

Die dritte Abbildung stellt einen echten zweisilbigen Akzent dar : beide Silben haben ungefähr die gleiche Intensität, und das Intervall zwischen ihnen ist auch erheblich kleiner ; dabei ist die Melodie in der ersten Silbe *schwachsteigend* und in der zweiten schwachfallend.

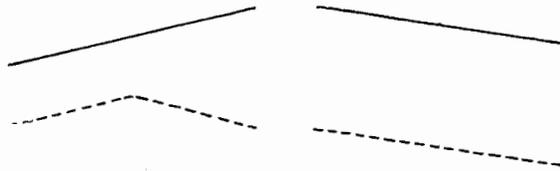


Abb. 4

Auf der vierten Abbildung ist der Akzent schon vollständig herübergezogen : die Intensität ist in der ersten Silbe grösser ; die Melodie ist *steigend* in der ersten Silbe, bzw. fallend in der zweiten Silbe. Doch bleibt die zweite Silbe höher als die erste.

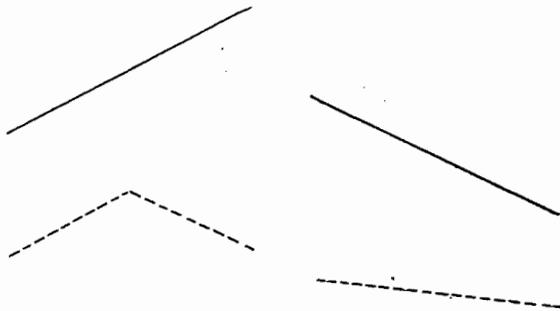


Abb. 5

Auf der fünften Abbildung endlich ist die Intensität gänzlich auf die vorhergehende Silbe übergegangen, und auch die Melodie ist erheblich höher als in der zweiten. Der Verlauf der Melodie ist ebenfalls vollkommen entwickelt : in der akzentuierten Silbe ist die Melodie *stark steigend* und in der nichtakzentuierten stark fallend. Das sind die allertypischsten sekundären Akzente.

Dies wären meiner Meinung nach die Hauptetappen durch die der skr. Akzent von seiner ältesten bis zur modernsten Gestalt hindurchgegangen ist. Das Bild, welches unsere gegenwärtigen Mundarten bieten, ist natürlich viel bunter : zwischen diesen Haupttypen bestehen zahlreiche Uebergangstypen, so dass die Entwicklung sozusagen Schritt für Schritt verfolgt

werden kann. Auf alle diese Einzelheiten, sowie auf die Frage nach der relativen Chronologie der einzelnen Erscheinungen kann ich hier selbstverständlich nicht näher eingehen.

Was endlich den *Grund* betrifft, weshalb gleichzeitig mit der Akzentverschiebung die Tonbewegung steigend wird, begnügte man sich früher einfach mit der Annahme, dass die Melodie in den Silben vor der akzentuierten steigend gewesen sei, so dass der Akzent, sobald er auf die vorangehende Silbe verlegt wurde, auch selbst steigend wurde. Alle experimentellen Untersuchungen stellten aber entschieden fest, dass die Annahme einer steigenden Melodie in den Silben, die der akzentuierten vorangehen, willkürlich ist : die Melodie ist in ihnen die gleiche wie in den Silben nach dem Akzent, d. h. fallend.

Die Ursache der Entstehung steigender Intonationen muss also anderswo gesucht werden. Meiner Meinung nach liegt sie in der Stärkeverschiebung. Es ist nämlich auffallend, dass, wie ich schon oben betonte, in den Mundarten mit vollkommen unverschobenem Akzent die steigende Melodie in den vortönigen Silben nie vorzukommen scheint. Das Steigen der Melodie scheint vielmehr an die Intensitätsverschiebung gebunden zu sein und mit ihr parallel zu gehen. Dass die Intensität in ähnlicher Weise auf die Tonbewegung einwirken kann, ist eine allgemein bekannte Tatsache, die nicht erst bewiesen werden muss. Dass der Satzakkent (also wieder in erster Linie die Verstärkung bzw. die Abschwächung) der Tonbewegung in unseren Akzenten eine vollkommen entgegengesetzte Richtung geben kann, zeigte ich erst neulich in einem Artikel (1). Unter dem Einflusse der im Laufe der Zeit immer mehr zunehmenden Stärke geht also der ursprünglich fallende Ton allmählich in einen ebenen und nach und nach in einen mehr und mehr steigenden über.

So entstehen im Serbokroatischen die sekundären Intonationen.

63. Dr. ANTTI SOVIJÄRVI (Helsinki) : *Die wechselnden und festen Formanten der Vokale, erklärt durch Spektrogramme und Röntgenogramme der finnischen Vokale.*

Die physikalische Klanganalyse ist in ausserordentlich beachtlicher Weise gefördert worden durch die von M. GRÜTZMACHER (2) und E. GERLACH (3) in Deutschland, sowie von C. R. MOORE und A. S. CURTIS (4) in Amerika entwickelte Suchtonanalyse,

(1) *Mélanges Belić*, Beograd, 1937, S. 218-224.

(2) *Z. f. Techn. Physik*, 8, 506, 1927.

(3) *Z. f. Techn. Physik*, 8, 515, 1927.

(4) *Bell Syst. Techn. Journ.*, 6, 216, 1927.

die jedoch phonetischen Zwecken verhältnismässig wenig angepasst worden ist. Dieses rührt wohl in erster Linie daher, dass das betreffende Verfahren eine verhältnismässig vielseitige und teure Apparatur erfordert. Meines Wissens haben bis jetzt folgende Forscher diese Methode zu phonetischen Zwecken verwendet: G. PANCONCELLI-CALZIA, E. THIENHAUS, A. BARCZINSKI, F. TRENDELENBURG und H. FLETCHER.

Als Vorteile dieses Verfahrens seien folgende erwähnt. Die Breite des Suchtonsiebtes kann man verändern und dadurch mehrere beliebige Frequenzbreiten (von 20-200 Hertz) zu Klang-, bzw. Geräuschanalysen benutzen. Wenn z. B. eine genaue Trennschärfe nötig ist, so verwendet man am besten die Siebbreiten von 20 oder 50 Hertz. (Bei der Analyse von gesungenen Vokalen zeigte es sich vorteilhafter, die Siebbreite von 50 Hertz zu verwenden, wohingegen bei der Analyse der geflüsterten Vokale schon eher die Siebbreite von 100 Hertz bessere Ergebnisse lieferte.) Ferner ist auch noch die Empfindlichkeit des Messinstruments regulierbar. Bei der Apparatur, mit welcher meine Versuche gemacht wurden, war es möglich beim Aufzeichnen drei Empfindlichkeitsstufen, nämlich die Stufen  $1/1$ ,  $1/5$  und  $1/25$  der Gesamtintensität zu benutzen. Diesen Vorteil kann man zur Analyse von Lauten und Stimmen so ausnutzen, dass die *unteren* Teiltöne oder Schallfeldkomponenten, die öfters die stärksten sind, mit der kleinsten ( $1/25$ ) oder kleineren ( $1/5$ ) Stufe und die *höheren* mit voller ( $1/1$ ) Empfindlichkeit aufgezeichnet werden.

