

AUDITORY EXCITATION PATTERNS
THE SIGNIFICANCE OF THE PULSATION THRESHOLD METHOD
FOR THE MEASUREMENT OF AUDITORY NONLINEARITY

The picture on the cover can be considered as a visual analogy to the continuity effect underlying the pulsation threshold method.

On the left-hand side one sees an interrupted picture of something that cannot easily be recognized. On the right-hand side the interruption is replaced by a "noisy" grid which induces the continuity effect, and shows an image of a tree.

The most clear effect can be obtained by covering the other half of the picture than that indicated for the demonstration.

The picture was drawn by Hans Kneefel after a design of a "pulsating tree" made by Prof. Dr. G. van den Brink.

AUDITORY EXCITATION PATTERNS

THE SIGNIFICANCE OF THE PULSATION THRESHOLD METHOD

FOR THE MEASUREMENT OF AUDITORY NONLINEARITY

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR
IN DE GENEESKUNDE AAN DE ERASMUS UNIVERSI-
TEIT TE ROTTERDAM OP GEZAG VAN DE RECTOR
MAGNIFICUS PROF. DR. B. LEIJNSE EN VOLGENS
BESLUIT VAN HET COLLEGE VAN DEKANEN. DE
OPENBARE VERDEDIGING ZAL PLAATS VINDEN
OP WOENSDAG 10 MEI 1978 DES NAMIDDAGS
TE 3.00 UUR PRECIES

DOOR

JOHANNIS VERSCHUURE

geboren te Goes

1978

W. D. MEINEMA B.V. - DELFT

PROMOTOR: PROF. G. VAN DEN BRINK
CO-REFERENTEN: PROF. DR. E. H. HUIZING
DR. IR. F. A. BILSEN,

Aan Jenny, Gerdy en Corrien

Aan mijn ouders

CONTENTS

	CONTEXT OF INVESTIGATION	11
	Introduction	11
	Structure and function of the auditory system	13
	Measures of the frequency selectivity	14
	Problems in interpretation of psychophysical patterns	15
	Other methods for measuring activity patterns	17
I	LITERATURE SURVEY	21
I-1	The pulsation threshold method	21
I-2	The continuity effect with speech	22
I-3	The continuity effect with tones and noise	25
I-4	Hypothetical explanations of the continuity effect	30
II	PILOT STUDIES	33
II-1	The continuity effect with pure tones	33
II-2	The continuity effect with speech	36
II-3	The continuity effect with music	38
III	SPECTRAL EFFECTS ON PULSATION MEASUREMENT	41
III-1	Introduction	41
III-2	Spectra of stimuli	42
III-3	Presentation conditions	46
III-4	Apparatus, observers and methods	47
III-5	Results	48
III-6	Discussion	53
IV	PULSATION PATTERNS	57
IV-1	Introduction	57
IV-2	Classification of methods of determination of masking and pulsation patterns	57
IV-3	Apparatus	61
IV-4	Results	62
IV-5	Discussion	71
V	ANALYSIS OF PULSATION PATTERNS	73
V-1	Introduction	73
V-2	The assumptions	73
V-3	Calculations	76
V-3-1	General description	76
V-3-2	Detailed description	77
V-4	Results	81
V-5	Consequences of the nonlinearity for shape and steepness of the patterns	86
V-6	Concluding remarks	90
VI	THE BANDWIDTH OF THE PULSATOR	91
VI-1	Introduction	91

VI-2	Apparatus	92
VI-3	Results	92
VI-3-1	Summit effects	93
a	On-frequency pulsation level	93
b	Summit pattern	94
VI-3-2	Patterns outside the bandwidth of the noise	95
VI-4	Spectral analysis of noise bands	97
VI-5	Analysis of noise-band patterns	98
VI-5-a	Analysis of the high-frequency side of the input pattern	98
VI-5-b	Analysis of the high-frequency side of the output pattern	102
VI-5-c	Analysis of the low-frequency side of the input pattern	103
VI-5-d	Analysis of the low-frequency side of the output pattern	105
VI-6	Discussion	106
VI-6-1	Summit effects	106
VI-6-2	Patterns outside the bandwidth of the noise	107
VI-6-3	Comparison with the literature	108
a	Slopes of sine-wave patterns	108
b	Slopes of sine-wave and noise-band patterns	109
c	Slopes of pulsation patterns and masking patterns	109
VII	NONLINEARITY IN NEUROPHYSIOLOGICAL EXPERIMENTS	113
VII-1	Introduction	113
VII-2	Methods and results	116
VII-3	Discussion	121
VIII	NONLINEARITY AND THEORY OF HEARING	127
VIII-1	Nonlinearity and masking patterns	127
VIII-2	Localization of the nonlinearity	129
VIII-3	Effect of nonlinearity on frequency selectivity	130
VIII-4	Slopes of the activity patterns	130
VIII-5	Comparison of slopes of activity patterns	132
VIII-6	Conclusions to be drawn from the comparison	134
VIII-7	Second filter	134
VIII-8	Models of the auditory system	136
IX	THE CONTINUITY EFFECT IN SPEECH	139
IX-1	Introduction	139
IX-2	Apparatus and methods	140
IX-3	Results	142
IX-3-a	Learning effect	142
IX-3-b	General effect of noise on intelligibility	143

IX-4	Discussion	147
X	CONCLUSIONS	151
	APPENDICES	155
A	The spectrum of the exponential amplitude envelope	155
B	The spectrum of the Gaussian transition function	157
C	The spectrum of the sinusoidal transition function	160
	SUMMARY	162
	SAMENVATTING	166
	REFERENCES	170

VOORWOORD

Aan het tot stand komen van dit proefschrift hebben velen bijgedragen. Sommigen wil ik persoonlijk noemen, me realiserend dat ik anderen tekort doe.

Allereerst noem ik mijn promotor, Prof. Dr. G. van den Brink. Gert, ik wil je bedanken voor de tijd die je voor mij hebt vrij gemaakt en voor je stimulerende opmerkingen, vooral betreffende de centrale processen achter de pulsatiedrempel, en voor het gebruik van jouw "pulserende" boom op de omslag.

Prof. Dr. E.H. Huizing wil ik bedanken voor de geboden gelegenheid dit onderwerp te bewerken en voor de kritische opmerkingen bij en interesse voor het proefschrift. Hierbij wil ik ook Prof. Dr. W.H. Struben betrekken die tot 1976 hoofd van de afdeling KNO was.

Dr. Ir. F.A. Bilsen bedank ik voor zijn kritische opmerkingen zowel bij het door-nemen van dit proefschrift, als bij eerdere gelegenheden waarbij stukken onderzoek gepresenteerd werden.

Dr. M. Rodenburg ben ik veel dank verschuldigd. Marius, vanaf mijn komst naar Rotterdam heb jij mijn onderzoek zeer kritisch en intensief begeleid. Dit heeft jou zeer veel tijd gekost en wezenlijk bijgedragen tot dit proefschrift. Ik wil je in het bijzonder bedanken voor het ter beschikking stellen van jouw analyses van de metingen van Dr. E.F. Evans, whom I also want to thank for his contribution to this thesis.

De gesprekken met Dr. Ir. T. Houtgast, de geestelijke vader van de pulsatiedrempel-methode heb ik zeer op prijs gesteld.

De technische kant van het onderzoek is met grote vaardigheid verzorgd door de heren A.J.J. Maas en M.P. Brocaar. Ronald en Michael, jullie bijdrage aan het onderzoek is zeer groot geweest doordat jullie naast de bouw van de apparatuur ook steeds als proefpersoon zijn opgetreden. Deze combinatie van proefpersoon en technicus is de apparatuur zeer ten goede gekomen en is aanleiding geweest tot vele, zeer diepgaande diskussies, ook over de achtergrond van het onderzoek.

Veel dank ben ik verschuldigd aan alle proefpersonen, maar in het bijzonder aan Drs. A.A. van Meeteren en J.N. Kroon. Aart en Han, jullie hebben de metingen met veel inzet uitgevoerd en de resultaten kritisch bediscussieerd.

Michael wil ik verder nog bedanken voor het uitvoeren van de spraakexperimenten uit hoofdstuk IX, waarvoor Mej. S. Hermer de banden heeft ingesproken. Simone, mijn hartelijke dank.

Bij het analyse werk op de rekenmachine heb ik veel hulp gehad van dhr. C. Verwey. Cees, je wist altijd weer oplossingen te vinden voor ingewikkelde problemen.

De statistische verwerking van de spraakproeven heeft problemen met zich meegebracht. Dank zij Dr. R. van Strik en Drs. H.J.A. Schouten van de afdeling Biostatistica konden deze worden opgelost.

De engelse tekst is zeer nauwgezet en met veel kennis van zaken gecorrigeerd door

Dr. R.H. Bathgate uit Eindhoven.

Tot slot wil ik mijn dank uitspreken aan hen die hebben bijgedragen aan de vormgeving van het proefschrift. Het tekenwerk werd verzorgd door Hans Kneefel van de Grafische studio en de afdrukken werden gemaakt door de fotografische dienst van de AVD. De eindversie werd getypt door Mevr. J.C. Wahouïé en de formules en appendices door Mej. M.G. van Kruining. Van een voorgaande versie werd een deel getypt door Mej. P. Maters en meerdere versies door Mevr. J. Verschuure-Gardenbroek. Jenny jouw bijdragen in het typewerk en het korrektiewerk van de definitieve versie zijn zeer tijdrovend geweest; je morele steun is voor mij nog belangrijker geweest.

Voor aanschaf van apparatuur werd financiële steun ontvangen van het Heinsius-Houbolt fonds.

Voor de vertaalkosten werd een vergoeding verkregen van het Ministerie van Onderwijs en Wetenschappen.

Het verschijnen van dit proefschrift werd mede mogelijk gemaakt door een bijdrage van de Stichting Keel-, Neus- en Oorheeskunde Rotterdam.

CONTEXT OF INVESTIGATION

Introduction

The auditory system is the total of organs that translates an acoustical signal into the perception of a sound. An acoustic signal is a vibration. It is described by physical parameters. The perception of sound is the awareness of a signal being present and the attribution of certain qualities to this awareness. These qualities are called percepts. Examples of percepts are pitch, loudness and timbre.

The percepts reflect the way in which the auditory system analyses and processes acoustic signals. Hence they also determine the most useful ways of describing the signals in physical terms.

Frequency analysis is one of the most important processes involved in hearing. It is described in mathematical form by Fourier theory. This theory states that every repetitive signal can be described uniquely as a sum of sine waves. These sine waves are called frequency components. The elementary signal, therefore, consists of only one frequency component and is a sine-wave signal. The sine wave thus is also the elementary signal for our ear; we call the perception of this signal a pure tone. Signals consisting of more than one component are perceived as complex sounds.

A pure tone has got three perceptual qualities: pitch, loudness and timbre. Pitch

is related to the frequency of the sine wave, loudness to its intensity and timbre to the spectral composition of the signal.

The frequency analysis of the auditory system is carried out with a limited accuracy. Two components separated far enough in frequency are resolved independently. If the two components are close in frequency, the perception of one interacts with that of the other. These interactions can differ considerably depending on the different conditions of presentation. The main theme of this thesis is a study of the accuracy of the ear's frequency analysis based on determinations of the threshold conditions for a special interaction phenomenon, viz:

— the perception of interrupted signals as if they were continuous —

A method based on determination of certain changes in perception is called a psychophysical method, because it requires a conscious interpretation of sounds of which the physical parameters are controlled.

One parameter is then varied and the observer has to report on the changes in his perception of the stimuli. These changes reflect in a direct way the processing by the auditory system; it does not reflect at what stage of processing this change is coded; the method merely presents us with input-output relations of the system as a whole.

Contrary to psychophysical studies, physiological studies reveal the action of certain organs in a direct way. They tell us what is going on in the organ in question, but they do not necessarily reveal the significance of this processing for hearing. The observed input-output relation may well be an artifact of the processing performed by this organ. Only a subsequent, related psychophysical study can tell us whether this relation is relevant to the process of hearing. The psychophysical data can further tell us more about the processes that take place in the organ.

In the present study we shall compare psychophysical and physiological data and combine these results in order to gain a better understanding of the frequency analysis performed by the ear.

Structure and function of the auditory system

In order to provide a basis for our further considerations we need to give a brief description of the auditory system here. For a more detailed description, the reader is referred to textbooks such as Green, 1976.

The auditory system can be divided into a number of subsystems. The first of these is the external ear (auricle and external auditory meatus); its function is primarily protective. The middle ear (with eardrum and ossicles) is the second subsystem; its function is an impedance transformation which avoids appreciable reflection of acoustic energy at the air-fluid boundary at the entrance to the third part: the inner ear or cochlea. The cochlea is a spiral shaped cavity in the petrous bone, containing a fluid-filled membranous structure dividing the space into three sub-spaces: the scala vestibuli, the scala media and the scala tympani. The scala media contains the sensory cells of the organ of Corti, situated upon the basilar membrane and covered by the tectorial membrane. A vibration of the middle-ear ossicles causes a pressure wave in the cochlea which results in vibration of the scala media. The amplitude of the vibration changes with distance along the cochlear duct for a sinusoidal signal. As a result of the mass and stiffness properties of the structures, every spot on the basilar membrane responds maximally to only one frequency thus giving rise to a place-frequency conversion. This reflects the first stage of the frequency analysis involved in hearing. However, each spot on the basilar membrane will respond not only to its "best" frequency but also to nearby frequencies. The frequency selectivity is thus reflected in the tuning properties of the individual spots. The vibration excites the sensory cells, giving rise to electric potentials. At some stage along the auditory pathway these potentials are coded into nerve-action potentials, the all-or-nothing coding used by the central nervous system. These potentials run along the brain stem up to the auditory cortex, passing a number of connecting and processing stations such as the cochlear nucleus, superior olive etc.

This peripheral processing is followed by the central processing which leads to the awareness of the sound and its interpretation. The neural activity patterns have to be recognized. This may be done by comparing them with patterns of sounds we heard before, which are stored somewhere in the brain. This recognition process is

at the same time a learning process in that it may lead to future recognition of what is now a new pattern. This recognition process is one of unconscious interpretation of sounds. Next we become aware of what the sounds mean to us, the conscious interpretation. We shall return to this interpretation process in Chapter IX with reference to various kinds of regularly interrupted speech.

Measures of the frequency selectivity

There are several ways of studying frequency selectivity. One method was introduced by Wegel and Lane in 1924. They measured the elevation of the audibility threshold of a sinusoidal signal (the maskee) caused by the presence of a stronger sinusoidal signal (the masker). The audibility threshold of the maskee can be determined at various maskee frequencies for a fixed level and frequency of the masker. The difference between the threshold of this maskee with and without the masker is called the amount of masking. The amount of masking as a function of maskee frequency is called a masking pattern.

Many authors have tried to relate this measure of frequency selectivity to other measures such as the just noticeable frequency difference (jnd-f), the spectral loudness summation, the critical band mechanism, the just noticeable intensity difference (jnd-L), and subjective pitch and loudness scales (Fletcher, 1940; Schäfer et al., 1950; Zwicker and Feldtkeller, 1967; Maiwald, 1967 a, b; Zwicker, 1970).

In these theories the masking pattern represents the activity pattern of the masker. The masking level is interpreted as the activity level of the masker at the maskee frequency (Chistovich, 1957). This activity pattern thus reflects the tuning properties of the auditory system and is a measure of the frequency selectivity. A change in some physical parameter of a signal causes a certain shift in the activity pattern. If this pattern shift can be detected by the auditory system (i.e. if it exceeds a certain threshold), the change may be heard.

This kind of theory implies the physical presence of an activity pattern similar to the masking pattern in the auditory system. The physiological patterns which could

conceivably play this role include the vibration pattern of the basilar membrane and the response pattern of the primary auditory nerve fibres. These two patterns and the masking patterns are very similar in shape but they differ in steepness. Masking patterns are considerably steeper than the basilar membrane vibration patterns (Chistovich, 1957, 1971; de Boer, 1969; Wilson & Evans, 1971; Evans & Wilson, 1973) but they are less steep than patterns constructed from auditory nerve recordings (de Boer, 1967, 1969; Wilson & Evans, 1971). Chistovich (1957), on the other hand, found no difference in the steepness of masking patterns and the patterns constructed from auditory nerve recordings.

The differences in tuning properties between the basilar membrane vibrations and the auditory nerve recordings require an additional physiological sharpening mechanism for their explanation. De Boer & Jongkees (1968) listed a few possible mechanisms, some of them located in the transducer mechanism of the organ of Corti, but lack of experimental data does not permit any conclusion about the actual mechanism so far. Evans & Wilson presented a processing model in 1973, using two filtering stages. The first filter is the mechanical cochlear filter. It is followed by an essential nonlinearity. This nonlinearity is responsible for the generation of combination tones (Goldstein 1967, 1970; Smoorenburg 1972, 1974) and does not affect the frequency selectivity. The second, linear filter causes the additional sharpening. It is followed by a spike-generating mechanism with nonlinearities such as a threshold and jitter.

In this model filters and nonlinearities are kept apart and do not interact. Our study will deal with possible interactions between filters and nonlinearities and will have a bearing on this model.

Problems in interpretation of psychophysical patterns

Returning to the psychophysical patterns and measures of frequency selectivity, we find a number of problems associated with activity-pattern theory.

-1 In the first place, values of such a measure determined by different authors vary

widely, sometimes differing by more than an order of magnitude. For example, the experimentally determined values of the just noticeable frequency difference (ind-f) vary from 0.1 Hz to 5 Hz at 1 kHz (Rakowski, 1971). Rakowski suggested that these differences were due to differences in experimental procedures. Experimental evidence for this suggestion was found by Verschuure & van Meeteren (1975). (For a more detailed discussion of this subject the reader is referred to Coninx, 1977).

The values of ind-f measured by Rakowski himself and by Verschuure & van Meeteren represent a much better frequency selectivity than that predicted by the masking pattern. This suggests that either there are grounds for doubting the validity of models such as those of Fletcher (1940) and Zwicker (1970) in general, or we must assume that while such a model is applicable, the masking pattern is the wrong measure to choose as representing the selectivity.

- 2 The second problem is one of interpretation of the masking patterns. It is generally assumed that the maskee level is a measure of the masker activity at the maskee frequency. This implies an invariable detection criterion over the entire frequency range covered by the pattern. Even the first patterns published (Wegel & Lane, 1924) show clear signs that this is not the case. Near the masker frequency and its harmonics the pattern is distorted by the perception of beats and roughness. The beats near the masker frequency cannot be avoided and are due to the temporal resolution of the auditory system: at small frequency separations, perception follows the temporal structure of the physical signal. It is somewhat more difficult to explain the beats at the harmonic frequencies as these are linked with a pitch effect called the sweep-tone effect (Plomp, 1967; Lamoré, 1975, 1977). These effects show that a different detection criterion is used in a restricted frequency range.

Another observed effect is even more damaging for the interpretation of masking patterns, viz the perception of odd-order combination tones (with a pitch corresponding to $f = f_1 - n(f_2 - f_1)$ where n is an integer). For a considerable part of the patterns the threshold of the weaker primary is higher than the threshold of the combination tone (Greenwood, 1971; Smoorenburg, 1972; Rodenburg et al., 1974), so the threshold is not even determined by the maskee but by the generation of combination tones.

Egan & Hake (1950) tried to overcome the complications by using narrow noise-band maskers. The notches in the patterns disappeared but Greenwood (1971) showed that combination bands still determined the detection threshold. The complications can only be avoided by nonsimultaneous presentation of masker and maskee as is the case in postmasking, premasking, gap masking, the pulsation threshold method etc.

- 3 The third problem is the nonlinearity of the patterns. Wegel & Lane's data (1924) showed a clear nonlinearity with level. At higher masker levels the steep, low-frequency edge of the masking pattern gets steeper, while the flatter high-frequency edge gets less steep. This tilting of the masking patterns was discussed by Chistovich (1957). She measured masking patterns by two different experimental methods. In the first method, called the direct method, the masker was fixed in frequency and level, as in the method used by Wegel & Lane. She also used another method, called the threshold method, in which the maskee level was kept constant and an upside down pattern was obtained. Turning the latter patterns downside up again does not cause them to coincide with the patterns, determined by the direct method, because of the nonlinearity involved.

Rodenburg et al. (1974) showed how the steepness of masking patterns obtained by two similar methods are quantitatively related. Their calculations show that the steepness of the patterns is influenced by the nonlinearity and depends upon the method of determination used.

Other methods for measuring activity patterns

It will be clear from the above that the use of masking patterns has some serious drawbacks and that results obtained with their aid have to be interpreted very carefully. This has been understood by many authors and some have proposed alternative methods to measure the extent of the activity.

-1 Nonsimultaneous masking

As we have already stated, some of the complications can be avoided by nonsimul-

taneous presentation of masker and maskee. Most of the authors who chose this approach presented the masker and maskee consecutively. It has been suggested (de Marée, 1939) that such a method actually measures fatigue. Lüscher and Zwislocki (1946, 1947, 1949) left the idea of fatigue and introduced the concept of momentary auditory threshold by using very short probesignals.

A clear exposition of the underlying concept was presented by Gardner (1947). He interpreted his postmasking results as the detection of the activity of the probe stimulus in the presence of some after-effect of the masker signal. In his concept the nonlinearity, which is also found in postmasking patterns, has been taken into account by the fact that we are comparing the probe activity pattern with the after-effect pattern. If the former pattern exceeds the latter pattern at any frequency, the probe tone is heard.

This postmasking method is also known under a variety of other names, such as residual masking, aftermasking, forward masking. It is widely used nowadays. The comparison of postmasking patterns with masking patterns involves a transformation. The need of this transformation may be regarded as one of the drawbacks of the method. The transformation is needed because the short probe pulse is shorter than the temporal integration time of the ear (Plomp & Bouman, 1959) so the detection level is rather high and moreover depends on the presentation conditions. On the other hand, the gap between masker and probe (usually about 20 ms) involves a decay of sensation (Plomp, 1964a). This decay is nonlinear, so a transformation curve has to be established to take the measuring conditions into account. This curve is usually determined for one and the same frequency for both signals. The probe detection level is then translated into an equivalent masker level via this transformation curve, on the assumption that this curve is frequency-independent.

Another drawback is the fact that we do not know the exact nature of this after-effect, nor at which stage in the auditory system it is established.

The measured postmasking patterns are considerably steeper than the simultaneous masking patterns. The two methods seem to be essentially different; for example postmasking shows suppression and masking does not (Houtgast, 1974a; Shannon, 1976).

The same kind of difference seems to exist between simultaneous masking and other nonsimultaneous masking methods such as premasking (backward masking) and gap masking. These methods will not be discussed further in this thesis.

-2 Pulsation threshold method

Another technique for measuring interactions between signals is the pulsation threshold method. We shall be using this method in this investigation. It was introduced by Houtgast (1971, 1972). It is a special case of gapmasking in that the signals are presented alternately. However, the threshold criterion is not detectability as in masking, but a perception of one of the two signals as continuous.

Houtgast (1974a, b) and Shannon (1976) showed that the results obtained with the pulsation threshold method are very similar to those given by postmasking. Both methods yield very steep patterns and show suppression. The advantage of the pulsation threshold method is that no transformation curve is needed (Houtgast, 1974a). A drawback is that neither the underlying mechanism, nor the location where it is established, is known.

Furthermore there are practical problems concerning the presentation conditions and the assumptions on which the interpretation is based. These problems will be dealt with in this study of auditory nonlinearity and the pulsation method.

I. LITERATURE SURVEY

I-1. The pulsation threshold method

The pulsation threshold method in its present form was introduced by Houtgast (1971). An alternation of two stimuli is presented to a listener at an alternation rate of 4 Hz, corresponding to a presentation time of 125 ms for each stimulus. The amplitude envelope at the switching-over is smoothed to avoid strong transient response.

When sine waves of the same frequency and level are used, a continuous tone is perceived. When one of the two is attenuated by more than a just noticeable difference, a loudness-modulated tone is perceived (we neglect possible phase effects which may be present but which can be easily recognized as such).

When sine waves of different frequencies and the same level are used a sequence of tone bursts is perceived. If the level of one sine wave (preferably the higher-frequency one) is lowered sufficiently, a stronger pulsating tone (the pulsator) and a weaker continuous tone (the probe) is perceived. If the probe level is now gradually raised, the probe will start pulsating again at a certain level which is called the pulsation threshold level of the pulsator at the probe frequency.

The pulsation (threshold) level (L_p) depends largely on the frequency separation Δf between pulsator and probe. Plotting L_p against Δf yields a "pulsation pattern"

very similar in shape to the masking pattern.

Pulsation patterns can also be determined for complex pulsator signals. The description of such patterns is more complicated because they reflect the spectrum of the pulsator as well as the auditory frequency selectivity of the listener.

The pulsation threshold method is not a real masking method in the strict sense as it is not based on determinations of a detection threshold but on a determination of a threshold at which one of two alternately presented signals sounds as if it were continuous.

The continuity effect on which it is based has been described for speech much earlier by Miller & Licklider (1950). They reported that periodically interrupted speech is perceived as continuous if the silent intervals are filled up with noise. This observation was looked upon for some time as a special case of the alternate presentation of speech to the two ears in turn. The first description of a similar effect for sine waves rather than speech was published by Thurlow in 1957. The two lines of investigation led a separate life for quite some time, because they received attention from two different groups of researchers with different interests and different views viz linguists, interested in the central processes underlying the effects of interruption and alternate presentation of speech to the two ears in turn on the intelligibility, and psychoacousticians, interested in the peripheral processes. The separate life of the two lines of investigation necessitates a separate review of the literature.

1-2. The continuity effect with speech

Miller & Licklider (1950) observed a continuity effect for speech interrupted about 10 to 15 times per second, if noise was presented in the gaps. They state that the voice of the speaker sounds hoarse and racous if no noise is presented in the gaps but becomes more natural and continuous when noise is presented in the gaps at a somewhat higher level than the speech. They called this phenomenon the picket-fence effect, after the visual effect of looking at a garden through a picket fence. Although the visual image of the garden is interrupted at regular intervals, the garden is still perceived as continuing behind the pickets and forming one total uninterrupted

picture. The authors observed that the speech seemed more intelligible when perceived as a continuous sound, but they could not demonstrate any difference in monosyllable intelligibility with and without noise at the interruption rates below 10 Hz. At higher interruption rates, the articulation scores depended on the level difference between speech and noise (the S/N ratio): the lower this ratio, the lower the articulation score. This observation can be understood in terms of a masking phenomenon, as was shown by Dirks & Bower (1970). They showed that the addition of noise causes post- and pre-masking effects on the speech, reducing its intelligibility.

The effect of the noise on the perception of speech is very strong and the first impression always seems to be that the noise improves the intelligibility. It is thus not surprising that in later years various investigators took up the subject again; however their publications show quite conflicting results.

Schubert and Parker (1955) found that the intelligibility of speech, interrupted at a rate of about 4 Hz, was about 64%. The presentation of noise in the silent intervals raised the score to about 77%. Their scores were determined by a "shadowing" technique, in which the observer had to reproduce the speech while listening to it; there was no silent interval for the reproduction of the speech. Intelligibility scores were determined for keywords in the text.

Hopkinson (1967) strongly criticized this method. She argued that the reproduction of the speech interferes with the listening and that it is this interference that causes the rather low score of 64% at the critical interruption rate of 4 Hz. She used monosyllables which were presented alternately to the two ears. The listener had to write down his response. The addition of noise tended to reduce the intelligibility of the speech but not significantly so.

Kreul (1971) used meaningful and nonsense monosyllables. The listener had to choose from a limited number of alternative responses and marked his choice on paper. Kreul found that the addition of noise reduced the intelligibility of the speech in both cases. His scores without noise at 4 Hz were 98% for meaningful speech and 92% for nonsense speech.

Dirks & Bower (1970) investigated the effect of noise, applied non-simultaneously, on the intelligibility of monosyllables. At an interruption rate of 1 Hz the intelligibility score was determined by the speech information received by the ear during the on period. The noise had no effect on the intelligibility. At interruption rates of 10 Hz and 100 Hz the maximum intelligibility score was about 90%, but the noise level had a clear masking effect on the interrupted speech reducing its intelligibility. The maximum score depended on the S/N ratio, the optimum value of which depended in its turn on the repetition rate.

Powers and Wilcox (1977) recently published results on the intelligibility of interrupted running speech with intervening noise for rates between 1 Hz and 4 Hz as a function of S/N ratio. They again determined the intelligibility score with reference to keywords. They found that noise improved the intelligibility. The degree of improvement depended somewhat on the S/N ratio, but always in such a way that the lower the S/N ratio, the better the intelligibility, down to a S/N ratio of -24 dB. The scores for S/N ratios between -6 dB and -24 dB were not statistically different; they were, however, better than the scores for positive S/N ratios.

Comparison of the results on intelligibility of interrupted speech with and without intervening noise led us to the following conclusions:

The noise does not improve intelligibility scores if monosyllables or nonsense speech is used. If anything, the noise reduces the score. This reduction is more marked, the higher the interruption rate. A low S/N ratio also reduces the intelligibility score at high interruption rates. Dirks & Bower showed that this was due to masking. If running speech is used, the intervening noise leads to improved intelligibility for moderate interruption rates (less than 10 Hz). The most appropriate S/N ratios are between -6 and -24 dB.

Huggins (1964) presented an explanation of the difference in intelligibility of interrupted speech with and without intervening noise. He argued that the speech signal is distorted at the moments of switching, producing false articulation cues. Noise masks these false cues, giving a higher intelligibility score. At that time no experimental data on monaurally presented interrupted speech were available. Such data became available only recently (Powers & Wilcox, 1977), and showed indeed that

intelligibility is higher with noise. Huggins regarded his explanation for monaural interrupted speech as a special case of that for speech presented alternately to the two ears with and without noise in the silent intervals. Schubert & Parker (1955) had investigated such stimuli and had presented a similar explanation, even involving the possibility of conflicting information coming from the two ears.

1-3. The continuity effect with tones and noise

At the same time other workers were investigating the continuity effect with tones and noises. Miller and Licklider (1950) also observed that a continuous tone can be perceived when a sine wave and a noise of a higher intensity are presented in alternation. Thurlow (1957) described the continuity phenomenon for an alternation of two sine waves. He wrote: "Under certain conditions the more intense of the two tones is heard clearly intermitted (somewhat as 'figure') and the less intense appears to sound continuously (somewhat as 'ground'). In a sense this is analogous to the situation in vision, where the ground is perceived as extending continuously behind the figure" (cf. the picket-fence effect of Miller and Licklider).

A more detailed paper on tone-on-tone continuity was published by Thurlow and Elfner in 1959. Their explanation of the continuity effect is based on the detectability of gaps between two tone pulses which they supposed to involve postmasking effects as described by Lüscher & Zwislocki (1946, 1947, 1949) and Miller (1948). In the older literature (see e.g. de Marée, 1939) those effects were ascribed to auditory fatigue, but later they were regarded as involving facilitation of ongoing neural discharge after the termination of the stimulus. The facilitation idea fits in nicely with the continuity effect. In view of their approach, these authors limited their experimental work to the measurement of minimum detectable silent intervals and the influence of repetition rate and intervening stimuli on these intervals.

Thurlow and Elfner used sine waves for both stimuli, and found that a strong sine wave in the gap reduced the detectability of the gap; the interval could be made longer before it was detected. The minimum detectable interval depended on the repetition (or interruption) rate and the frequency separation. These authors

also discussed in great detail other possible explanations than presented above. They abandoned the figure-ground concept because of the temporal sequence necessary for the continuity effect and because of the possibility of hearing figure-ground effects even without continuity. (These effects have recently been more systematically investigated by van Noorden (1975, 1977)). They considered von Békésy's 'funneling' of energy (von Békésy, 1958) as an explanation, but rejected it because the continuity effect described (the elongation of the gap-detection interval) was also found for stimuli presented to different ears. They conclude that the effect "may well take place centrally (and not at the level of first-order neurons)". Another observation that cannot be explained by a funneling process is the continuity effect for white noise induced by a sine wave of higher intensity.

The possibility of getting a continuity effect under dichotic conditions or when noise was presented alternately with a sine wave was investigated further by Elfner and various coworkers.

Elfner and Marsella* (1966) corroborated the existence of the continuity phenomenon under these two conditions, and Elfner and Caskey (1965) explored the existence region for "noise continuity" with monaural and dichotic presentation. They found the continuity threshold (the duration of the minimum detectable interval) significantly higher under these two conditions. The dichotic threshold (about 7.5 ms) was somewhat higher than the interruption threshold (about 6 ms) and did not depend much on the frequency of the intervening sine wave. Monaural thresholds were much higher (10-15 ms) and depended on the frequency separation. All thresholds depended on the rate of alternation (or interruption).

Elfner and Caskey relate their monaural findings to Plomp's rate-of-decay explanation (1964a), according to which an interruption can be detected if the drop in the "sensation" level exceeds a critical value (just noticeable difference). They explained their dichotic results in more general terms, assuming facilitation of ongoing neural discharge at a stage where binaural interaction takes place (Galambos et al., 1959; Bergeijk, 1962).

* This paper was presented to a meeting of the Midwestern Psychol. Assoc. in St. Louis in 1964.

The experiments with dichotic stimulation were repeated by Elfner and Homick (1966) at lower levels in order to exclude cross-ear leakage. They corroborated the above findings and found a larger frequency effect under monaural presentation conditions than that reported by Elfner and Caskey.

Elfner & Homick also investigated the effect of the duration of the presentation (or number of presented alternations) and the alternation rate. They found that the alternation rate does have an effect on the results but that this effect disappears at longer presentation times with monaural presentation, but not with dichotic presentation. The interruption threshold (minimum detectable silent interval) depends on the alternation rate. The two-factor model (peripheral-central) of Elfner and Caskey (1965) can fully explain these observations.

Elfner and Caskey (1965) and Elfner and Homick (1966) found that the noise continuity did not depend on the frequency of the intervening sine wave. Elfner and Homick (1967) investigated whether this was also true for tone-on-tone continuity with dichotic stimulation. They found that the continuity threshold depended on the frequency separation between the stimuli just as Thurlow & Elfner (1959) had found for monaural presentation.

Continuity effects under free-field stimulation conditions were investigated by Elfner (1969, 1971). The 1969 paper gives the results of noise continuity investigations which were very similar to the monaural results. The angular separation between the loudspeakers was 15° . The 1971 paper deals with tone-on-tone continuity investigations for various angles between the loudspeakers. A frequency effect was found again (c.f. monaural results of Thurlow and Elfner, 1959 and the dichotic results of Elfner and Homick, 1967). The angle between the loudspeakers producing the tones has only a minor effect.

The effect of source separation had been investigated earlier by Thurlow and Marten (1962) for alternating noise bursts. The major difference between these experiments and those reported above lies in the duty cycle of the presentation. All continuity thresholds so far had been measured as maximum undetectable gaps. The signal perceived as continuous was thus "on" for the greater part of the presentation time while the pulsator was only on for very short periods. In most of the experiments by Thurlow and Marten the on time was equal to the off time (duty cycle 50%). A duty

cycle of 50% is also used in pulsation threshold measurements.

The two alternating noise bursts gave rise to the perception of one steady noise or one steady and one pulsating noise or one steady and two pulsating noises, depending on the angle of separation of the sources and the alternation rate.

The most interesting part of the paper is that dealing with the explanation of the noise continuity at wider separation angles. The authors assume a neural distribution of disturbances caused by the noise bursts. If these distributions overlap, the authors imagine the continuous activity to be projected on to the converging neurons, which could produce excitatory sharpening. It is interesting to see that an overlap of excitation is assumed in order to get the continuity effect, in this case in the localization system.

Houtgast (1972) studied lateral inhibition and compared its appearance in various masking methods, including the gap masking method. Under special gap masking conditions the continuity effect is produced. In gap masking, the detection of the probe is used as criterion. If the probe is now presented at a supraliminal level, it can be heard as a continuous tone.

Houtgast found that the continuity effect can be clearly heard at repetition rates of about 4 Hz. He also reported a lower limit of about 1 Hz (Houtgast, 1974 a), where the phenomenon "fades away". He assumes that spectral spread will interfere at higher rates. Houtgast himself did not investigate these effects but he refers to unpublished data of van Meeteren and to Verschuure et al. (1974); these data were published in greater detail by Verschuure et al. (1976) and are also presented in this thesis.

From all his publications, it is clear that Houtgast used the pulsation threshold method simply as a tool to study inhibition or suppression effects, without going into details about the method (Houtgast, 1974 b). He showed that non-simultaneous masking procedures give such suppression effects, while simultaneous masking methods do not. Pulsation results are very similar to non-simultaneous masking results and can be obtained more easily and more accurately. The similarity of pulsation and postmasking results was corroborated by the experiments of Shannon (1976).

Houtgast's observation of differences between results from simultaneous masking ex-

periments and pulsation experiments was criticized by Fastl (1974 a) who found almost identical patterns for a critical-band-noise masker/pulsator in simultaneous masking and in pulsation. With a sinusoidal masker he found a steeper low-frequency edge to the pulsation pattern. Houtgast (1974 c) compared simultaneous masking patterns and pulsation patterns for narrow bands of noise and found them significantly different, the pulsation patterns having a steeper low-frequency edge. Fastl (1974b) commented on this finding by presenting some more results, however without specifying the spectrum of the noise band. Only a small, insignificant difference is found in his results.

Fastl (1975) also compared the pulsation patterns of a sine wave and critical-band-noise. He found that the patterns are similar but that they differ at the steep, low-frequency side, especially for higher frequencies. His steep edges are still less steep than Houtgast's. He ascribes this difference to two-tone interaction which, he says, occurs in Houtgast's paradigm (no gap between stimuli) and not in his (gap of 5 ms). Next he presents some data showing an effect of repetition rate on the steepness, corroborating our findings (Verschuure, 1974).

