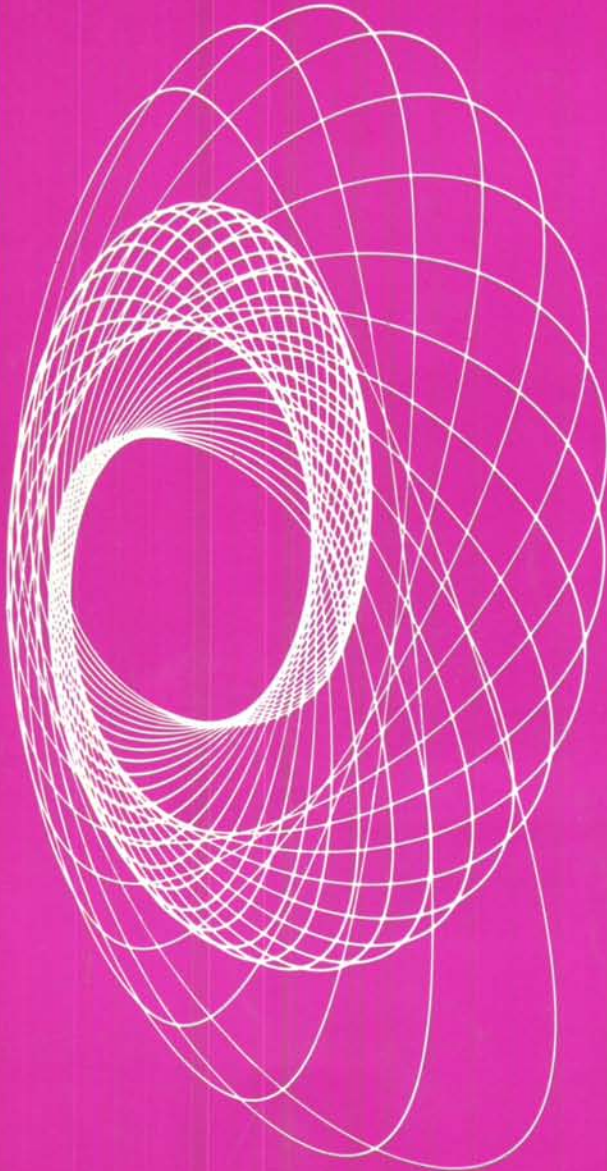


P 21976F

Audiologische Akustik

Audiological Acoustics

ISSN 0172-8261



20. JAHR/VOL. 20

JULI/JULY 1981

HEFT/NO. **4**

Super-Power

der »euphonischen Serie«



AD 426

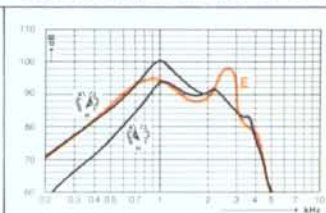
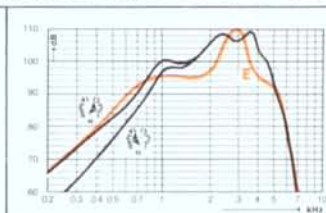
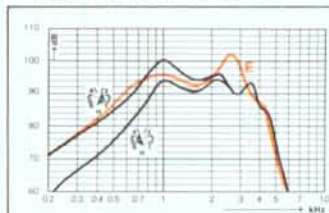
Super-Power • P • G • MT

AD 426-H

Hochton • Super-Power • P • G • MT

AD 426-L

Tieferton • Super-Power • P • G • MT



Normale akustische Wiedergabekurve mit Klangregelmöglichkeiten (rot = etymotisches Winkelstück E)

Indikation: Hochgradige Schalleitungs- und Schallempfindungsstörungen, auch progredienter Art. (Durch Verstärkervorwahl G auch für leichtere Fälle geeignet).

max. Lautstärkepegel 138 dB DIN 131 dB
desgl. bei 1000 Hz 138 dB
max. akust. Verstärkung 67 dB DIN 61 dB
desgl. bei 1000 Hz 67 dB
Frequenzbereich 320–4800 Hz

Indikation: Mittel- bis hochgradige Schalleitungs- und unkomplizierte Schallempfindungsstörungen im Hochtonbereich.

max. Lautstärkepegel 137 dB DIN 131 dB
desgl. bei 2000 Hz 132 dB
max. akust. Verstärkung 65 dB DIN 52 dB
desgl. bei 2000 Hz 62 dB
Frequenzbereich 470–5700 Hz

Indikation: Mittel- bis hochgradige Schalleitungs- und unkomplizierte Schallempfindungsstörungen im Tieftonbereich.

max. Lautstärkepegel 138 dB DIN 130 dB
desgl. bei 1000 Hz 138 dB
max. akust. Verstärkung 67 dB DIN 59 dB
desgl. bei 1000 Hz 67 dB
Frequenzbereich 280–4000 Hz

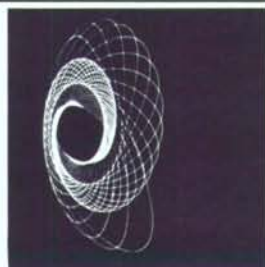


PHILIPS

3 - die es in sich haben

Audiologische Akustik

Audiological Acoustics



Unter Förderung durch / sponsored by H. L. Wullstein, Würzburg
herausgegeben von / edited by G. Flottorp, Oslo – J. Frenkiel, Liège
A. Glorig, Dallas – A. Götze, jun., Budapest – W. Güttner, Bad Tölz
J. D. Hood, London – F. Ingerslev, Lyngby – B. Johansson, Stockholm
R. Lehmann, Le Mans – E. Lehnhardt, Hannover
W. Niemeyer, Marburg – T. Palva, Oulu – H. J. Pichler, Wien
W. Pistor, Eutin – W. Reichardt, Dresden – L. Rüedi, Zürich
G. Sacerdote, Torino – V. Savelli, Milano – H. G. Schmitt, Essen
C. Wansdronek, Eindhoven – H. Weber, Zürich – F. Zöllner, Freiburg

median-verlag
Hans-Jürgen von Killisch-Horn

20. Jahr/Vol. 20

Juli/July 1981

Nr./No. 4

	Inhalt	Contents	
Lisa Guillory-Renner Roberta H. Hansen Donald G. Williamson	Die Wirkungen von vier verschiedenen Verfahren der Reizsignalverarbeitung zur Ermittlung der Reinton-Schwellen	The Effects of Four Different Stimulus Presentation Methods on Pure Tone Thresholds	106
Ernst Terhardt	Sprachsignal, Sprechvorgang und Hörwahrnehmung: Eine vergleichende Übersicht Teil 2	Speech, Speech Signal and Articulation Processes: A Comparative Synopsis Part 2	114
	Aktuelles	Forthcoming Meetings	127
	Personalia	Personalia	130

Die Wirkungen von vier verschiedenen Verfahren der Reizsignaldarbietung zur Ermittlung der Reinton-Schwellen

Lisa Guillory-Renner, Roberta H. Hansen, Donald G. Williamson

Zusammenfassung Vier verschiedene Verfahren der Reizsignaldarbietung wurden verglichen, um das Verfahren zu ermitteln, das die niedrigsten Reintonschwellen ergab. Es handelte sich um die folgenden vier Methoden: kontinuierlich steigend, kontinuierlich fallend, gepulst steigend und gepulst fallend. Für die Messungen zog man normalhörende Personen heran, wobei jede den vier Verfahren ungezielt unterworfen wurde. Für die folgenden Frequenzen wurden Reinton-Schwellen ermittelt: 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz und 6000 Hz. Es zeigte sich, daß das kontinuierlich steigende Verfahren die niedrigsten Schwellen bei den meisten Personen ergab.

Es war das Ziel dieser Untersuchung, ein bevorzugtes Verfahren zur Reinton-Schwellenprüfung unter Benützung der 5 dB-Schwächungsstufen, wie sie in der herkömmlichen Audiometrie eingesetzt werden, zu finden. In einem Bericht von Carhart und Jerger (1952) wurde auf die Notwendigkeit, daß sich die Audiologen auf ein Verfahren zur Reintonschwellenbestimmung einigen, besonders hingewiesen. Dabei beziehen sie sich auf Hirsh (1952), der dazu feststellte: »Da es kein allgemein anerkanntes Standardverfahren gibt, muß dem Arzt eingeschärft werden, daß die Unterschiede innerhalb dieser Ergebnisse möglicherweise zu Abweichungen innerhalb der Schwellen führen, die sich aus dem Verfahren selbst ergeben.« Obgleich Carhart und Jerger (1959) annahmen, daß das überarbeitete Hughson-Westlake-Verfahren das Problem der Abweichungen in den Ergebnissen beseitigen würde, stießen sie in ihren Untersuchungen auf keinen Beweis dafür, daß dieses Verfahren unter Benützung der 5 dB-Stufe etwa niedrigere Schwellen liefert als andere.

Dancer, Ventry und Hill (1976) waren sich darüber im klaren, daß das 5 dB-Meßintervall, wie es in der herkömmlichen Audiometrie verwendet wird, keine Abweichungen in den Schwellen infolge der verschiedenen Darbietung der Reizsignale aufweist. Obwohl in der Literatur weitgehend Einigung darüber herrscht, daß die 5 dB-Stufe zu groß sei, um Schwellenunterschiede auf Grund der verschiedenen Verfahren aufzuzeigen, bleibt die Tatsache bestehen, daß die Audiologen dieses Intervall verwenden und keine Bereitschaft erkennen lassen, davon in der nahen Zukunft abzugehen. Carhart und Jerger (1959) stellten dazu fest: »Wenn die 5 dB-Stufe verwendet wird, darf von jedem Verfahren, das eine klare Aussage ergibt, erwartet werden, daß es bei normalhörenden Personen im wesentlichen gleichwertige Ergebnisse liefert.« Die Forscher nahmen an, daß die 5 dB-Stufe, je nach dem Verfahren der Reizsignaldarbietung, zu keinen übereinstimmenden Schwellen führen sollte.

The purpose of this study was to find a preferred method of puretone threshold testing using the 5 decibel (dB) step of attenuation which is used in conventional audiometry. In a report by Carhart and Jerger (1959), the need for audiologists to agree on a standard method for determining pure-tone thresholds was emphasized. They quoted Hirsh (1952) as saying, "Since there is no accepted standard technique, the clinician must be warned that the differences among these results may be responsible for differences among the thresholds that result from their use." Though Carhart and Jerger (1959) believed adoption of the revised Hughson-Westlake method would eliminate the problem of differences in results, they found no evidence in their research to suggest that this method yielded lower thresholds than other methods, using the 5 dB interval.

Dancer, Ventry, and Hill (1976) agreed that the 5 dB measurement interval used in conventional audiometry does not show possible differences in thresholds by different stimulus presentations. Although much of the literature is in agreement that the 5 dB step is too large to show threshold differences caused by various techniques, the fact remains that audiologists are using this interval and do not appear to be ready to alter it in the near future. Carhart and Jerger (1959) have stated, "When the 5 dB step is used, any technique which elicits on-effect may be expected to yield essentially equivalent findings for subjects with normal hearing." The current researchers hypothesized that the 5 dB step should not yield equivalent thresholds related to various stimulus presentation methods.

Dancer, Ventry, and Hill (1976) also found that the effect of strict instructions to the subjects was greater than the effect of different stimulus modes. Strict instructions kept false-alarm responses to a minimum. However, the competent audiologist should be able to adequately deal with false-alarm responses, by careful observations of the subject; and therefore, keep the ef-

The Effects of Four Different Stimulus Presentation Methods on Pure Tone Thresholds

Summary *Four different stimulus presentation methods were compared to determine which method yielded the lowest pure-tone thresholds. The four methods were ascending-continuous, descending-continuous, ascending-pulsed, and descending-pulsed. Using normal hearing subjects, each subject was presented with the four methods in a random fashion. Puretone thresholds were obtained for the following frequencies: 250 Hertz (Hz), 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz, and 6000 Hz. The ascending-continuous method was found to produce the lowest thresholds in the majority of subjects.*

Dancer, Ventry und *Hill* (1976) stellten fest, daß die genauen, den Versuchspersonen erteilten Anweisungen sich stärker auswirkten als die verschiedenen Formen der Reizsignal darbietung. Genaue Anweisungen beschränkten Falschalarm-Reaktionen auf ein Mindestmaß. Ein erfahrener Audiologe sollte jedoch imstande sein, durch genaue Beobachtung der Versuchsperson solche Reaktionen richtig einzuschätzen und folglich die Auswirkung von Anleitungen möglichst gering zu halten. *Stevens* (1951) berichtete, daß Schwellen tiefer liegen, wenn die Versuchsperson angewiesen wird, »auf alles was nicht Stille ist, zu reagieren«. Er fügte ferner hinzu, daß höhere (schlechtere) Schwellen von Anleitungen herrühren, auf Schall mit Klang- und Tonhöhen eigenschaften zu reagieren.

In der Erörterung der steigenden gegenüber der fallenden Methode der Reintondarbietung besteht noch keine Klarheit, auch nicht hinsichtlich der Frage, ob das Reizsignal als kontinuierlicher oder gepulster Ton angeboten werden soll. In einem Artikel von *Green* (1972) wurde gesagt, daß *Carhart* und *Jerger* (1959) mit der fallenden Methode niedrigere Schwellen erzielten, wobei die Abweichungen allerdings weniger als 5 dB betragen. *Carhart* und *Jerger* (1959) verweisen in ihren Untersuchungen auf *Myers* und *Harris* (1949), die folgendes erklärten: »Bei einer fallenden Serie besteht die Tendenz, etwas bessere Schwellen zu liefern als bei einer steigenden«. Wobei hinzugefügt wurde, daß der Unterschied nur geringfügig 1 dB überschreitet. *Rosenblith* und *Miller* (zitiert bei *Hirsh*, 1952) gaben zu, daß das fallende Verfahren niedrigere Schwellen liefert, sofern ein unterbrochener Ton dargeboten wird.

Sie stellten jedoch fest, daß die Schwellen stets höher (schlechter) waren, sofern kontinuierlich fallende Töne dargeboten wurden, im Gegensatz zu kontinuierlich steigenden Tönen. Bei kontinuierlichen Tönen war in der fallenden Serie ein um 15 dB höherer Schalldruck zur Erzielung von Schwellen erforderlich, als bei der

Methoden der steigenden gegenüber der fallenden Methode der Reintondarbietung besteht noch keine Klarheit, auch nicht hinsichtlich der Frage, ob das Reizsignal als kontinuierlicher oder gepulster Ton angeboten werden soll. In einem Artikel von *Green* (1972), it was stated that *Carhart* and *Jerger* (1959) found the descending method gave lower thresholds, but differences were less than 5 dB. *Carhart* and *Jerger* (1959), in their own study, quote *Myers* and *Harris* (1949) as saying, "There is a tendency for descending series to yield a somewhat better threshold than an ascending series," adding that the difference is little more than 1 dB. *Rosenblith* and *Miller* (cited in *Hirsh*, 1952) agreed that the descending method yields lower thresholds, provided an interrupted tone is presented. However, they stated that when continuous tones were used with a descending series the thresholds were consistently higher (worse) than use of continuous tones with an ascending series. When the continuous tones were used, the descending series required 15 dB more sound pressure to obtain thresholds than the ascending series; however, "the magnitude of the discrepancy increased as the starting level of the descending sequence was raised" (*Carhart* and *Jerger*, 1959).

Methoden der steigenden gegenüber der fallenden Methode der Reintondarbietung besteht noch keine Klarheit, auch nicht hinsichtlich der Frage, ob das Reizsignal als kontinuierlicher oder gepulster Ton angeboten werden soll. In einem Artikel von *Green* (1972), it was stated that *Carhart* and *Jerger* (1959) found the descending method gave lower thresholds, but differences were less than 5 dB. *Carhart* and *Jerger* (1959), in their own study, quote *Myers* and *Harris* (1949) as saying, "There is a tendency for descending series to yield a somewhat better threshold than an ascending series," adding that the difference is little more than 1 dB. *Rosenblith* and *Miller* (cited in *Hirsh*, 1952) agreed that the descending method yields lower thresholds, provided an interrupted tone is presented. However, they stated that when continuous tones were used with a descending series the thresholds were consistently higher (worse) than use of continuous tones with an ascending series. When the continuous tones were used, the descending series required 15 dB more sound pressure to obtain thresholds than the ascending series; however, "the magnitude of the discrepancy increased as the starting level of the descending sequence was raised" (*Carhart* and *Jerger*, 1959).