Der einzige, aber auch bedeutende Nachteil dieser Methode ist derjenige, dass die Analyse, d. h. das Durchlaufen des Suchtons über den ganzen zu analysierenden Tonbereich, ziemlich lang, nämlich 3-5 Minuten dauern muss, die kurzen Atempausen mitgerechnet. Es ist deshalb für die Versuchsperson sehr schwer während der ganzen Analysenzeit möglichst genau dieselben Phonationsstellungen und genau dieselbe Tonhöhe beizubehalten.

Die Suchtonanalysen der finnischen Vokale, die im folgenden behandelt werden, sind im Akustischen Laboratorium der Physikalisch-Technischen Reichsanstalt (in Berlin-Charlottenburg) unter der Leitung von Regierungsrat Dr. M. GRÜZMACHER und Dr. W. LOTTERMOSER gemacht worden. Als Versuchsperson fungierte ich selbst. Bei den Analysenversuchen wurde ein sogenanntes Messkondensatormikrophon verwendet, dessen Frequenzkurve recht gleichmässig ist. Damit die Vp während der Lautgebung kontrollieren konnte, ob die Intensität ihrer Stimmgebung die ganze Zeit über möglichst gleichstark blieb, hatte sie einen Röhrenvoltmeter zu ihrer Verfügung, an dem sie jeden Augenblick die Stellung des Zeigers verfolgen und

bei Bedarf die Lautstärke etwas abschwächen oder vergrössern konnte.

Bei der Analyse der geflüsterten Laute dauerten die Lautgebungsabschnitte ca. 15-20 Sek., wohingegen sie bei den gesungenen länger nämlich ca. 30-35 Sek. waren. Die Atempausen dauerten nur so lange wie es zum Beginn eines neuen Lautgebungsabschnittes nötig war, d. h. einige Sekunden. Für die Pausen, gewöhnlich schon etwas vor Beendigung der Lautgebung, unterbrach der Assistent am Apparat die Registrierfähigkeit, um sie etwas nach Beginn des neuen Lautgebungsabschnittes wieder aufzunehmen. Die Anfangs- und Endstadien der einzelnen Lautgebungsabschnitte können nämlich hinsichtlich der Allgemeinintensität leicht zu stark oder zu schwach ausfallen.

Nun gehe ich über zum Schildern der verschiedenen Vokalfomanten.

Weil die Resonanzmaxima bei den Spektren der *geflüsterten* Vokale klarer und deutlicher zu erkennen sind als bei den der gesungenen Vokale (deren Teiltoneinteilung in dieser Hinsicht etwas störend ist), versuche ich mit Hilfe der erstgenannten die wechselnden und festen Formanten des Vokalschalles zu erklären. Zu den Spektren der acht geflüsterten finnischen Vokale habe ich die entsprechenden finnischen (von mir gezeichneten) Röntgenogramme beigefügt, um die verschiedenen Grössen und Formen der Ansatzrohr Räume bei den Vokalen zu veranschaulichen. Die Röntgenbilder stammen aber nicht von derselben Versuchsperson wie die Spektren, aber für den Zweck dieser Arbeit ist dieser Umstand nicht störend.

In der Abbildung 1 sehen wir das Spektrum und das Röntgenogramm des gefl. *ä* (1). Beim Spektrum sind die wechselnden Formanten  $R_M$ ,  $R_a$ ,  $R_b$ ,  $R_c$ ,  $M$  und  $L$  und die übrigen, festen Formanten sind  $B_1$ ,  $B_2$ ,  $T$ ,  $N_2$ ,  $N_1$ ,  $K_1$ ,  $V$ ,  $K_2$  und  $P$ . Die Buchstaben  $B_1$ ,  $B_2$  und  $T$  kennzeichnen die drei Brustresonanzmaxima.  $R_M$  bedeutet das Resonanzmaximum der ganzen Rachen-Mundhöhle,  $R$  die Rachenhöhlen-,  $R_a$ ,  $R_b$  und  $R_c$  die Rachteil-,  $M$  die Mundhöhlen-,  $N_2$  die Nasenrachen-,  $N_1$  die Nasenhöhlen- und  $L$  die Lippenraumresonanz. Dann folgen noch vier höhere feste Formanten, die im Kehlkopf entstehen.  $K_1$  bedeutet die Resonanz des oberen Kehlkopfraumes,  $V$  die der Vallekulenräume (*Valleculae glossoepiglotticae*),  $K_2$  die des mittleren Kehlkopfraumes und  $P$  die der kleinen *Sinus piriformis*-Räume.

(1) Die anderen Bilder der Vokalspektren und -Röntgenogramme, die ich im Zusammenhang des Vortrags gezeigt habe, veröffentliche ich nicht hier, weil sie vor kurzem (im Dezember 1938) in meiner Dissertationsarbeit („Die gehaltenen, geflüsterten und gesungenen Vokale und Nasale der finnischen Sprache. Physiologisch-physikalische Lautanalysen“) erschienen sind.