Kronberg et al. (1974) presented a parametric study on monaural and dichotic pulsation. They found that only repetition rates between 2 Hz and 10 Hz could be used. They found no effect of different transient functions on the observed pulsation patterns and a clear effect of repetition rate on the patterns. A variation of inter-stimulus gap duration caused considerable differences between patterns. They showed that the pulsation patterns of a two-frequency masker, the components being one octave apart, is the simple sum of the patterns for each component. This implies that there was no suppression. This difference between their findings and those of Houtgast is due to their particular choice of components.

Their dichotic pulsation patterns are most interesting. These patterns depend very little upon the pulsator level. The effect of the frequency separation is not clear, being blurred by the apparent influence of the audibility threshold on the shape and size of the pattern. In any case they found dichotic tuning to be far less sharp than that found in monaural pulsation. This is in contradiction with the finding of Elfner & Homick (1967), though we must bear the very different paradigms in mind.

Schreiner et al. (1977) published recently some results of an investigation of presentation conditions in pulsation. Their findings corroborate our earlier published results (Verschuure et al., 1976) to a great extent. Apparent discrepancies can be explained as resulting from different presentation conditions (see Chapter III).

1-4. Hypothetical explanations of the continuity effect

Various mechanisms for the continuity effect have been proposed. All these hypothetical models use overlap of neural excitation patterns as their basic assumption.

Houtgast needed a working hypothesis for the interpretation of his data. He put forward the following hypothesis (Houtgast, 1972):

"When a tone and a stimulus S are alternated (alternation cycle about 4 Hz) the tone is perceived as being continuous when the transition from S to tone causes no (perceptible) increase of nervous activity in any frequency region. This implies that the transition from tone to S causes no (perceptible) decrease of nervous activity for any frequency region".

This double criterion, applying to both transitions, was rephrased as follows in 1974 (Houtgast, 1974 a):

"A necessary condition for perceiving the pulsating character of the series of test-tone bursts would be that at the transition from masker to test-tone there is an increase of the response in the neural region associated with the test-tone frequency; if not, the series of test-tone bursts is perceived as a continuous tone".

It is clear from both formulations that we have to take full account of the cochlear filtering, and that the distribution of what we call the probe stimulus is compared with the activity of the pulsator stimulus at some level of the auditory nervous system. An overlap of these excitation patterns is necessary for the continuity effect. A similar hypothesis has been proposed by Thurlow and Marten (1962) for the localization system. It is not in conflict with the concept of facilitation of ongoing neural discharge (Thurlow and Elfner, 1959).

Warren et al. (1972) also published data on the continuity effect. They claimed that the patterns measured by what they called auditory induction reflect the ear's

frequency resolving power and pointed out the similarity of masking patterns and what we would call pulsation patterns. They noted the necessity for overlap of the spectral composition of the two stimuli and suggested:

"If there is contextual evidence that a sound may be present at a given time and the peripheral units stimulated by the louder sound include those which would be stimulated by the anticipated fainter sound, then the fainter sound may be heard as present".

All these models are more or less in line with one another as far as the physiological processes involved are concerned. They are all based on the fact that pulsation threshold patterns reflect the auditory frequency selectivity. The point where various authors diverge is when they consider the place where the decision about continuity is taken. Houtgast only talks in terms of a neural response at some level in the auditory pathway beyond the stage where lateral suppression operates (Houtgast, 1974 a, p. 79) and refrains from any speculation about the precise level involved (p. 80). Warren et al. use the words "contextual evidence", but they do not work out what they actually mean by this.

The dichotic pulsation thresholds (Kronberg et al., 1974) and the dichotic continuity effects (Elfner and Homick, 1967) indicate that the continuity decision is taken at a central place. This would also fit in with the improvement of the intelligibility of speech due to intervening noise (Schubert and Parker, 1955; Powers and Wilcox, 1977). We shall discuss our view on the continuity effect in Chapter IX.

II. PILOT STUDIES

II-1. The continuity effect with pure tones

Our first experiment explored the existence region of the continuity effect for pure tones. It is evident from general considerations that at very low alternation rates (e.g. one stimulus per five minutes) the continuity of the tone will not be perceived.

We presented two alternating sine waves monaurally to three trained observers. The equipment is described in detail in section III-4. It is sufficient to note here that the switching envelope of each sine wave was smoothed to avoid strong transient responses and that there was no silent interval between the two presentations. At first both stimuli were presented with an equal duration, given as the presentation time T (duty cycle of 50%). The observers were asked to adjust either the frequency or the level of one of the sine waves until the pulsation threshold was reached, and to describe their experiences in so doing.

At presentation times longer than about 400 ms, all observers had the experience that a unique continuity criterion could not be established: even after an adjustment had been made, the sound heard could change from continuous to pulsating. The shortest presentation time at which this effect was found, varied from 250 ms to 400 ms among the various observers.

Next we presented the stimuli at higher repetition rates. Up to about 10 Hz, no observer had any difficulty in making the adjustments. At higher rates the observers

reported difficulties. They could not separate the two tones any more: the pulsation of one tone could not be distinguished from that of the other. It was still possible to hear some kind of continuity in the tones but it was not possible to tell whether the weaker tone sounded completely continuous. We concluded from these observations that the continuity effect does exist and can lead to reproducible pulsation threshold determinations for presentation times between 50 ms and 250 ms (f_{rep} between 10 and 2 Hz).

There are only a few reports on repetition rates available in the literature. Houtgast (1974 a) mentions only a lower limit: he reports that the phenomenon fades away at rates lower than 1 Hz. Warren et al. (1972) used a rate of 1.7 Hz ($T = 300$ ms), which is just under our limiting value of 2 Hz. It is our experience that some observers will have severe difficulties at such low rates. Kronberg et al. (1974) and Schreiner et al. (1977) reported that it is generally possible to make pulsation adjustments at presentation times between 50 and 250 ms. Beyond these limits it becomes harder to make the adjustments. Their limits are thus the same as ours.

A comparison with other reported data is rather hazardous because of the differences in paradigms involved. The detectability of a short gap differs from the perceptibility of the continuity effect of tones with a duty cycle of 50%.

Thus, our conclusion remains that the continuity effect can be observed at rates between 2 Hz and 10 Hz (presentation times between 250 ms and 50 ms). Beyond these limits it will be difficult to hear continuity.

Next we varied the duty cycle of the signals. Again we did not introduce any gap between the presentations, so the sum of the duty cycles for the two stimuli was always 100%.

We found a clear continuity effect at presentation times longer than about 40 ms. If one of the presentations was shorter it was, again, difficult to separate the pulsation of one tone from that of the other. It seems that the lower limit for observation of the continuity effect can be generally given as a presentation time of about 50 ms for either one of the stimuli.

A similar result was found for the upper limit. At presentation times of more than about 250 ms, the perception of continuity becomes less clear.

The determining factor thus seems to be the presentation time rather than the alternation rate. For stimuli of equal duration, the lowest permissible rate is about 2 Hz. With a presentation time of 75 ms for one stimulus, the lowest rate is about 3 Hz. This rate depends only slightly on whether the continuous sound is induced by the longer or by the shorter stimulus. The condition with the sound heard as continuous presented during the longer period is a bit more favourable for continuity perception.

It is difficult to understand the implications of this finding for the work of Thurlow and coworkers and Elfner and coworkers. They measured the effect of one signal on the detectability of a small gap in the other signal. The results presented were measured in conditions above our pulsation threshold, otherwise they would have found much longer minimum detectable intervals. Only Elfner (1971) seems to have used a condition close to the pulsation threshold when he found a rather long continuity threshold of 90 ms near 300 Hz (his Fig. 2). This shows that their continuity threshold and our pulsation threshold are based on different paradigms although the effects may be related.

Concluding, we emphasize that continuity is well perceived if the presentation time of either stimulus is at least 50 ms and is not more than 250 ms.

We used equal durations for each stimulus; this makes a conversion for equal loudness unnecessary. It is a well known fact that our auditory system integrates the energy of a stimulus in time and in frequency (see e.g. Zwicker and Feldtkeller, 1967; Plomp, 1961). This time integration implies that a longer sinusoidal signal sounds louder. This may introduce the need for a conversion such as is generally accepted for postmasking (see Houtgast, 1974 a, p.30, 31 and Fig. 5-1).

11-2. The continuity effect with speech

Interrupted running speech can also be made to sound continuous by intervening noise (Miller and Licklider, 1950). We have seen that the continuity effect is closely linked with the ear's frequency selectivity. Since speech signals have a broad spectrum, it will be interesting to see whether the spectral qualities of the speech also appear in the pulsator signal. This would be expected if the ideas about continuity presented in section 1-4 were true, i.e. if continuity is perceived when the transition from pulsator to probe does not lead to a drop of activity in any auditory channel. This means that the activity pattern caused by the noise covers the activity patterns of the speech.

We tested this hypothesis by determining pulsation thresholds with running speech as a probe signal and different pulsator signals. The repetition rate was again 4 Hz, and signals were presented monaurally.

No continuity effect could be observed with sinusoidal pulsators. At very high pulsator levels, some kind of continuity effect was found, but it was very difficult to understand the speech: the sine wave tended to mask the speech by pre- and post-masking effects.

Next we used three types of noise, generated by a GR-1382 random noise generator. This generator produces a white noise (constant energy density over the total audio range), a "pink" noise (energy density falling off at a rate of 3 dB/oct) and a kind of noise resembling the long-term average of speech, called USASI*-noise, obtained from white noise by low-pass filtering with a 100 Hz cut-off frequency and high-pass filtering with a cut-off frequency of 320 Hz; both cut-off slopes are 6 dB/oct. We made spectrograms of speech and compared them with those of USASI noise. The two were alike, except at very low frequencies.

We compared the power levels of the three types of noise at the pulsation threshold. We found the lowest levels for USASI noise and the highest for white noise. The same order was found in comparing the loudnesses of the three types of noise.

* USASI is an acronym for United States of America Standards Institute, the body responsible for standardization of this USASI noise. The noise is shaped according to USASI standard S 1.4

The threshold level of the white noise was so high that the loud noise seemed to reduce the intelligibility of the speech by masking it.

We filtered out some frequency bands of the noise. Now some transitions were heard and some were not. Full continuity was not reached.

The differences in pulsation level can be understood in terms of the critical band concept (Fletcher, 1940). Studies of the critical bandwidth (Zwicker, 1961; Plomp, 1964 b; Zwicker and Feldtkeller, 1967) have shown that this quantity is roughly constant at frequencies below 500 Hz, and roughly proportional to frequency for frequencies above 1 kHz.

A comparison of the spectra shows that the pink noise is adjusted to a level where the mid-frequency-range energy density of noise and speech are about equal. This will result in too high a spectrum level at the low frequencies and to a smaller extent at the high frequencies. These low frequencies contribute to the loudness and thus cause the pink noise to be louder at the pulsation threshold.

For white noise the density level has to be adjusted to match the low-frequency density of the speech. Since white noise has a flat spectrum, it will cause the high-frequency density to be much too high, thus requiring a very loud noise at the pulsation threshold. The difference in slope between white noise and speech (6 dB/oct) leads to very unfavourable S/N ratios at the high frequencies.

A comparison of the USASI level and the speech level showed that a S/N ratio of about -10 dB was necessary for perceiving continuity. This ratio can be explained by the stronger irregularities of the speech signal both in time and in spectrum. We also found that the intelligibility of speech seems to fall for S/N ratios lower than -20 dB, probably due to masking. Powers and Wilcox (1977) found a lower speech intelligibility for S/N ratios of less than -24 dB and higher than -6 dB. They explained the lower limit as due to masking, just as we do, but gave no explanation for their upper limit. Given our continuity threshold of about -10 dB, their upper limit could well be interpreted as the continuity threshold.

We concluded that the pulsator spectrum has to contain the probe spectrum if the continuity effect is to be heard.

The continuity effect for speech sounds very striking. Interrupted speech sounds very unnatural, and is difficult to understand. As soon as the noise is introduced, the speech sounds very natural and seems to be much more intelligible. This aspect will be discussed further in Chapter IX.

11-3. The continuity effect with music

The most clear, striking demonstration of the continuity effect can be given with music, although the various aspects of music such as melody, rhythm etc. make a general description difficult.

Again we used a 4 Hz repetition rate. The pulsator signal was USASI noise with its level adjusted to about 15 dB above the average level of the loudest passages of music. Without intervening noise, the interrupted music sounds very distorted and unmusical. The melody can sometimes be recognized, but it is difficult to follow. The addition of noise in the gaps restores the normal perception of music. It sounds continuous and melodious and it seems much easier to follow the melody.

We saw in the previous sections that spectral matching of pulsator and probe is essential for the continuity effect with pure tones and speech. This is also the case for music. The various instruments require different pulsator levels. High-pitched instruments or instruments with a rich harmonic spectrum need higher levels for the continuity effect than low-pitched instruments with a pure-tone character. Different passages will therefore need different pulsator (noise) levels. The differences can easily be explained by the different spectra. We made spectrograms of the music and compared them with the spectrograms of the noise. The differences in level are a matter of matching the spectral densities.

A melody is easily followed, and even vibratos are heard to continue throughout pulsator bursts once the vibrato has been detected during the periods of presentation of the music. Rhythm can be severely distorted by the interruptions and is often not restored by the intervening noise. This is true of the melody rhythm, but is even more so for the rhythmic instruments such as percussion in popular music. It seems that we perceive the full extent of percussion sounds if there is some evidence during the

probe interval that such a percussion sound is or will be appearing. If there is no cue of its presence, a note on a percussion instrument can be totally missed. This kind of rhythmic occurrence is thus either heard or completely missed.

This brief account will have made clear the difficulties encountered in the description of continuous perception of interrupted music as a result of temporal patterns. We shall not investigate this field any further because of these complications. However, the continuity effect can be most spectacularly demonstrated with music.

III. SPECTRAL EFFECTS ON PULSATION MEASUREMENT

III-1. Introduction

The literature on the pulsation threshold method was presented in Chapter I and some qualitative pilot studies on the continuity effect in Chapter II. We saw that the frequency selectivity of the auditory system is reflected in the results. We want to know more about this frequency selectivity, and shall be using the pulsation threshold method for this purpose. However, the switching from one stimulus to the other might interfere with the auditory selectivity. Switching of signals is a temporal occurrence which will be reflected in the spectrum of the signals. If the spectrum gets so wide that its components can be resolved by the frequency selective mechanism, we will be measuring our input signal and not the auditory filter. We thus have to deal with two important properties:

- 1 Spectral broadening. The shorter the presentation time of a signal, the wider is its spectrum. The width of the spectrum is also determined by the amplitude envelope during transition. We shall derive these spectra analytically and compare their widths for different presentation conditions. In general we can say that the longer the signal, the less likely is it that spectral effects will appear in the pulsation patterns.
- 2 Temporal overlap of the stimuli. A gradual transition from one stimulus to the other will limit the frequency spread, but will introduce temporal overlap. The simultaneous presentation of two sine waves causes the generation of interaction products (see

page 16). The overlap should thus be minimized.

A compromise between the opposing demands has to be found, so we need some knowledge of their respective effects on the patterns. We therefore measured pulsation patterns for various values of the presentation parameters and checked whether spectral effects are found and for what values the influence on the patterns is negligible.

III-2. Spectra of stimuli

Monaural pulsation patterns were determined for a fixed pulsator. The presentation parameters considered are the repetition rate of the stimuli, quantified by the presentation time T of one stimulus, the shape and time constant τ of the transition function. We used two different smoothing functions: a Gaussian transition amplitude envelope and an exponential amplitude envelope. Fig. 1 shows the envelopes and

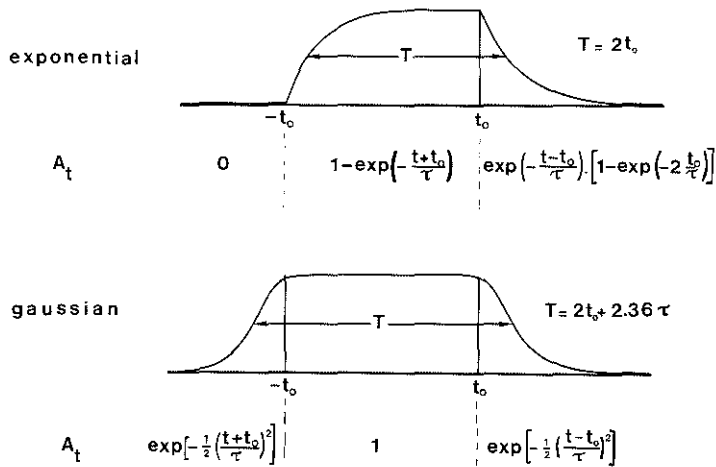


Fig. 1. Smoothing functions with their equations.

their mathematical equations for various time intervals. The presentation time T is defined as the time between half-amplitude points. Its equation is also given.

From the equations of Fig.1 we can derive the spectrum of a sinusoidal stimulus presented with the given amplitude envelope. The analytic derivation is presented in

Appendices. At this stage we only present the results.

The power function derived for an exponential envelope is (Appendix A):

$$W(\Delta\omega) = \frac{A_1^2 t_o^2}{2\pi} \cdot \frac{\sin^2(\Delta\omega t_o)}{(\Delta\omega t_o)^2} \cdot \frac{1}{1 + (\Delta\omega\tau)^2} \quad (1)$$

where $\Delta\omega$ is the frequency difference between signal and carrier, expressed in radians, and A_1 is the amplitude of the signal. This power spectrum is a product of three factors: the total energy, the energy distribution of a rectangularly gated signal and the smoothing factor. The last-mentioned factor will be denoted by A_e' and is equivalent to the effect of a low-pass filter, with Laplace transform $(1 + \tau s)^{-1}$.

The power function of the signal with a Gaussian transition envelope has also been derived (Appendix B). For the sake of clarity we give the Fourier transform of the signal:

$$F(\Delta\omega) = \frac{A_1 t_o}{\sqrt{2\pi}} \frac{\sin(\Delta\omega t_o)}{\Delta\omega t_o} \left(1 - 2 \frac{\Delta\omega\tau}{\sqrt{2}} D\left(\frac{\Delta\omega\tau}{\sqrt{2}}\right) \right) + \frac{A_1\tau}{2} \exp\left(-\frac{\Delta\omega^2\tau^2}{2}\right) \cdot \cos(\Delta\omega t_o) \quad (2)$$

The first two factors of the first term again represent the energy of the signal and the distribution of a rectangularly gated signal. The Gaussian transition envelope appears as a smoothing factor, to be denoted by A_g' , and as an additional spectral term representing the sum of the isolated Gaussian half-pulses at the moments of switching on and off. The symbol D in this factor represents Dawson's integral (Abramowitz and Segun, 1968).

We see from the equations that the substitution of a generalized coordinate $x = \frac{\Delta\omega\tau}{\sqrt{2}}$ simplifies the A' -functions. Next, we transform the factors to decibels and get:

$$A_e = 10^{10 \log A_e'} = -10^{10 \log (1 + 2x^2)} \quad (3)$$

$$A_g = 20^{10 \log A_g'} = 20^{10 \log (1 - 2xD(x))} \quad (4)$$

The Gaussian spectrum can also be expressed in terms of this coordinate x . We then get for the second term of the Fourier transform of the Gaussian transition function:

$$A_1 \frac{T}{2} e^{-x^2} \cos(\Delta\omega t_0) = A_1 \frac{T}{2} G'(x) \cos(\Delta\omega t_0) \quad (5)$$

The smoothing factors A_e and A_g were computed and are shown in Fig.2. The Gaussian spectrum is also shown. To facilitate the comparison we have also given the frequency separation in Hz for $\tau = 3.6$ ms and $T = 125$ ms. On this latter axis we have also plotted the pulse spectrum (S_p). The actual spectrum is the sum of the S_p and the smoothing factor.

One more curve (A_s) is shown in Fig.2; this is the smoothing factor for a sinusoidal transition function as used by Houtgast. We derive the spectrum equation in Appendix C. Using the same representation as in the lower part of Fig.1, we get $T = 2t_0 + \tau_0 = 2t_1$. There are five time intervals: $t < -(t_0 + \tau_0)$; $-(t_0 + \tau_0) < t < -t_0$; $-t_0 < t < t_0$; $t_0 < t < t_0 + \tau_0$ and $t > t_0 + \tau_0$

with

$$A_t = 0; \frac{A_1}{2} \left(1 + \cos \frac{\pi}{\tau_0} (t + t_0) \right); A_1; \frac{A_1}{2} \left(1 + \cos \frac{\pi}{\tau_0} (t - t_0) \right); 0$$

The power spectrum derived is given by

$$W(\Delta\omega) = \frac{A_1^2 t_1^2}{2\pi} \cdot \frac{\sin^2 \Delta\omega t_1}{\Delta\omega^2 t_1^2} \cdot \cos^2 \left(\frac{1}{2} \Delta\omega \tau_0 \right) \cdot \left(\frac{1}{1 - \frac{\Delta\omega^2 \tau_0^2}{\pi^2}} \right)^2 \quad (6)$$

This power spectrum is again a product of the total energy, the distribution of a rectangular pulse and a smoothing factor denoted by A'_s . We define $\tau = \tau_0/\pi$ and $x = \frac{\Delta\omega \tau}{\sqrt{2}}$, and obtain

$$A_s = 10 \log \left(\cos \left(\frac{1}{2} \pi x \sqrt{2} \right) / (1 - 2x^2) \right)^2$$

This function has a singular point for $x = \frac{1}{2}\sqrt{2}$. We can compute its limit at this point by approaching it from the left and from the right. The two limits converge to $A_s = -2.09$.

The attenuation factor including this singular point is plotted in Fig. 2.

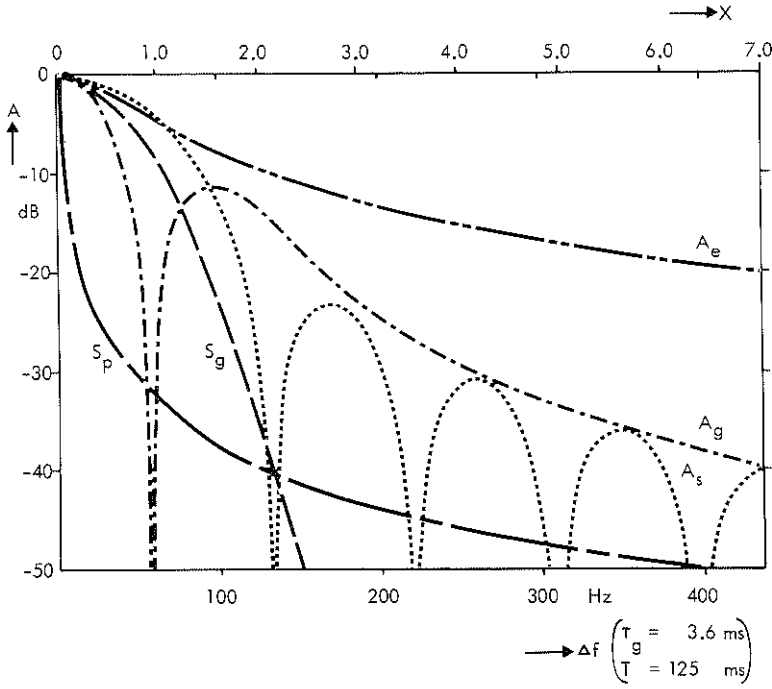


Fig. 2. Attenuation curves for various smoothing functions

- A_e exponential envelope
- A_g Gaussian switching envelope
- A_s sinusoidal switching envelope
- S_p spectrum of rectangular pulse
- S_g spectrum of Gaussian pulse

The spectrum has a strong cosine profile, but the tops are very close to the A_g curve.

For the comparison of the spectra, we computed their -20, -30, -40 and -50 dB bandwidths for a given 10-90% rise time of 6 ms (as used with the Gaussian function). These bandwidths are given in Table I. The top row of this table gives the real rise time constant corresponding to a 6 ms rise time.

Table I. Comparison of spectra for $\tau_{0.1-0.9} = 6$ ms			
	Gaussian	Sine	Exponential
τ (ms)	3.55	3.24	2.73
bandwidths (Hz) at			
-20 dB	12	12	12
-30 dB	29	36	37
-40 dB	41	73	73
-50 dB	105	105	150

It will be clear from Fig.2 and Table I that three regions are to be distinguished. For a separation between 0 Hz and about 25 Hz (-25 dB), the three spectra are much alike. Between about 25 Hz and 90 Hz the sine spectrum and the exponential spectrum are much alike (down to about 45 dB); the Gaussian spectrum is much narrower. From 90 Hz on, the Gaussian and sine spectra are more alike, apart from the lobes of the latter; the exponential spectrum is much wider. Houtgast used a τ_0 of about 20 ms. This gives a τ of 6.4 ms and a $\tau_{0.1-0.9}$ of 11.8 ms. This is a much longer time constant and thus gives a much narrower spectrum than our Gaussian time constant of 3.55 ms, corresponding to a $\tau_{0.1-0.9}$ of 6 ms.

III-3. Presentation conditions

The repetition rate, characterized by the presentation time T was varied between 12.5 Hz ($T = 40$ ms) and 2 Hz ($T = 250$ ms). The presentation time of 40 ms is just below the lower limit mentioned at the end of section II-1. It was used in an attempt to gain more precise information about the lower limit for observation of the continuity effect. The other values cover the entire range of continuity perception.

The rise time constant of the exponential envelope was 10 ms (10 - 90% rise time of 22 ms). This corresponds to just no audible click.

For the Gaussian transition function we used 10-90% rise times of 1.5, 3 and 6 ms

(rise time constants of 0.9, 1.8 and 3.6 ms). These conditions result in the perception of a strong click, a faint click and no audible click respectively.

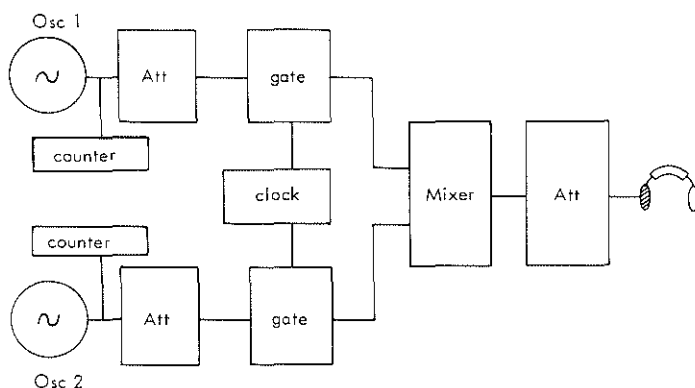


Fig. 3. Block diagram of our equipment

III-4. Apparatus, observers and methods

The block diagram of the equipment is shown in Fig.3. The stimuli were generated by HP 204 C generators. The frequencies were read from a GR 1192 B counter. The levels were checked with a B & K 2409 voltmeter. The stimuli were attenuated by HP 350 D attenuators. Two different sets of modulators and headphones were used.

The exponential amplitude envelope was generated by the switching of a small lamp. The light from this lamp fell on a light-dependent resistor which was part of a common-base amplifier. The system had a very high damping (over 85 dB) and a low noise level (less than -80 dB). The harmonic distortion was less than -50 dB. We used Sharpe HD 10 A and Beyer DT 48 S headphones.*

* The experiments with this equipment were performed at the Department of Medical Physics of Utrecht State University.

The Gaussian transition function was generated by shaping a trapezoid in a diode network. This signal was the source current of a differential amplifier. The system had a damping of about 70 dB and a signal-to-noise ratio of about -80 dB (open circuit). Harmonic distortion was less than -50 dB. We used Grason and Stadler TDH-39 headphones with this equipment.

The presentation time of the stimuli was defined as the time between the half-amplitude points; the gaps were defined as the time elapsed between the half-amplitude points. This definition has the additional advantage that we get minimum energy modulation when there is no gap between the stimuli. This has been checked by using only one oscillator for generating both signals and inspecting the amplitude envelope.

Three observers participated in the experiments. Their ages were between 20 and 30 years. They had no history of otologic diseases, their audiograms were normal and they were familiar with the background of the experiments. They were given ample time to familiarize themselves with the continuity effect. The standard deviation of the adjustments generally fell to a low, stable value after some 10 practice sessions of about $1\frac{1}{2}$ hour each. All results presented were collected after this practice period.

We used an adjustment method. All points were measured five times in a different order. Mean values and standard deviations were computed.

III-5. Results

The pulsation patterns for one observer and various presentation times using an exponential envelope are presented in Fig.4. The symbols show mean values. If the standard deviation is larger than the width of the symbol, this is indicated by a thin horizontal bar through the point. We see that the various patterns are more or less parallel and tend to shift downwards with increasing presentation time between 100 and 200 ms. At shorter presentation times, the patterns show a bulge. The spectra

of section III-2 are represented by heavy lines and symbols for two values of T .

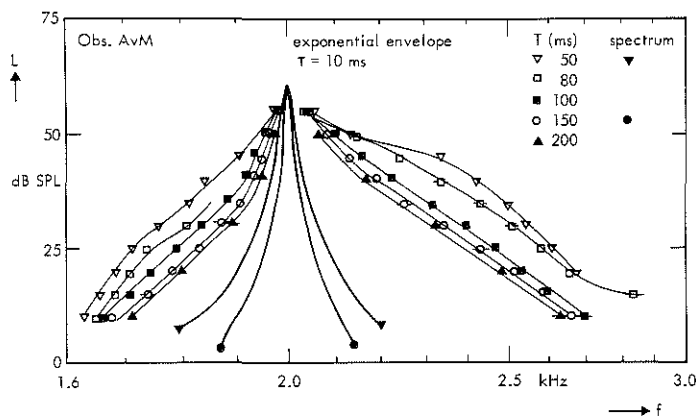


Fig. 4. Pulsation patterns and spectra for various presentation times and for an exponential amplitude envelope. The spectra are plotted as heavy lines.

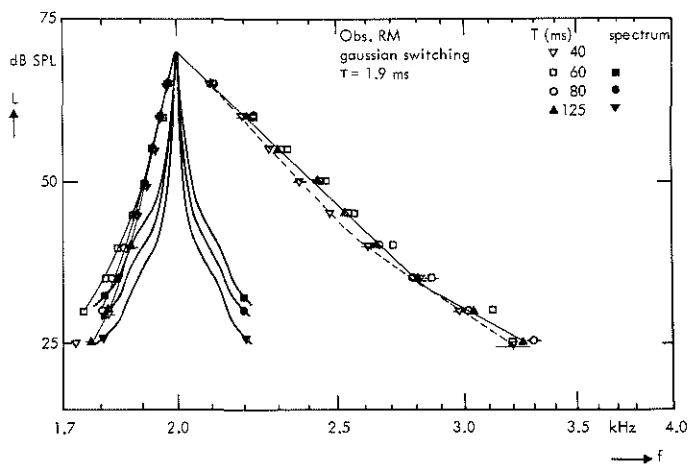


Fig. 5. Pulsation patterns and spectra for Gaussian transition functions as a function of presentation time. Spectra are plotted as heavy lines.

Fig.5 shows the results for various presentation times for a Gaussian transition envelope with $\tau = 1.9$ ms.

At the high-frequency side the points for the various presentation times coincide almost perfectly, except for the results for 40 ms which are shifted slightly towards the centre frequency. This tendency is also observed at the low-frequency side.

If we leave the 40 ms results out of consideration, we see that the points at the low-frequency side show a certain spread; the widest pulsation pattern is found under the conditions giving the widest spectrum. The same trend is observed at the high-frequency side. This fact suggests a spectral cause for the broadening of the pulsation patterns.

The deviation of the 40 ms results may be explained in terms of a detection shift. We have already seen that the observers reported difficulties in observing the continuity effect for presentation times of less than 50 ms (section II-1). This was also reported by the observer used for determining the data presented in Fig.5.

Fig.6 shows the effect of the rise time on the pulsation pattern. The patterns were measured with a presentation time of 60 ms and rise time constants of 0.9 ms, 1.8 ms and 3.6 ms.

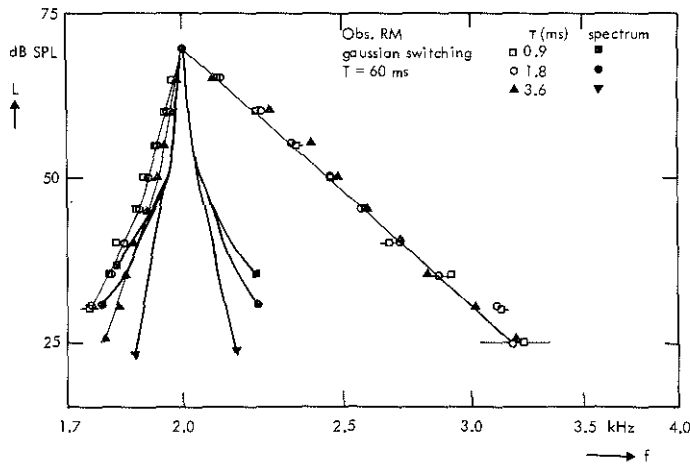


Fig. 6. Pulsation patterns and spectra for various rise time constants, Gaussian switching envelope and presentation time of 60 ms.

The spectra computed from the equations of the section III-2, are shown in heavy lines.

We see that the three measured patterns differ at the low-frequency side; the widest pattern is found under the conditions giving the widest spectrum. There is a clear difference between the 3.6 ms and 1.8 ms patterns but only minor differences between the 1.8 ms and 0.9 ms patterns. This may be due to interference between click perception and continuity perception.

We also measured patterns for presentation times of more than 100 ms. At these presentation times the patterns for the various time constants all coincided with the 3.6 ms pattern of Fig. 6. The absence of an effect of rise time at longer presentation times can be understood from the spectrum. At longer presentation times the stationary part of the spectrum contains more energy, while the energy contents of the transient part (seen as a shoulder in the spectrum) remains constant. This shows up as a downward shift of the shoulder in the spectrum and results in a narrower spectrum, the sidelobes of which cannot evidently be detected for presentation times longer than 100 ms.

Next we compare the patterns for different envelopes. Fig. 7 shows patterns for a presentation time of 125 ms. The two envelopes considered are those of Fig. 1: an exponential envelope and a Gaussian transition envelope. A rise time constant of 10 ms was chosen for the exponential envelope and one of 3.6 ms for the Gaussian transition envelope. This just avoids an audible click.

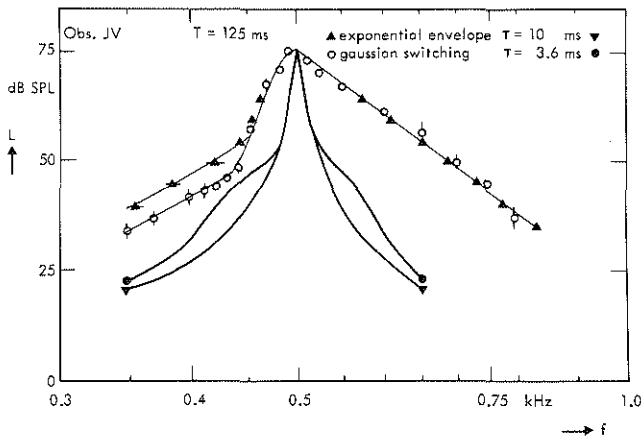


Fig. 7. Pulsation patterns and spectra for two smoothing functions: an exponential amplitude envelope and a Gaussian transition envelope. The rise time constants are chosen so that an audible click is just prevented.

We see that the high-frequency edges coincide. The low-frequency edges, however, differ considerably. In this case the wider pattern occurs under the conditions giving the narrower spectrum, so long-term spectral effects cannot account for the differences observed. A possible explanation is suggested by the observation that clearly audible combination tones were heard when the exponential envelope was used. These combination tones were not perceived with the Gaussian transition function for τ 's up to 3.6 ms nor for a sinusoidal transition function with $\tau = 6.4$ ms (Houtgast's condition).

An essential feature of combination tones is that they are audible only if the two primaries are presented simultaneously. This leads to the supposition that an overlap in the presentation might be responsible for the differences between the patterns. We, therefore, measured patterns with gaps between the presentations, using a Gaussian transition envelope. The results are shown in Fig.8. The gap parameters

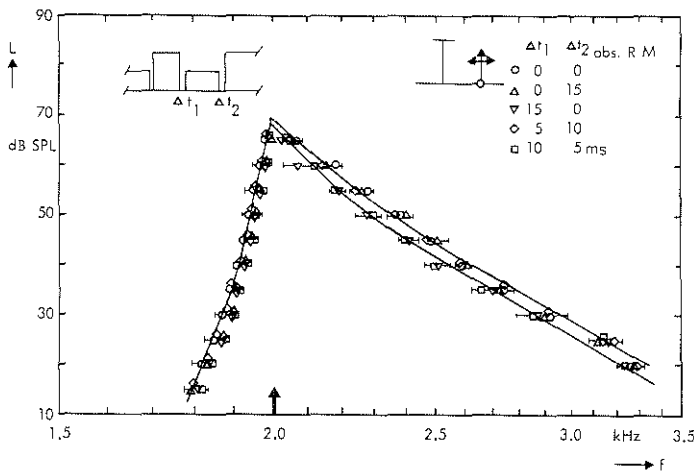


Fig. 8. Pulsation patterns for a number of gap conditions indicated in the figure.

are specified in the form $(\Delta t_1, \Delta t_2)$, where Δt_1 is the gap between pulsator and probe stimulus, and Δt_2 is that between probe and pulsator stimulus. Gap intervals are defined as the time between the half-amplitude points and are expressed in milliseconds. This means that the gap is zero under our normal conditions. We measured patterns for the gap parameters (0,15), (5,10), (10,5) and (15,0). We see that the patterns for (0,0), (0,15), and (5,10) are identical; there is no effect of overlap nor of decay. The patterns for (10,5) and (15,0) show a small decay effect at the high-frequency side.

