LATERAL SUPPRESSION IN HEARING

a psychophysical study on the ear's capability to preserve and enhance spectral contrasts

TAMMO HOUTGAST

VRIJE UNIVERSITEIT TE AMSTERDAM

LATERAL SUPPRESSION IN HEARING

a psychophysical study on the ear's capability to preserve and enhance spectral contrasts

Academisch proefschrift

ter verkrijging van de graad van
doctor in de Wiskunde en Natuurwetenschappen aan de
Vrije Universiteit te Amsterdam,
op gezag van de rector magnificus Mr. I. A. Diepenhorst,
hoogleraar in de faculteit der Rechtsgeleerdheid,
in het openbaar te verdedigen op woensdag 27 maart 1974
te 15.30 uur in het hoofdgebouw der universiteit,
De Boelelaan 1105

door

TAMMO HOUTGAST

geboren te Zeist

Academische Pers B.V. - Amsterdam

1974

Promotor: Prof. Dr. Ir. R. Plomp

Voor het tot stand komen van deze studie was mijn eigen inzet een wel noodzakelijke, maar geenszins een voldoende voorwaarde. Voor een voorspoedige groei is een juiste voedingsbodem en voldoende vrije ruimte onontbeerlijk; ik ben de Rijksverdedigingsorganisatie TNO, en in het bijzonder het Instituut voor Zintuigfysiologie, dankbaar voor het kreëren van deze voorwaarden. De lichtinval is veelal bepalend voor de richting waarin de groei zich ontwikkelt; de nimmer aflatende en immer stimulerende belangstelling van Reinier Plomp heeft een zeer grote invloed gehad op mijn werk. Het verwijderen van onkruid en het snoeien van vruchteloze loten zijn belangrijk voor een gezonde groei; de voortdurende gedachtenwisseling met mijn naaste kollega's, Louis Pols en Guido Smoorenburg, hebben mij vaak geholpen bij de interpretatie van de metingen en behoed voor het inslaan van doodlopende wegen. Wanneer de groeikracht dreigt te verslappen, is een verkwikkende regenbui onontbeerlijk; in deze zin dient de betekenis van mijn gezin, in het bijzonder van Willemien, voor mijn werk beschouwd te worden. Tenslotte, het verzorgen en het plukken van de vrucht was het werk van vele toegewijde handen; Herman Steeneken en Ton Mimpen hebben een belangrijke rol gespeeld bij het realiseren van de metingen, terwijl de Heer Huigen en Henny Gebbink van onschatbare waarde waren bij de vormgeving in beeld en schrift. Allen dank ik voor hun bijdragen.

The research reported in this study was carried out in the Institute for Perception TNO, Soesterberg, the Netherlands and supported by the Netherlands Organization for the Advancement of Pure Research (ZWO).

CONTENTS

CHAPTER 1		PRELIMINARY NOTES	1
CHAPTER 2		INTRODUCTION	4
CHAPTER 3		THE MEASURING TECHNIQUES, AND AN EXAMPLE OF AN "AUDITORY MACH BAND"	10
	3.1	General technical information	10
	3.2	Three time patterns of stimulus presentation	11
	3.3	Three measuring procedures	13
	3.4	Application to low-pass filtered noise	15
CHAPTER 4		PURE-TONE MASKERS AND NEURAL "TUNING CURVES"	19
	4.1	Four measuring techniques	19
	4.2	Extensive pulsation-threshold measurements	22
	4.3	Discussion	22
CHAPTER 5		TWO-TONE MASKERS AND "TWO-TONE INHIBITION"	26
	5.1	Five measuring techniques	26
	5.2	Extensive pulsation-threshold measurements	31
CHAPTER 6		INTERIM DISCUSSION: DIRECT MASKING VERSUS NON-SIMULTANEOUS MASKING	35
CHAPTER 7		TONE-PLUS-NOISE MASKERS, AND THE EFFECT OF NOISE-ON-TONE SUPPRESSION	38
	7.1	Three measuring techniques	38
	7.2	Noise-induced loudness reduction	41
	7.3	Tone plus band-rejected noise	43

CHAPTER 8 MASKING EXPERIMENTS WITH NOISE OF VARIABLE BANDWIDTH, AND THE "CRITICAL BAND" CONCEPT 48
8.1 The masker 48
8.2 Two measuring techniques 49
8.3 Discussion 50

CHAPTER 9 MASKING EXPERIMENTS WITH RIPPLED NOISE, AND THE "AUDITORY FILTER" 54
9.1 The masker 55
9.2 Four measuring techniques 56
9.3 Effect of modulation depth 60
9.4 Filter characteristics derived from rippled-noise masking data 62
9.5 Auditory bandwidth derived from ripple-resolution data 67

CHAPTER 10 DISCUSSION 74
10.1 Evaluation of the experimental results 74
10.2 Lateral suppression and neural data 80
10.3 Additional notes on the pulsation-threshold method 84

CHAPTER 11 AUDITORY SPECTRA OF VARIOUS STIMULI 89
11.1 General 89
11.2 Pulsation-threshold patterns of pure tones 91
11.3 Pulsation-threshold patterns of complex tones 95
11.4 Pulsation-threshold patterns of vowel-like sounds 99

SUMMARY AND CONCLUSIONS 103

SAMENVATTING EN KONKLUSIES 106

REFERENCES 110

CHAPTER 1

PRELIMINARY NOTES

This chapter is intended for the non-specialist reader. It presents an elementary framework of some common concepts and experimental approaches in auditory research and indicates the orientation of the present study within that framework.

For the purpose of this preliminary introduction we may consider the hearing mechanism as a particular type of input-output system. The input, typically air-pressure fluctuations, can be specified in physical terms, whereas the "output", the sound sensation, is to be described in perceptual terms. The general aim of auditory research is to establish the rules governing the input-output relationships and to understand the processes underlying these rules.

The wide variety of experimental approaches in auditory research can be divided into two main groups: (1) psychoacoustics, a typical "over-all" approach, based on the responses of subjects to well-defined sound stimuli, and (2) physiology, a typical analytical approach, in which the structure and organization of specific parts of the hearing pathway as well as their auditory functions are studied. Psychoacoustics is particularly suited to establish the input-output relationships, but it gives no direct information about the underlying processes. For instance, listening experiments may teach us that a frequency increment of 1% for a pure tone is just audible by a subject; however, there are many possible mechanisms which might explain this input-output rule. On the other hand, physiology provides us with detailed information about features and processes at various stages of the auditory pathway, but the question remains whether a specific physiological feature is relevant for perception. For instance, it has been found that a frequency increment of a pure-tone stimulus may result in at least two physiological changes: (1) a change in the temporal pattern of neural discharges and (2) a change in the distribution of the number of discharges over different neural elements. It is an open question which change is relevant for the change of the pitch of the tone.

A functional description of the hearing mechanism, in the form of a model, should both reflect the rules found in psychoacoustics and be in line with physiological data. I will present here only some basic properties of such a model, based on current knowledge of both psychoacoustics and physiology. The only purpose is to present a framework for indicating some main topics of to-day's auditory research as well as the subject of the present study.

A convenient step for arriving at a model is to distinguish two parts in the auditory pathway, a mechanical part and a neural part. The first part "simply" transmits the mechanical vibration, without a substantial modification of the original vibration. Within the cochlea the two parts meet. The "input" to the cochlea is the vibration of the stapes and its "output" is the activity in the ensemble of single nerve fibres leaving the cochlea. The stapes vibrations are transformed into a "pattern" of activity in the ensemble of primary fibres. (At a later stage, I will be more specific about the vague notion "pattern".) This pattern of neural activity is transmitted to higher centres, leading to the ultimate sound sensation. These latter processes might be described in terms of pattern transformation and pattern recognition.

The properties and limitations of the transformations in the cochlea, up to the level of the primary fibres, are considered to be of fundamental importance for the rules found in psychoacoustics, especially those in the frequency domain. To gain insight in these transformations is a most important goal of physiological research in audition. For the present moment I mention only two basic findings concerning the transformation from stapes vibration into a pattern of (peripheral) neural activity. The first one is a property in the *frequency domain*. For a pure tone as stimulus, the activity in a particular neural unit (number of neural discharges or spikes per time interval) is largest for a specific frequency of the tone (the so-called characteristic frequency or *best frequency* of the unit). When the frequency of the tone is shifted away from the unit's best frequency, the activity will drop considerably. This drop can be counterbalanced by an increase of the

level of the tone. This *frequency selectivity* of a unit's response is thought to be closely related to the filtering properties of the mechanical structures in the cochlea. The second important feature to be mentioned concerns the *time domain*. Analysis of the relation between the temporal pattern of a unit's activity and the time function of the stimulus reveals a high degree of *temporal resolution*. One aspect of this is that, when the input is a tone of not too high a frequency (up to 2 - 3 kHz), its periodicity is reflected in the temporal pattern of the neural discharges. Another aspect of this temporal resolution can be seen in a unit's response to a stimulus with fast intensity fluctuations.

I can now be more specific about the notion of an acoustic stimulus being transformed into "a pattern of neural activity" in the ensemble of primary fibres. Acoustic stimuli, such as speech or music, bring about a specific distribution of neural activity along two "dimensions", time and fibre-ensemble: (1) the time pattern of the spikes in a fibre (or in a restricted group of fibres) is related to the *time function* of the stimulus (within a restricted frequency region) and (2) the distribution pattern of the number of spikes within a (short) time interval in the fibre-ensemble is related to the (short-term) *sound spectrum* of the stimulus. This latter pattern, which may be named the "neural projection", or the "auditory projection", of a stimulus' sound spectrum, is of primary interest to the present study.

Within the framework presented here, this study is orientated as follows. From electrophysiological studies in animals some basic properties of the auditory projection of a stimulus' sound spectrum are well known. One intriguing aspect is the effect known as *lateral inhibition*. At this place, it may suffice to say that, possibly, lateral inhibition increases the "contrast" in the auditory projection of a stimulus' sound spectrum. Can such effects of contrast enhancement be revealed by listening experiments? This question lies at the roots of the present study.

CHAPTER 2

INTRODUCTION

Summary

When a steady-state stimulus is presented to the ear, the shape of its sound spectrum is reflected, to some degree, in the vibration pattern along the basilar membrane and in the distribution pattern of neural activity in the ensemble of primary auditory neurons. This *auditory projection* of a stimulus' sound spectrum is considered to form the basis for at least several perceptual attributes of the stimulus, such as timbre, loudness and, perhaps, pitch. How well does this projection reflect the original shape of the sound spectrum? In other words, what is the degree of "sharpness" of the projection and, above all, is there any evidence for "contrast enhancement", suggesting the operation of a lateral-inhibition mechanism? These are the basic questions of this study. An example of the effect of lateral inhibition in vision is given. It is indicated why I use the term lateral *suppression* rather than inhibition. It appears that, in contrast to some electrophysiological studies, traditional psychophysical data give essentially no indication for the existence of lateral suppression in audition. If it might appear that other than such traditional psychophysical methods do reveal effects of lateral suppression, how to decide which method, if any, gives the "correct" picture of the auditory projection of a sound spectrum?

The significance of frequency selectivity in hearing is beyond doubt. In this respect, the frequency-selective properties of the mechanical structure of the inner ear are considered most important. Each part of the basilar membrane, between the oval window and the helicotrema, is set into vibration most easily (thus, responds most heavily to) a stimulus with a specific frequency; this "resonance frequency" becomes progressively lower when moving upwards from oval window to helicotrema. Hence, in response to a sound stimulus, the maximum deflection at different positions along the basilar membrane shows a pattern which is related to the stimulus' sound spectrum; this

vibration pattern reflects, to some degree, the sound spectrum of the stimulus. Although there are still many unknown details in the processes leading from basilar-membrane vibrations to neural activity, initiated in the hair cells located along the basilar membrane and transmitted by the ensemble of single fibres leaving the cochlea, the result is such that the identity of a vibration pattern is preserved as a specific distribution pattern of neural activity in the ensemble of primary auditory neurons. Hence, the distribution pattern of neural activity evoked by a stimulus reflects, to some degree, the stimulus' sound spectrum.

Of course, as mentioned already in chapter 1, it should be realized that the distribution pattern of neural activity is not the *only* possible source of "information" about the actual stimulus, available at that peripheral stage in the auditory pathway. The temporal structure of the neural discharges in each individual fibre also carries information, namely about the temporal fine structure of the vibration of that part of the basilar membrane where the neural discharges are initiated. A very basic question in hearing theory is *what* source of "information" (distribution pattern or temporal structure of the neural activity) is relevant for the different perceptual attributes of a sound. Most commonly, *loudness* and *timbre* of steady-state sounds are thought to be derived from the distribution pattern, whereas the temporal characteristics of the neural discharges are considered highly relevant for the perception of the *direction* of a sound source. With respect to *pitch* perception, theories have not yet converged into a common opinion about which of the two "information carriers", distribution pattern or temporal structure, is the relevant one.

The above considerations are meant to illustrate that, at a very early stage in the auditory channel, the stimulus is transformed into a spatial *pattern* of vibration and, subsequently, into a distribution pattern of neural activity. At least several perceptual attributes of steady-state sounds are derived from specific features of the *shape* of this auditory projection of the sound spectrum. How *well* then does this shape reflect the contours, peaks and valleys of the original sound spectrum? In other words, how "sharp" is the projection, what is "lost" and what is still "resolved" in the first stages of auditory processing? It is clear that the degree of frequency resolution is limited simply because an "infinite" frequency resolution would imply an "infinite" time of analysis. Realizing that, in addition to resolution in the frequency domain, another important feature of hearing is resolution in the time domain (and that the inner ear as a frequency-selective mechanism cannot be "by-passed") it is clear that neither of these can be perfect. A given degree of resolution in the one domain sets a limit to the degree of resolution in the other domain. For the moment, I go no further than this simple qualitative notion: the auditory projection of a sound spectrum is submitted to smoothing, contrast-diminishing effects caused by the limited degree of frequency resolution of the mechanical structure of the inner ear.

The considerations given so far may be considered as very common and generally accepted notions on hearing. I now come to the question underlying the present study. Of course, some unsharpness is essential to any projection of a pattern, not restricted exclusively to the auditory projection of a sound spectrum. For several other sense modalities (vision, skin sensation) it is known that, besides the essential unsharpness, another important process, named *lateral inhibition*, is involved in the peripheral processing of a stimulus pattern (for instance, in the case of a light distribution pattern presented to the eye or a pattern of vibration along the skin). For a survey of such effects in vision I may refer to Ratliff (1965), and with respect to skin sensation to von Békésy (1958, 1960). The general feature of this process of lateral inhibition is that "strong" parts in a pattern have an inhibitory, suppressing effect on adjacent "weaker" parts. Hence, if the projection of a pattern is subjected to lateral inhibition, weak parts become still weaker relative to adjacent strong parts. Generally speaking, the effect of lateral inhibition is *contrast-enhancement*. An intriguing question is: does this phenomenon also play a role in hearing? In other words, is it possible that the contrast in the auditory projection of a sound spectrum is enhanced by a process like lateral inhibition? Such an effect of contrast enhancement would be of interest especially in the light of the contrast-*diminishing* effects introduced by the limited degree of frequency resolution of the mechanical structure of the inner ear.

As an illustration of the effect of lateral inhibition, let us consider a well-known demonstration of it in vision: the Mach bands. This phenomenon can be observed at the transition between two fields of different luminance. (For example, such a configuration can be seen in Fig. 3.1, page 17). Parallel to the boundary there appear two "bands": a dark one at the darker side of the boundary and a bright one at the brighter side. This is a typical effect of contrast enhancement: whereas the luminance is distributed according to a simple step function, the subjective brightness clearly shows "overshoot" and "undershoot". Apparently, the local brightness in a pattern is

not determined solely by the local luminance. The common concept is that the local brightness is subjected to negative influences from the near surround, or, in other words, subjected to lateral inhibition. The more the luminance in the surround, the stronger this negative influence of lateral inhibition.

The auditory parallel of the visual Mach band would occur if we consider a stimulus with a sound spectrum according to a step function, *e.g.* white noise passed through a sharp high-pass or low-pass filter. If the auditory projection of this sound spectrum is submitted to lateral inhibition, we would expect some enhancement near the edge of the noise spectrum. We will return to this in the next chapter.

In vision, it is generally accepted that the origin of the contrast-enhancement is a *neural* inhibitory process. However, in audition it cannot be excluded a priori that similar effects of contrast enhancement might have their origin in properties of the mechanical structure of the inner ear. Therefore, rather than using the "loaded" term lateral inhibition, I prefer to use the more neutral term lateral *suppression* to denote any process, of which the origin may be either neural or mechanical, leading to contrast enhancement in the auditory projection of a sound spectrum. The primary goal of the present study is to trace possible *effects* of lateral suppression in hearing (thus, generally, contrast enhancement) rather than to investigate or speculate on the nature of the underlying processes.

The idea that lateral suppression might play a role in audition is not new. It was promoted, for instance, by von Békésy, who pointed to the close similarity between hearing and skin sensation (von Békésy, 1958, 1960). However, psychophysically its operation has not been demonstrated convincingly. Some indications might be deduced from the phenomenon, reported by several authors, that listeners can easily match the frequency of a simple tone to the cut-off frequency of a band of noise (von Békésy, 1963; Small and Daniloff, 1967; Rakowski, 1968). This would indicate that some enhancement at the edge of the noise spectrum might indeed occur; however, this is far from conclusive (see also Rainbolt, 1968).

Is it possible to investigate the auditory projection of a sound spectrum more directly by psychophysical means? A traditional type of experiment is to superimpose on the sound to be investigated a *test tone* of variable frequency and to determine the *masked* threshold of the test tone as a function of its frequency. This *masking pattern* is often considered to be closely related to the auditory projection of the masker's sound spectrum. This method, besides others, will be used in this study and I will return to it in more detail. The point of interest here is that, generally, such traditional masking patterns do *not* reveal effects which are typical of the operation of a lateral-suppression mechanism (thus, any effect of contrast enhancement). This leads to three possible conclusions: (1) lateral suppression plays no role in audition, (2) lateral suppression plays a role, but the "unsharpness" caused by the limited degree of frequency resolution is already so severe that the over-all result does not show the typical contrast-enhancing effects, or (3) the traditional masking pattern gives no adequate picture of the auditory projection of the masker's sound spectrum. As will be seen, this latter possibility plays an important role in the present study.

Whereas there appears to be essentially no psychophysical evidence for the operation of lateral suppression in hearing, the results of some electrophysiological studies on the response of primary auditory neurons in a number of different mammals *do* show typical effects of lateral suppression. For example, it has been found that the response (average number of spikes in a time interval) to a pure tone at a neuron's best frequency may *decrease* when adding a second, strong, tone of appropriate frequency and level. (We will return to this in chapter 5.) This well-documented phenomenon, named two-tone inhibition, is a typical effect of a lateral-suppression mechanism: the strong tone in the spectrum suppresses the response to an adjacent weak tone. Thus, it would appear that at this peripheral neural level the auditory projection of such a sound spectrum is, indeed, subjected to the operation of a lateral-suppression mechanism.

Summarizing the present situation, we arrive at the following picture. In vision, and also in skin sensation, the operation of a lateral-suppression mechanism in the peripheral processing of a stimulation pattern is well accepted. The typical result of such a mechanism is contrast *enhancement*. With respect to hearing, the only positive indication concerning the operation of lateral suppression in the auditory projection of a sound spectrum is given by electrophysiological studies on neural responses to various sound stimuli. Results of traditional psychophysical measurements, *e.g.* the *masking pattern* produced by a sound, do *not* show effects of contrast enhancement.

The primary goal of the present study is to investigate the possible role of lateral suppression in hearing by psychophysical experiments. The principle, as in traditional masking experiments, of using a test tone of variable frequency, in an attempt to "scan" the auditory projection of a

sound spectrum, seems attractive. However, in addition to the traditional approach of measuring the masked threshold of a test tone *superimposed* on the sound to be investigated, I will use other methods as well. In the next chapter, the different methods are introduced and, as an example, applied to a stimulus which consists of white noise passed through a sharp low-pass filter in an attempt to obtain the auditory parallel of the visual Mach-band phenomenon.

The use of a number of different methods may raise a problem. Apparently, traditional masking patterns reveal no effects of lateral suppression. If it is found that the results of other methods *do* so, how to decide which method gives the most adequate picture? Apart from possible theoretical considerations, a most obvious approach is to apply the different methods to the type of stimuli which is commonly used in electrophysiological studies on single-nerve responses (*e.g.*, one or two pure tones). Generally speaking, that method which provides results in line with the single-nerve data can be considered to give the most adequate picture of the auditory projection of the sound spectrum. This notion underlies the experiments on one or two pure tones described in chapters 4 and 5. The subsequent chapters are directed to the auditory projection of stimuli with a more complex sound spectrum. Also in those cases, the incidentally available electrophysiological data are helpful in interpreting the results obtained with the different experimental methods.

CHAPTER 3

THE MEASURING TECHNIQUES, AND AN EXAMPLE OF AN "AUDITORY MACH BAND"

Summary

Three measuring techniques of the "masker-and-test-tone" type are described: direct masking, forward masking and pulsation threshold. The latter is a relatively new technique. As an example, these methods are applied to a masker consisting of sharply low-pass filtered noise. The main purpose is to show that the results of the three techniques are essentially *different*: a typical Mach-band phenomenon is revealed only by the latter two methods and not by the traditional direct-masking technique. In view of these differences, we shall apply the three measuring techniques to the type of maskers for which the neural response has been studied extensively.

Almost all experiments described in this study are of the "masker-and-test-tone" type. The underlying concept is that such experiments might give information concerning the auditory projection of the masker's sound spectrum. Within the group of masker-and-test-tone experiments, however, there are many possible experimental paradigms. Before describing the paradigms which are used in this study, I will first give some basic information, common to all experiments.

3.1. GENERAL TECHNICAL INFORMATION

All experiments were performed monaurally by means of a Beyer DT-48 telephone. Two types of stimuli were used most commonly: pure tones and noise. The level L of a pure tone is expressed in decibel (dB) relative to the level of a 200-msec 1000-Hz tone at hearing threshold: L in dB *re* 1000-Hz absolute threshold. (This absolute threshold is determined with the 2-AFC up-down procedure described below.) In some cases, a pure-tone level is ex-

pressed in dB relative to the level of a 200-msec tone of that *same* frequency at hearing threshold; then, it is indicated as dB SL (Sensation Level). The noise is characterized by its *spectral level* N_0: the intensity in 1-Hz intervals in dB relative to the intensity of a 200-msec 1000-Hz tone at hearing threshold (N_0 in dB/Hz *re* 1000-Hz absolute threshold).

Most figures include a schematic representation of the masker's sound spectrum (see, for example, the upper diagram in Fig. 3.1). In these diagrams, the test tone is always indicated by an interrupted line. The double-pointed arrow will always indicate the *independent* variable, set by the experimenter (thus, in Fig. 3.1, the test-tone frequency), whereas the single-pointed arrow indicates the *dependent* variable of which the threshold value is to be measured (in Fig. 3.1, the test-tone level). Several experimental methods can be applied to measure a threshold value, depending on the *time pattern* of presentation of masker and test tone and on the *procedure* used in manipulating the dependent variable.

3.2. THREE TIME PATTERNS OF STIMULUS PRESENTATION

a. Direct masking

The essential feature of direct masking (or *simultaneous* masking) is that the test tone is *superimposed* on the masker. A most simple configuration is given in Fig. 3.1, in which test-tone bursts are superimposed on a continuous masker. In direct masking, the threshold value to be measured always refers to the *detectability* of the test tone. Generally, over some range of the dependent variable, the detectability of the test tone changes from "perfect" to "impossible" and this simple psychophysical relation forms the basis for the different procedures to arrive at a threshold value.

b. Forward masking

In this case, a short test-tone burst is presented shortly after the termination of the masker, as in the second condition in Fig. 3.1. Again, as in direct masking, the threshold value to be measured refers to the *detectability* of the test-tone burst.

c. Pulsation threshold

In this method, masker bursts and test-tone bursts, both of 125 msec duration, are presented alternately, as indicated in the lower panel in Fig. 3.1. In this case, the threshold value to be measured does not refer to the effect of the masker on the *detectability* of the test tone. It refers to the effect of the masker burst on the *temporal character* of the series of test-tone burst as perceived by the listener, namely either as a *pulsating* tone or as a *continuous* tone. This effect of continuity, which is a general phenomenon when alternating fainter and louder sounds, has been described in the literature. We will return to this in section 10.3.1. It has not been applied systematically in masker-and-test-tone type of experiments. Its applicability may be illustrated by an example taken from Fig. 3.1. If, for example, the test-tone frequency is 1000 Hz, the series of test-tone bursts, alternated with the white-noise bursts, is perceived as a *pulsating* tone for high test-tone levels (above 40 dB) and as a *continuous* tone for a range of lower test-tone levels (roughly between 10 and 40 dB). Thus, depending on the value of the test-tone level (the dependent variable) the series of test-tone bursts is perceived either as pulsating or as continuous and this simple psychophysical relation forms the basis for arriving at a threshold value of the dependent variable: the pulsation threshold.

It is clear that the pulsation-threshold method is somewhat different from a "normal" masking experiment in which a threshold value always refers to the effectiveness of the masker in making the test tone *undetactable*. The pulsation threshold refers to the effectiveness of the masker in making the series of test-tone bursts sound as a *continuous* tone. Still, also in this case, I simply use the term "masker", although this goes beyond the usual definition of masking. Strictly speaking, "masking" refers to the audibility of the masked signal as a whole; pulsation threshold refers to the masking of only *one* specific feature of the series of test-tone bursts, namely its pulsating character.

Most data graphs are accompanied by a schematic presentation of the temporal envelope of masker and test tone. In these diagrams, the masker is always indicated by hatching, and the test tone by an interrupted line. It will be seen that the temporal envelopes have smooth rise and fall curves. All gating was performed with "slow gates" such that the temporal envelope of the gated signal changed gradually from zero to maximum, or from maximum to zero, according to half of a sinewave function. The time interval from zero to maximum (or from maximum to zero) could be set between 1 and 200 msec and is indicated in the diagrams.

3.3. THREE MEASURING PROCEDURES

a. Adjustment

In this case, the subject himself can control the dependent variable by turning a knob. The subject can manipulate it freely while the measuring condition is repeated on and on. In the case of a detection threshold (direct masking, forward masking), his instruction reads: adjust the knob to that position at which the test-tone bursts are *just not* perceived; in the case of a pulsation threshold: adjust the knob to that position at which the pulsating character of the series of test-tone bursts is *just not* perceived. The setting of the dependent variable thus obtained is considered the threshold value.

Though the directness of this procedure may be considered as an advantage, there are some serious objections against its "objectiveness". The basic problem is that the underlying psychophysical relation is always a more or less gradual function: in varying the dependent variable, the perceptual change may go from detectable via "less" detectable to undetectable, or from pulsating via "hardly" pulsating to continuous. In a simple adjustment procedure, both the subject's strategy in changing the dependent variable (for example, he can use large or small variations, he can approach the threshold from "above" or from "below") and his own interpretation of the criterion "just not perceived" remains obscure. However, in so far a subject uses a constant strategy and a constant criterion, the *differences* between the threshold values for different conditions are still meaningful.

b. Békésy up-down

In this procedure, as in the previous one, the measuring condition is repeated on and on. The dependent variable is changed automatically and the subject can only control the *direction* of that change by pushing or releasing a button. In this study, this procedure is only used for a level as the dependent variable. For example, if the test-tone level is the dependent variable, the subject's instruction in the case of a detection threshold reads: release the button (the level decreases) when you clearly hear the test-tone bursts and push it (level goes up) when you do no longer hear them. In the case of pulsation threshold his instruction reads: release the button when you perceive the series of test-tone bursts as clearly pulsating and push it when you perceive it as clearly continuous. The initial level, used as the starting position, was always adjusted by the subject as a first approximation of the threshold. The average value of the dependent variable during a fixed period of this "up-down tracking" was considered as the final threshold value, unless it deviated more than 6 dB from the first approximation. If this was the case, the procedure was repeated with the rejected threshold value as the new starting position.

In this case, the method in changing the dependent variable is, to a large extent, forced upon the subject. Still, the procedure is far from "objective" since the subject uses his own interpretation of criteria such as "clearly detectable" or "clearly pulsating". But, again, insofar as the influences of these unknown factors are constant from condition to condition, the *differences* between threshold values thus obtained are still reliable.

c. Two-alternative forced-choice (2-AFC up-down)

This is considered an efficient and reliable method in psychoacoustics (Levitt, 1971). It can only be applied properly to a detection threshold, thus, in case of direct masking or forward masking. In this study, it is exclusively used for a level as the dependent variable. The stimuli are presented in a sequence of so-called "trials", typically one trial every 3 to 4 sec. Each trial consists of *two* observation periods which, if necessary, can be marked by warning lights. Whereas the condition of the masker is identical for both observation periods, the test-tone burst is presented only in *one* of the two periods, chosen at random. The subject has two push buttons at his disposal. After the two observation periods, he has to indicate, by pushing one of the two buttons, during which of the two periods a test tone had been presented. The subject is immediately informed about the correctness of his decision. Between the trials, the dependent variable is changed according to the following rule: after an incorrect response the level is changed 2 dB in that direction for which detectability improves (thus, the test-tone level is increased or the masker level is decreased), whereas after two successive correct responses the level is changed 2 dB in the opposite direction. The initial level, used as the starting position, is adjusted by the subject as a first approximation of the threshold. The average value of the level during a fixed number of trials (typically, runs of 50 or 100) is considered as the threshold level, unless it deviates more than 6 dB from the starting position. If this is the case, the procedure is repeated with the rejected threshold level as the new starting position.

This procedure excludes to a large extent the unwanted and unknown effects of a subject's own strategy and criterion. The psychophysical relation between the detectability of the test tone and the level of the dependent variable is often expressed here as a psychometric function: the probability of a correct response as a function of the level of the dependent variable. This function goes from 50% (test tone not detectable) to 100% (test tone perfectly detectable). It can be shown that the threshold value arrived at with the present procedure is that level for which the probability of a correct response is about 71%. Thus, insofar as possible influences of a subject's concentration or fatigue on the underlying psychometric function can be neglected, the threshold values themselves as well as their differences are meaningful.

Summarizing, the following information is essential to all data graphs of masker-and-test-tone experiments:

- a specification of the spectral composition of masker and test tone which, in most cases, is illustrated by a schematic diagram which also indicates the dependent and independent variables.
- a specification of the temporal composition of masker and test tone (direct masking, forward masking or pulsation threshold), which is also illustrated by a schematic diagram in most figures.
- a specification of the procedure used to obtain the threshold value (adjustment, Békésy up-down or 2-AFC up-down).

3.4. APPLICATION TO LOW-PASS FILTERED NOISE

The aim of this experiment was to investigate possible effects near the edge of a sharply low-pass filtered noise, which would suggest the influence of lateral suppression. As a well-known visual illustration of such an effect, Fig. 3.1 (top left) presents a spatial light-dark distribution according to a simple step function, which shows the Mach-band phenomenon (a "dark" and a "bright" band parallel to the boundary), as discussed in the introduction.

3.4.1. Measurements

Two types of maskers were used, as indicated in the upper panel in Fig. 3.1: (a) white noise with a spectral level N_0 of 36 dB/Hz and (b) that same noise led through a sharp low-pass filter with a cut-off frequency of 1100 Hz (-3 dB at 1100 Hz and -20 dB at 1150 Hz). Additional to these two masker conditions, a third condition was included in which no noise was presented (unmasked threshold). As indicated, the test-tone frequency was the independent variable (settings from 700 to 1400 Hz) and the test-tone level was the dependent variable. Three measuring techniques were used.

a. Direct masking, 2-AFC up-down

The masker was presented continuously. Two observation periods, separated by an interval of 1 sec and marked by warning lights, were repeated every 3.5 sec. At one of the two periods (randomly) a 200-msec test-tone burst was presented with 17-msec rise and fall times. The 2-AFC up-down procedure with runs of 50 trials was used to measure the threshold value. Each data point is the average of two such threshold measurements by one subject.

