

Discrimination of changes in the spectral shape of noise bands

Citation for published version (APA):

Versfeld, N. J. (1997). Discrimination of changes in the spectral shape of noise bands. *Journal of the Acoustical Society of America*, 102(4), 2264-2275. <https://doi.org/10.1121/1.419599>

DOI:

[10.1121/1.419599](https://doi.org/10.1121/1.419599)

Document status and date:

Published: 01/01/1997

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Discrimination of changes in the spectral shape of noise bands

Niek J. Versfeld

Institute for Perception Research, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

(Received 31 January 1996; revised 24 June 1997; accepted 30 June 1997)

Discrimination experiments were performed for a change in the spectral shape of noise bands. The subject's task was to discriminate noise bands with a positive spectral slope from those with a negative spectral slope. Thresholds were measured at several bandwidths and center frequencies, as well as for several noise samples. Experiments were performed while roving the overall intensity. At a fixed center frequency of 1 kHz, sensitivity was best for bandwidths of 3–6 semitones (ST). At larger bandwidths, thresholds increased only slowly. At a fixed bandwidth of 1 ST, threshold hardly changed as a function of the center frequency. At a fixed bandwidth of 58 Hz, threshold was lowest near 500–1000 Hz. Model calculations show that the EWAIF model [Feth, *Percept. Psychophys.* **15**, 375–378 (1974)] can account for the present results if the signal's bandwidth does not exceed 1 ST. The IWAIF model [Anantharaman *et al.*, *J. Acoust. Soc. Am.* **94**, 723–729 (1993)] can account for the present results only if the signal's bandwidth is smaller than 1 ST but larger than about 25 Hz. Results obtained with broadband signals could be described only qualitatively with the multichannel model [Durlach *et al.*, *J. Acoust. Soc. Am.* **80**, 63–72 (1986)]. Then, the model needs the assumption that either the output of the different frequency bands cannot be optimally combined, or that only two bands are used in the discrimination process. The present results are compared with those obtained with two-tone complexes measured under identical conditions [Versfeld and Houtsma, *J. Acoust. Soc. Am.* **98**, 807–816 (1995)]. © 1997 Acoustical Society of America. [S0001-4966(97)04110-6]

PACS numbers: 43.66.Ba, 43.66.Fe [JWH]

INTRODUCTION

According to the concept of the critical band, the degree to which broadband signals can be discriminated (or are rated dissimilar) is related solely to the sum of differences in activity across a number of frequency bands (Plomp, 1976; Florentine and Buus, 1981). In this view frequency bands that remain unaltered upon switching from one signal to another do not contribute in the discrimination process. However, over the last decade many experiments have been reported clearly showing that the mere presence of energy in frequency bands remote from a target band may contribute to the detectability of a change in that target band. Thus, it was argued, the auditory system is able to make across-band comparisons and discriminates on the basis of a change in the spectral shape. The processing of the signal in this particular manner is often referred to as "profile analysis" (see Green, 1988, for a review).

Up to now most profile-analysis experiments have been performed with tonal stimuli; only a few papers address spectral-shape discrimination with noiselike stimuli (e.g., Farrar *et al.*, 1987; Moore *et al.*, 1989). To our knowledge, no papers exist that discuss the relation between tonal and noiselike stimuli. The present paper is an attempt to do so, and reports on experiments with noiselike stimuli. The results are compared with those presented in an earlier paper (Versfeld and Houtsma, 1995) containing data obtained with two-tone complexes, measured under identical conditions, and in one experiment even utilizing the same subjects.

Three experiments are reported, all dealing with the discriminability between two noise bands that are identical except for the sign of their spectral slope. These types of signal

have been chosen not only because they are closely related to the two-tone complexes and multitone complexes in previous experiments (Versfeld and Houtsma, 1991, 1995), but also because the results may be relevant for speech research since some speaker characteristics seem to be related to the spectral slope of speech sounds (Li and Pastore, 1995).

In the first experiment, discriminability for a change in the spectral slope is measured at several bandwidths; in the second experiment, the influence of different noise samples on threshold is investigated; in the last experiment, threshold is measured at several center frequencies where the bandwidth has been kept fixed to either 1 semitone (ST) or 58 Hz. Similarities and differences between the present data set and the one in Versfeld and Houtsma (1995) will be discussed in terms of models that have been used previously in describing profile-analysis data.

I. GENERAL PROCEDURE

A. Stimuli

Noise bands were generated digitally (at 16 bits resolution and a sample frequency of 10 kHz) by summation of sinusoids, spaced 1 Hz apart, of the appropriate amplitude and starting phase. For each experimental condition, that is, fixed bandwidth and center frequency, the starting phase was chosen randomly once, and subsequently preserved from trial to trial. The spectral slope was linear on a log-amplitude, log-frequency scale, and its magnitude was expressed as the level difference ΔL between the two spectral edges. The signal bandwidth was expressed in semitones (ST). As an example, the spectrum of a noise band with a bandwidth of 24 ST and a (positive) spectral slope of 8 dB/

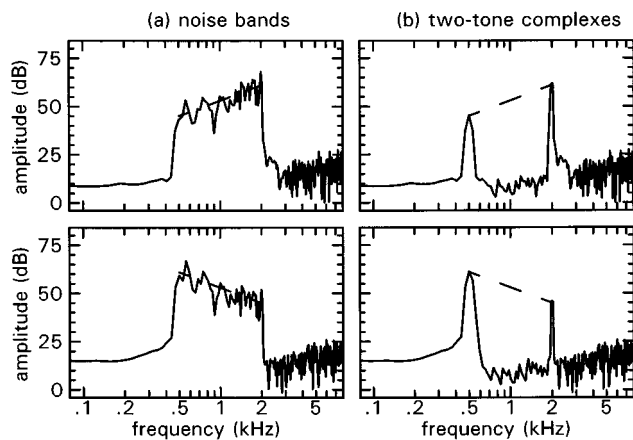


FIG. 1. Example of the amplitude spectra of stimuli similar to those used in the experiments: (a) noise bands and (b) two-tone complexes. All stimuli have a bandwidth of 24 ST ($f_1 = 500$ Hz, $f_2 = 2000$ Hz) and a level difference ΔL of 16 dB. The dashed lines indicate the imposed spectral slopes.

oct is plotted in the upper panel of Fig. 1(a). The lower panel of Fig. 1(a) displays the same noise band but with a negative spectral slope. The dashed line indicates the imposed spectral slope. Thus, the level difference ΔL between the edge components was 16 dB. Because the starting phase was kept fixed, departures from the imposed spectral slope were the same for both spectra (which is clearly illustrated in Fig. 1). Consequently, the difference between the two spectral shapes always is a straight line with a slope equal to twice the imposed spectral slope. In other words, the spectral change was well defined.