III-6. Discussion

The experiments described above confirm our preliminary finding that a 50 ms presentation time forms a lower limit for observation of the continuity effect. Below this limit, the observer is less sure about whether he hears continuity, and he tends to shift the experimental setting to one where he is more convinced of hearing continuity. This yields narrower patterns.

Apart from this shift, we also found spectral effects in the patterns. These effects are found for presentation times up to 100 ms, so we can only use presentation times between 100 ms and 250 ms if we want to avoid these spectral effects. Houtgast used a presentation time of 125 ms, which is thus well within the usable range. We shall be using the same presentation time of 125 ms in our further experiments unless otherwise stated.

We also found spectral effects related to the rise time constant. At presentation times longer than 100 ms, the patterns for various rise times are identical. At all presentation times, the 3.6 ms patterns are identical but the patterns widen for presentation times shorter than 100 ms and rise time constants less than 3.6 ms. We shall be using a rise time constant of 3.6 ms in our further experiments.

An unexpected effect was found in the results presented in Fig. 7. Here the widening was in a direction opposite to that of the spectral effects. The exponential envelope has a narrower spectrum but it results in a wider pattern. Two explanations were considered: short-term spectral effects and interaction phenomena.

We computed long-term spectra. However, the ear analyses the signals applied to it, with a certain time constant. The discontinuity in the derivative of the exponential envelope will thus cause a momentary widening of the spectrum: the shorter the time constant, the wider the spectrum. This means that the effect can also be explained as a spectral effect. It has no further consequences except that the exponential envelope is unsuited for the pulsation threshold experiments.

An interaction phenomenon involving two simultaneously presented sine waves is the perception of combination tones. These tones may affect the pulsation patterns. As we have mentioned, combination tones are clearly audible with the exponential

envelope. We wondered why this is so when no combination tones could be heard with the Gaussian transition function.

The stimuli were presented at 70 or 75 dB SPL, which is about 60 dB above threshold. Now the generation of combination tones requires the simultaneous presence of the two primaries –even at low intensities (Smoorenburg, 1972). Under our experimental conditions, therefore, combination tones can only be generated during the overlap of the different envelopes. We defined the overlap as the period of time during which the two components are physically presented simultaneously at a level higher than -60 dB with respect to the maximum level. For the exponential envelope ($\tau = 10$ ms) the overlap is 70 ms, about half the presentation time. For the Gaussian transition envelope on the otherhand, this period is 18 ms (Houtgast's sine transition, $\tau_0 = 20$ ms has an overlap of 19 ms) which is considerably shorter. Still, combination tones could be generated.

Fig. 8 shows that a gap of 5 ms after the test stimulus (gap parameters (5,10)) does not affect the patterns, although the overlap is reduced to 13 ms. Conditions with longer gaps (e.g. gap parameters (10,5)) cannot be used for the comparison, because decay effects start to play an appreciable role. The similarity in shape between the patterns does not suggest any great change in presence or absence of combination tones.

With negative gaps of -10 ms, clear combination tones are heard. The threshold is reached for gaps of about -5 ms.

We see that with a Gaussian envelope we are trapped between combination tones and decay effects. Our choice of normal presentation conditions with no gaps between the stimuli represents a compromise at which it is unlikely that combination tones influence the patterns. These tones may well have influenced the patterns measured with exponential envelopes.

Decay effects will complicate the interpretation of patterns determined with silent intervals between the stimuli, as was done by Kronberg et al. (1974) and Schreiner et al. (1977). Their results do show decay effects. Houtgast's results seem compatible with ours. His overlap is not significantly longer than ours and his spectrum is narrower than ours.

The results presented by Fastl (1975) show a very similar spectral widening of pulsation patterns. These results must, however, be treated with some reserve because two of his three presentation times are outside our range ($T = 30$ ms and $T = 300$ ms) so it is possible that criterion shifts as discussed in connection with Fig. 5 for short presentation times have influenced his results. The steepness of his patterns will be discussed in section VI-6.

Most of the results of Schreiner et al. (1977) cannot be compared directly with ours. They used a different set of parameters and variables (see section IV-2 for a detailed survey of possible pulsation patterns). The main difference between their presentation conditions and ours, however, is the variable gap of at least 12.5 ms they introduced between the stimuli. We have seen in our experiment that gaps of 10 ms already produce decay effects.

The authors discussed the decay effects on their figures in detail and concluded correctly that gaps longer than 10 ms between consecutive pulses complicate the interpretation of the patterns. Furthermore they suggested the use of gradual transients with time constants between 7 and 12 ms. We came to very similar conclusions, but we did not use gaps between the stimuli because pulsation patterns measured with small gaps (5 ms) do not differ from patterns measured without gaps. There is thus no reason to assume that the small overlaps have any effect on the patterns.

IV. PULSATION PATTERNS

IV-1. Introduction

After having established the conditions for the determination of pulsation patterns, we will now focus our attention on the frequency selectivity as it appears in these patterns. We are particularly interested in the dependence of the frequency selectivity on level or in other words in its linearity. Nonlinear behaviour implies that the various patterns that have been published are not mutually comparable because they are determined using probes and pulsators with different characteristics (cf. Houtgast, 1974 a; Schreiner et al., 1977).

We shall present a classification of the various methods used to determine masking and pulsation patterns (Rodenburg et al., 1974; Verschuure et al., 1974) and compare the methods with those used in physiological measurements. We will then describe our experiments on pulsation patterns with reference to the linearity.

IV-2. Classification of methods of determination of masking and pulsation patterns

The methods used for determination of masking patterns and pulsation patterns are very similar. Both the pulsator and the masker cause a distribution of activity, while the probe and maskee serve as indicators of the level of this distributed activity. The masking and pulsation patterns determined can thus be interpreted in very similar ways. The main differences between the methods are the presentation conditions and the threshold criteria. In simultaneous (direct) masking the two signals

are presented simultaneously and the threshold criterion is the detectability of the maskee. In the pulsation method the two signals are alternated and the criterion is continuity. It is assumed for the purposes of interpretation that the probe/maskee indicates the distributed activity at the frequency in question, i.e. that it represents the internal output level at some stage of the auditory system. The input signal is the pulsator/masker which is causing the distribution.

Using sine waves for the two signals, we get four variables: the two levels and the two frequencies of the sine waves. If we determine one of these variables as a function of a second we still have two parameters. Four variables can be combined in six different ways. Four of these six combinations are frequency-level dependences, which can be interpreted as patterns reflecting the ear's frequency selectivity. These four combinations are shown in Fig. 9. The small circle indicates the maskee/probe; arrows indicate variability of either frequency (horizontal arrows) or level (vertical ones).

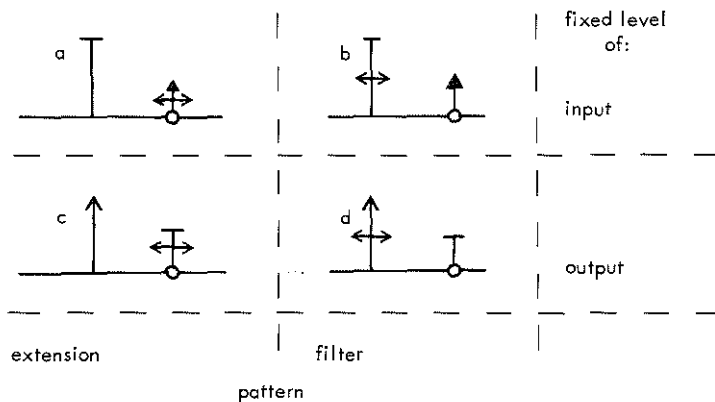


Fig. 9. Four possible methods for the determination of patterns reflecting the frequency selectivity. Arrows indicate the variability of a parameter. The small circle indicates the probe stimulus.

- method a. input extension pattern
- method b. input filter pattern
- method c. output extension pattern
- method d. output filter pattern

Wegel and Lane (1924) used a fixed masker and determined the distribution of activity with a scanning maskee. This is method a, in which one measures the extent of the

masker's activity pattern. We will call this method the determination of an extension pattern for a fixed input level. Chistovich (1957) calls it the direct method. It has been adopted by many investigators, specifically if complex maskers are used.

Another method, encountered in the masking literature under a variety of names, is method d. Now the maskee is kept constant and the masker is adjusted. In our terminology we have a fixed output level. The input signal is swept in frequency and adjusted to a level that produces an activity satisfying the criterion chosen at the maskee frequency. We call this method the determination of the filter pattern for a fixed output level. Chistovich (1957) calls it the threshold method. Rodenburg et al. (1974) call it the iso-response method. The measured patterns are also called iso- L_p -curves (Vogten, 1974) and psychophysical tuning curves (Zwicker, 1974).

Rodenburg et al. (1974) have also used method b, which is called the iso-intensity method in physiology. Here the output level at a fixed frequency is determined with a swept input signal of fixed level. We call the pattern determined in this way the filter pattern for a fixed input level, because it is the usual way of determination of filter characteristics in physics.

The remaining one of our four methods is method c, which yields what we call the extension pattern for a fixed output level. For the sake of brevity, we will generally refer to these various patterns as output extension pattern etc.

A similar classification can be applied to pulsation threshold methods. Houtgast usually measured output filter pulsation patterns. Fastl measured input extension pulsation patterns, as did we in Chapter III. The pulsation patterns of Kronberg et al. (1974) and most of the patterns of Schreiner et al. (1977) are also output filter patterns.

The relation between the patterns, measured by different methods is quite simple if the system is linear. Input patterns will be output patterns upside down; filter patterns will be extension patterns reversed from left to right. The relations will be more complicated if the system is nonlinear (Chistovich, 1957; Rodenburg et al., 1974).

It is evident that the same names for the different methods can be used in pre- and postmasking. They can also be used in physiological experiments (Rodenburg et al., 1974) although we have only one stimulating sine wave in this case: the input signal. The output signal is a vibration amplitude or the number of spikes in a fibre.

We know that the hydrodynamical system of the cochlea acts as a mechanical filter: a certain place along the cochlear partition responds maximally to a certain frequency. The vibration at a certain place can thus be interpreted as an output signal at the frequency to which this place gives a maximum response. Place thus relates to frequency and amplitude to level.

This is even more clear in neurophysiology. All measurements include the determination of the best frequency, i.e. the frequency at which the lowest threshold of neural activity is found. This is our output frequency. The number of spikes then relates to the level.

If we accept this nomenclature (and thus this interpretation) we can make a direct comparison between physiological and psychophysical experiments.

For this purpose we start by considering the measurements of cochlear partition motion. Von Békésy (1960) observed with a microscope the vibrations of the Reissner membrane which he showed to vibrate in phase with the basilar membrane. Determinations of the vibration amplitude at a given spot on the membrane as a function of frequency gave input filter patterns. Observations of the vibration at other spots on the membrane too, permitted the construction of input extension patterns.

In the Mössbauer technique a radioactive source is placed on the basilar membrane. Here too, we confine one set of observations to one spot of the membrane and thus measure a filter pattern. Most of the time (Johnstone & Boyle, 1967; Rhode, 1971) the input level was kept constant, so an input filter pattern was determined.

In their capacitive probe technique Wilson & Johnstone (1975) placed a transducer in a fixed position, so they were also measuring input filter patterns.

Kohlöffel (1972 a, b, c) used a laser technique. He illuminated a section of the basilar membrane with laser light and looked for fuzziness of the image. He showed that fuzziness, related to a lack of coherence of the light, appeared if the vibration amplitude exceeds a certain threshold value. He observed the spread of this

fuzziness along the basilar membrane. This gives an output extension pattern.

We see that vibration measurements have been performed using widely different techniques. The results may only be compared directly if the system is linear. Wilson & Johnstone have shown that it is for their capacitive probe method. Rhode (1971) and Rhode & Robles (1974) found clear nonlinearities with the Mössbauer technique. It follows that a direct comparison of input and output patterns or patterns measured for different levels is not permissible unless corrections are made.

Single-cell recordings involve only two different types of methods, since multiple electrodes are not available yet: we only determine filter patterns, either for fixed input (Rose et al.: response areas) or for fixed output (Kiang, Evans: tuning curve or frequency threshold curve). Some authors have compared the steepness of these patterns, although clear nonlinearities have been reported. We will return to this problem in Chapter VII.

IV-3. Apparatus

We began our experiments on the equipment described in Chapter III using Gaussian envelopes. This equipment only permits frequency adjustments. Under some conditions this can give erroneous results (see section VI-3-2).

We later designed a system in which we could also vary the level. This system was a semi-automatic one. The attenuators were replaced by GSC 1284 programmable attenuators (step size 0.5 dB; maximum attenuation 128 dB). The two frequencies were monitored with GR 1192 B counters. A punched-tape output of all attenuator and frequency settings could be obtained by depressing a push-button.

The level of one stimulus could be varied by means of a spring-loaded knob which returned to a neutral position after it was released. Turning of the knob resulted in a change in level; the further the knob was turned, the faster the level was altered. In the neutral position, the level is constant. Turning the knob to the right caused the level to rise; turning it to the left caused the level to fall. The level is only altered when the gate for the signal in question is closed.

We also introduced an interruption of the probe. The probe was presented seven times in the absence of the pulsator, and then omitted once. This interruption proved useful in fixing the observer's attention on the probe. With its aid we discovered a phenomenon which proved very useful in making pulsation threshold adjustments. There is a small range of levels (about 2 dB) where the probe is perceived as not being fully continuous, but as fading away and returning in an oscillatory way. This phenomenon dies out after a few hundred milliseconds, depending on intensity level. The observers were instructed to listen for this modulation effect; this effect was considered as a sign that we were just above the pulsation threshold.

We tested for differences between this semi-automatic equipment and the older system by repeating some earlier threshold determinations. No change in the standard deviations was found, and the differences between the average values were not significant. The results obtained with both sets of equipment are, therefore, directly comparable.

We always used a presentation time of 125 ms ($f_{\text{rep}} = 4 \text{ Hz}$) and a Gaussian rise time constant of 3.55 ms.

Four observers participated in the experiments. They had normal audiograms and no history of otologic diseases. They were familiar with the background of the experiments. New observers were given ample time to familiarize themselves with the procedure.

IV-4. Results

Pulsation patterns were measured for at least three levels at the frequencies 0.5, 1 and 2 kHz. All measurements were repeated five times on different days. The measured values were averaged and are shown in the figures.

Fig. 10 a gives the two extension patterns for observer RM at 1 kHz, and Fig. 10 b the two filter patterns for the same observer, reflected in a vertical line through 1 kHz and using a logarithmic frequency scale. The horizontal or vertical bars through the experimental points are the standard deviations in the adjustments and indicate whether frequency or level was adjusted. If no line is visible the standard deviation

is smaller than the symbol. The reflection of the filter patterns was performed to facilitate the comparison of extension and filter patterns; without this, the steep and the shallow edges of the patterns would be interchanged.

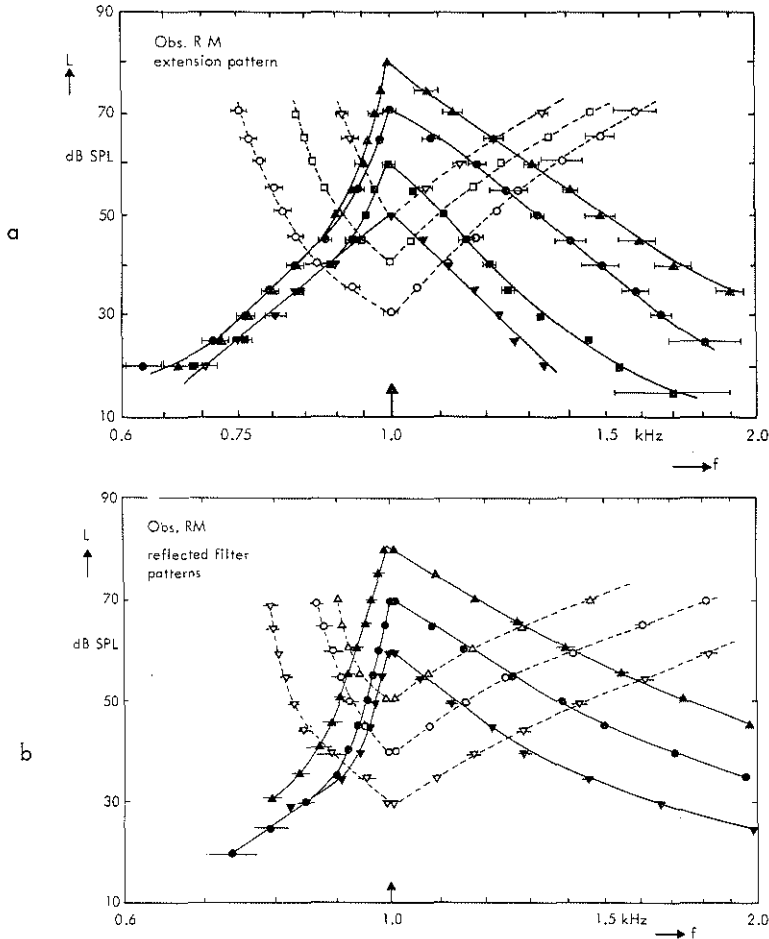


Fig. 10. Pulsation patterns for observer RM with a pulsator frequency of 1 kHz. The upper part of the figure (a) shows the extension patterns and the lower part (b) of the figure the filter patterns, reflected in a vertical line through 1 kHz and using a logarithmic frequency scale (full lines: input patterns; broken lines: output patterns).

Fig. 10 a shows some clear effects which were found in all results:

- Only the high-frequency edge of the input extension pattern is approximately a straight line; all other edges of the patterns are curved.
- The high-frequency edge of the input extension pattern becomes less steep at higher levels.

- The low-frequency edge of the input extension pattern becomes steeper at higher pulsator levels; at low probe levels the experimental points for different pulsator levels are sometimes not statistically different, which indicates a very strong nonlinearity.
- The high-frequency edges of the output extension patterns also show a nonlinearity; now the curves converge.
- The low-frequency edges of the output extension patterns show a divergent nonlinearity.

Similar effects are found in Fig. 10 b except that the steep (high-frequency) edges are generally steeper and the shallow (low-frequency) edges are generally less steep. This effect seems to be stronger for larger frequency separations from the fixed frequency (here 1 kHz).

In a linear system, input and output patterns would be identical in shape, except for reflection in a horizontal line. This can be understood by assuming that the response to a stimulus at a certain frequency is e.g. 10 dB lower than the stimulus itself (input method). This drop in output can be compensated by raising the input level by the same amount, 10 dB (output method).

Fig. 10 in fact shows a considerable difference between input and output patterns, in particular at the steep edges. These differences indicate the presence of a nonlinearity which makes it difficult to determine a pattern's steepness, because most of the lines are curved. Only the high-frequency edges of the extension patterns are straight lines. Their slopes depend on the level and range from 50 dB/oct at 80 dB SPL to 70 dB/oct at 50 dB SPL. We can fit this slope-level dependence with the linear equation: $S(L) = 0.78 L - 111$, where $S(L)$ is the slope in dB/oct and L is the pulsator level in dB SPL.

The curvature of the other edges of the pattern makes straightforward determination of their slope impossible, since this slope varies with frequency separation from the pulsator. We can however determine the initial slope of the input extension patterns near the peak, where the pattern is fairly close to a straight line. This slope can be extremely high. In Fig. 10 it ranges from 100 dB/oct at a pulsator level of 50 dB SPL to over 400 dB/oct at 80 dB SPL. The steep edge of the output patterns is even steeper. At low probe levels and high pulsator levels, this edge can have a

slope of more than 1000 dB/oct.

The shallow edge of the output patterns is also curved; however, its slope is always less than that of the input patterns.

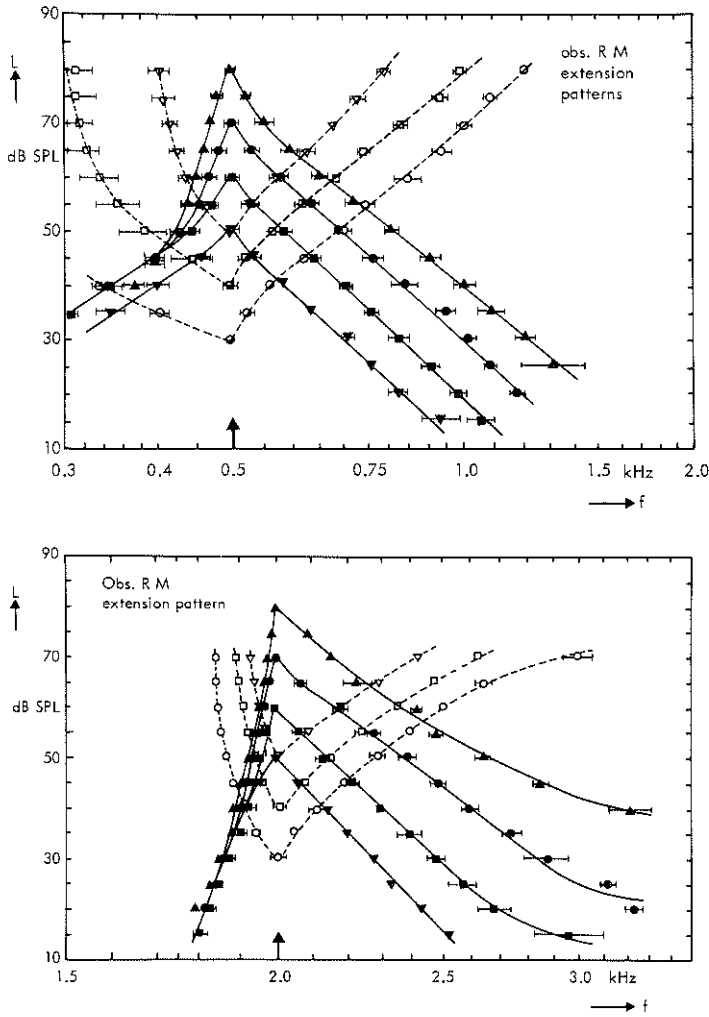


Fig. 11. Extension pulsation patterns for observer RM at 0.5 kHz and 2 kHz

This behaviour of the patterns, in particular the described curvature of the edges, is found for all frequencies as illustrated in Fig. 11 which shows the extension patterns for the same observer at 0.5 kHz and 2 kHz. The slope-level dependence of the

shallow edge of the input pattern is less pronounced at 0.5 kHz and more pronounced at 2 kHz. This frequency dependence of the nonlinear behaviour of the patterns may provide an explanation for the differences between extension and filter patterns. In the determination of the shallow edge of the input filter pattern the pulsator frequency, which is varied, is lower than the centre (probe) frequency. At lower frequencies the slope is less, so the pattern tends to level off, as is seen in Fig. 10 b. This frequency dependence also has its effect on the steep edge of the patterns. For the filter patterns the pulsator frequency is higher than the centre frequency. This will result in a steeper slope of the patterns. The effect is however slight, because the frequency separation between the sine waves is small as a result of the extreme steepness of the patterns.

Fig. 12 shows the filter patterns for observer RM at 0.5 kHz and 2 kHz. We see the same effects as described above.

Filter patterns are not the most appropriate patterns for studying nonlinearity, as they show both the nonlinearity and its frequency dependence at the same time. We shall, therefore, only give the extension patterns for the other observers.

Fig. 13 shows these patterns for observer JV. The steep edges of his input patterns are steeper than those of observer RM. This implies a stronger nonlinearity leading in turn to extremely steep edges in the output patterns. The steep edges of these patterns are practically vertical. Under these conditions, there is little point in determining the value of the slope, which reflects the strong nonlinearity and has a less strong connection with the frequency selectivity. In general, this observer shows a nonlinearity and a frequency dependence of the nonlinearity similar to the ones described for observer RM.

Fig. 14 and Fig. 15 give the patterns for observer MB and JK respectively. These also show the same general behaviour as described above, although the steepness of the patterns varies.

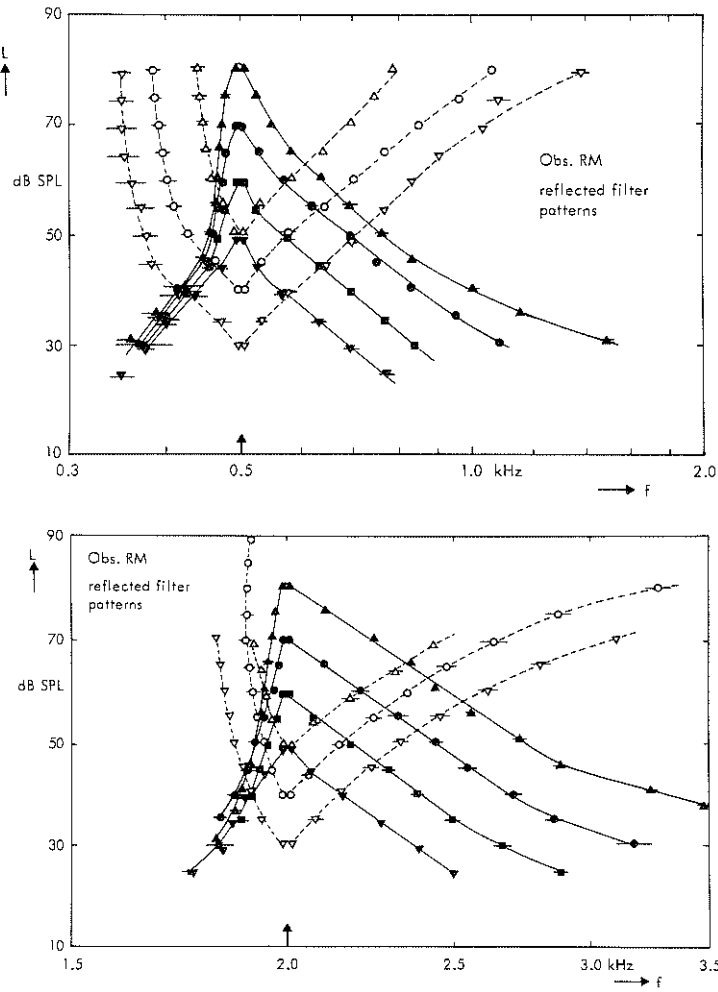


Fig. 12. Filter patterns for observer RM at 0.5 kHz and 2 kHz. The patterns are reflected in a line through the centre frequency, using a logarithmic frequency transformation.

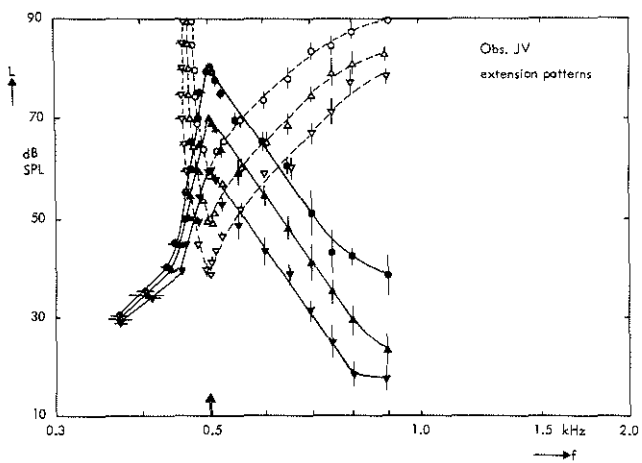
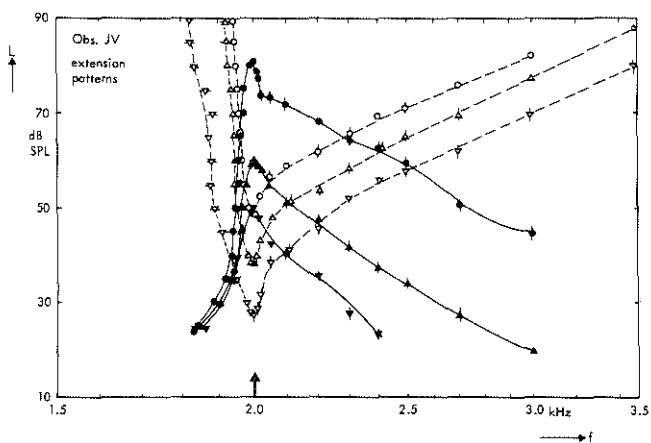
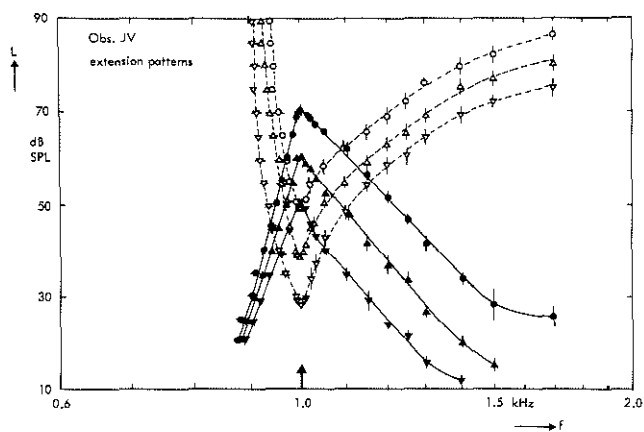


Fig. 13. Extension pulsation pattern for observer JV.



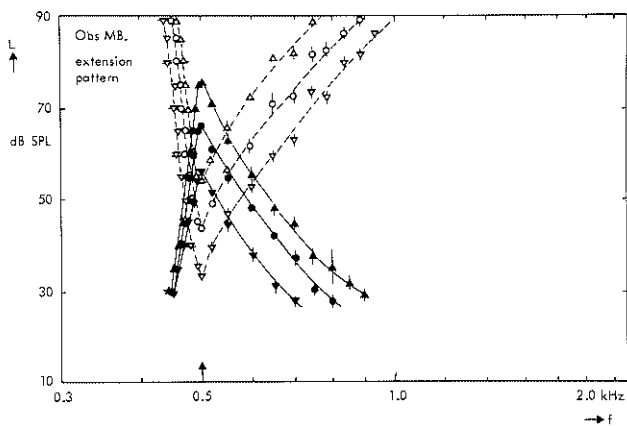
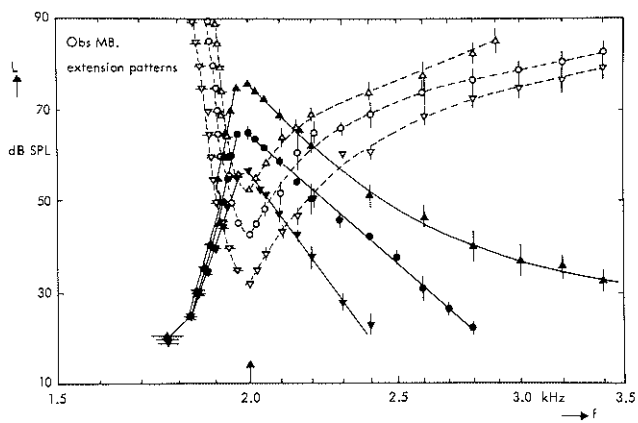
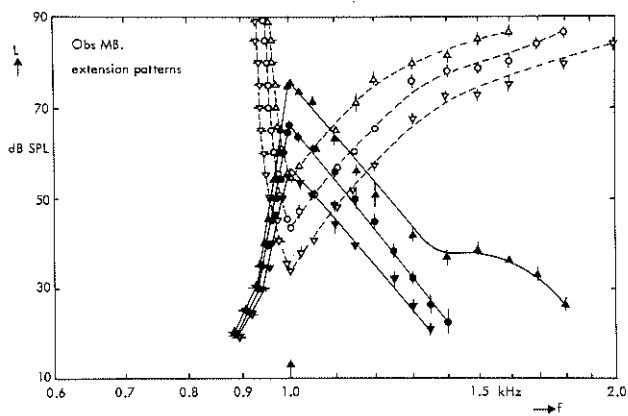


Fig. 14. Extension pulsation pattern for observer MB.



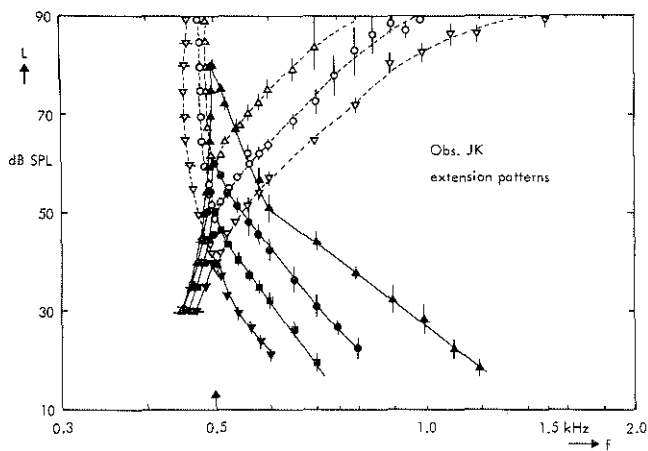
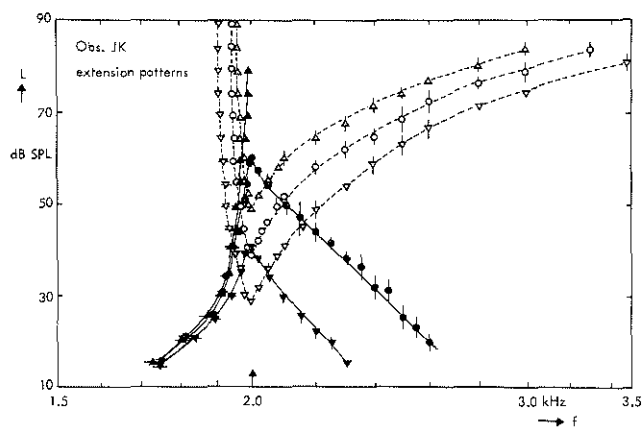
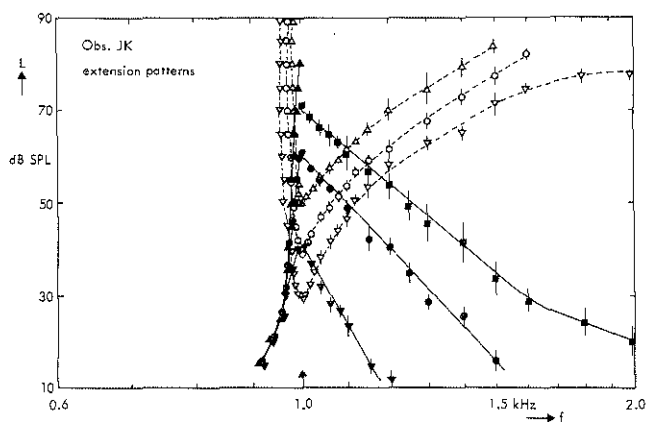


Fig. 15. Extension patterns for observer JK.



IV-5. Discussion

We found a clearly nonlinear behaviour of the pulsation patterns. The steep edge gets steeper as the level is raised, and the shallow edge gets less steep. This effect is also found in the direct masking results of Wegel and Lane (1924). It was analysed by Chistovich (1957), who stressed that results determined by different methods cannot be compared directly. With regard to the origin of the nonlinearity, she concluded that harmonic distortion cannot explain the effect, but that the nonlinearity must be located in the analyser itself i.e. the cochlear system.

A first attempt to compute the steepness of one kind of pattern from results measured with another method, was made by Rodenburg et al. (1974). Their calculations involved only two methods. The publication makes it clear that an absolute value of the steepness of the patterns must be considered with some reserve, and that one must be very careful in comparing results from different methods. In Chapter V we shall present our efforts to give a full description on the nonlinearity and its consequences for the steepness of the patterns.

It will be clear from the patterns presented here that small differences in level can result in very different slopes, even for the same method; this is particularly true for the steep edges of an input pattern.

V. ANALYSIS OF PULSATION PATTERNS

V-1. Introduction

In this chapter we shall describe a model for analysis of pulsation patterns. We shall start with an explicit statement of the assumptions involved in the model and see what their consequences are for the shape and steepness of the measured pulsation patterns.

V-2. The assumptions

1. Excitation pattern

We assume the existence of some internal activity pattern on which the continuity effect is based.

2. Continuity condition

Our assumption regarding the perception of the continuity effect is based on the concepts explained in section I-4. We assume that the continuity effect is observed when the central nervous system cannot detect the absence of the probe. This condition is met when the neural activity in any channel does not drop by a significant amount during the absence of the probe.

This somewhat complex formulation is illustrated in Fig. 16, which represents a pulsation procedure involving alternation of two stimuli, in this case two single-frequency stimuli (Fig. 16 a). Each stimulus has its own excitation pattern (assumption 1), so two excitation patterns will alternate in the auditory system (Fig. 16 b). One of these is due to the pulsator stimulus S_1 , which is clearly heard to pulsate and the other is due to the probe stimulus S_2 (dashed).

We can now consider all channels and see what the time pattern of the activity is. We distinguish between two conditions, one just above the pulsation threshold (dash-dot) and one just below it (dash). Fig. 16 c shows the time patterns at the three locations indicated in Fig. 16 b.

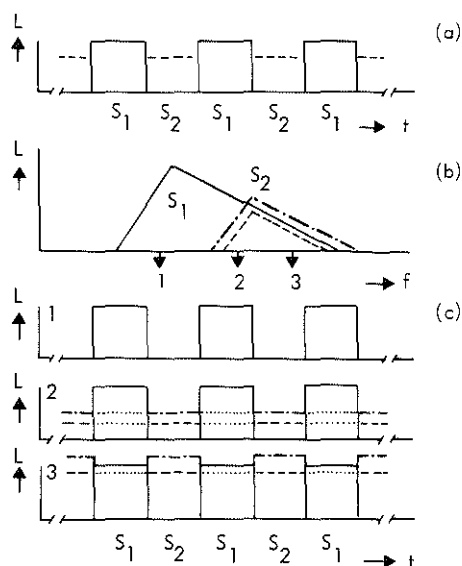


Fig. 16. Illustration of hypothesis underlying pulsation threshold measurement.