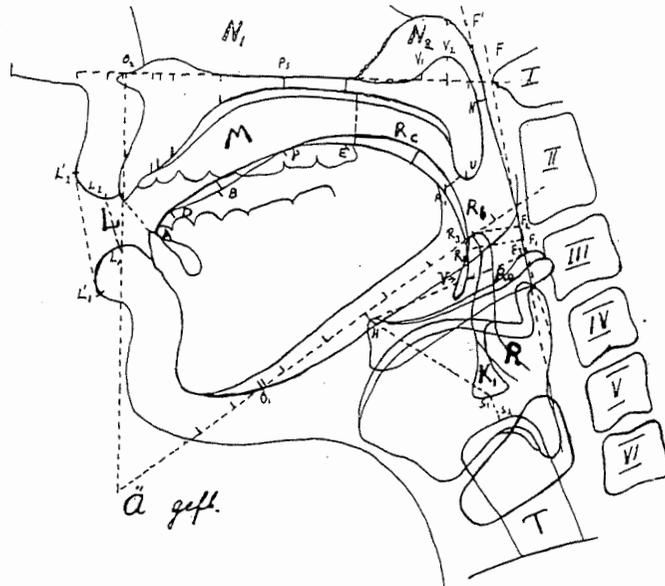
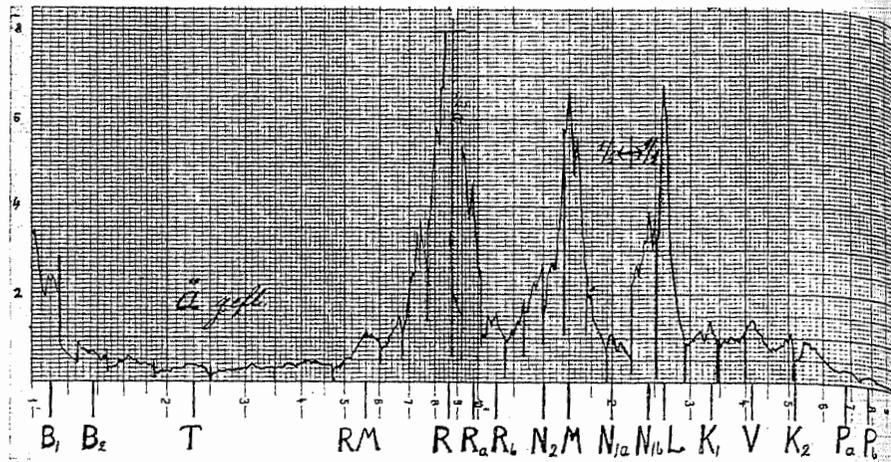


Abb. 1. Das Suchtonspektrum und Röntgenogramm des geflüsterten (finnischen) *ä*.

a. Die wechselnden Formanten.

Zuerst möchte ich die wechselnden Formantengebiete, die eben für die *Lautfarbe* massgebend sind, behandeln.

Wenn man die Spektren der geflüsterten Vokale in ihren

allgemeinen Zügen betrachtet, ist es augenfällig, dass es eine verhältnismässig lose Kopplung zwischen den verschiedenen Ansatzrohräumen gibt. So z. B. die Resonanzgebiete der beiden wichtigsten Räume, des Mund- und Rachenraums (M und R), kommen im Flüstergeräusch sehr scharf und deutlich hervor. (Bei den Röntgenogrammen habe ich die Grenze des M- und R-Raumes durch die Abstandsstriche zwischen den Spitzen der vorderen Ober- und Unterzähnen und zwischen dem Gaumen- und Zungenrücken bezeichnet.) Aber bei allen Spektren befindet sich (unterhalb des R) noch ein Maximum (RM), das ich von dem Gesamtraum der Mund- und Rachenhöhle herrührend erklären möchte (1). Seine Intensität ist — wie schon zu erwarten ist — nicht besonders stark und scheint ausserdem bei allen diesen Spektren ziemlich gleich stark zu sein. Die Grösse des RM-Raumes ist natürlich in einem nahen Abhängigkeitsverhältnis zu der R- und M-Räume und ihre Eigenfrequenz ist selbstverständlich immer niedriger als die des R. Ausserdem ist noch zu bemerken, dass seine Eigentonhöhen bei verschiedenen Vokalen sich zu einander auf dieselbe Weise verhalten wie die der entsprechenden R-Räume. Von den zu behandelnden Spektren habe ich die folgenden RM-Frequenzen bekommen: *a* 680, *ä* 560, *o* 360, *u* 325, *e* 365, *ö* 460, *i* 270 und *ü* 265 Hz. Zu den wechselnden Formanten gehört noch der Lippenformant (L), der in dem prädentalen, seine Form nach den Vokalen verändernden Lippenraum entsteht.

Die Höhen der M- und R-Formanten, die in diesen Spektren erscheinen, habe ich auf folgende Weise nachkontrollieren können. Die Eigentöne der M- und R-Höhle wurden durch Fingeranprall gegen die Backenwand, bzw. die Schildknorpelplatte bei den verschiedenen (finnischen) Vokalstellungen erzeugt und oszillographisch aufgenommen. Die von den Kurven gemessenen Eigenfrequenzen der R und M stimmen sehr gut mit den entsprechenden Maxima der Spektren überein. Die Hörbeobachtungen (2) der so erzeugten M-Töne gaben auch dieselben Resultate. Die Eigentonhöhen der L-Formanten habe ich durch Hören mit genügender Sicherheit feststellen können. Zur Erzeugung der L-Töne habe ich das sogenannte Lippensprengungsverfahren (3) benutzt. Indem man die Zungenspitze an die

(1) Es sei erwähnt, dass PIPPING auf die Möglichkeit hingedeutet hat, dass es einen besonderen Formanten der ganzen Mund- und Rachenhöhle neben der Mund- und Rachenformanten gäbe. Nach seiner Vermutung käme dieser Formant sicherer in denjenigen Vokalen in Frage, in welchen der Engpass zwischen dem Gaumen und Zungenrücken nicht sehr klein ist. S. „Ueber die Theorie der Vokale“, *Acta Societ. Scient. Fenn.* XX, 11, 1894.

(2) Sie wurden mit Hilfe eines normal gestimmten Klaviers ausgeführt.

(3) Durch dieses Verfahren hat W. TRENDELENBURG die Eigentöne

Die wechs. F.

RM	560	Hz
R	860	"
R <sub>a</sub>	980	"
R <sub>b</sub>	1100	"
R <sub>c</sub>	1400	"
M	1600	"
L	2620	"

Die festen F.

B <sub>1</sub>	110	Hz
B <sub>2</sub>	135	"
T	230	"
N <sub>2</sub>	1400	"
N <sub>1a</sub>	2350	"
N <sub>1b</sub>	2420	"
K <sub>1</sub>	3350	"
V	4150	"
K <sub>2</sub>	5200	"
P <sub>a</sub>	6800	"
P <sub>b</sub>	7800	"

oberen Schneidezähne ansetzt, so dass die Spalte zwischen den Ober- und Unterzähnen geschlossen wird, kann man durch die plötzlichen Lippenartikulationen *pa*, *pä*, *po*, *pu*, *pe*, *pö*, *pi* und *pü* Sprengungstöne hervorrufen, die den Eigentönen der verschiedenen Lippenräume entsprechen. Die Schwankungsgebiete ihrer Höhen, die nicht grösser als ein Ganztonbereich zu sein scheinen, stimmen wieder sehr gut mit den vermuteten L-Maxima der Spektren überein. Diese Maxima sind folgende: *a* 2850, *ä* 2620, *o* 1850, *u* 1650, *e* 2550, *ö* 2750, *i* 2970 und *ü* 2250 Hz.