The two panels in Fig. 1(b) display the spectrum of a two-tone complex, also with a bandwidth of 24 ST and a level difference ΔL of 16 dB. Versfeld and Houtsma (1995) reported experiments with these two-tone signals, and results reported in the present paper will be compared with theirs.

B. Procedure

The experimental procedure used throughout the experiments reported in this paper was identical to that adopted by Versfeld and Houtsma (1995). For that reason, the general procedure is described here only briefly.

In an adaptive, three-interval oddity task, the subject's task was to discriminate a noise band with a positive spectral slope from a noise band with a negative spectral slope [see Fig. 1(a)], by indicating which interval out of three contained the stimulus with the odd spectral slope. The magnitude of the spectral slope was varied adaptively and the adaptive rules were chosen such that most trials were conducted near the 70.7%-correct point on the psychometric function. For the present paradigm, the level difference ΔL to obtain this percentage of correct responses corresponds to a sensitivity d' of 2.13 (Versfeld *et al.*, 1996). In order to facilitate comparison with data in the literature, thresholds ΔL reported in the present paper correspond to $d' = 1$. A linear relationship between d' and ΔL was assumed, i.e., $d' \propto \Delta L$. Then, reported thresholds are obtained simply by dividing the originally obtained thresholds by 2.13. According to Versfeld

et al. (1996), a sensitivity $d' = 1$ corresponds to a 44.7% correct score.

Each of the three sound bursts in one trial lasted 400 ms (including a 20-ms linear onset and a 20-ms linear offset ramp). The bursts were separated by 100-ms silent intervals. There was no response-time limit, and visual feedback was provided after each response.

The absolute threshold for the stimulus (with a flat spectral envelope) was determined for each subject and each bandwidth in advance. In the actual experiment, the subjects were prevented from using loudness cues by randomly varying the overall level of *each sound burst* between 30 and 50 dB sensation level. It was verified both by measurements and computer simulations that, with the 20-dB roving level, discrimination thresholds based on loudness cues only could not produce thresholds lower than about 7 dB. Under practically all conditions the obtained thresholds were lower, indicating that loudness cues were indeed not utilized.

II. EXPERIMENT I: EFFECT OF BANDWIDTH

A. Stimuli and subjects

Stimuli were noise bands centered at 1 kHz, that is, the center frequency f_c , defined as the geometric mean of the two edge frequencies, was kept fixed at 1 kHz. The bandwidth was set at 0.5, 1, 2, 3, 6, 12, or 24 ST.

Six subjects (five university students, who were paid for their services, and the author) participated in this experiment. Some of them were experienced listeners. Before the actual data collection of a condition, subjects were trained until the threshold had stabilized. Next, per subject and per condition 750 trial responses were collected, resulting in a standard error of the threshold estimate of 3.6%.

B. Results

The results of experiment I are displayed in Fig. 2, where thresholds ΔL (dB) have been plotted as filled symbols as a function of the bandwidth (ST). The data have been averaged across subjects, and bars indicate the standard deviation between subjects. The data for the individual subjects can be found in Table I.

For all subjects discriminability was best at a bandwidth of 3–6 ST. Starting from the minimum, threshold only slowly increases with increasing bandwidth. With decreasing bandwidth, however, threshold below 1 ST shows a sharp upturn.

With narrow-band signals subjects reported that the perceptual cue was a change in pitch. With increasing bandwidth the pitch cue gradually changed into a timbre cue: noise bands with a positive spectral slope were perceived as sounding sharper (as opposed to dull) than bands with a negative slope. The reported percept is consistent with the findings of von Bismarck (1974), who studied the verbal attributes of steady-state signals with different spectral shapes.

Both threshold behavior and reported perceptual cues suggest that, around a bandwidth of 3 ST, a transition occurs. Since 3 ST is about the width of the critical band, it might be that for bandwidths smaller than 3 ST, within-channel cues

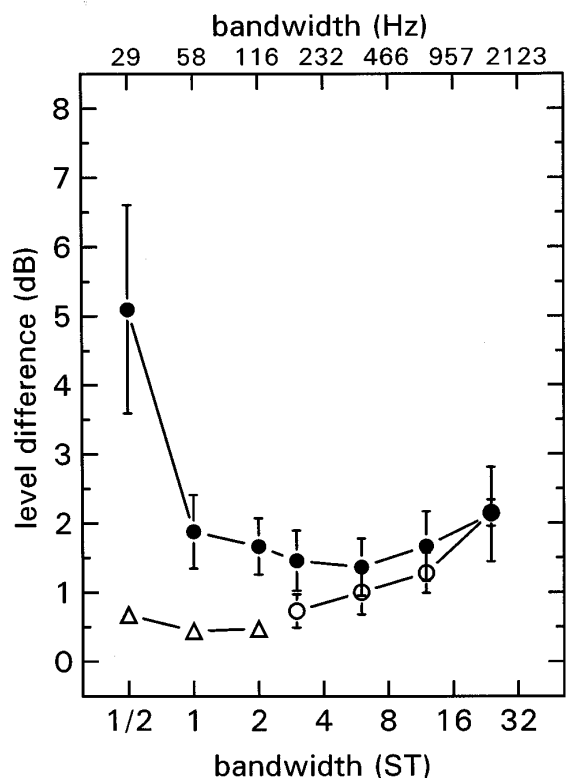


FIG. 2. Results of experiment I. Thresholds ΔL (dB) are plotted as a function of bandwidth (ST). Results are averaged across subjects. Filled and open symbols indicate thresholds obtained with noise bands and two-tone complexes, respectively. Error bars indicate the standard deviation between subjects. The signal bandwidth expressed in Hertz is given at the top of the figure.

dominate in the discrimination process, whereas across-channel cues dominate for larger bandwidths.