(a). Time pattern of presented stimuli. S_1 is pulsator; S_2 is probe stimulus.

(b). Excitation patterns resulting from an alternation of two sine waves. Two situations are presented, one just below the pulsation threshold and one just above the pulsation threshold.

(c). Time patterns of the excitation at the three positions indicated in figure (b).

At location 1 we observe excitation only during the presence of S_1 . This signal is perceived as pulsating.

At location 2 there is always more activity during S_1 than during S_2 . If the system

does not use time-coded information, there is no way of distinguishing between a continuous excitation level of S_2 and an alternation of S_1 and S_2 .

At location 3 the situation for the lower level is similar to that for location 2. For the higher level we see that the excitation level during S_2 is significantly higher than that during S_1 . The drop of the activity level at the transition from S_2 to S_1 can only be interpreted as absence of S_2 during presentation of S_1 . If this condition is not met in any channel, the central nervous system, analysing the incoming excitation, cannot decide whether S_2 is continuously present or not. Given the observed phenomenon and the shape of the pulsation patterns, it seems likely that the conclusion will be drawn that S_2 is continuously present. In terms of our patterns this assumption can be formulated as follows: the probe's excitation pattern may not exceed the pulsator pattern by more than a just noticeable difference.

A similar assumption was formulated by Gardner (1947) with regard to postmasking.

3. Shape of excitation pattern

We assume that the excitation pattern can be described by two straight lines on a dB-log f scale. This assumption seems to be justified for the shallow edge of the pattern; we found a straight line for the shallow edge of the input extension pattern, to which the excitation pattern is closely related. The steep edge of the input extension pattern was curved. This might be due to off-frequency detection, as assumed under 5. For the sake of simplicity we assume straight lines for the edges of the excitation pattern and thus get the triangular shape familiar from vibration patterns of the basilar membrane.

4. Nonlinearity

We found level-dependent slopes in our pulsation patterns. As a first step we introduce a first-order approximation $S(L) = aL + b$, where S is the slope of the pattern in dB/unit of the natural logarithmic scale and L the level in dB SPL. The two coefficients a and b are to be determined.

5. Off-frequency detection

As stated in assumption 2, we compare two excitation patterns. This makes it possible that pulsation is detected at another frequency than the frequency of the probe itself. This is called off-frequency detection, contrary to on-frequency detection when the pulsation is detected at the frequency of the probe. We thus have to limit the patterns at the bottom, because the excitation vanishes at some low level. We assume that this happens at the threshold of hearing which we shall be using as the lower limit of the excitation pattern.

6. Frequency dependence

We have seen that the slope of the patterns is a function of pulsator frequency; the magnitude of the nonlinear effects also depends on frequency (section IV-4). This frequency dependence, however, makes the analysis very complicated and cumbersome. For the time being we assume that the slope and nonlinearity do not depend on frequency when analysing the extension patterns at one centre frequency.

The effects of frequency dependence would have been small anyway in this analysis; on the other hand, they would make the general picture less clear.

We shall return to this point at a later stage of our analysis.

V-3. Calculations

V-3-1. General description

In this section we shall give a general description of our calculations. This will suffice to understand the computed results. A detailed description of methods and equations will be given in section V-3-2. For each value of a , b and the j nd we can compute theoretical points $\langle x_i \rangle$ in accordance with the above assumptions.

We define:

$$t_i = \left(\langle x_i \rangle - \bar{x}_i \right) / s_i$$

where \bar{x}_i is the averaged experimental value and s_i is the standard deviation of this value. Next we compute $TS = \sum t_i^2$ for all measured points on the two extension patterns for one observer at that frequency. The computation is performed separately for the low- and high-frequency edges of the patterns. The results are presented in section V-4.

V-3-2. Detailed description

We give the variables of the pulsator stimulus an index s (L_s, f_s) and those of the probe an index p (L_p, f_p). The j nd is denoted by k . We shall first give a description including the filter patterns and a frequency dependence for the slope (assumption 6). This dependence will be neglected in the actual analysis. The coefficients a and b are given an index referring to the frequency f_s or f_p .

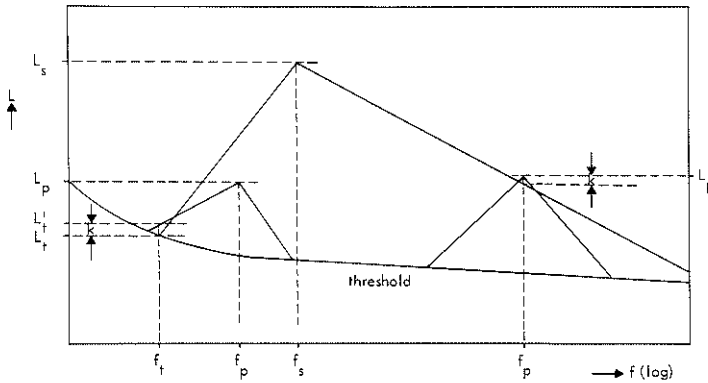


Fig. 17. Identification of stimulus parameters for the derivation of the equations of the theoretical pulsation patterns.
 p refers to the probe stimulus which is to be heard as a continuous tone
 s refers to the pulsator stimulus
 t refers to the threshold of hearing
 k is the just noticeable difference in activity level for pulsation detection

The high-frequency edge becomes less steep when we raise the stimulus level, so the detection of pulsation occurs at the probe frequency. We see from Fig. 17 that

$$S(L) = a_s L_s + b_s = \frac{L_p - k - L_s}{\ln f_p - \ln f_s}$$

This can be written:

$$L_p = L_s + k + (a_s L_s + b_s) \cdot \ln \left(\frac{f_p}{f_s} \right) \quad (7)$$

We assume that off-frequency detection occurs at the low frequency side. We define (f_t, L_t) as the point where the stimulus excitation pattern reaches the threshold level, which is assumed to be equal to the threshold of hearing.

We may now write:

$$L_t = L_s + (a_s L_s + b_s) \cdot \ln \left(\frac{f_t}{f_s} \right) \quad (8)$$

$$L'_t = L_p + (a_p L_p + b_p) \cdot \ln \left(\frac{f_t}{f_p} \right) \quad (9)$$

$$\text{Detection at threshold level gives } L'_t = L_t + k \quad (10)$$

The relationship between L_t and f_t is the threshold of hearing. We need this relationship as a continuous function, while in reality it is only determined at sample points over the total hearing range. Extra samples were taken near irregularities (dips) in the threshold, and the continuous function established by interpolation. However, interpolations tend to produce oscillating solutions. To suppress these oscillations we used a higher-order interpolation, viz a cubic spline interpolation*. In this kind of interpolation the third derivative of the solution is kept constant over an interpolation interval.

Equation (7) has only one variable, so it can be solved. One level and one frequency are the parameters. Either the level or the frequency is used as independent variable in the measurement. The solution for assumed values of a , b and k can be found.

*The reader is referred for details to a general textbook on numerical analysis such as "Numerical methods for scientists and engineers" by R.W. Hamming.

At the other side of the pattern we have four variables (f_t , L_t , L'_t and the dependent variable) and three equations ((8), (9) and (10)) for the solution of the problem, so we need one more equation. This is given by the threshold of hearing giving the relation between f_t and L_t ; so again a solution can be found for given values of a , b and k .

The problem is thus soluble and we can find best-fit values of a , b and k . The problem will be clearer if we present the equation in a more direct form for the various methods. We substitute f_o , L_o for the two parameters for the condition in question (in method a, $f_o = f_s$; $L_o = L_s$) and f and L for the variables. Representing the methods by

(a) = input extension pattern

(b) = input filter pattern

(c) = output extension pattern

(d) = output filter pattern

We get for the high-frequency side:

$$(a) \quad L = L_o + k + (a_o L_o + b_o) \cdot \ln \left(\frac{f}{f_o} \right) \quad (11)$$

$$(b) \quad L = L_o + k + (a L_o + b) \cdot \ln \left(\frac{f_o}{f} \right) \quad (12)$$

$$(c) \quad L = \left(L_o - k - b_o \ln \left(\frac{f}{f_o} \right) \right) / \left(1 + a_o \ln \left(\frac{f}{f_o} \right) \right) \quad (13)$$

$$(d) \quad L = \left(L_o - k - b \ln \left(\frac{f_o}{f} \right) \right) / \left(1 + a \ln \left(\frac{f_o}{f} \right) \right) \quad (14)$$

or

$$(a) \quad \ln \left(\frac{f}{f_o} \right) = - \frac{L_o - (L - k)}{a_o L_o + b_o} \quad (15)$$

$$(b) \quad \ln \left(\frac{f_o}{f} \right) = - \frac{L_o - (L - k)}{aL_o + b} \quad (16)$$

$$(c) \quad \ln \left(\frac{f}{f_o} \right) = \frac{L_o - (L + k)}{aL_o + b_o} \quad (17)$$

$$(d) \quad \ln \left(\frac{f_o}{f} \right) = \frac{L_o - (L + k)}{aL + b} \quad (18)$$

and low-frequency side:

$$(a) \quad L = \frac{L_t + k + bX_a}{1 - aX_a} \quad \text{with} \quad X_a = \ln \left(\frac{f}{f_o} \right) + \frac{L_o - L_t}{aL_o + b_o} \quad (19)$$

$$(b) \quad L = \frac{L_t + k + b_oX_b}{1 - a_oX_b} \quad \text{with} \quad X_b = \ln \left(\frac{f_o}{f} \right) + \frac{L_o - L_t}{aL_o + b} \quad (20)$$

$$(c) \quad L = \frac{L_t + b_oX_c}{1 - a_oX_c} \quad \text{with} \quad X_c = \ln \left(\frac{f}{f_o} \right) + \frac{L_o - (L_t + k)}{aL_o + b} \quad (21)$$

$$(d) \quad L = \frac{L_t + bX_d}{1 - aX_d} \quad \text{with} \quad X_d = \ln \left(\frac{f_o}{f} \right) + \frac{L_o - (L_t + k)}{aL_o + b_o} \quad (22)$$

or

$$(a) \quad \ln \left(\frac{f}{f_o} \right) = \frac{L - (L_t + k)}{aL + b} - \frac{L_o - L_t}{a_oL_o + b_o} \quad (23)$$

$$(b) \quad \ln \left(\frac{f_o}{f} \right) = \frac{L - (L_t + k)}{a_oL + b_o} - \frac{L_o - L_t}{aL_o + b} \quad (24)$$

$$(c) \quad \ln \left(\frac{f}{f_o} \right) = \frac{L_o - (L_t + k)}{aL_o + b} - \frac{L - L_t}{a_oL + b_o} \quad (25)$$

$$(d) \quad \ln \left(\frac{f_o}{f} \right) = \frac{L_o - (L_t + k)}{a_oL_o + b_o} - \frac{L - L_t}{aL + b} \quad (26)$$

In all equations we have $\ln(f/f_0)$ for methods (a) and (c) and $\ln(f_0/f)$ for methods (b) and (d). This represents a reflection in a vertical line through f_0 . This reflection was used for the comparison of experimental filter patterns and extension patterns in section IV-4 (Figs. 10 and 12).

V-4. Results

We are dealing with a three-parameter estimation problem. We use a grid-tabulation method to solve it. Tabulation will give useful information about the influence of each parameter on the fit. We make use of the generally accepted assumption that the sum of squares (TS) is a measure of the quality of the fit. A minimum value of TS represents the best fit.

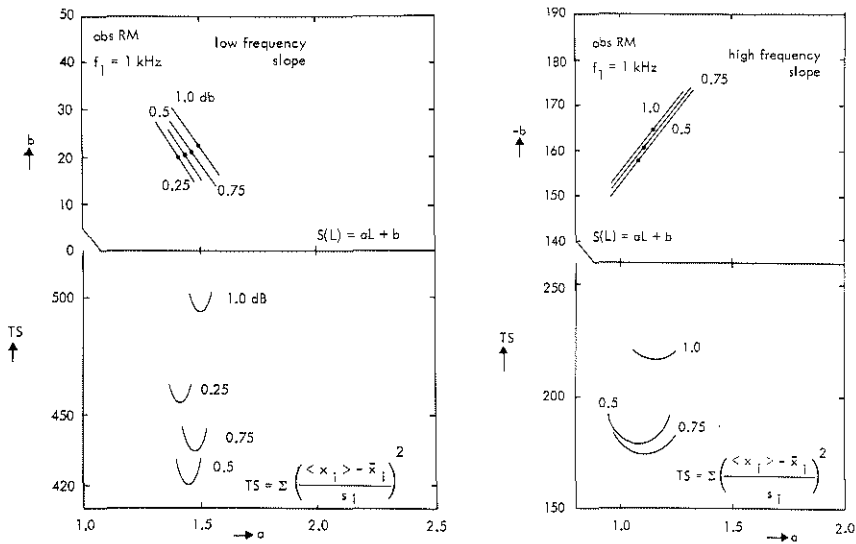


Fig. 18. An example of the results of the analysis based on our model. The slope-level dependence is assumed to be $S(L) = aL + b$. The figure on the left shows the analysis for the pulsation extension patterns at the low-frequency side and that on the right the analysis for the high-frequency side. The upper part of each figure shows the relationship between a and b and the lower part the relationship between a and the sum of squares of the differences between experimental and theoretical values.

The results of the analysis of the patterns of Fig. 10 a are shown in Fig. 18. Fig. 18 a gives the results for the low-frequency side, and Fig. 18 b, those for the high-frequency side. The low- and high-frequency sides show quite different behaviours. Most analyses show a single minimum without saddle points or local minima. The results presented in Fig. 18 are typical. In the lower part of each figure we give the minimum value of TS for a value of (a) with the jnd (k) as parameter. The value of (b) for which this minimum is found is shown in the upper part of the figure. The dots correspond to minimum value of TS.

The low-frequency results show that the value of (a) is very important for the fit. A small change in (a) raises the value of TS considerably. The analysis showed that the value of (b) is of minor importance. A considerable change in (b) causes only a minor change in TS. It can be seen from Fig. 18 that the jnd (k) has only a slight effect on the best values of (a) and (b). It has a considerable effect in the minimum value of TS, however. This observer shows a minimum TS for a jnd of about 0.5 dB.

The high-frequency results are less sensitive to changes in (a) as long as a linear relationship between (a) and (b) is maintained, as shown in the upper part of the figure. A small deviation from this linear relationship will raise TS considerably (not shown in Fig. 18). Again the jnd has only a minor effect on the best-fit values of (a) and (b) and a large effect on TS.

We present the best-fit values of (a), (b) and (k) for four observers at three frequencies in Fig. 19. The left-hand figure shows the results for the steep low-frequency side, and the right-hand figure those for the high-frequency side. The lines around the points show the sensitivity of the model to parameter variations. They represent the values of the parameter in question that will raise TS by not more than 10%. The other parameters are freely varied and may change considerably with change in the parameter in question.

The low-frequency fits of (a) show two points with a considerable range. For the 1 kHz point of observer HK this is due to the fact that the solution has two relative minima with about the same value of TS. The lowest value is obtained for an (a) of about 6, the other is found near 3.

The second point with considerable range is the 2 kHz point of observer JV. Now

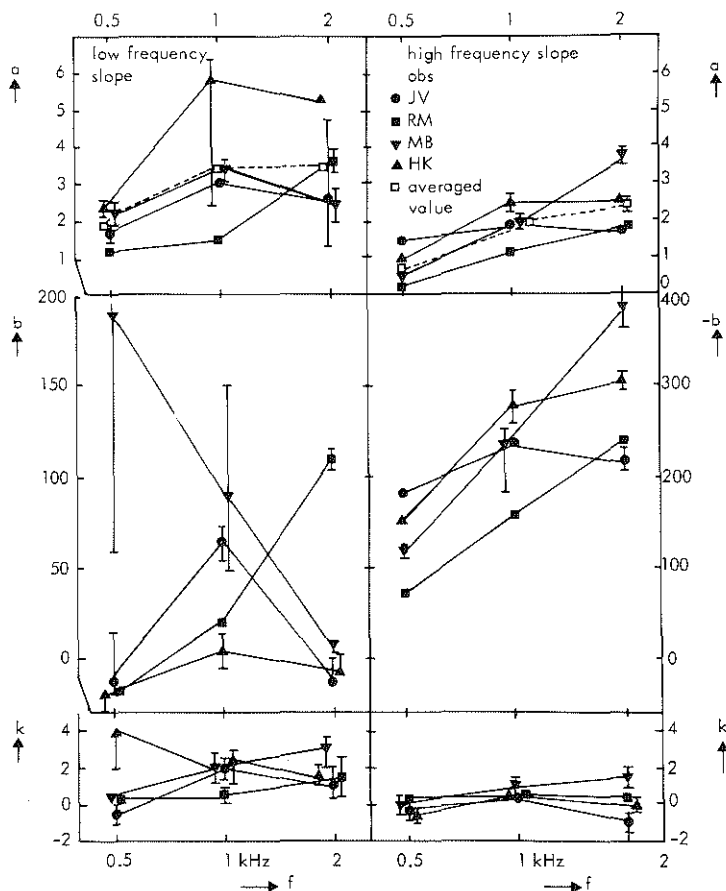
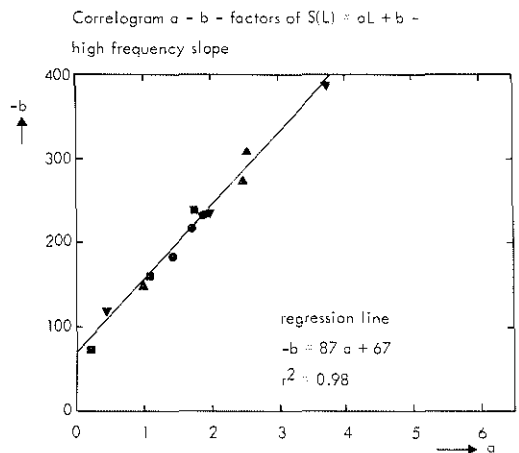


Fig. 19. Best-fit values of (a), (b) and (k) for three frequencies and four observers, computed with the aid of a model for analysis of pulsation threshold extension patterns. The lines through the best-fit values show the sensitivity of the model to parameter variations. They cover the range of values that will result in an increase of the value of TS (sum of squares) by not more than 10% of the minimum value.

TS has only one minimum, but the value of TS does not depend much on the value of (a). Unlike the typical results presented in Fig. 18, we get a very flat parabola for TS as a function of (a) in this case. We also show average values of (a) in Fig. 19. The average value at 1 kHz depends largely on the results for observer HK. The best fits have an average value of 3.4. If we take the second best fit for observer HK we get a value of 2.6. The former average value is exactly the same as that at 2 kHz. The value of (a) at 0.5 kHz is lower than the others. The averages make

Fig. 20. Correlogram of the best-fit values of $\langle a \rangle$ and $\langle b \rangle$, the coefficients of the equation $S(L) = aL + b$ for the high-frequency edge of the pulsation pattern. The equation of the regression line and the correlation coefficient are also given.



it impossible to draw any conclusions about the frequency dependence of $\langle a \rangle$. We see that observers JV and MB have about the same value of $\langle a \rangle$ at 1 kHz and 2 kHz and a smaller value at 0.5 kHz. Observer RM has the same value of $\langle a \rangle$ at 0.5 kHz and 1 kHz and a higher value at 2 kHz. The results for observer HK can be interpreted either way.

The best-fit values of $\langle b \rangle$ show a very different behaviour for the various observers, with large differences in the accuracy. We do not think that it is justified to speak of a general behaviour.

The values of $\langle a \rangle$ at the high-frequency side are all quite well determined. The frequency dependence is again different for the various observers. Observers RM and MB show the same frequency dependence as found for the average value: the higher the frequency, the higher the value of $\langle a \rangle$. Observer JV shows an almost constant $\langle a \rangle$, while for observer HK the best-fit value of $\langle a \rangle$ is the same at 1 kHz and 2 kHz and lower at 0.5 kHz.

The frequency dependence of $\langle b \rangle$ at the high-frequency side is like that of $\langle a \rangle$ at the same side of the pattern. We show the correlogram for $\langle a \rangle$ and $\langle b \rangle$ in Fig. 20 together with the equation of the linear regression line. We find a very high degree of correlation ($r^2 = 0.98$), indicating that the values of $\langle a \rangle$ and $\langle b \rangle$ are closely related. The accuracy lines of Fig. 19 show that this does not mean that all estimates can be replaced by one value for $\langle a \rangle$ and $\langle b \rangle$.

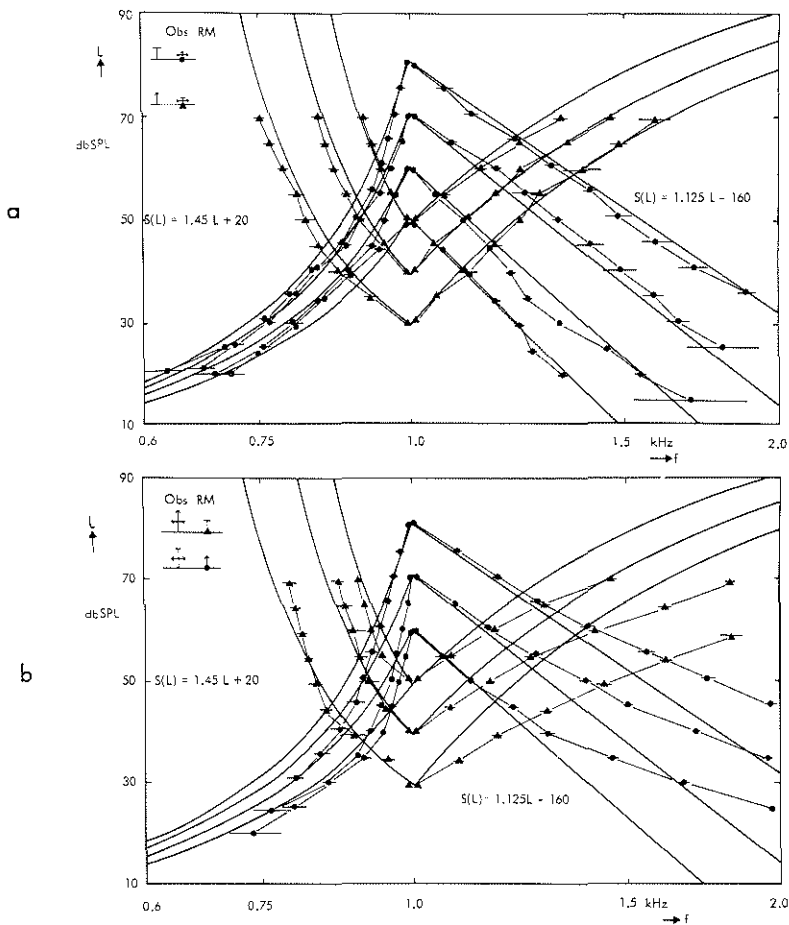


Fig. 21. Computed best-fit pulsation extension patterns (a) and the same theoretical patterns compared with the filter patterns reflected in a line through the centre frequency (b). The equations of the slope-level dependence are given in the figures. Circles denote input patterns and triangles denote output patterns.

The best-fit values of k (ind) are shown in the bottom panel. At the high-frequency side, the value of k varies little with frequency and from observer to observer. We averaged the values of k over all observers for the different frequencies and over all frequencies for the different observers. This analysis shows that the differences are not significant in a student- t test. The value of k at 1 kHz tends to be somewhat higher than the average, and observer RM tends to show a somewhat higher value. The average value over all frequencies and observers is 0.2 ± 0.3 dB ($n = 12$). At the low-frequency side the behaviour seems to be somewhat more complicated.

Averaging the results for different observers does not yield statistically different values for the three frequencies. The value of k averaged over all observers and frequencies is 1.6 ± 0.3 dB ($n = 12$). Averaging the value of k for the different frequencies yields statistically different results for the three observers, viz 1.3 ± 0.4 dB, 0.8 ± 0.4 dB, 3.0 ± 0.6 dB and 1.5 ± 0.6 dB for observers JV, RM, MB and HK respectively. This suggests differences in the criterion for pulsation between the various observers; the differences involved are rather small, however.

After obtaining the best-fit values, we can compute the shape of the patterns according to the model. Fig. 21 shows the best-fit theoretical patterns, for the results given in Fig. 10. We see that the model gives quite a good approximation for the extension patterns, but a poor one for the filter patterns. We already mentioned frequency dependence as a possible explanation for this discrepancy. Some of this dependence is also found in Fig. 19. We did not pursue this dependence further, because the analysis of the nonlinearity is our main interest.

V-5. Consequences of the nonlinearity for shape and steepness of the patterns

The above analysis shows that the nonlinearity is essential for the understanding of the shape of the pulsation patterns. A complicated calculation, however, is not needed for a qualitative understanding of the shape.

We should bear in mind that the low-frequency edge of the excitation pattern gets steeper and the high-frequency edge gets less steep as the level is raised. We may thus conclude on the basis of our continuity assumption (2 in section V-2), that the high-frequency side of the input extension pattern really is the high-frequency side of the excitation pattern itself, while at the low-frequency side pulsation is detected off-frequency near the threshold of hearing. Since the pulsation patterns represent relations between frequencies and levels, this means that the slope determined is a much steeper slope than that of the excitation pattern itself (see broken and full lines in Fig. 22). This in its turn implies a curvature of the patterns. A similar picture can be drawn for the output filter pattern (Fig. 23). The upper part of the figure shows the curvature of the shallow edge of the pattern. The broken lines in

this figure represent the excitation patterns, and the peak of these patterns determine the pulsation pattern. The lower part of the figure shows in a similar way the curvature and steepening effect for the steep edge of the patterns.

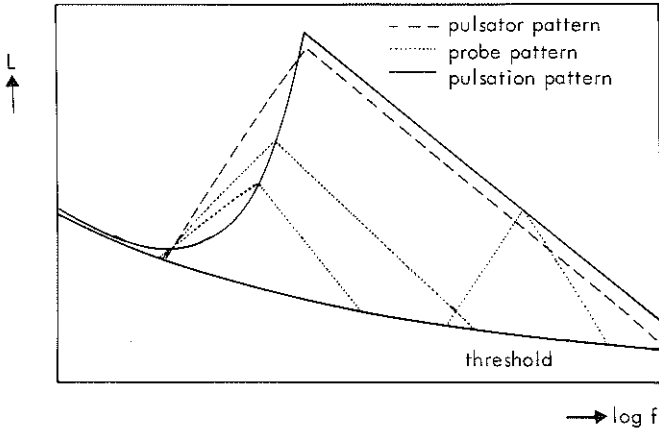
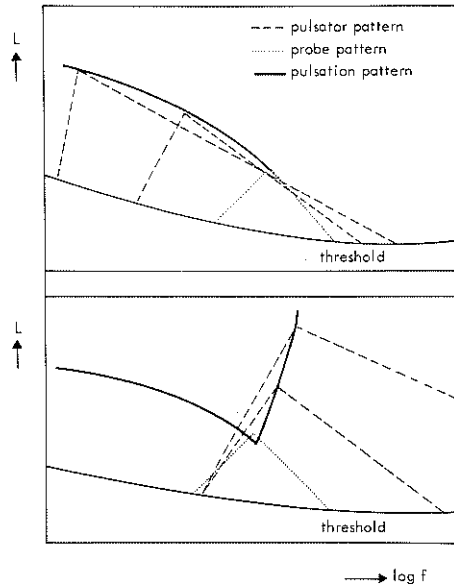


Fig. 22. Construction of the shape of the pulsation threshold input extension pattern from the nonlinear activity or excitation pattern.

Fig. 23. Construction of the shape of the output filter pulsation pattern from a nonlinear excitation pattern. The upper part of the figure shows the construction of the shallow low-frequency edge and the lower part the construction of the steep high-frequency edge.



Similar drawings can be made for the other types of patterns. We see that the steepness of most pulsation patterns does not reflect the tuning of the auditory system alone, but also its nonlinearity. This implies that the experimentally determined slopes are not adequate measures of the tuning.

This can also be demonstrated by deriving the equation for the steepness of the pulsation pattern. We assume again that the coefficients (a) and (b) are frequency-independent, in order to avoid very complicated equations. The slope of the low-frequency side of the input extension pattern (S_{iep}) can be derived from equation (23).

$$\ln \left(\frac{f}{f_o} \right) = \frac{L - (L_t + k)}{aL + b} - \frac{L_o - L_t}{a_o L_o + b_o}$$

Differentiation with respect to L gives:

$$\frac{d}{dL} \left(\ln(f) \right) = \frac{aL + b - aL + aL_t + ak}{(aL + b)^2}$$

$$\text{whence:} \quad S_{iep} = \frac{(aL + b)^2}{aL_t + b + ak} \quad (27)$$

We derive the slopes of the other patterns in the same straightforward way, but we have to remember that the value, obtained for the shallow edge of the two input patterns really does represent the slope of the excitation pattern. All other equations are very similar except that the k (jnd) appears in different positions. If we neglect this factor we find a general equation:

$$S_p = \pm \frac{(aL + b)^2}{aL_d + b} \quad (28)$$

where L_d is the level at which the detection takes place and L is the variable level. L_d is the threshold of hearing for the steep side and the probe level for the other side of the output patterns.

The (k) or jnd was neglected and will be neglected because its value is rather small (Fig. 19). It will however have a certain effect on the slope of the steep edge of the patterns, increasing this slope for the output pattern and reducing it for the input pattern.

Equation (28) can be written

$$S_p = \pm \frac{aL + b}{aL_d + b} \cdot (aL + b) \quad (29)$$

This can be interpreted as multiplication of the real slope (of the excitation pattern) by a factor $(aL + b) / (aL_d + b)$. At the steep side L_d is much smaller than L and the multiplication factor will be larger than one: the steep edge of the pulsation pattern is always steeper than the corresponding edge of the excitation pattern. This overestimation increases with the level, thus resulting in a curved pattern.

At the other side, L_d is smaller than L for output patterns. $S(L)$ is negative here, so the slope decreases at higher levels and the multiplication factor will be smaller than one. This will make the pulsation pattern less steep than the excitation pattern. The pattern will also be curved, due to the level dependence of the multiplication factor.

Derivation of the slope at the shallow side of the input filter patterns is not very complicated, even if we assume (a) and (b) to be frequency dependent. We find that the slope of the filter pattern is less than that of the excitation pattern owing to this frequency dependence. The rather complicated equations that can be derived for the steep side show a higher slope for the filter patterns than for the extension patterns.

Once more we want to stress the point that the slope of the pulsation threshold patterns does not reflect the tuning properties alone, but also the nonlinearity. The slopes at the steep side are considerably exaggerated. The slope of the excitation pattern for observer RM at 1 kHz and 70 dB SPL (Fig. 10) is about 100 dB/oct. His input and output patterns have slopes of 300 dB/oct and 250 dB/oct respectively. For observers with a stronger nonlinearity the difference is even larger. The computed excitation pattern for observer JV at 1 kHz and 80 dB SPL has a slope of 190 dB/oct; his input and output patterns have slopes of 400 dB/oct and 1500 dB/oct respectively.

This overestimation of the slope at the steep side (and underestimation of the other slope) may also occur in other determinations of cochlear tuning such as masking and neurophysiologic tuning. We must thus be very careful in interpreting experimental data whenever a similar nonlinearity is found.

We have seen that the nonlinearity is also found in masking experiments (section IV-5).

For a full analysis we need explicitly formulated assumptions concerning the masking criterion; only then we will be able to judge whether off-frequency detection is possible and what the effect of the nonlinearity on the masking patterns is.

Nonlinearity in single-fibre recordings will be discussed in Chapter VII.

V-6. Concluding remarks

Of the assumptions regarding the interpretation of pulsation patterns formulated at the start of this chapter, the assumptions 1, 2 and 3 represent generally accepted ideas about the functioning of the auditory system and the perception of continuity in pulsation conditions, while assumptions 4, 5 and the revised assumption 6 are derived from considering our experimental results.

On the basis of these assumptions, we derived a description of the shape of the pulsation patterns which agreed fairly well with patterns measured by various methods. It should be noted, however, that the assumptions presented here are not compatible with the widely held view that the probe level is a direct measure of the distributed activity of the stimulus at the probe frequency. In our opinion this view should be replaced by one based on comparison of activity patterns.

VI. THE BANDWIDTH OF THE PULSATOR

VI-1. Introduction

All of our experimental results presented so far were obtained with sinusoidal pulsators. In Chapter I we discussed, among other things, the literature on pulsation patterns of noise-band stimuli. Some authors compared these patterns with direct masking patterns. Fastl (1974a) found that the slopes of the pulsation patterns and masking patterns are about equal if narrow noise-band pulsators and maskers are used; these patterns differed for sinusoidal stimuli (Fastl, 1975). Houtgast (1974 a, c) found differences between simultaneous masking patterns and pulsation patterns, irrespective whether the pulsator and masker were sine waves or narrow noise bands. He attributed this difference to lateral suppression, a phenomenon not found in simultaneous masking. Fastl (1975) found that his pulsation input patterns for sine waves were less steep at the low-frequency side than Houtgast's corresponding pulsation patterns.

The main question underlying this discussion is whether or not there is a common mechanism for pulsation and masking. Houtgast concluded that masking and pulsation are established at different levels in the auditory system because they have different properties with regard to sharpness of tuning and suppression. In Fastl's view they both involve tuning at the peripheral level; while they may permit different experimental procedures, they will not show up different properties in the auditory system.

A good understanding of parallels and discrepancies between the results obtained with masking and pulsation can contribute to a better understanding of the processes underlying the phenomena on which these methods are based, and to what extent the methods will reveal different properties of the system.

The problems presented above can be reduced to two basic points: firstly the influence of the spectral width of the pulsator on the patterns, and secondly the relation between pulsation and masking. We shall investigate this subject in this chapter.

In the study of the influence of the bandwidth we distinguish between effects at the edges and effects at and near the summit of the pattern. The edge effects are related to the sharpness of tuning; the summit effects to the shape of the auditory filter and its integrating properties such as the critical band mechanism.

VI-2. Apparatus

We used the same equipment as described in section IV-3, except that one oscillator was replaced by a B & K type 1024 sine random generator. This generator produces narrow bands of noise with -3 dB bandwidths of 10, 30, 100 and 300 Hz. It is a beat oscillator with fixed filters. This means that the central frequency of the noise band can be measured by switching the generator to its "sine" mode and that the slope of the spectral cut-off does not depend on the central frequency of the noise. The generator's -40 dB bandwidths are approximately 70, 120, 160 and 400 Hz respectively. This corresponds to slopes of at least 250 dB/oct over the first 40 dB.

The rms power is kept constant for the various bandwidths. We adjusted this level to 1 V at the gate input, i.e. equal to the power of the sine wave used in the previous chapters.

VI-3. Results

We measured pulsation patterns for narrow noise bands centered at 1 kHz, viz an input extension pattern for a pulsator (noise-band) level of 70 dB SPL and an output

pattern for a probe (sine-wave) level of 40 dB SPL. All experimental points represent the average of five measurements.

VI-3-1. Summit effects

VI-3-1-a. On-frequency pulsation level

We first consider the pulsation level at the central frequency. We computed the difference between the pulsator level and the probe level. The differences were about equal for the input patterns and the output patterns. The average differences for two observers are shown in Fig. 24 for the four bandwidths, together with the corresponding values for a sine-wave pulsator.

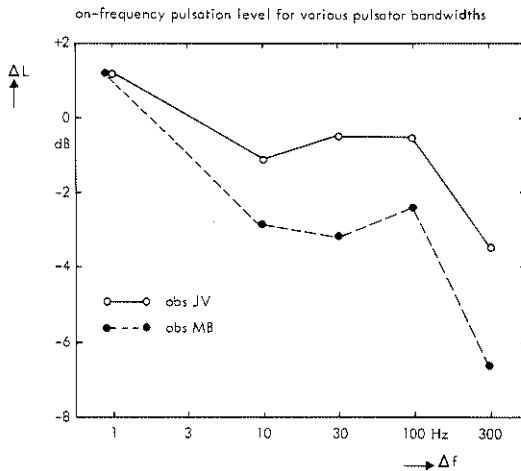


Fig. 24. The difference between the pulsation threshold level of a sine wave at the central frequency of the pulsator and the presentation level of the pulsator as a function of the bandwidth of the pulsator. The points represent the average of an input determination at 70 dB SPL and an output determination at 40 dB SPL. The results for two observers are presented.

The level difference for a sine-wave pulsator is equal for the two observers. The level differences for the 10, 30 and 100 Hz noise bands are more or less equal for each observer but differ between the observers. They differ by about -1.5 dB and -3.5 dB from the level differences for a sine-wave pulsator for the observers JV and MB respectively.

The difference between sine-wave and narrow-band level differences may be due to the limited width of the noise bands. The noise band actually sounds like an

irregularly modulated tone. An observer has to distinguish this irregular fluctuation from the regular 4 Hz modulation. This will result in a continuity criterion for noise bands that differs from that for a sine wave. A similar criterion effect for pulses in direct masking has been described by van den Brink (1964).

The 300 Hz level difference is some 3 dB down from those at the other bandwidths. This bandwidth effect cannot be due to a criterion effect, because the fluctuations are smaller for this wider band of noise and we would have expected the opposite effect on this ground. The effect can be explained by the critical band mechanism of the ear. For narrower bands of noise all energy is within one critical band and is summed to give a constant equivalent level. For the 300 Hz noise band some of the energy is outside the critical band at 1 kHz, resulting in a lower equivalent level.

VI-3-1-b. Summit pattern

We next consider the part of the pulsation pattern between the cut-off frequencies of the noise bands. There are no systematic differences between the input and output patterns, so we averaged the data for the two. The results for two observers are shown in Fig. 25. These results naturally show the level difference described in the previous section.