Subjects

Twenty subjects were used, ten males, ten females, between the ages of eighteen and thirty. All had normal hearing acuity established by a screening test at 20 dB HL ANSI for the frequencies 500 Hz, 2000 Hz, and 6000 Hz. Several of the subjects were familiar with pure-tone-testing as a procedure but had never been exposed to an actual testing situation. The majority were students at the University of Missouri-Columbia;

steigenden Serie. »Der Unterschied nahm jedoch zu, sowie der Anfangspegel der fallenden Reihe erhöht wurde« (Carhart und Jerger, 1959).

Versuchspersonen

Es wurden zehn männliche und zehn weibliche Versuchspersonen im Alter zwischen 18 und 30 Jahren ausgewählt. Alle verfügten über eine normale Hörschärfe, die durch eine Untersuchung bei einem Hörpegel von 20 dB ANSI für 500 Hz, 2000 Hz und 6000 Hz festgestellt wurde. Einige Versuchspersonen waren mit Reintonprüfungen als Verfahren vertraut, wenngleich sie sich nie einer solchen Prüfung unterzogen hatten. Die meisten Versuchspersonen setzten sich aus Studenten der Missouri-Columbia Universität zusammen, zwei Personen hatten jedoch mit der Universität nichts zu tun.

Verfahren

Zunächst nahmen die Prüflinge bequem in einer schalldichten IAC-Kabine, Modell IAC-402-ACT, Platz. Dann legte der Prüfer dem Kandidaten die Kopfhörer an und hängte sich selbst ein Mikrofon um, womit er die Reaktionen des Prüflings überwachen konnte. Bei den Kopfhörern handelte es sich um TDH Hörer in MX 41 AR Kissen, während das Mikrofon der Ausführung SLIM X Modell 777 Kristal angehörte. Für die Prüfung selbst wurde ein Maico MA-24 Audiometer verwendet. Die Eichung des Audiometers wurde während der Untersuchung wöchentlich psycho-akustisch kontrolliert. Jede Versuchsperson wurde bei zwei verschiedenen Gelegenheiten geprüft. Bei der ersten wurde die Hörprüfung zusammen mit zwei der vier verschiedenen Verfahren der Reizsignal darbietung für die Frequenzen 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz und 6000 Hz durchgeführt. Im zweiten Durchgang wurden die zwei restlichen Verfahren für dieselben Frequenzen durchgeführt. Es wurde jeweils nur das rechte Ohr geprüft. Das Reizsignal wurde bei allen 20 Prüflingen in zufälliger Reihenfolge dargeboten, um einer Störung durch einen Lernfaktor vorzubeugen (dies könnte eintreten, wenn eine Methode ständig als letzte an die Reihe kommt und die besten Ergebnisse liefert).

Die Anweisungen zur jeweiligen Methode wurden vorher verlesen. Sie lauteten für die vier Verfahren wie folgt:

1. Kontinuierlich steigend (A-C): Sie werden Töne hören, die etwa 2 Sekunden andauern. Sie werden sehr leise und schwer zu vernehmen sein. Wenn Sie die Töne hören, sagen Sie bitte »Ja«. Warten Sie nicht, bis sie lauter werden.

however, two of the subjects were non-university related persons.

Procedures

Subjects were seated comfortably in a sound-treated IAC booth, Model IAC-402-ACT. The examiner placed the headphones on the subject and attached a microphone around his neck to monitor the subject's responses. The headphones used were TDH 39 earphones mounted in MX 41 AR cushions, and the microphone was a Slim X Model 777 Crystal microphone. A Maico MA-24 audiometer was used for the testing. The calibration of the audiometer was psycho-acoustically checked weekly during the study.

Each subject was tested on two different occasions. During the first session, the screening test was administered along with two of the four different stimulus presentation methods for the frequencies 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz and 6000 Hz. At the second session, the remaining two methods were administered for the same frequencies. Only the subjects' right ears were tested. The order of the stimulus presentation was randomized for all twenty subjects to prevent interference from a learning factor (such as might occur if one method were continuously presented last and yielded the best results).

The instructions for a particular method were read prior to its administration. For the four methods, the instructions were as follows:

1. Ascending-Continuous (A-C): You are going to hear tones which last about two seconds. They will be very low and hard to hear. When you hear the tones I want you to say "yes"; do not wait for them to get loud.
2. Ascending-Pulsated (A-P): You are going to hear tones which beep. They will be very low and hard to hear. When you hear the beeps, I want you to say "yes"; do not wait for them to get loud.
3. Descending-Continuous (D-C): You are going to hear tones which last about two seconds. At first they will be loud enough for you to hear them easily, then they will gradually get softer. When you hear the tones, I want you to say "yes".
4. Descending-Pulsated (D-P): You are going to hear tones which beep. At first they will be loud enough for you to hear them easily, then they will gradually get softer. When you hear the beeps, I want you to say "yes".

A verbal response of "yes" was required to avoid the problem of the "halfup finger response" (Green,

2. Gepulst steigend (A–P): Sie werden Pieptöne hören. Sie werden sehr leise und schwer zu verstehen sein. Wenn Sie die Pieptöne hören, sagen Sie bitte »Ja«. Warten Sie nicht, bis sie lauter werden.

3. Kontinuierlich fallend (D–C): Sie werden Töne hören, die etwa zwei Sekunden andauern. Zunächst werden sie laut genug sein, um von Ihnen gehört zu werden. Sie werden dann allmählich leiser werden. Wenn Sie die Töne hören, sagen Sie bitte »Ja«.

4. Gepulst steigend (D–P): Zunächst werden sie laut genug sein, um von Ihnen mühelos gehört zu werden. Sie werden dann allmählich leiser werden. Wenn Sie die Pieptöne hören, sagen Sie bitte »Ja«.

Die verbale Antwort »Ja« war erforderlich, um das Problem zweideutiger Entscheidungen (Green, 1976) zu vermeiden und um dem Prüfer die Gewißheit zu geben, daß der Prüfling das Reizsignal auch sicher wahrnahm. Falls die Versuchsperson mit »Vielleicht« antwortete, wurde ihr gesagt, daß nur eindeutig bejahende Antworten gelten.

Jeder gepulste Reinton dauerte 500 ms. Sowohl der gepulste Ton als auch der Dauerton wurden mindestens eine Sekunde lang, aber nicht länger als zwei Sekunden dargeboten. Jeder Signaldarbietung folgte eine tonlose Pause von nicht weniger als 1 Sekunde Dauer, um einer Anpassung des Gehörs vorzubeugen. Anschließend wurden die Schwellen für die sechs Frequenzen bestimmt, wobei als Schwellen die zwei bejahenden — nicht unbedingt aufeinanderfolgenden — Antworten

1976), and to assure the examiner that the subject was certain of the presence of the stimulus. If the subject responded with "maybe", he was then told that only positive affirmative responses would be accepted.

Each pulsed pure tone beat lasted a duration of 500 milliseconds. Both the pulsed and the continuous tones were presented for at least one second but no more than two seconds. Each tonal presentation was separated by a toneless interval of not less than one second in order to prevent auditory adaptation. The thresholds were then established for the six frequencies, with threshold being defined as two affirmative responses (not necessarily consecutive) at the lowest intensity level the subject could perceive. This was independent of the number of crossings of actual threshold required to establish the above defined threshold.

The descending method was administered as follows:

1. the first stimulus was presented at 40 dB HL,
2. the stimulus was then presented in decreasing steps of 10 dB until the subject no longer responded,
3. threshold crossings were then carried out in a "10 dB up – 5 dB down" fashion.

The ascending method was administered as follows:

1. the first stimulus was presented at –25 dB and gradually increased until the subject responded,
2. threshold crossings were then carried out in a "10 dB down – 5 dB up" fashion.

Tabelle 1 Statistischer Vergleich der vier Methoden

Table 1 Statistical comparison of the four methods

		Mittelwert/ Mean	Standardabweichung/ Standard Deviation	Übereinstimmung/ Correlation	T-Test/ t-test
Vergleich I Comparison I	AC-DC:	* AC \bar{X} = –3.584 DC \bar{X} = –0.708	von/of AC = 4.259 von/of DC = 4.165	R = 0.716	**t = –4.124
Vergleich II Comparison II	AP-DP:	AP \bar{X} = –0.458 DP \bar{X} = –0.334	von/of AP = 3.582 von/of DP = 3.327	R = 0.597	t = +0.179
Vergleich III Comparison III	AC-AP:	* AC \bar{X} = –3.584 AP \bar{X} = –0.458	von/of AC = 4.259 von/of AP = 3.582	R = 0.704	**t = –4.536
Vergleich IV Comparison IV	DP-DC:	DP \bar{X} = –0.334 DC \bar{X} = –0.708	von/of DP = 3.327 von/of DC = 3.165	R = 0.634	t = +0.512
Vergleich V Comparison V	AC-DP:	* AC \bar{X} = –3.584 DP \bar{X} = –0.334	von/of AC = 4.259 von/of DP = 3.327	R = 0.664	**t = –4.512
Vergleich VI Comparison VI	AP-DC:	AP \bar{X} = –0.458 DC \bar{X} = –0.708	von/of AP = 3.382 von/of DC = 4.165	R = 0.812	t = +0.462

* Methode mit dem niedrigsten Mittelwert in dB/Method yielding lowest mean in decibels

** Auffallende Unterschiede/Significant differences

beim niedrigsten vom Prüfling wahrgenommenen Intensitätspegel galten.

Dies war unabhängig von der Anzahl der Überschreitungen der tatsächlichen Schwelle, um die vorher definierte Schwelle festzulegen.

Beim fallenden Verfahren wurde der Ton wie folgt dargeboten:

1. das 1. Reizsignal wurde bei einem Hörpegel von 40 dB dargeboten;
2. anschließend wurde das Reizsignal in Stufen von jeweils 10 dB abfallend angeboten, bis der Prüfling nicht mehr reagierte;
3. Überschreitungen der Schwelle um »10 dB nach oben – 5 dB nach unten«.

Beim steigenden Verfahren wurde der Ton wie folgt dargeboten:

1. das erste Reizsignal wurde bei – 25 dB angeboten und dann allmählich erhöht, bis die Versuchsperson antwortete;
2. Überschreitungen der Schwelle um »10 dB nach unten – 5 dB nach oben«.

Ergebnisse

Es wurden Reinton-Durchschnittswerte für jede der vier Methoden (für jede Versuchsperson) unter Verwendung aller sechs geprüften Frequenzen berechnet. Dies geschah, um die vier Verfahren leichter vergleichen zu können. Aus den Reinton-Durchschnittswerten der 20 Versuchspersonen sowie aus den folgenden Tonprüfungen: AC-DC (Vergleich I, siehe Tabelle 1), AP-DP (Vergleich II), AC-AP (Vergleich III), DP-DC (Vergleich IV), AC-DP (Vergleich V), und AP-DC (Vergleich VI) wurde ein Mittelwert für jedes Verfahren berechnet.

Von den 20 Prüflingen erzielten 10 niedrigere (bessere) Reinton-Durchschnitte bei der kontinuierlich steigenden Methode, fünf erzielten gleiche Reinton-Durchschnittswerte bei AC und eine Versuchsperson bei den anderen Methoden. Drei Versuchspersonen erreichten ihre niedrigste Schwelle beim DV-Verfahren, eine bei DP und ebenso eine bei AP. Die Ergebnisse besagen also, daß 15 Versuchspersonen ihre besten Reinton-Schwellen mit dem AC-Verfahren erzielten.

Diskussion

Läßt man unberücksichtigt, ob die steigende oder fallende Methode benutzt wurde, ergibt sich anhand der Ergebnisse, daß die Darbietung des Dauertons die besseren Schwellen lieferte (Tabelle 1, Vergleiche III und

Results

Pure tone averages were calculated for each of the four methods (for each subject) using all six frequencies tested. This was done in order to make an easier comparison of the four methods. A mean for each method was then calculated from the pure-tone averages of the twenty subjects, and the following t-tests run: AC-DC (comparison I, see Table 1), AP-DP (comparison II), AC-AP (comparison III), DP-DC (comparison IV), AC-DP (comparison V), and AP-DC (comparison VI).

Out of the twenty total subjects, ten yielded lower (better) puretone averages for the ascending continuous (AC) method, and five subjects yielded the same pure-tone average for AC and one of the other three methods. Three subjects yielded their lowest threshold for the DC method; one the DP method;

Tabelle 2 Reinton-Durchschnittswerte für die Frequenzen 250 Hz bis 6000 Hz.

Table 2 Pure-tone averages for frequencies 250 Hz – 6000 Hz

Versuchspersonen Subjects	AC	DP	DC	AP
1	– 1.67	0.00	– 3.33	– 0.83
2	– 5.00*	– 0.83	– 0.83	– 1.67
3	– 10.00*	– 2.50	– 2.50	0.00
4	– 9.17*	– 0.83	– 8.33	– 8.33
5	– 0.83*	+ 0.83	+ 1.67	0.00
6	– 11.67*	– 1.67	– 7.50	– 5.00
7	– 10.00*	– 5.00	– 1.67	– 2.50
8	– 6.67*	– 1.67	– 1.67	– 5.00
9	– 1.67*	+ 6.67	+ 8.33	+ 6.67
10	– 2.50*	0.00	– 2.50*	– 1.67
11	– 1.67*	+ 0.83	– 1.67*	+ 0.83
12	+ 2.50*	+ 3.33	+ 7.50	+ 2.50*
13	– 4.17*	– 1.67	– 0.83	– 4.17*
14	+ 0.83*	+ 2.50	+ 3.33	+ 5.00
15	– 7.50	– 10.00*	– 4.17	– 2.50
16	– 2.50*	– 0.83	0.00	+ 1.67
17	– 1.67	+ 0.83	– 2.50*	0.00
18	0.00*	0.00*	+ 2.50	+ 2.50
19	+ 2.50	+ 2.50	+ 1.67	+ 0.83*
20	– 0.83	+ 0.83	– 1.67*	+ 0.83
	* \bar{X} = 3.584	\bar{X} = 0.334	\bar{X} = 0.658	\bar{X} = 0.458

* Methode mit dem niedrigsten Mittelwert in dB
Method yielding lowest mean in decibels

Drei hervorragende Gründe, Knowles zu wählen



EF-HÖRER-SERIE

Dies ist ein Hörer in BK-Größe in einer sehr gefragten Sparte.

Er hat eine erweiterte Bandbreite bei guter Mittelbereichsempfindlichkeit. Die EF-Serie hat all die verbesserten Knowles-Eigenschaften einschließlich geringerer magnetischer Strahlung und verbessertem Stoßschutz. Sehr wichtig ist, daß sie wegen ihrer hervorragenden Tonqualität auch bei subjektiven Tests gewählt wurde.

* Sämtliche Fotografien in tatsächlicher Größe.

AN ALLE HÖRGERÄTE-HERSTELLER

Mit diesen zusätzlichen ausgezeichneten Hörern verfügt Knowles über ein unvergleichliches Angebot, das den meisten Hörstörungen gerecht wird. Dies bedeutet außerdem, daß es drei hervorragende Gründe gibt, Knowles Electronics zu wählen, setzen Sie sich deshalb baldmöglichst mit uns in Verbindung, und wir geben Ihnen nähere Informationen.



CI-HÖRER-SERIE

Mit dieser neuen Serie hat Knowles zum ersten Mal einen Hörer mit einer Spitzen-

Leistung von 140 dB und mit einer Größe, die für Hinter-dem-Ohr-Hörgeräte geeignet ist, während eine flache Wiedergabekurve über die ganze Bandbreite enthalten ist. Weitere Besonderheiten sind: geringere magnetische Strahlung, geringere Verzerrung, verbessertes Spitze-Tal-Verhältnis. Dies alles zusammengenommen bedeutet: ein neuer Hörer, der einem Konstrukteur die Möglichkeit geben sollte, ein Hinter-dem-Ohr-Hörgerät in Betracht zu ziehen, mit einer ähnlichen Leistungsfähigkeit wie ein Taschengerät.



ED-HÖRER-SERIE

Mit den Abmessungen 6,35 x 4,34 x 3,02 mm ist dies wahrscheinlich

der kleinste heutzutage erhältliche Hörer; er ist nur fast halb so groß wie ein Standardhörer für Hinter-dem-Ohr-Hörgeräte. Die ED-Serie hat eine große Bandbreite und gute Mittelbereichsempfindlichkeit. Durch ihre Größe und Empfindlichkeit sind diese Hörer besonders geeignet für Im-Ohr-Geräte und Hinter-dem-Ohr-Geräte.

Ebenso wie die EF-Serie hat die ED-Serie verbesserten Stoßschutz und geringere magnetische Strahlung bei besserer allgemeiner Tonqualität.

Knowles

Knowles Electronics Co.,
Victoria Road, Burgess Hill, Sussex RH15 9LP, England
Fernschreiber: 87460.