Nun werde ich einzeln die wichtigsten wechselnden Resonanzgebiete R und M sowie auch ihre in den bestimmten Zuständen erscheinenden Teilresonanzgebiete  $R_a$ ,  $R_b$ ,  $R_c$ ,  $M_a$  und  $M_b$  behandeln.

Beim Spektrum des geflüsterten *a* sind die Frequenzahlen der R und M 880 und 1100 Hz. Weil die R- und M-Räume beim *a* ziemlich gleich gross sind, ist es schon zu erwarten, dass ihre Eigentöne nicht weit voneinander liegen.

Bei dem *ä*-Spektrum (die Abb. 1) liegen R und M schon bedeutend weiter voneinander (860 und 1600 Hz), weil ja der M-Raum kleiner und gleichzeitig der R-Raum etwas grösser geworden ist. In bezug auf R muss man aber bemerken, dass seine Höhe beim *ä* doch ungefähr dieselbe ist wie beim *a* (860 gegen 880 Hz), was davon bedingt sein kann, dass durch die tiefere Stellung des Gaumensegels und die höhere Stellung des Kehlkopfes beim *ä* der R-Raum sich etwas verkleinert. Dieses hat also eine entgegengesetzte Wirkung wie die vordere Stellung der Zunge. Beim *ä*-Spektrum können wir noch die drei Rachenteilresonanzmaxima ( $R_a$ ,  $R_b$  und  $R_c$ ) beobachten. Die entsprechenden Frequenzahlen sind 980, 1100 und 1400 Hz. Nach meiner Vermutung entsteht der unterste  $R_a$ -Formant hauptsächlich in dem Raum, der sich zwischen den Gegenden des (kürzesten) Gaumenabstandes und des Zungenbeins befindet, und gleichzeitig vielleicht auch in dem ungefähr gleich grossen Raum, der zwischen der Uvulagegend und dem Rachengrund, d. h. den aryepiglottischen Falten liegt. Der mittlere  $R_b$ -Formant würde vorwiegend in dem Raum entstehen, der durch die Gegenden des Gaumen- und Rachenabstandes begrenzt wird. Ungefähr gleich grosse Räume sind noch die Rachenteilräume, die sich 1) zwischen den Gegenden der Uvula und des Zungenbeins und 2) des Rachenabstandes und Zungengrundes

des Lippenraumes bei *i* aufgezeichnet (s. „Frequenz und Dekrement der Eigenschwingungen der Mundhöhle bei Vokalstellungen“, Sonderausgabe aus d. Sitzungsber. d. Preuss. Ak. d. Wiss., Phys.-Math. Kl., 1936, XXII, S. 14).

befinden. Sie können also auch beim  $R_b$ -Maximum beteiligt sein. Der höchste Rachenteilformant  $R_c$  könnte am besten in dem Rachenteilraum entstehen, der sich zwischen den Gegenden des Gaumenabstandes und Uvula bildet. Ungefähr dieselbe Grösse haben noch die Rachenteilräume, die 1) zwischen den Gegenden der Uvula und des Rachenabstandes, 2) zwischen den Gegenden des Rachenabstandes und des Zungenbeins, und 3) zwischen den Gegenden des Zungenbeins und des Rachengrundes liegen. Der  $R_c$ -Formant könnte sich deswegen auch gleichzeitig in diesen Teilräumen bilden. Die Intensität der Rachenteilformanten ist immer kleiner als die der R- und M-Formanten, was dadurch sehr gut zu verstehen ist, dass die Grenzen der verschiedenen Rachenteilräume verhältnismässig unscharf sind (1).

Im *o*-Spektrum hat R 530 und M 800 Hz. Im Vergleich zum *a* und *ä* ist besonders zu bemerken, dass die beiden Oeffnungen der R- und M-Räume kleiner und dadurch die entsprechenden Eigentöne tiefer sind. Beim *o* (wie auch beim *u*) kommen die Mundteilformanten  $M_a$  und  $M_b$  zum Vorschein. Meiner Ansicht nach entsteht der  $M_a$ -Formant in dem hinteren Mundteilraum, der vorne zwischen den Gegenden der Zungenspitze und der Alveolen endigt, und der  $M_b$ -Formant in dem vorderen Mundteilraum, der sich oberhalb des Mundbodens und vor der Zungenspitzen- und Alveolengegend befindet. (Die Zungenspitze liegt ja beim *o* und *u* rund 10, bzw. 15 mm von den vorderen Unterzähnen entfernt.) Diese Mundteilformanten haben beim *o*-Spektrum die Frequenzahlen 915 und 1000 Hz. Von den Rachenteilformanten kann man beim *o* nur den tiefsten  $R_a$  erkennen (640 Hz), weil die  $R_b$ - und  $R_c$ -Formanten sehr wahrscheinlich mit den Formantengebieten  $M_a$ ,  $M_b$  und vielleicht auch  $N_2$  zusammenfallen und deshalb nicht mehr zu unterscheiden sind (2).

Beim *u* sind R und M im Vergleich zum *o* tiefer hauptsächlich deswegen, weil die Oeffnungen der beiden betreff. Räume kleiner sind (R 400 und M 630 Hz). Ebenso liegen die beiden Mundteilformanten tiefer aus demselben Grunde ( $M_a$  790 und  $M_b$  1020 Hz). Der  $R_a$ -Formant hat 495 Hz und ist also auch tiefer als beim *o*.

(1) Dass diese Rachenteilformanten wirklich in den Vokalen, besonders in den vorderen Vokalen vorkommen, beweisen noch die Fourier-Analysen, die ich von den Oszillogrammen der auf tieferen Tonhöhen (*c°* und *Fis*) gesungenen, bzw. gehaltenen Vokale (der zwei verschiedenen Versuchspersonen) ausgeführt habe. In ihnen erscheinen sie noch klarer als in den entsprechenden Suchtonspektren.

(2) Dasselbe ist in bezug auf die anderen hinteren Vokale *a* und *u* zu bemerken, bei welchen die R- und M-Formanten auch ziemlich nahe beieinander liegen.

Im *e*-Spektrum hat R 450 und M 2100 Hz. Im Vergleich zum *ä* ist der Rachenraum grösser und der Mundraum kleiner. Deswegen hat R eine tiefere und M eine höhere Frequenz. Weil die R- und M-Formanten ziemlich weit voneinander liegen, haben die drei Rachenteilformanten sozusagen mehr freies Frequenzgebiet, um zum Vorschein kommen zu können. R<sub>a</sub> hat 620, R<sub>b</sub> 730 und R<sub>c</sub> 900 Hz.