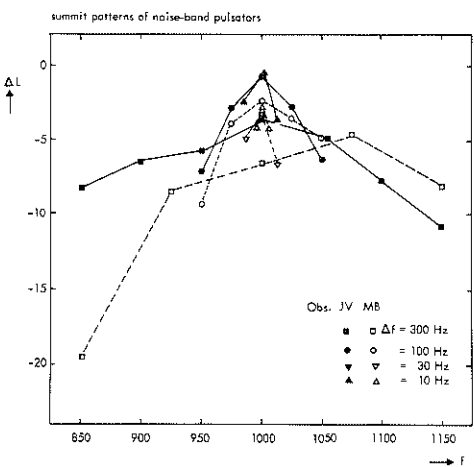


Fig. 25. Pulsation patterns for narrow bands of noise at frequencies within the noise bands. The experimental points presented for two observers are the averages for an input determination at 70 dB SPL and an output determination at 40 dB SPL. Various symbols represent different bandwidths; the solid symbols are for observer JV and the open ones for observer MB.

The summit patterns are peaked with their peaks near the central frequency. The height of the peak depends on the bandwidth of the noise. There are considerable differences between the patterns for the two observers. Observer JV has symmetrical patterns. The height of the peak is about 6 dB for a 100 Hz and a 300 Hz noise band. Observer MB has markedly asymmetrical patterns with a higher slope at the low-frequency side of the pattern.* This difference between the two observers is puzzling; we have not been able to find any explanation for it.

VI-3-2. Patterns outside the bandwidth of the noise

Next we consider the pulsation patterns outside the noise bands. We are particularly interested in this part of the pattern as it reflects the tuning properties.

The input patterns for one observer (JV) are presented in the lower part of Fig. 26 by full lines and solid symbols; the output patterns for the same observer are shown in the upper part of this figure by broken lines and open symbols. The respective patterns for a sine-wave pulsator at the same level are also shown.

At a first glance, the patterns seem to be shifted as a result of the enlarged bandwidth of the pulsator, except in the frequency region below 0.8 kHz; here the shape is entirely different.

A more thorough look at the patterns, however, reveals some remarkable differences.

Some patterns are within the sine-wave pattern in spite of the larger pulsator bandwidths. This could be due to the criterion difference described in section VI-3-1-a, resulting in a vertical shift of the patterns. This fact will be investigated in section VI-5.

Another difference is the shape of the output patterns. The curvature of the sine-wave patterns is different from the curvature of the narrow noise-band patterns (cf. the 10 Hz pattern). These patterns cannot be made to coincide by either a horizontal and/or

*Another observer (RM) determined these patterns in a somewhat different procedure. His patterns have the same shape as those found by observer JV.

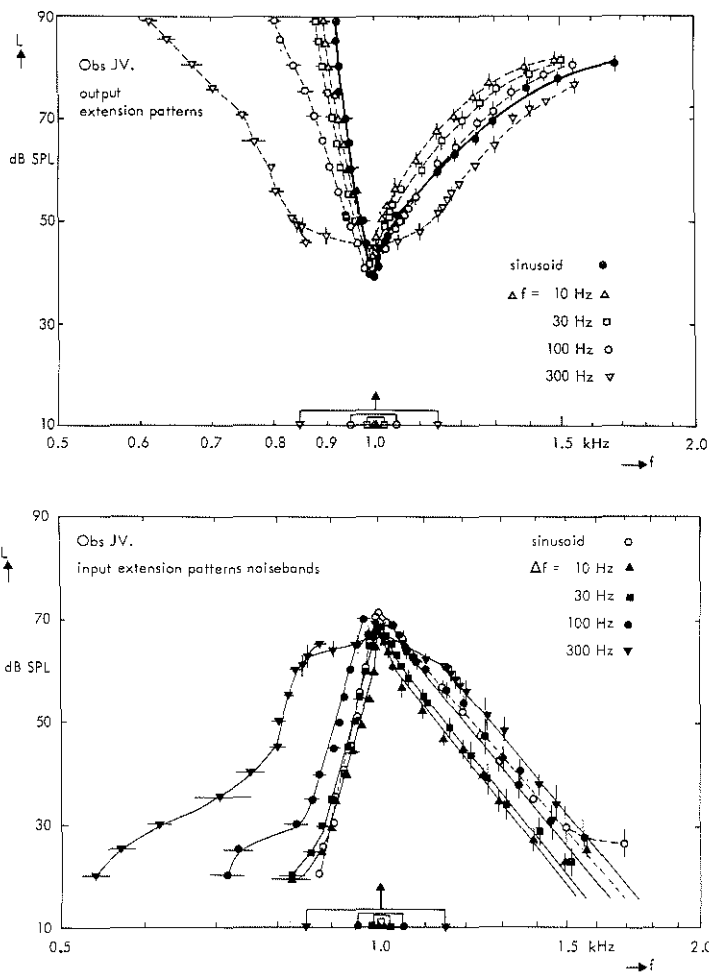


Fig. 26. Input and output extension pulsation patterns for observer JV with a sine-wave pulsator and four narrow noise-band pulsators with bandwidths of 10, 30, 100 and 300 Hz.

a vertical shift. The difference in shape could be due to the nonlinear behaviour described for sine-wave pulsators but such a statement requires a more careful analysis of the data. The analysis will be presented in section VI-5.

At the low-frequency side of the patterns we get different results near the summit of the pattern depending on whether the frequency or the level was adjusted to arrive at the pulsation threshold; this fact is particularly apparent in the 300 Hz pattern in Fig. 26. The variation of frequency was found to have an appreciable

drawback: near the summit, variation of the frequency results in the perception of a tone that sounds more or less continuous. The absence of a condition giving a clear continuity effect results in an unstable criterion, hence the observed difference in pulsation threshold. The instability of the criterion was found only for the points within the pulsator's bandwidth.

The findings concerning criterion shift and possible nonlinear behaviour made us decide to perform a more careful analysis of the data with the aid of the model described in Chapter V. Observer JV made his pulsation adjustments for equal frequency separations from the cut-off frequency or for equal levels for all bandwidths, thus permitting the detailed numerical analysis presented in section VI-5. Before we go on to deal with this analysis, however, we will first consider the bulge in the patterns below 0.8 kHz.

VI-4. Spectral analysis of noise bands

The low-frequency edge of the pulsation patterns shows a surprising bulge for frequencies below about 0.8 kHz. This effect is found for both the input and the output pattern for bandwidths of 100 Hz and 300 Hz. It represents such a striking departure from the general picture of more or less parallel patterns that we decided to reconsider the spectrum of the noise bands, which we had initially thought would not interfere with the frequency selectivity of the auditory system.

We measured the spectrum of the four noise bands with a HP 3580 spectrum analyser, using a 3 Hz resolution bandwidth and a corresponding time constant. This analysis shows a rather broad spectrum for the larger bandwidths and a narrow spectrum for the smaller bandwidths.

We then compared the spectra and the pulsation patterns. The comparison shows that the bulge in the pulsation patterns appears only in those parts where the spectrum of the pulsator is wider than the pulsation pattern that could be expected. This occurred only at the low-frequency edges of the patterns for 100 Hz and 300 Hz bandwidths. The spectrum of the narrower noise bands and all spectra at the high-

frequency side were narrower than the pulsation patterns and did not give rise to the bulge.

We conclude that the bulge in the patterns below 0.8 kHz is due to spectral effects. The parts concerned will be discarded in the analysis of differences between the patterns for sine waves and narrow noise bands. This effect shows once more the importance of careful consideration of the spectra involved.

VI-5. Analysis of noise-band patterns

The patterns for noise bands presented in Fig. 26 are not simply shifted sine-wave patterns. It was suggested in section VI-3-2 that nonlinearity might be responsible for the small differences in shape. This possibility can only be studied by analysing the patterns with the aid of the model described in Chapter V. This analysis has to take into account the findings presented in section VI-3 with regard to the criterion differences and the peaking of the summit patterns.

The criterion difference was found in the on-frequency pulsation levels presented in Fig. 24. We distinguish between the pulsation criterion for sine waves (k_s) and that for noise bands (k_n). The values of these variables used in the numerical analysis are taken from Fig. 24.

The sloping parts of the noise-band pulsator patterns seem to start at the point of the summit patterns at the cut-off frequency for the bandwidths of 100 Hz and 300 Hz. This point is about 6 dB down from the pulsation level at the central frequency. This level difference is denoted by ΔL_o in the equations.

VI-5-a. Analysis of the high-frequency side of the input pattern

The high-frequency side of the input pulsation pattern for a sine-wave pulsator was described by equation (11):

$$L = L_o + k_s + S_o \ln \left(\frac{f}{f_o} \right)$$

where L_o is the pulsator level, k_s the detection criterion, f_o the pulsator frequency, S_o the slope of the excitation pattern for the pulsator level and L and f' the level and frequency of the probe stimulus.

For a noise-band stimulus this equation becomes:

$$L_n = L_o + k_n + S_o \ln \left(\frac{f'}{f_o + \Delta f_n} \right) + \Delta L_o \quad (30)$$

Here Δf_n is half the bandwidth of the noise and S_o the slope of the excitation pattern for the noise band at stimulus level L_o . We assume that the slope of the excitation pattern for a sine wave is equal to that for a noise band.

Subtracting the pulsation level for a sine-wave pulsator at some frequency from the pulsation level for the noise band at the same frequency, we get:

$$\Delta L = L_n - L_s = \Delta L_o + (k_n - k_s) + S_o \ln \left(\frac{f_o}{f_o + \Delta f_n} \right) \quad (31)$$

The equation can be interpreted as a parallel shift of the patterns. The magnitude of this shift was computed, assuming a ΔL_o of -6 dB for all bandwidths.

The pulsation patterns calculated in accordance with the equation (30) are shown in the lower part of Fig. 26 as full lines. The calculated patterns fit the experimental points very well.

This kind of analysis can only be performed on the high-frequency side of the input pattern. A different kind of analysis must be used for the low-frequency side of this pattern and the two sides of the output pattern. We shall now present a method of analysis that is applicable to all four edges. In this analysis, we make use of the fact that observer JV determined experimental points either at a number of levels or at equal frequency separations from the cut-off frequency. We denote this frequency separation by Δf . Equation (11) can be written as:

$$L_s = L_o + k_s + S_o \ln \left(\frac{f_o + \Delta f}{f_o} \right) \quad (32)$$

and equation (30) as:

$$L_n = L_o + k_n + S_o \ln \left(\frac{f_o + \Delta f_n + \Delta f}{f_o + \Delta f_n} \right) + \Delta L_o \quad (33)$$

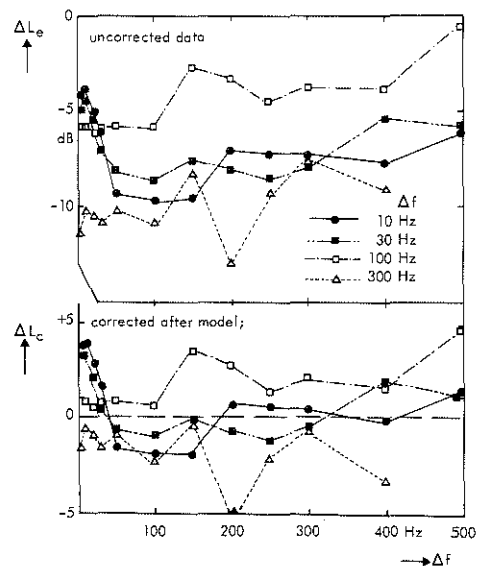
The difference between the levels at a given frequency separation is:

$$\Delta L = L_n - L_s = \Delta L_o + (k_n - k_s) + S_o \left(\ln \left(\frac{f_o}{f_o + \Delta f} \right) - \ln \left(\frac{f_o + \Delta f_n}{f_o + \Delta f_n + \Delta f} \right) \right) \quad (34)$$

The value of ΔL is the sum of the difference between the pulsation levels at the central frequency and the cut-off frequency, a criterion difference and a term taking the logarithmic scale into account. The value ΔL can be computed for each frequency at which we determined the pulsation levels; it represents the effect of changing from a sinusoidal stimulus to a noise-band stimulus we can expect on the grounds of this model.

We present the difference ΔL_e between the experimental pulsation levels of sine-wave and noise-band pulsator determined for the various frequency separations used in the upper part of Fig. 27, after correction for a pattern shift according to the bandwidth of the pulsator. Subtracting the computed values of ΔL from the

Fig. 27. Results of analysis of the high-frequency side of the input patterns for noise-band pulsators. The upper part shows the differences (ΔL_e) between the experimental values for the different bandwidths and the sine-wave pattern after correction for the bandwidth of the pulsator. The lower part shows the values (ΔL_c) after corrections for the criterion differences, the log scale used in the model and the peaking of the summit pattern ($\Delta L_o = -6$ dB). The broken line represents a ΔL_c of 0 dB.



experimental values of ΔL_e , we obtain the level differences ΔL_c which are presented in the lower part of Fig. 27. The values of ΔL_c represent the differences between the pulsation patterns for sine-wave and noise-band pulsators after corrections for

pattern changes due to criterion shift, the peaking of the summit pattern and the enlarged bandwidth of the noise.

We see a set of more or less flat curves for the 100 Hz and 300 Hz noise bands. The curves are near the broken line representing a ΔL_o of 0 dB. For the two smaller bandwidths we also get curves near the 0 dB line except at small frequency separations.

The accuracy with which we can compare the patterns is determined by the standard deviation of the determination of the individual points. This amounts to about 1 dB near the summit, and rises to 2 or 3 dB for the lower parts of the pattern. It follows that the accuracy of the points in Fig. 27 is 1 dB for small frequency separations and rises to several dB for larger separations. Consequently, the majority of the values are not significantly different from 0 dB. A few values differ significantly from 0 dB. They are found at frequency separations smaller than 50 Hz for bandwidths of 10 Hz and 30 Hz. Here, the summit patterns seem to extend into the sloping part of the patterns.

We averaged the differences in level (ΔL_o), shown in the lower part of Fig. 27 at frequency separations larger than 50 Hz for all bandwidths. The average value is 0.0 dB with a standard error of 0.4 dB ($n = 27$).

A small slope effect can be expected for the 300 Hz noise band because of the lower equivalent stimulation level at the central frequency, due to critical band mechanism. This level reduction is about 3 dB, resulting in a different slope S_o . The difference would be 5.7 dB per unit of the natural log scale, resulting in a 2.7 dB level difference at $\Delta f = 400$ Hz and a 1.3 dB difference at $\Delta f = 300$ Hz. This effect would bring the points closer to the 0 dB line, but its value is of the same order as the standard deviation.

We conclude that a ΔL_o of -6 dB should be reckoned with for all bandwidths. The difference between the high-frequency sides of the input patterns of narrow-band pulsators and sine-wave pulsators can be fully explained by a criterion difference, the peaked shape of the summit pattern ($\Delta L_o = -6$ dB) and a shift depending on the bandwidth of the noise bands. The excitation patterns for sine-wave and noise-band pulsators have equal slopes.

VI-5-b. Analysis of the high-frequency side of the output pattern

A similar analysis can be made of the high-frequency side of the output pattern.

For a sine-wave pulsator we use equation (13):

$$L_s = \frac{L_o - k_s - b \ln \left((f_o + \Delta f) / f_o \right)}{1 + a \ln \left((f_o + \Delta f) / f_o \right)}$$

The coefficients a and b are from the equation $S(L) = aL + b$ giving the dependence of the slope of the excitation pattern on the stimulus level.

For a narrow-band pulsator we get:

$$L_n = \frac{L_o - \Delta L_o - k_n - b \ln \left((f_o + \Delta f_n + \Delta f) / (f_o + \Delta f_n) \right)}{1 + a \ln \left((f_o + \Delta f_n + \Delta f) / (f_o + \Delta f_n) \right)} \quad (35)$$

Again we compute $\Delta L = L_n - L_s$, assuming a ΔL_o of -6 dB. The values of a and b are the best-fit values, shown in Fig. 19.

We present the difference between the experimental pulsation levels ΔL_e only corrected for the bandwidth of the pulsator in the upper part of Fig. 28 and the difference between the experimental values ΔL_e and the calculated values ΔL in

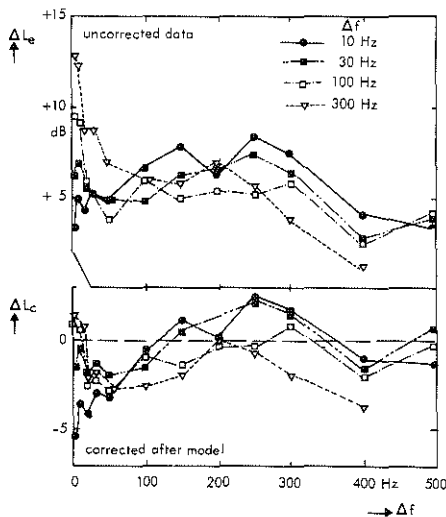


Fig. 28. Results of analysis for the high-frequency side of the output patterns of noise-band pulsators. The upper part represents the differences between the experimental values for the different bandwidths and the sine-wave pattern, corrected only for a shift over the noise bandwidth. The lower part shows the differences corrected in accordance with the model, taking into account differences in bandwidth and criterion and the peaking of the summit pattern. The broken line represents a difference of 0 dB.

the lower part of this figure. A difference of 0 dB is indicated by the broken line. The accuracy of these points is also about 1 dB for smaller frequency separations and 2 or 3 dB for larger frequency separations. In the present case, the points for frequency separations smaller than 100 Hz deviate from the zero-dB line for all bandwidths. The smaller bandwidths again show an extended summit pattern. For separations larger than 100 Hz the points are not statistically different from zero, except for some of the 300 Hz points. This may be due to the difference in equivalent level caused by the critical band mechanism.

We conclude that the difference between the pulsation levels for sine-wave and noise-band pulsators at the high-frequency side of the output patterns can also be explained by a criterion difference, the peaked shape of the summit pattern and a shift depending on the bandwidth of the noise. The differences in shape, mentioned in section VI-3-2 are not found here, so they are caused by an interaction of the nonlinearity and the above-mentioned differences.

VI-5-c. Analysis of the low-frequency side of the input pattern

A similar analysis can also be made of the low-frequency side of the input pattern. In the determination of this side of the patterns, the frequency of the probe was varied. In order to obtain comparable results, we shall transform these frequency differences into level differences.

For a sine-wave pulsator we use equation (23):

$$\ln \left(\frac{f'}{f_o} \right) = \frac{L - L_t}{S} - \frac{L_o - L_t}{S_o} - \frac{k_s}{S}$$

where L_t is the threshold of hearing.

For the noise-band pulsators we use modified forms of equations (8), (9) and (10).

Pulsator:

$$L_t = L_o + S_o \ln \left(\frac{f_t}{f_o - \Delta f_n} \right) + \Delta L_o$$

probe:

$$L_t' = L + S \ln \left(\frac{f_t}{f} \right)$$

at threshold:

$$L_t' = L_t + k_n$$

whence:
$$\ln \left(\frac{f}{f_o - \Delta f_n} \right) = \frac{L - L_t}{S} = \frac{L_o - L_t}{S_o} = \frac{k_n}{S} = \frac{\Delta L_o}{S_o} \quad (36)$$

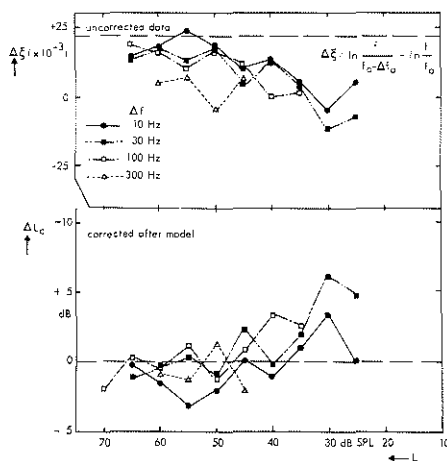
We define the frequency difference between the patterns for noise-band and sine-wave pulsators as:

$$\Delta \xi = \ln \left(\frac{f}{f_o - \Delta f_n} \right) - \ln \left(\frac{f}{f_o} \right) \quad (37)$$

whence:
$$\Delta \xi = - \frac{k_n - k_s}{S} = - \frac{\Delta L_o}{S_o} \quad (38)$$

We plot the values of $\Delta \xi$ for the experimental points in the upper part of Fig. 29.

Fig. 29. Results of analysis for the low-frequency side of the input patterns for various bandwidths. The points on the spectral bulge are not included. The upper part shows the uncorrected values of $\Delta \xi$ as defined in the figure. The broken line represents the value that could have been expected for a ΔL_o of -6 dB. The lower part shows the data, corrected for the criterion differences and the peaking of the summit pattern ($\Delta L_o = -6$ dB), transformed into level differences. The broken line represents a ΔL_c of 0 dB.



The broken line represents a ΔL_o of -6 dB if no correction for a criterion difference is made. The differences $\Delta \xi$ are now corrected for the criterion difference, for a ΔL_o of -6 dB and transformed to a level difference according to equation (38):

$$\Delta L_c = S_o \left(- \frac{k_n - k_s}{S} - \Delta \xi \right) + \Delta L_o \quad (39)$$

The computed level differences ΔL_c are plotted in the lower part of Fig. 29. The broken line represents a ΔL_c of 0 dB.

The accuracy of the points in this figure is rather poor because of the multiplication

by the high slope S_o . The standard deviation of the adjustments at the higher levels results in a standard deviation in the computed ΔL_c of about 3 dB. At lower levels the standard deviation is even larger. We conclude that the values of ΔL_c are not significantly different from 0 dB.

We conclude that the difference between the pulsation patterns for sine-wave and noise-band pulsators at this side of the input pattern can be explained just as at the high-frequency side, by a criterion difference, the peaking of the summit pattern and a shift depending on the bandwidth of the noise band.

VI-5-d. Analysis of the low-frequency side of the output pattern

The analysis of the low-frequency side of the output pattern follows the same lines as that presented in section VI-5-c. We again use $\Delta \xi$ as defined by equation (37) and derive in a very similar way:

$$\Delta L_c = S \left(- \frac{k_n - k_s}{S_o} - \Delta \xi \right) + \Delta L_o \quad (40)$$

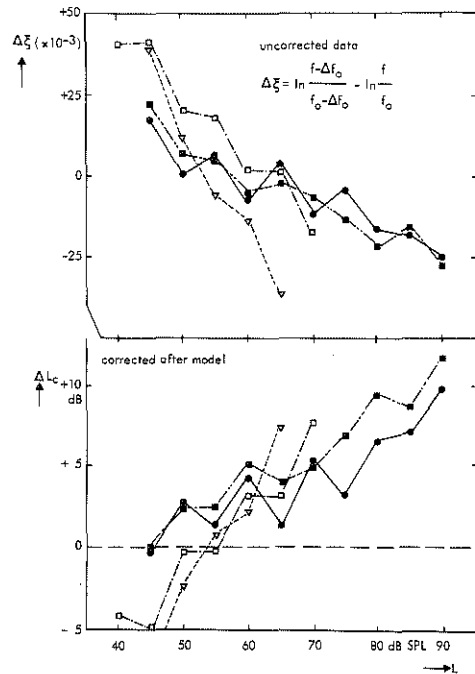
The values of $\Delta \xi$ for the experimental data are presented in the upper part of Fig. 30. We computed ΔL_c according to equation (40) and we present these values in the lower part of this figure.

The accuracy of the points is even worse than in Fig. 29. Near 40 dB the accuracy is again 3 dB but for higher stimulus levels we have to multiply the $\Delta \xi$ by the very high slope S , giving transformed standard deviations exceeding 10 dB.

The analysis is rather unreliable because of these large standard deviations. The trend does seem to indicate a difference between the slopes of the noise-band and sine-wave patterns. However, as a result of the large standard deviations, these differences can be considered as small.

We conclude that the data available do not provide evidence for a difference between the slopes of pulsation patterns for narrow noise-bands and sine-wave pulsators, if

Fig. 30. Results of analysis for the low-frequency side of the output patterns for various bandwidths. The upper part shows uncorrected values of $\Delta\xi$. The lower part shows values corrected for the criterion differences and for the peaking of the summit pattern ($\Delta L_o = -6$ dB), transformed into level differences. The broken line represents a ΔL_c of 0 dB.



the criterion difference, the peaking of the summit pattern and the width of the noise band are taken into account.

VI-6. Discussion

VI-6-1. Summit effects

We found a difference between the pulsation levels at the central frequency of 1 kHz for sine waves and noise bands of equal energy. Houtgast (1974) also found this difference. His two observers showed an average difference of -2.8 dB, which is very close to our average value of -2.5 dB for two observers.

Our pulsation level for the 300 Hz noise band is lower than the levels at other

bandwidths. Houtgast (1974 a) studied the dependence of the pulsation level at the central frequency on the bandwidth of the noise. He found a constant difference between noise-band levels and the sine-wave level for bandwidths between 20 Hz and 150 Hz. We also found a constant difference at 10 Hz, 30 and 100 Hz. Our pulsation levels for the 300 Hz noise band differed by -3.5 dB from the constant level for the smaller bandwidths. Houtgast found differences between -6 dB and -9 dB.

Houtgast explains this large drop in pulsation level between 150 Hz and 300 Hz as being due to suppression. It is not clear why we found a smaller drop; it could be due to differences either in critical bandwidth or in suppression. We did not investigate this further.

The summit patterns of our two observers differed considerably. We do not know of any explanation for this, nor could we find comparable data in the literature.

VI-6-2. Patterns outside the bandwidth of the noise

Our analysis of the pattern outside the bandwidth of the noise showed that the differences between the patterns for a noise-band and a sine-wave pulsator can be explained for the high-frequency side by a criterion difference, a peaking of the summit pattern and a shift depending on the bandwidth of the noise. This is even true for the output pattern, although the shapes seem slightly different in these cases. This agreement, in spite of the difference in shape, shows once more that nonlinearity has unexpected consequences for the shape of the pulsation patterns. The analysis of the low-frequency side is less reliable because of the poor accuracy involved. Nevertheless, the differences between the various input patterns are well explained by the above-mentioned three factors. The least reliable and most widely differing results were found for the output patterns. The trend of the results points to a certain difference in slope between noise-band and sine-wave patterns, but the observed differences are rather small in comparison with the standard deviations in the adjustments. We cannot therefore consider this as evidence for a different slope in noise-band patterns.

We thus find no evidence in our data for a difference between the slopes of sine-wave and narrow-band pulsation pattern, nor for a difference between the dependence of the slope on level in the two cases.

The spectral broadening shows that we have to be very careful with the spectra of our signals. The slopes of the pulsation patterns are very high and especially the low-frequency side of the extension pattern is easily contaminated by spectral effects.

VI-6-3. Comparison with the literature

VI-6-3-a. Slopes of sine-wave patterns

Data on the slopes of pulsation patterns have been presented by Fastl (1974 b; 1975) and by Houtgast (1974 a, b). These authors determined the slopes of the steep side of input extension patterns. Houtgast found a slope of the sine-wave pattern at 1 kHz and 60 dB SL of 250 dB/oct or, in the terms of Fastl, of 75 dB/Bark. Our slope at 1 kHz and 70 dB SPL is 240 dB/oct or 70 dB/Bark (Fig. 13). Fastl (1974 b) found an initial slope of 50 dB/Bark at 60 dB SPL.

Houtgast's slopes and our slopes are quite comparable, while the value found by Fastl is somewhat lower. This difference could be due to the different levels used, which may be expected to give rise to a different slope as a result of the nonlinearity.

Fastl (1975) found a slope of 75 dB/Bark at 2 kHz and 70 dB/SPL, where we find a slope of 70 dB/Bark (Fig. 13) , so here the slopes at equal levels are nearly equal.

We conclude that the slopes of the various authors are quite comparable under comparable conditions.

VI-6-3-b. Slopes of sine-wave and noise-band patterns

The data published on this matter are conflicting. Houtgast found that the steep sides of the sine-wave and noise-band pulsation input extension patterns have the same slope. We also came to the conclusion that the slopes do not differ and further that they showed the same nonlinear behaviour. Fastl (1974 b, 1975), however, found the slopes of critical band pulsation patterns to be lower than the corresponding slopes of sine-wave patterns.

The only explanation for the difference we could think of was the spectral composition of the critical band noise. The slope of the spectrum of the critical band noise is determined by the absence of a broadening effect in masking patterns. The slopes of masking patterns are lower than those of pulsation patterns (see next section). It is, therefore, conceivable that the spectral composition of the critical band noise does not show up in the edge of the masking patterns, but that it does so in pulsation patterns. We could not check this point because the spectral composition of the critical band noise was not specified in Fastl's paper.

VI-6-3-c. Slopes of pulsation patterns and masking patterns

One of the matters raised in section VI-1 was whether pulsation patterns and masking patterns reflect the same frequency selectivity or not. Fastl (1974 a) found equal low-frequency slopes for pulsation input patterns and masking patterns if a critical band noise is used as pulsator and masker. The slopes differed if a sine wave was used as pulsator and masker. Houtgast (1974 a, c) found differences between the low-frequency slopes of pulsation patterns and masking patterns, irrespective whether the pulsator and masker were sine waves or narrow noise bands. We found no difference between the slopes of the pulsation patterns for sine waves and narrow noise band. This finding applied to both the low-frequency and the high-frequency sides. In view of the above discussion between Fastl and Houtgast it seemed appropriate to compare our pulsation patterns with masking patterns for the same observer.

Observer MB measured masking filter patterns and pulsation filter patterns for three levels at 1 kHz with a sinusoidal input stimulus. We preferred a sine wave to a narrow noise band in the interests of avoiding any spectral effects.

The masking patterns were transformed to permit comparison with the pulsation patterns. This transformation was performed with reference to the dynamic characteristic at 1 kHz, which was measured by determining the detectability threshold of a 1 kHz sine wave in the presence of a louder signal of the same frequency. It allows the transformation from a maskee (output) level to the masker (input) level at that particular frequency. The sensitivity curve is given in the inset of Fig. 31. This curve is a straight line which can be represented by the equation $L_m = 1.24 L_p^{-0.4}$, where L_m is the masker level and L_p the maskee or probe level.

Our output filter patterns were measured for a probe with a constant level at the centre frequency of 1 kHz. This probe level has to be transformed; the whole pattern thus has to be shifted in accordance with the sensitivity curve. The pulsation output patterns and the shifted masking output patterns are shown in the upper part of Fig. 31.

Our input filter patterns were measured using a probe with a variable level at the centre frequency. In this case all experimental points have to be transformed in accordance with the sensitivity curve. The pulsation input patterns and the transformed masking input patterns are shown in the lower part of Fig. 31.

The masking patterns can be divided into three distinct regions: one below 0.8 kHz, a middle region between 0.8 kHz and the centre frequency of 1 kHz and one above 1 kHz. This distinction between three regions has already been made by Rodenburg et al. (1974).

The regions below 0.8 kHz and above 1 kHz show clear differences between the pulsation and masking patterns. Houtgast and Fastl only considered the steep (in this case the high-frequency) side. We conclude that our results corroborate Houtgast's findings and that they are in disagreement with those of Fastl.

In the region between 0.8 kHz and 1 kHz the pulsation patterns and the masking patterns agree quite well. Greenwood (1971) showed that the masking pattern in this region is based on the detection of combination tones, which are not heard in these pulsation experiments. An argument against the involvement of combination

tones in pulsation measurement is the absence of a break in the pulsation patterns corresponding to that in the masking pattern at 0.8 kHz. We think that the close correspondence between pulsation patterns and the threshold of the audibility region for combination tones (which is what the masking pattern is in this region) suggests a common mechanism underlying combination-tone generation and the nonlinearity causing the slope-level dependence of pulsation patterns.

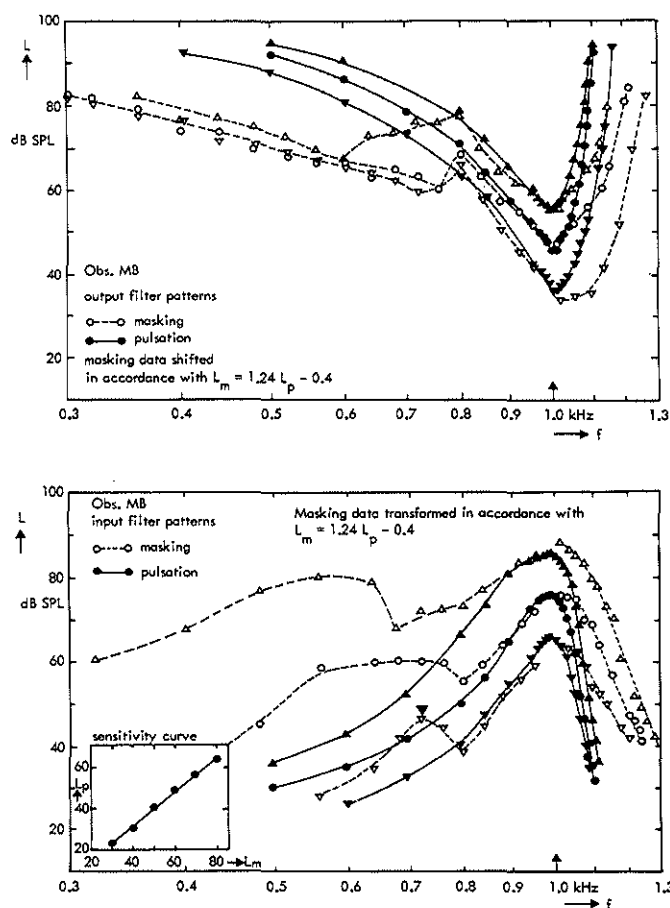


Fig. 31. Comparison of masking and pulsation input filter patterns. The open symbols give the masking pattern transformed to the equivalent masking level of a 1 kHz sine wave, in accordance with the sensitivity function $L_m = 1.24 L_p - 0.4$, shown in the inset. The solid symbols give the pulsation patterns. The upper part of the figure gives the output patterns and the lower part the input patterns for observer MB.

The nonlinearity of the masking patterns is also apparent from Fig. 31. The nonlinearity in all three regions is of the same kind as that found in pulsation. The nonlinearity in the low-frequency region is particularly striking: the points of the output patterns for different levels differ by less than a standard deviation in the adjustments and cannot be considered as different.

The difference in tuning between the excitation patterns calculated from the pulsation patterns and masking patterns outside the combination-tone region is remarkably large, and does not point to a common mechanism underlying the two methods.

We may conclude that both pulsation and masking patterns show auditory nonlinearity. The slopes of the excitation patterns are much steeper on either side than those of masking patterns if the combination-tone region is not considered. Masking patterns thus reflect a much poorer tuning than pulsation patterns do. Both reflect a nonlinear behaviour. Pulsation patterns and masking patterns coincide in the region where combination tones determine the masking threshold.

VII. NONLINEARITY IN NEUROPHYSIOLOGICAL EXPERIMENTS

VII-1. Introduction

The pulsation patterns presented in Chapter IV and VI revealed a nonlinearity of the auditory system. Our analysis of this nonlinearity in mathematical terms in Chapter V indicated some consequences of the nonlinearity for the shape and steepness of the patterns but did not show at what stage of processing in the auditory system this nonlinearity is located. The localization can only be found by comparing psychophysical and physiological results.

The first stage of sharp auditory tuning is found in the mechanism of the basilar membrane vibrations. The linearity of these vibrations has received a fair amount of attention in the literature. The data presented so far have been rather conflicting.

Rhode (1971) used a Mössbauer technique for the measurement of the vibrations. He placed a tiny radioactive source on the basilar membrane and measured the Doppler shift of the radiation caused by the velocity of the motion of the membrane. He found nonlinear vibrations, at least in a restricted frequency range near the best frequency of the spot to which the source was attached.

This nonlinearity was investigated further by Rhode & Robles (1974). It was theoretically explained by Kim et al. (1973) as a nonlinear damping of the basilar membrane motion.

Wilson and Johnstone (1972, 1975) used a capacitive probe technique. The probe served as one plate of a capacitor and the basilar membrane as the other. Displacement of the membrane caused a change in the capacitance. They covered a wide range of levels with this technique and found the vibrations to be fully linear.

The next stage of the auditory processing is the detection of the vibrations by the sensory cells and the generation of neural activity. The neural activity can be measured in the primary nerve fibres. The recordings of spike activity in primary nerve fibres show clear nonlinear effects such as suppression and combination tones (specifically odd-order difference tones), but there are only a few publications on the subject of linearity of response patterns.

The first response patterns were published by Kiang (1965). After applying an electrical contact to a nerve fibre he determined the level of a sine wave of different frequencies that produced threshold activity in the fibre. The tuning curves determined in this way have a shape very similar to that of our psychophysical output patterns, described in Chapter IV. In fact, they are output filter patterns in our terminology, since the input (acoustic stimulus) is varied in level and frequency while the output (activity level in a certain fibre) is kept constant. The curved shape and extreme steepness of the steep (high-frequency) side of the tuning curve suggests that a similar nonlinearity is involved here and in the psychophysical experiments.

Stronger evidence for the involvement of a nonlinearity is found in the response areas described by Anderson et al. (1971). They kept the level of the stimulating sine wave constant and swept its frequency. The output was the neural activity recorded in a fibre. The activity level was expressed as the number of spikes in some time interval. In our terminology the response areas* are input filter patterns. The response areas are nonlinear. This is very clear at the steep (high-frequency) side. The frequency at which the activity is at the threshold level for different stimulation levels is almost independent of the stimulation level, while the activity levels at the best frequency show large differences. This results in steep tuning curves.

*Also called iso-intensity areas by other authors.

The effect of level on the other (low-frequency) slope cannot be described as clearly in terms of linearity or nonlinearity. The part of this pattern that is within the dynamically interesting range is small and shifts in frequency. The peak of the patterns is in the saturation area for quite a number of levels. These facts do not permit the determination of a slope at this side of the pattern.

Saturation is, in fact, one of the basic problems in the comparison of physiological data of this kind with psychophysical data.