Tel.: (04446) 5432. Telegramme: Eleknoj, Burhill.
Für technische Einzelheiten: David M. Wadiow.

IV). Ein Vergleich der Dauerton-Verfahren läßt erkennen, daß die steigende Methode eine um etwa 3 dB niedrigere Schwelle ergab (Vergleich I). Wenn die Verfahren mit gepulstem Ton verglichen werden, ergibt sich für die steigende Methode ebenfalls eine geringfügig bessere Schwelle (Vergleich II). Die Vergleichswerte aus den 6 Vergleichen zeigen, daß sich die 4 Methoden nicht wesentlich voneinander unterscheiden. Die hohe Übereinstimmung zwischen den Vergleichen I und III deuten auf eine geringfügig bessere Leistung der Dauerton-Darbietung (Vergleich I) und der steigenden Methode hin (Vergleich III). Die höchste Übereinstimmung ergab der Vergleich zwischen AP und DC (Vergleich VI). Dies bedeutet jedoch nur, daß die beiden Verfahren für die Versuchspersonen ähnliche Ergebnisse lieferten.

Eine weitere Untersuchung der Mittelwerte führt zur interessanten Feststellung, daß die AC-Methode im Vergleich zu anderen eine wesentlich geringere mittlere Schwelle lieferte (um etwa 3 dB niedriger) und bei allen diesen Vergleichen ergab der statistische T-Test, daß die AC-Methode wesentlich von den anderen Verfahren abweicht.

Schlußfolgerung

In der herkömmlichen klinischen Audiometrie, wo das Reizsignal in Stufen von 5 dB angeboten wird, ergaben sich bei der kontinuierlich steigenden Tondarbietung niedrigere Schwellenwerte als bei der gepulst steigenden, kontinuierlich fallenden oder gepulst fallenden. Diese Untersuchung widerlegt die Literatur, in der festgestellt wird, daß die 5 dB-Stufe zu groß sei, um individuelle Schwellenunterschiede auf Grund der verschiedenen Verfahren aufzuzeigen. Dieser Widerspruch soll zu weiteren Untersuchungen anregen in der Erwartung, daß es eines Tages zu einer einzigen standardisierten Prüfmethode zur Ermittlung der Reizton-Schwellen kommt.

Literatur/References

- Carhart, R. and Jerger, J. 1959. Preferred method for clinical determination of pure-tone thresholds. *J. Speech. Hearing Dis.*, 24, 330.
- Dancer, J., Ventry, I. and Hill, W. 1976. Effects of stimulus presentation and instructions on pure-tone thresholds and false-alarm responses. *J. Speech Hearing Dis.*, 41, 315.
- Green, D. S. 1972. Pure-tone air conduction thresholds. In *Handbook*

and one, the AP method. The results thus indicate that fifteen of the twenty subjects achieved their best pure tone thresholds when the AC method was used.

Discussion

In analyzing the results, regardless of whether the ascending series or descending series was used, the continuous stimulus gave better thresholds. (Table I, comparisons III and IV). It appears that when the continuous stimulus modes were compared, the ascending method yielded a lower threshold by approximately 3 dB (comparison I). Also, when comparing the pulsed modes, the ascending method gave a slightly better threshold (comparison II). The correlation figures derived from the six comparisons indicate that none of the four methods is grossly different from the rest; however, the high correlation of comparisons I and III are indicative of the slightly higher performances of the continuous stimulus mode (comparison I) and the ascending method (comparison III). The highest correlation resulted from the comparison between AP and DC (comparison VI); however, this only indicates that the two methods yielded similar results for the subjects used in this study.

In further analysis of the means, it is interesting to note that whenever the AC method was compared to another method, it yielded a considerably lower mean threshold (by approximately 3 dB); and in each of those comparisons the t-test indicated that the AC method was significantly different from the other methods.

Conclusion

In conventional clinical audiometry, where the 5 dB increment is used, the ascending-continuous method has yielded lower thresholds than the ascending-pulsed, descending-continuous, or descending-pulsed. This study contradicts the literature indicating that the 5 dB step is too large to show individual threshold differences caused by various methods. This controversy invites further exploration in hopes that one day a single method of pure tone threshold testing will be standardized.

of Clinical Audiology (ed. J. Katz): Williams & Wilkins, Co., Baltimore.

Hirsh, I. J. 1952. *The Measurement of Hearing*. McGraw-Hill Book Co., Inc. New York.

Licklider, J. C. R. 1951. Basic correlates of the auditory stimulus. In *Handbook of Experimental Psychology* (ed. S. Stevens). John Wiley & Sons, Inc., New York.

»Sonderheft 1979«

der »Zeitschrift für Hörgeräte-Akustik«

Dieses Sonderheft enthält zahlreiche Referate zum Themenkreis »Retrocochleäre Hörstörungen« sowie freie Vorträge in teilweise erweiterter Fassung.

Sämtliche Referate entstammen der Tagung der »Arbeitsgemeinschaft Deutscher Audiologen und Neurootologen«, die vom 8. bis 10. März 1979 in München abgehalten wurde.

Das »Sonderheft 1979« umfaßt auf 100 Seiten 14 Arbeiten mit über 110 zum Teil mehrfarbigen Abbildungen. 20 Autoren haben an diesem Sonderheft mitgewirkt.

Ein Blick aufs Inhaltsverzeichnis wird auch Sie überzeugen!

H.-M. Theopold	Morphologische Grundlagen retrocochleärer Hörstörungen
Reinhard Schunicht	Untersuchungen zur Altersabhängigkeit der späten ERA-Reizantwort
Dieter Mrowinski	Hirnstamm-Audiometrie bei cochleären Funktionsstörungen
Konrad Maurer	Veränderungen der Frühen Akustisch Evozierten
Ernst Schäfer	Potentiale (FAEP) beim Akustikusneurinom (AN)
Mohamed Y. Abdel Aziz	
Harald Leitner	
Raimund Brix	Der »Objektive Frequenz Dekrement Test« (OFDT): seine diagnostische Aussagekraft für retrocochleäre Hörstörungen
Michael Handrock	Die Bedeutung des Sympathikotonus auf den temporären
Rudi L. Fischer	und bleibenden Hörverlust nach Beschallung
Arndt v. Lüpke	Zur audiometrischen Ermittlung der MdE
Ulrich Eysholdt	Psychoakustische Tuningskurve – eine Verfeinerung der audiometrischen Diagnostik?
Friedrich Keller	Sprachverständlichkeit und Hörgerät
Ludwig M. Moser	Die Kenndaten eines Hörgerätes in einheitlicher, räumlicher Darstellungsform als Voraussetzung für vergleichende Gruppierungen
Thomas Spillmann	Zur quantitativen Darstellung von Parametern des akustischen Reflexes
Norbert Dillier	mit verschiedenen Impedanzmeßbrücken
Jürgen Kießling	Frequenzspezifische Stapediusreflex-Auslösung zur objektiven Hörgeräte-Anpassung
Günter Esser	Hörgeräte-Anpassung bei Kindern mit Hilfe der
Hartmut Rodina	Stapediusreflex-Audiometrie
Friedrich Pfander	TTS und Rückwanderungszeit im Audiogramm als Maßstäbe für die Beurteilung hörgefährdender akustischer Einwirkung (Kurzfassung)

Ernst Terhardt

Zusammenfassung Das Sprachsignal, zweckmäßigerweise durch das Kurzzeit-Spektrum repräsentiert, weist Merkmale auf, die sich einerseits mit dem Sprechvorgang, andererseits mit der Hörwahrnehmung in Zusammenhang bringen lassen. Man kann sinnvollerweise die Merkmale in solche der Zeitstruktur, der Spektralstruktur und des Frequenz-Zeitmusters einteilen. Bezüglich der Zeit- und der Spektraldimension kann man jeweils Makro- und Mikrostrukturen unterscheiden. Es wird ein Überblick über die sprachakustischen Ursachen der Merkmale sowie ihre Zusammenhänge mit der Hörwahrnehmung gegeben. Dabei zeigt sich ein hohes Maß wechselseitiger Anpassung und Beeinflussung von Sprachsignal, Sprechvorgang und Hörwahrnehmung. Das zentrale Gehör zeigt eine Anpassung seiner Verarbeitungsalgorithmen an die Sprachsignaleigenschaften einerseits und die Eigenschaften des peripheren Gehörs andererseits. Die zentralen, ganzheitlichen Verarbeitungsvorgänge scheinen durch die sprachliche Kommunikation stark beeinflusst zu sein.

3. Spektralstrukturmerkmale

Auch hinsichtlich der spektralen Intensitätsverteilung des Sprachsignals kann man eine Makro- und eine Mikrostruktur unterscheiden. Das in Abb. 1 gezeigte Spektrogramm gibt nur die erstere wieder, weil die Bandbreite des Analysefilters verhältnismäßig groß ist. Dafür veranschaulicht das so gewonnene Spektrogramm wichtige Globalstrukturen, vor allem die Formanten der Vokale. Verzichtet man auf Darstellung der Zeitabhängigkeit, so kann man die Amplitudenspektren einzelner Laute wie in Abb. 2 darstellen. Hier wird deutlich, daß der Vokal /a/ durch ein periodisches Signal (die Schallflußimpulse durch die Glottis) erzeugt wird, denn das Spektrum besteht in seiner Mikrostruktur aus äquidistanten Teilschwingungen. Das Amplitudenspektrum der Glottisimpulse nimmt ungefähr mit $1/f^2$ nach höheren Frequenzen ab, woraus sich im wesentlichen der Höhenabfall des Vokalspektrums ergibt (vgl. [4], [8]). Die Formanten (das heißt, die Maxima der Makrostruktur) entstehen durch die Dämpfungsminima, welche die vom Vokaltrakt gebildete akustische Leitung aufweist. Dieser Übertragungsweg erstreckt sich von der Glottis bis zur Mundöffnung und ist 17 bis 18 cm lang. Seine örtliche Querschnittsverteilung — und damit seine akustischen Übertragungseigenschaften — können durch die Artikulationsorgane (Lippen, Zunge, Kiefer, Gaumensegel) stark verändert werden, wodurch sich die Vielfalt möglicher Formantkombinationen ergibt (vgl. *Fant* [4]). Wäre der Vokaltrakt ein zylindrisches Rohr (das heißt, eines mit überall demselben Öffnungsquerschnitt), so lägen seine Übertragungsresonanzen bei den ungeradzahigen Vielfachen derjenigen Frequenz, bei welcher gerade eine Viertelwellenlänge hineinpaßt, das heißt, die Formantfrequenzen wären $F_n = (2n-1)c/(4L)$, wobei L die Vokaltraktlänge, c die Schallgeschwindigkeit und n eine positiv ganze Zahl

3. Characteristics of the spectral structure

A macrostructure and microstructure can also be distinguished with regard to the *spectral* intensity distribution of the speech signal. The spectrogram (Fig. 1) only reproduces the macrostructure because the bandwidth of the analyzing filter is relatively large. But, in turn, the resulting spectrogram illustrates important overall structures especially the formants of the vowels. If we dispense with the representation of the time dependence, the amplitude spectra of single sounds can be presented as in Fig. 2. From this, we can conclude that the vowel /a/ is generated by a periodic signal (sound flow pulses through the glottis), for the microstructure of the spectrum consists of equidistant component vibrations. The amplitude spectrum of glottic pulses decreases by roughly $1/f^2$ toward the higher frequencies and this mainly accounts for the decrease in height of the vocal spectrum (see [4], [8]). The formants (i.e. the maxima of the macrostructure) are produced by the attenuation minima which characterize the acoustic path formed by the vocal tract. This transmission path extends from the glottis to the mouth opening and measures 17 to 18 cm in length. Its local cross-section distribution — and thus its acoustic transmission properties — can be changed to a great extent by the organs of articulation (lips, tongue, jaw, soft palate) and this leads to the multitude of possible formant combinations (see *Fant* [4]). If the vocal tract were a cylindrical tube (i.e. with the same cross-section over its entire length), then its transmission resonances would be an odd multiple of that frequency at which a quarter wavelength just fits in, i.e. the formant frequencies would be $F_n = (2n-1)c/(4L)$, in which L denotes the vocal tract length, c the velocity of sound and n a positive integer ($n=1, 2, 3, \dots$). This gives in round figures $F_1 = 500$ Hz, $F_2 = 1,500$ Hz, $F_3 = 2,500$ Hz and so forth. For the "neutral vowel" /e/, these

Speech, Speech Signal and Articulation Processes: A Comparative Synopsis

Part 2

Summary The speech signal, conveniently represented by the short-time spectrogram possesses characteristics which can be related both to the articulation process and to auditory perception. It appears logical to classify them into characteristics of the temporal structure, the spectral structure and of the frequency time pattern. As to the temporal and the spectral dimensions, both macro- and micro structures can be distinguished. An overview of the speech-acoustic causes of the characteristics and their interrelationships with auditory perception is given. In this context, a high degree of reciprocal adjustment and interaction between speech signal, articulation and auditory perception is seen to exist. The central hearing manifests an adaptation of its processing algorithms to speech signal properties on the one hand and to the properties of peripheral hearing on the other. The central integral processing operations appear to be strongly influenced by speech communication.

($n = 1, 2, 3, \dots$) ist. Daraus ergeben sich die runden Zahlenwerte $F_1 = 500$ Hz, $F_2 = 1500$ Hz, $F_3 = 2500$ Hz usw. Für den »neutralen Vokal« / ∂ / sind diese Formantfrequenzen in der Tat typisch, denn bei seiner Artikulation ist der Vokaltraktquerschnitt über der Länge angenähert konstant. Da alle anderen Vokale artikulatorisch als »Abweichungen« von diesem mittleren Vokal angesehen werden können, liefern die oben genannten Daten zugleich repräsentative Angaben über die mittleren Spektraleigenschaften der Vokale im allgemeinen: Die Formantfrequenzen haben im Mittel ungefähr einen Frequenzabstand von 1000 Hz vonein-

formant frequencies are indeed typical for when the vowel is articulated, the vocal tract cross-section is roughly constant over its whole length. Since all other vowels can be regarded as articulatory "deviations" from this mean vowel, the above-mentioned figures simultaneously furnish representative information on the mean spectral properties of the vowels in general: On the average, the formant frequencies are spaced at approximately 1000 Hz from each other and the existence range of the first formant lies between 0 and 1 kHz, that of the second formant between 1 and 2 kHz and so forth.

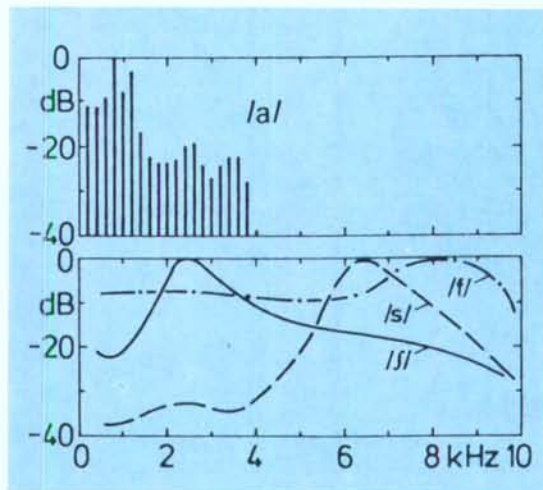


Abb. 2 Beispiele einiger Sprachlautspektren, leicht schematisiert

Fig. 2 Examples of some speech sound spectra, schematic representation

It can be clearly seen from Fig. 2 that the vowel / a / (represented here with a vocal cord frequency of 200 Hz) also satisfies these requirements: The first two formants lie at 750 and 1,200 Hz, respectively, the third and fourth are close to 2.5 and 3.5 kHz. Owing to the acoustics of the vocal tract, the third formant (and this applies the more so to all further formants) is only slightly influenced by articulation which means it is only slightly shifted from its central position of 2,500 Hz. In other words, the information characterizing the respective vowel is conveyed almost exclusively by the first two formants. In the spectrogram (Fig. 1) the third formant of all occurring vowels is found in the range from 2,500 Hz to 3,000 Hz and the fourth formant, to be expected at 3,500 Hz, hardly appears at all.

In the lower part of Fig. 2, the density spectra of intensity of the unvoiced fricatives / f /, / s / and / $\ʃ$ / are represented in a slightly schematic form, as further examples of stationary speech sound spectra. Owing to the manner in which these sounds are generated (air-flow noise at a narrow point of the vocal tract) the spectra are without systematic microstructures. Three macrostructures of the spectra which are typical of and distinguishable for the fricatives indicated can be

ander und der Existenzbereich des ersten Formanten liegt zwischen 0 und 1 kHz, derjenige des zweiten Formanten zwischen 1 und 2 kHz, usw.

