstellung der sprechenden Stimme ist es auch zu verstehen, daß in Protokoll 1 zwei Einfälle (in manchen Protokollen auch noch mehr), die doch zwei verschiedenen Menschen zugehören, nacheinander auftreten können und einander durchaus nicht stören. Es erklärt sich daraus, daß 1. die Zweiheit von eingefallener und sprechender Person immer festgehalten wird (es heißt: „Hat auch ein rötliches Gesicht“) und 2. die Einheit der wahrgenommenen Stimme für den ganzen Deutungsablauf das Primäre ist und die ähnliche Stimme nur soweit erinnert wird als die Ähnlichkeit reicht.

Ad 2. Die *nicht-charakterologische* Verwertung: Innerhalb der Bevorzugung des Optischen ist Art und Intensität der Verwertung verschieden, von der individuellen Persönlichkeit des Hörers abhängig. Meistens ist es offenbar so, daß das angegeben wird, was an der eingefallenen Person ein besonderes Charakteristikum darstellt und was durch die Stimme des Sprechers irgendwie ausgelesen wird: d. h. der spezielle Einfall wird seiner typischen Bedeutung nach ausgewertet. Nun verstehen wir auch die typische Bedeutung der optischen Details in den Zusätzen zum Ravag-Experiment, die ja größtenteils auf Einfälle zurückgehen.

Die Frage nach der *Berechtigung* der Verwertung ist offenbar am ehesten vom Erfolg her zu beantworten. Leider kann ich hier keine definitive Antwort geben, weil für meine Protokolle, wo ich den Einfluß der Einfälle genau ersehen kann, keine Erfolgskontrolle durchgeführt wurde; in den Ravag-Experimenten haben wir wohl den Erfolg der Deutung, aber es ist aus den Zusendungen nicht reinlich zu erkennen, was auf Grund direkter Deutung und was durch die Einfälle zustande kam. Allerdings konnte ein überwiegender Anteil der Einfälle an den Zusätzen über optische Details festgestellt werden. Es erscheint also vom Erfolg her die unbefangene Auswertung der Einfälle durch den Hörer gerechtfertigt; es scheint so, als ob mit einer Stimmähnlichkeit, einer Ähnlichkeit des Stimmtypus auch eine Ähnlichkeit des Körperbau-

typus zu verbinden wäre, wofür die Treffer der Deutung sprechen, die ja immer solche typische Details darstellen. Jedenfalls ergibt sich von hier aus die Notwendigkeit, zu untersuchen, ob die KRETSCHMERSche Lehre auch nach der Stimme auszubauen ist. Natürlich liegt aber in der speziellen Natur des Einfalles die Gefahr einer 1. zu speziellen, 2. von der Stimme her nicht berechtigten Auswertung. Sie ist offenbar um so größer, je weniger der Einfall als solcher bewußt ist (da das Bewußtsein des Einfalles seine Verwertung einschränkt oder zumindest zur Vorsicht mahnt), je mehr also die Deutung nach dem „Gefühl“, das mit echter Resonanz nicht zu verwechseln ist, vor sich geht. Und wir haben offenbar in den Einfällen auch die Hauptfehlerquelle in den Deutungen zum Ravag-Experiment gefunden.

Wir können also sagen: In den Einfällen kommt die *Individualität* des Hörers zur Geltung: 1. weil in ihnen individuelle Stimmerlebnisse aktualisiert werden, 2. weil die Art der Auswertung individuell verschieden ist. Was die Stellung der Einfälle innerhalb des Stimmerlebnisses anlangt, so tragen sie in erster Linie zu seiner *Lebendigkeit* bei. Der Vergleich des Sprechers mit einer eingefallenen Person führt sachlich vielleicht nicht allzu weit. Er ist, wie Protokoll 3 zeigte, für den Deutungsfortschritt oft geradezu hinderlich, weil er das unbefangene Deuten der Stimme durch das Ähnlichkeitsbewußtsein, das zur Lösung drängt, stört. Aber die Einreihung in die Welt der Bekannten, die durch den Einfall vorgenommen wird, macht den eigentlichen individuellen Erlebnischarakter der Stimmdeutung aus.

Die quantitative und auch die qualitative Betrachtungsweise haben sich also für unsere Fragestellung: „Stimme als Ausdruck der Persönlichkeit“ als fruchtbar erwiesen. Wir haben unser Ziel weniger in letzten Ergebnissen gesehen als in einer ersten Sondierung des Gebietes, in einer Aufdeckung der Punkte, wo es sich als fruchtbar erweisen würde, weiter einzusetzen und detailliertere, auf genauere Ergebnisse zugespitzte Fragen zu stellen.

(Eingegangen am 2. März 1933)