In psychophysics we measure level as a function of frequency; in the determination of response areas we measure the neural activity by counting the number of spikes in some time interval as a function of frequency. We thus need a transformation from the number of spikes to a stimulation level in order to make a comparison. The most appropriate choice of transformation curve seems to be the curve giving the number of spikes as a function of stimulus level (the spike-rate curve) at the best frequency.

The linearity of the responses has been investigated by de Boer (1967). He used Gaussian white noise as the stimulating signal and recorded fibre responses. The transfer characteristic, or at least its linear part, was determined by reversed correlation of the white noise signal and the fibre response. De Boer showed that the transfer characteristic does not depend upon the level of the white noise.

Evans (1972) determined frequency threshold curves, which are similar to tuning curves. In a recent paper (Evans, 1977) he found a linear behaviour up to a level of about 70 dB SPL, which was in his experiment about 45 dB above threshold, and a nonlinear behaviour at higher levels.

Møller (1977) also investigated the linearity of the primary fibre response, using cross-covariance functions of a stimulating noise and the response it evokes in the fibre. He concluded that this type of analysis, which is only sensitive to nonlinearities in the tuning properties, reveals a nonlinearity. The sharpness of tuning depends on level, primarily because of the change in slope of the shallow edge of the pattern. The slope is the highest at the lowest levels. He also found a shift of the frequency of maximum response with changing level.

The results of de Boer and of Evans suggest the absence of a nonlinearity in primary

fibre recordings at the levels we used in our psychophysical experiments. The results of Anderson et al. and Möller suggest, however, that a nonlinearity is present. Further analysis of neurophysiological data with regard to linearity would thus seem to be useful. An essential step in such an analysis will be the transformation of the data into a form which is directly comparable with our psychophysical results.

The next stage of processing in the auditory system is found in the cochlear nucleus. Here many authors have observed strong nonlinear effects like suppression and inhibition (Möller, 1976). One also finds a much wider dynamic range in the cochlear nucleus than in the primary fibre recordings (Evans & Palmer, 1975). Smoorenburg & Linschoten (1977) showed very recently that a nonlinearity of the kind found in our psychophysical experiments is also present at this stage.

The necessary primary nerve data for an analysis were given to us by Dr. M. Rodenburg of our department, who has stayed at Keele University in the U.K. for a while where he has been analysing primary nerve data collected by Dr. E.F. Evans.

VII-2. Methods and results

We use primary response data recorded on cats by Evans. For details of recording equipment and physiological conditions we refer to Evans (1972). Rodenburg constructed spike-rate curves from these recordings. These curves show the average response of a fibre, characterized by its frequency of best response, as a function of stimulus level at a fixed stimulus frequency. The response is given as the number of spikes in an arbitrary defined time interval. The spike-rate curve for each fibre was constructed for a number of stimulus frequencies. Some examples of these curves are shown in Fig. 32.

We constructed input and output filter patterns from these spike-rate curves. For this purpose we assumed that the spike-rate curve at the best frequency represents the transformation from neural activity to an equivalent level of a stimulus at the best frequency, this stimulus being taken as the probe in pulsation experiments, on

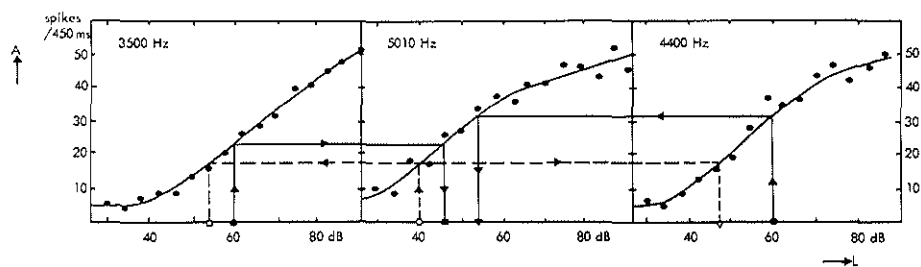


Fig. 32. Some spike-rate curves of unit 266-18 in the cat with a best frequency of 5.0 kHz constructed by Rodenburg from the results of Evans. We show the procedure for construction of input and output filter patterns from the spike-rate curves of the primary auditory fibres. The symbols on the L axis refer to the points shown in Fig. 33. For full explanation see text.

the assumption that there is no off-frequency detection. An example of the construction procedure is shown in Fig. 32 for unit 266-18, which had a best frequency of 5010 Hz.

For the construction of the e.g. 60 dB input filter pattern we read the number of spikes from the spike-rate curve for some frequency, e.g. 23 spikes/450 ms from the 3500 Hz curve. At the best frequency, 5010 Hz, the same number of spikes is elicited by a sine wave of 46 dB. It follows that a 3500 Hz, 60 dB sine wave elicits the same activity in this fibre as a sine wave of 5010 Hz and 46 dB.

This procedure is in accordance with our criterion for the pulsation threshold for on-frequency detection and a jnd of 0 dB. The probe corresponds to the fibre's best frequency, 5010 Hz. Now an alternation of a 5010 Hz, 46 dB sine wave and a

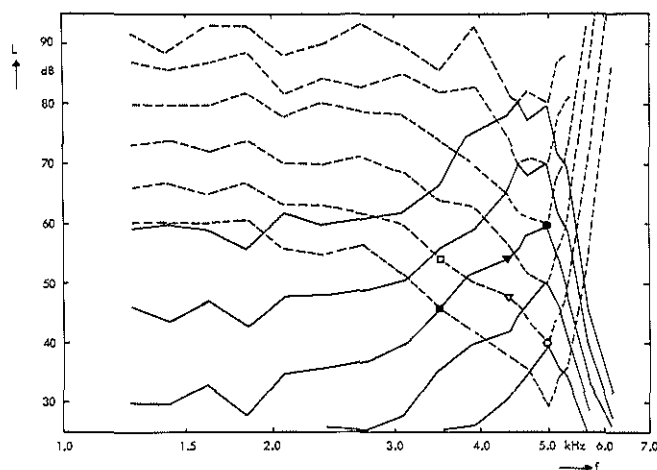


Fig. 33. Input and output filter patterns of unit 266-18 (Evans) with a best frequency of 5.0 kHz. The symbols refer to the explanation of the construction procedure (see text). The full lines represent the input patterns and the broken lines the output patterns.

3500 Hz, 60 dB sine wave would cause a continuous activity of 23 spikes/450 ms, which is assumed to be necessary for perception of the probe as continuous; we assume further that the absence cannot be detected at some other frequency (or in physiological terms, in some other fibre), so off-frequency detection is impossible.

In the same way we find the 4400-Hz point of the 60 dB input pattern at 54 dB, since a 60 dB sine wave of 4400 Hz elicits 30 spikes/450 ms in this 5010 Hz fibre and a sine wave of 5010 Hz and 54 dB would elicit the same 30 spikes/450 ms.

The output pattern is constructed in a very similar manner, except that the reasoning is reversed. Let us for example find the 40 dB point of the output pattern at 3500 Hz. A 40 dB sine wave at the best frequency elicits 17 spikes/450 ms. The same activity is elicited by a 3500 Hz sine wave at 54 dB. The 40 dB point of the output pattern at 3500 Hz is thus to be plotted at 54 dB. Similarly we find a level of 48 dB for a frequency of 4400 Hz.

These points can be constructed for every frequency for which we have a spike-rate curve. Fig. 33 shows the constructed input and output patterns of unit 266-18 (5.0 kHz). The full lines represent the input patterns and the broken lines the output patterns. The symbols show the points used for the explanation of the construction procedure. Fig. 34 gives some more patterns for fibres with best frequencies of 1.1, 2.27, 8.62 and 12.7 kHz. In all cases the full lines represent the input patterns and the broken lines the output patterns.

The patterns show clear signs of nonlinearity. This is particularly clear for the high-frequency side of the input patterns. The steepening effect is very clear and its level dependence can again be approximated by a linear function $S(L) = aL + b$, just as in the pulsation analysis. The slope-level functions are given in Table II on page 125. The coefficient (b) given in Table II, cannot be compared with one another, or with the psychophysical results, because of the use of different reference levels for the dB scale at different frequencies.

The output patterns have the familiar curved shape at the high-frequency side. This curvature is a result of the nonlinearity, as was shown in Chapter V for the

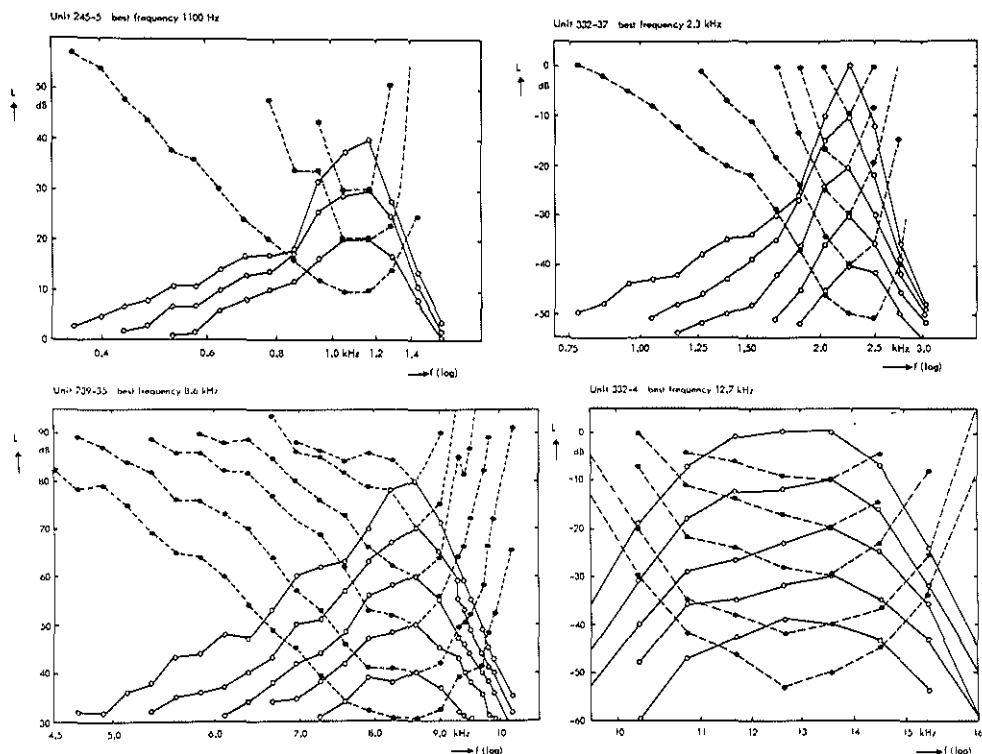


Fig. 34. Input and output filter patterns constructed from spike-rate curves from primary auditory fibres in the cat (data collected by Evans). The full lines represent the input patterns and the broken lines the output patterns.

pulsation patterns. The derived equation for the slope of the patterns (equation (29)) is also applicable to these output patterns, except that the detection level now corresponds to the number of spikes at the best frequency for which the curve was constructed. This level is 40 dB for the 40 dB pattern. Calculations of the slope of the patterns agree very well with the experimentally observed values.

The analysis of the low-frequency (shallow) edge of the patterns is more complicated. For low-frequency units (245-5, 1.1 kHz and 332-37, 2.3 kHz) saturation effects interfere with the construction procedure. Saturation causes a levelling-off of the spike-rate curve, resulting in a maximum number of spikes that can be elicited

by a sine wave. For these low-frequency units this maximum depends on the frequency. For unit 245-5 (1.1 kHz) this maximum is about 50 spikes/450 ms at 640 Hz, about 60 spikes/450 ms at 860 Hz and about 76 spikes/450 ms at 1160 Hz. The maximum activity at 860 Hz corresponds to an equivalent level of about 19 dB. This means that an input pattern can never exceed this 19 dB at 860 Hz. This causes an inflection in the curve for unit 245-5 at 860 Hz as seen in Fig. 34, and a similar inflection at 1.87 kHz for unit 332-37 (2.3 kHz). Saturation, possibly in combination with inhibition phenomena, thus makes our construction procedure unreliable and leads to results that are difficult to interpret.

Higher-frequency units do not show this dependence of the maximum spike-rate on frequency, but here the results are not very consistent.

The patterns for unit 332-4 (12.7 kHz) show a rather broad summit of the pattern and a gradual increase in slope as the frequency separation increases. This makes good definition of the slope impossible. In any case there is no clear sign of nonlinearity, either in the input or in the output patterns.

The input patterns for unit 239-35 (8.6 kHz) seem to be almost parallel. There is no evidence of nonlinearity. The output patterns seem to level off, but this might be due to the fact that we are considering filter patterns instead of extension patterns. Filter patterns in pulsation threshold experiments also showed this levelling-off. This finding was explained as being due to the effect of frequency on the nonlinearity.

The last set of data presented is for unit 266-18 (5.0 kHz). These data show clear levelling-off, explained above as due to frequency dependence of the nonlinearity. Apart from this, we observe nonlinearity in the upper parts of the pattern. The slope-level dependence can be described by the equation $S(L) = 0.27 L - 59$ with $S(L)$ expressed in dB per unit of the natural logarithmic scale (dB/e unit). The nonlinearity is also clear from the flat part of the patterns near 1.5 kHz. Raising the input stimulation level by 10 dB results in a 15 dB rise in the output level. This means that the dynamic characteristic has a slope of 1.5. For the output pattern this would imply a rise of 6.7 dB per 10 dB rise in stimulation level, which is actually found.

VII-3. Discussion

For the discussion of the physiological evidence for nonlinearity, we have to distinguish between the two sides of the activity patterns. At the steep side the non-linear behaviour is clear; at the shallow side we encountered some problems.

A spike-rate curve can be considered as a kind of dynamic characteristic. A non-linearity on the steep side, as found in section VII-2, would mean a dynamic characteristic with a slope of less than one; that on the shallow side, if it exists, would mean a slope of more than one. In our construction procedure the reference spike-rate curve is the curve at the best frequency. Hence this curve will have a slope of one by definition. We can rephrase the above considerations as follows: the spike-rate curve for frequencies above the best frequency should be less steep than the curve at the best frequency; for frequencies below the best frequency, the spike-rate curve should be steeper.

Sachs and Abbas (1974) analysed spike-rate curves and actually found a slope-frequency dependence of the spike-rate curves of this nature. The major difference between their analysis and ours is that they normalized the maximum firing rate, which we did not. At all frequencies the maximum rate was set to one and then the slope was determined. This would mean that our low-frequency spike-rate curves have to be transformed into curves with the same maximum firing rate. If we do this, we find a slope-frequency dependence as described by Sachs and Abbas with a nonlinearity for both sides of the patterns.

Fig. 35 shows the correction factor for the normalization of the spike-rate curves, and the input patterns constructed from these normalized curves. The nonlinearity at the steep side is still present although we cannot follow the course of the curve down to near the threshold because the spike-rate curves do not level-off any more and the normalization factor is unknown. The summit of the patterns is rather broad and its edges are straight lines, that show a slight level dependence as could be expected from the slope of the normalized spike-rate curves. We see that the procedure proposed by Sachs and Abbas introduces the nonlinearity, although this is better visible in the spike-rate curves than in the input pattern.

These normalized input patterns must be interpreted with caution. The firing rate is no longer the determining factor, but the ratio of firing rate to maximum rate for the fibre in question at a particular frequency. The coding is no longer simply a one-to-one relation of spike-rate and level but a complicated procedure involving frequency ratio, best frequency and level. A further analysis of these possible coding features goes beyond the scope of this thesis and would end in conjectures.

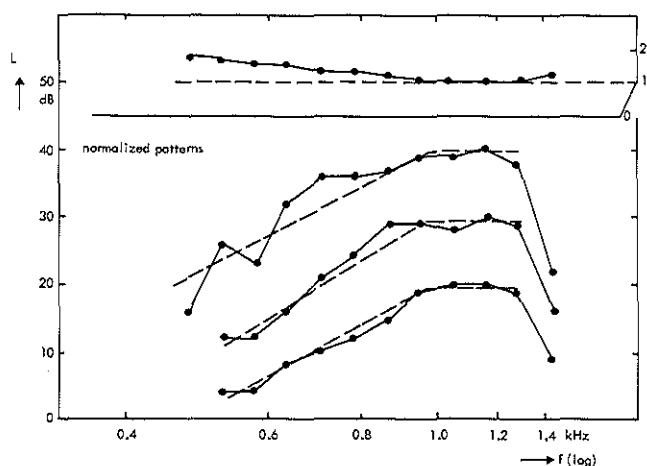


Fig. 35. Input filter patterns of unit 245-5, best frequency 1.1 kHz, after normalization of the spike rate curves. The normalization factor is shown in the upper part and the resulting patterns in the lower part.

We see in any case that the patterns are nonlinear at the steep side and possibly at the shallow side too. This nonlinear behaviour could be observed indirectly in the shape of the tuning curves and in the response areas, as we have argued in the introduction to this chapter.

The fact that de Boer and Evans did not find a nonlinearity, however, remains. The procedure used by de Boer is entirely different from ours. His "revcor" technique (de Boer, 1969) is only partly sensitive to nonlinearities and this sensitivity depends entirely on the way the nonlinearity is coded.

The technique used by Møller (1977) is comparable with that of de Boer and Møller found evidence for a nonlinearity of the kind we found in our psychophysical experiments. He also found evidence for a shift in the frequency of maximum response.

A shift in maximum response frequency was also found by Smoorenburg and Linschoten (1977) in the cochlear nucleus, and may contribute to the psychophysical nonlinearity. The shift would be translated into a steepening of the steep side and a decrease in slope of the shallow side, because the activity patterns of probe and

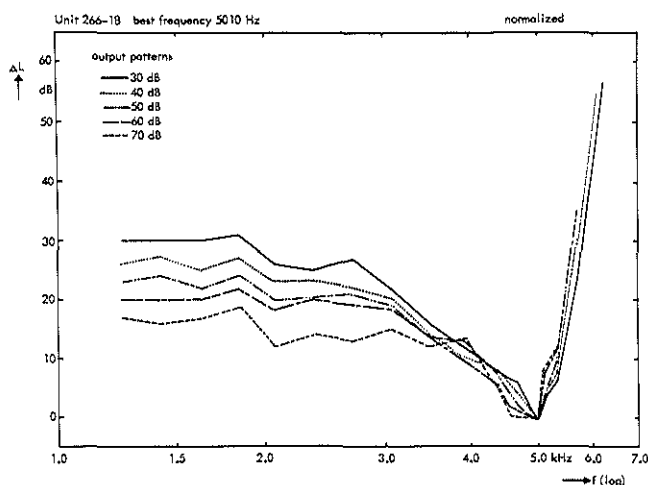
pulsator would be displaced. The pulsation threshold method does not permit the detection of pattern shifts, except as changes in the slope of the patterns.

Response areas do not permit the detection of a shift in the frequency of maximum response because of their limited dynamic range which causes that the peak of the pattern is hidden in the saturation region for higher levels. The technique of Möller and of Smoorenburg & Linschoten permits the detection of the shift because they perform a Fourier transform on a kind of impulse response.

We did not find a clear shift in the frequency of the maximum response in our analysis. If it does exist, it must be small; smaller than the shifts found by Möller and Smoorenburg & Linschoten.

Evans found linear frequency threshold curves for lower levels while we found nonlinear patterns for these levels. This contradiction might be due to our use of input curves. We have seen that the clearest nonlinear effect is found at the steep side of the input patterns. At its foot the dynamic range is limited to only a few dB for large differences in stimulation level. The consequence of this finding is a very steep output pattern. This steepness, together with the associated curvature, makes the definition of steepness somewhat arbitrary and tends to hide the nonlinearity. Fig. 36 shows this effect. The output patterns are shifted in order to give coincidence of the 5.0 kHz points.

Fig. 36. Output filter patterns for primary auditory fibres in cat. The curves are shifted in level to permit comparison of the forms at different levels.



This approach was suggested by Evans (1977). It shows only minor differences and seems to justify the conclusion of linearity at lower levels. The input patterns,

however, show that this is not correct, as can be seen from Fig. 33. The method of comparing output patterns does not seem to be sufficiently sensitive to show the nonlinearity for small differences in stimulation level. This is particularly so because the nonlinearity has an appreciable effect on the shape of the patterns.

In this connection it is interesting to read the discussion between Pfeiffer, Evans and Zwislocki in Möller (1973) about the use of input and output patterns. Pfeiffer preferred the use of input patterns; Evans preferred the use of output patterns because the activity level in one fibre is constant in that case and it would not be influenced by a nonlinearity. The conclusion of the discussion was that there were no data available to justify any preference. We think that our analysis of the data justifies a preference for input patterns. In these patterns the nonlinearity is constant for each level and clearly visible at different levels. In an output method tuning and nonlinearity interact a complicated way.

We conclude from the above that the primary nerve responses are nonlinear and that the results of Evans do not exclude the possibility of nonlinearity. The results of de Boer and Möller using similar methods, seem to conflict with each other; we do not know of any explanation for this.

Numerical analysis of the nonlinearity shows it to be almost frequency-independent. The value of the coefficient (α) in the slope-level dependence is about 3 on the steep side and about 0.5 on the other side (see Table II). A value of 0.5 is also found at 8.6 kHz if we do not take the two summit points of the patterns into account. If we only consider the points of the sloping parts of the patterns there is no inflection; they are on straight lines and show the nonlinearity in differences of the slopes of the patterns.

Under the above restriction, and the restriction of normalizing the low-frequency spike-rate curve, we find the nonlinearity on both sides of the patterns. The value of (α), describing the nonlinearity, does not depend on the best frequency of the fibre. The frequency-independence of the nonlinearity does not imply that the slopes of the patterns are frequency-independent.

In Fig. 37 we show the range of the slopes found in our patterns between background

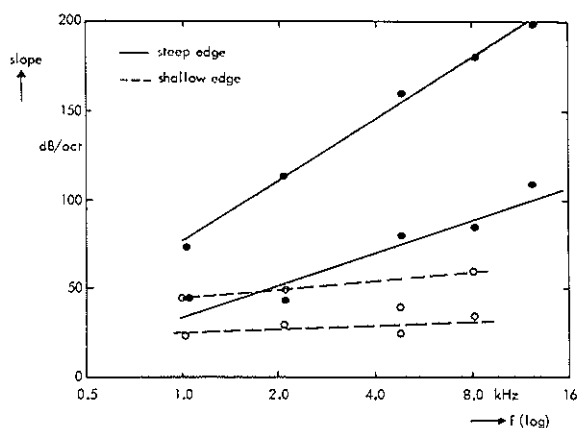


Fig. 37. Range of slopes of input patterns of primary auditory nerve recordings in the cat, obtained from the figures 33 and 34. The open circles and broken lines represent the minimum and maximum slopes of the low-frequency side of the input patterns. The dots and full lines represent the minimum and maximum slopes of the steep side of the pattern.

nervous activity and saturation. The dynamic range is about 40 to 50 dB. This figure shows that the steep edges of the patterns are steeper for higher frequencies than for lower frequencies; the dependence is approximately a straight line in a dB/oct vs $\log f$ plot. The slope of the shallow edge shows only a slight dependence on frequency, but again the higher slopes are found at higher frequencies. A comparison of the tuning at the various stages of the auditory system will be presented in the next chapter, where we shall try to place our findings in the general context of theory of hearing.

Table II. Slope-level functions for physiological data. The slope is expressed in dB per unit of the natural logarithmic scale.			
unit	best frequency	high-frequency side	low-frequency side
245-5	1.1 kHz	$2.89 L + 10$	$1.15 L - 25$
332-37	2.3 kHz	$2.73 L + 163$	$0.77 L - 46$
266-18	5.0 kHz	$2.88 L$	$0.27 L - 59$
239-35	8.6 kHz	$3.25 L$	$0.50 L - 113$
332-4	12.7 kHz	$3.05 L + 286$	-----

VIII. NONLINEARITY AND THEORY OF HEARING

It seems appropriate at this point to consider how our findings relate to previous publications and what their consequences are for the theory of hearing. The literature on nonlinearity in physiological determinations of tuning has already been discussed in Chapter VII.

VIII-1. Nonlinearity and masking patterns

Nonlinearity has been described in connection with masking experiments. The first quantitative masking data were published by Wegel & Lane (1924). They presented their data primarily as dynamic characteristics, measuring the maskee level at the detection threshold as a function of masker level for a number of frequency combinations.

Chistovich (1971) described the nonlinearity of the auditory system as a dependence of the slope of the dynamic characteristic on the ratio of maskee frequency and masker frequency. The slope is one for frequency ratios near one if the beat region is avoided, less than one for frequency ratios below one, and greater than one for ratios above one. This dependence is also found in the results of Wegel & Lane.

A study of the nonlinearity of dynamic characteristics and a more explicit description of the dependence of their slope on the frequency ratios has recently been published

by Schöne (1977).

Wegel and Lane also published a masking pattern with a nonlinearity, involving a steepening of the low-frequency (steep) side of the pattern with increasing level and an opposed effect at the high-frequency side. This tilting of the masking pattern has been found by many authors too e.g. Chistovich (1957) and Rodenburg et al. (1974). Zwicker and Feldtkeller (1967) found a nonlinear high-frequency edge of the masking patterns measured for narrow noise-band maskers, while the slope of the steep low-frequency edge of their masking patterns did not depend on level.

The nonlinearity of the masking pattern suggests that masking patterns can be analysed in a way similar to that described in Chapter V for our pulsation patterns. This is not possible in the absence of an explicitly stated, uniquely determined detection criterion.

Wegel & Lane (1924) suggested that detection occurs if the vibration amplitude at the spot along the basilar membrane corresponding to the maskee frequency exceeds that of the masker at the same spot (S/N ratio of 0 dB). This criterion resembles our continuity criterion (assumption 2 of section V-2) and Gardner's detection criterion for postmasking (Gardner, 1947). It can take the pattern nonlinearity into account, but not other nonlinear phenomena.

Nonlinearity presents three main complications to an analysis of masking patterns like the one for pulsation patterns. In the first place, Greenwood (1971) has shown that cubic combination tones determine the shape of the upper part of the masking pattern at the side of the shallow edge of the pattern (see section VI-6-3-c and Fig. 31). Another complication of nonlinearity is the invalidity of the superposition theorem: we may no longer assume that the combined effect of two superimposed sine waves is the sum of the effects of each sine wave separately. The processing of one sine wave, may thus interfere with the processing of the other (e.g. owing to suppression effects).

A third complication is found in the different behaviour of the steep side of the masking pattern for sine-wave and narrow-band maskers. Zwicker & Feldtkeller (1967) claim that the slope is constant for narrow-band maskers and depends on

level for sine-wave maskers. This suggests a dependence of the slope of the pattern on the bandwidth of the masker, which has to be formulated before we can analyse the patterns. On the above considerations we refrained from attempting full analysis of masking patterns.

If we consider these three complications with reference to our pulsation analysis, we can say that it is very unlikely that combination tones are of importance (section III-6). Interactions are unlikely, as there is no superposition of the stimuli. The third complication is concerned with differences in behaviour between sinusoidal and narrow-band stimuli. We have seen in Chapter VI that the tuning properties of the auditory system are equal for both kinds of stimuli. Pulsation patterns thus do not suffer from these complications.

VIII-2. Localization of the nonlinearity

An interesting point is where the "slope-level" nonlinearity is generated in the auditory system. We discussed the linearity of the basilar membrane vibrations and of the primary fibre response in Chapter VII and found conflicting data at both stages of processing. Our analysis showed that some of the discrepancies in primary fibre results can be understood and that we had to accept a nonlinear primary fibre response. This was also clear from the normalized spike-rate curves of Sachs & Abbas (1974). The slope of the curves depended on the frequency ratio, which is in agreement with the psychophysical findings of Wegel & Lane (1924) and of Chistovich (1971). A similar dependence was found in our pulsation patterns.

In this respect it is interesting to consider Rhode's nonlinearity (Rhode, 1971; Kim et al., 1973). Rhode found a slope of one for the dynamic characteristic except for frequencies near the best frequency, where the slope is less than one. This implies that the nonlinearity found by Rhode in the basilar membrane vibrations is neither the nonlinearity found in the primary fibre responses, nor that found in the pulsation patterns.

We conclude that the nonlinearity is located between the cochlear mechanics and the generation of spike activity in the primary fibres.

VIII-3. Effect of nonlinearity on frequency selectivity

The nonlinearity has consequences for the measurement of and the measures chosen for the frequency selectivity. A comparison of measures determined at different stages of the auditory system must include a comparison of the various methods that have been used and a consideration of the limitations of the method with regard to the nonlinearity. Various experimental methods for the determination of frequency selectivity have been considered in the literature, e.g. grating acuity, various correlation techniques, determinations of Q and comparisons of slopes of activity patterns. The nonlinearity may have its effect on these methods; in most cases the effect of the nonlinearity on the measures of tuning considered is difficult to understand. We therefore prefer at present to compare the slopes of activity patterns determined at the different stages of auditory processing. By doing this for different levels, we can distinguish between tuning properties and nonlinear effects. It is evident that we have to compare the two sides of the patterns separately for this purpose.

VIII-4. Slopes of the activity patterns

The nonlinearity complicates the comparison of the slopes of activity patterns. It is not sufficient to present a single figure; one should rather present the range of slopes found at different stages.

Such a comparison is shown in Fig.38. It gives minimum and maximum slopes, thus indicating the range of slopes that can be found.

The open symbols represent the psychophysically determined slopes. The squares and solid lines indicate results obtained for the steep edge of the patterns; the circles and broken lines those for the shallow edge. The symbols represent the average slopes of the excitation patterns for the four observers at 20 dB SPL and at 100 dB SPL at the three frequencies. These excitation patterns were computed from the pulsation data of Chapter IV. The individual values show a wide variability among observers, and with frequency for the observers.

The solid symbols represent the slopes of the input patterns for primary fibre responses

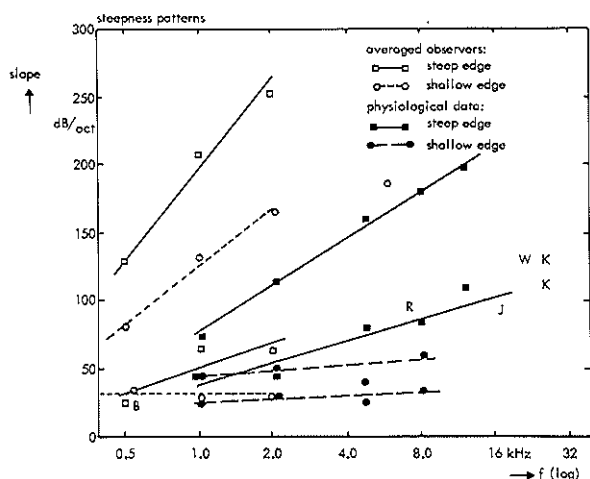


Fig. 38. Range of slopes of patterns reflecting the tuning. The maximum and minimum slopes of the steep edge are given by squares and of the shallow edge by circles. Open symbols were obtained from pulsation threshold measurements by the analysis of Chapter V. The points are the average values for the four observers at 20 dB SPL and 100 dB SPL. The solid symbols are from the analysis of the primary fibre responses presented in Chapter VII and represent individual minimum and maximum values. The letters refer to determinations of the vibration pattern of the basilar membrane. B stands for von Békésy (1960), R for Rhode (1971), W for Wilson & Johnstone (1975), J for Johnstone & Boyle (1967), and K for Kohliöffel (1972 c).

presented in Chapter VII, Fig. 37. The results for the steep side of the vibration patterns of the basilar membrane are indicated by letters referring to the first author of the publication from which they were taken.

Two regions have to be distinguished here. Von Békésy (1960) performed measurements at low frequencies (between 200 and 700 Hz) on the ears of dead animals. He found slopes of about 30 dB/oct. Kohliöffel (1972 c) has shown that the slope depends on the time elapsed after the death of the animal, so one has to be very cautious in interpreting von Békésy's results.

The second region is between 7 and 30 kHz. All the data here were collected from living animals. The R and the W represent average values because the authors found a range of values. Rhode found slopes ranging from 50 to 140 dB/oct with an average of about 100 dB/oct. Wilson & Johnstone also reported variable slopes, ranging from 100 to 150 dB/oct with an average of 132 dB/oct.

The slopes of the other side of the vibration patterns are not given in Fig. 38 because

of the dependence of their value on the method of determination. Wilson & Johnstone (1975) distinguished between constant voltage input, constant sound-pressure input and constant incus displacement. For a constant SPL input they found a slope of about 5 dB/oct up to 400 Hz and further a flat curve up to the cut-off frequency. For a constant incus displacement they found a slope of 6 dB/oct up to 3.5 kHz and a slope of 12 dB/oct up to the cut-off frequency near 20 kHz. These differences make it clear that the method of determination is very important and can cause considerable differences in results, due to the middle ear mechanics.

Kohlöffel used a somewhat different experimental technique. He presented his data as a function of the stimulus input level, which makes the results comparable with those of the constant SPL method. He found slopes of 7 dB/oct and 9 dB/oct just below his cut-off frequency of 28 kHz.

Rhode (1971) found a low-frequency edge consisting of two parts. The part for frequencies below about 3 kHz had a slope of 6 dB/oct; between 3 kHz and the frequency of maximum response (about 8 kHz) the slopes ranged between 18 and 30 dB/oct with an average of 24 dB/oct. He presented his results for constant malleus displacement.

We conclude from this discussion that the low-frequency edge of the vibration pattern of the basilar membrane is rather flat and that its slope depends on the reference level and the frequency ratio, ranging between 0 dB/oct and 30 dB/oct.

VIII-5. Comparison of slopes of activity patterns

Comparison of the results presented in Fig. 38 reveals some interesting points.

The minimum slopes of the steep edges of the excitation patterns, the primary fibre response patterns and the cochlear vibration patterns show a very similar frequency dependence.

The minimum slopes of the shallow edges of the excitation patterns and the primary fibre response patterns are also comparable with a value of about 30 dB/oct.

Comparison of this value with the slope of the vibration pattern is complicated, as we explained above. However, Rhode found slopes up to 30 dB/oct. The discrepancy

between the slopes presented here and the mechanical slopes is thus small, if it exists at all.

We may conclude that both edges of psychophysically and physiologically determined patterns have comparable minimum slopes. This minimum slope is also comparable with that of the vibration pattern at the steep side. The correspondence between the slopes of the psychophysically and physiologically determined patterns and those of the vibration pattern at the shallow side, is less clear, because of the different methods of measurement used; however it cannot be concluded that there is a considerable difference.

There is on the other hand a considerable difference between the psychophysically and physiologically determined maximum slopes at both sides of the patterns. This can be due either to a much wider dynamic range for the psychophysical values, or to a stronger nonlinearity expressed in the coefficient (a) of the linear approximation for the slope-level dependence. We checked this by plotting in Fig. 39 the values of the coefficient (a) against frequency. This figure shows the values for the five primary fibres and the average values for the observers at the three frequencies.

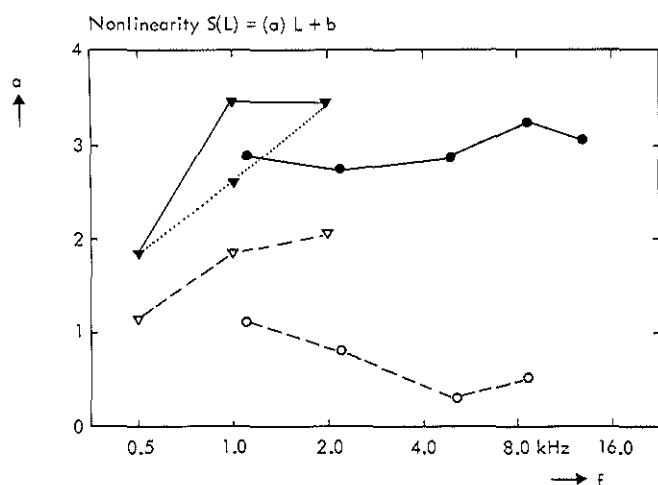


Fig. 39. The value of the coefficient (a) of the slope-level equation $S(L) = aL + b$ for psychophysically determined pulsation patterns (triangles) and the response patterns of primary auditory fibres in cats (circles). The solid symbols represent the values at the steep side of the patterns and the open symbols those at the shallow side.

The nonlinearity at the steep side corresponds to a value of (a) of about 3 for the physiological data and for the psychophysical data at 1 kHz and 2 kHz. The magnitude of the nonlinearity thus appears to be roughly equal in the psychophysical and in the physiological results. A similar analysis for the shallow edge of the patterns shows evidence of a stronger nonlinearity in the psychophysical patterns than in the physiological ones. The wide range of slopes here must be due to a stronger nonlinearity and to a wider dynamic range.

VIII-6. Conclusions to be drawn from the comparison

We have seen in the previous section that the minimum slopes at the different stages of processing do not differ much. This means that the tuning properties at the different stages cannot differ much either. The apparent discrepancies reported in the literature seem to be largely due to differences in the nonlinearity.

This nonlinearity can be described as suppressive. The mechanical vibration pattern determines the total possible extent of the activity patterns. The vibration pattern is processed by a nonlinear detection mechanism, which suppresses the activity at the steep side of the pattern at high stimulation levels, while it has little effect on the other side. At low stimulation levels, however, the nonlinear mechanism has little effect on the steep side of the pattern, but it suppresses the activity at the other side. This kind of behaviour is rather strange, as most nonlinearities are operative only at higher stimulation levels.

VIII-7. Second filter

The differences between mechanical tuning, primary fibre tuning and psychophysical tuning have been discussed in many publications (see the introductory chapter, Context of investigation). This discussion led to the idea that there may be a second filter in the processing chain (Evans & Wilson, 1973).