The open circles in Fig. 2 are the averaged data of experiment I of Versfeld and Houtsma (1995), who, for four subjects, measured thresholds ΔL for a relative change in the amplitudes of two-tone complexes [see Fig. 1(b)]. Error bars

indicate the standard deviation between subjects. The triangles denote thresholds for the same type of two-tone complexes and were taken from Versfeld (1993). These data have been collected under very similar conditions to the present experiment. Only one subject (the author) participated, but the data are in good agreement with earlier results of Versfeld and Houtsma (1991). With two-tone complexes, too, a minimum occurs, although this is not too apparent from Fig. 2. The results of Versfeld and Houtsma (1991) and Versfeld (1993) show that this minimum is near 1 ST, and thus is situated at a smaller bandwidth than with noise bands. With two-tone complexes, too, threshold increases sharply with decreasing bandwidth. The upturn, however, starts at bandwidths below 0.5 ST (cf. Versfeld and Houtsma, 1991). Except for a bandwidth of 24 ST, thresholds for two-tone complexes are smaller than those for noise bands.

III. EXPERIMENT II: EFFECT OF NOISE SAMPLE

Since in the previous experiment the starting phase was preserved from trial to trial, in other words, only one noise sample per bandwidth was taken, no information was obtained on the variability of the threshold as a function of the noise sample. To get an impression of the variability, in this experiment thresholds for different noise samples were measured.

A. Stimuli

Stimuli were noise bands centered at 1 kHz and had a bandwidth of 1 ST. Five noise samples were taken at random and labeled A–E. They differed from each other only with respect to the starting phase of the components. Sample A was used in experiment I. Thresholds were measured using six subjects (those who participated in experiment I). At least 450 trials were taken per subject and per condition, resulting in a standard error of the threshold of 4.5% or less.

TABLE I. Results of experiment I. Thresholds ΔL (dB), Δ_{EWAIF} (Hz), and Δ_{IWAIF} (Hz) for several bandwidths (ST).

Bandwidth	Subject 1			Subject 2			Subject 3			Subject 4		
	ΔL	Δ_{EWAIF}	Δ_{IWAIF}	ΔL	Δ_{EWAIF}	Δ_{IWAIF}	ΔL	Δ_{EWAIF}	Δ_{IWAIF}	ΔL	Δ_{EWAIF}	Δ_{IWAIF}
0.5	5.87	6.18	6.32	4.99	5.34	5.41	4.18	4.59	4.56	7.35	7.47	7.78
1	1.74	5.52	3.85	1.58	4.94	3.49	2.21	7.15	4.88	2.31	7.47	5.10
2	1.58	8.94	6.99	1.48	8.39	6.55	1.73	9.75	7.65	2.08	11.55	9.19
3	1.48	10.29	9.83	1.15	7.53	7.64	1.53	10.67	10.16	1.92	13.32	12.73
6	1.27	18.17	16.88	1.26	17.98	16.75	1.60	22.68	21.25	1.27	18.20	16.88
12	1.59	47.14	42.38	1.79	52.43	47.68	1.87	54.75	49.80	1.34	40.21	35.74
24	2.13	109.21	114.91	2.17	111.37	117.05	2.48	127.31	133.65	1.67	86.38	90.20
Bandwidth	Subject 5			Subject 6			Average					
	ΔL	Δ_{EWAIF}	Δ_{IWAIF}	ΔL	Δ_{EWAIF}	Δ_{IWAIF}	ΔL	Δ_{EWAIF}	Δ_{IWAIF}	ΔL	Δ_{EWAIF}	Δ_{IWAIF}
0.5	2.90	3.32	3.19	5.29	5.62	5.72	5.10	5.42	5.53			
1	1.01	2.77	2.24	2.42	7.83	5.34	1.88	5.95	4.15			
2	1.01	5.66	4.47	2.11	11.68	9.32	1.66	9.33	7.34			
3	0.78	5.51	5.19	1.89	13.11	12.53	1.40	10.1	9.30			
6	0.77	11.12	10.25	2.01	28.47	26.67	1.36	19.4	18.08			
12	0.98	29.61	26.15	2.44	70.82	64.85	1.67	49.2	44.50			
24	1.16	61.74	62.72	3.15	160.86	169.32	2.13	109.0	114.91			

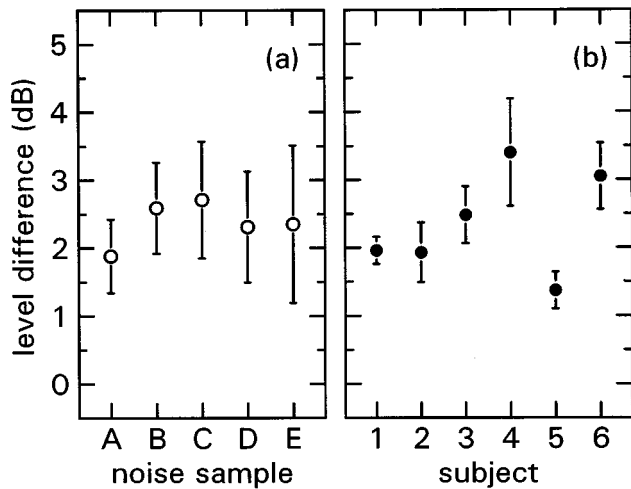


FIG. 3. Results of experiment II. Thresholds ΔL (dB) are plotted for (a) the noise samples A–E and (b) the six subjects #1–#6. Error bars indicate the standard deviation between subjects (a) and noise samples (b).

B. Results

Figure 3 displays the results of experiment II, where thresholds have been plotted as a function of the five noise samples [Fig. 3(a)], and as a function of the six subjects [Fig. 3(b)]. The individual data are given in Table II. Error bars denote the standard deviations between subjects [Fig. 3(a)] or between samples [Fig. 3(b)].