Wenn die *ö*- und *e*-Spektren miteinander verglichen werden, hat das erstere ein höheres R-Maximum (510 Hz) und ein tieferes M-Maximum (1630 Hz). Das ist dadurch zu erklären, dass beim *ö* der R-Raum kleiner, der M-Raum grösser und die Lippenöffnung kleiner ist als beim *e*. Die Rachenteilformanten liegen etwas höher als beim *e* (660, 840 und 1000 Hz), was auch wegen des höheren R beim *ö* schon zu erwarten ist.

Beim *i* sind R und M noch weiter auseinander gewichen als beim *e*. R hat nämlich 330 und M 2520 Hz. Die Ursachen sind natürlich dieselben, die schon beim Vergleichen der *ä*- und *e*-Vokale erwähnt worden sind. Auch die Rachenteilformanten liegen beim *i* entsprechend tiefer als beim *e* (440, 600 und 900 Hz).

Beim *ü* ist zu bemerken, dass R und M tiefer liegen als beim *i*, denn R hat 310 und M 1750 Hz. Die sehr kleine Lippenöffnung beim *ü* scheint also ebenso gut das tiefere M-Maximum wie auch das tiefere R-Maximum zu verursachen, obgleich der R-Raum beim *ü* etwas kleiner ist als beim *i*. Dasselbe Verhältnis ist hinsichtlich der Rachenteilformanten der beiden Vokale zu beobachten. Beim *ü* hat nämlich R<sub>a</sub> 370, R<sub>b</sub> 580 und R<sub>c</sub> 770 Hz.

Bei den Spektren der *gesungenen* Vokale weisen die wechselnden (sowie auch die festen) Formanten im grossen und ganzen dieselben Frequenzgebiete wie bei den der *geflüsteren* Vokale auf. An Hand eines grösseren Analysenmaterials, das von zahlreichen Filterspektren und Fourier-Analysen der Oszillogramme der *gesungenen* und *gehaltenen* Vokale gewonnen worden ist, kann man bestimmte Schlussfolgerungen hinsichtlich der Unterschiede zwischen den Formanten der *geflüsteren* und *gesungenen*, bzw. *gehaltenen* Vokale ziehen. Sie hier zu behandeln würde zu weit führen. Nur als ein Beispiel von den Suchtonspektren der *gesungenen* Vokale zeige ich die der auf *c<sup>o</sup>* und *c<sup>i</sup>* *gesungenen* *ä*-Vokale. Bei ihnen sehen wir wieder dieselben wechselnden Resonanzgebiete wie bei dem Spektrum des *geflüsteren* *ä*. In bezug auf die festen Formanten sei zu bemerken, dass die nebeneinander stehenden Resonanzgebiete der drei (verschieden grossen) Nasengänge (N<sub>1a</sub>, N<sub>1b</sub>, N<sub>1c</sub>) sehr deutlich erscheinen. Diese Nasenhöhlenformanten befinden

sich bei allen Spektren der Vokale (sowie auch der Nasale) immer innerhalb des Frequenzgebietes 2000-3300 Hz.

#### b. Die festen Formanten.

Von den verschiedenen festen Formantengebieten, von denen zum grossen Teil die *individuelle Stimmfarbe* abhängig ist, möchte ich dann eine übersichtliche Erklärung geben.

Die niedrigsten drei festen Maxima (B<sub>1</sub>, B<sub>2</sub> und T), deren Schwankungsgebiete und Durchschnittswerte 105-130, Dw. 112 Hz (laut 9 Spektren), 125-145, Dw. 136 Hz (9) und 187-265, Dw. 221 Hz (14) betragen, entstehen ohne Zweifel aus den verschieden hohen Resonanzen des Brustkorbes, des Bronchialbaumes und der Luftröhre. Ausser diesen deutlicheren Hauptmaxima der Brust sind noch innerhalb des unteren Frequenzgebietes (etwa von 100 bis 300, bzw. 400 Hz) kleinere unregelmässige Nebenmaxima zu beobachten, die wahrscheinlich aus den verschiedenen Teilräumen des Bronchialbaums entstehen. Aus den wichtigsten Arbeiten, die die Brustresonanz behandeln, geht hervor, dass das Resonanzgebiet der Brust ung. zwischen 100 und 300 Hz liegt. Z. B. BASS (1) fand bei Zuleiten sinusförmiger Fremdtöne vom Mund aus ein Mitschwingen der Thoraxwand in normalen Fällen zw. 90 und 400 Hz, mit Maximum bei 190 Hz. In einer anderen Arbeit (2) behauptet er, dass bei normalem Lungengewebe die Resonanz des Brustraumes zw. etwa 80 und 350 Hz liegt und entsprechend dem zusammengesetzten Bau des Bronchialbaums mehrere Maxima aufweist. Ferner hat W. TRENDELENBURG (3) die Wandschwingungskurven von den Lauten *a*, *e*, *i*, *o*, *u*, *m*, *n* und *s* von dem Brustkorb und der Trachea abgenommen. Die Grundtonhöhe war bei diesen Versuchen immer 110 Hz. „Die vom Brustkorb vorn, seitlich oder hinten aufgenommenen Kurven zeigen im wesentlichen nur die Grundschwingung (erste Teilschwingung) und die zweite Teilschwingung und sind sich recht ähnlich.“ Aus der Form der Brustkorbkurven kann man weiter noch den genaueren Schluss ziehen, dass der Grundton (110 Hz) stärker ist als der zweite Teilschwingung (220 Hz), d. h. dass das Maximum der Brustkorbresonanz eben in der Nähe von 110 Hz liegt. Die Fourier-Analyse einer *a*-Kurve

(1) BASS, E., „Das akustische Frequenzspektrum des normalen und infiltrierten Lungengewebes“, *Verh. Ges. inn. Med.*, 40, 166, 1928.

(2) BASS, E., „Ueber Schalleitung und Resonanz im normalen Lungengewebe“, *Z. f. d. ges. exp. Med.*, 77, 303-322, 1931.

(3) TRENDELENBURG, W., „Physiol. Untersuchungen über die Stimmklangbildung“, *Sonderausg. aus d. Sitzungsber. d. Preuss. Ak. d. Wiss., Phys.-Math. Kl.*, 1935, XXXI, 32, Abb. 3 und 4, Taf. II.

(Grundton auch 110 Hz), die von der Luftröhre (dicht oberhalb des Brustbeins) abgenommen wurden, zeigt sehr deutlich, dass der 2. Teilton überwiegend stärker ist als die anderen Teiltöne und deswegen muss das Resonanzmaximum der Trachea bei 220 Hz liegen. Das zwischen  $B_1$  und T liegende dritte Maximum  $B_2$  (Dw. 136 Hz) vermute ich von der Resonanz des Bronchialbaums (d. h. des Gesamtraums der Bronchien und Luftröhre) herzurühren, denn dieser Raum ist ja seinem Volumen nach grösser als der blosser Luftröhrenraum.