This line of reasoning assumes that the mechanical cochlear system is the first stage

of filtering. The displacement pattern of the basilar membrane reflects the acuity of the filter at this stage. The displacement is detected by the sensory cells and the displacement signal passes through the "essential" nonlinearity responsible for the generation of combination tones. This nonlinearity is assumed not to influence the tuning properties and cannot thus be the nonlinearity we have been discussing so far. The signal is next assumed to be processed by a second filter, which sharpens the response of the mechanical filter, resulting in the tuning found in the primary fibres. It then passes through a nonlinearity responsible for spike generation. The difference between the tuning found in the primary fibres and in the cochlear mechanics is entirely ascribed to the action of the second filter. The nonlinearities do not influence the tuning properties.

Evans & Wilson used the slope of their frequency threshold curves (our output filter patterns) as a measure of the primary fibre tuning. These slopes can only be determined at rather low stimulation levels because of the limited dynamic range. The discrepancy between these slopes and those of the vibration pattern of the basilar membrane is about 300 dB/oct versus 100 dB/oct at the steep side and about 70 dB/oct versus 20 dB/oct at the shallow side, respectively.

We have shown in the preceding chapters that the steep edge of an output filter pattern gives an exaggerated estimate of the tuning acuity because of the nonlinearity involved. We think that the discrepancy just mentioned is entirely due to the nonlinearity.

At the shallow side, the nonlinearity is particularly evident at low levels. We have seen that the slopes observed at higher levels are lower. We found input patterns with slopes of about 50 dB/oct at the lower stimulus levels. At intermediate levels the slope was reduced to about 30 dB/oct and an extrapolation to higher levels (within the saturation region) will bring this slope in line with the mechanical slopes.

Our data are fully in line with Evans & Wilson's findings. The major difference is the interpretation of the data. They explain a number of phenomena by assuming a linear second filter. We think that there is substantial evidence for a nonlinearity, which is frequency dependent. This nonlinearity can account for phenomena for which the second filter has been assumed, such as the steep slopes of the tuning curves. Some

properties of the second filter, such as its physiological vulnerability, can equally well be ascribed to the nonlinearity. The effect of the vulnerability of the second filter can be considered to be a linearization effect which goes hand in hand with a reduction of the dynamic range, primarily by elevating the threshold of hearing. We have not pursued this line of thought; it requires a verification based on measurement of patterns on patients with cochlear hearing disorders.

VIII-8. Models of the auditory system

Localization of the "slope-level" nonlinearity in the mechano-electric transducer makes it conceivable that our nonlinear behaviour and other nonlinear phenomena like combination tones and suppression are products of some common mechanism. Smoorenburg (1974) showed that the generation of combination tones could be described by an exponential transfer characteristic with an exponent smaller than one. The "slope-level" nonlinearity can be considered as a dynamic characteristic, which can be described by an exponential function. The exponent of this function is less than one for frequencies below the centre frequency. If the dynamic characteristic can be interpreted as a transfer characteristic, the slope-level nonlinearity and the combination-tone nonlinearity may be the same.

Houtgast (1974 a, c) showed clear evidence for suppression in the auditory system. His findings can also be explained to some extent by the slope-level nonlinearity, if we make an additional assumption about the interaction of the excitation patterns of the two sine waves. If we accept the idea that the nonlinearity is based on suppression of activity as described in section VIII-6, the additional assumption would be that the excitation pattern of the second, lower-intensity sine wave would also be suppressed. This suppression would give rise to a dynamic characteristic with a slope of less than one at the steep side, which could be responsible for the generation of combination tones as explained above.

A model based on similar principles has been presented by Hall (1974, 1977 a, b). He used the mechanical nonlinearity found by Rhode (1971) and analysed by Kim et al. (1973) as the basic nonlinearity. A derivative of this nonlinear motion is assumed to be detected. This model shows a "slope-level" nonlinearity, it generates

combination tones and it shows suppression. It is not essential for the model that the nonlinearity is of a mechanical nature. A physiological nonlinearity would also show these effects, provided the excitation pattern of the second sine wave was subject to the same nonlinear effect as the pattern of the more intense sine wave.

An essential property of Hall's model is the detection of a derivative of the vibrations by the detection mechanism. The nature of the input signal of the detector is one of the unsolved questions in hearing. There is no conclusive experimental evidence available about this. Various workers have presented models involving different assumptions regarding the input signal.

Von Békésy (1953) distinguished between radial and longitudinal waves. Khanna et al. (1968) showed that those two vibrations have different amplitude patterns, which could cause interactions (Zwislocki & Sokolich, 1974).

Steele (1973) indicated the possibility that fluid velocity might be detected.

Duifhuis (1976) presented a model in which two different tuning elements are present. This model has a certain nonlinearity, but does not explain our slope-level nonlinearity. It assumes a second filter with only a very moderate tuning characteristic (Duifhuis, 1977)

This brief survey may suffice to illustrate the various conjectures presented as regards the transducer mechanism. We must conclude that we lack decisive data on which a physiological description of the slope-level nonlinearity could be based.

IX. THE CONTINUITY EFFECT IN SPEECH

IX-1. Introduction

The main theme of this thesis is auditory nonlinearity. Our psychophysical method for determining the nonlinearity described is the pulsation threshold method, which is based on the perception of one of two alternately presented signals as continuous. The literature on this continuity effect was reviewed in Chapter I and some pilot studies on the subject were described in Chapter II. In the other chapters the phenomenon was simply accepted and used as a tool for determining the distribution of activity in the auditory system. We think it justified at the end of this thesis to pay some attention to the continuity effect underlying the pulsation method and to present our thoughts on the phenomenon.

The continuity effect was first described for interrupted speech, alternated with noise. Without the noise the interrupted speech sounds very unnatural; the intervening noise makes the speech sound natural and fluent again. This observation is strikingly clear and was also reported by the observers participating in the pilot study described in section II-2. They not only reported the natural sound of this speech, they also had the impression that the speech could be understood better and more easily.

We concluded from the literature on this subject, reviewed in section I-2, that it was quite possible that the intervening noise improves the intelligibility of running speech but not that of monosyllables. This difference can be interpreted as the result

of differences in redundancy.

Improved intelligibility of redundant speech under conditions in which it is heard as continuous has theoretical implications. It suggests that continuity is established at the same central stage of the auditory system, where the differences between redundant and nonredundant speech are also established.

We decided to test the improvement of interrupted speech by intervening noise again for two kinds of speech signals: with higher and lower redundancy. The same experimental procedure will be used in both cases in order to exclude procedural differences, to which differences between results in the literature have often been ascribed.

IX-2. Apparatus and methods

We presented two types of speech to a group of listeners. Both kinds consisted of short sentences with a correct grammatical structure. We call one kind "meaningful speech"; this consisted of normal sentences with a clear meaning, e.g. "All bottles are in the cupboard". We call the second kind of speech "nonsense speech". This consisted of strings of words without a meaning e.g.: "The church bell drank a sailor".

The sentences were recorded on tape on a two track reel recorder. They were spoken by a trained female speaker from our phoniatric department. She was seated in an IAC acoustic room, and was asked to speak at as constant a level as possible.

The average level was adjusted to a -3 dB reading on the VU meter of the recorder, which gives a peak reading of about 0 dB. A 1 kHz calibration tone was recorded at a level of 0 dB on the tape before the recorded speech.

The sentences were rearranged and presented as an alternation of two lists of nonsense speech followed by two lists of meaningful speech and so on. Each list consisted of ten sentences. The grouping in two lists of each kind was realized to permit a direct comparison of speech with and without noise.

The recorder output was adjusted to $1 V_{rms}$ for the calibration tone and the speech was processed by the equipment described in section IV-3 in order to get the interrupted speech. A GR-1382 noise generator set to USASI noise provided the noise signal.

The noise level was adjusted to $1 V_{rms}$. The equipment was set to a switching rate of 4 Hz. The two output signals (speech and noise) were not summed but applied to the input sockets of another tape recorder. The speech was recorded at a level giving a -15 dB level for the continuous calibration tone. The noise was recorded at a level of 0 dB for non-interrupted noise.

This method made it possible to replay the speech in the test set-up as interrupted speech (replay channel 1) or as interrupted speech with intervening noise at a S/N ratio of -15 dB (replay in mode parallel).

The signals were replayed through a Peters AP-6 audiometer with Grason & Stadler TDH-39 headphones. They were presented monaurally. The calibration tone was used to calibrate the audiometer in the manner customarily used in speech audiometry. This calibration procedure results in a 100% intelligibility score for normal PB lists at an intensity level of 40 dB in persons with normal hearing. We presented the signal at 70 dB, which means that the speech is at 55 dB, 15 dB above the level of full intelligibility for monosyllables. The intervening noise makes the speech sound continuous at these levels.

All listeners were persons with normal hearing working in the ENT Department or in the Physics Department. The majority of the listeners were not familiar with the background of the experiments. The presentation of each sentence was followed by a seven-second silent interval in which the listeners were required to repeat the sentence. The errors made by the listeners were noted and the intelligibility score was calculated. In these calculations, articles were neglected and combined words like brand-new, were treated as if the component words were separate.

The listeners were divided into two groups, to which the lists were presented in a different sequence. The total of sixteen lists was presented in two sessions (i.e. eight lists per session) that were never on the same day. The first group of listeners started their first session with nonsense speech without intervening noise followed by nonsense speech with noise. Then they were given two lists of meaningful speech, the first one without noise and the next one with noise. The order of presentation with and without noise was then reversed so the listeners were presented with nonsense speech with noise, nonsense speech without noise, meaningful speech with noise and

meaningful speech without noise. In the second session the lists presented comprised nonsense speech with noise and without noise, meaningful speech with and without noise; again nonsense speech, now first without noise and next with noise and finally meaningful speech without and with noise.

The second group received the lists in the same order except that wherever group 1 had a list with noise they had one without noise, and vice versa.

The presentation of the lists in groups of two with the same kind of speech permits direct comparison between speech with and without noise. The alternation of nonsense speech and meaningful speech spreads the influence of increasing experience (learning effect; see the next section), if it exists, over the two kinds of speech.

IX-3. Results

IX-3-a. Learning effect

After the data for the first eight listeners had been collected, we performed a tentative analysis. This showed a remarkable learning effect for the nonsense speech without intervening noise and a much less marked effect for the nonsense speech with intervening noise. We therefore preceded the formal presentations by two extra recorded lists of nonsense speech, one without noise and one with noise, in order to achieve some familiarizing with the nonsense speech and to reduce initial effects. The learning effect was almost absent for the normal speech.

Data were collected for 18 listeners. We still found a learning effect in the results, even after the introduction of the two extra lists. This effect is illustrated in Fig. 40. The open symbols represent intelligibility scores without noise and the solid symbols with noise. The data are plotted for the lists in the order in which they were presented. We see a steady increase in the nonsense score without noise and a certain increase (though not so clear) in the nonsense score with noise. This effect was not found for the meaningful speech either with or without noise for this observer.

All data were examined for a learning effect. We considered a learning effect to be present when the score for list pair 2 was higher than that for list pair 1, and

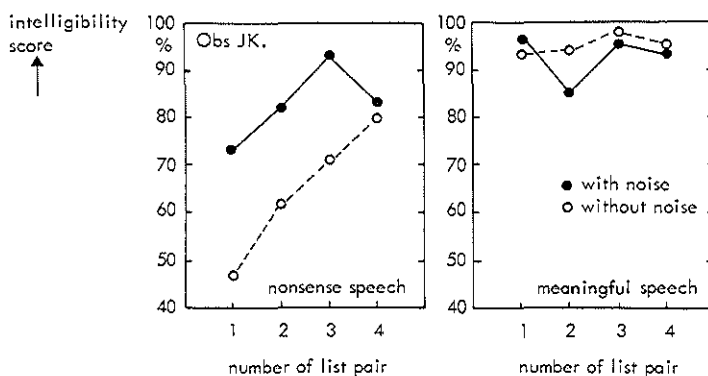


Fig. 40. Intelligibility scores of one observer for consecutive presentations of lists of interrupted connected speech. The left-hand panel shows the scores for nonsense speech and the right-hand panel those for meaningful speech. The interruption rate was 4 Hz and the speech-noise duty cycle 50 %.

that for list pair 4 higher than for pair 3. The analysis showed that the learning effect was present in two-thirds of the listeners for nonsense speech without noise, in half of the listeners for nonsense speech with noise and for one-third of the listeners for meaningful speech without noise. It was not found for meaningful speech with noise. The learning effect for meaningful speech without noise was always very small; for nonsense speech it could correspond to an increase in the score from 40 % up to 80 %.

IX-3-b. General effect of noise on intelligibility

As we mentioned above, the speech was presented in paired lists, one without intervening noise and the other with. This mode of presentation permits a direct comparison of speech with and without intervening noise. Differences between the lists are compensated by presenting the first list of the pair, with noise, to half the listeners and the same list without noise to the other half.

We determined the effect of the noise on the intelligibility for each pair. The results are shown in Table III, where we present in three columns the number of listeners that had an improved, equal or worse intelligibility score. The fourth column shows

the statistical significance of the improvement determined with the test of Wilcoxon. This test does not make use of the paired presentation of the lists; it compares the score with and without noise for all observers and does make use of the fact that we reversed the presentation with and without noise for half the number of listeners. The fifth column shows the conditions in which we found a significant difference ($p < 0.05$) between the lists (Wilcoxon).

The results presented in Table III show that the noise improves the intelligibility of both kinds of speech. The improvement is more significant for nonsense speech than for the meaningful speech, contrary to the results presented in the literature. Our interpretation of the literature was that the noise improves the intelligibility of running (redundant) speech but not that of monosyllables, which have a low level of redundancy.

Table III. The effect of intervening noise on the intelligibility of interrupted speech for 18 observers and 4 pairs of lists.					
Number of listeners with intelligibility score	improved	equal	worse	significance of improvement (Wilcoxon)	condition with difference between lists ($p < 0.05$; Wilcoxon)
nonsense speech					
pair no 1	17	0	1	$p < 0.01$	--
2	14	1	3	$p < 0.01$	--
3	15	1	2	$p < 0.01$	with noise
4	14	0	4	$p < 0.01$	---
meaningful speech					
pair no 1	13	2	3	$p < 0.01$	--
2	12	0	6	$p > 0.05^*$	--
3	14	0	4	$p < 0.05$	with noise
4	13	0	5	$p < 0.05$	--

* not significant

We also analysed the quantitative differences between the score with and without noise. The computed average improvements with noise are smaller than the computed standard deviations. This would have led to the conclusion that the improvement is not statistically significant if a student-t test were applied. The data, however, do not permit the use of such a test, nor of standard deviations because the differences between the scores with and without noise do not have a normal distribution. An analysis of variance of the data is thus not permissible under these circumstances.

The quantitative differences between the scores with and without noise indicate a great variability amongst the listeners. The largest improvements with noise are found for listeners with a low intelligibility score without noise. This suggests a correlation between the difference score and the intelligibility of the interrupted speech without intervening noise. The correlograms of these two scores in fact suggest the presence of such correlation for each pair of lists. The computed regression lines differed only slightly for the same kind of speech. We therefore present the correlograms for all four pairs of lists and all 18 observers, with the equations of the regression line. Fig. 41 is the correlogram for meaningful speech and Fig. 42 that for nonsense speech. There are three broken lines in the figures. The horizontal line represents no improvement of the intelligibility with noise. The vertical line represents the maximum score of 100 % intelligibility and the line of slope 45° the maximum possible improvement of the intelligibility with noise up to a score of 100 %.

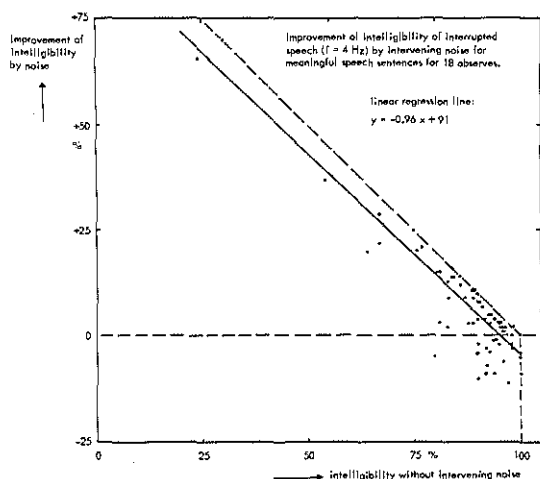


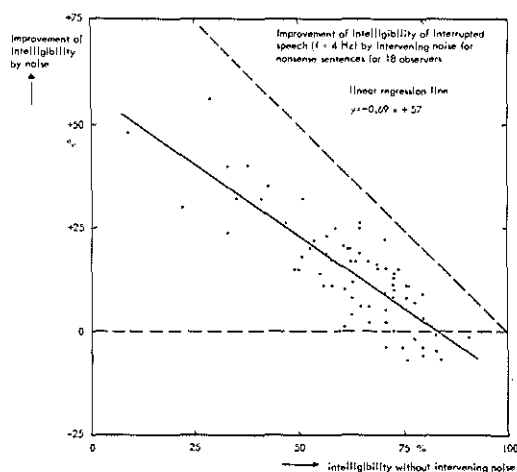
Fig. 41. Correlogram of improvement of intelligibility of meaningful speech by intervening noise as a function of the intelligibility of interrupted speech without noise.

The regression line of Fig. 41 indicates that the intelligibility of interrupted meaningful speech with intervening noise is always close to 95 %, although the intelligibility without noise can be very different for the various listeners.

The regression lines for each separate list are very much alike. Table III shows that the intelligibility of two lists within one pair differ only significantly for pair no. 3 presented with noise. The regression lines for these two lists are parallel. The average scores are 89 % and 98 % with standard errors of 2 % and 1 % respectively ($n = 9$). The differences between the lists do not seem to be important. The regression lines always indicate that intervening noise improves the intelligibility of interrupted speech up to a high score of about 95 %, irrespective of the widely differing intelligibility score of interrupted meaningful speech without intervening noise for the various listeners.

The equation of a regression line is rather sensitive to large departures in a variable. We therefore computed a regression line for the points with a score of more than 70 % for speech without noise. This regression line does not differ much from that of Fig. 41, which implies that the points with the low score do not indicate a different behaviour of these listeners from the others.

Fig. 42. Correlogram of improvement of intelligibility of nonsense speech by intervening noise as a function of the intelligibility of interrupted speech without noise.



The regression line of Fig. 42 for nonsense speech indicates the same general effect as that of Fig. 41 for meaningful speech. The noise improves the score up to a higher value, near 80 %. This maximum value of 80 % is reached if the score without noise

is rather high (near 70 %). If the score without noise is lower, we get an improvement up to a score of slightly less than 80 % with intervening noise. The difference between the regression lines for the various lists is again rather small and only significant for the third pair of lists. The average scores for the two lists are 67 % and 78 % with standard errors of 3 % ($n = 9$). The regression lines are parallel and show the same general effect as described above. The only difference is the point where the regression lines meet the line of no improvement, which we shall refer to as the maximum score.

There are two main differences between the results for nonsense speech and for meaningful speech.

The first is in the value of the maximum scores. These differences are due to the redundancy level of the kind of speech involved.

The second difference is in the slope of the regression line. The observed learning effect could be responsible for the lower slope of the regression line for the nonsense speech. The initial scores with and without noise are both lower for about half our listeners. If the learning effect were involved, however, we would expect a steepening of the regression line for lists presented later; this was not found.

IX-4. Discussion

The findings described show that the learning effect and differences between lists are of minor importance. The general picture for both kinds of speech is of a rather individual intelligibility score with large variations between listeners without noise and a fairly constant score with noise. The value of this constant score depends on the degree of redundancy of the speech material used.

This suggests that the speech signals are processed in different ways with and without noise. With intervening noise a maximum speech intelligibility is achieved, depending only on the properties of the speech (redundancy level). The performance of the listener, reflecting his or her motivation, vocabulary, intelligence and experience, plays only a very small role. Without intervening noise, the performance is of

major importance. It seems as if different processes of speech analysis are involved in the two cases. With noise the speech is perceived as continuous and some speech recognition process fills in the parts of the missing signal which are interpreted as being heard. The redundancy level determines the possibilities, offered to the auditory system in this filling-in process. Without intervening noise, we are aware of the silent gaps and seem to be forced to guess consciously what has been said in them.

Such a difference in processing can also account for the comparatively high scores for meaningful speech without noise. The high degree of redundancy helps the listener to make a reasonable guess in this case. This is not so possible with the nonsense speech, and explains the extremely low scores for some listeners.

It also accounts for the subjective finding that speech with noise is much easier to understand. People even seem to relax when interrupted speech with noise is presented. The unconscious process of filling-in or interpolating a masked period with the most probable signal is a normal process in daily life.

The continuity effect with sine waves can also be understood on this basis. As soon as the head and tail of a sine wave are detected there is an unconscious assumption that this sine wave was present during the entire period in which its absence could not be detected. This view is fully in line with the concepts presented in section 1-4.

This concept of unconscious interpolation differs from the explanation suggested by Huggins (1964), which seems to be accepted by other authors. Huggins noted that the waveform of the speech signal is distorted at the moments of switching. This distortion may result in false articulation cues. The intervening noise masks the distortion, so the interruption does not lead to false cues.

This explanation also leads to a constant maximum score with noise. If no noise is presented the speech is distorted and the intelligibility is lower. However, this view does not account for the large variability we found in the absence of noise, nor does it take into account the reduction of speech information by masking noise. The large variability cannot be caused by differences in speech information. Our speech was taped so the switching conditions were identical for all listeners. If we accept the idea of Huggins, some listeners seem to ignore all false cues while others fall into every pitfall presented by the interrupted speech. This difference in behaviour cannot

be explained by differences in information content nor by differences in peripheral processing but only by some difference in performance; nor does it explain the relaxed attitude of the listeners when the speech is presented with intervening noise as opposed to their tense attitude without noise.

Though we cannot be sure that Huggins' explanation is incorrect, we think that we have presented evidence against it.

Other authors report intelligibility scores of about 80 % at an interruption rate of 4 Hz with intervening noise, which is equal to our score for nonsense speech, or between 90 and 100 %, our score for meaningful speech.

A score of 80 % for PB words is not very surprising, since this type of speech also has a low degree of redundancy. The scores obtained by Schubert & Parker and by Powers & Wilcox showed improvement with noise. We found that the degree of improvement depends largely on the performance of the listeners, so the value obtained depends on the listeners. The low intelligibility score for running speech cannot be explained by our results. It must be due partly to properties of the speech material (e.g. number of words per minute) or of the experimental procedure (e.g. shadowing of the speech). The deterioration of the intelligibility by intervening noise as found by Kreul must be attributed to masking, as shown by Dirks & Bower. The form in which their data are presented does not permit a conclusion about the effect of the noise apart from the masking effect.

Nearly all authors reported a great variability in their results. We found this, too. We observed this variability particularly in the absence of intervening noise and explained this as being due to variability in performance.

We conclude that the different results presented in the literature are not in conflict with our data, nor with our interpretation.

The idea of interpolation presented here and its relation to the continuity effect accounts for a wide variety of observed effects and obtained results. It also accounts for the existence of the continuity effect with sine-wave stimuli, and is not inconsistent with the effects observed with music (see section 11-3).

This idea of interpolation can be included in the hypotheses presented in section 1-4. The continuity effect is understood to be the result of the inability to detect the interruption. The brain has then to find a feasible sound for the situation where the actual signal cannot be perceived. This view can be formulated as follows:

When an acoustic signal is temporarily interrupted and another signal is presented during the (short) interruption period, we perceive the interruption only if the activity level in some auditory channel drops by a significant amount during the interruption. If the interruption is not perceived, the brain interprets the activity not as resulting from a sequence of signals but as resulting from simultaneous presentation of a weaker continuous signal and an interrupted louder signal. The gap appearing in the signal perceived as continuous is filled in with the most probable signal.

This formulation covers not only regular interruptions as in pulsation experiments but also irregular short interruptions which are so often experienced in daily life. An example of such a situation is somebody coughing while we listen to a speaker. In most instances we are hardly aware of the cough; we just hear a continuous stream of words.

The main difference between this formulation and those presented in section 1-4 is our attention for central processes. Our view with reference to the peripheral processes is in line with Houtgast's view, but we think that a formulation of a view on continuity perception must include such central processes.

X. CONCLUSIONS

Presentation conditions of the pulsation threshold method

- The continuity effect on which the pulsation threshold method is based can be perceived at repetition rates between 2 Hz and 10 Hz.
- The switching necessary for the alternate presentation of signals can result in a spectral broadening appearing in the pulsation patterns.
- With Gaussian transition envelopes, spectral broadening may occur at higher repetition rates than 5 Hz and time constants of the Gaussian transition envelope of less than 3.5 ms.
- An exponential amplitude envelope is not suitable for pulsation threshold measurement owing to the long overlap of the two alternating stimuli.
- There is no need for a silent interval between successive presentations for pulsation measurements.

Pulsation patterns and excitation pattern

- Pulsation patterns can be determined by four different methods; we call the patterns determined by these methods the input extension pattern, the output extension pattern, the input filter pattern and the output filter pattern.
 - Pulsation patterns reveal a nonlinearity in the auditory system. As a result of this nonlinearity the patterns determined by the different methods show basic differences.
 - A model describing the experimental data can be based on assumptions regarding the existence and shape of an excitation pattern, including a nonlinear dependence of the slope of this pattern on stimulation level, and a formal hypothesis concerning the conditions under which the continuity effect is perceived.
 - The nonlinearity can account for the shape of the pulsation patterns.
 - The differences between extension and filter patterns are explained as being due to a dependence of the tuning properties and/or the nonlinearity on frequency.
 - The differences between input and output patterns are due to the nonlinearity.
- The nonlinearity in combination with off-frequency detection causes an overestimation

of the slope of the steeper edge of the pulsation patterns and an underestimation of the slope of the shallow edge of the output pulsation pattern. The shallow edge of the input extension pattern is the only one that reflects the tuning properties properly.

- Our model contradicts the assumption that the probe level is a direct measure of the activity elicited by the stimulus at the probe frequency.

Pulsation patterns for noise bands

- The pulsation patterns reflect the spectrum of the pulsator and the critical band mechanism.
- The spread of activity to frequencies outside the spectrum as revealed by the pulsation patterns, is the same for the sine waves and noise bands. An identical nonlinear behaviour is found in both cases.
- The summit of the noise patterns is not flat but peaked. The peak is 6 dB high.
- The parts of the masking patterns determined by combination tones resemble the pulsation patterns. This fact suggests a common mechanism underlying these two nonlinear effects.

Nonlinearity of physiologically determined activity patterns

- A nonlinearity similar to that found in psychophysical experiments is found in the responses of the primary auditory fibres of cats.
- This nonlinearity causes a steepening of the steep edge of the fibre's response pattern with a rise in level. At very low levels, the slope of this edge is comparable with that of the motion pattern of the basilar membrane. The magnitude of the physiologically determined nonlinearity is comparable with that found in psychophysical experiments.
- The nonlinearity results in a steepening of the shallow edge of the fibre's response pattern with a drop in level. At high levels the slope of this edge approaches the measured slope of the shallow edge of the vibration pattern of the basilar membrane. The magnitude of the physiologically determined nonlinearity is lower than that found by psychophysical methods.
- The correspondence between the mechanical and physiological data at some stimulation level, either high or low, casts doubt on the existence of the second filter

postulated by some authors. It seems more likely that the discrepancy noted in the tuning properties can be explained as being due to a nonlinearity.

The continuity effect in speech

- The intelligibility of speech under pulsation conditions depends on the redundancy level of the speech.
- There is an interaction between the continuity effect and the intelligibility of the interrupted speech.
- These findings have led to the following formulation of a hypothesis for perception of the continuity effect:

When an acoustic signal is temporarily interrupted and another signal is present during the (short) interruption period, we perceive the interruption only if the activity level in some auditory channel drops by a significant amount during the interruption. If the interruption is not perceived, the brain interprets the activity not as resulting from a sequence of signals but as resulting from simultaneous presentation of a weaker continuous signal and an interrupted louder signal. The gap appearing in the signal perceived as continuous is filled in with the most probable signal.

APPENDIX A

The spectrum of the exponential amplitude envelope

The time function is given in Fig. 1:

$$t \leq -t_0 : f(t) = 0$$

$$-t_0 \leq t \leq t_0 : f(t) = A_1 \left\{ 1 - \exp \left[-\left(\frac{t + t_0}{\tau} \right) \right] \right\} \cos \omega_0 t$$

$$t > t_0 : f(t) = A_1 \exp \left[-\left(\frac{t - t_0}{\tau} \right) \right] \left[1 - \exp \left(-2 \frac{t_0}{\tau} \right) \right] \cos \omega_0 t$$

The Fourier transform is:

$$F(\omega) = \frac{A_1}{\sqrt{2\pi}} \left[\int_{-t_0}^{t_0} \left\{ 1 - \exp \left[-\left(\frac{t + t_0}{\tau} \right) \right] \right\} \cos \omega_0 t e^{-i\omega t} dt + \right. \\ \left. + \int_{t_0}^{\infty} \exp \left[-\left(\frac{t - t_0}{\tau} \right) \right] \times \left[1 - \exp \left(-2 \frac{t_0}{\tau} \right) \right] \cos \omega_0 t e^{-i\omega t} dt \right]$$

Now we define $\Delta\omega_- = \omega - \omega_0$ and $\Delta\omega_+ = \omega + \omega_0$. We see that:

$$F(\Delta\omega_+, \Delta\omega_-) = F_0(\Delta\omega_+) + F_0(\Delta\omega_-)$$

It follows that:

$$F_0(\Delta\omega) = \frac{A_1}{\sqrt{2\pi}} \left[\int_{-t_0}^{t_0} \exp(-i\Delta\omega t) dt - \int_{-t_0}^{t_0} \exp \left[-\left(\frac{t + t_0}{\tau} \right) \right] \exp(-i\Delta\omega t) dt - \right. \\ \left. - \int_{t_0}^{\infty} \exp \left[-\left(\frac{t - t_0}{\tau} \right) \right] \exp(-i\Delta\omega t) dt + \int_{t_0}^{\infty} \exp \left[-\left(\frac{t - t_0}{\tau} \right) \right] \exp(-i\Delta\omega t) dt \right]$$

It is a straightforward matter to derive:

$$F_0(\Delta\omega) = \frac{A_1}{\sqrt{2\pi}} \left[\frac{\sin \Delta\omega t_0}{\Delta\omega} - \frac{\sin \Delta\omega t_0}{\Delta\omega - i/\tau} \right] = \frac{A_1 t_0}{\sqrt{2\pi}} \frac{\sin \Delta\omega t_0}{\Delta\omega t_0} \frac{i}{i - \tau \Delta\omega}$$

The power function is:

$$W_0(\Delta\omega) = F_0(\Delta\omega)F_0^*(\Delta\omega) = \frac{A_1^2 t_0^2}{2\pi} \frac{\sin^2 \Delta\omega t_0}{\Delta\omega^2 t_0^2} \frac{1}{1 + \Delta\omega^2 \tau^2}$$

First we calculate the top of the spectrum:

$$\lim_{\Delta\omega \rightarrow 0} W(\Delta\omega) = \frac{A_1^2 t_0^2}{2\pi} = W(0)$$

So the envelope of the spectrum can be described by:

$$W(\Delta\omega) = \frac{W(0)}{\Delta\omega^2 t_0^2} \frac{1}{1 + \Delta\omega^2 \tau^2}$$

We computed the spectra of Figs. 2, 4 and 7 according to this equation.

APPENDIX B

The spectrum of the gaussian transition function

The time function is (Fig. 1):

$$t \leq -t_0 : f(t) = A_1 \exp \left[-\frac{1}{2} \left(\frac{t+t_0}{\tau} \right)^2 \right] \cos \omega_0 t$$

$$-t_0 \leq t \leq t_0 : f(t) = A_1 \cos \omega_0 t$$

$$t \geq t_0 : f(t) = A_1 \exp \left[-\frac{1}{2} \left(\frac{t-t_0}{\tau} \right)^2 \right] \cos \omega_0 t$$

The Fourier transform is:

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt$$

Again we use $\Delta\omega \pm = \omega \pm \omega_0$ and

$$F(\Delta\omega+, \Delta\omega-) = F_0(\Delta\omega+) + F_0(\Delta\omega-)$$

This leads to:

$$F_0(\Delta\omega) = \frac{A_1}{2\sqrt{2\pi}} \left\{ \int_{-\infty}^{-t_0} \exp \left[-\frac{1}{2} \left(\frac{t+t_0}{\tau} \right)^2 - i\Delta\omega t \right] dt + \int_{-t_0}^0 e^{-i\Delta\omega t} dt + \right. \\ \left. + \int_0^{\infty} \exp \left[-\frac{1}{2} \left(\frac{t-t_0}{\tau} \right)^2 - i\Delta\omega t \right] dt \right\}$$

Next we perform the time translations $u = t + t_0$ and $u' = t - t_0$. We get:

$$F_0(\Delta\omega) = \frac{A_1}{2\sqrt{2\pi}} \left\{ e^{i\Delta\omega t_0} \int_0^{\infty} \exp \left[-\frac{1}{2} \left(\frac{u}{\tau} \right)^2 + i\Delta\omega u \right] du + \frac{e^{-i\Delta\omega t}}{-i\Delta\omega} \Big|_{-t_0}^{t_0} + \right. \\ \left. + e^{-i\Delta\omega t_0} \int_0^{\infty} \exp \left[-\frac{1}{2} \left(\frac{u}{\tau} \right)^2 - i\Delta\omega u \right] du \right\}$$

Completing the squares in the exponents, we obtain:

$$F_0(\Delta\omega) = \frac{A_1}{2\sqrt{2\pi}} \exp\left(i\Delta\omega t_0 - \frac{\tau^2}{2} \Delta\omega^2\right) \int_0^\infty \exp\left[-\frac{1}{2\tau^2}(u - i\Delta\omega\tau^2)^2\right] du + \\ + 2 \frac{\sin \Delta\omega t_0}{\Delta\omega} + \exp\left(-i\Delta\omega t_0 - \frac{\tau^2}{2} \Delta\omega^2\right) \int_0^\infty \exp\left[-\frac{1}{2\tau^2}(u + i\Delta\omega\tau^2)^2\right] du$$

The integral we got is of the form:

$$\int_0^\infty \exp\left[-a(u + i\beta)^2\right] du \quad \text{with } a, \beta \text{ real constants.}$$

This is an integration in the complex plane over a line parallel to the real axis.

According to Cauchy's theorem:

$$\int_0^\infty \exp\left[-a(u + i\beta)^2\right] du = \int_0^\infty \exp(-au^2) du + \int_{-i\beta}^0 \exp(-az^2) dz = \\ = \frac{1}{2} \sqrt{\frac{\pi}{a}} + i \int_0^\beta \exp(-au^2) du \quad \text{with } u, \beta \text{ real}$$

This result is substituted. We find:

$$F_0(\Delta\omega) = \frac{A_1}{2\sqrt{2\pi}} \left(\frac{2 \sin \Delta\omega t_0}{\Delta\omega} + \exp\left(-\frac{\tau^2}{2} \Delta\omega^2\right) \tau \sqrt{2\pi} \cos \Delta\omega t_0 - \right. \\ \left. - 2 \sin \Delta\omega t_0 \exp\left(-\frac{\tau^2}{2} \Delta\omega^2\right) \int_0^{\Delta\omega\tau^2} \exp\left(-\frac{u^2}{2\tau^2}\right) du \right)$$

The substitution of $t = u/\tau\sqrt{2}$ in the integral yields:

$$F_0(\Delta\omega) = \frac{A_1\tau}{2} \left[\exp\left(-\frac{\Delta\omega^2\tau^2}{2}\right) \cos \Delta\omega t_0 + \right. \\ \left. + \frac{\sin \Delta\omega t_0}{\sqrt{\pi}} \left(\frac{\sqrt{2}}{\Delta\omega\tau} - 2 \exp\left(-\frac{\Delta\omega^2\tau^2}{2}\right) \int_0^{\Delta\omega\tau/\sqrt{2}} \exp(-t^2) dt \right) \right]$$

The integral thus found has the general form:

$$D(x) = \exp(-x^2) \int_0^x \exp(-t^2) dt$$

and is called the Dawson's integral (Abramowitz and Segun, 1968).

The most important properties of the Dawson's integral are:

1) For $x \rightarrow 0$: $D(x) \rightarrow 0$.

Whence:

$$F_0(0) = \frac{A_1 t_0}{\sqrt{2\pi}} \left[1 + \frac{\tau}{2t_0} \sqrt{2\pi} \right]$$

2) For $x \rightarrow \infty$ (asymptotic behaviour):

$$D(x) \sim \frac{1}{2x}$$

So if $(\Delta\omega\tau/\sqrt{2}) \rightarrow \infty$ (large τ or large $\Delta\omega$) the spectrum is:

$$F_0(\Delta\omega) \sim \frac{A_1 \tau}{2} \exp\left(-\frac{\Delta\omega^2 \tau^2}{2}\right) \cos \Delta\omega t_0$$

We use, however, small τ and small $\Delta\omega$, so we find:

$$F_0(\Delta\omega) = \frac{A_1 t_0}{\sqrt{2\pi}} \frac{\sin \Delta\omega t_0}{\Delta\omega t_0} \times \left[1 - 2 \frac{\Delta\omega \tau}{\sqrt{2}} D\left(\frac{\Delta\omega \tau}{\sqrt{2}}\right) \right] + \\ + \frac{A_1 \tau}{2} \exp\left(-\frac{\Delta\omega^2 \tau^2}{2}\right) \cos \Delta\omega t_0$$

We computed the envelope of the power spectrum by comparing:

$$\frac{\tau}{2} \exp\left(-\frac{\Delta\omega^2 \tau^2}{2}\right)$$

and

$$\frac{1}{\sqrt{2\pi}} \frac{1}{\Delta\omega} \left[1 - 2 \frac{\Delta\omega \tau}{\sqrt{2}} D\left(\frac{\Delta\omega \tau}{\sqrt{2}}\right) \right]$$

and using the larger number in the computation. In the Figs. 2, 5, 6 and 7 the spectra are shown. The "shoulder" in the spectrum is that part of the spectrum, where the gaussian spectrum ($\cos \omega t_0$) is larger.