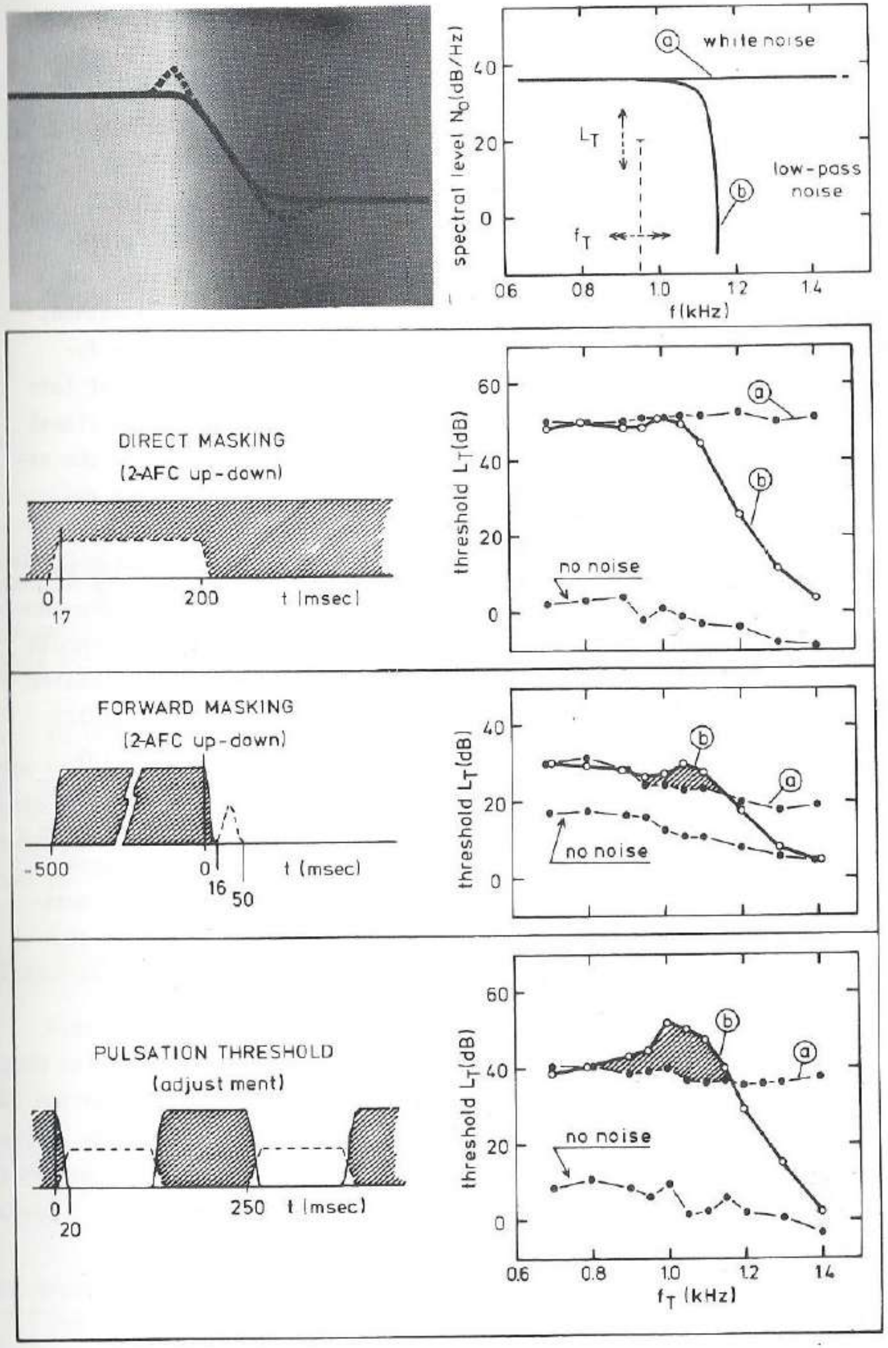
b. Forward masking, 2-AFC up-down

Two observation periods, separated by an interval of 1 sec, were repeated every 3.5 sec. In both periods, a 500-msec masker burst was presented, with 16-msec rise and fall times. After one of the two masker bursts (randomly) a 34-msec test-tone burst was presented with 17-msec rise and fall times. The 2-AFC up-down procedure with runs of 50 trials was used to measure the threshold value. Each data point is the average of two such threshold measurements by one subject.

c. Pulsation threshold, adjustment

125-msec masker bursts and 125-msec test-tone bursts were presented in continuous alternation. All rise and fall times were 20 msec. The subject's concentration on the test tone was facilitated by leaving out each fourth test-tone burst. Then, the condition "continuity" corresponds to the perception of "long" test-tone bursts in a one-sec cycle (the three successive test-tone bursts are perceived as one continuous tone), whereas the condi-

Fig. 3.1 (opposite page). "Sharpening" at the edge of a noise spectrum. As a visual example, the figure in the top-left panel presents a light-dark pattern in which a bright and a dark Mach band can be observed; whereas the luminance is distributed as indicated by the continuous curve, the brightness (interrupted curve) clearly shows edge effects, typical of lateral suppression. The masking experiments aim at an auditory parallel of the bright Mach band. The spectral composition of the masker and the test tone is indicated in the upper diagram, and the temporal composition according to the three different masking paradigms in separate diagrams at the left side of each data graph. (The test tone is always indicated by an interrupted line.) An edge effect (shaded areas in the data graphs) is only shown by the second and third methods. (Data of one subject.)



tion "pulsating" corresponds to the perception of series of three "short" test-tone bursts. The subject could control the test-tone level by turning a knob and adjusted it to pulsation threshold. Each data point is the average of two such adjustments by one subject.

3.4.2. Discussion

The interesting point of the data in Fig. 3.1 concerns the difference for the two conditions white noise (a) and low-pass noise (b). It will be seen that, both in forward masking and with the pulsation-threshold method, the threshold values near the edge of the low-pass noise exceed those for the white noise. This edge effect, which is typical of the operation of lateral suppression and may be compared with the bright Mach band in the visual example, is *not* present in direct masking. This result deviates from the results of similar experiments reported by Carterette *et al.* (1969) who applied a direct-masking technique and did find edge effects, which were interpreted in terms of auditory Mach bands. However, this interpretation was subject to severe criticism (see Carterette *et al.*, 1970). A replication of these direct-masking experiments showed essentially no edge effect (Rainbolt *et al.*, 1972), which agrees well with our present result. For more extensive data on the subject of "Mach bands in hearing", I refer to Houtgast (1972).

The main purpose of this example is to show that different measuring techniques of the masker-and-test-tone type may give essentially different results. How to investigate then which method, if any, gives a correct picture of the auditory projection of the masker's sound spectrum? This question led us to apply the different measuring techniques to the type of maskers for which the activity evoked in primary auditory neurons has been studied extensively.

CHAPTER 4

PURE-TONE MASKERS AND NEURAL "TUNING CURVES"

Summary

The different masking paradigms (direct masking, forward masking, and pulsation threshold) are applied to a pure-tone masker. This enables a comparison between the frequency selectivity revealed by the different masking techniques and the well-documented frequency selectivity in the neural response to a pure tone. In this respect, the pulsation-threshold data agree well with "tuning curves" of primary auditory neurons.

In this chapter, the different masking techniques are applied to the simple case of a pure-tone masker. The aim of these experiments was to investigate a possible correspondence between psychophysical masking data and electrophysiological single-nerve data, with respect to frequency selectivity in the case of a pure-tone stimulus.

4.1. FOUR MEASURING TECHNIQUES

This experiment is illustrated in Fig. 4.1. The test-tone was fixed at 1000 Hz and 23 dB. The masker was a pure tone with frequency f_2 and level L_2. As indicated, f_2 was the independent variable (settings from 400 to 1350 Hz) and L_2 was the dependent variable. For the condition f_2 = 1000 Hz the test tone and the masker were presented in phase. Four measuring techniques were used.

a. Direct masking, 2-AFC up-down (long test tone)

The masker was presented continuously. Two observation periods, separated by an interval of 1 sec and marked by warning lights, were repeated every 3.5

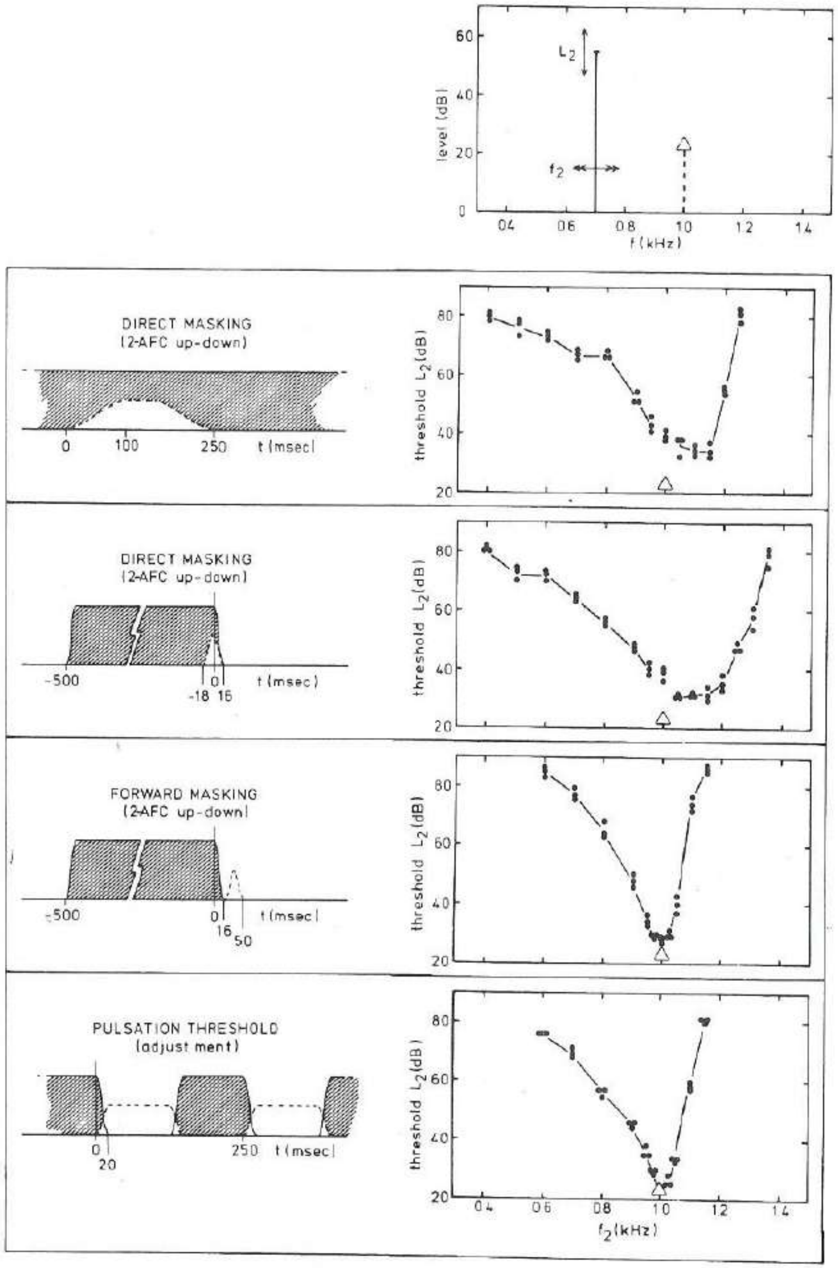


Fig. 4.1. Masking experiments with a fixed test tone (indicated by the position of the triangle) and a pure-tone masker with variable frequency f_2 and level L_2. (Data of one subject.)

sec. At one of the two periods (randomly) a 250 msec test-tone burst was presented with 100-msec rise and fall times. The 2-AFC up-down procedure with runs of 50 trials was used to measure the threshold value. For each setting of f_2, three such measurements were performed with one subject and each individual value is plotted in the data graph.

b. Direct masking, 2-AFC up-down (short test tone)

Two observation periods, separated by an interval of 1 sec, were repeated every 3.5 sec. In both periods a 500-msec masker burst was presented with 16-msec rise and fall times. At the tail of one of the two masker bursts (randomly) a 34-msec test-tone burst was presented with 17-msec rise and fall times, such that the end of the test-tone burst coincided with the end of the masker burst. The 2-AFC up-down procedure with runs of 50 trials was used to measure the threshold value. The results of three such measurements for one subject at each setting of f_2 are plotted as individual data points.

c. Forward masking, 2-AFC up-down

This measurement was identical to the previous one, except that the test-tone burst was "shifted" over an interval of 34 msec, such that its start, rather than its end, coincided with the end of one of the two masker bursts.

d. Pulsation threshold, adjustment

The pulsation-threshold method was applied with the continuous alternation of 125-msec masker bursts and 125-msec test-tone bursts with, again, each fourth test-tone burst missing. The subject adjusted L_2 to pulsation threshold (the lowest value of L_2 for which the pulsating character of the series of test-tone bursts was just not perceived). The results of three such adjustments for one subject at each setting of f_2 are plotted as individual data points.

Each of the four curves in Fig. 4.1 can be considered as representing an ensemble of pure-tone stimuli with a common feature, namely just masking a particular test tone in a particular experimental paradigm. It will be seen that the frequency selectivity revealed by these curves depends on the experimental paradigm. Typically, the two curves obtained in direct masking are broader than the two other curves. However, the difference is not so dramatic as to justify extensive experiments, involving more subjects and more conditions for the fixed test tone, with the explicit purpose to differentiate between the various experimental techniques by comparing the results with electrophysiological data obtained on animals. Quantitative differences of the same order of magnitude may very well exist between the auditory systems of various specimens, which would make any conclusion based on a detailed quantitative comparison rather dubious. The similarity between equal-masking curves, obtained in direct masking, and tuning curves of primary auditory neurons in cat (*e.g.*, Small, 1959) plays an important role in auditory theories. I shall, therefore, restrict the more extensive experiments to the new pulsation-threshold technique to show that also the results of that experimental paradigm agree reasonably well with neural data.

4.2. EXTENSIVE PULSATION-THRESHOLD MEASUREMENTS

This experiment is illustrated in Fig. 4.2. A series of measurements was performed for six conditions for the fixed test tone, each indicated by the position of the triangle in the data graph. As indicated, the level L_2 of the tonal masker was the independent variable and its frequency f_2 was the dependent variable. The subject adjusted f_2 to pulsation threshold (the lowest value of f_2 below the test-tone frequency and, respectively, the highest value above the test-tone frequency for which the pulsating character of the series of test-tone bursts was just not perceived). Seven subjects participated in the experiment, five of them on both ears individually. Of the twelve adjustments thus obtained in each condition, the average and the $\pm$ one sigma interval are plotted.

4.3. DICUSSION

As a typical example of tuning curves measured on primary auditory neurons in cat, Fig. 4.3 reproduces some well-known data of Kiang (1965). It will be seen that the general features of the curves in Fig. 4.2 agree reasonably well with those of the curves in Fig. 4.3. As mentioned before, a similar agreement holds for the results of direct-masking experiments as well. Therefore, the confrontation of Figs. 4.2 and 4.3 is merely meant to indicate that the similarity between psychophysics and electrophysiology does not exist only for the well-known direct-masking technique, but also for other measuring techniques. Since, in this particular case of tone-on-tone masking, the nature of the differences between the results for the various masking techniques is only *quantitative* (magnitude of the slopes or

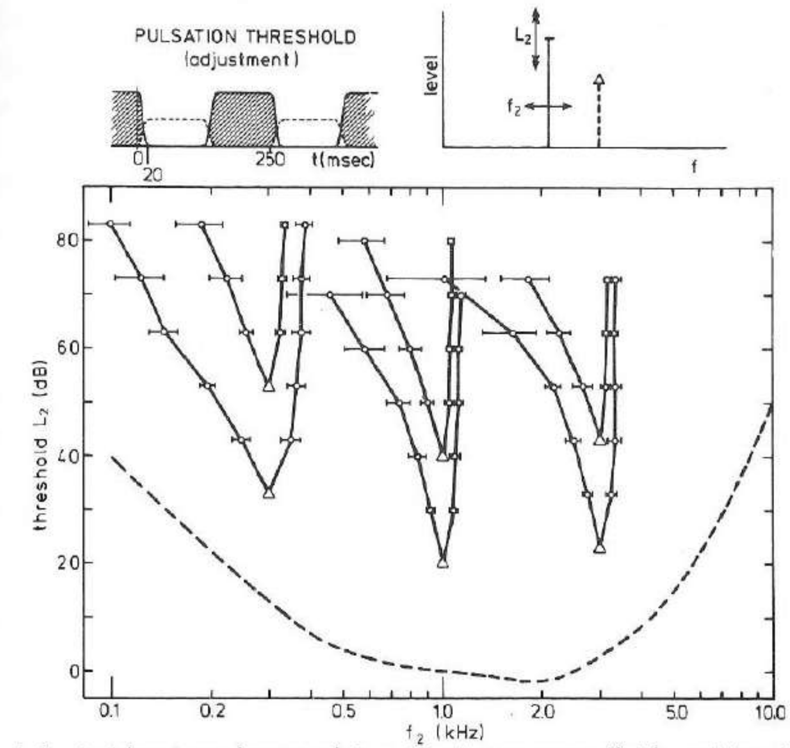


Fig. 4.2. Masking experiments with a fixed test tone (indicated by the positions of the triangles) and a pure-tone masker with variable frequency f_2 and level L_2. (Data of seven subjects.)

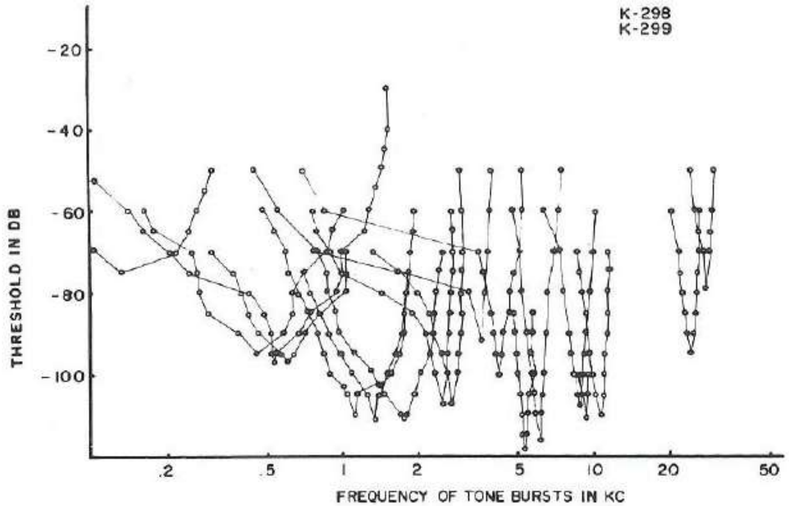


Fig. 4.3. Tuning curves of single auditory-nerve fibres in cat (Kiang, 1965, by permission of the publisher).

width of the curves), any differentiation between the "neural relevance" of the several masking techniques would be based here on the explicit assumption of a *quantitative* agreement between the auditory processes in man and cat. Instead of continuing along this line, it seems more appropriate to base a possible differentiation on *qualitative* differences, as will be shown in the next chapter.

The various masking techniques have a more practical aspect which deserves further attention. The use of a direct-masking paradigm, with the test tone superimposed on the masker, is complicated by possible interactions between masker and test tone, leading to the perception of beats, roughness or combination products which may very well have been involved in the threshold values measured for some direct-masking conditions in Fig. 4.1 (*e.g.*, Egan and Hake, 1950; Greenwood, 1971). Additionally, the condition f_2 = 1000 Hz is a very specific one since the phase relation between masker and test tone has great influence on the measured threshold value. Therefore, different portions of the curves obtained in direct masking may reflect different aspects of auditory signal processing, some of which are perhaps not directly related to the auditory projection of the masker's sound spectrum. These possible confounding implications of the direct-masking technique are excluded when the test tone and the masker are presented *non*simultaneously, as in forward masking or with the pulsation-threshold method. Although the possibility of masker-test-tone interactions in some specific conditions cannot serve as a conclusive argument against the use, in general, of the direct-masking technique for investigating the auditory projection of a masker's sound spectrum, the fact that such confounding interactions can simply be avoided by using other techniques (pulsation threshold, forward masking) favours the latter types as more appropriate.

The conclusion of this chapter is that the frequency-selectivity measures in tone-on-tone masking, as revealed by the various experimental paradigms, are not dramatically different. The frequency selectivity in both direct-masking data (in the literature) and in the present pulsation-threshold data agree reasonably well with the frequency selectivity disclosed by tuning curves of single fibres in the cat's auditory nerve. Therefore, these experiments do not differentiate between the degree of "neural relevance" of the various experimental techniques. However, for a more practical reason, namely to avoid possible confounding interactions between masker and test tone (beats, combination products), methods which make use of a nonsimultaneous test tone (pulsation threshold, forward masking) may be considered preferable.

CHAPTER 5

TWO-TONE MASKERS AND "TWO-TONE INHIBITION"

Summary

Several masking techniques are applied to a two-tone masker with a test tone centred at the weaker of these two tones. In contrast to the traditional direct-masking technique, the results obtained with forward masking and with the pulsation threshold agree well with the phenomenon of "two-tone inhibition" revealed by single-nerve studies.

5.1. FIVE MEASURING TECHNIQUES

This experiment is illustrated in Fig. 5.1. The two masker components are labeled as 1 and 2. The frequency of component 1 was always fixed at 1000 Hz, which was also the frequency of the test tone. (Masker component 1 and the test tone were presented in phase.) Two conditions of the masker were considered, corresponding to the two upper diagrams of Fig. 5.1. First, as a kind of reference condition, only component 1 was presented as the masker. The level L_1 of this 1000-Hz masker was the independent variable (settings from 20 to 60 dB) and the level L_T of the 1000-Hz test tone was the dependent variable. Second, the level of component 1 was fixed at 40 dB and the masker component 2 was added at a level of 60 dB. Now, the frequency f_2 of component 2 was the independent variable (settings from 1000 to 2500 Hz) and, again, L_T was the dependent variable. (For the condition f_2 = 1000 Hz this component was presented in phase with the 1000-Hz test tone.) Five measuring techniques were used.

a. Direct masking, Békésy up-down (long test tone)

The masker was presented continuously. Test-tone bursts with a duration of 300 msec and rise and fall times of 150 msec were repeated every 600 msec. The level of the test-tone bursts was varied automatically, with a 2-dB step after each test-tone burst. The direction of these 2-dB steps, up or down, was controlled by the subject who was instructed to push a button (2-dB steps up) when he did not hear the test-tone bursts, and to release the button (2-dB steps down) when he heard them. The average level during 50 test-tone presentations was considered the threshold value. In each condition, four such measurements were performed with two subjects and the average values are presented in Fig. 5.1, first row.

b. Direct masking, 2-AFC up-down (short test-tone)

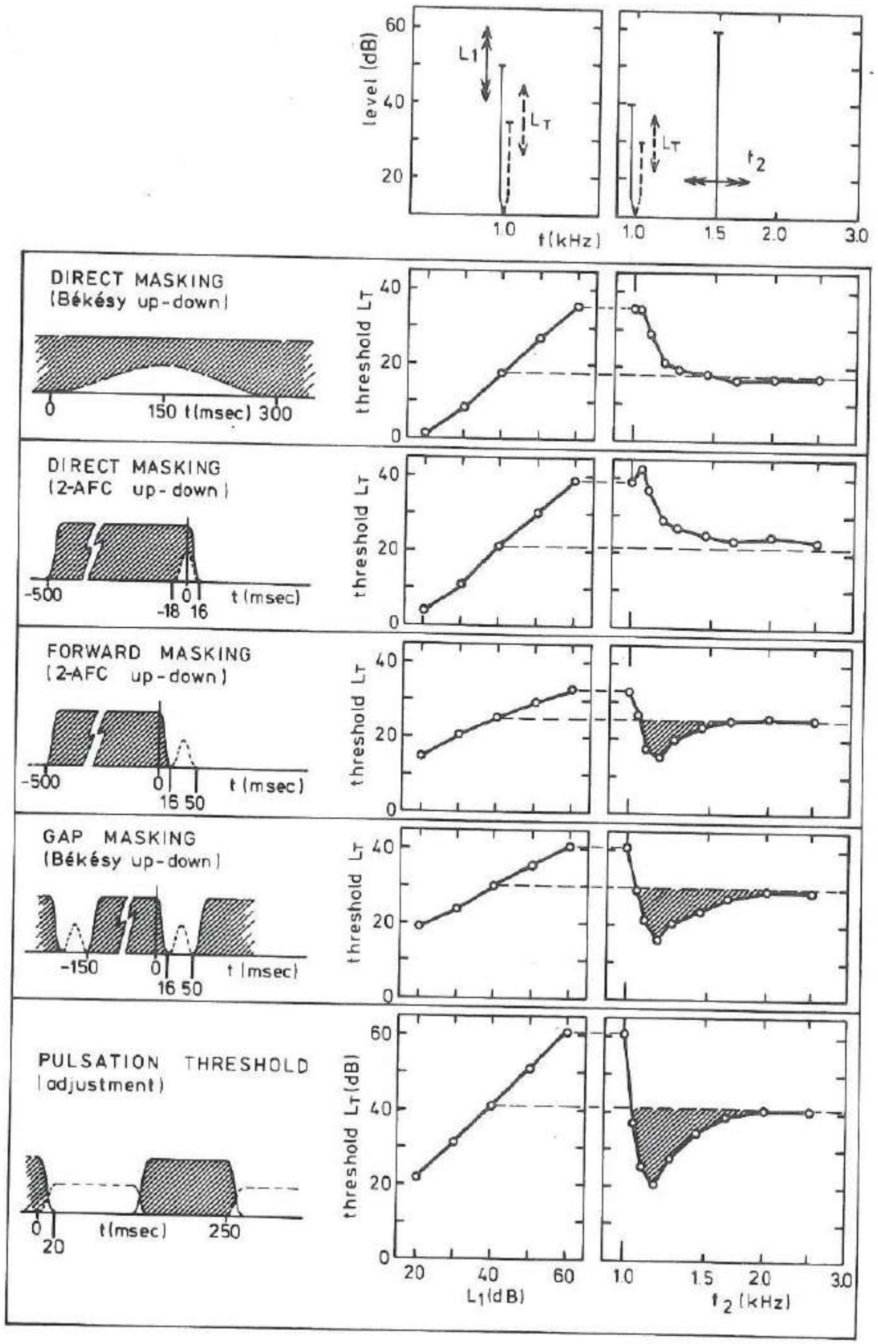
Two observation periods, separated by an interval of 1 sec, were repeated every 3 sec. In both periods, a 500-msec masker burst was presented with 16-msec rise and fall times. At the tail of one of the two masker bursts (randomly) a 34-msec test-tone burst was presented with 17-msec rise and fall times, such that the end of the test-tone burst coincided with the end of the masker burst. The 2-AFC up-down procedure with runs of 50 trials was used to measure the threshold value. For each condition, four such measurements were performed with two subjects and the average values are presented in Fig. 5.1, second row.

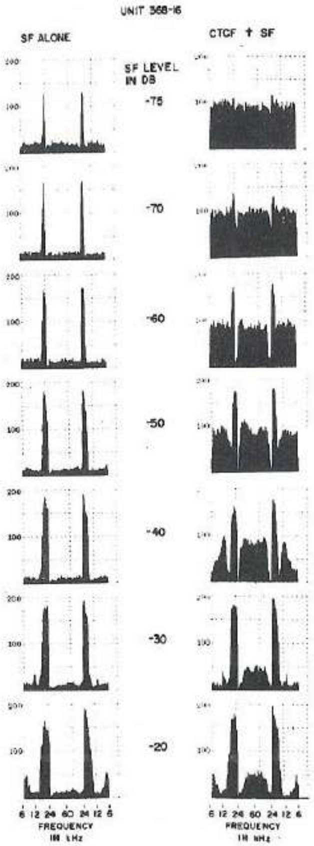
c. Forward masking, 2-AFC up-down

These measurements were identical to the previous ones, except that the test-tone burst was "shifted" over an interval of 34 msec, such that its start, rather than its end, coincided with the end of one of the two masker bursts (Fig. 5.1, third row).

d. Gap masking, Békésy up-down

This technique was used exclusively in this experiment. The masker was gated periodically, 150 msec on and 50 msec off, with 16-msec rise and fall times. During the silent masker gaps, 34-msec test-tone bursts, with 17-msec rise and fall times, were presented according to the following pattern: four gaps *with* test-tone bursts, two gaps *without*, four gaps with, etc. Each series of four successive test-tone bursts, when presented at a level just above detection threshold, is perceived as *one* continuous tone burst. (This is the same phenomenon as the effect of continuity which plays an essential





Left column: histograms of the responses to a sweep-frequency signal (SF) at seven stimulus levels. *Right column:* histograms of the responses to the same SF series added to a continuous tone at the CF of the fiber. Each histogram was obtained from the responses to two sweeps. Sweep rate: one decade per 36 sec. CTCF level: −75 dB; CF: 22.2 kHz; reference level: 200 V p-p into condenser earphone. 180 bins. Bin width: 400 msec.

Fig. 5.2. Example of tone-on-tone inhibition in auditory nerve fibres in cat. (Note the "double" frequency scales.) The data in the right column indicate that the addition of a second frequency component (SF), in particular with a frequency just above the frequency of the fixed frequency component (CTCF), may reduce the fibre's discharge rate below the baseline set by the CTCF component *alone*. Our condition in Fig. 5.1, with the level of the second component 20 dB above that of the first component, would correspond to a SF level of −55 dB in this figure. (Sachs and Kiang, 1968, by permission of The Journal of the Acoustical Society of America.)

Fig. 5.1. (opposite page). Masking experiments on tone-on-tone suppression. Two masker conditions were investigated with a 1000-Hz test tone (upper diagrams): (1) a *single*-tone masker (1000 Hz, variable level L_1) and (2) a *two*-tone masker (1000 Hz, 40 dB *plus* a second tone of variable frequency F_2 at a level of 60 dB). The shaded areas in the three lower data graphs indicate that the addition of the second tone may reduce the masking effectiveness. (Average of two subjects.)

role in the pulsation-threshold technique.) Hence, the pattern of short test-tone bursts (4 on, 2 off, etc.) is *perceived* as a series of single long test-tone burst with a 1.2-sec repetition cycle. The usual Békésy up-down procedure was applied to measure the detection threshold of these test-tone bursts, similar as in case *a.* above. The test-tone level was changed 2 dB after each series of four successive test-tone bursts and the direction of that change, up or down, was controlled by the subject. Again, each data point in Fig. 5.1, fourth row, is the average of four threshold measurements with two subjects.

e. Pulsation threshold, adjustment

The pulsation-threshold method was applied with the continuous alternation of 125-msec masker bursts and 125-msec test-tone bursts of which, again, each fourth test-tone burst was left out. The subject adjusted L_T to pulsation threshold: the highest value of L_T for which the pulsating character of each series of three successive test-tone burst was just not perceived. (The reference condition, with a single-tone masker of the same frequency as the test tone, is rather trivial in this case; since these measurements essentially refer to the amplitude-modulation threshold, the adjustments are about 2 dB above the level of the masker.) Again, the average of four adjustments of two subjects are presented in Fig. 5.1, fifth row.

The left column of data graphs in Fig. 5.1 indicates the "sensitivity" of the various masking techniques in reflecting level changes of a tonal masker centred at the test-tone frequency. The steeper the slope of the curve, the more "sensitive" the method is in reflecting level changes of the masker. (Of course, an additional aspect in relation to this "sensitivity" is the variability in the data points which is not considered here.) The pulsation-threshold method has a relatively high degree of sensitivity and, additionally, it has the attractive feature that the measured threshold value directly represents, within a few dB's, the actual level of the tonal masker.