Standard deviations in Fig. 3(a) are about equal. For all subjects thresholds for sample A were lower than those for samples B–D. Similarly, thresholds for sample D were lower than those for sample C. Standard deviations in Fig. 3(b) are generally much smaller than those in Fig. 3(a) and roughly increase with increasing threshold. This clearly indicates that the variance in the data is caused mainly by between-subject differences: for a single subject, thresholds for the five noise samples are very similar. A 6 (subjects) \times 5 (samples) analy-

sis of variance (ANOVA) confirmed this: the differences between subjects accounted for 72% of the total variance (which is highly significant: $F[5,20]=19.2$, $p<0.0001$), whereas the differences between the noise samples accounted for only 8% of the total variance (which is also significant: $F[4,20]=4.1$, $p<0.05$). There was no significant interaction between subject and noise sample.

In conclusion, the results show that between-subject variability is much larger than between-sample variability, the latter being small enough to ensure that the shape of the threshold curve in Fig. 2 is preserved.

IV. EXPERIMENT III: EFFECT OF CENTER FREQUENCY

The results of experiment I showed that the use of noise bands instead of two-tone complexes (i.e., addition of frequency components) increased threshold and caused a shift in the bandwidth for which a minimum is reached from 1 to 3–6 ST. In experiment III it is investigated how threshold for a change in the sign of the spectral slope of noise bands behaves when it is measured as a function of center frequency f_c . The obtained results can be compared to those obtained by Versfeld and Houtsma (1995) with two-tone complexes.

A. Stimuli

In complete analogy with Versfeld and Houtsma (1995), the bandwidth was fixed at 58 Hz for one condition, whereas it was fixed at 1 ST in the other condition. Center frequencies for both conditions were 125, 250, 500, 1000, 2000, and 4000 Hz. At $f_c=1$ kHz a bandwidth of 58 Hz corresponds to a bandwidth of 1 ST. Yet, at this center frequency thresholds were measured for two different noise samples, both different from those used in experiments I and II.

Two subjects participated. They had also participated in the two previous experiments [subjects #2 and #5 (the au-

TABLE II. Results of experiment II. Thresholds ΔL (dB), ΔE_{WAIF} (Hz), and ΔI_{WAIF} (Hz) for several noise samples.

Sample	Subject 1			Subject 2			Subject 3			Subject 4		
	ΔL	ΔE_{WAIF}	ΔI_{WAIF}	ΔL	ΔE_{WAIF}	ΔI_{WAIF}	ΔL	ΔE_{WAIF}	ΔI_{WAIF}	ΔL	ΔE_{WAIF}	ΔI_{WAIF}
A	1.74	5.52	3.85	1.58	4.94	3.49	2.21	7.15	4.88	2.31	7.47	5.10
B	2.27	5.49	5.01	2.47	5.95	5.45	2.59	6.01	5.71	3.68	8.51	8.06
C	1.99	3.69	4.40	2.33	4.28	5.14	3.07	5.63	6.75	3.51	6.45	7.70
D	1.93	3.84	4.26	1.76	3.52	3.89	2.58	5.03	5.69	3.06	5.85	6.73
E	1.86	4.15	4.11	1.51	3.41	3.34	1.97	4.37	4.35	4.45	9.69	9.70
Average	1.96	4.54	4.33	1.93	4.42	4.26	2.48	5.64	5.48	3.40	7.59	7.46
s.d.	0.20	0.90	0.43	0.44	1.06	0.97	0.42	1.05	0.91	0.79	1.55	1.70
Sample	Subject 5			Subject 6			Average					
	ΔL	ΔE_{WAIF}	ΔI_{WAIF}	ΔL	ΔE_{WAIF}	ΔI_{WAIF}	ΔL	ΔE_{WAIF}	ΔI_{WAIF}			
A	1.01	2.77	2.24	2.42	7.83	5.34	1.88	5.95	4.15			
B	1.65	3.95	3.65	2.87	6.84	6.32	2.59	6.13	5.71			
C	1.61	3.02	3.56	3.72	6.83	8.15	2.71	4.98	5.97			
D	1.22	2.49	2.70	3.34	6.32	7.33	2.31	4.51	5.10			
E	1.38	3.13	3.05	2.92	6.35	6.42	2.35	5.18	5.18			
Average	1.37	3.07	3.04	3.05	6.83	6.71	2.37	5.35	5.22			
s.d.	0.27	0.55	0.59	0.49	0.61	1.07	0.32	0.68	0.70			

thor) in experiments I and II], as well as in the experiments with two-tone complexes (Versfeld and Houtsma, 1995). Although the number of subjects is small, results from previous experiments show that their behavior is similar to that of the other subjects in the group. Their performance, however, is somewhat better, which is probably due to greater experience. Per subject and per condition 600 trials were taken, resulting in a standard error of the threshold estimate of about 4%.

B. Results

Figure 4 displays, with filled symbols, the individual thresholds for two subjects for noise bands with a 58-Hz [Fig. 4(a)] or a 1-ST [Fig. 4(b)] bandwidth. Different symbols indicate the different subjects (subject #2 is indicated with circles; #5 with triangles). Also plotted (with open symbols) are thresholds for changes in the amplitude of two-tone complexes, as obtained by Versfeld and Houtsma (1995). The secondary axis at the top of the figure indicates the corresponding bandwidth of the noise band in semitones [if the bandwidth was kept fixed at 58 Hz, Fig. 4(a)], or in Hertz [if the bandwidth was 1 ST, Fig. 4(b)]. As noted before, the stimuli in the two conditions at $f_c=1$ kHz had the same bandwidth, but were different samples. Nevertheless, thresholds are very similar [cf. thresholds at $f_c=1$ kHz in Fig. 4(a) with those at $f_c=1$ kHz in Fig. 4(b)]. They are also close to those obtained in experiment II. This supports the conclusion of experiment II that, at least for these two subjects, threshold hardly depends on noise sample.