Eher als unveränderlich denn als veränderlich ist das schwache Maximum zu betrachten, das in den betreffenden Spektren zwischen 1350 und 1700 Hz liegt, Dw. 1520 Hz (8), und von dem ich angenommen habe, dass es aus dem Nasenrachenraum stammt. Eine solche Schlussfolgerung wird durch die Hörbeobachtungen, die die geflüsterten Nasale betreffen, gestützt, nach denen das stärkste Gebiet ca. zwischen  $f^3 \cdot a^3$  (1350-1700 Hz) liegt (1), sowie andererseits durch das starke Nasenrachen-Resonanzgebiet 1320-1650 Hz, Dw. 1475 Hz (6), das in den Spektren der geflüsterten Nasale auftritt.

Für das unveränderliche Maximum ( $N_{1a}$ ), dessen Frequenz 2050-2350 Hz beträgt, Dw. 2230 Hz (8), nehme ich an, dass es der untersten Nasenhöhle entstammt. Auf Grund der Hörbeobachtungen habe ich die Schlussfolgerung gezogen, dass der Resonanzton der  $N_{1a}$ -Höhle zwischen ca.  $c^4 \cdot d^4$  (2050-2300 Hz) liegt (2). Ferner zeigen die Suchtonspektren der geflüsterten Nasale, dass oberhalb eines starken  $N_2$ -Maximums meistens ein etwas schwächeres Maximum zwischen 2000-2150 Hz, Dw. 2085 Hz (6) liegt. Das kann man wohl nicht besser anders erklären als so, dass es gerade von dem  $N_{1a}$ -Raum entstammt. Den Eigenschwingungen des mittleren Nasengangs scheint wiederum das in den gleichen Nasalspektren auftretende niedrige schwächere Maximum ( $N_{1b}$ ) zu entsprechen, dessen Schwingungszahl bei 2420-2700 Hz, Dw. 2540 Hz (6) liegt. Das der oberen Nasenhöhle entsprechende Maximum ( $N_{1c}$ ) schwankt sich zwischen 2800-3300 Hz, Dw. 3040 Hz (7). In den Klanganalysen der gesungenen Nasale zeigen sich die Resonanzmaxima der drei Nasengänge noch deutlicher.

Das folgende nächsthöhere, einigermaßen bestimmte For-

(1) Beim Flüstern der *a*-, *u*- und *i*-farbigen, wichtigsten Nasallautvarianten (*a*)*m*, (*u*)*m*, (*i*)*m*; (*a*)*n*, (*u*)*n*, (*i*)*n*; (*a*)*η*, (*u*)*η*, (*i*)*η* habe ich von dem Lautgeräusch als den stärksten Ton immer innerhalb des genannten Frequenzgebietes gehört.

(2) Das Resonanzgebiet der Nasenhöhle habe ich durch Hören so feststellen können, dass ich die Tonhöhe des stärksten Komponenten des Luftstromgeräusches, der beim Blasen durch die Nase entsteht, bestimmt habe.

mantengebiet ( $K_1$ ) in den betr. Vokalspektren ist das, welches ich als dem oberen Kehlkopf entstammend erklären möchte. Sein Eigenschwingungsgebiet liegt bei 2950-3680, Dw. 3390 Hz (13). Dieses verhältnismässig grosse Formantengebiet ist so zu erklären, dass die Bewegungen des Kehldeckels eine Veränderung des Eigentones des  $K_1$ -Raumes zur Folge haben. Wenn nämlich der Kehldeckel sich nach hinten senkt, hat dieses eine Verkleinerung der oberen Oeffnung des  $K_1$ -Raumes und eine darauf beruhende Senkung des Eigentones zur Folge, und umgekehrt.

Ein guter Beweis für das Eigentongebiet des  $K_1$ -Raums ist folgender: W. TRENDELENBURG und H. WULLSTEIN (1) haben die Luftschwingungen und die Stimmlippenschwingungen des Kalbskehlkopfes ohne jeden Ansatzstutzen und Ansatzraum, aber mit erhaltener Epiglottis und Schildknorpel, nacheinander aufgenommen. In der erhaltenen Luftklangkurve gibt es höhere Teiltönschwingungen, die „vom oberen Kehlkopfraum (subepiglottischer Raum) an die äussere Luft abgegeben werden“. Da in der Periode der betreffenden Luftklangkurven 22 voneinander gleich weit entfernte Gipfel auftreten, bedeutet dies, dass bei dem 22. Teilton ein starkes Resonanzmaximum liegt, das ohne Zweifel durch den  $K_1$ -Raum verursacht ist. Da der Grundton des Klanges 147 Hz ist, beträgt die Schwingungszahl des 22. Teiltönes 3234 Hz, welcher Wert nun ausserordentlich gut mit dem Durchschnittswert (3390 Hz) des betr. Resonanzgebietes übereinstimmt. Die  $K_1$ -Räume sind ja in dem Kehlkopf von Kalb und Menschen ungefähr gleich gross, obwohl die Morgagnischen Taschen beim Kalb fehlen. Ausserdem ist die Kurve des von dem menschlichen Kehlkopf entwickelten Luftklanges in der betr. Untersuchung dargestellt. In seinen Perioden treten ca.  $5 \frac{1}{2}$  voneinander gleich weit entfernte Gipfel auf, so dass bei dem 5. und 6. Teiltone das  $K_1$ -Maximum liegt. Die entsprechenden Schwingungszahlen sind, da der Grundton des Klanges 565 Hz beträgt, 2825 und 3390 Hz. Die Uebereinstimmung ist also wiederum sehr überzeugend.