The right column of data graphs in Fig. 5.1 indicates the effect of the addition of a variable 60-dB tone to a fixed 1000-Hz 40-dB tone on the threshold value of a 1000-Hz test tone. This effect can be estimated by a comparison of the curve in each graph with the horizontal dashed line, which represents the condition when the 60-dB tone is absent. There appears an interesting qualitative difference among the various masking techniques: whereas for the last three methods the threshold values in case of the *two*-tone masker

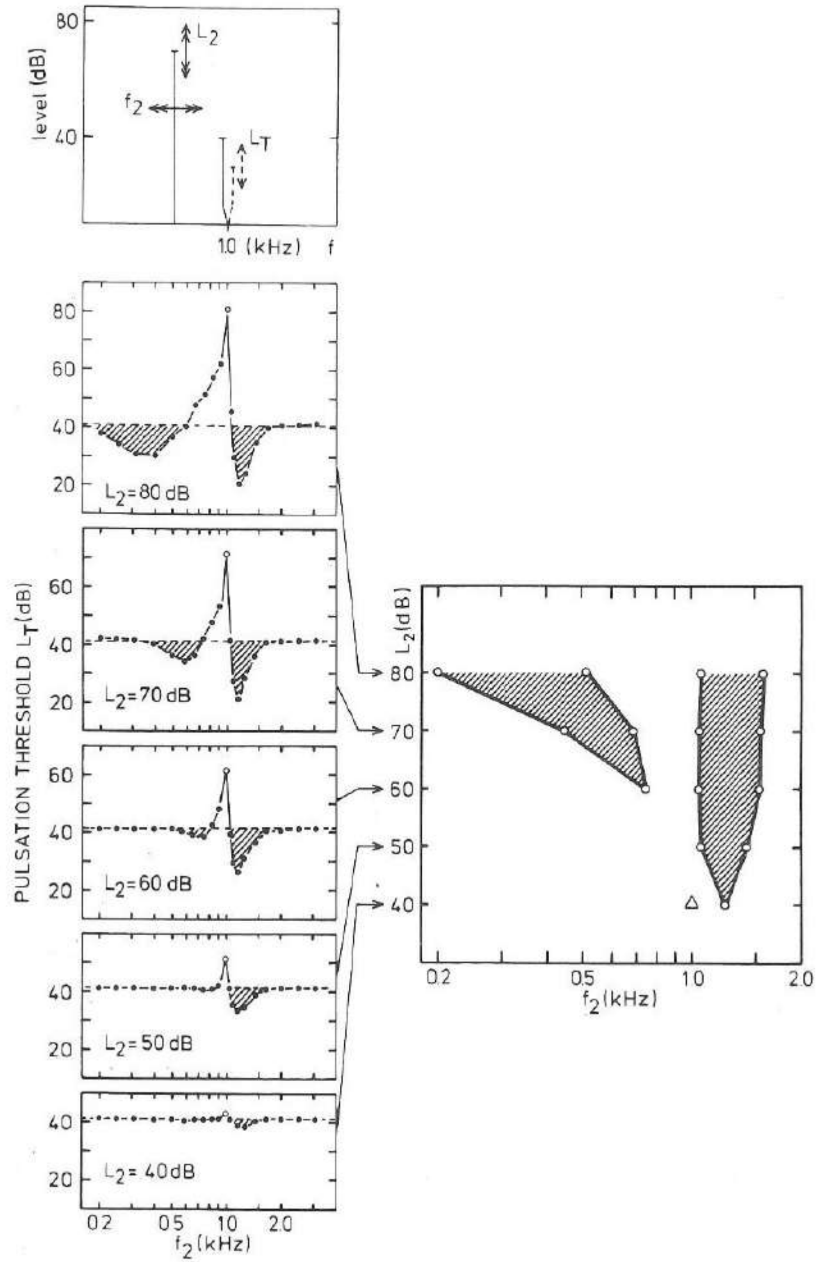
may drop considerably below the baseline set by the *single*-tone masker alone (hatched areas), this effect does not occur for the two direct-masking techniques. Thus, only for the last three methods, the *addition* of an appropriate strong tone to a first weak tone causes a *decrease* of the threshold for the test tone centred at the weak tone. Can this be related to electrophysiological data on the neural coding of two simultaneous pure tones? Effects of "two-tone inhibition" have been demonstrated by Nomoto *et al.* (1964), Sachs and Kiang (1968), Liff and Goldstein (1970) and Arthur *et al.* (1971). For example, Fig. 5.2 reproduces data obtained on primary fibres in the cat's auditory nerve (from Sachs and Kiang, 1968). A fixed weak tone is centred at the "best frequency" of a primary auditory neuron and set at a level which produces a baseline activity in the neuron well above spontaneous activity. When a second tone is added, the response may drop considerably below the baseline, set by the first tone alone, when the frequency of the second tone is just above or (only at higher levels) just below the frequency of the first tone. If we accept a qualitative correspondence between auditory processing in man and in cat, the comparison of Fig. 5.1 and Fig. 5.2 suggests that the results of the last three masking techniques give an adequate picture of the way in which the two-tone masker is coded neurally.

5.2. EXTENSIVE PULSATION-THRESHOLD MEASUREMENTS

The aim of these more extensive measurements, with more subjects and a wider range of conditions, was to investigate further the correspondance between the results of the pulsation-threshold technique when applied to a two-tone masker , and electrophysiological data on "two-tone inhibition".

The experiment is illustrated in Fig. 5.3. Again, the masker consisted of two tones of which one was fixed at 40 dB and 1000 Hz. The second tone was varied both in level (settings from 40 to 80 dB) and in frequency (from 200 Hz to 3100 Hz). The level of the 1000-Hz test tone was the dependent variable. The usual pulsation-threshold method was applied. Seven subjects participated, five of them on both individual ears. The average values of the twelve adjustments are plotted in the data graphs at the left side in Fig. 5.3.

The graphs indicate, again, that the addition of the second masker component to the first one may cause a considerable drop of the pulsation-threshold level below the value in case of the first masker component alone (hatched areas below the baseline at 41 dB). The f_2-values, at which the drop below the



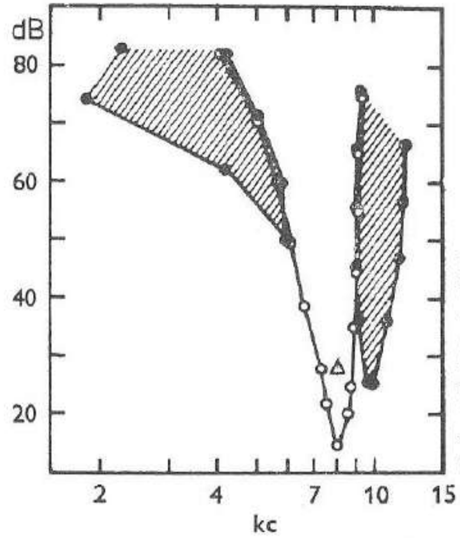


Fig. 5.4. Example of excitatory and inhibitory areas (hatched) of a primary-auditory neuron in cat. The response to a first frequency component at the best frequency (triangle) *plus* a second component within the hatched area is more than 20% *below* the response to the first one alone. (From Arthur *et al.*, 1971, by permission of the Journal of Physiology.)

baseline just amounts to 3 dB, are plotted, for each of the L_2-values used, in one common graph at the right side in Fig. 5.3. Thus, the hatched areas in this f_2 - L_2 plane indicate for what second components the drop is 3 dB or more. In other words: the pulsation threshold for a 1000-Hz test tone in case of a first masker component at 1000 Hz and 40 dB (triangle) *plus* a second masker component within the hatched area is more than 3 dB *below* the pulsation threshold in case of the first masker component alone.

In order to compare this psychophysical result with electrophysiological data on "two-tone inhibition", Fig. 5.4 reproduces data of Arthur *et al.* (1971). The hatched areas indicate the "inhibitory areas": the response to a first frequency component at the best frequency of the primary auditory neuron (triangle) *plus* a second component within the hatched area is more than 20% *below* the response to the first one alone. Accepting a qualitative similarity between auditory processing in man and in cat, the correspondence

Fig. 5.3. (opposite page). Pulsation-threshold measurements on tone-on-tone suppression. The masker consists of a fixed tone at 1000 Hz and 40 dB *plus* a second tone with variable frequency f_2 and level L_2. The pulsation threshold L_T for a 1000-Hz test tone is measured as a function of f_2 (abscissa) and L_2 (parameter). The hatched areas in the figure at the right side indicate the region of f_2 - L_2 values for which the *addition* of the second tone causes a *reduction* of the pulsation threshold by more than 3 dB. (Average of seven subjects).

between Fig. 5.3 and Fig. 5.4 strongly suggests that the pulsation-threshold method correctly reflects the neural response to a two-tone masker in the frequency region of the weaker tone.

The conclusion of this chapter is that various masking techniques, when applied to a two-tone masker, reveal a qualitative difference. Only those methods which use a *non*simultaneous test tone show an effect which corresponds to the well-established phenomenon of "two-tone inhibition" in auditory neurons. Extensive experiments with the pulsation-threshold paradigm revealed that this correspondence holds for a wide range of conditions of a "strong" component added to a fixed "weak" component. These results suggest that nonsimultaneous masking techniques, as the pulsation-threshold method, disclose an adequate picture of the auditory projection of a masker's sound spectrum.

CHAPTER 6

INTERIM DISCUSSION: DIRECT MASKING VERSUS NONSIMULTANEOUS MASKING

Summary

The data presented so far indicate that the results of masking experiments with the test tone superimposed on the masker (direct masking) are systematically different from those for a test tone presented nonsimultaneously with the masker (forward masking, pulsation threshold); only the latter type reveals effects typical of lateral suppression. A comparison with neural data favours *non*simultaneous masking as an adequate method for disclosing the auditory projection of a masker's sound spectrum. In the next chapter, we shall consider more complex stimuli, and the main questions are: (1) does the pulsation-threshold method (and also the forward-masking method) consistently reveal effects of lateral suppression, (2) are, in this respect, the results obtained in direct masking systematically different and (3) if so, does the comparison with available neural data for such stimuli further underline the relevance of the pulsation-threshold method?

The purpose of this discussion is to emphasize some of the main features of the experimental results presented so far, in particular concerning the differences between the results obtained with the various experimental methods. This may serve as a basis for designing further experiments. At a later stage (in chapter 10), I will speculate about possible mechanisms which may *underly* the differences among the several experimental methods.

There are three main experiments which allow a comparison between the various types of masker-and-test-tone experiments: Fig. 3.1 for a masker consisting of sharply low-pass noise, Fig. 4.1 for a pure-tone masker, and Fig. 5.1 for a two-tone masker. There appear important differences, either qualitative, in the case of "Mach bands" and "two-tone suppression", or quantitative in the case of "tuning curves". These differences are associated with the way in which the masker and the test tone are presented temporally: *di-*

rect masking versus *nonsimultaneous* masking. This finding raises the question anticipated in the introduction. The various masker-and-test-tone experiments were performed here explicitly to study the effect of auditory processing (filtering and, possibly, lateral suppression) on a masker's sound spectrum; however, different experimental methods provide different results, and there is no obvious a priori reason why one method would be more relevant than another method. Insight concerning the relevance of the various methods, direct versus nonsimultaneous masking, can be gained by a comparison with electrophysiological data. As mentioned before, two-tone inhibition is a well-documented phenomenon, found throughout the auditory nervous pathway. Of course, we should be aware of the danger in ascribing too much weight to a correspondence between electrophysiological data obtained in animals and psychophysical data from men. Still, a comparison between Figs. 4.2 and 4.3, and also between Figs. 5.3 and 5.4, strongly suggests that the results of nonsimultaneous-masking techniques are closely related to the way in which a masker's sound spectrum is coded neurally.

Summing up, the present situation is the following. Our primary interest is to study what we called "the auditory projection of a masker's sound spectrum", with special interest in possible effects of lateral suppression. In applying various methods of the masker-and-test-tone type, we touched upon a systematic difference between direct masking and nonsimultaneous masking: only in the latter case the results strongly suggest the operation of a lateral-suppression mechanism. Furthermore, such results appear to agree with electrophysiological data from auditory-nerve studies. The general idea underlying the next experiments, with other types of maskers, is to expand further on these findings. Thus, in the following experiments both direct masking and nonsimultaneous masking will be applied to maskers of which the sound spectra are interesting from the point of view of lateral suppression. Again, if results of electrophysiological studies with the same type of stimuli are available, these are helpful in interpreting the psychophysical masking data.

Within the class of nonsimultaneous-masking techniques, I will concentrate on the pulsation-threshold method because this method is relatively new (see also section 10.3.1) and, in the light of the previous results, appears to be promising. However, it might be considered a drawback that the determination of a pulsation threshold cannot be fitted into a more "objective" procedure, such as the two-alternative forced-choice procedure (2-AFC). I realize that many psychoacousticians may have their reservations in accepting results obtained with an adjustment procedure as "hard facts" in the sense of more

common psychophysical detection-threshold data obtained with the well-accepted 2-AFC procedure. Therefore, in order to meet such possible reservations, some main conditions will be investigated also with the forward-masking technique, which is based on the detection threshold of the test tone and hence, allows the application of a 2-AFC procedure. It has been shown already that this method has a drawback as well, namely its lack of "sensitivity" (*e.g.*, Fig. 5.1 shows that a change of the level of the masker is reflected only partly in the change of the forward-masked threshold level). Thus, the important question is whether the qualitative similarity between pulsation-threshold data and forward-masking data, as revealed by the previous experiments, also holds for other types of masker. If so, this would further support the view that data obtained with the pulsation-threshold method, despite the use of an adjustment procedure, are not exclusive, but that they belong to a class of methods of which the essential feature is that masker and test tone are presented nonsimultaneously.

CHAPTER 7

TONE-PLUS-NOISE MASKERS, AND THE EFFECT OF NOISE-ON-TONE SUPPRESSION

Summary

Masking experiments are performed with a 1000-Hz test tone and a masker consisting of a 1000-Hz tone plus white noise at a variable level. Results obtained with nonsimultaneous-masking techniques (pulsation threshold, forward masking) indicate that the addition of noise may *reduce* the masking effectiveness of the 1000-Hz masker. This effect, which is typical of the operation of a lateral-suppression mechanism, is not found in direct masking. It appears that the effect corresponds quantitatively to the well-known effect of noise on the *loudness* of a tone.

7.1. THREE MEASURING TECHNIQUES

The experiment is illustrated in Fig. 7.1. The tone in the tone-plus-noise masker was fixed at 1000-Hz. The test tone, also with a frequency of 1000 Hz, was presented in phase with the tonal masker. Two conditions of the masker were considered, corresponding to the two diagrams at the top of Fig. 7.1. First, as a reference condition, only the tone, without the noise, was presented as the masker. The level L_1 of this 1000-Hz masker was the independent variable (settings from 20 to 60 dB) and the level L_T of the 1000-Hz test tone was the dependent variable. Second, white noise was added to the tonal masker of which the level was fixed at 40 dB. The spectral level of the noise, N_0, was the independent variable (settings from -14 to 46 dB/Hz) and, again, L_T was the dependent variable. Additionally, for a few values of N_0, the case of the noise masker *alone*, without the 40-dB tone, was also considered. Three measuring techniques were used.

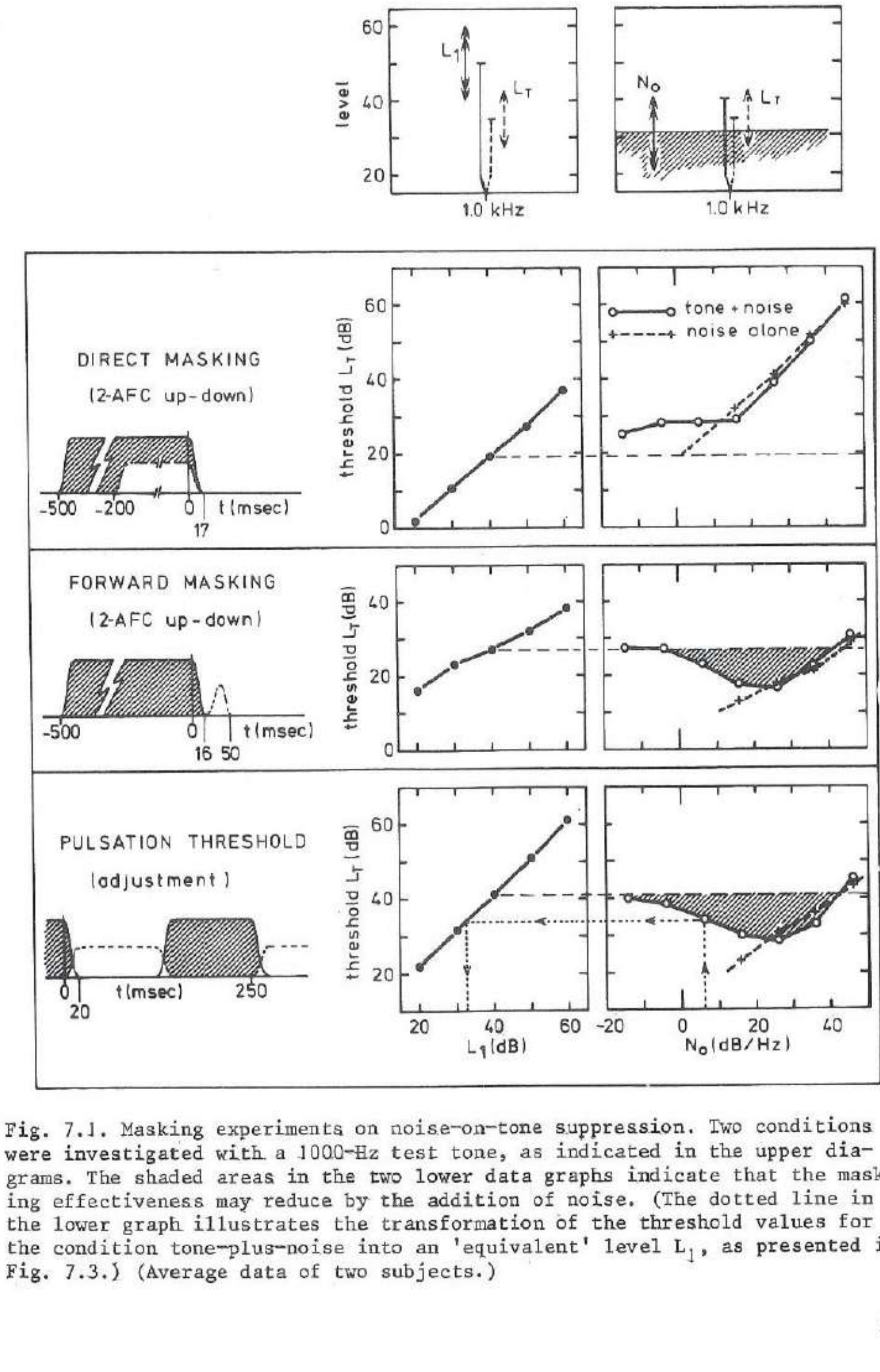


Fig. 7.1. Masking experiments on noise-on-tone suppression. Two conditions were investigated with a 1000-Hz test tone, as indicated in the upper diagrams. The shaded areas in the two lower data graphs indicate that the masking effectiveness may reduce by the addition of noise. (The dotted line in the lower graph illustrates the transformation of the threshold values for the condition tone-plus-noise into an 'equivalent' level L_1, as presented in Fig. 7.3.) (Average data of two subjects.)

a. Direct masking, 2-AFC up-down

Two observation periods, separated by an interval of 1 sec, were repeated every 3 sec. In both periods, a 500-msec masker burst was presented with 17-msec rise and fall times. At the end of one of the two masker bursts (randomly) a 200-msec test-tone burst was presented with 17-msec rise and fall times, such that the end of the test tone burst coincided with the end of the masker burst. The 2-AFC up-down procedure with runs of 50 trials was used to measure the threshold value. In each condition, three measurements were performed with two subjects and the averages are presented in Fig. 7.1, first row.

b. Forward masking, 2-AFC up-down

This measurement was identical to the previous one, except that a short (34 msec) test-tone burst was used which immediately followed one of the masker bursts (Fig. 7.1, second row).

c. Pulsation threshold, adjustment

The pulsation-threshold method was applied with the continuous alternation of 125-msec masker bursts and 125-msec test-tone bursts with, again, each fourth test-tone burst left out. The averages of three adjustments by two subjects are presented in Fig. 7.1, third row.

As in a previous experiment (Fig. 5.1), the left column of data graphs in Fig. 7.1 indicates the "sensitivity" of the various masking techniques to level changes of the tonal masker. The dashed horizontal line, based on the threshold measured for a 40-dB tonal masker alone, serves as a reference line for the data graphs in the right column. The orientation of the data points for tone-*plus*-noise (circles) relative to this reference line (tone alone) directly indicates the effect of the addition of the noise to the 40-dB tonal masker. It will be seen that both in forward masking and for pulsation threshold the *addition* of the noise to the 1000-Hz masker may *reduce* the masking effectiveness at 1000 Hz (hatched areas). This is not the case in direct masking; here the addition of noise only causes an increase of the masked threshold. (This increase is found even for relatively low noise levels, which illustrates the well-known effect that the increment threshold for a pure tone is very sensitive to the addition of noise; see, for instance, Zwicker, 1956.) Finally, in all three cases, the addition of high-level noise gives about the same threshold values as for the condition "noise alone".

As in previous experiments, there appears an essential difference between

the results of the two types of experimental techniques: direct masking versus nonsimultaneous masking. Only experiments of the latter type reveal an effect which strongly suggests the operation of a lateral-suppression mechanism. This difference can be compared with electrophysiological data on the response of single fibres in the cat's auditory nerve to tone-plus-noise stimuli, in which broad-band noise of variable level is added to a fixed tone at the fibre's best frequency (Kiang, 1965). A typical result is that the response to a tone-*plus*-noise stimulus is *less* than the response to the tone alone.

7.2. NOISE-INDUCED LOUDNESS REDUCTION

A well-known effect of the addition of noise to a tone is the reduction of the *loudness* of the tone (*e.g.*, Zwicker, 1963; Scharf, 1964). A short loudness-matching experiment was performed, with the same subjects and the same tone-noise combinations as in the previous experiment in order to investigate whether there exists a quantitative agreement between this phenomenon

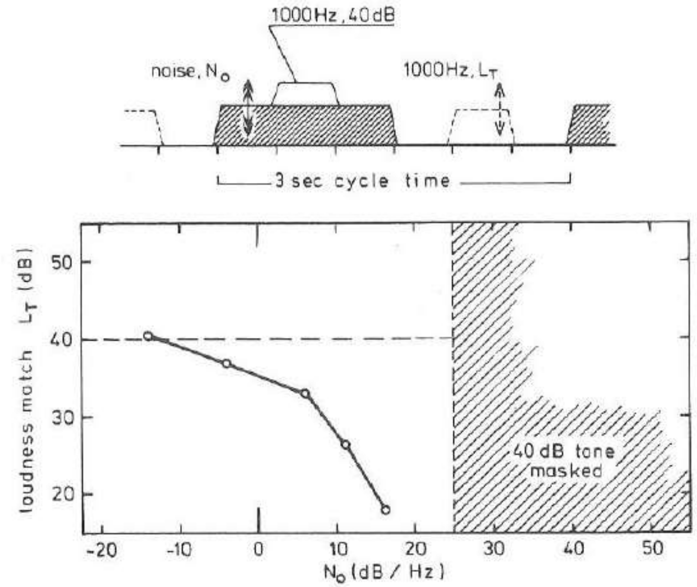


Fig. 7.2. The effect of noise on the loudness of a tone. The 1000-Hz tone bursts with a fixed level of 40 dB, superimposed on white noise of variable level N_0 (abscissa), sound equally loud as the isolated tone bursts with level L_T (ordinate). (Average data of two subjects.)

and the noise-induced reduction of the *masking* effectiveness of a tone, as found in the previous section.

The experiment is illustrated in Fig. 7.2. White-noise bursts, with on and off times of 1.5 sec, were presented continuously. Both during the noise bursts and in the silent intervals, 1000-Hz tone bursts of 0.5 sec duration were presented. The subject was instructed to adjust the level L_T of the tone bursts in the silent intervals such that they sounded equally loud as the 40-dB tone bursts superimposed on the noise. The median value of six adjustments was taken. The data points in Fig. 7.2 are the averages of two subjects.

In order to facilitate a comparison of these data with the results of the various masking experiments, the data should be plotted along a common axis. Such a graph is presented in Fig. 7.3. The abscissa gives, as in the previous graphs, the spectral level of the noise added to the 1000-Hz, 40-dB tone. The ordinate presents, rather than the actual threshold values measured, the level of a 1000-Hz masker *alone* which would cause that same threshold value. Thus, the effect of the addition of the noise to the fixed 1000-Hz masker is "translated" into an "equivalent level" of the 1000-Hz masker when presented alone. (These transformations of the actual threshold values are performed as indicated by the example in the lower panel in Fig. 7.1.) The results of the loudness-matching experiment are also reproduced in Fig. 7.3.

A most interesting aspect of Fig. 7.3 is the similarity, over a considerable range of noise levels, between the equal-loudness curve (a) and the two curves based on nonsimultaneous masking (b) and (c). Of course, this similarity is lost for high levels since the curves based on masking reflect the masking produced by tone *plus* noise, whereas the equal-loudness curve reflects the loudness of the tone *itself* in the noise. Given this restriction, it appears that the effect of noise on the loudness of a tone is closely related to its effect on the masking effectiveness of the tone in nonsimultaneous masking. If the latter is understood as reflecting an effect of lateral suppression of the tone by the noise, it is not unreasonable to assume that the same mechanism plays a part in the loudness reduction of the tone as well. The consequences of this assumption with respect to current theories on noise-induced loudness reduction (or "partial masking") are discussed elsewhere (Houtgast, 1974).

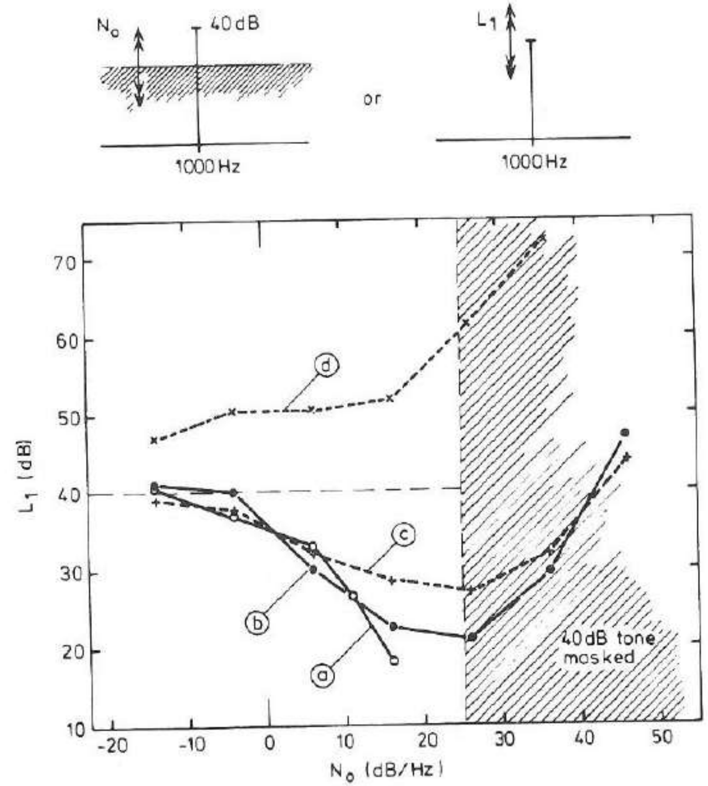


Fig. 7.3. The curves indicate the relation between N_o and L_1 for which the two stimuli, presented schematically in the top diagram (tone-plus-noise and tone alone), are equivalent with respect to:
(a) the loudness of the 1000-Hz tone,
(b) the forward-masked threshold of a 1000-Hz test tone,
(c) the pulsation threshold of a 1000-Hz test tone,
(d) the direct-masked threshold of a 1000-Hz test tone.
(Data adapted from Figs. 7.1 and 7.2.)

7.3. TONE PLUS BAND-REJECTED NOISE

In the previous sections, we considered the effect of adding uniform (white) noise to a 1000-Hz tone. In view of a possible interpretation of the results in terms of lateral suppression, it is of interest to investigate the spectral region over which suppression is found. For that purpose band-rejected noise was used, with a spectral gap of variable width centred around the tonal masker.

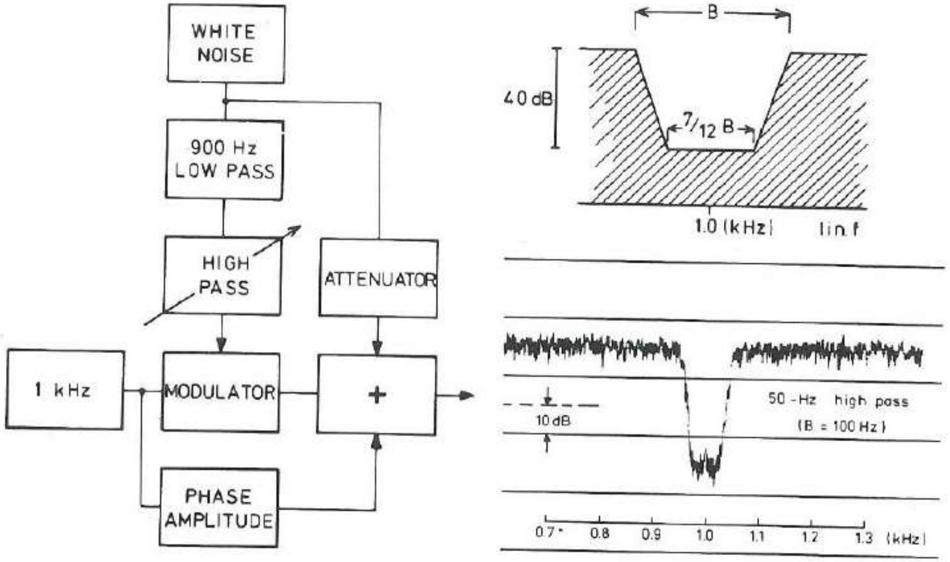


Fig. 7.4. Band-rejected noise, used in the experiment of Fig. 7.5, was obtained by modulation of a 1000-Hz carrier with white noise filtered by a 900-Hz low-pass filter and a variable high-pass filter (slopes 48 dB/oct). After cancellation of the carrier and addition of low-level noise, a spectrum was obtained as indicated in the diagrams.

Fig. 7.4 indicates that the band-rejected noise was obtained by carrier-suppressed modulation of a 1000-Hz component with a noise band, determined by a fixed low-pass filter (900 Hz, 48 dB/oct) and a variable high-pass filter (0-500 Hz, 48 dB/oct). By this procedure, the noise band was modulated "upwards" over a distance of 1000 Hz and appeared as two side bands around the (suppressed) 1000-Hz carrier. The resulting spectrum was flat between 100 and 1900 Hz , with a sharp gap around 1000 Hz. The -3 dB width of the gap was equal to two times the setting of the variable high-pass filter; its -48 dB width was half its -3 dB width. By adding extra low-level noise, the "depth" of the gap was fixed at -40 dB. The noise added to the 1000-Hz masker was thus defined by two parameters: its spectral level N_0 in the two pass bands and the -3 dB width B of the spectral gap. (For the example in Fig. 7.4, the high-pass filter was set at 50 Hz; the spectrum was measured with a Rohde & Schwarz wave analysér type FAT 3, bandwidth 4 Hz and speed 10 min/1000 Hz.)