With noise bands, threshold for the 58-Hz condition [Fig. 4(a)] decreases at first as center frequency increases. At 500 Hz, a minimum of about 1.2 dB is reached. Further increase of the center frequency causes the threshold to increase, with a sharp upturn between 2000 and 4000 Hz. Thresholds for two-tone complexes show a similar trend, albeit that their absolute values are smaller and that the location of the minimum is situated near 1 kHz. Furthermore, it appears that subject #2 is able to maintain low thresholds even at 4 kHz. For both the noise bands and the two-tone complexes, the largest between-subject difference is situated at $f_c=4$ kHz.

Threshold for noise bands in the 1-ST condition [Fig. 4(b)] seems to be independent of center frequency for subject #2. Threshold behavior for subject #5 tends to show a bowl-like shape. The same is the case for the thresholds obtained with two-tone complexes. The between-subject differences are strikingly similar for thresholds obtained with noise bands and two-tone complexes, which is visible in both panels of Fig. 4.

V. MODEL PREDICTIONS

So far, no model has been successful in accounting for all profile-analysis data. Instead, the belief is that discriminability is based on different perceptual cues, or a combination of more cues (Richards and Nekrich, 1993), depending on the type of spectral change and on the signal's bandwidth. The use of a specific cue even can be manipulated by the instruction to the subject (Southworth and Berg, 1995). A

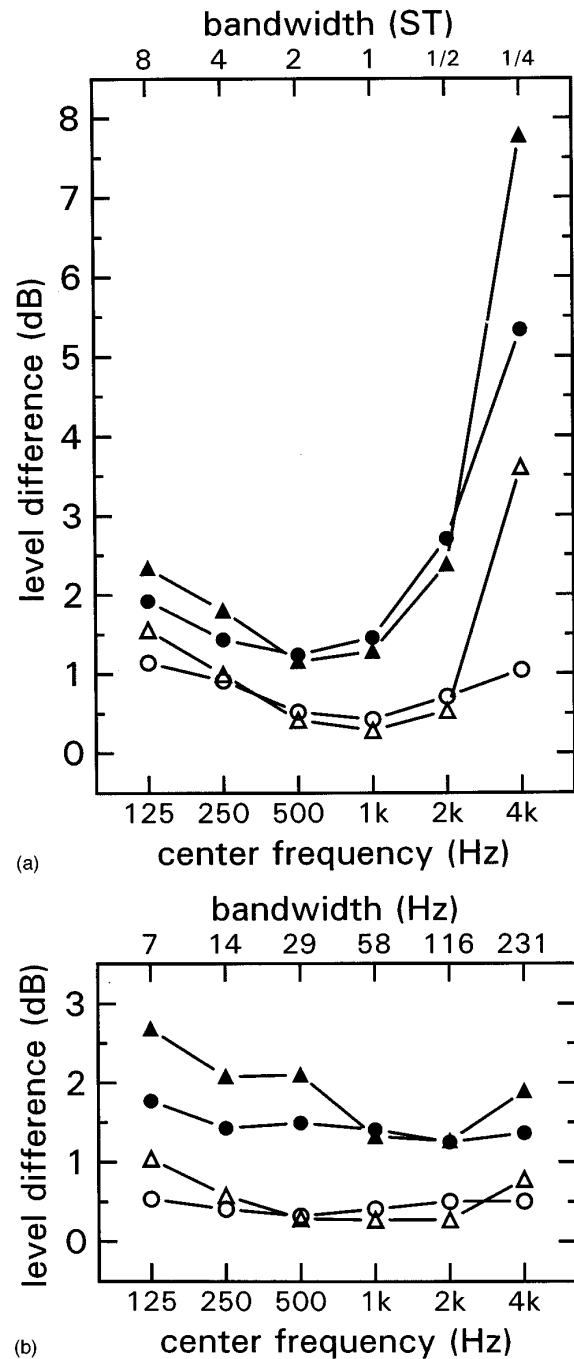


FIG. 4. Results of experiment III. Thresholds ΔL (dB) for signals with a bandwidth of 58 Hz (a) and 1 ST (b) are plotted as a function of center frequency. Filled and open symbols indicate thresholds obtained with noise bands and two-tone complexes, respectively. Different symbols indicate the individual subjects.

change in pitch may be the perceptual cue if the spectral shape changes in an asymmetrical manner (Green *et al.*, 1992), which is, for example, the case with the stimuli in the present experiment. A change in some of the characteristics of the temporal envelope may also provide the listener with a potential cue (Richards, 1992; Kidd, Jr. *et al.*, 1993). Lastly, if the signal's bandwidth is large enough in comparison with the critical bandwidth, changes in the spectral envelope can be detected by comparing the output of the different bands in a relative fashion—profile analysis in the true sense of the

word. Whether one or more of these cues could potentially be utilized by the subjects in the present experiment, and whether this behavior can be described with a model, will be discussed in this section.

A. Pitch cues

As noted earlier, subjects reported using pitch cues when discriminating between a (narrow) noise band with a positive and one with a negative spectral slope. A model that might describe the data is the EWAIF model (Feth, 1974). Another model, closely related to the EWAIF model is the IWAIF model (Anantharaman *et al.*, 1993; Dai, 1993).

1. The EWAIF model

The EWAIF model states that a change in the spectral shape is perceived as a change in pitch. The mapping of the signal's spectrum onto the pitch axis is done by calculation of the envelope-weighted average of the instantaneous frequency, or EWAIF,

$$\text{EWAIF} = \frac{\int_0^T E(t)f(t) dt}{\int_0^T E(t) dt}. \quad (1)$$

In this equation $E(t)$ and $f(t)$ are the (temporal) envelope function and the instantaneous frequency, respectively. Averaging is done over some time interval T . The EWAIF is expressed in Hertz. If, for two signals, the difference in the associated values for EWAIF increases, the discriminability between these two signals increases. Feth (1974) proposed the EWAIF model to describe threshold behavior of complementary pairs of two-tone complexes, similar to those depicted in Fig. 1(b). The model could account for thresholds obtained with narrow-band signals, i.e., signals that fell within one critical band. In order for the auditory system to calculate some quantity like EWAIF, the instantaneous frequency and the temporal envelope need to be extracted from the signal as a whole, thus it is to be assumed that the model breaks down as soon as the signal bandwidth exceeds the critical bandwidth. Just this assumption led Feth and O'Malley (1977) to make use of two-tone complexes to measure the width of the critical band.