In Abb. 2 ist leicht zu erkennen, daß auch beim Vokal /a/ (hier mit 200 Hz Stimmbandfrequenz dargestellt) diese Regeln erfüllt sind: Die ersten beiden Formanten liegen bei 750 bzw. 1200 Hz, der dritte und vierte dicht bei 2,5 bzw. 3,5 kHz. Bedingt durch die Akustik des Vokaltrakts wird allgemein der dritte Formant (und erst recht alle weiteren) durch die Artikulation nur noch wenig beeinflusst, das heißt, aus seiner mittleren Lage von 2500 Hz verschoben. Dies bedeutet, daß die den betreffenden Vokal kennzeichnende Information nahezu ausschließlich von den ersten beiden Formanten übermittelt wird. Im Spektrogramm Abb. 1 findet man den dritten Formanten sämtlicher vorkommender Vokale im Bereich 2500 Hz bis 3000 Hz und der bei 3500 Hz zu erwartende Formant tritt kaum in Erscheinung.

Als weitere Beispiele stationärer Sprachlautspektren sind im unteren Teil von Abb. 2 die Intensitätsdichtespektren der stimmlosen Frikativlaute /f/, /s/ und /ʃ/ leicht schematisiert dargestellt. Wegen der Art der Schallerzeugung dieser Laute (Strömungsgeräusch an einer Engstelle des Vokaltrakts) handelt es sich hierbei um Spektren ohne systematische Mikrostruktur. Man erkennt deutlich drei für die angegebenen Frikativlaute typische und unterscheidbare Makrostrukturen der Spektren. Die Unterschiede sind darauf zurückzuführen, daß die Art der erwähnten Engstelle und vor allem ihr Ort im Vokaltrakt bei den drei Lauten verschieden sind (Oberzähne/Unterlippe bei /f/, Zunge/Zahndamm bei /s/, Zunge/Gaumen bei /ʃ/). Dadurch unterscheiden sich einesteils schon die an der betreffenden Engstelle angeregten Geräuschspektren, andernteils die Einflüsse der angekoppelten Vokaltrakträume und der Schallübertragung nach außen.

Die von den Plosivlauten erzeugten Geräuschspektren weisen ebenfalls systematische Unterschiede der Makrostruktur auf, welche zur Unterscheidbarkeit dieser Laute beitragen. Auch diese Unterschiede beruhen auf der Artikulationsart und -Stelle. Die Nasallaute /m/, /n/, /ŋ/ schließlich haben eine weitgehend einheitliche, konstante spektrale Makrostruktur, welche durch feste Formanten in der Umgebung der Frequenzen 250, 1000, 2000 und 4000 Hz gekennzeichnet ist. Sie entsprechen den durch Artikulation nahezu gar nicht veränderbaren Übertragungsresonanzen des Nasaltrakts.

Von einer spektralen Mikrostruktur kann man unter gewissen Voraussetzungen in Bezug auf die stimmhaf-

clearly discerned. The differences are attributable to the fact that the nature of the narrowing mentioned and above all its location in the vocal tract are different for the three sounds (upper teeth/lower lip with /f/, tongue/tooth with /s/, tongue/palate with /ʃ/). Thus, the noise spectra excited at the point of narrowing differ, and the influences of the coupled vocal tract spaces and the sound transmission to the outside also differ.

The noise spectra generated by plosives also show systematic differences in the macrostructure which contribute to the distinguishability of these sounds. These differences again are due to the type and point of articulation. Finally, the nasal sounds /m/, /n/, /ŋ/ possess a largely uniform constant spectral macrostructure which is characterized by fixed formants near the frequencies 250, 1000, 2000 and 4000 Hz. They correspond to the transmission resonances of the nasal tract which is hardly changed at all by articulation.

When certain prerequisites exist with respect to voiced speech sounds, we can speak of a spectral microstructure. The periodic pulse sound source, the glottis vibration, generates a harmonic partial tone spectrum. The partial tones are related to each other in the frequencies ratios 1:2:3:4:5 and so forth. This means that their mutual frequency spacing is equal to the fundamental frequency (vocal cord frequency). Pre-supposing that the vocal cord frequency is small as compared with the frequency spacings of the formants, then a microstructure and macrostructure can be reasonably differentiated. The figures indicated above for fundamental frequencies and formant spacings show that this preliminary condition is in many instances fulfilled or approximately so in natural speech. However, this situation can be completely reversed in singing: The fundamental frequency can become higher than the formant spacing, so that the partial tone spectrum becomes the macrostructure, whereas the formant structure is greatly diminished or disappears altogether. Voiced sounds are the vowels (for instance /a/, /e/, /i/, /o/, /u/), the nasal sounds (/m/, /n/, /ŋ/), the voiced fricatives (for instance /z/, /ʒ/, /v/) and the "soft" plosives (/b/, /d/, /g/). The spectral macrostructure of these sounds is represented, as indicated above, by "samples" at the spacing of the vocal cord frequency. In order to gain macrostructure, interpolation is necessary between neighbouring samples. In the spectrogram in Fig. 1, this is done simply by making analysis bandwidth greater than the highest vocal cord frequency occurring. In natural speech vocal cord frequencies above 300 Hz are relatively rare, and for this reason an analysis bandwidth of 300 Hz has proved to be

ten Sprachlaute sprechen. Die periodische Impulschallquelle, die Glottisschwingung, erzeugt ein harmonisches Teiltonspektrum. Die Teiltöne stehen in den Frequenzverhältnissen 1:2:3:4:5 usw. zueinander. Das heißt, ihr gegenseitiger Frequenzabstand ist gleich der Grundfrequenz (Stimmbandfrequenz). Unter der Voraussetzung, daß die Stimmbandfrequenz klein ist gegenüber den Frequenzabständen der Formanten, kann man also sinnvoll Mikro- und Makrostruktur unterscheiden. Die zuvor angegebenen Zahlen über Grundfrequenzen und Formantabstände zeigen, daß diese Voraussetzung bei natürlicher Sprache in vielen Fällen bzw. angenähert erfüllt ist. Beim Gesang können sich dagegen die Verhältnisse durchaus umkehren: Die Grundfrequenz kann höher werden als der Formantabstand, so daß das Teiltonspektrum zur Makrostruktur wird, während die Formantstruktur stark beeinträchtigt ist oder verschwindet. Stimmhafte Laute sind die Vokale (z.B. /a/, /e/, /i/, /o/, /u/), die Nasallaute (/m/, /n/, /ŋ/), die stimmhaften Frikativlaute (z.B. /z/, /ʒ/, /v/) und die »weichen« Plosivlaute (/b/, /d/, /g/). Die spektrale Makrostruktur dieser Laute wird nach dem zuvor Gesagten durch »Stützwerte« im Abstand der Stimmbandfrequenz repräsentiert. Um die Makrostruktur zu gewinnen, muß man zwischen benachbarten Stützwerten interpolieren. Dies geschieht beim Spektrogramm nach Abb. 1 einfach dadurch, daß die Analysierbandbreite etwas größer als die höchste vorkommende Stimmbandfrequenz gemacht wird. In natürlicher Sprache sind Stimmbandfrequenzen oberhalb 300 Hz schon relativ selten, weshalb sich eine Analysierbandbreite von 300 Hz als günstig bewährt. Dadurch wird, zugunsten der Formantstruktur, die von den harmonischen Teiltönen gebildete Mikrostruktur verwischt. Diese Art der Spektraldarstellung ist allerdings ein Kompromiß. Sie ist nicht völlig gehöradäquat, denn in der auditiven Wahrnehmung spielen die einzelnen Teiltöne der stimmhaften Sprachklänge, also die spektrale Mikrostruktur, durchaus eine Rolle. Man kann allgemein sagen, daß der tonale Charakter der stimmhaften Sprachlaute darauf beruht, daß das periphere Gehör diese Klänge in gewissem Umfang in ihre Teiltöne zerlegt.

Die Wahrnehmung der spektralen Mikro- und Makrostruktur kann anhand des sogenannten Erregungspegel-Frequenzmusters und des Spektraltonhöhenmusters veranschaulicht werden. Abb. 3 zeigt Beispiele für diese Muster, zusammen mit den Teiltonspektren der beiden Vokale /a/ und /e/, wobei die Stimmbandfrequenz 200 Hz beträgt. Der Erregungspegel L_E ist eine aus Mithörschwellen abgeleitete psychoakustische Modellgröße (vgl. Zwicker u. Feldtkeller [26]). Sein Verlauf als Funktion der Frequenz spiegelt das frequenzse-

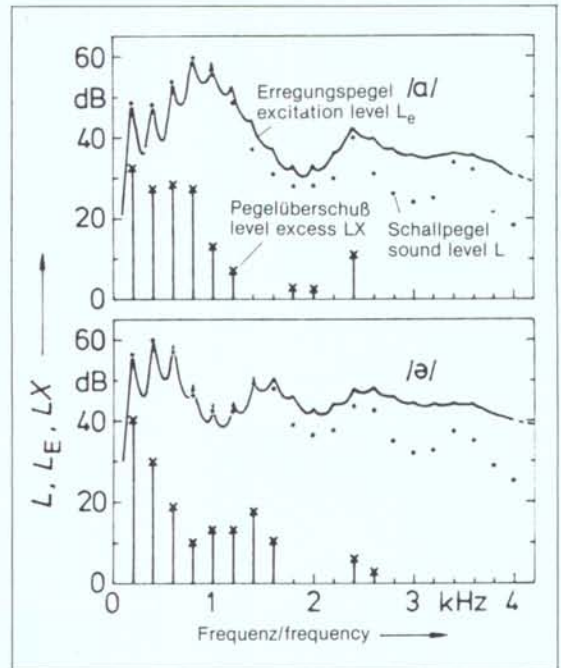


Abb. 3 Schallpegel L , Erregungspegel L_E und Pegelüberschuß LX der beiden Vokale /a/ und /e/. Stimmbandfrequenz jeweils 200 Hz

Fig. 3 Sound level L , excitation level L_E and level excess LX of the two vowels /a/ and /e/. Vocal cord frequency 200 Hz

favourable. This has the effect that the microstructure formed by the harmonic partial tones is blurred to the advantage of the formant structure. This type of spectral representation is admittedly no more than a compromise. It is not completely adequate for hearing because in auditory perception, the individual partial tones of the voiced speech sounds i.e. the spectral microstructure, do play a rôle. Generally it can be said that the tonal character of the voiced speech sounds relies to a certain extent on the peripheral analysis breaking these sounds down into their partial tones.

The perception of the spectral micro- and macrostructure can be illustrated by the so-called excitation level-frequency pattern and the spectral pitch pattern. Fig. 3 shows examples of these patterns together with the partial tone spectra of the two vowels /a/ and /e/ in which the vocal cord frequency amounts to 200 Hz. The excitation level L_E is a psycho-acoustic model magnitude derived from monitoring thresholds (see Zwicker & Feldtkeller [26]). Its curve, as a

lektive Verhalten des Gehörs mit großer Zuverlässigkeit wider. Die in Abb. 3 dargestellten Verläufe $L_E(f)$ der beiden Vokale wurden nach einem früher beschriebenen Verfahren aus den Teiltonspektren berechnet [19]. Sie zeigen neben der Formant- (also Makro-) Struktur im Frequenzbereich unterhalb etwa 2 kHz sehr deutlich die durch die einzelnen Harmonischen gebildete Mikrostruktur. Beim Vokal /ə/ sind übrigens die oben dafür genannten Formantfrequenzen von 500, 1500, 2500 und 3500 Hz gut erkennbar. Die scharfen Spitzen im Erregungsverlauf veranschaulichen, daß die entsprechenden Harmonischen in der Hörwahrnehmung eine Rolle spielen. In den meisten Fällen ist man sich allerdings nicht bewußt, einzelne Teiltöne eines stimmhaften Sprachlauts wahrzunehmen. Vielmehr hört man den Sprachklang als Ganzes und ordnet demselben unbewußt Empfindungsmerkmale höherer Ordnung zu, ohne sich der komplexen Verarbeitungsprozesse bewußt zu werden, welche dies ermöglichen. Dies ist als ganz natürliches Verhalten des Hörsystems anzusehen, denn die Sinnesorgane sind offensichtlich alle derart beschaffen, daß sie im »Normalmodus« eher die ganzheitliche Bedeutung eines Umweltreizes als dessen Detailstruktur zur bewußten Wahrnehmung bringen. Es bedarf im allgemeinen einer gewissen Aufmerksamkeit bzw. Anstrengung des Wahrnehmenden, um vom »normalen«, ganzheitlichen (synthetischen) Wahrnehmungsmodus auf den analytischen Modus (das heißt die Wahrnehmung der Mikrostruktur) überzugehen.

Man kann den analytischen Hörmodus unter geeigneten Bedingungen gut herstellen. Dabei gelingt es tatsächlich, die einzelnen Harmonischen in demselben Maße als reine Töne mit entsprechenden Tonhöhen (den sogenannten Spektraltonhöhen) bewußt wahrzunehmen, wie dies durch die Höhe der entsprechenden Spitzen im Erregungsverlauf (Abb. 3) angedeutet wird [12], [15], [19]. Um den Aspekt der Wahrnehmbarkeit der individuellen Spektraltonhöhen bequem und deutlich beschreiben zu können, wurde der sogenannte Pegelüberschuß der einzelnen Harmonischen herangezogen. Als Pegelüberschuß eines bestimmten Teiltönen wird die Differenz zwischen dem betreffenden Teiltonpegel und demjenigen Schallpegel, welcher der maskierenden Wirkung aller übrigen Teiltöne zusammen entspricht, bezeichnet, wobei zusätzlich die Hörschwelle Berücksichtigung findet [19], [20]. Stellt man die Pegelüberschüsse der einzelnen Harmonischen als Funktion ihrer Frequenz dar, so erhält man das sogenannte Spektraltonhöhenmuster, wie in Abb. 3 gezeigt. Es veranschaulicht, daß nur ein gewisser Teil der Harmonischen individuelle Spektraltonhöhen erzeugt. Je höher die Ordnungszahl der Harmonischen, umso

function of frequency, reflects with great fidelity the frequency-selective behaviour of hearing. The curves $L_E(f)$ of the two vowels represented in Fig. 3 were calculated from the partial tone spectra, in accordance with a method described earlier [19]. In addition to the formant (macro-) structure in the frequency range below roughly 2 kHz, they show very clearly the microstructure formed by the individual harmonics. In the case of the vowel /e/ the formant frequencies of 500, 1500, 2500 and 3500 Hz, indicated above, are readily discernible. The sharp peaks in the excitation curve show that the corresponding harmonics play a rôle in auditory perception. In most instances, however, we are not aware of hearing the individual partial tones of a voiced speech sound. What happens is that we hear the speech sound as a whole and that we subconsciously assign perception characteristics to it of a higher order without being consciously aware of the complex processing operations involved which make this possible. This can be regarded as a natural behaviour of the auditory system because the sensory organs are obviously constructed so as to normally transmit the total significance of an environmental stimulus to our conscious perception rather than its detail structure. Generally, it requires a certain amount of attention or effort on the part of the listener to change over from the "normal" synthetic mode of perception to the analytic mode of perception (i.e. perception of the microstructure).

Under suitable conditions the analytic hearing mode can be readily established. It is in fact possible to consciously perceive the individual harmonics to the same extent as pure tones, with their corresponding pitches (the so-called spectral pitches) as is suggested by the height of the corresponding peaks [12], [15], [19] (Fig. 3). In order to be able to describe the perceptibility aspect of the individual spectral pitches conveniently and clearly, recourse has been made to the so-called level excess of the individual harmonics. The level excess of a specific partial tone is defined as the difference between the respective partial tone level and that sound level which corresponds to the masking effect of all remaining partial tones together, while also making allowance for the hearing threshold [19], [20]. If the level excesses of the individual harmonics are represented as a function of their frequency, we obtain the so-called spectral pitch pattern, as shown in Fig. 3. It demonstrates that only a certain proportion of the harmonics generate their own individual spectral pitches. The higher the ordinal number of the harmonics, the lower the probability that they will appear with individual spectral pitches.

The spectral pitch pattern also reflects to a certain ex-

geringer die Wahrscheinlichkeit, daß sie überhaupt mit individuellen Spektraltonhöhen in Erscheinung treten.

Das Spektraltonhöhenmuster spiegelt bis zu einem gewissen Grade auch die Formantstruktur wider, denn naturgemäß erreichen diejenigen Harmonischen, welche mit Formanten zusammenfallen, besonders hohe Pegelüberschüsse. Abb. 4 zeigt am Beispiel des Vokals /a/, daß das Spektraltonhöhenmuster gegenüber Änderungen des Schallpegels recht stabil ist. Nur der Pegelüberschuß der ersten Harmonischen (des Grundtones) ändert sich deutlich, und zwar gleichsinnig mit dem Schallpegel (während infolge des Verdeckungseffekts bei anderen Harmonischen durchaus auch gegensinnige Änderungen vorkommen). Bei mittleren und höheren Pegeln (60 bzw. 80 dB) dominiert der Grundton über alle anderen Harmonischen. Dies veranschaulicht die Hörerfahrung, daß die »Klangfülle« bzw. »Sonorität« einer tiefen Stimme mit wachsendem Schallpegel deutlich zunimmt.