Den sichersten Beweis für die Höhe des  $K_1$ -Formanten liefern die Resultate der folgenden Versuche. In dem physiologischen Institut der Universität Berlin habe ich unter der Leitung Professor W. TRENDELENBURGS die seltene Gelegenheit gehabt, Versuche, welche Vokalschwingungen galten, anstellen zu können, wobei ich selbst die Versuchsperson war. Auf demselben Filmband wurden mit Hilfe des Oszillographen g l e i c h

(1) „Untersuchungen über die Stimmbandschwingungen“, *Sonderausg. aus d. Sitzungsber. d. Preuss. Akad. d. Wissensch., Phys.-Math. Kl.*, 1935, XXI, 17, Abb. 6, 28, Abb. 8, 1.

zeitig Kehlkopfwand- und Luftschwingungen aufgenommen. Die ersteren wurden mit dem Riegger-Mikrofon und die letzteren mit dem Sell-Mikrofon aufgenommen. Da die Schwingungszahl des Grundtons der gehaltenen Vokale ca. 144 Hz ( $d^0$ ) war, so erstreckten sich die Klanganalysen von 30 Teiltönen bis zur Schwingungszahl 4320 Hz ( $cis^5$ ), welche der 30. Teilton hat. Wenn somit meine Voraussetzung, dass die Eigenfrequenz des  $K_1$ -Raums durchschnittlich ca. 3400 Hz ist, stichhaltig wäre, müsste bei der Klanganalyse der beiden Kurven das deutliche Resonanzmaximum in irgendeinem von den höchsten Teiltönen zum Vorschein kommen, deren Schwingungszahl ungefähr in der Nähe von 3400 Hz liegt. Man hat schon im voraus vermuten können, dass die  $K_1$ -Resonanz in den Kehlkopfklangen deutlicher zum Ausdruck kommen müsse als in den Luftklängen. Denn der Leitungsweg der Eigenschwingungen des  $K_1$ -Raums zu der Membran des Wandmikrophons ist ja äusserst kurz. Die Kehlkopfwand- und Luftklang-Spektren der analysierten *a*- und *i*-Vokale beweisen auch die Richtigkeit dieser Vermutung. In den Kehlkopfklangspektren der beiden Vokale taucht in dem höheren Teiltongebiet plötzlich ein alleinstehendes, ziemlich starkes Maximum auf und zwar in beiden Fällen im 26. Teilton, der 3744 Hz hat. Diese Frequenzzahl stimmt wieder gut mit den schon erwähnten, vom oberen Kehlräum herrührenden Eigenfrequenzen überein. Auch bei den Luftklangspektren bildet sich in demselben 26. Teilton bei beiden Vokalen dasselbe Maximum, das jedoch nicht mehr allein massgebend für das höhere Frequenzgebiet dieser Spektren ist.

Etwa 800 Hz oberhalb des  $K_1$ -Maximums befindet sich ein seiner Intensität nach höhere Maximum, dessen Schwingungszahlgebiet in den fraglichen Spektren 4100-4900, Dw. 4340 Hz (11) ist. Ich vermute, dass dieses Maximumgebiet dem kleinen Raum entspricht, welcher sich zwischen der Zungenwurzel und dem oberen Teil des Kehldeckels bildet, d. h. dem *Sinus glossoepiglotticus*. Weil diesen Raum das *Ligamentum glossoepiglotticum* in zwei gleichgrosse Teile (*Valleculae*) teilt, kann man annehmen, dass somit zwei gleich, einander verstärkende Resonanzgebiete entstehen. Ihrem Volumen nach sind die Valleculen mehr als doppelt kleiner als der  $K_1$ -Raum, weshalb man für wahrscheinlich halten kann, dass das fragliche Maximumgebiet von den Eigenschwingungen der Valleculen stammt.

Dem Resonanzgebiet der Valleculenräume folgt das Maximumgebiet ( $K_2$ ), dessen Entstehen ich aus den Eigenschwingungen des mittleren Kehlkopfraumes erkläre. Dessen Schwingungszahlgebiet ist laut den fraglichen Spektren 5200-5800, Dw. 5520 Hz (13). Wenn es sich um Flüstern handelt, meine

ich mit dem  $K_2$ -Raum den Luftraum, der von der Knorpelritze (*Glottis intercartilaginea*) gebildet wird, und zugleich die Spalte, die von den hinteren Teilen der (unmittelbar oberhalb derselben gelegenen) Taschenlippen gebildet wird. Wenn hingegen die Rede vom Sprechen oder Singen ist, so bildet das *Ventriculum Morgagni* den  $K_2$ -Raum, der beim Flüstern fehlt, indem die Taschenlippen gegen die Stimmlippen drücken (1). W. TRENDELENBURG (2) schätzt das Volumen vom *Ventriculum Morgagni* auf  $\frac{1}{2}$ -1  $cm^3$ . Der beim Flüstern entstehende  $K_2$ -Raum ist vielleicht etwas kleiner. Darauf weist der Umstand hin, dass das  $K_2$ -Maximum beim Flüstern durchschnittlich höher als beim Singen liegt. Im Vergleich zum Volumen der Valleculenräume ist der  $K_2$ -Raum durchaus kleiner, weshalb man annehmen kann, dass das fragliche Maximum gerade durch die Eigenschwingungen des mittleren Kehlräume entsteht.

Das allerhöchste Maximum in den zu behandelnden Spektren ist das, dessen Frequenzgebiet 6800-7900, Dw. 7530 Hz (8) ist. Meiner Ansicht nach ist es wahrscheinlich, dass es durch die Eigenschwingungen der ganz kleinen *Sinus piriformis*-Räume zustande kommt. Diese Räume sind nämlich (in einem normalen Kehlkopf) ihrem Volumen nach gewöhnlich kleiner als die  $K_2$ -Räume, und gerade auch dann, wenn die Phonation vor sich geht. Dann ist nämlich der *Sinus piriformis* etwas enger als bei der Ruhelage (3). Ausserdem kann die eine Raumhälfte durch die Speichelfüllung kleiner werden als die andere, wodurch zwei verschiedene Resonanztöne entstehen. Auf diesen Umstand weisen auch die Spektren der gesungenen Vokale hin. Wenn man ausserdem in Betracht zieht, dass es weder im Kehlkopf noch im Ansatzrohr irgend welche andere kleine Räume gibt, so stützt auch dieser Umstand seinerseits meine Annahme.

Die Eigentönehöhen der kleinsten Räume des Ansatzrohres (der V-,  $K_2$ - und P-Räume) können aber erst durch besondere Kehlkopfklangversuche mit Hilfe von Kehlkopfpräparaten exakt bestimmt werden, wobei man die verschiedenen Räume des Kehlkopfes nach Belieben füllen, bzw. entfernen und die jeweils entsprechenden Luftklänge aufnehmen und analysieren kann. Hoffentlich können schon in der nächsten Zukunft dergleichen

(1) Diesen Unterschied zwischen Flüstern (und Singen (bzw. Halten)) kann man deutlich bemerken, wenn man die Schattengebiete miteinander vergleicht, die auf den Röntgenbildern der geflüsterten und gesungenen (bzw. gehaltenen) Laute in der Nähe der Stimmlippen auftreten.

(2) W. TRENDELENBURG, „Physiologische Untersuchungen über die Stimmklangbildung“, *Sonderausgabe aus d. Sitzungsber. d. Preuss. Ak. d. Wiss., Phys.-Math. Kl.*, 1935, XXXI, 14.