For all bandwidths, noise samples and values of ΔL used in experiments I–III the instantaneous frequency and envelope function were extracted from the sampled waveform by means of a discrete Hilbert transformation (Feth and Stover, 1987; Kidd, Jr. *et al.*, 1991). $E(t)$ and $f(t)$ were calculated for the noise-band pairs with a positive and negative spectral slope and were subsequently used to calculate ΔEWAIF , the difference in EWAIF between a noise band with a positive slope and one with a negative slope. The integration time T was set to 400 ms, the duration of the stimulus. The relation between ΔEWAIF and ΔL is plotted in Fig. 5(a) as solid lines for the bandwidths used in experiment I, and in Fig. 5(b) for the five noise samples from experiment II. Note that the ordinate in panel (a) is logarithmic, whereas (for clarity) it is linear in panel (b). Figure 5 shows that for narrow-bandwidth signals, relationships are not entirely monotonic. This probably is due to errors intro-

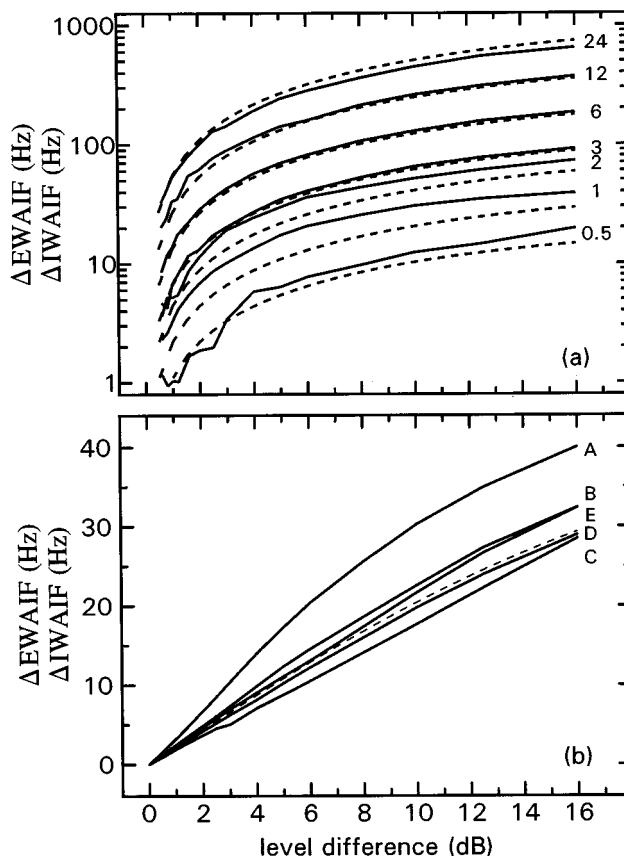


FIG. 5. Relationship between ΔL and ΔEWAIF (solid lines) or ΔIWAIF (dashed lines). (a) Relationship for the different bandwidths used in experiment I. (b) Relationship for the different noise samples used in experiment II. Note that the ordinate in panel (a) is logarithmic, whereas it is linear in panel (b).

duced by the calculation of the instantaneous frequency, involving a differentiation which is a highly noise-sensitive process (cf. Anantharaman *et al.*, 1993). The functions for the different bandwidths are, apart from the irregularities, rather similar and seem to lie parallel. Indeed, if ΔEWAIF is normalized with respect to the stimulus bandwidth (i.e., $\Delta\text{EWAIF}/[f_2 - f_1]$), the different curves practically coincide. Relationships between ΔL and ΔEWAIF were also determined for the stimuli from experiment III, but they are not shown in a figure, because they hardly provide additional information (as will be shown below).

With the aid of Fig. 5, thresholds ΔL were next converted to thresholds ΔEWAIF . Tables I–III yield, for the individual subjects, thresholds ΔL (dB) and the corresponding ΔEWAIF thresholds (Hz) for experiments I–III, respectively.

Since ΔEWAIF is a measure for discriminability, it is expected to be constant at threshold. The values in Table I for experiment I show that this might be true for bandwidths of 0.5 and 1 ST, but certainly not for larger bandwidths: ΔEWAIF rapidly increases with increasing bandwidth.

The EWAIF model is phase dependent and might therefore account for the differences in threshold with experiment II. Ideally, ΔEWAIF for the five noise samples should be the same for the individual subject. The calculations presented in Table II show that this seems not to be the case. An ANOVA

TABLE III. Results of experiment III. Thresholds ΔL (dB), Δ_{EWAIF} (Hz), Δ_{IWAIF} (Hz), and ΔF (Hz) as a function of center frequency f_c (Hz), for Subjects #2 and #5. Stimulus bandwidth was either 58 Hz or 1 ST.

Bandwidth	f_c	Subject #2				Subject #5			
		ΔL	Δ_{EWAIF}	Δ_{IWAIF}	ΔF	ΔL	Δ_{EWAIF}	Δ_{IWAIF}	ΔF
58 Hz	125	1.92	3.94	4.25	1.04	2.33	4.74	5.14	1.82
58 Hz	250	1.43	3.08	3.17	2.05	1.80	4.11	3.98	2.06
58 Hz	500	1.24	4.12	2.75	1.39	1.16	3.88	2.57	1.22
58 Hz	1000	1.45	3.27	3.21	2.49	1.28	2.87	2.83	1.72
58 Hz	2000	2.70	5.85	5.95	4.29	2.38	5.10	5.25	3.75
58 Hz	4000	5.33	10.23	11.53	7.43	7.78	13.44	16.39	15.87
1 ST	125	1.77	1.07	0.49	1.04	2.68	1.47	0.74	1.82
1 ST	250	1.43	1.37	0.79	2.05	2.08	1.97	1.15	2.06
1 ST	500	1.49	1.56	1.65	1.39	2.10	2.27	2.32	1.22
1 ST	1000	1.41	3.44	3.12	2.49	1.32	3.24	2.92	1.72
1 ST	2000	1.25	4.64	5.53	4.29	1.27	4.69	5.62	3.75
1 ST	4000	1.36	12.47	12.04	7.43	1.89	17.28	16.71	15.87

shows that the differences between the noise samples hardly decrease upon converting thresholds ΔL to Δ_{EWAIF} : the different samples still account for 7.6% of the variance ($F[4,20]=4.1, p<0.05$), whereas the between-subject variability does not decrease significantly, indicating that the differences between the thresholds of the different samples cannot be explained in terms of differences in EWAIF.