Die in Abb. 3 an Beispielen veranschaulichten Erregungspegelverläufe gestatten übrigens zutreffende Abschätzungen der eben wahrnehmbaren Formant- und Grundfrequenzänderungen. Nach dem von Zwicker und Feldtkeller [26] beschriebenen Prinzip, wonach die Änderung eines Schalles dann wahrnehmbar ist, wenn sie eine Änderung des zugehörigen Erregungspegelverlaufs bei irgend einer Frequenz von mindestens 1 dB bewirkt, ergeben sich eben wahrnehmbare Grundfrequenzänderungen von etwa 0,6% (bei allen Sprachgrundfrequenzen) und eben wahrnehmbare Formantfrequenzänderungen von 5% bis 20% (je nach Nummer und Frequenzlage des Formanten; vgl. [19]). Diese Werte sind in ausgezeichneter Übereinstimmung mit den Ergebnissen entsprechender Hörversuche.

Die hohe Frequenzselektivität, welche dem Gehör in dieser Weise die spektrale Mikrostruktur aufzulösen erlaubt, ist von erheblicher Bedeutung. Wie erwähnt, erstreckt sich der Bereich optimaler Frequenzauflösung auf Frequenzen unterhalb 2 kHz. Dies ist gerade derjenige Frequenzbereich, in welchem die beiden wichtigsten (das heißt, am meisten Information tragenden) Formanten der Vokale liegen. Daraus ergibt sich, daß das Gehör die unzweifelhaft fundamental wichtige Extraktion der Formantlage nicht nach dem einfachen Prinzip des Verwischens der Mikrostruktur zugunsten der Makrostruktur durchführt. Wie Abb. 1 zeigt, gestattet diese Methode zwar in vielen Fällen, die Formanten zu verdeutlichen, jedoch nur dann, wenn die Grundfrequenz tief genug und der Frequenzabstand der Formanten groß genug sind. In natürlicher Sprache sind diese beiden Voraussetzungen aber gerade für die ersten beiden Formanten im allgemeinen nur angenä-

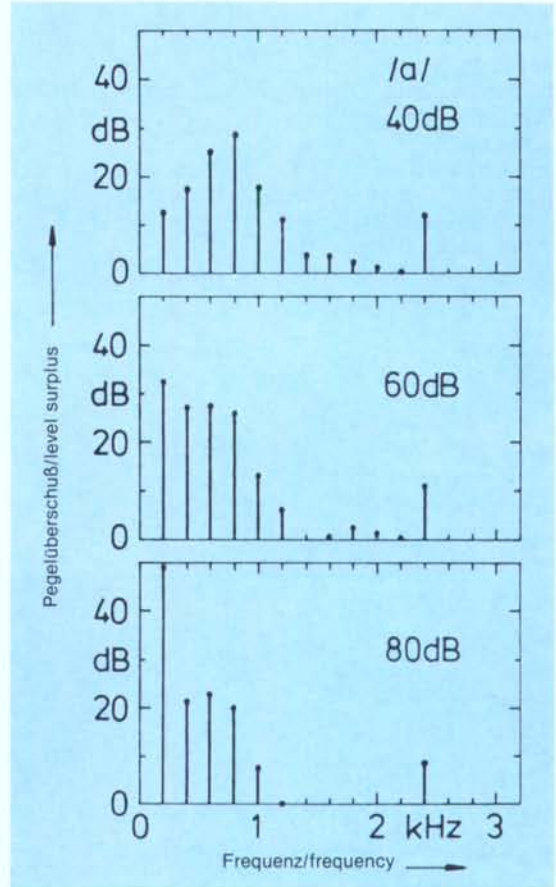


Abb. 4 Abhängigkeit des Pegelüberschuß-Musters vom Schallpegel beim Vokal /a/. Stimmbandfrequenz 200 Hz

Fig. 4 Dependency of the level excess pattern on the sound level in the case of vowel /a/. Vocal cord frequency 200 Hz

tent the formant structure because, for obvious reasons, those harmonics coinciding with formants reach particularly high level surpluses. Using the example of vowel /a/, Fig. 4 shows that the spectral pitch pattern is very resistant to changes in the sound level. Only the level excess of the first harmonic (the fundamental) alters significantly, i.e. in the same sense as the sound level (whereas, owing to the masking effect, with other harmonics, changes in the opposite sense can also occur). At medium and high levels (60 and 80 dB respectively) the keynote dominates all other harmonics. This illustrates the auditory experience that the

hert erfüllt. So sind beispielsweise im Spektrogramm (Abb. 1) die ersten beiden Formanten der Laute /o/ und /u/ jeweils weitgehend miteinander verschmolzen. Das Gehör leistet sich offensichtlich nicht den Luxus, unter derart erschwerten Bedingungen auch noch durch zu grobe Spektralanalyse Information zu »verschwenken«. Vielmehr scheint das auditive System darauf angelegt zu sein, aus dem Sprachsignal durch Auswertung sowohl der spektralen Makrostruktur (hier repräsentiert durch die Grobstruktur des Erregungspegelverlaufs) als auch der Mikrostruktur (hier repräsentiert durch das Spektraltonhöhenmuster) so viel Information wie physikalisch überhaupt möglich zu gewinnen. Brauchbare Modelle für diesen Prozeß gibt es allerdings bisher kaum. Für die technische Sprachsignalverarbeitung wären sie von großem Nutzen.

Fletcher und Galt [9] haben auf Grund von Silbenverständlichkeitsmessungen eine realistische Abschätzung der »Wichtigkeit der verschiedenen Spektralgebiete« durchgeführt, das heißt, der relativen Beiträge zur Silbenverständlichkeit, welche von schmalen Frequenzbändern in Abhängigkeit von deren Lage auf der Frequenzachse geleistet werden. Dabei ergab sich die in Abb. 5 dargestellte Gewichtsfunktion. Sie veranschaulicht, daß das zentrale Gehör sich bei der Sprachwahrnehmung auf den bereits mehrfach erwähnten Spektralbereich unterhalb 2 kHz konzentriert, mit einem Maximum des Gewichts bei 700 Hz. Diese Frequenz markiert recht genau das Zentrum des von den beiden wichtigsten Formanten überstrichenen Frequenzbereichs, so daß hier eine Anpassung des zentralen Gehörs an wichtige Sprachsignaleigenschaften vorzuliegen scheint. Darüber hinaus ist der durch die Gewichtsfunktion als »dominant« ausgewiesene Frequenzbereich, der sich ungefähr von 300 bis 2000 Hz

»tone volume« or »sonority« of a low voice unmistakably increases with growing sound level.

The excitation level curves based on examples which are shown in Fig. 3, moreover permit true estimates of the just perceptible changes in the formant and fundamental frequencies. From the principle described by Zwicker and Feldtkeller [26], according to which the change in a sound can be perceived, if there is an alteration in the respective excitation level curve of at least 1 dB at any frequency, just perceptible changes of roughly 0.6% occur in the fundamental frequency (at all fundamental speech frequencies) and just perceptible changes of 5% to 20% in the formant frequencies (depending on the number and frequency level of the formants; see [19]). These values tally perfectly with the results of the respective hearing tests.

Of considerable importance is the high frequency selectivity which enables the hearing to resolve the spectral microstructure in this manner. As mentioned earlier, the range of optimum frequency resolution extends to frequencies below 2 kHz. This is exactly the frequency range in which the two most important vowel formants (i.e. those carrying most information) fall. This means that the hearing does not effect the extraction of the formant position which is undoubtedly of fundamental importance, according to the simple principle of blurring the microstructure in favour of the macrostructure. As made clear in Fig. 1, in many instances, this method does allow the formants to be represented but only if the fundamental frequency is low enough and the frequency spacing of the formants is large enough. However, in natural speech, these two prerequisites are in general only approximately fulfilled, specifically for the first two formants. Thus, for

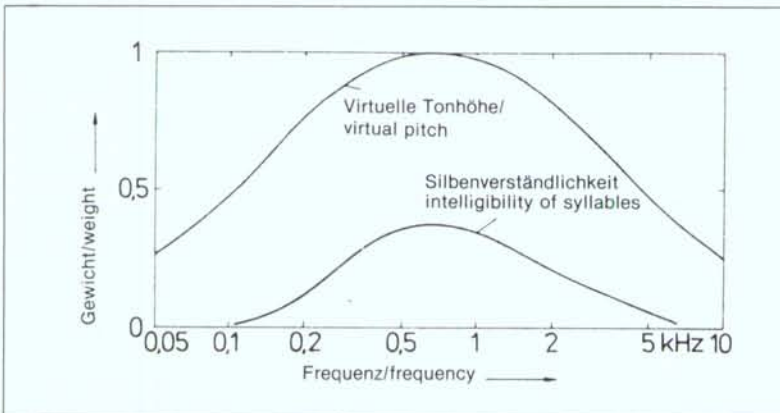


Abb. 5 Wichtigkeit der verschiedenen Spektralbereiche bei der Silbenverständlichkeit (nach Fletcher und Galt [9]) und bei der Tonhöhenwahrnehmung (nach Terhardt et al. [21])

Fig. 5 Importance of the various spectral ranges for the intelligibility of syllables (Fletcher & Galt [9]) and for pitch perception (Terhardt et al. [21])

erstreckt, gerade auch derjenige, in welchem das periphere Gehör optimal arbeitet: Die Normalhörschwelle liegt in diesem Bereich praktisch auf ihrem bestmöglichen Niveau und das Frequenzauflösungsvermögen ist am größten. Es scheint also zugleich eine Anpassung des zentralen Gehörs an das periphere Gehör vorzuliegen.

Die vom Gehör durchgeführte Analyse der von den harmonischen Teiltönen gebildeten spektralen Mikrostruktur ist darüber hinaus von fundamentaler Bedeutung für alle Hörwahrnehmungen, die durch das Stichwort »tonal« angedeutet werden. Dazu gehört zunächst die Wahrnehmung der Stimmhaftigkeit von Sprachlauten, das heißt, die Unterscheidung der phonetischen Merkmale »stimmhaft« und »stimmlos«. Sie wird ermöglicht durch Wahrnehmung der Tonhöhe (oder auch, im analytischen Wahrnehmungsmodus, mehrerer Tonhöhen). Ein Sprachlaut, der tonale Hörempfindungen hervorruft, ist eben dadurch als stimmhaft gekennzeichnet.

Die von stimmhaften Sprachlauten hervorgerufene Tonhöhenwahrnehmung ist ein eindrucksvolles Beispiel auditiver Gestaltwahrnehmung. Bekanntlich entspricht diese Tonhöhe der Stimmbandfrequenz. »Entspricht« heißt hier, daß ein Sinuston, dessen empfundene Tonhöhe man durch Frequenzänderung mit der Tonhöhe des betrachteten Sprachlauts in Übereinstimmung gebracht hat, dieselbe Frequenz aufweist wie die Stimmbandschwingung bzw. die erste Harmonische des Teiltonspektrums. Die Frage, warum die Tonhöhenempfindung in diesem Sinne gerade der ersten Harmonischen entspricht und nicht irgend einer anderen, und warum sie nicht — entsprechend der großen Anzahl hörbarer Harmonischer — weitgehend unbestimmt ist, konnte durch die Psychoakustik erst im Verlauf der vergangenen 10 Jahre befriedigend beantwortet werden. Die Antwort lautet, daß die Tonhöhe eines komplexen Schallsignals, wie es ein stimmhafter Sprachlaut ist, ein ganzheitliches, aus vorbewußt existierenden elementaren Wahrnehmungsdetails abgeleitetes Attribut ist. Sie entsteht also im oben als natürlich bezeichneten ganzheitlichen (synthetischen) Wahrnehmungsmodus, ebenso wie beispielsweise die wahrgenommenen Vokalqualitäten. Dies bedeutet, daß die Tonhöhenempfindung zwar, wie erwähnt, der Stimmbandfrequenz bzw. der Frequenz der ersten Harmonischen entspricht, aber keineswegs durch einen dieser beiden physikalischen Parameter direkt bestimmt sein muß. Weder die zeitliche Periodizität, welche das Sprachsignal infolge der Stimmbandschwingung aufweist, noch die erste Harmonische des Frequenzspektrums sind für die Tonhöhenempfindung des Sprachlauts notwendige physikalische Merkmale. Vielmehr

example, in spectrogram (Fig. 1) the first two formants of the phonemes /o/ and /u/ are largely merged together. The hearing, it appears, cannot afford to give away information, in such difficult conditions, by an overly coarse spectral analysis. It seems much more that the auditory system is designed to extract as much information as is physically possible from the speech signal by evaluating both the spectral macrostructure (represented by the coarse structure of the excitation level curve) and the microstructure (represented here by the spectral pitch pattern). So far, there are indeed scarcely any usable models for this process. They would be a great advantage for technical speech signal processing.

Fletcher and Galt [9] carried out a realistic evaluation of the "importance of the various spectral areas" based on measurements of syllabic intelligibility, i.e. they assessed the relative contribution to syllabic intelligibility made by narrow frequency bands depending on their location on the frequency axis. This resulted in the weight function in Fig. 5 which shows that the central hearing, in its speech perception function, concentrates on the repeatedly mentioned spectral range below 2 kHz with a weight maximum at 700 Hz. This frequency marks very accurately the center of the frequency range covered by the two most essential formants so providing evidence that the central hearing appears to adapt to important speech signal properties. In addition, the frequency range extending roughly from 300 to 2,000 Hz, indentified by the weight function as "dominant" is exactly the one within which the peripheral hearing functions optimally: the normal hearing threshold lies in this range practically at its optimum level and the frequency resolving power is at its maximum. Thus, there also seems to be a simultaneous adaptation of the central hearing to the peripheral hearing.

The auditory analysis of the spectral microstructure formed by the harmonic partial tones is, moreover, of fundamental importance for all auditive perceptions which come under the heading "tonal". This encompasses first of all the perception of the voiced quality of speech sounds, i.e. the distinction between the phonetic characteristic "voiced" and "unvoiced". This distinction is due to the perception of the pitch (or even of several pitches in the analytical perception mode). A speech sound which provokes tonal auditory sensations, is thereby designated as voiced.

The pitch perception provoked by voiced speech sounds represents an impressive example of auditory *Gestalt* perception. As is known, this pitch corresponds to the vocal cord frequency. "Correspond" here means that a sine wave tone whose sensed pitch

sind dies die höheren Harmonischen des Spektrums, und zwar bemerkenswerterweise gerade diejenigen, die im oben erwähnten dominanten Spektralbereich liegen. Weil also im allgemeinen die Tonhöhenempfindung komplexer Signale nicht direkt mit einer bestimmten Spektralkomponente zusammenhängt, wird diese Art der Tonhöhe als virtuelle Tonhöhe bezeichnet [16]. Im Gegensatz dazu steht die bereits erwähnte Spektraltonhöhe, welche stets einer bestimmten Spektralkomponente zugeordnet werden kann und von deren Vorhandensein abhängt. Der Spektralbereich, der vom zentralen Gehör normalerweise bei der Wahrnehmung der virtuellen Tonhöhe bevorzugt wird, ist mit dem in Bezug auf die Silbenverständlichkeit dominanten Bereich identisch. In Abb. 5 wurde zur Veranschaulichung eine Gewichtsfunktion eingezeichnet, welche nach einer Theorie der Tonhöhenwahrnehmung die Beiträge der verschiedenen Spektralbereiche zur virtuellen Tonhöhe beschreibt (Terhardt et al. [21]). Auch diese Gewichtsfunktion hat ihr Maximum bei 700 Hz. Sie macht besonders deutlich, daß die der Stimmbandfrequenz entsprechende virtuelle Tonhöhe im allgemeinen durch Spektralkomponenten bestimmt wird, die in einem deutlich höheren Frequenzbereich liegen (Wie erwähnt, bewegen sich die Stimmbandfrequenzen natürlicher Sprache zumeist im Bereich unterhalb 300 Hz, also unterhalb des nach Abb. 5 dominanten Bereichs). Dies ist der Grund dafür, daß man hinsichtlich der meisten für die Sprach- und Musikübertragung wichtigen Merkmale (einschließlich der Tonhöhe) auf den Frequenzbereich unterhalb etwa 300 Hz ohne weiteres verzichten kann. In der Tat findet ein solcher »Verzicht« bei vielen technischen Geräten und Einrichtungen (Fernsprecher; kleine Radio- und Magnetbandgeräte; Wechselsprechanlagen) statt.