The first experiment is illustrated in Fig. 7.5. In this case, the spectral level of the noise in the two side bands was kept constant (N_0 = 26 dB/Hz).

The independent variables were the level L_1 of the tone in the noise (settings from 10 to 70 dB) and the new parameter B (settings from 0 to 1000 Hz). The test-tone level L_T was the dependent variable. Only the pulsation-threshold method was used (see Fig. 7.1). Thus, the masker (noise plus 1000-Hz tone) and the 1000-Hz test-tone were presented alternately and the subject adjusted the test-tone level L_T to the highest value at which the 1000-Hz tone sounded continuously. In each condition, three such adjustments were made by two subjects and the averages are presented in Fig. 7.5.

The second experiment is illustrated in Fig. 7.6. Here, for one value of B (200 Hz), the effect of the higher or the lower noise band was investigated separately. (This experiment was performed at 1100 Hz instead of 1000 Hz because of the availability of very steep low- and high-pass filters, needed for the separation of the two noise bands, with cutoff frequencies of 1100

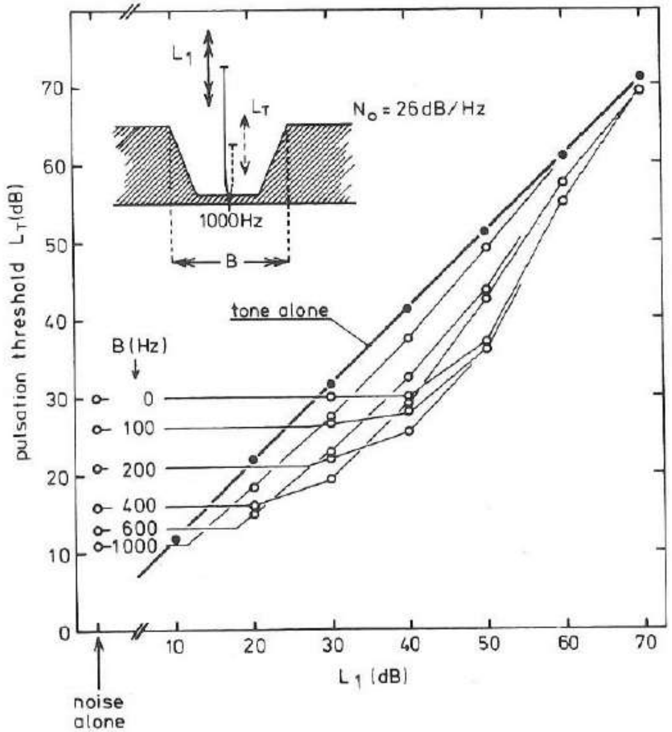


Fig. 7.5. Pulsation-threshold measurements with a 1000-Hz test tone and a masker which consisted of a 1000-Hz tone with variable level L_1 *plus* band-rejected noise with variable gapwidth B. All data points which fall below the curve 'tone alone' indicate that the noise suppresses the masking effectiveness of the tone. (Average data of two subjects.)

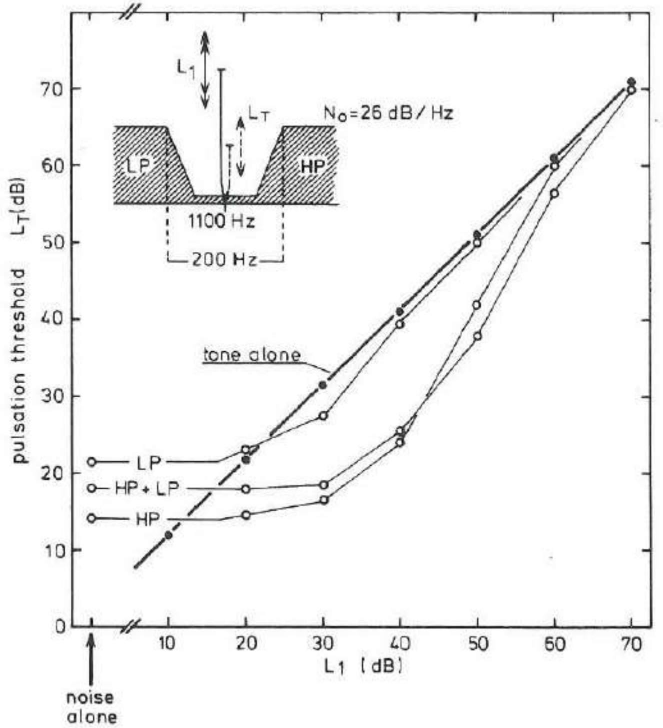


Fig. 7.6. This is an extension of the experiment in Fig. 7.5. For one value of the width of the spectral gap (200 Hz), the effect of the higher or the lower noise band was investigated separately. (Average data of two subjects.)

Hz.) For the rest, the procedure was identical to that of the previous experiment.

The interesting aspect in Figs. 7.5 and 7.6 is the effect of the *addition* of the noise to the tonal masker. Thus, in each figure, the family of curves should be related to the curve indicated as "tone alone". It appears that, for a wide range of conditions, the effect of adding the noise is such that the pulsation threshold *reduces*, but never below the value obtained for the condition "noise alone". Fig. 7.5 shows that, given this limit set by the "noise alone", the reducing effect of the noise is essentially the same for values of B from 0 to 200 Hz. For B-values beyond 200 Hz, the reducing effect of the noise becomes progressively smaller. In terms of lateral suppression, this would indicate that the suppression region around a 1000-Hz tone is concentrated at a "distance" of about 100-200 Hz. Additionally, Fig. 7.6 indicates

that such a suppression region is located mainly at the high-frequency side of the tone.

The conclusion of this chapter is that in all respects the present results of noise-on-tone suppression agree with previous results of tone-on-tone suppression: (1) effects of suppression are found only with *non*simultaneous masking techniques, both in forward masking and with the pulsation threshold method, and (2) the location of the "suppression region" around a 1000-Hz tone agrees well with the more precisely estimated "suppression areas" in case of tone-on-tone suppression (Fig. 5.3).

MASKING EXPERIMENTS WITH NOISE OF VARIABLE BANDWIDTH, AND THE "CRITICAL BAND" CONCEPT

Summary

Masking experiments are performed with a 1000-Hz test tone and a masker consisting of a band of noise of variable width centred at 1000 Hz. In contrast to traditional direct-masking data, the results for pulsation threshold indicate that an *increase* of the masker's bandwidth may result in a *reduction* of the masking effectiveness in the centre of the noise band. This result is typical of the operation of lateral suppression and is compared with an electrophysiological study on neural responses to noise bands of variable width.

8.1. THE MASKER

Fig. 8.1 illustrates that the masking noise was obtained by the same technique of carrier-suppressed modulation as applied in the previous chapter. In this case, a 1000-Hz carrier was modulated with low-pass filtered noise (variable cutoff frequency, 48 dB/oct). After cancellation of the carrier, a band of noise is obtained with a constant spectral level within the range 1000 Hz ± the cutoff frequency of the low-pass filter; beyond that range the spectral level falls off sharply and reaches -48 dB at the positions 1000 Hz ± *two* times the cutoff frequency of the low-pass filter.

The band of noise is defined by its width B and the spectral level N_0 within the pass band; both were varied in the experiment. The example in Fig. 8.1 refers to the case of 50-Hz low-pass noise and the resulting spectrum was measured with a Rohde & Schwarz wave analyzer (type FAT 3, bandwidth 4 Hz, speed 10 min/1000 Hz). It can be seen that the spectral shape closely fits the theoretical one, except for the appearance of residual noise at a level

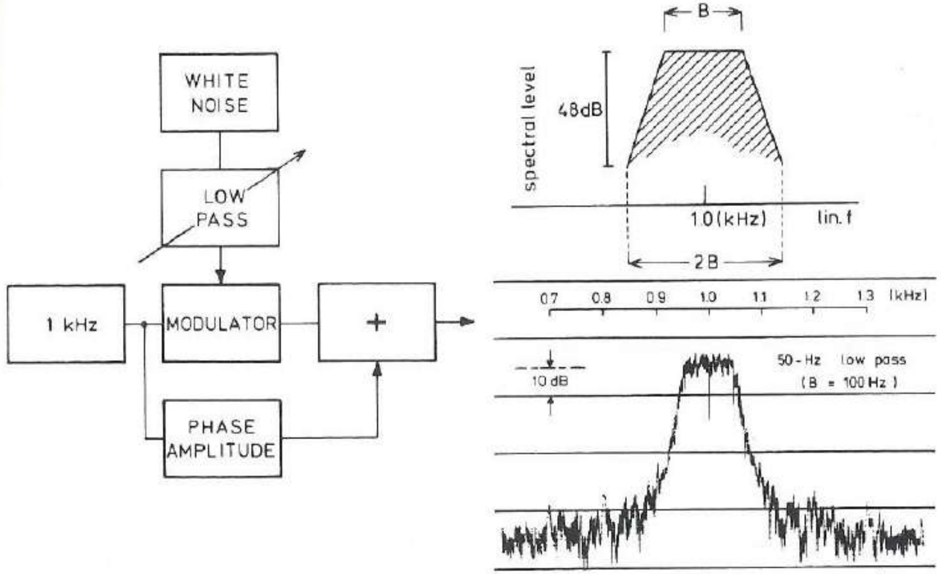


Fig. 8.1. Band-pass noise, used in the experiment of Fig. 8.2, was obtained by modulation of a 1000-Hz carrier with white noise filtered by a variable low-pass filter (48 dB/oct). After cancellation of the carrier, a spectrum was obtained as indicated in the diagrams.

of about 60 dB below the level within the pass band.

8.2. TWO MEASURING TECHNIQUES

The experiment is illustrated in Fig. 8.2. The masking noise was defined by its bandwidth B and its spectral level N_0. Both were treated as independent variables (settings of B from 20 to 2000 Hz and settings of N_0 from 21 to 51 dB/Hz). The level L_T of the 1000-Hz test tone was the dependent variable. As a lower limit, for the condition B = 1 Hz, a 1000-Hz pure-tone masker was used with a level equal to N_0. (In the case of direct masking it is of importance to specify the phase relation between pure-tone masker and test tone; in view of the random-phase relation for the noise masker, we used a 90^0 phase relation for the pure-tone masker.) Two measuring techniques were used.

a. Direct masking, Békésy up-down

The masker was presented continuously. Test-tone bursts with a duration of 500 msec, and 25-msec rise and fall times, were repeated every 1000 msec.

The level of the test-tone bursts was varied automatically, with a 2-dB step after each test-tone burst. The subject was instructed to push a button (2-dB steps up) when he did not hear the test-tone bursts, and to release the button (2-dB steps down) when he did hear them. The average level over a period of 50 test-tone presentations was considered as the threshold value. For each condition, two measurements were performed with two subjects and the average values are presented in Fig. 8.2, upper panel.

b. Pulsation threshold, Békésy up-down

The pulsation-threshold method was used with the continuous alternation of 125-msec masker bursts and 125-msec test-tone bursts with, again, each fourth test-tone burst left out. The level of the test-tone bursts was varied automatically, with a 2-dB step after each series of three test-tone bursts. The subject was instructed to push a button (2-dB steps up) when he perceived each series of three successive test-tone bursts as one *continuous* tone burst, and to release the button (2-dB steps down) when he perceived the series of three successive test-tone bursts as a *pulsating* tone. The average level over a period of 50 presentations of three successive test-tone bursts was considered as the threshold value. For each condition, three measurements were performed with two subjects and the average values are presented in Fig. 8.2, lower panel.

8.3. DISCUSSION

Fig. 8.2 shows several interesting aspects. Our main interest is the effect of bandwidth on the masking effectiveness in the centre of the noise band. The two graphs reveal two different effects. When, starting with a large bandwidth (B = 2000 Hz), the noise band is made progressively narrower, the direct-masking curves only show a *decrease* of the masking effectiveness, whereas the pulsation-threshold curves first show an *increase* of the masking effectiveness, followed by a decrease for small values of B.

The nature of the direct-masking curves is very common. They illustrate the traditional concept that the masking of a tone is determined mainly by noise components within a restricted frequency region around the tone frequency. The curves are not in conflict with the usual estimate of such a "critical bandwidth", in the 1000-Hz frequency region, of about 160 Hz (*e.g.*, Scharf, 1970). However, the pulsation-threshold curves cannot be understood on the basis of a simple critical-band concept. The curves indicate, addi-

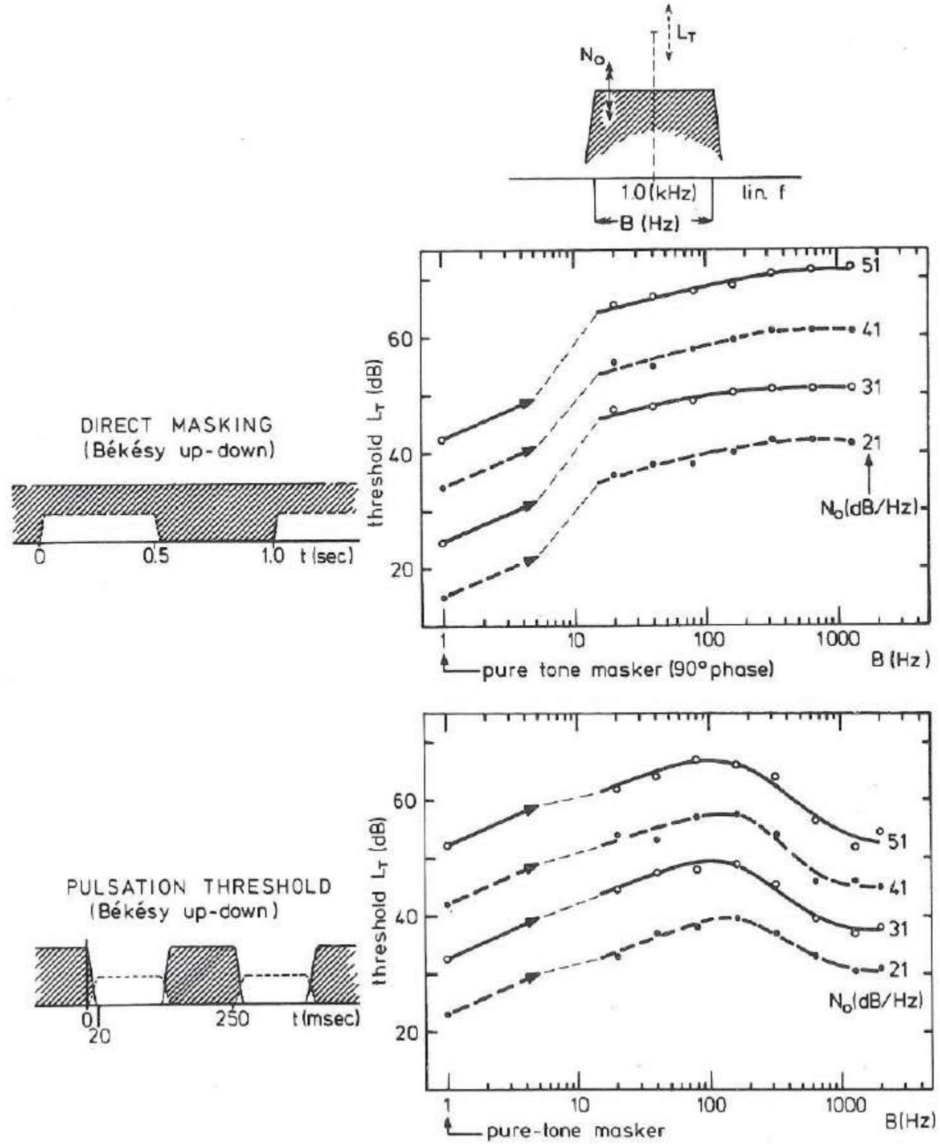


Fig. 8.2. Masking experiments with a 1000-Hz test tone located in the centre of a band of noise with variable spectral level N_o (parameter) and variable bandwidth B (abscissa). The lower graph indicates that the masking effectiveness decreases when the bandwidth is widened beyond a value of about 150 Hz. (average data of two subjects).

tional to the positive contribution of noise components close to the tone frequency, a *negative* contribution to masking effectiveness of more remote noise components. This agrees well with the concept of lateral suppression and it is in line with the results presented in the previous chapters.

There is an electrophysiological study which is closely related to the present psychophysical experiment. Greenwood and Goldberg (1969) performed experiments in which bands of noise were widened around the best frequency of single units in the cochlear nuclear complex of the cat. Two types of effects were observed: (1) "summation", in which increasing bandwidth from narrow width (constant spectrum level) resulted in an increase in firing rate and (2) "suppression", in which increasing bandwidth beyond the range in which summation occurred resulted in a systematic reduction in firing. Though the degree of the summation and suppression effects appeared to vary markedly for different units, the general characteristics of these data are very similar to our results obtained with the pulsation-threshold method. Even their observation that "The bandwidth at which summation ceased and suppression began decreased somewhat at higher spectrum levels" seems to be consistent with a slight shift of the position of the maximum in the pulsation-threshold curves for higher values of N_0. Thus, it would appear that the masking effectiveness in the centre of a noise band, as measured with the pulsation-threshold method, is closely related to the neural coding of such stimuli.

Fig. 8.2 reveals also another aspect which may be related to lateral suppression. This concerns the relation between the threshold level L_T and the spectral level N_0 in the case of the broad-band masker (data for B = 2000 Hz). In direct masking, this relation is traditional: a change in N_0 causes an equal change in L_T. However, in the pulsation-threshold approach, the relation is different: when N_0 is increased by 30 dB (from 21 to 51 dB/Hz), L_T increases only about 22 dB. This is typical of pulsation-threshold data with a broad-band masker (see, for example, the curve "noise alone" in the lower panel of Fig. 7.1). It should be realized that in the trivial case of a pure-tone masker (data points at B = 1 Hz), the pulsation-threshold level exactly follows the level of the pure-tone masker. Thus, for pulsation threshold, a change of the level of a broad-band masker is less effective than a change of the level of a pure-tone masker. This might be related to an observation of Møller (1970) that, for the response of units in the cochlear nucleus of the rat, a change of the level of a broad-band noise stimulus is less effective, (*i.e.*, causes a smaller change in the unit's response) than a change of the level of a pure-tone stimulus of a frequency equal to the unit's characteristic frequency. In a way, one might say that a broad-band stimulus is subjected to its own effects of lateral suppression. This demonstrates once again the similarity between pulsation-threshold data and neural-response data.

Besides aspects related to lateral suppression, Fig. 8.2 reveals an additional point which deserves further attention. The pure-tone masker, representing the limit case for B = 1 Hz, is included in the data graphs merely for the sake of curiosity. It should be realized that it is not in all respects a "true" limit case: the temporal structure of a small band of noise, with its strong envelope fluctuations, is very different from the steady-state character of a pure tone. This aspect is totally neglected in the prediction lines in the data graphs, which are drawn through each pure-tone data point with a slope of +3 dB for each doubling of the masker's bandwidth. Hence, the fact that, in the case of direct masking, these lines do *not* fit the curves referring to "true" bands of noise is not surprising; it means that the masking effectiveness of a narrow-band masker is not determined only by its average, long-term intensity but that the temporal character of the masker, steady-state or fluctuating, is important as well. It is surprising that the prediction lines do reasonably well fit the curves in case of pulsation threshold. This implies that the pulsation threshold in case of narrow-band maskers, a steady-state tone or a fluctuating 20-Hz band of noise, is determined essentially by the average, long-term, intensity of the masker, regardless of its temporal character.

The conclusion of this chapter is that the masking effectiveness in the centre of a band of noise of variable width, when measured with the pulsation-threshold method, strongly suggests the influence of lateral suppression: the masking effectiveness decreases when the bandwidth is increased beyond some critical value. In direct masking, no such effect is found. Furthermore, it was illustrated that the pulsation-threshold data agree well with neural-response data, both with regard to variations of noise bandwidth and of noise level.

CHAPTER 9

MASKING EXPERIMENTS WITH RIPPLED NOISE, AND THE "AUDITORY FILTER"

Summary

The masker is "rippled" noise, with a spectrum (intensity on a linear frequency scale) shaped according to a sinusoidal function. The masking effectiveness depends on the localization of the test tone relative to the peaks and valleys of the ripple, but this dependence decreases when the ripple is made progressively finer (thus, when the peak-to-peak distances along the frequency scale decrease). The results of the masking experiments allow an estimation of the degree of frequency resolution, in terms of a filter, involved in the auditory processing of such rippled-noise spectra. It appears that the degree of frequency resolution estimated from pulsation-threshold data, and also from forward-masking data, is higher than that estimated from direct-masking data. When expressed in terms of bandwidth of an auditory filter, the difference is about a factor two. This difference is interpreted in terms of a lateral-suppression mechanism which, again, manifests itself only in *non*simultaneous masking techniques.

Rippled noise is a stimulus which is particularly suited for investigating the "resolution power" of a system in the frequency domain. Generally speaking, a limited resolution power will manifest itself by a decrease of the peak-to-valley ratio at the output of the system when the ripple spacing is made progressively finer. This is a common approach for investigating resolution power in some other domains: in optics (unsharpness, or resolution in the place domain), and also in perception, one often uses test stimuli of which the intensity is sine-wave shaped along the relevant dimension.

In the present experiment, the rippled noise will be used as a masker. In principle, the difference between the masking effectiveness at the peak and in the valley of the ripple may be used to estimate the extent to which the ripple is "resolved" by the ear. In this case, a typical effect of lateral

suppression would be an increase of "ripple resolution" for specific ripple spacings.

9.1. THE MASKER

The masker was obtained by the well-known technique of the summation of white noise and its delayed repetition (Fig. 9.1). White noise was sampled by an analogue-to-digital converter and regenerated by two digital-to-analogue converters (sample and regeneration time about 70 μsec). The delay and polarity of one of these signals was controlled digitally, whereas its amplitude was controlled by an attenuator. After summation and appropriate low-pass filtering, the resulting power spectrum (intensity on a linear frequency scale) is sine-wave shaped. Four parameters are of importance (see Fig. 9.1): (1) $\bar{I}$, the average intensity, which is the intensity at the half-way points between a peak and a valley, (2) the modulation depth m of the ripple, which is determined by the attenuation, (3) the peak-to-peak distance Δf of the ripple, which is determined by the delay time, and (4) the phase of the ripple at the point f = 0 (a peak or a valley), which is determined by the polar-

$$\text{Int.}(f) = \bar{I}\,(1 \pm m \cos 2\pi \frac{f}{\Delta f})$$

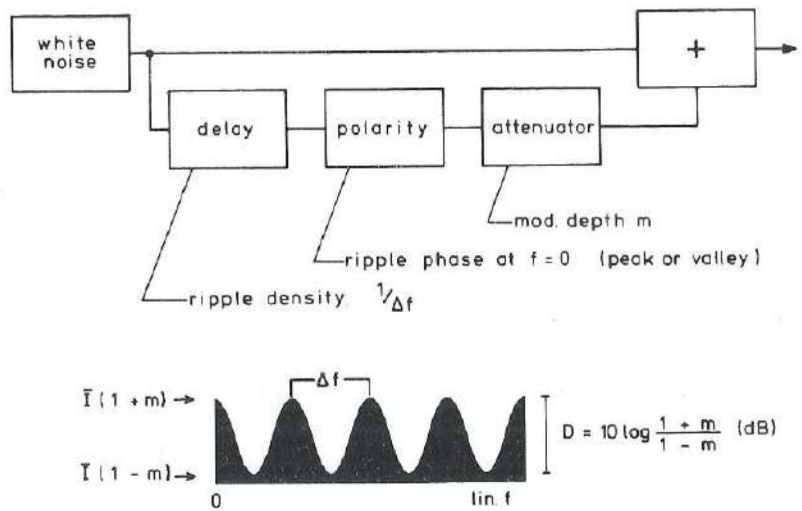


Fig. 9.1. Rippled noise was obtained by summation of white noise and its, delayed, repetition. The ripple density is determined by the *delay time*, the phase of the ripple at f = 0 (peak or valley) by the *polarity* and the peak-to-valley ratio D by the *attenuation* of the added noise.

ity of the delayed noise.

In this chapter, we will use the following notation to characterize a rippled-noise masker: (1) the spectral level $\bar{N}_0$ in dB/Hz at the half-way points between a peak and a valley, (2) the peak-to-valley level ratio D in dB, (3) the *ripple density* $1/\Delta f$ (the number of ripples per Hz bandwidth) or the *relative* ripple density $f_T/\Delta f$ (the number of ripples between f = 0 and the test-tone frequency $f = f_T$), and (4) the phase of the ripple at $f = f_T$. In all cases, the rippled-noise masker was passed through a low-pass filter (48 dB/oct) with a cutoff frequency equal to three times the test-tone frequency.

9.2. FOUR MEASURING TECHNIQUES

The experiment is illustrated in Fig. 9.2. The primary purpose was to investigate possible differences among the results of the various masking techniques. As indicated, the modulation depth D was fixed at 24 dB, the spectral level $\bar{N}_0$ was the dependent variable and the relative ripple density $f_T/\Delta f$ was the independent variable (settings between 1 and 12). A test tone with a fixed frequency (f_T = 1000 Hz) and a fixed level L_T was used. For each of the four masking techniques, the fixed test-tone level L_T was chosen such that the threshold values obtained for $\bar{N}_0$ were of the same order of magnitude in the four cases. Two phases of the ripple at the test-tone frequency 1000 Hz were used: (a) a *peak* centred at 1000 Hz, and (b) a *valley* centred at 1000 Hz. Four measuring techniques were used.

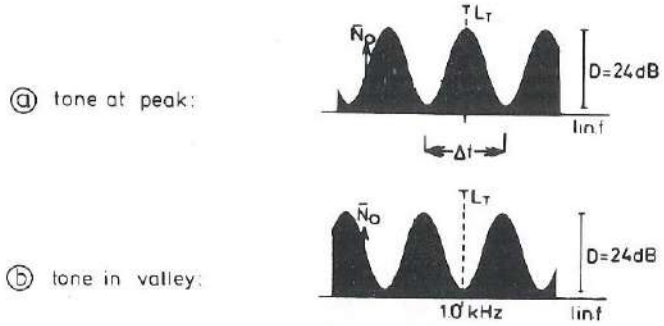
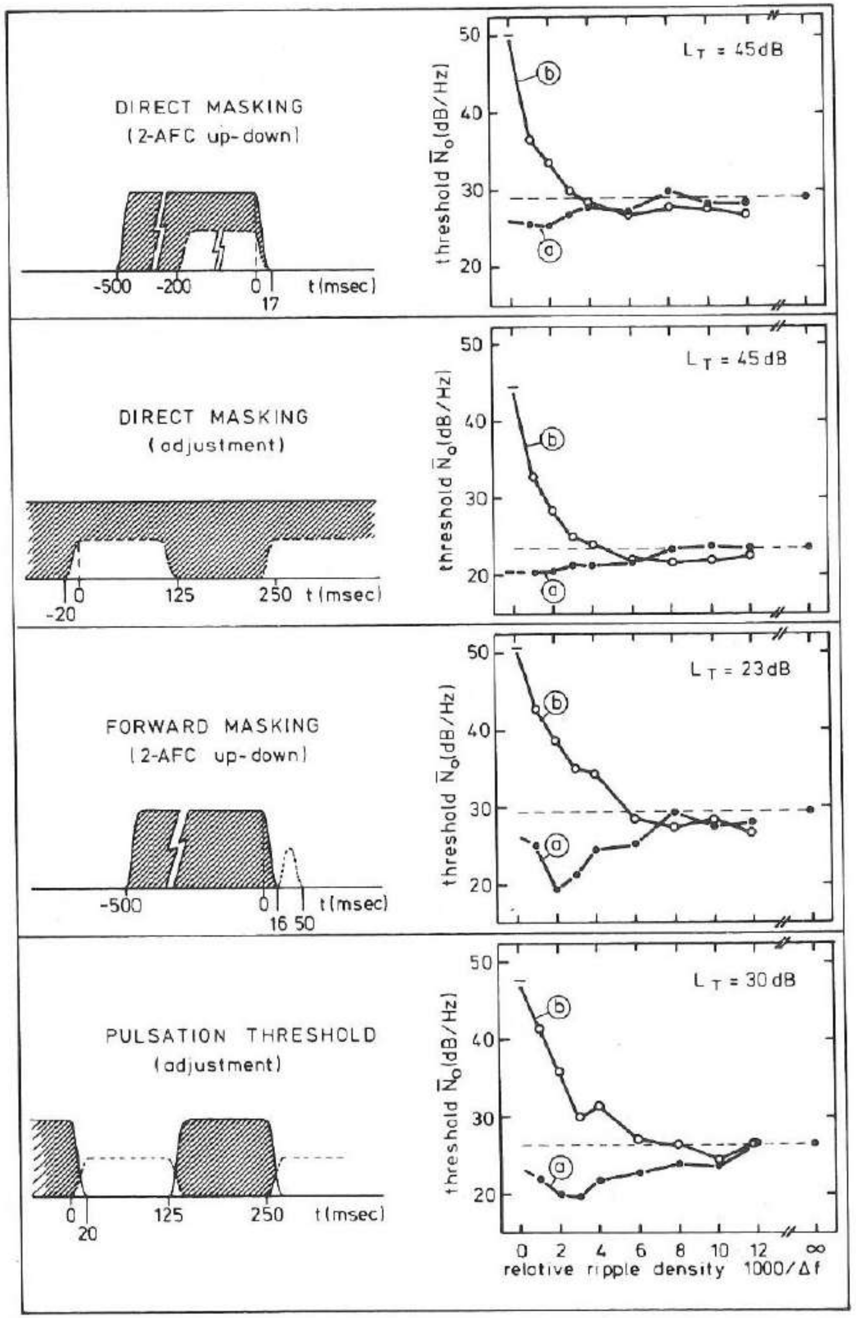


Fig. 9.2. Masking experiments with a rippled-noise masker.
Above: schematic presentation of the two conditions considered. Note that the level of the 1000-Hz test tone was fixed, that Δf was the independent, and $\bar{N}_0$ the dependent variable.
Opposite page: results of four types of masking experiments. The differences between the two conditions (a) and (b) indicate the degree of auditory ripple resolution. (Data of one subject.)



a. Direct masking, 2-AFC up-down

Two observation periods, separated by an interval of 1 sec, were repeated every 3 sec. In both periods a 500-msec masker burst was presented with 17-msec rise and fall times. At the end of one of the two masker bursts (randomly) a 200-msec test-tone burst was presented with 17-msec rise and fall times, such that the tail of the test-tone burst coincided with the tail of the masker burst. The 2-AFC up-down procedure with runs of 50 trials was used to measure the threshold value. In each condition four measurements were performed with one subject and the average values are presented in Fig. 9.2.

b. Direct masking, adjustment

The masker was presented continuously. The test tone was gated periodically, 125 msec on, 125 msec off, etc., with 20-msec rise and fall times. The subject controlled the dependent variable $\bar{N}_0$ and adjusted it to the lowest value for which the test-tone bursts were just not perceived. In each condition four measurements were performed with one subject and the average values are presented in Fig. 9.2.

c. Forward masking, 2-AFC up-down

Two observation periods, separated by an interval of 1 sec, were repeated every 3 sec. In both periods a 500-msec masker burst was presented with 16-msec rise and fall times. Immediately after one of the two masker bursts (randomly) a 34-msec test-tone burst was presented with 17-msec rise and fall times. The 2-AFC up-down procedure with runs of 100 trials was used to measure the threshold value. In each condition, four measurements were performed with one subject and the average values are presented in Fig. 9.2.

d. Pulsation threshold, adjustment

The pulsation-threshold method was used with the continuous alternation of 125-msec masker bursts and 125-msec test-tone bursts and with each fourth test-tone burst left out. The subject adjusted $\bar{N}_0$ to the lowest value for which the pulsating character of each series of three successive test-tone bursts was just not perceived. In each condition, eight measurements were performed with one subject and the average values are presented in Fig. 9.2.