Table III yields, for two subjects, the thresholds Δ_{EWAIF} for noise bands as a function of the center frequency f_c obtained from the results of experiment III. Thresholds were measured for a bandwidth of 58 Hz or 1 ST. It can be shown that in calculating the difference in EWAIF between two signals, all terms containing f_c are cancelled out. Thus, two pairs of noise bands that differ with respect to center frequency but have the same bandwidth if expressed in Hertz, and the same starting phase, have identical values for Δ_{EWAIF} . The ΔL greatly varies with center frequency [cf. Fig. 4(a)], thus Δ_{EWAIF} also greatly varies with center frequency. Because spectral changes are transformed to pitch changes, it is likely that Δ_{EWAIF} varies as a function of center frequency in much the same way as does the pure-tone frequency difference limen, ΔF . Consequently, $\Delta_{EWAIF}/\Delta F$ rather than Δ_{EWAIF} is expected to be constant at threshold. Versfeld and Houtsma (1995) measured such pure-tone jnd's for the present two subjects (#2 and #5). Thresholds ΔF , based on 600 trials, are given in Table III. Figure 6(a) displays, with filled symbols, and for the two individual subjects, the ratio $\Delta_{EWAIF}/\Delta F$ for noise bands with a 58-Hz bandwidth as a function of the center frequency f_c . The two subjects are indicated with different symbols. Figure 6(b) does the same for 1-ST bandwidth noise bands. The standard error is about 6%. Open symbols in Fig. 6 represent model calculations with two-tone complexes for the same two subjects. If the perceived change in pitch, described by the change in EWAIF, was the discrimination cue, this ratio should be close to unity. Figure 6(b) shows that, on average, this is true for the noise bands in the 1-ST condition. For the noise bands with a bandwidth of 58 Hz [Fig. 6(a)], the ratio $\Delta_{EWAIF}/\Delta F$ is systematically larger than unity for center frequencies below 1000 Hz but is otherwise similar to ratios of the 1-ST condition. Values for $\Delta_{EWAIF}/\Delta F$ are systemati-

cally larger for two-tone complexes than for noise bands, despite the fact that thresholds ΔL are much smaller for two-tone complexes than for noise bands. Two phenomena may account for these results. First, since the EWAIF model is a within-channel model, it might be that the pitch cue cannot be used in full because the signals are partially resolved by the auditory periphery (which is true for the 58-Hz signals at the lowest center frequencies in terms of equivalent rectangular bandwidths). Second, signals like noise bands or even two-tone complexes are more complex than pure tones, and may need a greater pitch difference to obtain threshold. This may cause the systematic increase in the value of $\Delta_{EWAIF}/\Delta F$.

A perhaps more direct test of the EWAIF model, that is, minus the assumption that Δ_{EWAIF} is coupled to ΔF , is to calculate at each center frequency the ratio Δ_{EWAIF} (58 Hz)/ Δ_{EWAIF} (1 ST), which ideally should be unity. For noise bands, this ratio has been plotted with filled symbols as a function of the center frequency in Fig. 7(a). The open symbols were obtained with two-tone complexes. The different

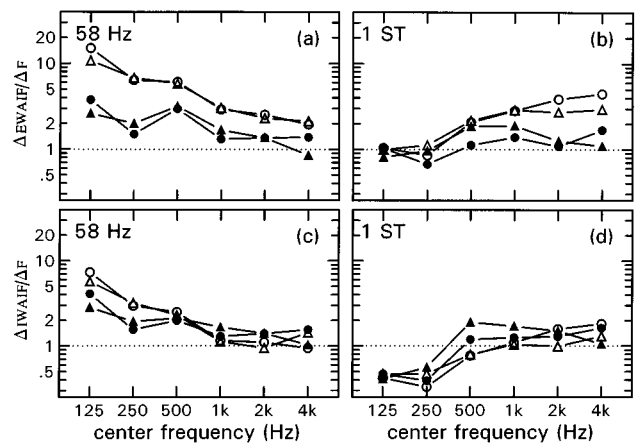


FIG. 6. Ratio $\Delta_{EWAIF}/\Delta F$ (a),(b) and $\Delta_{IWAIF}/\Delta F$ (c),(d) at threshold as a function of center frequency for the results of experiment III. Signal bandwidth was 58 Hz (a),(c) or 1 ST (b),(d). Filled and open symbols indicate results obtained with noise bands and two-tone complexes, respectively. The individual subjects are indicated by different symbols.

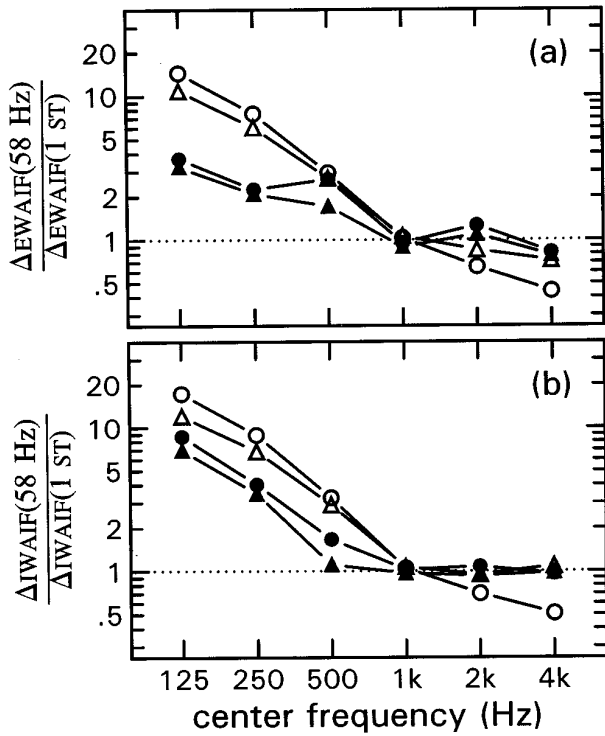


FIG. 7. Ratio $\Delta EWAIF(58 \text{ Hz})/\Delta EWAIF(1 \text{ ST})$ (a) and $\Delta IWAIF(58 \text{ Hz})/\Delta IWAIF(1 \text{ ST})$ (b) at threshold as a function of center frequency for the results of experiment III. Filled and open symbols indicate results obtained with noise bands and two-tone complexes, respectively. The individual subjects are indicated by different symbols.