Die Wahrnehmung stimmhafter, stationärer Sprachlaute hat man sich demnach als einen Prozeß vorzustellen, in welchem durch aktive Verarbeitung der spektralen Makro- und Mikrostruktur einerseits die lauttypischen Klangfarbenmerkmale, andererseits die virtuelle Tonhöhe gebildet werden. Die für das Funktionieren dieses Verarbeitungsprozesses erforderlichen Systemparameter eignet sich das Hörsystem wahrscheinlich in der frühesten Lebensphase, das heißt in derjenigen des Spracherwerbs an. Diese Schlußfolgerung hat weitreichende Konsequenzen für einen ganz anderen Bereich der akustischen Kommunikation, nämlich die Musik. Für diese sind bestimmte, mit dem Begriff »tonal« verbundene Hörempfindungen, nämlich die Tonverwandtschaft der musikalischen Grundintervalle Oktave, Quinte, Quarte, Terz von fundamentaler Bedeutung. Diese Intervalle werden natürlicherweise von den ersten sechs Harmonischen eines jeden periodischen

has been brought into agreement with the pitch of the speech sound concerned by changing the frequency, has the same frequency as the vocal cord vibration, or the first harmonic of the partial tone spectrum. The question why the pitch sensation in this sense corresponds to just the first harmonic and not to any other, and why it is not largely indeterminate, corresponding to the large number of audible harmonics, has only been answered satisfactorily by psycho-acoustics in the course of the last ten years. The answer is that the pitch of a complex sound signal such as a voiced speech sound is a synthetic attribute consisting of pre-consciously existing elementary perception details. Hence, it develops in the synthetic perception mode designated above as natural, just as, for example, the vocal qualities perceived. This means, indeed, that the pitch sensation corresponds, as already mentioned, to the vocal cord frequency or the frequency of the first harmonic, but must by no means be directly determined by one of these two physical parameters. Neither the temporal periodicity of the speech signal due to the vocal cord vibration, nor the first harmonic of the frequency spectrum are necessary physical characteristics for the pitch perception of the speech sound. They are rather the higher harmonics of the spectrum and particularly those lying in the above-mentioned dominant spectral range. Since, generally, the pitch sensation of complex signals is not directly related to a certain spectral component, this type of pitch is called virtual pitch [16]. In contrast to this stands the spectral pitch which can always be assigned to a certain spectral component and which depends on the existence of this component. The spectral range which is normally preferred by the central hearing in the perception of the virtual pitch, is identical to the dominant range with regard to syllabic intelligibility. To illustrate this, a weight function was plotted in Fig. 5, which according to a theory of the pitch perception describes the contributions to the virtual pitch made by the various spectral ranges (Terhardt et al. [21]). This weight function has its maximum also at 700 Hz. It makes it particularly plain that the virtual pitch corresponding to the vocal cord frequency is in general determined by spectral components which lie in a distinctly higher frequency range. (As already mentioned, the vocal cord frequencies of natural speech are usually in the range below 300 Hz, i.e. below the dominant range according to Fig. 5). This is the reason why the frequency range below about 300 Hz can be dispensed with regard to most characteristics (including the pitch) which are important for the transmission of speech and music. Indeed, many technical devices and facilities (telephone, small radio-sets and magnetic tape recorders, intercom systems) dispense with this frequency range.

Signals gebildet. Da die stimmhaften Sprachlaute periodische Signale sind, stellen sie die natürlichste Quelle der genannten Tonintervalle dar. Dieselben sind ursprünglich nichts anderes als typische, stets wiederkehrende spektrale Strukturmerkmale der stimmhaften Sprachlaute. Sie prägen sich als solche wahrscheinlich dem zentralen Gehör in der Spracherwerbsphase ein und spielen insbesondere bei der Bildung der virtuellen Tonhöhe (die ihrerseits primär ein Bestandteil der Sprachwahrnehmung ist) eine wichtige Rolle als »Systemparameter«. Die für die Entwicklung der tonalen Musik grundlegende Empfindung der Tonverwandtschaften (und der Harmonieempfindung, vgl. [18]) findet demnach in den angedeuteten Prinzipien der Hörwahrnehmung und ihren Wechselbeziehungen zum Sprachsignal eine einfache Erklärung.

4. Merkmale des Frequenz-Zeitmusters

Das Sprechen kann man als »Gestik« der Artikulationsorgane bezeichnen. Die Artikulationsvorgänge stellen somit die primäre (das heißt, am wenigsten durch irgendwelche störende Einflüsse beeinträchtigte) Quelle der Sprachinformation dar. Das dabei entstehende Schallsignal spiegelt diese ursprünglichen Vorgänge nur mit eingeschränkter Genauigkeit wider. Das Gehör des Angesprochenen ist auf das Schallsignal angewiesen. Sprachwahrnehmung besteht also zu einem erheblichen Teil darin, aus dem Schallsignal die demselben zugrunde liegenden Artikulationsvorgänge zu rekonstruieren. Die Artikulation ist ein mechanisch-akustischer Vorgang (der physiologische Aspekt wird hier außer Acht gelassen) und als solcher kontinuierlich. Dies bedeutet, daß lauttypische, aufeinanderfolgende Artikulationskonfigurationen nicht übergangslos aneinandergereiht sein können. Die Artikulationsabläufe können durch eine Folge diskreter Schallsignalmerkmale, wie zum Beispiel die zuvor betrachteten einzelnen Lautspektren, nicht vollständig beschrieben werden. Vielmehr müssen zwischen solchen quasi diskreten Signalmerkmalen bestimmte kausale Zusammenhänge bestehen und kontinuierliche Übergänge vorhanden sein, die einen beträchtlichen Teil der Gesamtinformation tragen. Dies sei am Beispiel des dem Plosivlaut /k/ vorausgehenden Vokals /e/ im Spektrogramm (Abb. 1) erläutert. Es ist deutlich zu erkennen, daß die beiden ersten Formanten des /e/ während der ganzen Dauer des Vokals aufwärts gleiten. Dies wird dadurch hervorgerufen, daß schon während der Dauer des /e/ der Zungenrücken dem Gaumen angenähert und der Vokal durch völligen Verschluss des Sprechtrakts an dieser Stelle beendet wird, weil dies für die Artikulation des /k/ nötig ist. Im Spektrogramm wird der Laut /k/ also in der Tat lediglich indirekt repräsentiert.

The perception of voiced, stationary speech sounds can therefore be imagined as a process in which by active processing of the spectral macro and microstructure, both the typical timbre characteristics of the sound and the virtual pitch are formed. Probably as early as the first phase of life, i.e. the period of speech acquisition, the hearing acquires the system parameters required for the functioning of this processing operation. This conclusion greatly influences another area of acoustic communication, namely music. Certain auditory sensations associated with the conception "tonal", that is to say tone affinity of the basic musical intervals octave, fifth, fourth, third, are of fundamental importance. These intervals are obviously formed by the first six harmonics of each periodical signal. Since the voiced speech sounds are periodic signals, they represent the most natural source of the tone intervals mentioned. Originally, these are no more than typical, constantly recurring spectral structural features of the voiced speech sounds. As such they are likely to impress themselves on the central hearing in the phase of speech acquisition and play an important role as system parameters in the formation of the virtual pitch (which in turn is primarily a component of speech perception). The perception of tone affinities (and the sense of harmony, see [18]) which is of fundamental importance in the development of the tonal music, therefore finds a simple explanation in the principles of auditory perception and their interrelationships with the speech signal.

4. Characteristics of the frequency/time pattern

Speaking can be called "gesticulation" of the organs of articulation. The articulatory processes thus represent the primary source of speech information (i.e. the one least affected by disturbing influences). The resulting speech signal reflects these primary processes only to a limited degree of accuracy. The hearing of the listener depends on the sound signal. Perception of speech consists to a considerable extent in the reconstruction of the articulation processes underlying the speech signal. Articulation is a mechanical acoustic process (the physiological aspect is here left out of account) and as such *continuous*. This means that sound typical successive configurations of articulation cannot follow one another without transitions. The articulation processes cannot be completely described by a sequence of discrete sound signal characteristics as is the case with the individual sound spectra dealt with earlier. It is rather that between such quasi discrete signal characteristics certain causal next and continuous transitions must exist which convey a considerable proportion of

tiert: Durch die geschilderten Formantübergänge des vorausgehenden /e/ wird die Artikulationsstelle angezeigt und durch die nachfolgende Pause wird signalisiert, daß es sich um einen Plosivlaut handelt. Da der Artikulationsort der beiden Plosivlaute /k/ und /g/ derselbe ist, reichen die beiden genannten Merkmale nicht aus, diese beiden Laute voneinander zu unterscheiden. Es gibt jedoch weitere Merkmale des Frequenz-Zeitmusters, die dies in vielen Fällen ermöglichen.

Lieberman [11] und seine Kollegen haben als erste versucht, die wichtigen Merkmale des Frequenz-Zeitmusters systematisch zu erfassen und ihre Bedeutung zu erforschen. Abb. 6 gibt als Beispiel einige schematisierte, auf die interessierenden Merkmale reduzierte Spektrogramme von Konsonant-Vokalfolgen wieder, wobei der Vokal stets /a/ ist. Als Abszisse hat man sich also auch in diesen Schemata der Frequenz-Zeitmuster die Zeit, als Ordinate die Frequenz vorzustellen. Die horizontal punktierten bzw. gestrichelten Linien deuten das Auftreten des Grundtones der Stimmbandschwingung bzw. höherer Harmonischer im Bereich der stärksten Formanten an. Die starken Linien repräsentieren die Zeitverläufe des ersten und zweiten Formanten des Vokals /a/. Die vertikal punktierten Linien bedeuten Geräuschanteile. Man erkennt, daß die einander entsprechenden, das heißt, am selben Ort artikulierten Plosivlaute /p,b/, /t,d/, /k,g/ praktisch übereinstimmende Formantübergänge zeigen, wie dies auf Grund der übereinstimmenden Artikulation auch zu erwarten ist. Sie unterscheiden sich durch zusätzliche Frequenz-Zeitmerkmale: Die weichen Plosiv-

laute /k/ ist nur indirekt: The point of articulation is indicated by the formant transitions of the preceding /e/, and the subsequent interval signals that a plosive is involved. As the site of articulation for the two plosives /k/ and /g/ is identical, the two characteristics mentioned do not suffice to distinguish between the two sounds. However, there are other characteristics of the frequency/time pattern which make this possible in many instances.

Lieberman [11] and his colleagues were the first to try to determine the important characteristics of the frequency/time pattern systematically and to investigate their importance. In Fig. 6 are some examples of schematic spectrograms showing consonant-vowel sequences, reduced to the characteristics of interest; /a/ being the vowel in each case. Also, in this schematic representations of the frequency/time pattern, time is represented by the abscissa and frequency by the ordinate. The horizontally dotted or dashed lines indicate the occurrence of the basic tone of vocal cord vibration or of the higher harmonics in the range of the predominant formants. The bold lines represent the time curves of

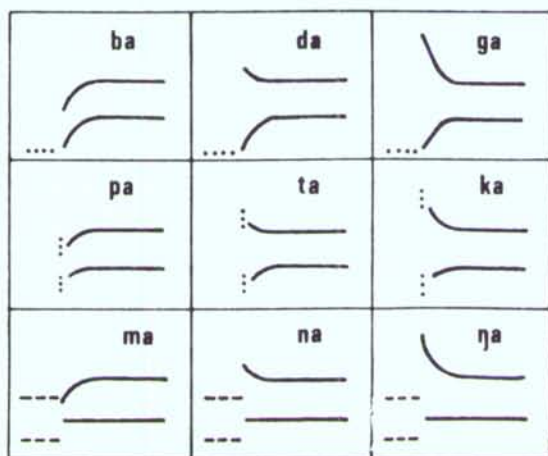


Abb. 6 Schematische Spektrogramme als Veranschaulichung wichtiger Merkmale des Frequenz-Zeitmusters (nach Liberman [11]). Horizontal: Zeit; vertikal: Frequenz. Starke Linien: Verläufe der ersten beiden Formanten. Horizontal punktierte Linien: Stimmhaftigkeit bei weichen Plosivlauten. Vertikal punktierte Linien: Geräuschanteile bei harten Plosivlauten. Horizontal gestrichelte Linien: Formanten der Nasallaute.

Fig. 6 Schematic spectrograms illustrating important characteristics of the frequency/time pattern (Lieberman [11]). Horizontal: time; Vertical: frequency. Bold lines: curves of the first two formants. Horizontal dotted lines: Voicedness with soft plosives. Vertical dotted lines: noise component with hard plosives. Horizontal dashed lines: formants of the nasal sounds

laute sind stimmhaft, während die harten stimmlos und mit Geräuschanteilen (Reibegeräusch beim Öffnen des Sprechtraktverschlusses) versehen sind. Die Nasallaute /m, n, ŋ/ unterscheiden sich hinsichtlich ihrer Spektren so gut wie gar nicht voneinander, wohl aber im Zeitverlauf der zweiten Formantfrequenz des nachfolgenden Vokals /a/.

Das Gehör ist nicht nur auf Grund der Analyseleistungen seines peripheren Teils in der Lage, solche Frequenz-Zeitmuster zu detektieren, sondern wertet die darin enthaltene Information mit Sicherheit auch aus. Man kann dies beispielsweise daran erkennen, daß die Lautwahrnehmung völlig umschlagen kann, wenn einzelne Merkmale des Frequenz-Zeitmusters willkürlich geändert oder unterdrückt werden, ohne den zum Laut selbst gehörenden Signalanteil zu verändern [11]. Beispielsweise hört man anstelle von /ja/ die Lautfolge /ga/, wenn von /ja/ das j ausgeblendet wird.

Über die funktionellen Aspekte der Wahrnehmung von Frequenz-Zeitmustern ist bisher wenig bekannt. Die Untersuchung dieser Vorgänge ist schwierig, weil sie weitgehend unbewußt ablaufen. Man kann sie als typische Beispiele ganzheitlicher Wahrnehmung verständlich zu machen suchen in demselben Sinne, wie dies im Zusammenhang mit den Spektralstrukturen erörtert wurde. Während allerdings die Spektralmuster eindimensional sind (Intensitätsverteilung längs der Frequenzachse), tritt bei den Frequenz-Zeitmustern die Zeit als zweite Dimension hinzu. Es ist kaum vorstellbar, daß die Wahrnehmungsalgorithmen, nach denen Frequenz-Zeitmuster vom Gehör verarbeitet werden, nicht an der Sprache orientiert beziehungsweise durch dieselbe geprägt sein könnten. Es ist oft die Vermutung ausgesprochen worden, daß die Wahrnehmung von Sprache nach ganz anderen Gesetzen ablaufe als diejenige nichtsprachlicher Schallsignale, daß also ein »sprachspezifischer Wahrnehmungskanal« existiere. Diese Vermutung erscheint in dem Sinne einleuchtend, als das Gehör – im Rahmen der durch die analytischen Möglichkeiten seines peripheren Teils vorgegebenen Grenzen – sprachbezogene Algorithmen der zentralen Verarbeitung entwickelt haben dürfte, die den raschen und unbewußten Ablauf komplexer Erkennungsvorgänge ermöglichen. Ein derartiger ganzheitlicher Wahrnehmungsmodus unterscheidet sich in der Tat fundamental von einem neutralen, das heißt im wesentlichen analytischen Modus. Es scheint, daß ganzheitliche Hörprozesse im wesentlichen *sprachbezogene* Hörprozesse sind und daß diese spätestens dann nahezu unvermeidlich wirksam werden, wenn ein Schallsignal typische Merkmale des *Sprachsignals* aufweist, beispielsweise kontinuierlich gleitende Grundfrequenz und kontinuierliche, von der Grundfrequenz

the first and second formants of the vowel /a/. The vertical dotted lines denote the noise components. One can see that the plosives /p,b/, /t,d/, /k,g/, which correspond to one another and which are articulated at the same site, show practically conforming formant transitions, as is to be expected, in view of the identical articulation. They are distinguished by additional frequency/time characteristics: the soft plosives are voiced, while the hard are unvoiced and have noise components (a rubbing noise when the vocal tract closure opens). The nasal sounds /m,n,ŋg/ are practically indistinguishable with regard to their spectra but the second formant frequency time curve of the succeeding vowel /a/ provides some distinction.

The hearing is able not only to detect such frequency/time patterns because of the analytic ability of its peripheral section but also to reliably evaluate the information contained in them. It is possible to see this in the phenomenon that the perception of sound can be completely upset, if some characteristics of the frequency/time pattern are deliberately changed or suppressed, without changing the signal component belonging to the sound itself [11]. For example, instead of hearing /ja/, the vocal sequence /ga/ is heard, if j is faded out from /ja/.