The data in Fig. 9.2 reveal some interesting aspects. First, it will be seen that, also in the case of a rippled-noise masker, different experimental techniques give different results. Apart from minor differences, there appear

essentially two types of curves, which are associated with the temporal presentation of masker and test tone: the two graphs for *non*simultaneous masking show greater differences between the peak and valley conditions (thus, suggest greater "ripple resolution") than the two graphs for direct masking. I will return to this later. For the moment, a most important implication of the results in Fig. 9.2 is that the further experiments in this chapter can be restricted to only two measuring techniques (we will use the second and fourth of Fig. 9.2), the results of which may still be considered as typical of the two classes involved: direct masking versus nonsimultaneous masking.

There is a second aspect of the data in Fig. 9.2 which deserves further attention. Often, such data on the limited resolution power of sine-wave modulated stimuli are used to estimate the width or the shape of a weighting function, an "integration window", which could cause that limitation. In our case such an "integration window" would take the form of an intensity-weighting function in the frequency domain, or a *filter*. However, we have to do here with *psychophysical* data and we should formulate precisely the assumption which underlies the use of such psychophysical data for estimating the characteristics of that hypothetical filter. Because of the use of a test tone with *fixed* frequency and level, such an assumption can be very simple: all threshold values measured with a fixed test tone and with a fixed experimental paradigm reflect an ensemble of conditions for which the noise masker has the *same* intensity within a filter around the test-tone frequency. In other words: all noises which just mask a fixed test tone in a fixed paradigm have the same intensity within a certain filter. Under this "equal-filter-output" assumption, the psychophysical data can be used to specify that particular filter which would satisfy the data. I will return to this later. Here, I would like to indicate that some details of the data in Fig. 9.2 are not in line with the equal-filter-output assumption. The extreme right point in each of the four data graphs (for the limit condition $1000/\Delta f = \infty$) refers to a condition with flat, unmodulated noise with spectral level $\bar{N}_0$. The equal-filter-output assumption implies a specific relation between this flat-noise threshold and the peak and valley threshold for a rippled spectrum. These values are not independent; given two of these values and given the equal-filter-output assumption, the third value is completely determined. For instance, the equal-filter-output assumption does not allow that the threshold values for both the peak and the valley condition fall below the flat-noise threshold, which appears to be the case very systematically for higher ripple densities. Despite this, the equal-filter-output assumption will be used to interpret the masking data, especially the threshold *difference* between the peak and valley conditions as a function of ripple density, in terms of a filter bandwidth. It should be kept in mind that this interpretation does not account correctly for the *orientation* of the threshold values for the peak and valley conditions relative to that for the flat-noise condition.

9.3. EFFECT OF MODULATION DEPTH

This experiment is illustrated in Fig. 9.3. Its primary purpose was to investigate possible effects of the modulation depth D on the "ripple resolution" as revealed by masking experiments. As indicated, three values of the modulation depth D were used: 24 dB, 11 dB and 5.7 dB. The measurements were identical to those described in the previous section except that, out of the four experimental procedures, only direct masking (adjustment) and pulsation threshold (adjustment) were used. The data graphs in Fig. 9.3 always present the *difference* between the threshold values of $\bar{N}_0$ for the two conditions "tone at peak" and "tone in valley".

I will discuss the data in Fig. 9.3 in the light of the equal-filter-output concept as described in the previous section. Thus, a threshold difference between a peak and a valley condition is interpreted as reflecting just that level difference for which the two rippled-noise maskers have the same intensity within a filter around 1000 Hz. The question here is: can the degree of ripple resolution, as revealed for the three different values of D, be understood as reflecting the action of one and the same filter? This requires a specific relation between the three curves, namely that their general characteristics should be the same, apart from a different "scale factor" corresponding to the value of D. [This requirement can be formalized as follows: when all threshold differences (valley-peak) are converted into an apparent intensity-modulation depth m according to: threshold difference dB = 10 log (1+m) - 10 log (1-m) (see also Fig. 9.1), and when these m-values are normalized with respect to the physical modulation depth of the ripple (m_0 = 0.992, 0.853 and 0.575, respectively), and when this ratio m/m_0 is plotted as a function of ripple density, this should result in three identical curves.]

It can be seen, even without the formal elaboration, that the three curves in the left panel of Fig. 9.3 (direct masking) are in reasonable agreement with the above requirement. This means that the limited degree of ripple resolution, as revealed by the three curves for different modulation depths, can be understood as reflecting the action of one and the same filter. How-

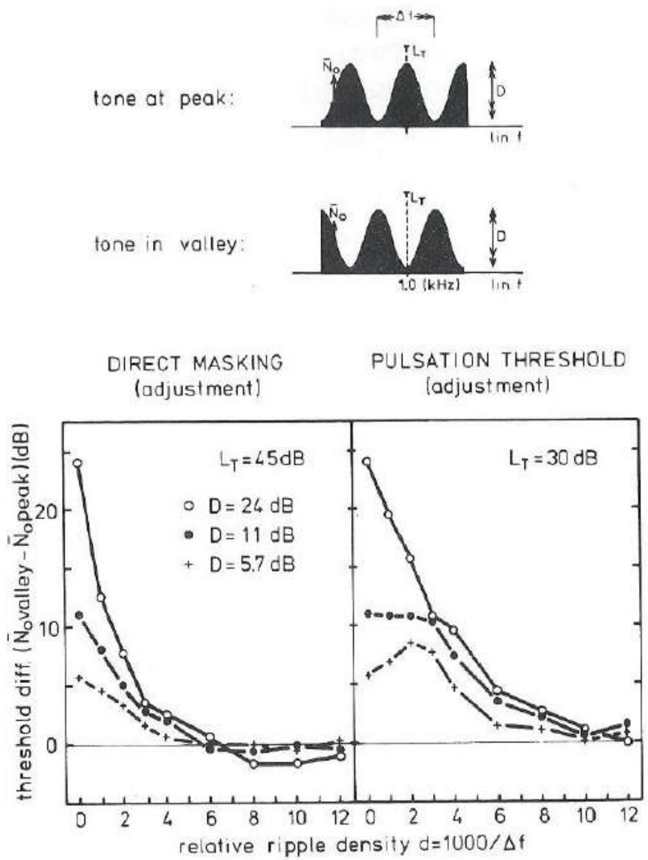


Fig. 9.3. Two masking experiments (direct masking and pulsation-threshold method) with a 1000-Hz test tone and a rippled-noise masker for various peak/valley ratios D (parameter). Each data point presents the *difference* between the threshold values for the two conditions 'tone in valley' and 'tone at peak'. (Data of one subject.)

ever, it can be seen that this is *not* the case for the three curves in the right panel of Fig. 9.3 (pulsation threshold). Here, the shape of the curve essentially depends on the modulation of the ripple. For D = 24 dB, the curve decreases monotonically for increasing ripple density, whereas for D = 5.7 dB the curve shows a distinct maximum around a relative ripple density of about 2 to 3. Thus, these three curves can *not* be understood as reflecting the action of one and the same filter. This would suggest that, within the equal-filter-output concept, the shape of the filter depends on the modulation depth of the masker. A filter with such a specific behaviour is not very realistic.

It seems more realistic to conclude that in ripple resolution, as measured with the pulsation-threshold technique, also processes other than simple intensity weighting along the frequency scale (filtering) are involved (as, for example, lateral suppression) and that the over-all result cannot be described, in detail, by the action of a filter.

There is one additional aspect in Fig. 9.3 which deserves further attention. For the small modulation depth (D = 5.7 dB), the two measuring techniques reveal an essential difference: whereas the curve for direct masking decreases monotonically for increasing ripple density, the curve for pulsation threshold shows a distinct maximum. The occurrence of such a maximum is very typical of the operation of a lateral-suppression mechanism. For the interpretation of this difference between the two measuring techniques it might be helpful, as we have done before in this study, to compare these psychophysical data with electrophysiological results for the same type of stimuli. Ten Kate *et al.* (1974) studied single-unit responses in cochlear nuclei of the cat to rippled-noise stimuli, in which the modulation depth D was treated as a parameter. For large values of D, the peak-valley difference in the firing rate decreases gradually for increasing ripple density. However, for smaller values of D (11 dB, 5.7 dB or 3.1 dB), the peak-valley differences show a maximum for a relative ripple density (unit's CF/Δf) of about 3 to 4. This suggests that the auditory representation of a rippled-noise spectrum, as revealed by the pulsation-threshold method, is closely related to the neural response to such stimuli.

9.4. FILTER CHARACTERISTICS DERIVED FROM RIPPLED-NOISE MASKING DATA

In this section, the equal-filter-output concept is used for transforming the psychophysical masking data into the shape of a filter, although the previous sections indicated a typical non-linear aspect which is not in agreement with that concept. Still, I will continue along this line a little further. Using the same two measuring techniques (direct masking, pulsation threshold) and the same three modulation depths for the rippled-noise masker as in the previous section, we shall arrive at six filter characteristics. If we find that the differences between those filters associated with the three modulation depths are only small in comparison with the differences associated with the two measuring techniques, the non-linear aspect might be considered a second-order problem. Then, it would still be possible to speak of a filter characteristic "typical" of direct masking and another characteristic

"typical" of nonsimultaneous masking, and this is what we are interested in primarily.

It cannot be excluded a priori that the auditory filter we would like to specify is *asymmetric*. Therefore, besides the two conditions "tone at peak" and "tone in valley", we need two additional conditions: "tone at positive slope" and "tone at negative slope". A difference between the thresholds for these two conditions would illustrate the asymmetry of the underlying filter.

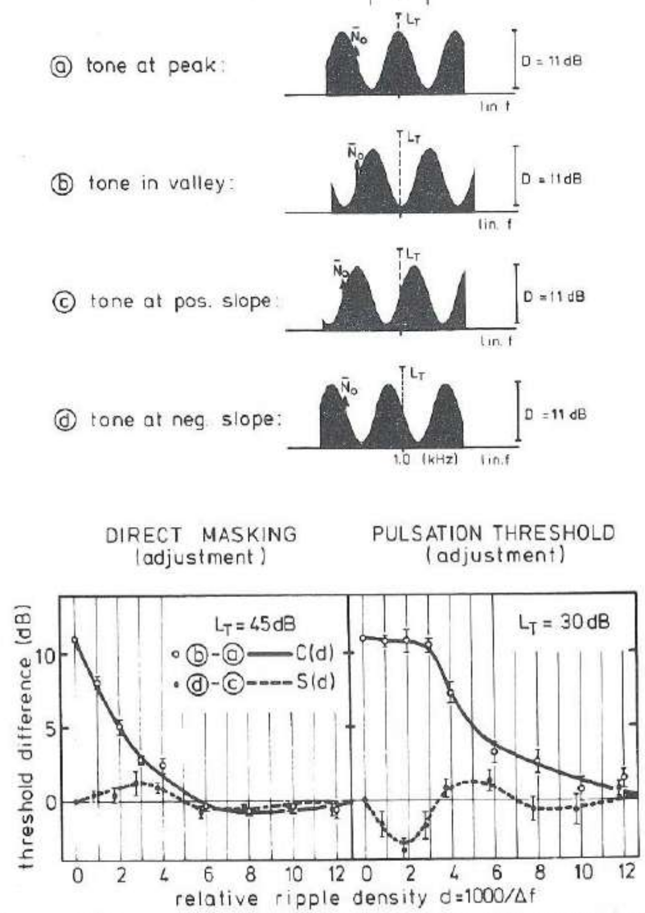


Fig. 9.4. Two masking experiments (direct masking and pulsation-threshold method) with a rippled-noise masker with a peak/valley ratio of 11 dB. Four conditions were considered (upper diagrams). The data points present the differences between the two conditions 'tone in valley' and 'tone at peak' (open symbols), and between 'tone at negative slope' and 'tone at positive slope' (closed symbols). The vertical intervals represent standard errors. The curves C(d) and S(d) are used to derive the filter characteristics in Fig. 9.5 middle panel. (Data of one subject.)

Apart from this extension of the masker conditions, the measurements were identical to those in the previous section. Fig. 9.4 presents the data for the condition D = 11 dB. Again, the *differences* between the threshold values for two conditions (b-a and d-c) are plotted as a function of the relative ripple density d. The two curves drawn through the data points will be referred to as C(d), the level difference for the two *cosine* conditions, and S(d), the level difference for the two *sine* conditions, respectively. The same experiment was performed for D = 5.7 dB and D = 24 dB, respectively.

For each of the six conditions (two measuring techniques and three modulation depths) a filter characteristic was derived from the curves C(d) and S(d) on the basis of the equal-filter-output concept. Thus, given a combination of C(d) and S(d) curves, the question was *what* filter, when presented with the rippled-noise signals, would have given those *same* curves when the criterion would have been to keep the filter output *constant* rather than "just mask" a fixed test tone. The results are presented in Fig. 9.5. The formal derivation is given below.

A filter can be defined as an intensity-weighting function W(f). It is assumed that the filter, in case of a 1000-Hz test tone, is located within the range f = 500 - 1500 Hz, thus, formally:

$$W(f) = 0, \text{ for } f < 500 \text{ Hz and } f > 1500 \text{ Hz}.$$

Any filter within that range can be written as

$$W(f) = A_0 + \sum_{n=1}^{\infty} A_n \cos 2\pi n \frac{f-1000}{1000} + \sum_{n=1}^{\infty} B_n \sin 2\pi n \frac{f-1000}{1000} .$$

Since our primary interest is the *shape* of the filter, all coefficients can be divided by A_0:

$$w(f) = 1 + \sum_{n=1}^{\infty} a_n \cos 2\pi n \frac{f-1000}{1000} + \sum_{n=1}^{\infty} b_n \sin 2\pi n \frac{f-1000}{1000} . \qquad (1)$$

The coefficients a_n and b_n can be derived from the curves C(d) and S(d), respectively. Consider, for example, the rippled-noise spectra for the peak and valley conditions:

$$I(f) = \bar{I} \left(1 \pm m_0 \cos 2\pi \frac{f-1000}{\Delta f}\right) .$$

Or, with relative ripple density $d = 1000/\Delta f$:

$$I(f) = \bar{I} \left(1 \pm m_0 \cos 2\pi d \frac{f-1000}{1000}\right) .$$

The intensity of the rippled noise within the filter w(f) is given by the convolution of the noise spectrum and the weighting function w(f). Hence, the level difference (in dB) between the peak and valley conditions, C(d), can be written:

$$C(d) = 10 \log \frac{\int_{f=500}^{1500} \left(1 + m_0 \cos 2\pi d \frac{f-1000}{1000}\right) w(f)\, df}{\int_{f=500}^{1500} \left(1 - m_0 \cos 2\pi d \frac{f-1000}{1000}\right) w(f)\, df} . \qquad (2)$$

By substitution of Eq. (1) in Eq. (2) it can easily be seen that, for the specific case d = n, this reduces to

$$C(n) = 10 \log \frac{1 + \frac{1}{2} m_0 a_n}{1 - \frac{1}{2} m_0 a_n} .$$

From this:

$$a_n = \frac{2}{m_0} \; \frac{10^{\frac{C(n)}{10}} - 1}{10^{\frac{C(n)}{10}} + 1} . \qquad (3)$$

Similarly:

$$b_n = \frac{2}{m_0} \; \frac{10^{\frac{S(n)}{10}} - 1}{10^{\frac{S(n)}{10}} + 1} . \qquad (4)$$

With Eqs. (3) and (4), a set of a_n and b_n coefficients can be derived from the C(d) and S(d) curves for each of the six conditions. (D = 24 dB, 11 dB and 5.7 dB corresponds to m_0 = 0.992, 0.853 and 0.575, respectively.) Then w(f) can be obtained from Eq. (1). Since no finer gratings were used than those with a relative ripple density d = 12, only the coefficients a_n and b_n up to n = 12 can be derived in this way. (Formally, the higher coefficients in Eq. (1) are set zero.) This limits the details of w(f) which can be revealed in this way, and, accordingly, w(f) is calculated only for 24 positions along the frequency scale between 500 and 1500 Hz.

Thus, under the assumption that, for a fixed test tone, "threshold" is always associated with a certain, fixed, signal-to-noise ratio within an audi-

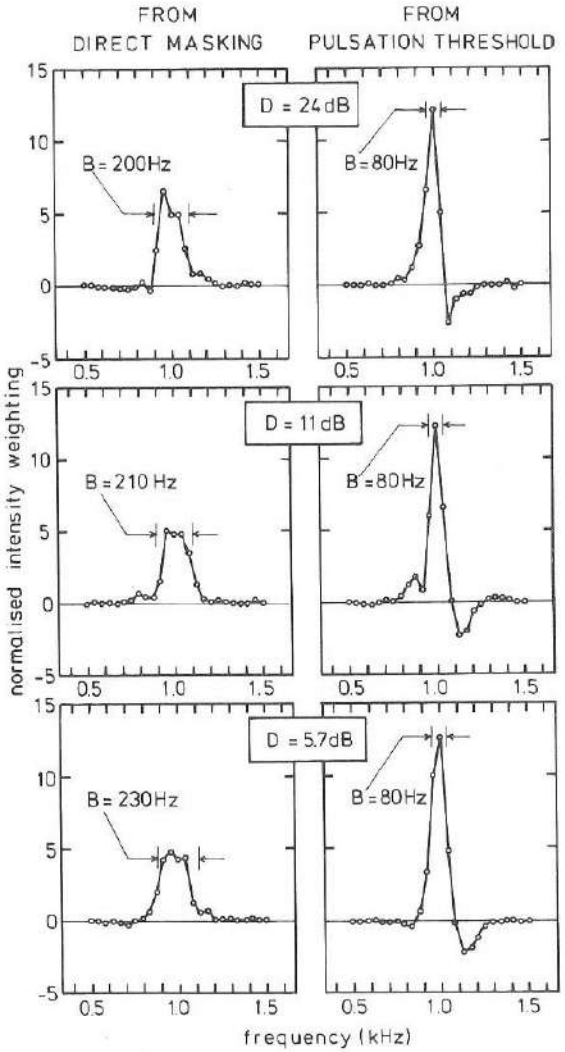


Fig. 9.5. Filter characteristics, in the form of intensity weighting functions, as derived from two types of masking experiments with the rippled-noise masker. Note that the differences associated with the peak / valley ratio D are small in comparison with the differences associated with the two types of masking experiments.

tory filter, w(f) is the characteristic of that very auditory filter. Fig. 9.5 shows that the main difference among the six filters is associated with the experimental technique: direct masking versus pulsation threshold. Compared with this, the effect of the modulation depth D of the rippled noise is only of secondary importance. Thus, it appears still to be possible to interpret differences between direct-masking and pulsation-threshold data with respect to ripple resolution in terms of typical differences between the underlying

auditory filters. The direct-masking data reveal a relatively "simple" auditory filter, with a bandwidth B of about 210 Hz. [B is the "equivalent rectangular bandwidth": the width of a rectangular filter with the same area and the same value w(1000 Hz) as the actual function w(f). For white noise, it gives the same S/N ratio for a 1000-Hz tone as the filter w(f).] The pulsation-threshold data reveal a much-sharper filter with a negative-going part at the high-frequency side and a bandwidth B of 80 Hz. The negative-going part implies that noise within that range *reduces* the output of the filter. It is highly suggestive of the operation of lateral suppression which acts, in accordance with what was found in previous chapters, mainly from higher towards lower frequencies. As was mentioned before, it is questionable whether the operation of such a lateral-suppression mechanism can be accounted for correctly by a negative-going part in an intensity-weighting function. The site and nature of such a mechanism are obscure, and it may very well operate at a level where "intensity summation along the frequency scale" has totally lost its meaning. Therefore, it is not surprising that a simple intensity-weighting-function concept leads to nonlinearities of the kind seen in the previous section.

This section has shown that an estimation of an auditory bandwidth from ripple-resolution data is well possible and that the modulation depth of the ripple has only a minor influence on the magnitude of that estimate. In the next section, we will further proceed along this line.

9.5. AUDITORY BANDWIDTH DERIVED FROM RIPPLE-RESOLUTION DATA

In this section, more extensive threshold measurements with the rippled noise will be described, for a number of additional test-tone frequencies and with more subjects. The main purpose is to determine the order of magnitude of the auditory bandwidth, or "critical bandwidth", which might underly such data. We are not interested here in the exact filter shape, but primarily in the differences between the bandwidths derived from direct-masking data and from pulsation-threshold data.

It will first be demonstrated that one needs only a limited amount of data to estimate the bandwidth of the auditory filter with some accuracy, without further considerations about the exact shape of that filter. In Fig. 9.6, five filters, or intensity-weighting functions, are considered which differ markedly in shape but all have the same bandwidth B. (Again, B is defined as the bandwidth of a *rectangular* filter with the same area and the same top val-

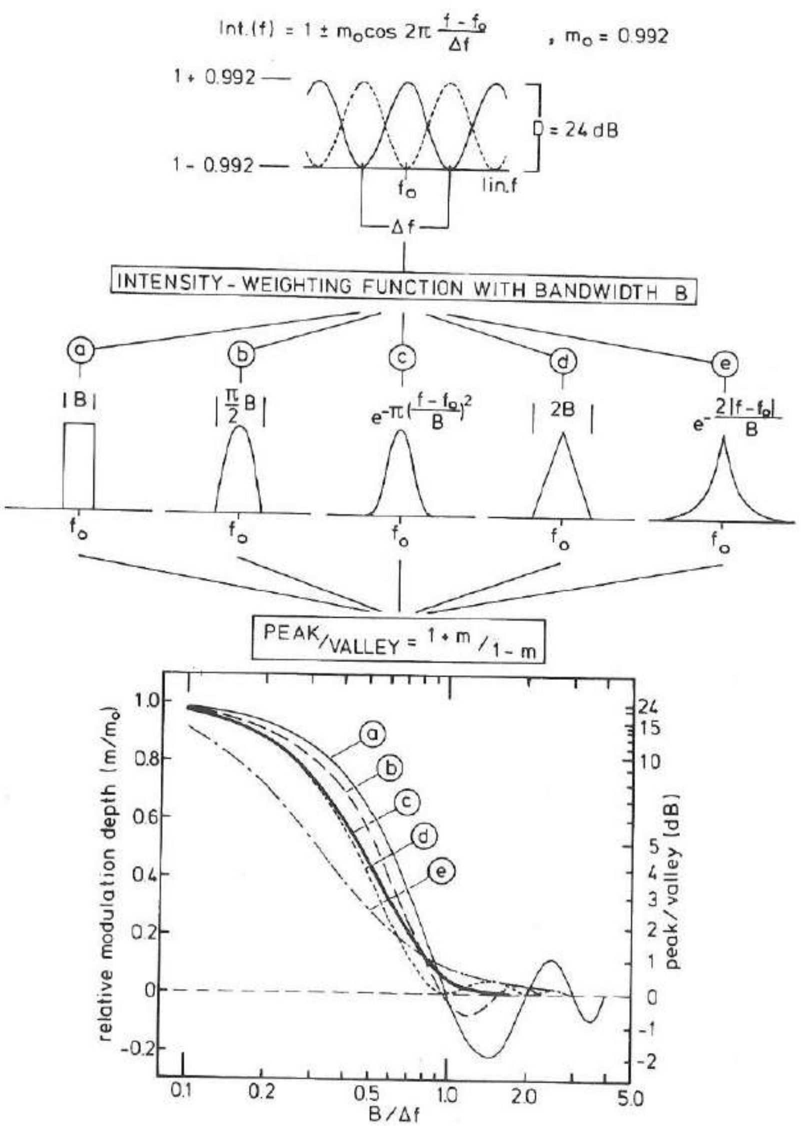


Fig. 9.6. The effect of filter shape on ripple resolution. Five different intensity-weighting functions, with the same 'equivalent rectangular bandwidth' B, are applied to the two rippled-noise spectra as indicated at the top. The difference between the 'peak' and 'valley' conditions, as a function of B/Δf (abscissa), is expressed in terms of the relative modulation depth m/m_o (left side ordinate), and in terms of the ratio in dB (right-side ordinate). Curve (c) will be used to interpret the results of masking experiments with the rippled-noise in terms of an 'auditory bandwidth'.

ue as the actual filter.) When the rippled-noise signal, with ripple density 1/Δf and D = 24 dB, is applied to such a filter, the level difference at the filter output for the two conditions "filter centred at top" and "filter centred at valley" can be calculated. This level difference is a function of the ratio B/Δf. The results for the five filter types are given in the lower panel in Fig. 9.6. On the basis of these results it was decided to perform the threshold measurements for only four values of ripple density 1/Δf, with intervals of a factor two. Furthermore, it was decided to perform the bandwidth estimation by fitting the data points thus obtained with the curve corresponding to the Gaussian-shaped filter (curve c), unless it would appear that its slope deviates systematically from the slope suggested by the data points.

The measurements will provide threshold differences between the peak and valley conditions for four values of ripple density 1/Δf. Fig. 9.7 presents a graph in which such ripple-resolution data can be plotted. The family of curves in this graph is based on curve (c) in Fig. 9.6; each curve indicates how the level difference depends on 1/Δf for a Gaussian filter with the indicated bandwidth B. Hence, this family of curves, together with the B-scale in the graph, provides an easy estimation of the order of magnitude of the bandwidth B associated with the data points. As an example, the four data points

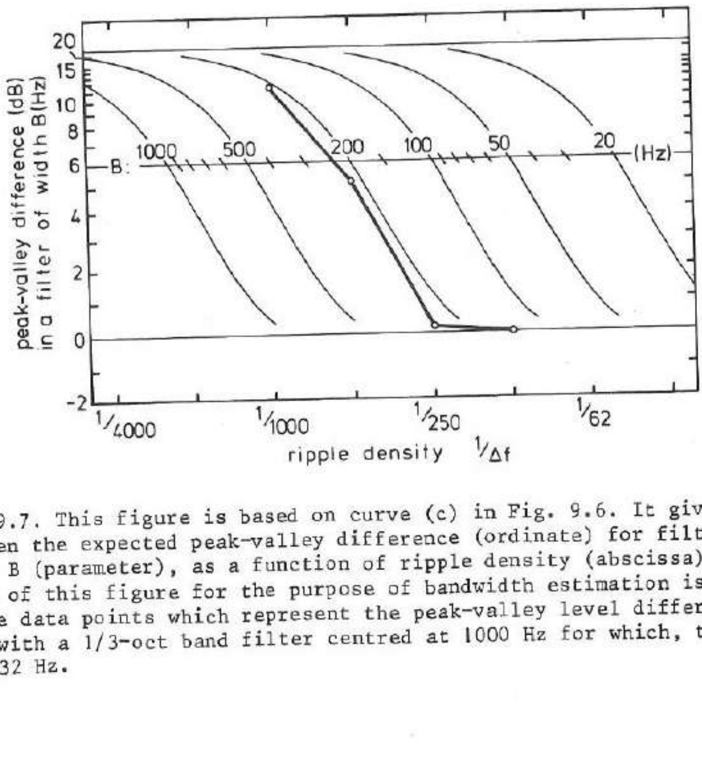


Fig. 9.7. This figure is based on curve (c) in Fig. 9.6. It gives the relation between the expected peak-valley difference (ordinate) for filters with bandwidth B (parameter), as a function of ripple density (abscissa). The relevance of this figure for the purpose of bandwidth estimation is illustrated by the data points which represent the peak-valley level difference as measured with a 1/3-oct band filter centred at 1000 Hz for which, theoretically, B = 232 Hz.

in Fig. 9.7 refer to physical measurements with the rippled noise, performed with a Bruël and Kjaer 1/3-octave band-pass filter with 1000 Hz centre frequency. They present the level difference at the filter output between the two conditions "top centred at 1000 Hz" and "valley centred at 1000 Hz" for the four values of $1/\Delta f$ as indicated. From these data points one would estimate the filter bandwidth to be about 230 Hz; the theoretical "rectangular bandwidth" of that filter is 232 Hz. This illustrates the validity of this approach.

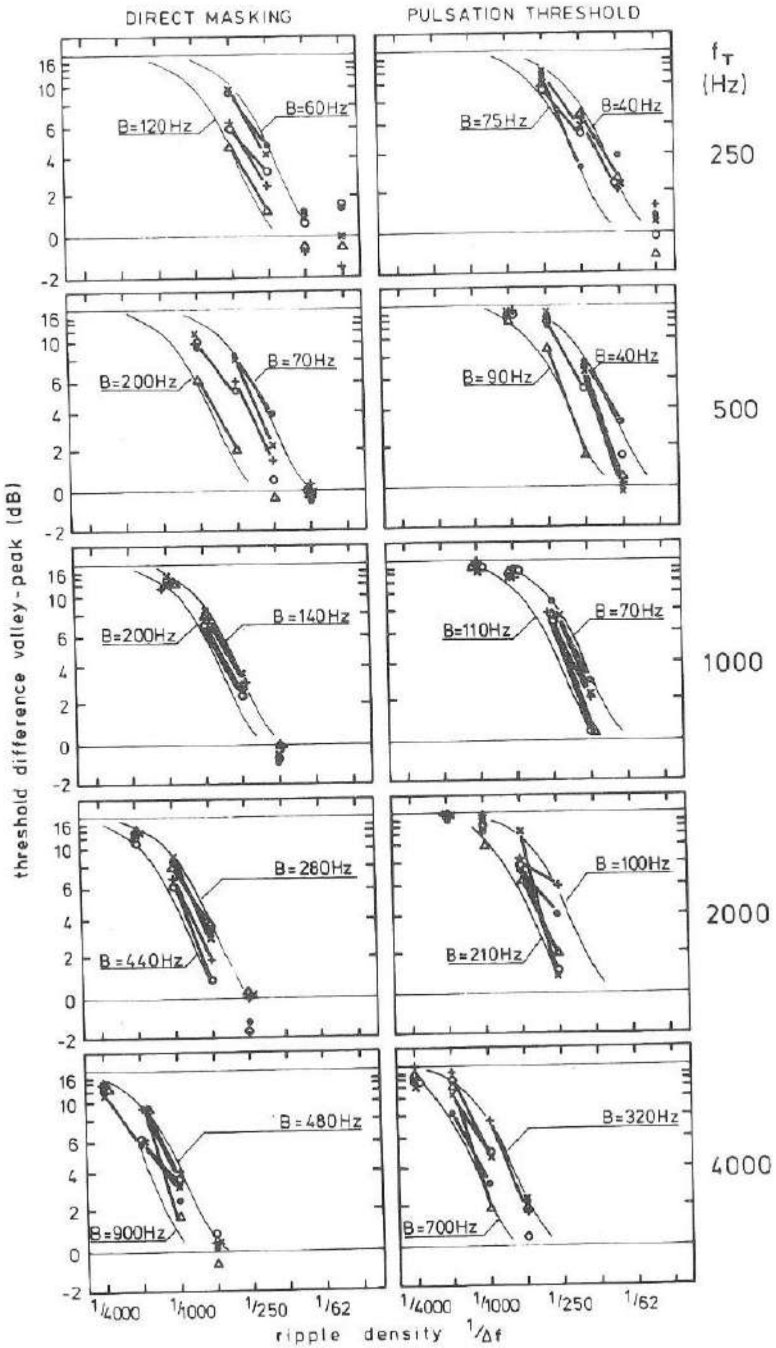
9.5.1. Measurements

Two measuring techniques were used: direct masking and pulsation threshold, as in the second and fourth examples in Fig. 9.2. In this case, five frequencies for the fixed test tone were used: f_T = 250, 500, 1000, 2000 and 4000 Hz, respectively. Each of the ten conditions (two measuring techniques, five frequencies) was investigated in a separate session.