(3) R. WALDAFFEL, „Methodik der Röntgenuntersuchung des Kehlkopfes“, *Fortschritte auf d. Gebiete d. Röntgenstrahlen*, Ergänzungsband 53, 1938, 80.

Versuche angestellt werden, deren Ergebnisse ein bestimmteres Wort in dem fraglichen Problem sagen würden.

64. Miss URSULA FEYER (Berlin) : *Synkope und Apokope in einigen nordwestdeutschen Mundarten des nordniedersächsischen Raumes.*

Bekanntlich sind in manchen Gruppen des Nordniedersächsischen infolge des Schwundes von unbetontem Auslauts-e, auch von e der Endung, soweit es vor Sonoren stand, die langen Vokale zur Ueberlänge gedehnt worden, wenn sie vor ursprünglich sth. Verschluss- oder Engenlauten standen, oder die Endung ihnen unmittelbar folgte. So entsteht aus mnd. dage (Tage) mundartliches  $\text{dāo:x}$ ; aus mnd. bedrege (betrüge) ein  $\text{bedrer:ç}$ ; aus mnd. kōie (Kühe) ein  $\text{ka:i}$ . Die Ueberlängung betrifft also Langvokale, Diphthonge und in manchen Maa. auch offene Kurzvokale, die vor alter g-Geminata standen : mnd. egge (Egge) >  $\text{ε:x}$ . Bei der gleichen Synkope oder Apokope zeigen jedoch kurzer Vokal + Nasal + sth. Konsonanz sowie kurzer Vokal + geminiertem Nasal Kürze des Vokals und Längung des Nasals : aus. mnd. danne (Tanne) wird  $\text{dan̄}$ ; aus mnd. ende (Ende) ein  $\text{en̄}$  u. s. w. Die gleiche Dehnung erfährt der Nasal auch nach langem Vokal :  $\text{bloumm̄}$  (Blume) < mnd. blōme.

Die beiden eben beschriebenen Phänomene hat man mit dem Schwunde des Endungs-e in Verbindung gebracht, etwa folgendermassen : Folgt der Stammsilbe ein sth. Konsonant, so wird sie nach Verlust der Endungs-e gedehnt; es tritt Morenersatz ein. Entweder wird der Vokal gedehnt oder — bei Nasalen — der Konsonant. CONSTANTIN NÖRRENBURG, der als erster die Gesetze der zirkumflektierten Betonung der rheinischen Maa. darstellte, hat im Anz. 13, 384 ff. die Ueberlängen des Nd. zum Vergleich herangezogen. Ausschlaggebend dafür war die Vereinigung von Zweigipfligkeit der Stammsilbe mit der Apokope, „obwohl im Mecklenburgischen der Rumpf seine musikalische und expiratorische Einheitlichkeit behält; andererseits bekommen seine Sonore einen quantitativen Zuwachs, der im Rheinischen nicht eintritt und die Ausdrücke Konsonantenverdoppelung und Ersatzdehnung rechtfertigt“. Tatsächlich ist die „Zirkumflekction“ der Ueberlängen auf nd. Gebiet konstatiert worden, besonders von TH. RABELER in seinen *Mundarten des Kreises Bleckede*.

Indem man den zirkumflektierten Vokal musikalisch in einen Haupt- und einen Nebenton zerlegte, expiratorisch in Haupt- und Nebeniktus, suchte man die Erklärung dafür im Schwunde der folgenden Silbe, indem nun die allein übrig ge-

bliebene Silbe die einstige Akzent- und Melodieabstufung in sich zu vereinen schien.

FRINGS hat dann in seiner Untersuchung über *Die rheinische Akzentuierung* die Verhältnisse für das Rheinische geklärt, indem er die Bedeutung der Druckverhältnisse herausstellte : Die Energie des Vokaleinsatzes bedinge vielmehr ein Erschlaffen des Druckes und des Expirationsstromes, und namentlich damit auch „die zeitliche Reduktion der Konsonanz wie namentlich der ihr folgenden Silben“, also Synkope und Apokope. Daneben läuft im Rheinischen ein anderer Intonationstypus : eine mit Ueberlänge verbundene Zweigipfligkeit.

Die Fringsschen Untersuchungen legen es nahe, auch im nd. Sprachgebiet die Dehnungsverhältnisse nicht über einen Kamm zu scheren. Die von RABELER und auch BREMER (*Nd. Jb.* 53, 1 ff.) z. T. experimentell bewiesene, z. T. nur akustisch konstatierte Zweigipfligkeit fehlt z. B. in der Mundart von Burg in Dithmarschen, deren Quantitätsverhältnisse von STAMMERJOHANN (*Zeitschrift für dtische Maa.* 1914) sehr sorgfältig, auch experimentell, untersucht wurden (vgl. besonders S. 72 ff.).

Die gleiche Eigenschaft, nämlich überwiegend ebene Intonation mit leichter Senkung gegen das Silbenende hin ( $\frac{1}{2}$  Ton — eine Sekunde etwa) zeigen die Maa., auf die ich jetzt zu sprechen komme. Es sind gewisse hannöversche Maa. im Kreise Verden.

Verfolgt man im Sprachatlas die Apokopierungslinie in der Bremer Gegend, so findet man das Wesergebiet von der Mündung der Aller in die Weser abwärts innerhalb des Apokopegebietes; doch gruppieren sich um Bremen in dichter Masse eine Reihe von Maa., die das Endungs-e bewahrt haben. Genaueres Studium mehrerer dieser Maa. hat gezeigt, dass sie — mit einigen Ausnahmen, wo Apokope nachgewiesen werden kann — unter den bekannten Bedingungen Ueberlänge zeigen, obwohl nicht apokopiert, wohl aber synkopiert wurde. Ich habe einstweilen nur die Ma. von Baden genauer untersuchen können. Die quantitative Analyse, die aufgrund der Abschreibung der Schallplatte LA 1449 (Institut für Lautforschung, Berlin) im *Deutschen Spracharchiv*, Berlin - Buch vorgenommen wurde, ergab drei - stufige Quantität. Ueberlänge tritt auf :

I. Vor sth. Konsonanz und folgender Silbe.

1. bei den mundartlichen Vertretern der wg. Längen und Diphthonge und den durch Tondehnung entstandenen :  $\text{my:zə}$  (Mäuse);  $\text{ou:gə}$  (Auge);  $\text{hardə}$  (Heide);  $\text{vɛ:r:gə}$  (Wege).

2. bei alten Kürzen vor geminiertem g :  $\text{ry:gə}$  (Rücken);  $\text{pla:ḡ}$  (Heidsode).

3. bei alten Kürzen vor geminierter Liquida oder Liquida