As yet, little is known about the functional aspects associated with the perception of frequency/time patterns. It has proved difficult to investigate these processes because they largely take place unconsciously. One can try to interpret them as typical examples of synthetic perception in much the same way as the connection with the spectral structures was explained. However, whereas the spectral patterns are unidimensional (distribution of intensity along the frequency axis), time appears as a second dimension in the frequency/time pattern. It is hardly imaginable that the perception algorithms by which the frequency/time patterns are processed by our hearing, are not speech-oriented or are not coined by speech. Frequently it was conjectured that the perception of the speech is governed by completely different laws to those which apply to non-verbal sound signals, meaning that a "speech-specific perception channel" exists. This conjecture appears plausible insofar as the hearing – within the limits set by the analytical possibilities of its peripheral section – might have developed speech-related algorithms in the central processing system which permit the rapid and unconscious functioning of complex identification processes. Such a synthetic mode of perception indeed differs fundamentally from a neutral, i.e. basically analytical mode. It seems that synthetic hearing processes are mainly speech-related and these almost unavoidably become effective at the latest,

unabhängige Änderungen der spektralen Makrostruktur.

Zusammenfassend kann man feststellen, daß offenbar ein hohes Maß wechselseitiger Anpassung und Beeinflussung von Sprechvorgang, Sprachsignal und Hörwahrnehmung besteht. Die vorstehend diskutierten Ergebnisse deuten insbesondere darauf hin, daß das zentrale Gehör stark von der sprachlichen Kommunikation beeinflusst ist, so daß selbst solche Wahrnehmungsmerkmale, die scheinbar mit der Sprache nichts zu tun haben wie die Empfindung der musikalischen Tonverwandtschaften, ihren Ursprung letztlich in der sprachlichen Kommunikation haben.

Herrn Prof. Dr.-Ing. E. Zwicker und Herrn Dr.-Ing. H. Fastl sei für nützliche Hinweise und Anregungen gedankt. Dieser Beitrag entstand im Sonderforschungsbereich 50 »Kybernetik«, gefördert durch die Deutsche Forschungsgemeinschaft.

whenever a sound signal has the typical characteristics of a speech signal, such as a continuously varying fundamental frequency and continuous changes in the spectral macro structure, independent of the fundamental frequency.

Summarizing, it can be said that a high degree of mutual adaptation and interaction appear to exist between articulation, speech signal and auditory perception. The results discussed strongly suggest that the central hearing is greatly influenced by linguistic communication, so that even such characteristics of perception which are seemingly unrelated to speech, such as the perception of musical tone affinities, ultimately have their origin in speech communication.

My thanks go to Prof. Dr.-Ing. E. Zwicker and Dr.-Ing. H. Fastl for useful hints and suggestions. This contribution has been sponsored by the Deutsche Forschungsgemeinschaft within the Sonderforschungsbereich 50 "Kybernetik".

Literatur/References

- [1] Békésy, G. von: Experiments in Hearing. McGraw-Hill, New York (1960).
- [2] Burghardt, H.: Über die subjektive Dauer von Schallimpulsen und Schallpausen. *Acustica* 28, 284-290 (1973).
- [3] Cole, R., Rudnický, A., Zue, V., and Reddy, D. R.: Speech as patterns on paper. In "Perception and Production of Fluent Speech" (R. Cole, Edit.). Hillsdale, NJ, Erlbaum Assoc. (1978).
- [4] Fant, G.: Acoustic Theory of Speech Production. Mouton, Den Haag (2. Aufl. 1970).
- [5] Fastl, H.: Mithörschwelle und Subjektive Dauer. *Acustica* 32, 288-290 (1975).
- [6] Fastl, H.: Temporal masking effects: I. Broad band noise masker. *Acustica* 35, 287-302 (1976).
- [7] Fastl, H.: Temporal masking effects: II. Critical band noise masker. *Acustica* 36, 317-331 (1977).
- [8] Flanagan, J. L.: *Speech Analysis, Synthesis and Perception*. Springer, Heidelberg (1972).
- [9] Fletcher, H., and Galt, R. H.: The perception of speech and its relation to telephony. *J. Acoust. Soc. Am.* 22, 89-151 (1950).
- [10] Fujimura, O.: Syllable as a unit of speech recognition. *IEEE/ASSP-23*, 82-87 (1975).
- [11] Liberman, A. M.: Some results of research on speech perception. *J. Acoust. Soc. Am.* 29, 117-123 (1957).
- [12] Plomp, R.: The ear as a frequency analyzer. *J. Acoust. Soc. Am.* 36, 1628-1636 (1964).
- [13] Port, E.: Die Lautstärke von Tonimpulsen verschiedener Dauer. *Frequenz* 13, 242-245 (1959).
- [14] Ruske, G. und Schotola, T.: An approach to speech recognition using syllabic decision units. *IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Tulsa OK, 722-725 (1978).
- [15] Stoll, G.: Psychoakustische Messungen der Spektraltonhöhenmuster von Vokalen. In »Fortschritte der Akustik« (DAGA '80), VDE-Verlag, Berlin, 631-634 (1980).
- [16] Terhardt, E.: Zur Tonhöhenwahrnehmung von Klängen. *Acustica* 26, 173-199 (1972).
- [17] Terhardt, E.: On the perception of periodic sound fluctuations (roughness). *Acustica* 30, 201-213 (1974).
- [18] Terhardt, E.: Ein psychoakustisch begründetes Konzept der Musikalischen Konsonanz. *Acustica* 36, 121-137 (1976).
- [19] Terhardt, E.: On the perception of spectral information in speech. In "Hearing Mechanisms and Speech" (O. Creutzfeldt, H. Scheich, C. Schreiner, eds.), Springer, Berlin, 281-291 (1979).
- [20] Terhardt, E.: Calculating virtual pitch. *Hearing Research* 1, 155-182 (1979).
- [21] Terhardt, E., Stoll, G., and Seewann, M.: Quantitative evaluation of pitch and pitch-salience for complex tonal signals. In Vorbereitung.
- [22] Zwicker, E.: Ein Beitrag zur Lautstärkemessung impulshaltiger Schalle. *Acustica* 17, 22-25 (1966).
- [23] Zwicker, E.: Subjektive und objektive Dauer von Schallimpulsen und Schallpausen. *Acustica* 22, 214-218 (1970).
- [24] Zwicker, E.: Zusammenhänge zwischen neueren Ergebnissen der Psychoakustik. In »Akustik und Schwingungstechnik«, VDI-Verlag, Düsseldorf, 9-21 (1971).
- [25] Zwicker, E.: Die Zeitkonstanten (Grenzdauern) des Gehörs. *Z. Hörger. Akustik* 13, 82-102 (1974).
- [26] Zwicker, E. und Feldtkeller, R.: Das Ohr als Nachrichtenempfänger. Hirzel, Stuttgart (2. Aufl. 1967).

Aktuelles

8. Kolloquium audiologisch tätiger Physiker und Ingenieure

Am 26. und 27. März 1981 fand in Erlangen zum achten Mal das Kolloquium audiologisch tätiger Physiker und Ingenieure statt, ein Kreis, der von Mitarbeitern sowohl aus Kliniken als auch aus der Industrie aufgesucht wird und der von der Firma Siemens als Mäzen und Katalysator in achtenswerter und völlig neutraler Weise initiiert ist. Der in solchem Rahmen auch wohl-erfüllbare Sinn des Kolloquiums ist der, im kleinen Kreis von etwa drei Dutzend Technikern einigermaßen repräsentativen Querschnitts aktuelle persönliche Arbeitsrichtungen einander näherzubringen — bei den häufigen 1 Mann-Elfenbeintürmen in der Audiologie eine durchaus rühmliche Aufgabe.

In diesem Jahr war das Kolloquium gleichzeitig eine Art wissenschaftlicher Würdigung und Dank einer 31 Jahre währenden Tätigkeit von Herrn Diplom-Physiker Clemens *Starke* auf dem Hörgerätesektor, der nunmehr zu Mitte dieses Jahres auf einen ergebnisreichen Abschluß seiner Arbeit zurückblicken kann.

Es war auch dies ein Grund, daß die Vortragsreihe von seinem Kollegen und Gast aus den USA eröffnet wurde, M. D. *Burkhard* aus Elk Grove Village bei Chicago, der über den Ohrsimulator auf Grundlage des Zwislocki-Kupplers referierte, ebenso wie über die Anforderungen und Entwicklungen, die zu dem jetzt bestehenden IEC-Vorschlag führten. Leider wird von diesem Simulator der wichtigste Aspekt, nämlich die Dämpfung des menschlichen Ohres (Realteil der Impedanz), nicht ausreichend nachgebildet*), und der Referent vergaß nicht zu betonen, daß individuelle Gehörgänge und Ohrimpedanzen beträchtlichen weiteren Streuungen unterworfen sind.

H. *Fastl*, München, experimentierte mit dem Pegelunterscheidungsvermögen von Patienten mit unterschiedlichen Hörstörungen und fand, daß, im Gegensatz zu modulierten, gepulste Töne bei 30 dB SL keine zusätzlichen Frequenzabhängigkeiten bringen, und gab um den Faktor 3 höhere Unterschiedsschwellen an bei retrocochleären Hörstörungen — im Vergleich zu den-

*) Siehe auch A. Guldstein, In-Situ-Messungen. »Der Hörgeräte-Akustiker« 5/81, Abb. 18, S. 3.

Forthcoming Meetings

8th Colloquium of Physicists and Engineers Concerned with Audiology

On the 26th and 27th March 1981, the Colloquium for audilogically active physicists and engineers was held in Erlangen, for the 8th time. This meeting is attended by delegates from both clinics and industry and promoted by Messrs Siemens acting as a patron and catalyst in an estimable and completely neutral manner. The colloquium was intended to bring together a small circle of about three dozen technicians representing a good cross-section of specialist audiological fields. The meeting fulfilled this purpose; an altogether praiseworthy achievement which deserves our acknowledgement all the more since audiologists too often tend to work in one-man ivory towers.

This year, the Colloquium was simultaneously intended to pay scientific tribute to Dipl.-phys. Clemens *Starke* for his 31 years' activity in the hearing-aid sector and to thank him for his dedicated work. When Mr. *Starke* retires in the middle of this year, he can look back on three decades of successful work.

This was also the reason why the series of lectures was opened by a colleague and guest from the United States, M.D. *Burkhard* from Elk Grove Village, near Chicago. His paper dealt with the ear simulator based on the Zwislocki coupler, as well as with the requirements and developments which led to the existing IEC proposal. Unfortunately, this simulator does not make adequate allowance for the most important aspect*), namely the attenuation of the human ear (effective part of the impedance) and the lecturer did not omit to emphasize that individual auditory canals and ear impedances are subjected to further considerable scatter.

H. *Fastl*, Munich, reported on his experiments with the level discrimination of patients suffering from different hearing disorders. He found that in contrast to modulated tones, pulsed tones at 30 dB SL bring no additional frequency dependence and gave difference thresholds higher by factor 3 for retrocochlear hearing disturbances, as compared with those of persons with normal hearing, conductive hearing loss, presbycusis,

*) See also A. Guldstein: In-Situ-Messungen. "Der Hörgeräte-Akustiker" 5/81, Fig. 18, page 3.



Blick in den Hörsaal beim diesjährigen Kolloquium audiologisch tätiger Physiker und Ingenieure in Erlangen (Foto Mohrbacher)

Glance into the lecture room during the Colloquium of physicists and engineers concerned with audiology of this year in Erlangen

jenigen von Normalhörenden, Schalleitungs-, Alters-, Lärmschwerhörigkeit, toxischen Schäden, Hörsturz und M. Ménière.

Tägliche und durchaus verbreitete Kümernisse des klinisch tätigen Technikers bei der Erstellung audiologischer Dokumentation beschrieb M. Berg, Erlangen, für die er die On-Line-Datenerfassung vom Audiometer als für Routinefragen unangemessen aufwendig und den Assistenten unzumutbar hielt und lieber, sozusagen für die Nach-Verarbeitung am Nachmittag, die rechnergestützte Off-Line-Dokumentation vorschlug.

Die zukünftige Hörgeräte-Anpaß-Meßtechnik verlangt, nach den Worten von L. M. Moser, Würzburg, sowohl nach dem (ungläubigen) SAUL (Symbolic Audiometric Language) als auch gleichzeitig nach dem (gläubigen) PAUL (Programmed Audiometer Logic), so daß, vergleichsweise, auch Ungläubigen durch die Botschaft der Bibel die Bekehrung zu umgestalteter Technik nahegebracht wurde, sofern sie nicht auf eine Weiterentwicklung des Hörgerätes verzichten.

Unter dem Aspekt des zeitlichen Ablaufs von Sprachphonemen diskutierte H. v. Wedel, Bonn, das Zeitaufhebungsvermögen des Gehörs und die entsprechenden Wünsche an die Übertragungstechnik von Hörgeräten; insbesondere das Ausschwingverhalten — akustisches »Loch« — von AGC-Regelschaltungen mag zu Verständnisschwierigkeiten leise nachfolgender Konsonan-

hearing impairment of toxic origin, sudden loss of hearing and Ménière's disease.

The daily and ubiquitous woes of a technician in a clinic were described by M. Berg, Erlangen, in connection with the preparation of an audiological documentation. He felt that the on-line data acquisition from the audiometer is inadmissibly expensive when routine problems are involved and unacceptable for the assistant, and therefore he proposed computer-assisted off-line documentation as post-processing in the afternoon.

In the opinion of L. M. Moser, Würzburg, the future hearing-aid fitting and measuring procedure calls both for the (doubting) SAUL (Symbolic Audiometric Language) and the (believing) PAUL (Programmed Audiometer Logic) so that, by way of comparison, also non-believers are converted, through the message of the bible, to the reorganised procedure, provided they did not renounce the further development of the hearing aid.

H. v. Wedel, Bonn, discussed the temporal resolution of the hearing from the aspect of the temporal sequence of speech phonemes as well as the requirements to be satisfied by the hearing aids with regard to transmission technology. Especially the decay behaviour — the acoustic "gap" — of AGC control circuits may lead to difficult understanding of soft succeeding consonants and transients. With a success rate of only 10

ten und Transienten führen. Mit einer Erfolgsrate von nur 10 Prozent (8 von 80 Patienten) beurteilte er im weiteren die Versorgung von Patienten mit Ohrgeräuschen durch Tinnitus-Masker recht kritisch, ähnlich wie J. Kießling, Gießen, der Langzeiterfolge (länger als 1 Jahr) bei 34 ausgesuchten, stark leidenden Patienten von weniger als 10 Prozent erreichte; wenn schon, war eher, bei gleichzeitig bestehender Schwerhörigkeit, ein normales Hörgerät das Mittel der Wahl (von 58 Patienten 37 mit gutem, 13 mit partiellem Erfolg). Offensichtlich, und leider, erfüllen die reinen Masker-Geräte, zwar vielleicht als technisches Angebot, aber doch nicht die weiter in sie gesetzten pathophysiologischen Erwartungen. Erschwerend ist die Feststellung, daß subjektive Angaben über das Ohrgeräusch schlecht mit nachfolgenden Tinnitus-Analysen korrelieren.

Über die Bestimmung der Normal-Hörschwelle für Knochenleitung subjektiv und objektiv am neuen B&K-Mastoid (das alte ist nicht mehr verwendbar) berichtete K. Brinkmann, Braunschweig. Bleibt nur die (geringe) Hoffnung, daß ein für den Techniker gut kalibrierbarer Knochenleitungshörer ebenso einfach zu Routinemessungen an Patienten brauchbar wäre.

Individuelle, durch Körperschallabnehmer objektiv ermittelte Streuungen bei Aufnahme der Knochenleitungsschwelle liegen, so D. Röser, Frankfurt, bei ± 4 dB (tiefere Frequenzen) bis ± 7 dB (höhere Frequenzen ab 800 Hz). Die Abhängigkeit vom Andruck betrage 12 dB beim Übergang von 200 auf 1200 p (der erstere Wert dürfte dem in der Praxis gebräuchlichen Andruck entsprechen; der Norm-Andruck von 400 p ist bereits sehr hoch).

Ausgehend von der Beobachtung, daß das Auge eine 30mal höhere Verarbeitungsgeschwindigkeit hat als das Ohr, kam G. Esser, Düsseldorf, auf den Gedanken, beispielsweise für das Sprachtraining Hörgestörter, Sprache nach Amplitude, Zeitverlauf und Frequenzinhalt in einen langsamen, farbigen Bildablauf zu transformieren, wobei über Dreiecksfilter den tiefen, mittleren und hohen Frequenzen die Farben rot, grün und blau äquivalent zugeordnet werden; die Beispiele zeigten sehr hübsche und anschauliche Wort-Bild-Muster.