The four ripple densities used in a session were adapted to the test-tone frequency, such that the *relative* ripple density $f_T/\Delta f$ was 1, 2, 4 or 8. For both the top and valley conditions, each subject made four threshold adjustments of the noise level $\bar{N}_0$. Hence, one session consisted of 32 adjustments. As we were interested in threshold *differences* between the peak and valley conditions, these two conditions were always presented in immediate succession (in random order). The 16 pairs in a session were presented in a random sequence. Hence, for each ripple density, four values of the peak-valley threshold difference were obtained and the average values are plotted in Fig. 9.8. Five subjects participated and their data are presented separately.

Finally, the level of the fixed test tone used in each of the ten conditions in Fig. 9.8 has to be specified. For direct masking, the levels for the different f_T-values were chosen such that all tones were about 40 dB above absolute hearing threshold. (We used, for f_T = 250, 500, 1000, 2000 and 4000 Hz, levels of 50, 44, 40, 38 and 35 dB *re* 1000-Hz absolute threshold, respective-

Fig. 9.8 (Opposite page). Two types of masking experiments (direct masking and pulsation-threshold method) were performed with a test tone of various frequencies (f_T as indicated at the utmost right side). Again $\bar{N}_0$ of the rippled noise was the dependent variable. For four values of the ripple density (abscissa), the threshold value of $\bar{N}_0$ was measured for the two conditions 'tone at peak' and 'tone in valley', and the *differences* between these two conditions are presented. Each panel is a graph similar to Fig. 9.7, and the upper and lower estimates of the bandwidth B, associated with the positions of the data points, are indicated. (Individual data of five subjects.)



ly.) For pulsation threshold somewhat lower levels were used in order to arrive at adjustments of the noise level $\bar{N}_0$ in the same range as those for direct masking. (For 250, 500, 1000, 2000 and 4000 Hz this level reduction was 0, 2, 5, 5 and 0 dB, respectively.) The exact values of the test-tone level are not critical with respect to the data in Fig. 9.8, since these always refer to threshold *differences*.

9.5.2. Discussion

Each panel in Fig. 9.8 is a graph similar to Fig. 9.7. However, instead of presenting the whole family of curves for different bandwidths B, only two curves are plotted which form the upper and lower boundary of the range covered by the data points of the five subjects. The bandwidth B associated with the position of each curve is indicated in the figure. These upper and lower estimates of the bandwidth for each of the five test-tone frequencies are presented in Fig. 9.9, both for the direct-masking data and for the pulsation-threshold data.

Fig. 9.9 shows that the bandwidths derived from ripple-resolution data obtained in direct masking, are of the same magnitude as the traditional "critical bandwidths" which play a role in a wide variety of auditory phenomena (*e.g.*, Scharf, 1970). Compared with this, the bandwidths derived from the pulsation-threshold data are considerably smaller. The difference is about a factor of two. It seems justified to ascribe this difference to effects of lateral suppression which, according to all previous experiments, is not revealed by direct-masking experiments, but only when masker and test tone are presented *non*simultaneously. This point will be discussed further in the next chapter.

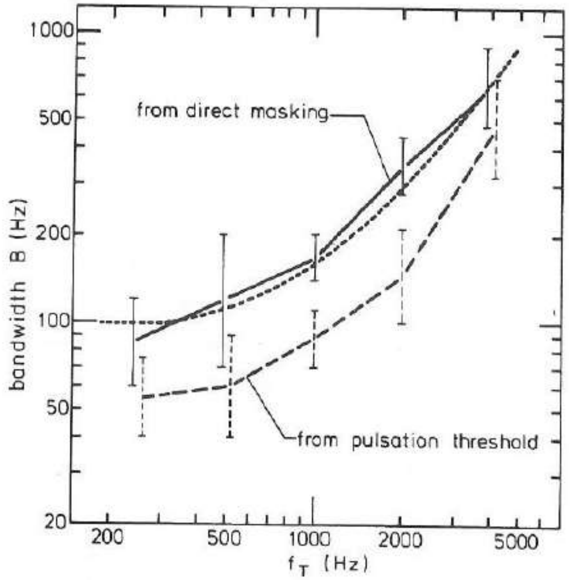


Fig. 9.9. The ranges between the upper and lower estimates of the bandwidth B, as indicated in each panel in Fig. 9.8, are plotted as a function of the frequency of the test tone used in the corresponding masking experiment. The dotted line represents the 'critical bandwidth' after Scharf (1970).

I would like to underline the simplicity of the experimental method used here for the estimation of a "critical bandwidth". Only a few threshold measurements of tone in noise are needed and, furthermore, these can be obtained by a simple procedure without special considerations about systematic deviations, since only threshold *differences* are considered. Although this falls beyond the scope of this study, this method might be well applicable, at least in principle, as an audiometric test to provide information about the status of a patient's ear in terms of "critical bandwidth", or "resolution power along the frequency scale", which is a fundamental property of the hearing mechanism. It might even appear that, when both the direct-masking method and the pulsation-threshold method could be applied, the difference between the two bandwidths (possibly associated with a lateral-suppression mechanism) will give valuable additional information.

CHAPTER 10

DISCUSSION

Summary

The experimental results strongly suggest that lateral suppression is involved in the auditory projection of a sound spectrum. A major question is why one experimental paradigm (direct masking) does not show this effect, while other techniques do so consistently (*non*simultaneous masking). It is shown that this can be understood within a framework of reasonable assumptions, of which the most important are: (1) the effect of lateral suppression is comparable with that of an attenuation *factor* in the suppressed frequency region (such that the S/N *ratio* in direct masking remains unchanged), and (2) this effect ends almost instantaneously at the end of the masker (such that the argument of constant S/N ratio does not apply to a *non*simultaneous test tone).

Additionally, the relevance of the pulsation-threshold method is discussed in some detail. A theory about the possible underlying mechanism, based upon the close correspondence between pulsation-threshold data and neural-response data, implies that the pulsation-threshold method provides a most simple way to investigate what we called earlier the "auditory projection" of a masker's sound spectrum. The last chapter gives some examples of this application of the pulsation-threshold method.

10.1. EVALUATION OF THE EXPERIMENTAL RESULTS

Table I gives a survey of the experimental results on contrast enhancement, suggesting the operation of a lateral-suppression mechanism. The main conclusions to be drawn from this table are:

(a) direct-masking data consistently show *no* effects of lateral suppression,

(b) forward-masking data do consistently show effects of lateral suppression,

(c) pulsation-threshold data also consistently show effects of lateral suppression,

(d) with respect to these effects of lateral suppression, the results do not

Table I. Survey of experimental data on effects of lateral suppression. A + indicates that the masking effectiveness (or the neural response) *decreases* when the masker's intensity in an adjacent frequency region is *increased* (or vice versa). A - indicates that no such effect was found.

EFFECT OF	FIG.	DIRECT MASKING	FORWARD MASKING	PULSATION THRESHOLD	NEURAL DATA
enhancement at the edge of a noise band	3.1	(c) -	(c) +	(a) +	
tone-on-tone suppr.	5.1 5.3	(b,c) -	(c) +	(a) + (a) +	Nomoto *et al.* (1964) + Sachs and Kiang (1968) + Arthur *et al.* (1971) +
noise-on-tone suppr.	7.1 7.5	(c) -	(c) +	(a) + (a) +	Kiang (1965) +
suppression in the centre of a noise band	8.2	(b) -		(b) +	Greenwood and Goldberg (1970) +
contrast enhancement for rippled noise with small mod. depth	9.3	(a) -		(a) +	ten Kate *et al.* (1974) +

measuring procedures: (a) adjustment, (b) Békésy up-down, (c) 2-AFC up-down.

depend on the procedure used in measuring the threshold of the test tone (2-AFC up-down, Békésy up-down or adjustment),

(e) in those cases in which electrophysiological data on the neural response to the same type of stimuli are available, these consistently show effects of lateral suppression.

Conclusion (d) speaks for itself. The other conclusions (b, c and e) strongly suggest that lateral suppression is indeed involved in the auditory processing of a stimulus' sound spectrum. Then, the main question remains: can we understand why *non*simultaneity of masker and test tone is an essential condition for revealing the effect of lateral suppression?

In this section, I will develop a framework of some simple assumptions about features of the lateral-suppression mechanism and about processes associated with the different experimental techniques, which can account for the conclusions given above.

10.1.1. Implications of direct-masking data

Let us consider the case of a masker and a test tone in a condition of simultaneous (direct) masking. A basic empirical finding is that the threshold of the test tone depends primarily upon the signal-to-noise ratio (S/N ratio) in a restricted spectral region around the test-tone frequency. When the masker's intensity in that spectral region is changed, the test-tone threshold changes accordingly with the same *factor*, and the S/N *ratio* at threshold remains constant. (Although this rule has some exceptions, those will not be considered here.)

We found experimentally that, when according to other measures (nonsimultaneous masking, neural response) the degree of lateral suppression in the frequency region of the test tone changed, this had essentially *no* effect on the test-tone threshold in direct masking. For the lateral-suppression mechanism this implies that, although it reduces the neural response to the masker in the test-tone frequency region, it does *not* affect the apparent S/N ratio in that frequency region. This means that, at whatever level of auditory processing the lateral-suppression mechanism may operate, it does not influence the specific feature which, at that very level, is associated with constant S/N ratio. Such a mechanism can be easily conceived. A most simple model, which is more illustrative than realistic, is the following. Consider the case that lateral suppression operates at a very peripheral level, at which it is still relevant to speak of a S/N ratio in terms of intensities within the test-tone frequency region. Then, if lateral suppression operates such that addition of intensity in an adjacent frequency region causes a reduction of the effectiveness in the test-tone frequency region by a certain *factor*, this would not affect the S/N ratio and, hence, would leave the threshold of the test tone unchanged. Another, more physiological but also highly schematic, model is the following. Consider the case that lateral suppression operates at a neural level where the rate of discharges is proportional to the *logarithm* of the intensity in the corresponding frequency region. Then, the feature associated with a constant S/N ratio is a certain absolute, not relative, increment of the discharge rate. If lateral suppression operates such that addition of intensity in an adjacent frequency region would accomplish a reduction of the discharge rate in the suppressed frequency region by a certain *amount*, this would not affect the *absolute* increment brought about by a test tone at threshold and, hence, would leave the threshold of the test tone unchanged.

These considerations are meant to illustrate how our first conclusion, "direct masking does not reveal effects of lateral suppression" can still be reconciled with the existence of such a mechanism. It simply implies that the mechanism is such that the effect of lateral suppression is comparable with

a reduction of the intensity in that frequency region by a certain *factor*, affecting *both* the masker and the test tone. This implication is illustrated in Fig. 10.1, upper panel (stimuli and response patterns). Here, two conditions are compared, a single-tone and a two-tone stimulus, and two of the main conclusions, as mentioned before, are accounted for: conclusion (e), the

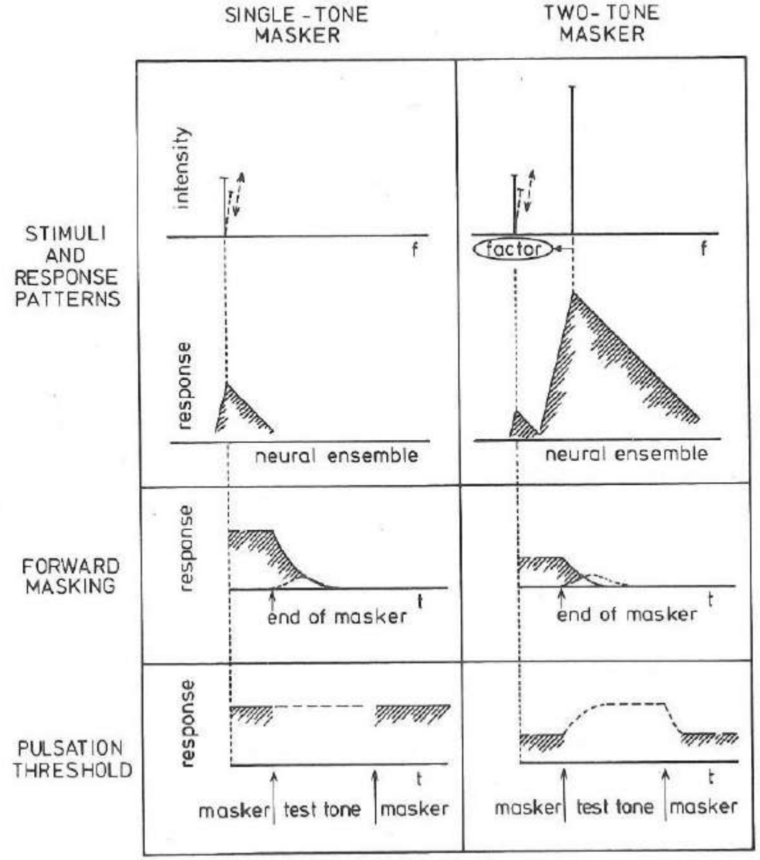


Fig. 10.1. Illustration of the framework of assumptions which may account for the main results of the present study. The addition of a second (strong) component to a first (weak) component reduces the neural response to the weak component (literature data on lateral suppression). If this lateral suppression behaves as a reduction *factor* in the suppressed frequency region (upper panel), it has no effect on the threshold for a test tone *superimposed* on the weaker component (direct masking), when this threshold is associated with a certain S/N *ratio*. If, additionally, the effect of lateral suppression stops instantaneously at the end of the masker, the addition of the strong component will have an effect in case of a *non*simultaneous test tone (forward masking, pulsation threshold), when the first reflects a process of decay of (or recovery from) neural effects of the preceding masker (middle panel) and when the latter is associated with continuity of the neural response (lower panel).

neural response to a single-tone stimulus is reduced by the addition of a second strong tone, and conclusion (a), lateral suppression, comparable with a reduction of the effectiveness in the suppressed frequency region by a certain *factor*, does not affect the test-tone threshold in direct masking.

10.1.2. Implications of forward-masking data

In contrast to direct masking, forward-masking data do reflect the effect of lateral suppression. This agrees with the model developed so far, if we assume that the lateral-suppression mechanism is a "fast" mechanism, such that the effectiveness of a test tone presented shortly *after* the masker is not affected by it. The argument goes as follows. The origin of forward masking may be described either in terms of "decay of response" or in terms of "recovery from adaptation". This distinction, which is of interest for hearing theories, is not of importance for the argument here. Let us consider the first concept. Thus, forward masking reflects a decay of the response caused by the preceding masker in the frequency channel corresponding to the test tone. Then, a change of the degree of lateral suppression in that frequency region will bring about a change in the forward-masked threshold if two conditions are fulfilled: (1) lateral suppression operates at a stage *preceding* the decaying process involved in forward masking, and (2) the test tone itself, presented shortly after the termination of the masker, is *not* subjected to that same effect of lateral suppression. This is illustrated in Fig. 10.1, middle panel (forward masking). The first condition means that the initial level of the decay process at the end of the masker is lower in case of the two-tone masker than in case of the single-tone masker. The second condition means that the effectiveness of the test tone is the same in both situations.

Thus, in order to account for the forward-masking data, our model should be extended with two more assumptions: (1) lateral suppression set up by the masker stops almost instantaneously at the end of the masker, and (2) forward masking reflects the decay of "something" (either a response level or an adaptation level set by the preceding masker) at a level in the auditory pathway beyond the stage were lateral suppression operates.

10.1.3. Implications of pulsation-threshold data

The pulsation-threshold method is not a common method and there does not exist a theory about the possible underlying processes. It appeared that the pulsation threshold correctly reflects the effect of two-tone suppression (see, for instance, Fig. 5.3) and, for that specific case, the theory can be

very simple: when alternating a test tone and a two-tone stimulus (of which the weaker tone has the same frequency as the test tone), the test tone will sound as a continuous tone if the neural response to the test tone is of the same "height" as the neural response to the weaker tone in the two-tone stimulus. Thus, a tone is perceived as continuous when the corresponding neural response is continuous. This condition of continuity should refer to the neural response at a level in the auditory pathway beyond the stage where lateral suppression operates. Fig. 10.1, lower panel, illustrates the concept. For the trivial case of a single-tone masker alternating with a test tone of the same frequency, "pulsation threshold" is reached when the test tone has the same level as the masker. When the second tone is added, suppressing the neural response to the weaker tone, the condition of continuity is no longer fulfilled until the test-tone level is reduced appropriately. As in the case of forward masking, it is essential that the lateral suppression brought about by the second tone acts "instantaneously", since the response to the test tone itself should, of course, *not* be subjected to the suppression effect of the second tone.

These considerations illustrate that conclusion (c) of section 10.1, referring to pulsation-threshold data, does not imply any additional assumption about the lateral-suppression mechanism. Within the model presented so far, the assumption that perceptual continuity (pulsation threshold) is associated with continuity in the neural response accounts for conclusion (c).

Our framework of assumptions, which can account for the main conclusions of this study, has taken the following form.

Two assumptions concerning the lateral-suppression mechanism:

(1) the effect of lateral suppression is comparable with a reduction of the effectiveness in the suppressed frequency region by a certain *factor*,
(2) lateral suppression acts essentially *instantaneously*, such that its effect stops immediately at the end of the frequency components which cause that suppression.

Three further assumptions concerning the different experimental techniques:

(3) the direct-masked threshold is associated with a certain S/N *ratio* in the frequency region of the test tone,
(4) the forward-masked threshold is associated with a process of recovery or decay at a level in the auditory pathway *beyond* the stage where lateral suppression operates,
(5) the pulsation threshold is associated with *continuity* of the neural re-

sponse at a level beyond the stage where lateral suppression operates.

This framework of assumptions has a descriptive character and remains rather vague in many respects. For instance, the notion "neural response", which is frequently used here, remains unspecified with respect to the level in the auditory pathway to which it refers. However, it must be kept in mind that the present study is a *psychophysical* one. As mentioned in chapter 1, psychophysics is particularly suited for investigating "input-output" rules of the hearing mechanism as a whole, but does not reveal detailed processes at the different stages of auditory processing; such information can only be obtained from physiological studies. The psychophysical rules may only suggest some general assumptions concerning the nature of the underlying processes. The purpose of this section was to indicate at least one consistent framework of such assumptions, which can account for the main results of this study.

A most important conclusion of this section is that, within the framework developed here, the contribution of lateral suppression in "sharpening" the auditory projection of a masker's sound spectrum may be estimated from the difference between pulsation-threshold data and direct-masking data: the first do and the latter do not reveal its action. According to this, the various experiments indicate that lateral suppression contributes significantly to the preservation (in some cases even an enhancement) of spectral contrasts. The rippled-noise data indicate the order of magnitude by which lateral suppression contributes to the resolving power in the frequency domain: the bandwidth of the "auditory filter" *including* lateral suppression is about a factor of two smaller than the classical "critical bandwidth" derived from the direct-masking data. Thus, it would appear that the auditory projection of a stimulus' sound spectrum is about a factor of two "sharper" than according to classical theories. This will be illustrated further in chapter 11.

10.2. LATERAL SUPPRESSION AND NEURAL DATA

The evaluation of the experimental data suggested a lateral-suppression mechanism which acts almost *instantaneously*, whereas its effect is comparable with a reduction *factor* in the suppressed frequency region. Do such features agree with the nature of lateral suppression as revealed by electrophysiological studies? This question involves more than a qualitative agreement between masking data and neural-response data, as shown in some of the previous chapters. Here, we are concerned with the question of whether the *nature* of the lateral-suppression mechanism, as derived from our results, is not in conflict with results of physiological studies. We will consider this question only briefly.

Recently, an extensive review of physiological data on the coding of sound in the lower levels of the auditory system is given by Møller (1972). The temporal aspect of a possible sharpening mechanism is illustrated most clearly by a comparison between the frequency-response curve of neural units in the cochlear nucleus of the rat in case of steady-state pure tones and the impulse response as derived from the unit's response to paired clicks (Møller, 1970). The selectivity derived from the impulse response is not less than in case of long pure tones. Hence, Møller concludes that, if a lateral-suppression mechanism is responsible for a sharpening of tuning, it must act instantaneously. Also in studies on two-tone inhibition no important differences between the latencies involved in the excitatory effect and the inhibitory effect of a pure tone have been found, at least in primary auditory-nerve fibres (Nomoto *et al.*, 1964 and Arthur *et al.*, 1971); at higher levels, as in units in the inferior colliculus in rat (Vartanian, 1973), substantial latencies in inhibition have been found.

The instantaneous character of lateral suppression, as revealed by most electrophysiological studies, makes it highly improbable that a lateral-inhibitory mechanism in the traditional sense is involved, as a neural network involving one or more synapses. It is considered most likely that lateral suppression is associated with processes within the cochlea. However, further research will be needed to specify the site and nature of the mechanism more precisely.

With respect to our second point, namely that lateral suppression manifests itself, psychophysically, as a reduction by a certain *factor*, there appears to be no clear physiological parallel. The main difficulty is a "translation" of this psychophysical rule in terms of neural responses. Some incidental indications concerning a possible agreement are given, for example, in the work of Sachs (1969) in which neural data on two-tone inhibition are described by introducing the concept "inhibitory-multiplier", and perhaps in the work of Hind *et al.* (1967) and Hind (1970), who consider two-tone inhibition as a kind of modulation mechanism. This may illustrate that the neural data at least are not in conflict with this psychophysical aspect of lateral suppression.

Finally, a few words about linearity. The phenomenon of two-tone suppression clearly reveals a non-linear aspect of the neural coding of a stimulus' sound spectrum. On the other hand, it has been demonstrated that the relation between the neural response in case of a broad-band stimulus (noise) and in case of a narrow-band stimulus (pure tone) can be described rather well in terms of a *linear* frequency-selective mechanism. For example, de Boer (1969) has shown, by correlating the temporal pattern of nerve discharges in a single primary fibre in the cat with the temporal wave-form of a noise stimulus, that the unit is stimulated by noise components within a narrow frequency region around the unit's best frequency, which appeared to correspond closely to the selectivity as revealed by the unit's tuning curve. It has been argued (de Boer, 1967) that, in case of a broad-band noise stimulus, lateral suppression cannot contribute to frequency selectivity since the stimulation pattern is essentially flat. However, this is only true in terms of the time-averaged intensity spectrum; in view of the instantaneous nature of the lateral-suppression mechanism, we should also consider the statistical, short-term, fluctuations of the stimulating waveform at the different positions along the basilar membrane. It is not completely inconceivable that lateral suppression, responding to such statistical level differences between adjacent positions, may still contribute to frequency selectivity in case of a broad-band stimulus. Hence, the similarity between the degree of frequency selectivity as revealed by the tuning-curve approach and by the correlation approach does not necessarily imply that lateral suppression plays no role in frequency selectivity, but rather that the frequency-selective mechanism as a whole, including possible effects of lateral suppression, behaves substantially linearily.

Another example of linearity is presented by Wilson and Evans (1971). For the same type of rippled-noise stimuli as we used in chapter 9, it was shown that the degree of ripple resolution in the response of single cochlear-nerve fibres in the cat corresponds closely to that expected on the basis of the unit's spectral selectivity as revealed by its tuning curve.

These examples illustrate that in many electrophysiological experiments the frequency-selective mechanism may be considered, as a first approximation, to be linear. Nonlinearity manifests itself only in extreme cases. How does this relate to our results on frequency selectivity with the pulsation-threshold method, of which we claim that it includes effects of lateral suppression? We have seen, for example, that the data with the rippled-noise masker with different modulation depths could not be accounted for in *detail* by one linear filter. On the other hand, the three filter characteristics in Fig. 9.5

(right column) are nearly identical, indicating that the nonlinear aspect is only of minor importance. To illustrate this further, we shall compare the filter characteristics, derived from pulsation-threshold measurements with the rippled-noise masker, with the frequency selectivity as revealed by pulsation-threshold measurements with a *pure-tone* masker. This is illustrated in Fig. 10.2. The upper panel reproduces the pulsation-threshold data for a pure-tone

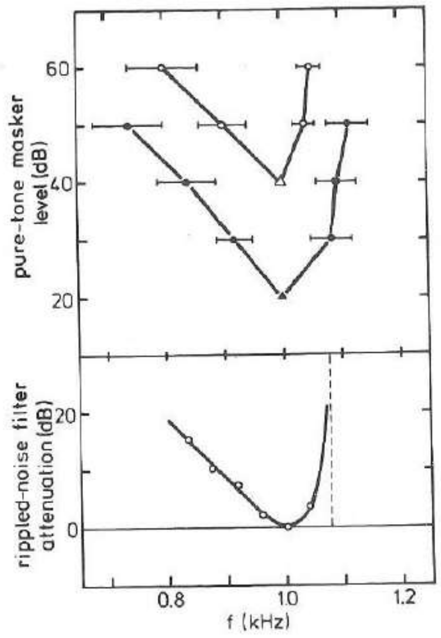


Fig. 10.2. The upper panel reproduces some pulsation-threshold data of Fig. 4.2, with a fixed test tone (triangles) and a variable pure-tone masker. The lower panel reproduces the (average) characteristic of the 'rippled-noise filter', as derived from pulsation-threshold measurements with the rippled-noise masker (Fig. 9.5). Note that the degree of frequency selectivity revealed by the two figures is very similar.

masker of Fig. 4.2. All tones on such a curve have the same effect on the 1000-Hz test tone, indicated by the little triangle; thus, when the frequency-selective mechanism involved is linear, these curves should reflect the filter characteristic at 1000 Hz as derived from the pulsation-threshold measurements with the rippled-noise masker. The lower panel gives the average of the three filter characteristics in Fig. 9.5. (Of course, the negative-going parts of the original intensity-weighting functions cannot be represented on this dB-scale.) It will be seen that the frequency selectivity in case of the pure-tone masker is not dramatically different from that in case of the rippled-noise masker. Thus, also on the basis of such psychophysical data, the frequency-selective mechanism appears to be substantially linear. Nonlinear aspects become manifest only for extreme types of stimuli (two-tone suppression), or when a more precise analysis is carried out, as we did for the rippled-noise masker with different modulation depths.

Summarizing, it appears that the nature of the lateral-suppression mechanism, as revealed by the present psychophysical study, is not in conflict with the way in which lateral suppression manifests itself in electrophysiological studies.

10.3. ADDITIONAL NOTES ON THE PULSATION-THRESHOLD METHOD

10.3.1. Related studies

In view of the close correspondence between pulsation-threshold data and neural-response data, as shown in this study, this method deserves further attention.

The historical background is very limited. In 1957, Thurlow described the effect of continuity for two alternately sounding tones in terms of an auditory "figure-ground" effect: "Under certain conditions the more intense of the two tones is heard as clearly intermittent (somewhat as 'figure'), and the less intense appears to sound continuously (somewhat as 'ground')". It was noted that this effect is only observed when the two tones are brought near together in frequency. The quantitative relation between the frequency difference and the intensity difference between the two tones, for which continuity of the fainter tone is perceived, was investigated by Thurlow and Elfner (1959). This approach comes close to one of our experiments (Fig. 4.2), although the temporal alternation pattern was somewhat different (in our terms 67-msec test-tone bursts alternated with 21-msec masker bursts). Additional experiments were concerned primarily with the question to what extent the perception of *short* interruptions in a fainter sound can be masked by short bursts of a louder sound (*e.g.*, Elfner, 1971). This is not related directly to the present study.

The application of the phenomenon of illusory continuity, for relatively long durations (> 100 msec), within the framework of masker-and-test-tone experiments was introduced by Warren *et al.* (1972) and Houtgast (1971, 1972). Warren *et al.* refer to the effect of continuity as "auditory induction" and they indicate that this is found generally in alternating a fainter and a louder sound when the louder sound has a broader spectrum, which includes that of the fainter sound. This observation lead them to the following general rule: "If there is contextual evidence that a sound may be present at a given time, and if the peripheral units stimulated by a louder sound include those which would be stimulated by the anticipated fainter sound, then the

fainter sound may be heard as present". It is interesting to note that this rule is very similar to the interpretation of the effect of continuity proposed by Houtgast (1971, 1972), which will be further developed in this discussion. However, before aiming at such a theory underlying the effect of continuity and, thus, underlying the pulsation threshold, we should first consider possible effects of the *rate* at which masker and test tone are alternated.

In this study, we always used an alternation rate of 4 Hz: 125-msec masker bursts and 125-msec test-tone bursts. In other studies, different alternation rates were used (*e.g.*, 1.7 Hz as used by Warren *et al.*, 1972), and in one study the alternation rate was treated as a parameter (van Meeteren, 1972). It appears that the effect of continuity of the test tone occurs for a wide range of alternation rates, of which the boundaries are rather vague. At the lower side (slow alternation rates), the phenomenon becomes less clear, it "fades away". For example, for two tones with an alternation rate below 1 Hz, the average subject will perceive "simply" an alternation of two tones without any effect of continuity. Apparently, if the masker does not actually contain the test-tone frequency, the duration of the masker bursts, which can be bridged by the effect of continuity, is limited to a value of the order of 0.5 sec. At the higher side (fast alternation rates), the limit is of a different nature. Temporal gating has an unwanted effect in the spectral domain: generally, the faster the temporal gating, the more the long-term spectra of both the masker and the test tone are broadened. Since this might complicate the interpretation of the results, the alternation rate should be taken as slowly as the effect of continuity permits.

These considerations underly our choice of a 4-Hz alternation rate. It is important to note that the parametric study (van Meeteren, 1972), with alternation rates from 2.5 to 10 Hz, indicated that the results are essentially independent of the alternation rate; the differences which were found could be accounted for, in principle, by the spectral broadening for the faster alternation rates. Thus, it would appear that the results with the 4-Hz alternation rate apply *generally* to the range of alternation rates for which, at one side, the phenomenon of continuity is unambiguous and, at the other side, the effect of spectral broadening can be neglected. (See also Verschuure *et al.*, 1974.)

10.3.2. A theory on the pulsation threshold

In section 10.1.3 of this discussion, the pulsation threshold was considered for the specific case in which the test tone alternated with a masker which *did* contain a frequency component of the same frequency as the test tone. However, this is not a necessary condition for the phenomenon of continuity. Therefore, the very simple theory put forward in section 10.1.3 should be generalized.

Consider the general case of a test tone and a masker, not necessarily containing a frequency component of the same frequency as the test tone, which are alternated according to the pulsation-threshold method (Fig. 10.3 gives the example for a low-pass noise masker). When varying the test-tone

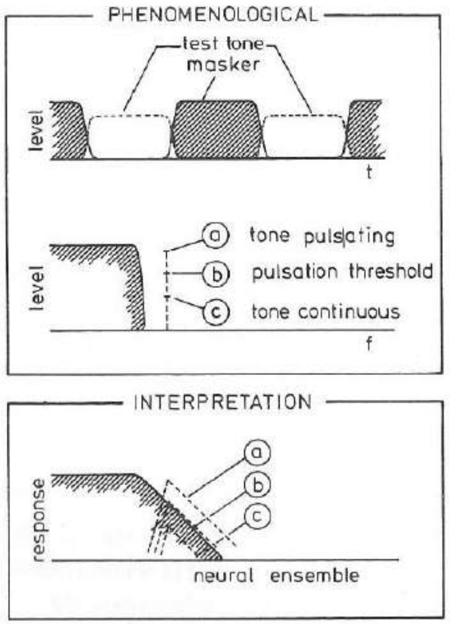


Fig. 10.3. Illustration of the theory underlying the interpretation of pulsation-threshold data. The basic assumption is that the pulsation threshold reflects a condition for which the neural response pattern of the test tone *just* reaches that of the masker.

level, one observes two different ways in which the series of test-tone bursts is perceived: at high levels it sounds as a pulsating tone, in accordance with the alternation rhythm, and at a range of lower levels it sounds as a continuous tone, completely "separated" from the masker bursts. The boarder line between these two level regions is the pulsation-threshold level. The theory about the underlying process refers again to the domain of "neural responses" and is essentially the same as that in section 10.1.3: pulsation threshold is that particular test-tone level at which the neural response to

the test tone has the same "height" as the response to the masker in the neural region associated with the test-tone frequency (see Fig. 10.3). In other words: *a necessary condition for perceiving the pulsating character of the series of test-tone bursts would be that at the transitions 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.