A comparison of the computed spectra and spectra measured at the output of our shapers, shows that these spectra are identical within 1 dB down to about -60 dB re top.

APPENDIX C

The spectrum of the sinusoidal transition function

The time function is:

$$t < -(t_o + \tau_o) \quad : \quad f(t) = 0$$

$$-(t_o + \tau_o) < t < -t_o : f(t) = \frac{1}{2} A_1 \left(1 + \cos \left(\pi \left(\frac{t + t_o}{\tau_o} \right) \right) \right) \cos \omega_o t$$

$$-t_o < t < t_o \quad : \quad f(t) = A_1 \cos \omega_o t$$

$$t_o < t < t_o + \tau_o \quad : \quad f(t) = \frac{1}{2} A_1 \left(1 + \cos \left(\pi \left(\frac{t - t_o}{\tau_o} \right) \right) \right) \cos \omega_o t$$

$$t > t_o + \tau_o \quad : \quad f(t) = 0$$

Once again we use $\Delta\omega_{\pm} = \omega \pm \omega_o$ and

$$F(\omega) = F'(\Delta\omega_+, \Delta\omega_-) = F_o(\Delta\omega_+) + F(\Delta\omega_-)$$

It is a straightforward matter to derive:

$$F_o(\Delta\omega) = -\frac{A_1}{2\sqrt{2}\pi} \left[\sin(\Delta\omega_o t) + \sin\left(\Delta\omega(t_o + \tau_o)\right) \right] \cdot \left(\frac{\pi^2/\tau_o^2}{\Delta\omega(\Delta\omega^2 - \pi^2/\tau_o^2)} \right)$$

Substituting $t_1 = t_o + \frac{1}{2}\tau_o$ and rearranging the factor with the sinusoids yields:

$$F_o(\Delta\omega) = \frac{A_1 t_1}{\sqrt{2\pi}} \cdot \frac{\sin(\Delta\omega t_1)}{\Delta\omega t_1} \cdot \cos\left(\frac{1}{2}\Delta\omega\tau_o\right) \cdot \frac{1}{1 - \frac{\Delta\omega^2 \tau_o^2}{2\pi}}$$

This function has a singular point at $\Delta\omega = 0$.

We get, however:

$$\lim_{\Delta\omega \rightarrow 0} F_o(\Delta\omega) = \frac{A_1 t_1}{\sqrt{2\pi}}$$

The spectrum is shown in Fig. 2 as a dotted line.

SUMMARY

The main theme of this study is the auditory frequency selectivity and its nonlinearity. These were studied by means of the pulsation threshold method for which we determined the existence and presentation conditions.

The introductory chapter ("Context of investigation") deals with the various psychophysical and physiological methods which have been used for determination of the frequency selectivity, and with the theory of the peripheral processing of acoustic signals used to explain the results obtained with these methods. Some advantages and drawbacks of the various psychophysical methods are presented.

Chapter I surveys the literature on the pulsation threshold method and the underlying continuity effect.

In the next chapter, some pilot studies on the continuity effect are presented. These studies show that the frequency-selective mechanism of the auditory system is involved in the processes that lead to the continuity effect; central processes also seem to be involved.

The effects of the presentation conditions of sine-wave stimuli on the results of the pulsation threshold method are studied in Chapter III. It is shown that fast switching of the signals causes a spectral broadening which may show up in the patterns measured by the pulsation threshold method (pulsation patterns). Slow switching on the other hand may lead to overlap of the stimuli, resulting in the perception of combination tones which may affect the patterns too. A repetition rate of 4 Hz, a Gaussian transition envelope with a time constant of 3.6 ms and presentation without silent intervals between the stimuli seem to represent a good compromise. We will be using these presentation conditions in our further experiments.

A classification of the various possible ways of determining pulsation patterns is given in Chapter IV. We see that there are four possible methods, giving pulsation

patterns for which we introduce a nomenclature. In a linear system the relations between the four methods are simple, but the results presented indicate a nonlinearity in the system which obscures the relations between the various types of patterns.

A model for analysis of the patterns is described in Chapter V. The assumptions on which this model is based are formulated. They concern the existence, the shape and the level dependence of an excitation pattern. The criterion for the perception as continuous of one stimulus of a pair presented alternately, is assumed to be based on comparison of the excitation patterns of the two stimuli. The level dependence of the excitation pattern is approximated to with a first-order equation for the dependence of the slope of the pattern on level; this introduces a nonlinearity into the model.

The model explains the different shapes of the various types of patterns. The patterns determined for four observers at three frequencies are analysed with reference to the model, yielding numerical results for the slopes and the nonlinearity of the excitation patterns. The results show that the slope of the pulsation patterns is not generally a good measure of the frequency selectivity of the auditory system. This is particularly so for the steep edge of the pulsation patterns, whose slope is considerably higher than that of the excitation patterns. We assume that the excitation pattern reflects the real frequency selectivity.

Pulsation patterns for narrow noise bands are presented in Chapter VI. The patterns are analysed with the aid of the model of Chapter V. The analysis shows that there is no evidence for a difference in frequency selectivity for sine waves and noise bands, provided we taken into account a difference in pulsation criterion for the two types of pulsators and a difference of 6 dB between the pulsation levels at the central frequency and at the cut-off points of the pattern. The small difference in shape between the pulsation patterns for sine waves and noise bands are due to the nonlinearity.

We also compared pulsation patterns with direct masking patterns. These two types of patterns coincide in the region where the detection threshold of the maskee is determined by combination tones; the patterns differ considerably outside this region, although both patterns show a similar kind of nonlinear behaviour. The slopes of the excitation patterns computed from our pulsation patterns differ considerably

from those of the masking patterns. The results obtained with these two methods cannot thus be considered to reflect the same tuning or to be established at the same stage of processing in the auditory system.

In Chapter VII we study the nonlinearity of responses recorded from primary auditory fibres. These responses show a nonlinearity of the same type as we found in our psychophysical experiments. We conclude that the "response areas" reflect the tuning properties and the nonlinearity in a way which makes it possible to separate them; "tuning curves" in general also reflect the auditory tuning and the nonlinearity, but in such a complicated way that it is not possible to separate them.

The consequences of our findings for the theory of hearing are discussed in Chapter VIII. A comparison of the slopes of the vibration pattern of the basilar membrane with those of the response patterns of the primary fibres and of the excitation patterns computed from pulsation patterns shows that the edges of the excitation patterns and of the response patterns of primary fibres have about the same minimum slopes as the vibration patterns of the basilar membrane. The nonlinearity of the excitation pattern and the response pattern of primary fibres is comparable for the steep side of the patterns but not for the shallow side.

Our findings cannot be understood in terms of the concept of a linear second filter; they suggest the involvement of a frequency-dependent nonlinearity in the auditory system.

The continuity effect underlying the pulsation threshold method, is studied in Chapter IX for speech. We used the same presentation conditions for speech as for sine waves ($f_{\text{rep}} = 4$ Hz; duty cycle 50 %; Gaussian transient envelope with $\tau \approx 3.6$ ms; no gap), and compared the intelligibility of nonsense speech and meaningful speech with periodic interruptions, with and without noise presented during the interruptions. The noise was presented at a level 15 dB above the level of the speech; this leads to perception of the speech as if it were continuous. The noise improves the intelligibility of the speech, particularly of nonsense speech, in most cases. The intelligibility of speech with intervening noise is almost equal for all listeners and is determined by the degree of redundancy of the speech (the intelligibility score for nonsense speech is 80 %, for meaningful speech 95 %). The intelligibility score for speech without intervening noise differs considerably from listener to listener. The noise

improves the intelligibility of interrupted speech appreciably for listeners who have a low score without intervening noise, but has little or no effect on the intelligibility for listeners who have a high score without intervening noise. This variability in the results can account for the conflicting results reported in the literature.

Chapter X contains a list of the most important findings and conclusions from all chapters.

SAMENVATTING

Het beschreven onderzoek heeft als hoofdtema de auditieve frekwentie selektiviteit en de alineariteit daarvan. We bestuderen deze m.b.v. de pulsatiedrempelmetode, waarvan we bestaans- en presentatiekondities bepalen.

In het inleidende hoofdstuk ("Context of investigation") worden de verschillende psychofysische en fysiologische methoden behandeld die gebruikt worden voor het bepalen van de frekwentie selektiviteit en wordt de theorie van de perifere verwerking van signalen beschreven die ontwikkeld is om de eksperimentele resultaten te verklaren. Enkele voor- en nadelen van de psychofysische methoden worden besproken.

In hoofdstuk I wordt een beschrijving van de pulsatiedrempelmetode gegeven en een overzicht van de literatuur over deze methode en over het continuïteitsfenomeen waarop deze methode berust.

In het volgende hoofdstuk worden enkele verkennende onderzoekjes beschreven naar dit continuïteitsfenomeen waaruit blijkt dat het frekwentie selektieve mechanisme van het auditieve systeem een rol speelt bij het tot stand komen van de waarneming van continuïteit. Tevens blijkt hieruit dat centrale processen daarbij meespelen.

De effecten van de aanbiedingskondities van sinusvormige signalen op de resultaten verkregen met de pulsatiedrempelmetode worden beschreven in hoofdstuk III. Hieruit blijkt dat de met deze methode bepaalde patronen (pulsatiepatronen) enerzijds beïnvloed kunnen worden door spektrale effecten ten gevolge van het te snel schakelen van de signalen; anderzijds kan een overlap in de aanbiedingskondities ontstaan bij te langzaam schakelen waardoor de patronen beïnvloed worden door combinatie-tonen. Als compromis kozen we een repetitiefrekwentie van 4 Hz, een Gaussische transient omhullende met een tijdsconstante van 3,6 ms en een presentatie zonder stille periodes tussen de afwisselend aangeboden stimuli.

Een indeling van de verschillende mogelijkheden om pulsatiepatronen te bepalen wordt gegeven in hoofdstuk IV. Er blijken vier methoden te zijn waarvoor we een

benaming invoeren. De verbanden tussen de verschillende patronen zijn eenvoudig in een lineair systeem. Uit de gemeten pulsatiepatronen, verkregen volgens de verschillende methoden, blijkt echter dat het systeem alineaair is zodat de onderlinge verbanden niet meer eenvoudig te leggen zijn.

Om deze verbanden te kunnen leggen, presenteren we in hoofdstuk V een model voor de analyse van deze patronen. De aannames die ten grondslag liggen aan deze modelbeschrijving worden geformuleerd. Zij betreffen het bestaan, de vorm en de nivo-afhankelijkheid van een eksitatiepatroon. Er wordt verder aangenomen dat de continuïteitswaarneming berust op een vergelijking van de eksitatiepatronen van de twee alternerend aangeboden stimuli. De nivo-afhankelijkheid van het eksitatiepatroon wordt benaderd met een eerstegraads vergelijking die het verband beschrijft tussen de helling van het eksitatiepatroon en het nivo van de stimulus. Hiermee is een alineaairiteit geïntroduceerd in het model. Dit model blijkt de vorm van de verschillende pulsatierempelpatronen goed te verklaren. De gemeten pulsatiepatronen worden met dit model geanalyseerd zodat we de alineaairiteit en de steilheid van de aan deze pulsatiepatronen ten grondslag liggende eksitatiepatronen numeriek kunnen bepalen. Hieruit blijkt dat de hellingen van de pulsatiepatronen in de meeste gevallen geen goede schatting geven van de frekwentie selektiviteit. Met name de steile helling van het pulsatiepatroon is veel steiler dan de overeenkomstige helling van het eksitatiepatroon, waarvan we aannemen dat het een juiste afspiegeling is van het auditieve filter.

In hoofdstuk VI worden de pulsatiepatronen van smalle ruisbanden beschreven. De gemeten patronen worden geanalyseerd met behulp van het model, zoals beschreven in hoofdstuk V. Het blijkt dat deze patronen geen aanleiding geven een andere frekwentie selektiviteit te veronderstellen dan die bij de sinusvormige signalen, mits we rekening houden met een gebleken criteriumverschil en met een verschil van 6 dB tussen de pulsatienivo's bij de centrale frekwentie van de ruisband en bij het begin van de hellingen van het patroon van de ruisband. De geringe verschillen die worden gevonden tussen de vormen van de pulsatiepatronen voor sinusvormige signalen en voor smalle ruisbanden zijn het gevolg van de alineaairiteit.

Vervolgens vergelijken we maskeerpatronen en pulsatiepatronen. Hieruit blijkt dat deze beide patronen samenvallen in het gebied waar het maskeerpatroon bepaald

wordt door de detektie van combinatie-tonen. Buiten dit gebied zijn er grote verschillen tussen de beide patronen. Wel laten beide patronen een zelfde soort alineariteit zien. Er bestaan grote verschillen tussen de steilheden van de uit de pulsatiedrempel bepaalde eksitatiepatronen en de maskeerpatronen. De resultaten verkregen met deze twee psychofysische methoden kunnen dus niet direkt vergeleken worden en kunnen dus ook niet beschouwd worden een direkt verband met elkaar te hebben.

In hoofdstuk VII gaan we na in hoeverre de responsies van de primaire neuronen alineaair zijn. Het blijkt dat deze responsies een zelfde soort alineariteit vertonen als we gevonden hebben met onze psychofysische experimenten. Hieruit volgt dat de "response areas" de frekwentie selektiviteit en de alineariteit laten zien op een manier waarin deze te scheiden zijn; de "tuning curves" laten deze twee aspecten ook zien, maar nu op een gekompliceerde manier zodat de effecten van de alineariteit niet te scheiden zijn van die van het auditieve filter.

In hoofdstuk VIII gaan we de gevolgen van onze bevindingen voor de theorie van het horen na. Uit een vergelijking van de steilheid van de bewegingspatronen van het basilaire membraan, van de responsie patronen van de primaire neuronen en van de eksitatiepatronen blijkt dat de minimale hellingen van de eksitatie patronen en van de primaire neuron-patronen vergelijkbaar zijn met die van de bewegingspatronen van het basilaire membraan. De alineariteit van eksitatie- en primaire neuron-patronen aan de steile kant van de patronen is vergelijkbaar; dit geldt niet aan de andere zijde van de patronen.

Onze bevindingen zijn moeilijk te beschrijven met een lineair tweede filter; ze geven aanleiding tot het veronderstellen van een frekwentie-afhankelijke alineariteit in het auditieve systeem.

In hoofdstuk IX gaan we in op het continuïteitsfenomeen dat ten grondslag ligt aan de pulsatiedrempel methode, met name voor spraak. De spraak wordt aangeboden volgens dezelfde presentatiekondities als we gebruikt hebben in de pulsatiedrempel experimenten ($f_{rep} = 4 \text{ Hz}$; duty cycle 50 %; Gauss-vormige transient omhullende met $\tau = 3,6 \text{ ms}$). We vergelijken de spraakverstaanbaarheid van onderbroken zinnige en nonsens spraak zonder en met ruis in de gaten. De ruis wordt zo hard gemaakt (15 dB boven het spraaknivo) dat de spraak met ruis kontinu klinkt. We vinden dat

de ruis de spraak meestal beter verstaanbaar maakt. De spraak met ruis is voor alle luisteraars vrijwel even goed verstaanbaar; de verstaanbaarheids-score wordt voornamelijk bepaald door de redundantiegraad van de spraak (verstaanbaarheid van nonsens spraak 80 %, van zinnige spraak 95 %). In de konditie zonder ruis treden erg grote verschillen in spraakverstaan op. Luisteraars die zonder ruis de onderbroken spraak slecht verstaan, tonen een sterk verbeterd spraakverstaan met ruis; luisteraars die de onderbroken spraak zonder ruis al goed verstaan, tonen weinig of geen verbetering. Op grond van deze grote variabiliteit in gedrag zijn de grote verschillen in resultaten uit de literatuur begrijpelijk.

In hoofdstuk X geven we een opsomming van de belangrijkste resultaten en konklusies uit alle hoofdstukken.

REFERENCES

- ABRAMOWITZ M., SEGUN I.A., (1968), Handbook of mathematical functions, Dover Publications Inc., New York, p. 319.
- ANDERSON D.J., ROSE J.E., HIND E.H., BRUGGE J.F., (1971), Temporal position of discharges in single auditory nerve fibres within the cycle of a sine-wave stimulus : frequency and intensity effects, *J. Acoust. Soc. Amer.* 49 : 1131-1139.
- BEKESY VON G., (1953), Shearing microphonics produced by vibrations near the inner and outer hair cells, *J. Acoust. Soc. Amer.* 25 : 786-790.
- BEKESY VON G., (1958), Funneling in the nervous system, *J. Acoust. Soc. Amer.* 30 : 399-412.
- BEKESY VON G., (1959), Similarities between hearing and skin sensation, *Psychol. Rev.* 66 : 1-22.
- BEKESY VON G., (1960), Experiments in hearing, McGraw-Hill Book Company, Inc., New York.
- BERGEIJK VAN W., (1962), Variation on a theme of Békésy : a model of binaural interaction, *J. Acoust. Soc. Amer.* 34 : 1431-1437.
- BOER DE E., (1967), Correlation studies applied to the frequency resolution of the cochlea, *J. Aud. Research* 7 : 209-217.
- BOER DE E., (1969), Reversed correlation II. Initiation of nerve impulses in the inner ear, *Proc. Kon. Ned. Akad. Wetenschappen*, 72, Serie C : 129-151.
- BOER DE E., JONGKEES L.B.W., (1968), On cochlear sharpening and cross-correlation methods, *Acta oto-laryngologica* 65 : 97-104.
- BRINK VAN DEN G., (1964) Detection of tone pulses of various durations in noise of various bandwidths, *J. Acoust. Soc. Amer.* 36 : 1206-1211.
- CHISTOVICH L.A., (1957), Frequency characteristics of the masking effect, *Biophysics* 2 : 714-725.
- CHISTOVICH L.A., (1971), Auditory processing of speech stimuli-evidences from psychoacoustics and neurophysiology, *Proc 7th ICA.*, vol. 1 : 27-41 (21G1) Akademiai Kiado, Budapest.
- CONINX F., (1977), On the perception of changes in amplitude and frequency of acoustic signals, Thesis Utrecht State University, Dept of Physics.
- DIRKS D.D., BOWER D., (1970), Effect of forward and backward masking on speech intelligibility, *J. Acoust. Soc. Amer.* 47 : 1003-1008.
- DUIFHUIS H., (1976), Cochlear nonlinearity and second filter : possible mechanism and implication, *J. Acoust. Soc. Amer.* 59 : 408-423.
- DUIFHUIS H., (1977), Cochlear nonlinearity and second filter : a psychophysical evaluation, in "Psychophysics and physiology in hearing", E.F. Evans & J.P. Wilson Eds, Academic Press, London, 153-163.
- EGAN J.P., HAKE W., (1950) On the masking of a simple auditory stimulus, *J. Acoust. Soc. Amer.* 22 : 622-630.
- ELFNER L.F., (1969), Continuity in alternately sounded tone and noise signals in a free field, *J. Acoust. Soc. Amer.* 46 : 914-917.
- ELFNER L.F., (1971), Continuity in alternately sounded tonal signals in a free field, *J. Acoust. Soc. Amer.* 49 : 447-449.
- ELFNER L.F., CASKEY W.E., (1965), Continuity effects with alternately sounded

- noise and tone signals as a function of manner of presentation,
J. Acoust. Soc. Amer. 38 : 543-547.
- ELFNER L.F., HOMICK J.L., (1966), Some factors affecting the perception of continuity in alternately sounded tone and noise signals, J. Acoust. Soc. Amer. 40 : 27-31.
- ELFNER L.F., HOMICK J.L. (1967), Continuity effects with alternately sounding tones under dichotic presentation, Perception and Psychophysics 2 : 34-36.
- ELFNER L.F., MARSELLA A., (1966), Continuity effects with alternately sounded tone and noise signals, Med. Res. Eng. 5 : 22-23.
- EVANS E.F., (1972), The frequency response and other properties of single fibres in the guinea pig cochlea, J. Physiol. (London), 226 : 263-287.
- EVANS E.F., (1974), Auditory frequency selectivity and the cochlear nerve, in: "Facts and Models in Hearing", E. Zwicker & E. Terhardt Eds., Springer Verlag, Berlin.
- EVANS E.F., (1977), Frequency selectivity at high signal levels of single units in cochlear nerve and nucleus, in "Psychophysics and Physiology of hearing", E.F. Evans & J.P. Wilson Eds, Academic Press, London, p 185-192.
- EVANS E.F., PALMER A.R., (1975), Responses of units in the cochlear nerve and nucleus of the cat to signals in the presence of the bandstop noise, J. Phys. (London) 252 : 60-62 P.
- EVANS E.F., WILSON J.P., (1971), Frequency sharpening of the cochlea: the effective bandwidth of cochlear nerve fibres, Proc. 7th ICA, Budapest 3 : 453-456 (23H4) Akademiai Kiado Budapest.
- EVANS E.F., WILSON J.P., (1973), Frequency selectivity of the cochlea, in: "Basic mechanisms in hearing", A.R. Møller Ed., Academic, New York : 519-55
- FASTL H., (1974 a), Transient masking pattern of narrow band maskers, in: "Facts and models in hearing", E. Zwicker & E. Terhardt Eds., Springer Verlag, Berlin : 251-257.
- FASTL H., (1974 b), Comment on : "The slopes of masking patterns" by T. Houtgast in: "Facts and models in hearing", E. Zwicker & E. Terhardt Eds., Springer Verlag, Berlin : 273-274.
- FASTL H., (1975), Pulsation patterns of sinusoids vs. critical band noise, Perception and Psychophysics 18 : 95-97.
- FLETCHER H., (1940), Auditory patterns, Rev. Mod. Physics 12 : 47-65.
- GALAMBOS R., SCHWARTZKOPFF J., RUPERT A., (1959), Microelectrode study of superior olivary nuclei, Am. J. Physiol. 197 : 527-536.
- GARDNER M.B., (1947), Short duration auditory fatigue as a method of classifying hearing impairment, J. Acoust. Soc. Amer. 19 : 178-190.
- GOLDSTEIN J.L., (1967), Auditory nonlinearity, J. Acoust. Soc. Amer. 41 : 676-689.
- GOLDSTEIN J.L., (1970), Aural combination tones in "Frequency analysis and periodicity detection in hearing". R.Plomp & G.F. Smoorenburg Eds., A.W. Sijthoff, Leiden, p. 230-245.
- GREEN D.M., (1976), An introduction to hearing, Lawrence Erlbaum Associates, Publishers, Distributed by John Wiley & Sons, New York.
- GREENWOOD D.D., (1971), Aural combination tones and auditory masking, J. Acoust. Soc. Amer. 50 : 502-543.
- HALL J.L., (1974), Two-tone distortion products in a nonlinear model of the basilar membrane, J. Acoust. Soc. Amer. 56 : 1818-1828.
- HALL J.L., (1977 a), Two-tone suppression in a nonlinear model of the basilar membrane,

- J. Acoust. Soc. Amer. 61 : 802-810.
- HALL J.L., (1977 b), Spatial differentiation as an auditory "second filter" : assessment on a nonlinear model of the basilar membrane, J. Acoust. Soc. Amer. 61 : 520-524.
- HOPKINSON N.T., (1967), Combined effects of interruption and interaural alternation on speech intelligibility, Language and Speech 10 : 234-243.
- HOUTGAST T., (1971), Psychophysical evidence for lateral inhibition in hearing, Proc. 7th ICA, Vol. 3 : 521-524 (24H15), Akademiai Kiado, Budapest.
- HOUTGAST T., (1972), Psychophysical evidence for lateral inhibition in hearing, J. Acoust. Soc. Amer. 51 : 1885-1894.
- HOUTGAST T., (1974 a), Lateral suppression in hearing, Thesis Free University of Amsterdam, Academische Pers B.V., Amsterdam.
- HOUTGAST T., (1974 b), Masking patterns and lateral inhibition, in : "Facts and models in hearing", E. Zwicker & E. Terhardt Eds, Springer Verlag, Berlin, : 258-265.
- HOUTGAST T., (1974 c), The slopes of the masking pattern, in : "Facts and models in hearing", E. Zwicker & E. Terhardt Eds, Springer Verlag, Berlin : 269-272.
- HUGGINS A.W.F., (1964), Distortion of the temporal pattern of speech: Interruption and alternation, J. Acoust. Soc. Amer. 36 : 1055-1064.
- JOHNSTONE B.M., BOYLE A.J.F., (1967), Basilar membrane vibration examined with the Mössbauer technique, Science 158 : 389-390.
- KHANNA S.M., SEARS R.E., TONNDORF J., (1968), Some properties of longitudinal shear waves : a study by computer simulation, J. Acoust. Soc. Amer. 43 : 1077-1084.
- KIANG N.Y.-S., (1965), Discharge patterns of single fibres in the cat's auditory nerve, M.I.T. Press, Cambridge, Mass. U.S.A.
- KIM D.O., MOLNAR C.E., PFEIFFER R.R., (1973), A system of nonlinear differential equations modeling basilar-membrane motion, J. Acoust. Soc. Amer. 54 : 1517-1529.
- KOHLLOEFFEL L.U.E., (1972, a), A study of basilar membrane vibrations I. Fuzziness-detection : a new method for the analysis of microvibrations with laser light, Acustica 27 : 49-65.
- KOHLLOEFFEL L.U.E. (1972 b), A study of basilar membrane vibrations II. The vibratory amplitude and phase pattern along the basilar membrane (post-mortem), Acustica 27 : 66-81.
- KOHLLOEFFEL L.U.E. (1972 c), A study of basilar membrane vibrations III. The basilar membrane frequency response curve in the living guinea pig. Acustica 27 : 81-89.
- KREUL E.J., (1971), Speech intelligibility for interaural alternated speech with and without intervening noise for words and nonsense, Language and Speech 14 : 99-107.
- KRONBERG H., MELLERT V., SCHREINER CH., (1974), Dichotic and monaural pulsation thresholds, Proc. 8th ICA, London, Vol. 1 : 143.
- LAMORE P.J.J., (1975), Perception of two-tone octave complexes, Acustica 34 : 1-14.
- LAMORE P.J.J., (1977), Pitch and masked threshold in octave complexes in relation to interaction phenomena in two-tone stimuli in general, Acustica 37 : 249-257.
- LUESCHER E., ZWISLOCKI J., (1946), Ueber Abklingvorgänge des Ohres, Pract. Oto-Rhino-Lar. 8 : 531-533.
- LUESCHER E., ZWISLOCKI J., (1947), The decay of sensation and the remainder

- of adaptation after short pure-tone impulses on the ear, *Acta Oto-Laryngol.* 35 : 428-445.
- LUESCHER E., ZWISLOCKI J., (1949), Adaptation of the ear to sound stimuli, *J. Acoust. Soc. Amer.* 21 : 135-139.
- MAIWALD D., (1967 a), Beziehungen zwischen Schallspektrum, Mithörschwelle und der Erregung des Gehörs, *Acustica* 18 : 69-80.
- MAIWALD D., (1967 b), Die Berechnung von Modulationsschwellen mit Hilfe eines Funktionsschemas, *Acustica* 18 : 193-207.
- MAREE DE G., (1939), Audiometrische Untersuchungen: Ueber das Verhalten des normalen und schwerhörigen Ohres bei funktioneller Belastung nebst Bemerkungen zur Theorie des Gehörs. *Acta Oto-Laryngologica Suppl.* XXXI.
- MILLER G.A., (1948), The perception of short bursts of noise, *J. Acoust. Soc. Amer.* 20 : 160-170.
- MILLER G.A., LICKLIDER J.C.R., (1950), The intelligibility of interrupted speech, *J. Acoust. Soc. Amer.* 22 : 167-173.
- MØLLER A.R., (1976), Dynamic properties of the responses of single neurones in the cochlear nucleus of the rat, *J. Physiol.* 259 : 63-82.
- MØLLER A.R., (1976), Dynamic properties of excitation and two-tone inhibition in the cochlear nucleus studied using amplitude-modulated tones, *Exp. Brain. Res.* 25 : 307-321.
- MØLLER A.R., (1977), Frequency selectivity of the basilar membrane revealed from discharges in auditory nerve fibres, in "Psychophysics and Physiology of Hearing" E.F. Evans & J. P. Wilson Eds. Academic Press, London, : 197-205.
- NOORDEN VAN L.P.A.S., (1975), Temporal coherence in the perception of tone sequences, Thesis Technical University Eindhoven/ Institute for Perception.
- NOORDEN VAN L.P.A.S., (1977), Minimum differences of level and frequency for perceptual fission of tone sequences abab. *J. Acoust. Soc. Amer.* 61 : 1041-1045.
- PLOMP R., (1961), Hearing threshold for periodic tone pulses, *J. Acoust. Soc. Amer.* 33 : 1561-1569.
- PLOMP R., (1964 a), Rate of decay of auditory sensation, *J. Acoust. Soc. Amer.* 36 : 277-282.
- PLOMP R., (1964 b), The ear as a frequency analyzer, *J. Acoust. Soc. Amer.* 36 : 1628-1636.
- PLOMP R., (1967), Beats of mistuned consonances, *J. Acoust. Soc. Amer.* 42 : 462-474.
- PLOMP R., BOUMAN M.A., (1959), Relation between hearing threshold and duration for tone pulses, *J. Acoust. Soc. Amer.* 31 : 749-758.
- POWERS G.L., WILCOX J.C., (1977), Intelligibility of temporally interrupted speech with and without intervening noise, *J. Acoust. Soc. Amer.* 61 : 195-199.
- RAKOWSKI A., (1971), Pitch discrimination at the threshold of hearing, *Proc. 7th ICA, Vol.3 : 373-376 (20H6) Akademiai Kiado, Budapest.*
- RHODE W.S., (1971), Observations of the vibration of the basilar membrane in squirrel monkey using the Mössbauer technique. *J. Acoust. Soc. Amer.* 49 : 1218-1231.
- RHODE W.S., ROBLES L., (1974), Evidence from Mössbauer experiments for non-linear vibration in the cochlea, *J. Acoust. Soc. Amer.* 55 : 588-596.
- RODENBURG M., VERSCHUURE J., BROCAAR, M.P., (1974), Comparison of two masking methods, *Acustica* 31 : 99-106.

- SACHS M.B., ABBAS P.J., (1974), Rate versus level functions for auditory-nerve fibres in cats: tone burst stimuli, *J. Acoust. Soc. Amer.* 56 : 1835-1847.
- SCHAFER T.H., GALES R.S., SHEWMAKER C.A., THOMPSON P.O., (1950), The frequency selectivity of the ear as determined by masking experiments, *J. Acoust. Soc. Amer.* 22 : 490-496.
- SCHOENE P., (1977), Nichtlinearitäten in Mithörschwellen-Tonheitsmuster von sinustönen, *Acustica* 37 : 37-44.
- SCHREINER CH., GOTTLOB D., MELLERT V., (1977), Influences of the pulsation threshold method on psychoacoustical tuning curves, *Acustica* 37 : 29-36.
- SCHUBERT E.D., PARKER C.D., (1955), Addition to Cherry's findings on switching speech between the two ears, *J. Acoust. Soc. Amer.* 27 : 792-794.
- SHANNON R.V., (1976), Two-tone unmasking and suppression in a forward-masking situation, *J. Acoust. Soc. Amer.* 59 : 1460-1470.
- SMOORENBURG G.F., (1972), Audibility region of combination tones, *J. Acoust. Soc. Amer.* 52 : 603-614.
- SMOORENBURG G.F., (1974), On the mechanism of combination tone generation and lateral inhibition in hearing, in "Facts and Models in Hearing", E. Zwicker & E. Terhardt Eds., p 332-242.
- SMOORENBURG G.F., LINSCHOTEN D.H., (1977), A neurophysiological study on auditory frequency analysis of complex tones, in "Psychophysics and Physiology of hearing", E.F. Evans & J.P. Wilson Eds, Academic Press, London, 175-183.
- STEELE C.R., (1973), A possibility for sub-tectorial membrane fluid motion, in "Basic Mechanisms in Hearing", A.R. Møller Ed., Academic Press, New York, 69-90.
- THURLOW W.R., (1957), An auditory figure-ground effect, *Am. J. Psychol.* 70: 653-654.
- THURLOW W.R., ELFNER L.F., (1959), Continuity effects with alternately sounding tones, *J. Acoust. Soc. Amer.* 31 : 1337-1339.
- THURLOW W.R., MARTEN A.E., (1962), Perception of steady and intermittent sounds with alternating noise-burst stimuli, *J. Acoust. Soc. Amer.* 34 : 1853-1858.
- VERSCHUURE J., (1974), Transient phenomena in pulsation threshold measurement, *Audiology* 13 : 90.
- VERSCHUURE J., (1975), Presentation conditions in the pulsation threshold measurement, NAG-Publication 33 : 58-63.
- VERSCHUURE J., (1977), Pulsation threshold patterns and neuro-physiological tuning in : "Psychophysics and Physiology of hearing", E.F. Evans & J.P. Wilson Eds, Academic Press, London.
- VERSCHUURE J., MEETEREN VAN A.A., (1975), The effect of intensity on pitch, *Acustica* 32 : 33-43.
- VERSCHUURE J., RODENBURG M., MAAS A.J.J., (1974), Frequency selectivity and temporal effects of the pulsation threshold method, *Proc. 8th ICA*, London, Vol. 1 : 131.
- VERSCHUURE J., RODENBURG M., MAAS A.J.J., (1976), Presentation conditions of the pulsation threshold method, *Acustica* 35 : 47-54.
- VOGTEN L.L. M., (1974), Pure-tone masking; a new result from a new method, in : "Facts and Models in Hearing", E. Zwicker & E. Terhardt Eds., Hirzel Verlag, Berlin, : 142-155.
- WARREN R.M., OBUSEK C.J., ACKROFF J.M., (1972), Auditory induction : Perceptual synthesis of absent sounds, *Science* 176 : 1149-1151.

- WEGEL R.L., LANE C.E., (1924), The auditory masking of one pure tone by another and its probable relation to the dynamics of the inner ear, *Phys. Rev.* 23 : 266-285.
- WILSON J.P., EVANS E.F., (1971), Grating acuity of the ear: psychophysical and neurophysiological measures of the frequency resolving power, *Proc 7th ICA*, Vol. 3 : 397-400 (20H14), Akademiai Kiado, Budapest.
- WILSON J.P., JOHNSTONE J.R., (1972), Capacitive probe measures of basilar membrane vibration, *Hearing Theory* 1972, IPO Eindhoven: 172-181.
- WILSON J.P., JOHNSTONE J.R., (1975), Basilar membrane and middle ear vibration in guinea pig measured by capacitive probe, *J. Acoust. Soc. Amer.* 57: 705-723.
- ZWICKER E., (1961), Subdivision of the audible frequency range into critical bands (Frequenzgruppen), *J. Acoust. Soc. Amer.* 33 : 248.
- ZWICKER E., (1970), Masking and psychological excitation as consequences of the ear's frequency analysis, in : "Frequency analysis and periodicity detection in hearing", R. Plomp, G.F. Smoorenburg Eds., : 376-394, A.W. Sijthoff, Leiden.
- ZWICKER E., (1974), On a Psychoacoustical equivalent of tuning curves, in : "Facts and Models in Hearing", E. Zwicker & E. Terhardt Eds., Hirzel Verlag, Berlin : 132-139.
- ZWICKER E., FELDTKELLER R., (1967), *Das Ohr als Nachrichtenempfänger*, 2nd ED. Hirzel Verlag, Stuttgart.
- ZWISLOCKI J., SOKOLICH W.G., (1974), Neuro-mechanical frequency analysis in the cochlea, in "Facts and Models in Hearing", E. Zwicker & E. Terhardt Eds., Springer Verlag, Berlin : 107-117.

CURRICULUM VITAE

De schrijver van dit proefschrift werd geboren op 3 april 1946 te Goes. Hij bezocht het Chr. Lyceum voor Zeeland te Goes en het Eerste Chr. Lyceum te Zeist waar hij het einddiploma Gymnasium β behaalde in 1964.

In 1964 begon hij met zijn studie in de wis- en natuurkunde aan de Rijksuniversiteit te Utrecht. In 1966 legde hij het kandidaatseksamen natuurkunde en wiskunde met geofysika af, in 1970 gevolgd door het doktoraal examen experimentele natuurkunde met bijvakken wiskunde en meteorologie. Het laatste deel van het experimentele werk voor het doktoraal examen werd verricht op de afdeling Medische en Fysiologische Fysika o.l.v. Prof. Dr. M.A. Bouman en Dr. G.F. Smoorenburg. Het betrof onderzoek aan combinatie-tonen.

Na het doktoraal examen werd hij bij dezelfde afdeling aangesteld in dienst van Z.W.O. voor het verrichten van onderzoek aan toonhoogte-luidheids interacties, waarbij een intensief contact werd onderhouden met de auditieve werkgroep van het Instituut voor Zintuigfysiologie RVO-TNO (hoofd Prof. Dr. Ir. R. Plomp).

Per 1 december 1971 trad de schrijver in dienst van de afdeling KNO van de toenmalige Medische Fakuliteit Rotterdam, later de Medische Fakuliteit van de Erasmus Universiteit (hoofd Prof. Dr. W.H. Struben; na 1976 Prof. Dr. E.H. Huijzing). Het beschreven onderzoek werd verricht o.l.v. Dr. M. Rodenburg binnen het kader van de auditieve werkgroep van de vakgroep "Natuurkunde en Technologie", waarin o.a. de fysika groep van de afdeling KNO participeert.

De schrijver is hiernaast betrokken geweest bij de klinische audiometrie en bij enkele andere onderzoekactiviteiten van de afdeling KNO, waarin hij de natuurkundige inbreng verzorgde.