M. Zollner, München, erörterte die elektronische Vorverarbeitung und mechano-elektrische Einleitung von Sprachsignalen mit den Schwierigkeiten der Elektroden-Übergangsschicht zur elektrischen Stimulation der Hörnerven (Reizung der Cochlea). Zur Codierung verwendete er das Vocoder-Prinzip mit 12-6-3 modulierten Frequenzlinien (letztere: 50 Prozent Verständlichkeit für Freiburger Einsilber) und führte akustische Beispiele hierzu vor. Für Implantat-Hörgeräte erachte-

per cent (8 of 80 patients) he assessed rather critically providing patients with ear noise generated by tinnitus maskers. No less critical was J. Kießling, Gießen, who reached long-term (longer than a year) scores of less than 10 per cent in 34 selected, severely suffering patients; if the condition was to be remedied, a normal hearing aid was rather the means of choice (of 58 patients, 37 used it with good success, 13 with partial success). Unfortunately, it appears that the pure maskers, which may be acceptable as a technical proposal, have failed to meet the pathophysiological expectations placed in them. What makes matters worse, is the fact that subjective data on ringing in the ears poorly correlates with subsequent tinnitus analyses.

K. Brinkmann, Braunschweig, reported on the determination of the normal hearing threshold for bone conduction subjectively and objectively on the new B&K mastoid (the old version is no longer suitable for use); thus there remains only the (slight) hope that a bone-conduction receiver which can be readily calibrated proves equally simple to handle by the technician in routine measurements on patients. According to D. Röser, Frankfurt, the determination of the bone-conduction threshold, revealed the scatter objectively ascertained with body-sound pickups, to range between ± 4 dB (lower frequencies) and ± 7 dB (higher frequencies from 800 Hz); the dependence on the pressing-on force he reported to be 12 dB on changing from 200 to 1200 p (this value might correspond to the pressing-on force used routinely; the standard pressing-on force of 400 p is already very high).

Based on the observation that the human eye has a processing speed 30 times higher than that of the ear, G. Esser, Düsseldorf, had the idea that for the speech training of the hearing-impaired, speech could be converted into a slow-moving colour-coded pictorial sequence broken down into amplitude, temporal sequence and frequency content, where, via triangular filters the colours red, green and blue are assigned to the low, medium and high frequencies; the word/picture patterns he presented were very attractive and illustrative.

M. Zollner, Munich, discussed the electronic preprocessing and mechano-electric feed of speech signals and the difficulties encountered in the electrode transition layer to the electric stimulation of the auditory nerves (stimulation of the cochlea); for the coding, he used the Vocoder principle with 12-6-3 modulated frequency lines (the latter: 50% intelligibility for Freiburg monosyllables) and presented acoustic examples. He regarded at least 12 modulation frequencies as necessary for implanted hearing aids because not all

te er wenigstens 12 Modulationsfrequenzen als notwendig, weil nicht alle Elektroden beim Einführen auf der Cochlea richtig »landen«.

Über die Strategie eines computerisierten Testsystems machten sich N. Dillier und J. Güntensperger, Zürich, Gedanken, um Hörempfindungen und sogar Sprachdiskrimination elektrisch zu stimulieren.

Besonders erfreulich bei dem gesamten Ablauf des Programms: Die Zeit, die nicht nur den Vortragenden, sondern insbesondere auch den Diskutierenden gegeben war, voneinander zu hören. *Friedrich Keller*

electrodes are placed correctly on the cochlea when they are introduced. N. Dillier and J. Güntensperger, Zurich, investigated the strategy of a computerized testing system for the electrical stimulation of hearing perception and even speech discrimination.

It was especially gratifying to observe that the participants had ample opportunity to make personal contacts and to exchange views on the fields discussed. *Friedrich Keller*

Personalia

Amplifon CRS Preis 1980 für Prof. Dr. H. L. Wullstein

In Würdigung seiner besonderen Verdienste um die Bekämpfung der Schwerhörigkeit wurde am 8. April dieses Jahres in Mailand Prof. Horst Ludwig Wullstein, Würzburg, mit dem Amplifon CRS Preis 1980 ausgezeichnet. Die Verleihungszeremonie fand im Napoleon-Saal des Mailänder Presse-Clubs statt.

Prof. Ettore Pirodda, Direktor der 1. HNO-Klinik der Universität von Bologna, beleuchtete die Persönlichkeit des renommierten deutschen Wissenschaftlers. Einen Einblick in die Ursprünge der von Prof. Wullstein erarbeiteten Tympanoplastik-Operationstechniken gab Prof. Carlo Zini, Direktor der 2. HNO-Klinik der Universität von Parma. Am gleichen Nachmittag hielt Prof. Wullstein im Amplifon Entwicklungs- und Forschungszentrum einen Vortrag zum Thema »Die Entstehung der Tympanoplastik und der technische Fortschritt, der ihre Verwirklichung ermöglichte«.

Prof. Wullsteins Überlegung, das Wissen um die Funktion des Mittelohrs bei der operativen Behandlung der chronischen Otitis media anzuwenden, machte es in den frühen 50er Jahren möglich, das Prinzip der Tympanoplastik zu definieren. Den Begriff »Tym-

Prof. Dr. H. L. Wullstein receives Amplifon CRS Preis 1980

Prof. Horst Ludwig Wullstein (Würzburg-Germany) was awarded, on April 8th last in Milan, the Amplifon Research and Studies Centre International Prize for his essential contribution to the fight against deafness. The prizegiving ceremony was held in the Napoleon Room of the Milan Press Club.

The figure of the renowned German scientist was illustrated by Prof. Ettore Pirodda, Director of the I ENT Clinic of Bologna University. Prof. Carlo Zini, Director of the II ENT Clinic of Parma University described the original aspects of the tympanoplastic surgical techniques worked out by Prof. Wullstein.

The same afternoon Prof. Wullstein gave at the Amplifon Research and Study Centre, a lecture on: "The birth of tympanoplasty and the technical progress which has made its realization possible".

Prof. Horst Ludwig Wullstein had the original idea of applying the knowledge of the middle ear audiology function to the surgery of chronic otitis media, thus making it possible — in the early fifties — to rationalize the principles of "tympanoplasty", a definition he used in 1953 for the first time.

A. C. Holland, Präsident des Amplifon Entwicklungs- und Forschungszentrums (links), und Prof. Dr. Horst Ludwig Wullstein, Inhaber des Internationalen Amplifon CRS Preises 1980, im Mailänder Presse-Club. (Foto Argenta Primo)

A. C. Holland, President of Amplifon Research and Studies Center (on the left), and Prof. Dr. Horst Ludwig Wullstein, winner of the International Amplifon CRS Prize 1980, at the Milan Press Club.



panoplastik« verwandte Prof. Wullstein 1953 zum ersten Mal. Ein Operationsmikroskop erleichterte es dem Chirurgen, die für die Mittelohrentzündung verantwortlichen entzündlichen Herde auszuräumen und einen vereinfachten Mechanismus zu rekonstruieren, der noch imstande war, Töne mit dem verbliebenen und nun nicht mehr infektiösen Mittelohr-Rest zu verstärken.

Dem Verleihungsausschuß des Amplifon CRS Preises gehören folgende Personen an: Prof. Antonio Antonelli, Prof. Emanuele Biondi, Prof. Silvano Boccardi, Dr. Marcello Cantoni, Prof. Giuseppe Cavallazzi, Prof. Massimo Del Bo, Prof. Aldo Dufour, Dr. Franco Ferrero, Prof. Fabio Giaccai, Anna Maria Holland, Ing. Angelo Zanibelli, Ing. Charles Holland, Prof. Antonio Ottaviani, Prof. Giovanni Rossi, Prof. Gino Sacerdote, Prof. Carlo Sirtori, Prof. Luigi Vignolo und die Präsidenten von: American Speech-Language Hearing Association (ASHA), International Society of Audiology (AUDI), American Auditory Society (AAS), Società Italiana di Otorinolaringoiatria, Società Italiana di Foniatria, Società Italiana di Audiologia.

In der Vergangenheit wurde der Preis verliehen an: Prof. Tokuro Suzuki (Japan), Prof. Erik Wedemberg (Schweden), Prof. Hallowell Davis (USA), Suzanne Borel Maisonny (Frankreich), Prof. James Jerger (USA), Prof. Jozef J. Zwislocki (USA), Prof. Michele Arslan (Italien), Prof. Aram Glorig (USA) und Prof. Ettore Bocca (Italien).

Red.

Using an operating microscope the surgeon was able to eliminate more easily the septical foci responsible for the chronic middle ear infection and therefore to reconstruct a simplified mechanism still capable of amplifying sound by means of the residual, not yet infected middle ear structure.

The awarding Committee was as follows: Prof. Antonio Antonelli, Prof. Emanuele Biondi, Prof. Silvano Boccardi, Dr. Marcello Cantoni, Prof. Giuseppe Cavallazzi, Prof. Massimo Del Bo, Prof. Aldo Dufour, Dr. Franco Ferrero, Prof. Fabio Giaccai, Mrs. Anna Maria Holland, Ing. Charles Holland, Prof. Antonio Ottaviani, Prof. Giovanni Rossi, Prof. Gino Sacerdote, Prof. Carlo Sirtori, Prof. Luigi Vignolo, Ing. Angelo Zanibelli and the Presidents of: American Speech-Language Hearing Association (ASHA), International Society of Audiology (AUDI), American Auditory Society (AAS), Società Italiana di Otorinolaringoiatria, Società Italiana di Foniatria, Società Italiana di Audiologia.

In the past the Prize was awarded to: Prof. Tokuro Suzuki (Japan), Prof. Erik Wedemberg (Sweden), Prof. Hallowell Davis (U.S.A.), M.me Suzanne Borel Maisonny (France), Prof. James Jerger (U.S.A.), Prof. Jozef J. Zwislocki (U.S.A.), Prof. Michele Arslan (Italy), Prof. Aram Glorig (U.S.A.), Prof. Ettore Bocca (Italy).

IV. Audiologische Fortbildungstagung der Technischen Universität München

Vom 19. bis 22. September 1981 findet an der HNO-Klinik der Technischen Universität München ein Fortbildungskurs für HNO-Fachärzte und Audiometrie-Assistentinnen statt.

Auskunft und Anmeldung erbeten an Prof. Dr. H. Chüden, c/o HNO-Klinik der Technischen Universität München, Ismaningerstraße 22, 8000 München 80, Telefon 089/41 40 23 83. *Red.*

Die Mitarbeiter dieses Heftes / Our Contributors



Lisa Guillory-Renner, BA, MA, erlangte den BA (Bachelor of Arts) 1977 am Stephens College in Columbia, Missouri, und den MA (Master of Arts) 1979 in Audiologie an der Universität von Missouri, Columbia. Zur Zeit tätig als Audiologin in einer Klinik für Allgemeinmedizin in Columbia, Missouri.

Lisa Guillory-Renner, BA, MA, obtained the Bachelor of Arts degree from Stephens College, Columbia (Missouri) in 1977 and the Master of Arts degree in Audiology from the University of Missouri (Columbia) in 1979. She is currently employed as a Clinical Audiologist in a multi-specialty medical clinic in Columbia, Missouri.



Roberta H. Hansen, MA, erlangte den MA (Master of Arts) 1979 in Audiologie an der Universität von Missouri, Columbia. Zur Zeit tätig als klinische Audiologin an der Missouri School for the Deaf in Fulton, Missouri.

Roberta H. Hansen, MA, received the MA degree in Audiology from the University of Missouri, Columbia, in 1979. She is currently employed as a Clinical Audiologist at the Missouri School for the Deaf in Fulton, Missouri.

Ernst Terhardt, Prof. Dr.-Ing., zur vita autoris siehe »Audiologische Akustik« Nr. 3/1981.

Ernst Terhardt, Professor, Doctor of Engineering, as to the biographical dates we refer to "Audiological Acoustics" No. 3/1981.

Donald G. Williamson, Dr. phil., zur vita autoris siehe »Zeitschrift für Hörgeräte-Akustik« Nr. 2/1979.

Donald G. Williamson, Dr. phil., as to the biographical dates we refer to "Journal of Audiological Technique" No. 2/1979.

Audiologische Akustik

Verlag und Redaktion
median-verlag
Hans-Jürgen von Killisch-Horn
Hauptstr. 64, Postfach 103964
6900 Heidelberg 1,
Fernruf (06221) 24665
Schriftleiter:
Dr. phil. Werner Güttner
Bachstraße 11, 8170 Bad Tölz
Redaktion: Stephen C. Mohrbacher
Anzeigen: Karin Ball
Layout: Friedrich Vogt
Satz und Druck: Dietz-Druck,
6900 Heidelberg

Erscheinungsweise: sechsmal
jährlich alle zwei Monate
Bezugspreis
jährlich DM 42.— netto
Einzelpreis
je Heft DM 9.— netto
Zur Zeit hat
Anzeigenliste Nr. 8
vom 1. Januar 1978 Gültigkeit
Nachdruck, Übersetzungen,
Rundfunksendungen
nur mit Genehmigung des Verlages
© median-verlag 1961

Audiological Acoustics

Owned and published by
median-verlag
Hans-Jürgen von Killisch-Horn,
Hauptstrasse 64, 6900 Heidelberg
Editor-in-chief:
Dr. phil. Werner Güttner
Bachstraße 11, 8170 Bad Tölz
Editorial department:
Stephen C. Mohrbacher
Advertising: Karin Ball
Layout: Friedrich Vogt

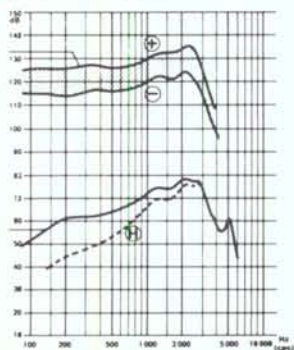
Published bi-monthly 6 issues
per annum
Annual Subscription DM 42.— no
Single Copies DM 9.— no
Supplied directly by
median-verlag, 6900 Heidelberg 1,
P.O. Box 103964
Current advertisement
rates No. 8 January 1st, 1978
All rights reserved by
median-verlag
© median-verlag 1961
Printed in Western Germany

Viennatone Hörbrillen, zuverlässige Partner für gutes Hören



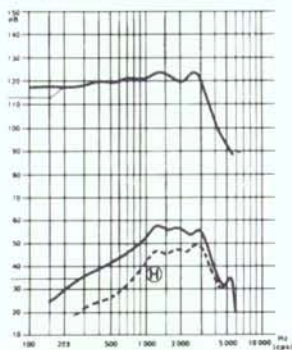
Modisches Design
Ausgereifte Schaltungstechnik
Servicefreundlicher Innenaufbau

Exzellent 77 AL-PP



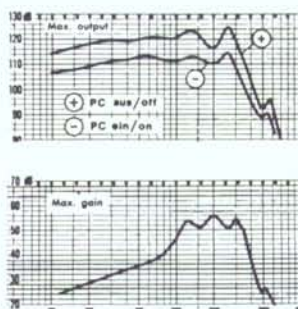
AL-PP: kontinuierlich regelbare Output-Absenkung ca. 10 dB/Tonblende (Schraubenziehereinstellung). (Electret-Mikrofon)

Exzellent 77 ALC (AVC)



ALC: kontinuierlich regelbare Dynamik-Kompression/Tonblende (Schraubenziehereinstellung). (Electret-Mikrofon)

Exzellent 77 ALPC



ALPC ist das Standardgerät der neuen Reihe. (Keramik-Mikrofon)

viennatone HÖRGERÄTE

ein Name mit Weltgeltung - ein Begriff für Qualität

SIEMENS

Thema: Diagnose von Hörschäden

Die Ton-Sprach- audiometer-Kombination:

Ton - Sprache - Geräusche - Tests.

BEOMAT von Siemens

Siemens bietet ein abgerundetes Audiometerprogramm für alle Anwendungen. Vom einfachen Siebtestaudiometer über Ton-Sprachaudiometer-Kombinationen bis zur objektiven Audiometrie mit Impedanzmeßgeräten und ERA-Systemen.

Für alle von uns gelieferten Audiometer steht Ihnen ein schneller und zuverlässiger Kundendienst überall in der Bundesrepublik zur Verfügung.



Auskünfte durch:
Siemens AG
Bereich Medizinische Technik
Geschäftsgebiet Hörgeräte
Gebbertstraße 125,
8520 Erlangen
Tel. 091 31/84-2831

Ihr Partner für alle audiometrischen Einrichtungen: Siemens