Admittedly, this is an ad-hoc theory; at the other side, it must also be admitted that *given* the phenomenon of continuity and *given* the close correspondence between pulsation-threshold data and neural-response data, it seems a most simple and straight-forward theory.

The theory presented here would imply that the pulsation-threshold method is capable of revealing, in a most simple way, the features of the neural response pattern associated with a masker's sound spectrum. Consider the general case of a masker with some sound spectrum. The pulsation threshold measured as a function of the test-tone frequency can be named the *pulsation-threshold pattern* of that particular masker. There exists an interesting relation between the masker's sound spectrum and its pulsation-threshold pattern: according to the theory put forward here, each *single* tone with a level according to the pulsation-threshold pattern gives the same response in the neurons for which that tone is the "best frequency", as the masker itself. In other words: *the pulsation-threshold pattern constitutes the spectrum of just that hypothetical sound which, when presented to an "ideal" ear in which neurons respond only to spectral components corresponding to their "best frequency" (thus, an "infinite" frequency selectivity and no lateral interactions), gives the same neural-response pattern as the masker itself presented to the "real" ear.* Thus, the pulsation-threshold pattern of a masker reflects the effects of the "nonperfectness" of the actual ear (limited frequency resolution, lateral interactions) on the masker's sound spectrum. It may be considered the "apparent" or "equivalent" spectrum of the masker (at a neural level to which our theory of "continuity of neural response" applies, which remains unspecified).

In conclusion: when accepting the theory on the phenomenon of continuity, the pulsation-threshold pattern of a masker discloses just that what we called in the introduction the "auditory projection" of the masker's sound spectrum; it reflects the "apparent" shape of the masker's sound spectrum after the first stages of auditory processing. It was also mentioned in the introduction that this shape is probably highly relevant for several perceptual attributes of a sound, such as loudness, timbre and, perhaps, pitch. Therefore, in the next chapter, we will proceed a little further along this line and give some examples of this application of the pulsation-threshold method.

CHAPTER 11

AUDITORY SPECTRA OF VARIOUS STIMULI

Summary

This chapter is based on the results and the theory put forward in the preceding chapters: the pulsation-threshold method reveals an adequate picture of the auditory projection of a sound spectrum, including effects of lateral suppression. Three types of stimuli, commonly used in psychoacoustics, are investigated: pure tones, complex tones and vowel-like sounds. Features of the auditory spectra of these stimuli are related to several perceptual aspects, such as loudness, pitch and timbre. Of course, the question remains in how far such perceptual attributes are derived from the auditory projection of a stimulus' sound spectrum. The present results may contribute to the discussion on current theories about hearing processes, as theories on pitch perception and theories on vowel perception.

11.1. GENERAL

The nature of this chapter is somewhat different from that of the preceding ones. In those chapters we typically used a *fixed* test tone and we were interested in the ensemble of maskers which had a specific effect on that test tone. This facilitated the interpretation of the results, since such data reflect the properties of auditory processing for one specific frequency region (generally, such properties are not invariant, but depend on the frequency region considered). Furthermore, the approach was attractive in view of a comparison with neural-response data, referring to the ensemble of stimuli which have a specific effect on a neural unit. In the present chapter, we return to the approach used in the very first experimental chapter (chapter 3), where we touched upon the auditory "Mach band". Thus, we will use a fixed masker and a variable test tone with the explicit purpose to map out the "auditory spectra" of some stimuli which are commonly used in psychoacoustics.

The "auditory spectrum" is defined as follows: *the auditory spectrum of a sound constitutes the spectrum of a hypothetical sound which, when applied to an "ideal" ear in which neurons respond only to spectral components corresponding to their "best frequency" (thus, an "infinite" frequency selectivity and no lateral interactions), would give the same neural-response pattern as the actual sound applied to the actual ear.* (The stage in the auditory channel to which this neural-response pattern applies remains unspeficied.) Thus, the auditory spectrum reflects the effect of the actual ear being *not* "ideal" on the stimulus' sound spectrum. According to the theory presented in the preceding chapter, this auditory spectrum can be obtained by measuring the pulsation-threshold pattern of the sound. Although the foundation of that theory, namely the close correspondence between pulsation-threshold data and neural-response data, could be verified only for a limited number of cases, its validity is plainly accepted in this chapter. The reason for this is twofold: (1) insight in the shape of the auditory spectrum of, for instance, complex tones or vowels is of interest to theories on pitch perception and vowel perception, and (2) besides the pulsation-threshold method, which *might* give a correct picture, there exists no obvious alternative.

With respect to this second point, I will briefly summarize my objections against more traditional masking methods. With regard to direct masking, there are two major points. First, we have seen in this study several cases in which direct-masking data were essentially different from neural-response data. Second, as mentioned in Chapter 4, it is well known that in many cases confounding interactions between a masker and a superimposed test tone occur (beats, combination products), which might play a role in the detection threshold of the test tone. In such cases, the detection threshold is probably not directly related to the neural response to the masker in the neural region associated with the test-tone frequency. Although these two objections can be avoided by using a forward-masking technique, this introduces the problem of a low "sensitivity" (see, for example, Figs. 5.1 and 7.1): according to the model presented in the discussion, forward masking reflects an after effect which, at the moment the test tone is presented, has recovered already considerably. Additionally, both direct masking and forward masking would require extensive reference measurements with single-tone maskers, relating masked threshold to the level of a single-tone masker, in order to transform such masking patterns into an auditory spectrum according to our concept. Finally, apart from these considerations, it should be realized that for *any*

psychophysical method the interpretation of a masking pattern in terms of an auditory spectrum always requires a theory about underlying neural processes and, consequently, remains vulnerable.

Briefly: the pulsation-threshold method appears to be the most acceptable approach to disclose the auditory spectrum of a stimulus.

11.2. PULSATION-THRESHOLD PATTERNS OF PURE TONES

The experiment is illustrated in Fig. 11.1. The pulsation-threshold method was applied, with, again, each fourth test-tone burst left out. The masker

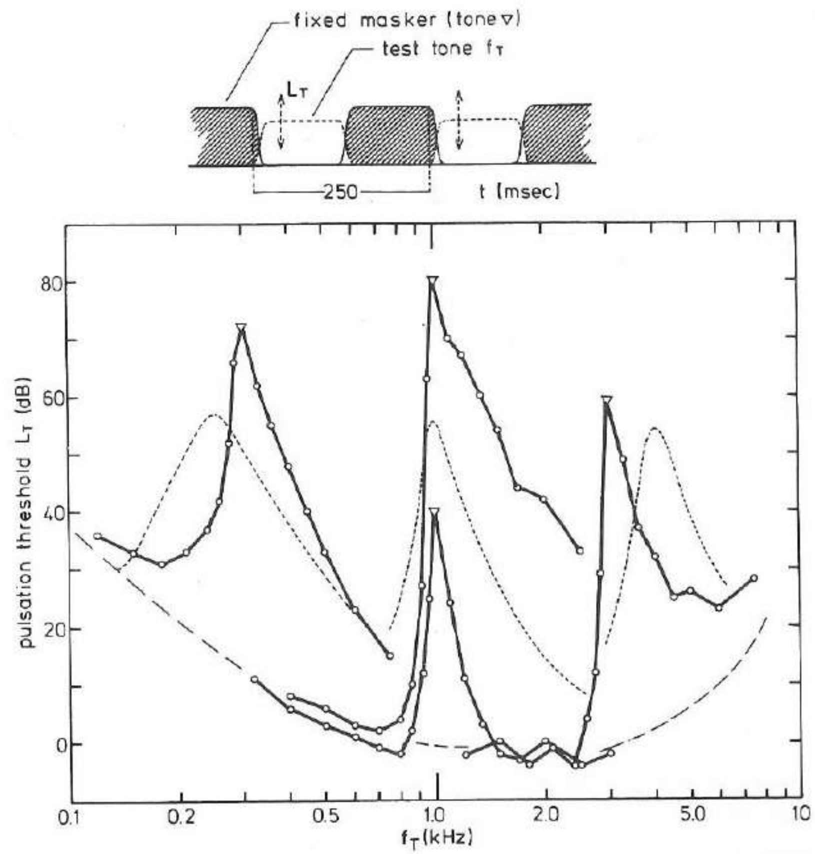


Fig. 11.1. Pulsation-threshold patterns of various pure-tone maskers (indicated by the positions of the triangles in the data graph). The dotted curves are direct-masking patterns for three narrow-band maskers (redrawn from Maiwald, 1967). (Average data of three subjects.)

was a fixed pure tone. Four such maskers were used, of which the levels and frequencies are indicated by the little triangles in the data graph. The test-tone frequency f_T was the independent, and the test-tone level L_T the dependent variable. For each setting of f_T, the subject was instructed to adjust L_T to the highest value at which the pulsating character of the series of test-tone bursts was just not perceived. Each of three subjects made one such adjustment and the median values are presented in Fig. 11.1. The curves connecting the data points are referred to as pulsation-threshold patterns.

According to the theory presented before, each pulsation-threshold pattern constitutes the auditory spectrum of the corresponding pure-tone masker. It is interesting to compare these with the results of more traditional masking procedures, such as direct-masking patterns. The confounding interactions in the case of tone-on-tone masking can be avoided to a large extent by using a narrow-band noise masker, rather than a tonal masker, of not too high a level (Egan and Hake, 1950; Maiwald, 1967). Some direct-masking patterns produced by such narrow-band maskers are reproduced in Fig. 11.1 (from Maiwald, 1967). It will be seen that the main difference between pulsation patterns and direct-masking patterns concerns the steepness of the low-frequency slope. This agrees well with some of the concepts developed in the present study: (1) lateral suppression acts mainly in the direction from higher towards lower frequencies and (2) pulsation threshold does, and direct masking does not, reflect its action. Thus, it would appear that the extreme steepness of the low-frequency slope of the pulsation-threshold patterns reflects the effect of lateral suppression by the peak of the pattern towards lower frequencies. The high-frequency slopes of the pulsation-threshold patterns and the direct-masking patterns are in reasonably good agreement. Also the typical nonlinear behaviour, in that this slope depends on the level of the tonal masker, is traditionally found in direct masking as well (Zwicker, 1958).

The results in Fig. 11.1 should be considered in the light of the important role of the direct-masking patterns in traditional theories on hearing (Zwicker, 1958; Maiwald, 1967). According to these theories, a direct-masking pattern of a narrow-band masker can be transformed, by a vertical shift over a number of dB's, into a so-called "psychoacoustical excitation pattern" of a pure tone; this pattern takes the position of the auditory spectrum according to our concept. Thus, perceptual attributes of a steady-state pure tone, as loudness and pitch, are thought to be associated with features of this excitation pattern. (The excitation pattern also involves a transformation of

the frequency scale, but this is not essential for the present discussion.) The implication of the present study is that this theory is not correct since it is founded on an incorrect assumption: a direct-masking pattern does not give a complete picture of the effects of the first stages of auditory processing because it does not reflect the effect of lateral suppression. According to our concept, such a theory should be built on the auditory spectrum as measured with the pulsation-threshold method.

Admittedly, this new insight does not immediately contribute to our understanding of pure-tone perception. On the contrary, it seems to raise a problem rather than solving one: at one side, we have reason to dispute the relevance of the excitation-pattern model whereas, at the other side, it has been shown (*e.g.*, Zwicker, 1970) that the application of that model appears to be rather successful. For instance, the steepness of the low-frequency slope of the excitation pattern (dB/Hz) might explain the ratio between just-noticeable differences (JND's) in level and frequency of a pure tone. Since this may be considered a strong point in favour of the excitation-pattern model, it is interesting to specify exactly what assumptions are involved in this explanation: (1) the excitation pattern of a pure tone can be derived from the direct-masking pattern of a narrow-band noise, and (2) a JND in frequency or in level of a pure tone reflects a condition in which the level in the excitation pattern, anywhere along the frequency scale, changes by a certain dB-step. This is a good example of a simple framework of assumptions which can account correctly for a psychophysical rule, *i.e.*, the relation between JND's of frequency and level of a pure tone measured under certain conditions. I am in the unfortunate position that I dispute the validity of this model without offering a clear alternative. A similar reasoning applied to the pulsation-threshold pattern rather than to the direct-masking pattern, would imply, for instance, a much smaller JND in frequency. Although there are data in the literature which might suggest that the accuracy in pitch perception is indeed much greater than that explained by the excitation-pattern model (*e.g.*, Rakowski, 1971; Verschuure and van Meeteren, 1974), such arguments cannot lead to a fruitful discussion. It should also be noted that the very concept itself, in relating JND's of level and frequency to the slope of a pattern set up by the tone, leans on the hypóthesis that the pitch of a pure tone is derived from the distribution pattern (place theory) and not from the temporal pattern of neural response (periodicity theory).

In conclusion: according to the present study, the pulsation-threshold pattern of a pure tone reflects the distribution pattern of the neural response to the tone. A (monaural) pure tone has two perceptual attributes: loudness and pitch. An interesting aspect of the relation between loudness and pulsation threshold was discussed in chapter 7. With respect to pitch, the extreme steepness of the slope of the pulsation-threshold pattern might play a role, in the light of the place theory for pitch perception, in the great accuracy in pure-tone pitch perception as reported in the literature.

There is an additional aspect in Fig. 11.1 which deserves further attention. A comparison of the two pulsation-threshold patterns for the 1000-Hz tone at the two different levels suggests that the low-frequency slope is steeper for the high-level tone than for the low-level tone. Similar results have been found in other studies as well (Verschuure *et al.*, 1974). This may have consequences for the interpretation of pulsation-threshold data. This interpretation leans on the concept that pulsation threshold is reached when the neural-response pattern of the test tone reaches that of the masker (see Fig. 10.3), and it was assumed implicitly that this would apply to the *top* of the test-tone pattern. In other words, the test-tone pattern was assumed to be not less peaked, not broader, than that of the masker. However, Fig. 11.1 would suggest that the slope of the pattern of a high-level tone (masker) is steeper than that of a low-level (test) tone. Thus, the pulsation thresholds at the low-frequency side of a pure-tone masker may refer to a situation in which it is not the top of the test-tone pattern, but some other part of it which first reaches the masker's pattern. In that case, the interpretation of that part of the pulsation-threshold pattern in terms of an auditory spectrum, according to our definition, is not correct. We now are in a difficult situation: if the steep slopes of the pulsation-threshold patterns in Fig. 11.1 are interpreted in terms of neural patterns, it can be inferred that the pattern of a (low-level) test-tone is broader than that of the masker and, consequently, that the interpretation is probably incorrect.

Consequently, in view of the ambiguity of the interpretation of pulsation thresholds at the low-frequency side of a pure-tone masker, we should not ascribe too much weight to the exact shape of a pulsation-threshold pattern in that frequency region. Fortunately, this has no important consequences for the present study. In all cases considered, there is no reason to suppose that the finite width of the test-tone pattern has caused serious complications. However, we should be aware of the possibility that in extreme cases, like the present one, this might be the case and lead to incorrect interpretations.

11.3. PULSATION-THRESHOLD PATTERNS OF COMPLEX TONES

This experiment is illustrated in Fig. 11.2. In this case, the masker was a complex tone consisting of the first 10 harmonics of the fundamental frequency 250 Hz. The harmonics had equal amplitudes and were added in sine phase. The results obtained with two such maskers, with a level difference of 30 dB, are presented in the figure. At each setting of the test-tone frequency f_T (in random order), each of three subjects made two adjustments (in separate sessions) of the test-tone level L_T to pulsation threshold. The average values are presented in Fig. 11.2.

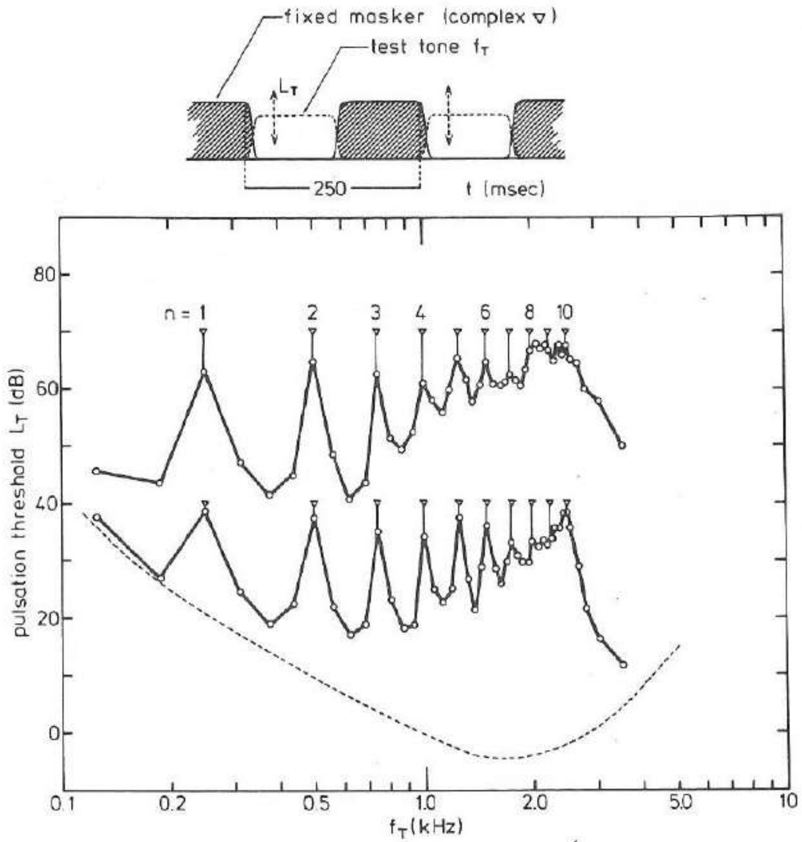


Fig. 11.2. Pulsation-threshold patterns of two complex tones, consisting of the first ten harmonics of the fundamental frequency 250 Hz (positions of the triangles in the data graph). Note that the lower harmonics are associated with distinct peaks in the pulsation-threshold patterns. (Average data of three subjects.)

According to our theory, the pulsation-threshold patterns in Fig. 11.2 constitute the auditory spectra of the two complex tones. It is interesting to note that the peaks of the auditory spectra may fall considerably below the levels of the individual harmonics (little triangles). This may be interpreted as the result of lateral suppression of the individual harmonics by adjacent parts of the complex, which is in line with the concept developed in this study. The most interesting aspect of the auditory spectra in Fig. 11.2 is the extent to which the individual harmonics of the complex are reflected in the auditory spectra. Generally, only the lower harmonics appear to be associated with distinct peaks in the auditory spectra. This result is, of course, not unexpected. The specific contribution of the present result is the quantitative nature of the data. Can these data be related to perceptual aspects of such complex tones?

The many experiments on complex tones are mainly concerned with two issues: (1) the number of (lower) harmonics in a complex that can be perceived as individual tones and (2) the pitch of complex tones. With respect to the first issue, a basic type of experiment was performed by Plomp and Mimpen (1968). That experiment was set up in such a way that the score of a two-alternative recognition task (ranging from 100% to 50%) was related to the audibility of an individual harmonic, ranging from "perfect" to "unaudible". For the complex tone with a fundamental frequency of 250 Hz, their data are reproduced in Fig. 11.3, upper panel. Although there are some minor differences between the complex tone used in that experiment and the one used in the present experiment (the harmonics were added in cosine phase, the amplitude of each harmonic was adjusted to equal sensation level and the number of harmonics extended beyond $n = 10$), it seems interesting to compare their results with features of the present auditory spectra. We determined the relative "height" of the peak in the auditory spectra, associated with each individual harmonic n, by taking the difference between the pulsation threshold for $f_T = n.250$ Hz and the average for $f_T = (n \pm \frac{1}{2}).250$ Hz. This "peak-to-valley ratio" for n from 1 to 9 is plotted in Fig. 11.3, lower panel. (The four symbols refer to four different levels L_c of the complex, of which only two, the 70-dB and 40-dB conditions, have been represented in Fig. 11.2.) For the lower harmonics, the hearing threshold has a marked influence on the peak-to-valley ratio, especially for the complexes at the lower levels. However, the interesting aspect of Fig. 11.3 is the general correspondence between the graphs in the upper and lower panels. (The upper panel refers to a complex tone with a sensation level of 60 dB, which agrees perhaps best with the 55-

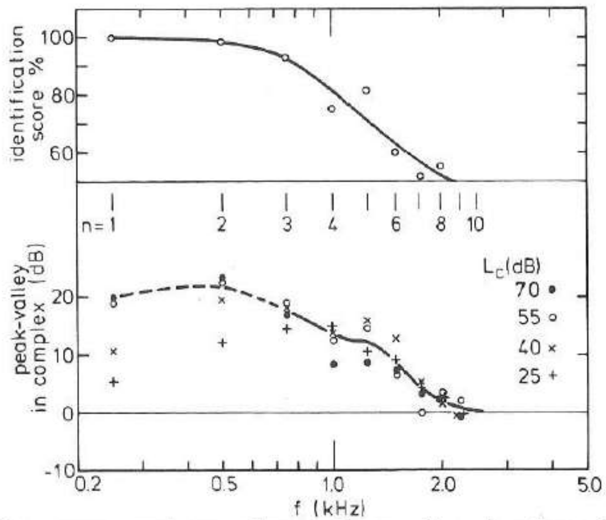


Fig. 11.3. The upper panel reproduces the results of a forced-choice experiment, in which the identification score (from 100% to 50%) is related to the audibility of each individual harmonic (from 'perfect' to 'inaudible') in a complex tone with a fundamental frequency of 250 Hz (redrawn from Plomp and Mimpen, 1968). The lower panel presents the 'height' of the peaks in the pulsation pattern of the complex tone for each individual harmonic. The parameter L_c is the level of the harmonics in the complex tone. Fig. 11.2 represents the pulsation-threshold patterns for only two of the four conditions investigated (L_c = 70 dB and L_c = 40 dB).

dB condition in the lower panel; this would suggest a very detailed correspondence, but we are not inclined to ascribe much weight to that.) The similarity suggests that the audibility of an individual harmonic in a complex tone is related to the "height" of the peak associated with that harmonic in the auditory spectrum of the complex. It supports the view that the pulsation-threshold pattern gives a relevant picture of the auditory projection of a sound spectrum.

A second main issue of psychophysical experiments on complex tones concerns pitch. In this field, a great variety of experiments are performed and the discussion and interpretation of the results is focussed mainly on one central theme: is the pitch of a complex tone derived from the neural projection of the *spectrum* of the complex tone (place or "pattern" theory) or from the neural correlate of the *periodicity* of the complex tone (periodicity theory)? At the present time, the discussions on this theme are still very animated and it would go beyond the scope of this study to take part in that. Briefly, one might say that, after some decades during which the periodicity theory found general favour, pattern theories (see also de Boer, 1956) are rapidly gaining grounds as very serious alternatives (*e.g.*, Houtsma and Goldstein, 1972; Terhardt, 1972 a, b; Wightman, 1973;van de Brink,1974). Within the framework of the present study, the auditory spectrum of a complex tone, as revealed by the pulsation-threshold pattern,reflects the very pattern to which such pattern theories might apply. Can some features of the auditory spectra be related, then, to psychophysical results of pitch experiments with complex tones? Let us consider, briefly, two aspects: the dominance region and the existence region.

The dominance region refers to the harmonics in a complex which appear to have a major contribution to the pitch of a complex tone. For complexes with a fundamental frequency of about 250 Hz, this dominance region is located around the fourth harmonic, thus in the frequency region around 1000 Hz (Ritsma, 1967; Ritsma *et al.*, 1967; Plomp, 1967). In terms of the pattern theory, one would expect this dominance region to be associated with that frequency region in which a change of the fundamental frequency of the complex leads to the most significant change in the auditory spectrum of the complex. Realizing that a change of the fundamental frequency accomplishes a change of the frequency of each individual harmonic by a certain factor, the dominant region would be associated with that frequency region at which the auditory spectra in Fig. 11.2 (with a logarithmic frequency scale) have the most prominent and sharp peaks. It appears that this notion is not in conflict with the location of the dominance region around 1000 Hz.

The existence region refers to the highest harmonics in a complex which can still evoke a low pitch. According to a basic study of Ritsma (1962), this limit, for a fundamental frequency of 250 Hz, is reached at harmonic numbers around $n = 16$. However, the interpretation of such experiments with complex tones, consisting of only a few higher harmonics, is often confounded by combination tones, generated in the ear, such that the "effective" spectrum of the stimulus may include a number of lower harmonics beyond the "physical" spectrum. Recent experiments (Smoorenburg, 1970; Houtsma and Goldstein, 1972; Moore, 1973) indicate that the "effective" spectrum of a complex tone (thus, including possible aural combination tones) should contain some harmonics below about $n = 10$ in order to evoke a low pitch. In terms of a pattern theory, one would expect this limit to be associated with that frequency region at which the peak-to-valley ratio for the individual harmonics in the auditory spectrum has essentially reduced to zero dB. It should be noted that, theoretically, one might expect that the peak-to-valley ratio for high harmonics in the auditory spectrum *approaches* zero dB, but never *reaches* it. Hence, a quantitative comparison of this limit with the peak-to-valley ratio in the auditory spectrum requires a specification of what is meant by "essentially reduced to zero": is the limit of pitch perception associated with a peak-to-valley ratio of the order of 1 dB, or perhaps 0.1 dB? Given this fundamental uncertainty, it would seem that the data on the peak-to-valley ratio, as presented in Fig. 11.3, agree reasonably well with a limit of about $n = 10$.

In conclusion: the pulsation-threshold patterns of complex tones, interpreted as the auditory spectra of such stimuli, correspond well to psychophysical data on the audibility of the individual harmonics. Furthermore, within the framework of a pattern theory for pitch perception, the features of the auditory spectra agree with the dominance region and the existence region for pitch perception of complex tones.

11.4. PULSATION-THRESHOLD PATTERNS OF VOWEL-LIKE SOUNDS

This experiment is illustrated in Fig. 11.4. Two maskers were considered. Each masker consisted of the first 32 harmonics of the fundamental frequency 125 Hz, all added in sine phase. The two line spectra, which are presented in Fig. 11.4, differed only with respect to the exact locations of three formants (positions along the frequency scale where the envelope of the line spectrum has a maximum). The first sound, labelled "*a*", had its formants at 720, 1180 and 2600 Hz, and the second sound, labelled "ε", at 580, 1700 and 2460 Hz. The normal pulsation-threshold method was used. The test-tone frequency f_T was set randomly at a frequency corresponding to one of the harmonics from $n = 2$ to $n = 26$ and each of three subjects made two adjustments (in separate sessions) of the test-tone level L_T to pulsation threshold. The average values are presented in Fig. 11.4.

It should be noted that, in this case, the test-tone frequency was always set at a frequency component of the masker. Consequently, the curves connecting the data points should be interpreted here as auditory *line* spectra, reflecting the effect of the first stages of auditory processing on the shape of the original line spectra.
These results may play a role in the discussion on theories about vowel perception. There exist various theories on the auditory mechanisms underlying vowel recognition and discrimination. Commonly, it is assumed that the audi-

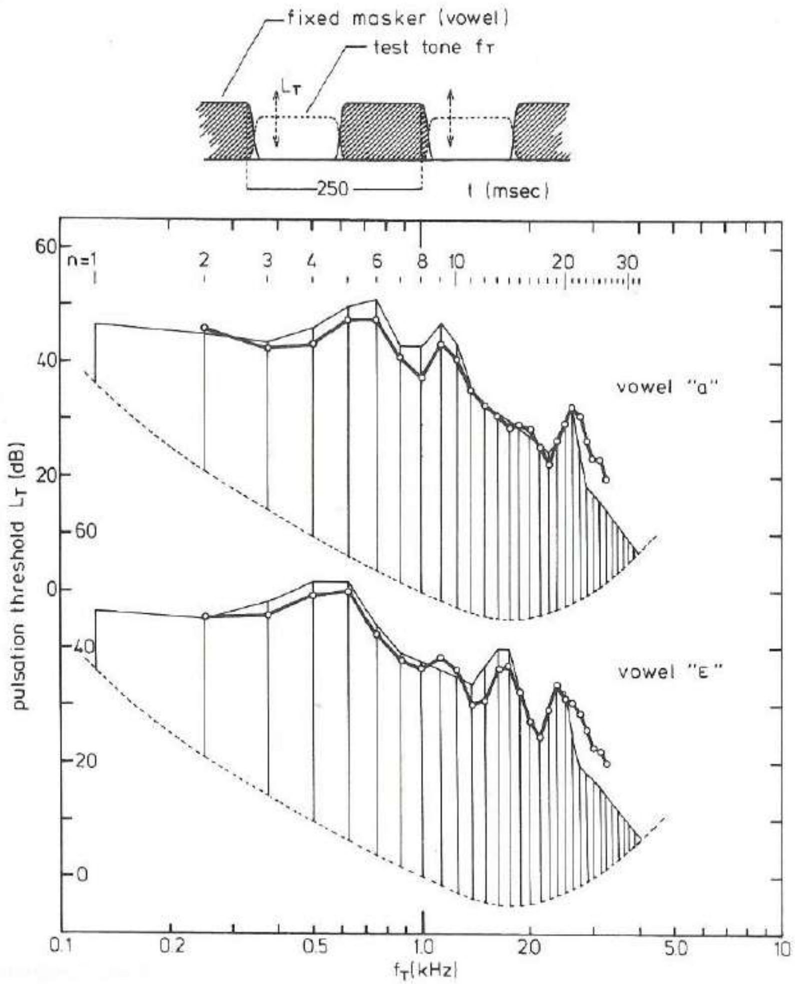


Fig. 11.4. Pulsation-threshold patterns of two vowel-like sounds with a fundamental frequency of 125 Hz. The line spectra of the two 'vowels', with a typical formant structure, are indicated in the data graph. Note that the test tone was always centred at one of the harmonics (from n = 2 to n = 26). (Average data of three subjects.)

tory projection of a vowel's line spectrum forms the basic source of information, but with respect to the subsequent processing there exists a major difference of opinion. According to one theory, the positions of the formants play a major role in vowel perception (*e.g.*, Chistovich, 1971) and, consequently, it is assumed that the processing of the auditory spectra is primarily a process of "formant extraction". According to another theory (*e.g.*, Pols *et al.*, 1969), the whole shape of the auditory spectrum is considered to be important, without any specific weight ascribed to the peaks, and, consequently, the processing of the auditory spectra is assumed to take the form of a more general type of "pattern processing".

The data in Fig. 11.4 indicate that the formants in the original line spectra are clearly reflected in the auditory line spectra. It should be noted that, in general, any effect of lateral suppression, by which weaker parts in a spectrum are suppressed by adjacent stronger parts, stresses the significance of the formants. Does this imply that the present study, which underlines the role of such a lateral-suppression mechanism, would favour the "formant-extraction" theory as the more relevant one? The answer is negative. Admittedly, the effect of lateral suppression is such that the positions of maxima in a spectrum are marked more clearly. On the other hand, one also might say that, generally, lateral suppression helps to preserve, and perhaps even to enhance, the original shape of a stimulus' line spectrum. This favours a general type of "pattern processing" as much as it would favour a specific "formant-extraction" type of process.

There is one aspect which deserves further attention. It is often assumed that the relation between the original line spectra and the auditory line spectra might be approximated by 1/3 octave-band filtering, in line with traditional ideas about the filter properties of the ear (Plomp *et al.*, 1967). Accordingly, within the framework of "pattern-processing" theories, perceptual differences between two vowels, or between two sounds in general, are related to the spectral differences as measured in a number of adjacent 1/3 octave-band filters (Plomp, 1970). The present data indicate that this does not account correctly for the high degree of spectral resolution as revealed by the auditory spectra. This is illustrated in Fig. 11.5, in which the spectral differences between the vowel-like sounds "ε" and "a" are presented. The upper panel gives the original level difference for each of the 32 harmonics. The middle panel gives the difference between the pulsation-threshold values measured at each of the harmonics from n = 2 to n = 26. The lower panel gives the level difference as measured with a 1/3 octave-band filter when centred, successively, at each of the harmonics from n = 2 to n = 28. It can be seen that the difference between the two auditory line spectra (middle panel) contains considerably more details of the original spectral difference (upper panel) than after a process of 1/3 octave-band filtering (lower panel).

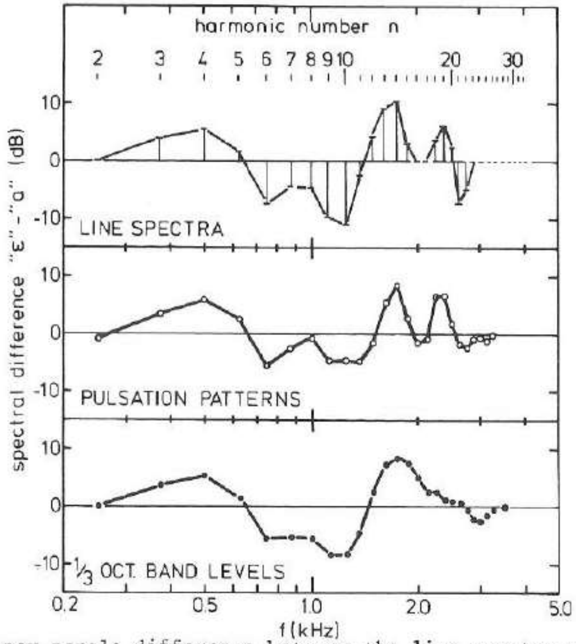


Fig. 11.5. Upper panel: difference between the line spectra of the two 'vowels' in Fig. 11.4. Middle panel: difference between the two corresponding pulsation-threshold patterns. Lower panel: difference between the levels measured with a 1/3-octave band-pass filter when centred at each individual harmonic successively. Note that the middle panel mirrors the original spectral differences in greater detail than after spectral analysis with a 1/3-octave band-pass filter.

In conclusion: the pulsation-threshold patterns of vowel-like sounds, interpreted as auditory line spectra, are of interest to theories about vowel perception. The auditory line spectra reveal a high degree of resolution of the shape of the original line spectra. Although the present study strongly suggests that lateral suppression plays a role here, this should not be interpreted as being in favour of a "formant-extraction" type of theory exclusively. Finally, it appears that a simulation of the effect of the first stages of auditory processing by 1/3 octave-band filtering can serve only as a very first approximation, which does not account for the high degree of spectral resolution revealed by the pulsation-threshold patterns.

SUMMARY AND CONCLUSIONS

The present study should be considered in the light of the common concept that the sound spectrum of a steady-state auditory stimulus is represented as a specific stimulation pattern along the basilar membrane in the cochlea, and also as a specific distribution pattern of neural activity in the auditory pathway. The properties of this auditory projection of a stimulus' sound spectrum are considered to play a fundamental role in perception (as in the perception of loudness, timbre and, perhaps, pitch). Therefore, it is of interest to study the features of that projection.

The results of many electrophysiological experiments in animals suggest that the auditory projection of a sound spectrum, in a way comparable with the visual projection of a light-dark distribution, is subjected to both *unsharpness* (such that fine details of the original shape of a sound spectrum are not resolved), and, in some cases, *sharpening* (such that the original spectral contrasts are enhanced). This latter phenomenon is often understood as an effect of lateral inhibition or *lateral suppression*: when "weaker" parts in a pattern are suppressed by adjacent "stronger" parts, the contours of the projection may become more pronounced than the contours of the original frequency spectrum. The main issue of the present study is to gain insight into the role of sharpening in the auditory projection of a sound spectrum.

The experimental approach is based on the usual concept that *masking experiments*, with a test tone of variable frequency, can be used to investigate the auditory projection of a masker's sound spectrum: the effect of the masker on the threshold of the test tone, as a function of the test-tone frequency (a *masking pattern*), is often interpreted as reflecting the auditory projection of the masker's sound spectrum.

In this study, three different measuring techniques are applied for investigating the auditory projection of a masker. *Direct masking* is a very traditional method. The test tone is superimposed on the masker and the criterion always refers to the detectability of the test tone. *Forward masking* is also a commonly used technique. A brief test tone is presented shortly after the termination of the masker and, again, the criterion refers to the detectability of the test tone. The *pulsation-threshold* method is a relatively new method. The masker and the test tone are presented in continuous alternation, without silent intervals, with an alternation rate of 4 Hz. The criterion refers to the way in which the test tones are perceived: either as a *pulsating* tone corresponding to the alternation rhythm (for high test-tone levels), or as a *continuous* tone (for low test-tone levels).

In chapter 3, these three methods are applied to investigate the auditory projection of a masker consisting of sharply filtered low-pass noise. In this case, lateral suppression would manifest itself as an enhancement at the edge of the noise spectrum, in a way similar to the well known *Mach bands* in vision. The results are typical of the present study: whereas the traditional direct-masking pattern shows no such edge effect, the patterns obtained with the two other methods clearly do so.

The purpose of the subsequent experiments (up to chapter 9) is twofold: (1) to compare the results of the three different types of masking experiments with published data on the neural response to various auditory stimuli, and (2) to investigate maskers with sound spectra which are interesting from the point of view of lateral suppression. (The main results are summarized in Table I on page 75.) It appears that the two methods in which the masker and the test tone are presented *non*simultaneously (forward masking and pulsation threshold) consistently reveal effects of lateral suppression: the masking effectiveness in a frequency region may *decrease* when the intensity of the masker in an adjacent (higher) frequency region is *increased*. These results, especially those obtained with the pulsation-threshold method, agree well with neural-response data. This suggests that this method gives a correct picture of the auditory projection of a masker's sound spectrum. On the other hand, the results obtained with direct masking consistently show *no* effects of lateral suppression.

In the discussion in chapter 10 it is indicated that the results suggest a lateral-suppression mechanism with the following properties: (1) the suppression acts mainly in the direction from higher towards lower frequencies, (2) suppression in some frequency region does not affect the signal-to-noise *ratio* for a test tone superimposed on the masker in that frequency region,

and (3) the suppression acts almost instantaneously. The second property would explain why effects of lateral suppression cannot be revealed by traditional direct-masking methods.

If indeed the pulsation-threshold method *does*, and the direct-masking method does *not* reveal the effect of lateral suppression, the degree of sharpening by lateral suppression can be derived from the difference between the results of these two experimental techniques. According to this, the results presented in chapter 9, obtained with a rippled-noise masker (noise of which the intensity as a function of frequency is sine-wave shaped), indicate that lateral suppression substantially increases the frequency selectivity in the auditory projection of a sound spectrum. When this is expressed in terms of the bandwidth of an auditory filter, it appears that the bandwidth, derived from pulsation-threshold data, is about a factor of two *smaller* than the classical "critical bandwidth" derived from the direct-masking data. It is also shown that this high degree of frequency selectivity, including the effect of lateral suppression, can be described roughly in terms of a (narrow) *linear* filter. Nonlinearities, which are encountered at several places in this study, can be considered as second-order effects.

Briefly, the present study indicates that lateral suppression contributes substantially to the preservation of spectral contrasts in the auditory projection of a sound spectrum and, furthermore, that the pulsation-threshold method, in contrast to the traditional direct-masking method, gives a correct picture of this projection.

In the final chapter, this new insight is applied to investigate the auditory projection of several stimuli which are commonly used in psychoacoustics: pure tones, complex tones and vowel-like sounds. It is shown that various traditional psychoacoustical data, as, for example, the number of (lower) harmonics in a complex tone which can be heard individually, can be related to specific features of the auditory projection of the stimulus. The results are discussed in the light of the fundamental question in how far the various perceptual attributes of a sound (timbre, pitch) can be understood as being derived from the auditory projection of the sound spectrum. The results obtained with the pulsation-threshold method contribute to the discussions on several current theories in this field, such as theories on pitch perception and theories on vowel perception.

SAMENVATTING EN KONKLUSIES

De onderhavige studie moet gezien worden tegen de achtergrond van de algemeen aanvaarde gedachte dat, bij een auditieve stimulus met een konstant karakter, het geluidspektrum wordt "afgebeeld" in de vorm van een specifiek stimulatiepatroon langs het basilair membraan in de cochlea, en ook als een specifiek distributiepatroon van zenuwaktiviteit in de gehoorbaan. Inzicht in de eigenschappen van deze afbeelding is van belang daar deze eigenschappen een fundamentele rol spelen bij verschillende aspekten van de geluidwaarneming, zoals bij de perceptie van luidheid, klankkleur en misschien ook toonhoogte.

Uit velerlei elektrofysiologische dierexperimenten kan worden afgeleid dat bij de auditieve afbeelding van een geluidspektrum, enigszins in analogie met de visuele afbeelding van een licht-donkerverdeling, enerzijds een zekere mate van *onscherpte* optreedt (waarbij fijne details van de oorspronkelijke vorm van een geluidspektrum verloren gaan) terwijl anderzijds in bepaalde gevallen juist een *opscherping* kan optreden (waarbij de oorspronkelijk aanwezige spektrale kontrasten kunnen worden versterkt). Dit laatste verschijnsel wordt veelal beschreven als een effekt van *laterale suppressie* of zijdelingse onderdrukking: doordat "zwakke" gedeelten in een patroon worden onderdrukt door naastliggende "sterkere" gedeelten kan de uiteindelijke afbeelding een overdreven beeld van de oorspronkelijke vorm van het geluidspektrum geven. Het centrale thema van de huidige studie is in welke mate opscherping zich manifesteert in de auditieve afbeelding van een geluidspektrum.

De experimentele benadering sluit aan bij de, in de psychoakoestiek algemeen aanvaarde gedachte dat met behulp van *maskeermetingen*, met als testsignaal een zuivere toon van variabele frekwentie, de auditieve afbeelding van het frekwentiespektrum van de maskeerder kan worden afgetast. De drempel-

verhoging van de testtoon als funktie van diens frekwentie (het *maskeerpatroon*) wordt veelal geacht een afspiegeling te zijn van de auditieve afbeelding van de maskeerder.

In deze studie worden drie verschillende typen maskeermetingen naast elkaar toegepast. *Direkte maskering* is een klassieke methode waarbij de testtoon op de te onderzoeken maskeerder wordt gesuperponeerd en waarbij als meetkriterium geldt het al dan niet waarnemen van de testtoon. *Voorwaartse maskering* is een eveneens niet ongebruikelijke methode waarbij een (korte) testtoon vlak na de maskeerder wordt gepresenteerd en waarbij eveneens het al dan niet kunnen waarnemen van de testtoon als meetkriterium wordt gehanteerd. De *pulsatiedrempel*methode is een nieuwe methode waarbij de maskeerder en de testtoon in voortdurende snelle afwisseling, zonder stille intervallen, worden gepresenteerd (met een afwisselingsritme van 4 Hz) met als meetkriterium of hierbij de testtoon als een *pulserende* toon, overeenkomstig het afwisselingsritme, dan wel als een doorlopende, kontinue toon wordt waargenomen.

In hoofdstuk 3 worden deze meetmethoden toegepast voor het "aftasten" van een maskeerder die bestaat uit scherp gefilterde ruis, zodat het frekwentiespektrum een scherpte knik vertoont. De invloed van laterale suppressie zou hierbij tot uitdrukking kunnen komen, in analogie met de bekende Machbanden bij de visuele waarneming, als een opscherping ter plaatse van de knik in het frekwentiespektrum. De resultaten zijn typerend voor die van de verdere metingen: het klassieke maskeerpatroon, verkregen met direkte maskering, vertoont geen opscherping terwijl de twee andere methoden wel duidelijke effekten van opscherping te zien geven.

De verdere experimenten (t/m hoofdstuk 9) hebben een tweeledig doel: (1) het vergelijken van de resultaten van de drie typen maskeermetingen met uit de literatuur beschikbare gegevens betreffende de neurale responsies voor verschillende stimuli en (2) het aftasten van maskeerders met frekwentiespektra die interessant lijken met het oog op effekten van laterale suppressie. (De belangrijkste resultaten zijn samengevat in de tabel op blz. 75.) Hieruit komt het volgende beeld naar voren. De resultaten van metingen volgens de methode waarbij de maskeerder en de testtoon *niet* gelijktijdig worden aangeboden (voorwaartse maskering en pulsatiedrempel) geven zeer systematisch effekten van laterale suppressie te zien: de maskerende werking in een bepaald frekwentiegebied kan *afnemen* door het *verhogen* van de intensiteit van de maskeerder in een naastliggend (hoger) frekwentiegebied. Deze resultaten, vooral die volgens de pulsatiedrempelmethode, sluiten goed aan bij beschikbare elektrofysiologische gegevens. Dit wijst er op dat deze metingen een juist beeld geven van de auditieve afbeelding van het geluidspektrum van de maskeerder. De resultaten van de klassieke methode van direkte maskering geven daarentegen in alle gevallen *geen* effekten van laterale suppressie te zien.

In de diskussie in hoofdstuk 10 wordt uiteengezet dat deze resultaten begrepen kunnen worden in het kader van laterale suppressie met de volgende eigenschappen: (1) de suppressie werkt voornamelijk in de richting van hogere naar lagere frekwenties, (2) suppressie van een bepaald frekwentiegebied tast de signaal-ruis*verhouding* voor een testtoon, die in dat frekwentiegebied op de maskeerder wordt gesuperponeerd, niet aan en (3) de suppressie werkt nagenoeg *momentaan*. Op grond van de tweede eigenschap kan worden begrepen dat effekten van laterale suppressie zich niet manifesteren bij metingen volgens de klassieke methode van direkte maskering.

Indien het juist is dat metingen volgens de pulsatiedrempelmethode de invloed van laterale suppressie *korrekt* weergeven en de metingen volgens direkte maskering deze invloed *niet* weergeven, dan kan de mate van opscherping door laterale suppressie worden afgeleid uit het verschil tussen deze twee metingen. Wanneer de metingen in hoofdstuk 9 met de "geribbelde ruis" (ruis waarvan de intensiteit als funktie van de frekwentie een sinusvormig verloop heeft) op deze wijze worden geïnterpreteerd, blijkt dat het frekwentie-oplossend vermogen bij de auditieve afbeelding van een frekwentiespektrum belangrijk toeneemt door de invloed van laterale suppressie. Als dit wordt uitgedrukt in termen van bandbreedte, dan is de bandbreedte die kan worden afgeleid uit de pulsatiedrempelmetingen een faktor twee *smaller* dan de klassieke "kritieke bandbreedte" zoals die volgt uit de metingen volgens de methode van direkte maskering. Hierbij blijkt bovendien dat de frekwentieselektiviteit, inklusief de bijdrage van laterale suppressie, in grote trekken beschreven kan worden als het effekt van een (smal) *lineair* filter. De niet-lineariteiten, die zich alleen bij een nauwkeurige analyse manifesteren, kunnen worden beschouwd als tweede-orde effekten.

Samenvattend leidt de huidige studie tot het inzicht dat opscherping door laterale suppressie wezenlijk bijdraagt tot het behoud van de spektrale kontrasten in de auditieve afbeelding van een geluidspektrum en, bovendien, dat de metingen volgens de pulsatiedrempelmethode, in tegenstelling tot de klassieke methode volgens direkte maskering, hiervan een juist beeld geven.

In het laatste hoofdstuk wordt dit nieuw verworven inzicht toegepast voor het onderzoeken van de auditieve afbeelding van stimuli die een belangrijke

rol spelen in de psychoakoestiek: zuivere tonen, komplexe tonen en klinkerachtige signalen. Het blijkt dat enkele traditionele gegevens uit de psychoakoestiek, zoals bijvoorbeeld het aantal (lage) harmonischen van een komplexe toon dat afzonderlijk kan worden waargenomen, goed aansluiten bij de specifieke eigenschappen van de auditieve afbeelding van de geluidspektra van dergelijke stimuli. Hierbij speelt steeds de fundamentele vraag op de achtergrond in hoeverre de auditieve afbeelding van het frekwentiespektrum van een stimulus een rol speelt bij de perceptie, zoals bij toonhoogtewaarneming of bij de diskriminatie van klinkers. Het "aftasten" van de auditieve afbeelding volgens de pulsatiedrempelmethode kan een bijdrage leveren tot de diskussie betreffende de verschillende theorieën op dit gebied.

REFERENCES

ARTHUR, R.M., PFEIFER, R.R. and SUGA, N. (1971). "Properties of 'Two-Tone Inhibition' in Primary Auditory Neurons", J. Physiol. *212*, 593-609.

BEKESY, G. von (1958). "Funneling in the Nervous System and its Role in Loudness and Sensation Intensity on the Skin", J. Acoust. Soc. Amer. *30*, 399-412.

BEKESY, G. von (1959). "Similarity between Hearing and Skin Sensation", Psychol. Rev. *66*, 1-22.

BEKESY, G. von (1960). "Frequency Analysis and the Law of Contrast", in *Experiments on Hearing* (Mc Graw-Hill Book Co., New York).

BEKESY, G. von (1963). "Hearing Theories and Complex Sounds", J. Acoust. Soc. Amer. *35*, 588-606.

BOER, E. de (1956). *On the Residue in Hearing*, Doctoral Thesis, University of Amsterdam.

BOER, E. de (1967). "Correlation Studies Applied to the Frequency Resolution of the Cochlea", J. Audit. Res. *7*, 209-217.

BOER, E. de (1969). "Encoding of Frequency Information in the Discharge Patterns of Auditory Nerve Fibers", Intern. Audiol. *8*, 547-556.

BRINK, G. van den (1974). Study to be published.

CARTERETTE, F.C., FRIEDMAN, M.P. and LOVELL, J.D. (1969). "Mach Bands in Hearing", J. Acoust. Soc. Amer. *45*, 986-998.

CARTERETTE, F.C., FRIEDMAN, M.P. and LOVELL, J.D. (1970). "Mach Bands in Auditory Perception", in *Frequency Analysis and Periodicity Detection in Hearing*, R. Plomp and G.F. Smoorenburg, Eds.(Sijthoff, Leiden, the Netherlands).

CHISTOVICH, L.A. (1971). "Auditory Processing of Speech Stimuli-Evidence from Psychoacoustics and Neurophysiology", Proc. 7th I.C.A., Budapest, 21 G 1.

EGAN, J.P. and HAKE, W. (1950). "On the Masking of a Simple Auditory Stimulus", J. Acoust. Soc. Amer. *22*, 622-630.

ELFNER, L.F. (1971). "Continuity of Alternately Sounded Tonal Signals in a Free Field", J. Acoust. Soc. Amer. *49*, 447-449.

GREENWOOD, D.D. (1971). "Aural Combination Tones and Auditory Masking", J. Acoust. Soc. Amer. *50*, 502-543.

GREENWOOD, D.D. and GOLDBERG, J.M. (1970). "Response of Neurons in the Cochlear Nuclei to Variations in Noise Bandwidth and to Tone-Noise Combinations", J. Acoust. Soc. Amer. *47*, 1022-1040.

HIND, J.E. (1970). "Two-Tone Masking Effects in Squirrel Monkey Auditory Nerve Fibers", in *Frequency Analysis and Periodicity Detection in Hearing*, R. Plomp and G.F. Smoorenburg, Eds. (Sijthoff, Leiden, the Netherlands).

HIND, J.E., ANDERSON, D.E., BRUGGE, J.F. and ROSE, J.E. (1967). "Coding of Information Pertaining to Paired Low-Frequency Tones in Single Auditory Nerve Fibers of the Squirrel Monkey", J. Neurophysiol. *30*, 794-816.

HOUTGAST, T. (1971). "Psychophysical Evidence for Lateral Inhibition in Hearing", 7th I.C.A., Budapest, 24 H 15.

HOUTGAST, T. (1972). "Psychophysical Evidence for Lateral Inhibition in Hearing", J. Acoust. Soc. Amer. *51*, 1885-1894.

HOUTGAST, T. (1973). "Psychophysical Experiments on 'Tuning Curves' and 'Two-Tone Inhibition'", Acustica *29*, 168-179.

HOUTGAST, T. (1974). "Lateral Suppression and Loudness Reduction of a Tone in Noise", to be published in Acustica.

HOUTSMA, A.J.M. and GOLDSTEIN, J.L. (1972). "The Central Origin of the Pitch of Complex Tones: Evidence from Musical Interval Recognition", J. Acoust. Soc. Amer. *51*, 520-529.

KATE, J.H. ten, BILSEN, F.A., RAATGEVER, J. and BUUNEN, T.J.F. (1974). "Single Unit Responses in Acoustic Nuclei of Cat to Noise and Its Attenuated Repetition", submitted to 8th I.C.A., London.

KIANG, N.Y.S. (1965). *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve* (M.I.T. Press, Cambridge, Mass.).

LEVITT, H. (1971). "Transformed Up-Down Methods in Psychoacoustics", J. Acoust. Soc. Amer. *49*, 467-477.

LIFF, H.J. and GOLDSTEIN Jr., M.H. (1970). "Peripheral Inhibition in Auditory Fibers in the Frog", J. Acoust. Soc. Amer. *47*, 1538-1547.

MAIWALD, D. (1967a). "Beziehungen zwischen Schallspektrum, Mithörschwelle und der Erregung des Gehörs", Acustica *18*, 69-80.

MAIWALD, D. (1967b). "Ein Funktionsschema des Gehörs zur Beschreibung der Erkennbarkeit kleiner Frequenz- und Amplitudenänderungen", Acustica *18*, 81-92.

MEETEREN, A.A. van (1972). Unpublished data.

MØLLER, A.R. (1970a). "Unit Responses in the Cochlear Nucleus of the Rat to Noise and Tones", Acta Physiol. Scand. *78*, 289-298.

MØLLER, A.R. (1970b). "Studies on the Damped Oscillatory Response of the Auditory Frequency Analyzer", Acta Physiol. Scand. *78*, 299-314.

MØLLER, A.R. (1972). "Coding of Sounds in Lower Levels of the Auditory System" Quarterly Reviews of Biophysics *5*, I, 59-155.

MOORE, B.C.J. (1973). "Some Experiments Relating to the Perception of Complex Tones", Quart. J. Exper. Psychol. *25*, 451-475.

NOMOTO, M., SUGA, N. and KATSUKI, Y. (1964). "Discharge Patterns and Inhibition of Primary Auditory Nerve Fibers in the Monkey", J. Neurophysiol. *27*, 768-787.

PLOMP, R. (1967). "Pitch of Complex Tones", J. Acoust. Soc. Amer. *41*, 1526-1533.

PLOMP, R. (1970). "Timbre as a Multidimensional Attribute of Complex Tones", in *Frequency Analysis and Periodicity Detection in Hearing*, R. Plomp and G.F. Smoorenburg, Eds. (Sijthoff, Leiden, the Netherlands).

PLOMP, R. and MIMPEN, A.M. (1968). "The Ear as a Frequency Analyzer. II", J. Acoust. Soc. Amer. *43*, 764-767.

PLOMP, R., POLS, L.C.W. and GEER, J.P. van de (1967). "Dimensional Analysis of Vowel Spectra", J. Acoust. Soc. Amer. *41*, 707-712.

POLS, L.C.W., KAMP, J.T. van der and PLOMP, R. (1969). "Perceptual and Physical Space of Vowel Sounds", J. Acoust. Soc. Amer. *46*, 458-467.

RAINBOLT, H.R. (1968). "Use of Noise to Establish Noise Pitch", J. Acoust. Soc. Amer. *43*, 316-323.

RAINBOLT, H.R. and SMALL, A.M. (1972). "Mach Bands in Auditory Masking: An Attempted Replication", J. Acoust. Soc. Amer. *51*, 567-574.

RAKOWSKI, A. (1968). "Pitch of Filtered Noise", 6th I.C.A., Tokyo, A-5-7.

RAKOWSKI, A. (1971). "Pitch Discrimination at the Threshold of Hearing", 7th I.C.A., Budapest, 20 H 6.

RATLIFF, F. (1965). *Mach Bands. Quantitative Studies on Neural Networks in the Retina* (Holden-Day, San Francisco).

RITSMA, R.J. (1962). "Existence Region of the Tonal Residue. I", J. Acoust. Soc. Amer. *34*, 1224-1229.

RITSMA, R.J. (1967). "Frequencies Dominant in the Perception of Pitch in Complex Sounds", J. Acoust. Soc. Amer. *42*, 191-199.

RITSMA, R.J., BILSEN, F.A., JONG, T.A. de and VERKOOYEN, C.J. (1967). "Spectral Regions Dominant in the Perception of Repetition Pitch", IPO Annual Progress Report *2*, 24-30.

SACHS, M.B. (1969). "Stimulus-Response Relation for Auditory-Nerve Fibers: Two-Tone Stimuli", J. Acoust. Soc. Amer. *45*, 1025-1036.

SACHS, M.B. and KIANG, N.Y.S. (1968). "Two-Tone Inhibition in Auditory-Nerve Fibers", J. Acoust. Soc. Amer. *43*, 1120-1128.

SCHARF, B. (1964). "Partial Masking", Acustica *14*, 16-23.

SCHARF, B. (1970). "Critical Bands", in *Foundations of Modern Auditory Theory*, J.V. Tobias, Ed. (Academic Press, New York).

SMALL Jr., A.M. (1959). "Pure-Tone Masking", J. Acoust. Soc. Amer. *31*, 1619-1625.

SMALL, A.M. and DANILOFF, R.G. (1967). "Pitch of Noise Bands", J. Acoust. Soc. Amer. *41*, 506-512.

SMOORENBURG, G.F. (1970). "Pitch Perception of Two-Frequency Stimuli", J. Acoust. Soc. Amer. *48*, 924-942.

TERHARDT, E. (1972a). "Zur Tonhöhenwahrnehmung von Klängen. I. Psychoakustische Grundlagen", Acustica *26*, 173-186.

TERHARDT, E. (1972b). "Zur Tonhöhenwahrnehmung von Klängen. II. Ein Funktionsschema", Acustica *26*, 187-199.

THURLOW, W.R. (1957). "An Auditory Figure-Ground Effect", Amer. J. Psychol. *70*, 653-654.

THURLOW, W.R. and ELFNER, L.F. (1959). "Continuity Effects with Alternately Sounding Tones", J. Acoust. Soc. Amer. *31*, 1337-1339.

VARTANIAN, I.A. (1973). "On Mechanisms of Specialised Reactions of Central Auditory Neurons to Frequency-Modulated Sounds", Paper presented at the Symposium *Auditory Analysis and Perception of Speech*, Leningrad.

VERSCHUURE, J. and MEETEREN, A.A. van (1974). "The Effect of Intensity on Pitch", submitted to Acustica.

VERSCHUURE, J., RODENBURG, M. and MAAS, A.J.J. (1974). "Frequency Selectivity and Temporal Effects of the Pulsation-Threshold Method", submitted to 8th I.C.A., London.

WARREN, R.M., OBUSECK, C.J. and ACKROFF, J.M. (1972). "Auditory Induction: Perceptual Synthesis of Absent Sounds", Science *176*, 1149-1151.

WIGHTMAN, F.L. (1973). "The Pattern-Transformation Model of Pitch", J. Acoust. Soc. Amer. *54*, 407-416.

WILSON, J.P. and EVANS, E.F. (1971). "Grating Acuity of the Ear: Psychophysical and Neurophysiological Measures of Frequency-Resolving Power", 7th I.C.A., Budapest, 20 H 14.

ZWICKER, E. (1956). "Die elementaren Grundlagen zur Bestimmung der Informatio[n]kapazität des Gehörs", Acustica *6*, 365-381.

ZWICKER, E. (1958). "Ueber psychologische und methodische Grundlagen der Lautheit", Acustica *8*, 237-258.

ZWICKER, E. (1963). "Ueber die Lautheit von ungedrosselten und gedrosselten Schallen", Acustica *13*, 194-211.

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 and G.F. Smoorenburg, Eds. (Sijthoff, Leiden, the Netherlands).

STELLINGEN

I

De verklaring dat de donkere Mach-band in visuele experimenten niet tot uitdrukking komt in een overeenkomstige drempelverlaging in direkte maskering doordat "In the region of the dark Mach Band a superimposed light must be made quite strong to be noticed since inhibition also reduces the effect of the superimposed (test) light" (1) impliceert dat de heldere Mach-band evenmin tot uitdrukking zal komen in een overeenkomstige drempelverhoging, hetgeen in strijd is met experimentele resultaten.

(1) G. von Békésy, J. Opt. Soc. Amer. *58*, 1-8 (1968).

II

Metingen betreffende het frekwentieanalyserend vermogen van het oor met behulp van sinusrasters in het frekwentiedomein, zoals in dit proefschrift beschreven, kunnen een waardevolle aanvulling betekenen bij audiologisch onderzoek.

III

Dat, onder bepaalde omstandigheden, een toonhoogte overeenkomstig de (afwezige) grondfrekwentie van twee opeenvolgende harmonischen ook kan worden waargenomen wanneer de twee harmonischen niet gelijktijdig doch alternerend worden aangeboden, heeft belangrijke konsekwenties voor theorieën over de toonhoogtewaarneming.

IV

Bij het oordeel van een luisteraar of een bepaalde geluidbron zich recht vóór of recht achter hem bevindt kan een rol spelen dat, ten gevolge van de akoestische eigenschappen van de oorschelp en de oriëntatie van het slakkehuis in het hoofd, het verschil tot uitdrukking komt in een overeenkomstig verschil van het zwaartepunt van de projektie van het geluidspektrum in het slakkehuis.

V

De mogelijkheid tot het waarnemen van de richting van akoestische waarschuwingssignalen in het verkeer, zoals gevoerd door politie-, brandweer- en ziekenwagens, kan wellicht worden bevorderd door een geschikte keuze van zowel de spektrale als de temporele samenstelling van dergelijke signalen.

VI

De ongunstige invloed van nagalm op de verstaanbaarheid van spraak kan worden gekwantificeerd door de "onscherpte in de tijd" waaraan het signaal bij de transmissie van bron naar luisteraar onderworpen is, te meten met behulp van sinusrasters in het tijddomein (modulatie-overdrachtsfunktie).

VII

Bij het toepassen van lawaaibeoordelingskrommen (Noise Rating Curves) voor het bepalen van de mate waarin stoorlawaai het verstaan van spraak beïnvloedt (1), kan een betrouwbaarder waarde worden afgeleid uit het geluiddrukniveau van het lawaai in de oktaafband met de *gunstigste* signaal/ruis verhouding dan, zoals gebruikelijk, uit die met de *on*gunstigste signaal/ruis verhouding.

(1) Proposal of the International Organization for Standardization, Technical Committee 43, Acoustics.

VIII

In het licht van de diskussies omtrent, enerzijds, de lawaaibelasting rondom vliegvelden en, anderzijds, de bevordering van het openbaar (rail-)vervoer is het interessant te bedenken dat de oppervlakte van de zône van overmatige lawaaibelasting rondom een vliegveld van dezelfde grootteorde is als die langs een baanvak met een lengte van 100 km, indien daarbij eveneens een zônegrens volgens de 'Kosten-formule' (1) wordt vastgesteld.

(1) G.J. Kleinhoonte van Os, Publ. nr. 12 v.h. Ned. Akoest. Gen. (1968).

IX

Dat juist de zoom van een bos veelal wordt gekenmerkt door een weelderige boomgroei kan worden opgevat als een 'Mach-band', die de invloed weerspiegelt van 'zijdelingse onderdrukking' waaraan aan alle zijden ingesloten bomen zijn onderworpen.

X

Voor wegen binnen de bebouwde kom die geen funktie hebben als verkeersader (te karakteriseren als 'woonstraten'), dient de toegestane maximum snelheid aanzienlijk lager gesteld te worden dan 50 km/uur.

T. Houtgast

Amsterdam, 27 maart 1974