symbols indicate the two subjects. Figure 7(a) shows that, especially for center frequencies below 1000 Hz, the ratio $\Delta EWAIF(58 \text{ Hz})/\Delta EWAIF(1 \text{ ST})$, deviates from unity for both noise bands and two-tone complexes.

The results in Figs. 6(a), 6(b), and 7(a) thus indicate that the EWAIF model can account for the data obtained with two-tone complexes and noise bands if the signal's bandwidth is smaller than about 1 ST, that is, the EWAIF model can probably account for the results only if the signal's bandwidth is small in comparison with the critical bandwidth, i.e., only if the signal is completely unresolved. The relative complexity of the noise bands and two-tone complexes in comparison with the pure tones probably does play a role, causing the ratio $\Delta EWAIF/\Delta F$ to be systematically larger than unity.

2. The IWAIF model

A model closely related to, and in fact derived from, the EWAIF model is the so-called IWAIF model (acronym for intensity-weighted average of the instantaneous frequency; Anantharaman *et al.*, 1993; Dai, 1993). The IWAIF of a signal is given by

$$IWAIF = \frac{\int_0^T E^2(t) f(t) dt}{\int_0^T E^2(t) dt}. \quad (2)$$

Anantharaman *et al.* (1993) have shown that the IWAIF is in fact equal to

$$IWAIF = \frac{\int_0^\infty |S(f)|^2 f df}{\int_0^\infty |S(f)|^2 df}, \quad (3)$$

where $S(f)$ is the Fourier transform of the (time) signal. Thus, IWAIF represents the "center of gravity" of the energy spectral density function. Although the EWAIF and the IWAIF models in origin are closely related, Eq. (3) shows that neither instantaneous frequency nor envelope function are required to determine IWAIF, thus in principle obviating the restriction that it can be applied only to signals that cannot be resolved by the auditory system. Moreover, the EWAIF model operates in the temporal domain, whereas the IWAIF model can be viewed as one operating in the spectral domain. Also, the EWAIF model is phase sensitive, whereas the IWAIF model essentially is not. Thus, although the EWAIF and the IWAIF model calculations probably will resemble each other, the interpretation can be essentially different.

Departing from Eq. (3), the difference in IWAIF between a noise band with a positive and a negative spectral slope can be derived analytically, and can be written as (see the Appendix)

$$\Delta IWAIF = \frac{1+m}{2+m} \left[\frac{f_2^2 \alpha^2 - f_1^2}{f_2 \alpha^2 - f_1} \right] - \frac{1-m}{2-m} \left[\frac{f_2^2 - \alpha^2 f_1^2}{f_2 - \alpha^2 f_1} \right], \quad (4)$$

where $\alpha = 10^{\Delta L/20}$, $m = 2 \ln(\alpha)/\ln(f_2/f_1)$ (where the logarithm has base e), f_1 and f_2 are the lower and upper frequency component of the signal, respectively, α represents the amplitude ratio of the two edge components of the noise bands, and m is simply related to the spectral slope S (in dB/oct) by $S \approx 3m$. One property of Eq. (4) is that $\Delta IWAIF$ (just like $\Delta EWAIF$) does not depend on the absolute frequencies, but rather on the frequency difference $f_2 - f_1$, and thus is independent of center frequency. Moreover, it can be shown that $\Delta IWAIF/(f_2 - f_1)$ is virtually independent of the signal bandwidths used in the present experiments. Thus, for the present noise bands $\Delta IWAIF/(f_2 - f_1)$ is only dependent on ΔL .

Again, the relationship between ΔL and $\Delta IWAIF$ has been calculated for the noise samples of the experiments, using Eq. (4). Their relationship is represented in Fig. 5 by dashed lines. Figure 5(a) shows that the shape of the functions is similar for the EWAIF model and the IWAIF model, but their relative position is sometimes different. Figure 5(b) shows only one curve for the IWAIF model (or stated differently: the five curves coincide), since the model is phase independent.

Thresholds $\Delta IWAIF$ for noise bands were determined, and are given in Tables I–III for experiments I–III, respectively. The calculations using the data of experiment I show that $\Delta IWAIF$ decreases with decreasing bandwidth but levels off at a bandwidth of about 1 ST. With narrower bandwidths, $\Delta IWAIF$ seems to increase again. Computations further indicate that the IWAIF model cannot account for the differences in threshold for the various noise samples in experiment II, simply because the model is phase independent. An ANOVA shows that the different samples still account for 8.1% of the variance ($F[4,20] = 4.2$, $p < 0.05$).

Figure 6(c) and (d) displays the ratio $\Delta I_{WAIF}/\Delta F$ for the 58-Hz condition and the 1-ST condition of experiment III, respectively. Again, open and filled symbols denote calculations with two-tone complexes and noise bands, respectively. The figures show that for low center frequencies none of the four conditions yields ratios that correspond to the frequency difference limen. For center frequencies of 500 Hz and above, the ratios obtained with noise bands are slightly higher than unity. Thus, the change in the center of gravity needs to be somewhat larger than the change in frequency of a pure tone. The reason for this is similar to that mentioned with the EWAIF model. The finding that ratios $\Delta I_{WAIF}/\Delta F$ are remote from unity in Fig. 6(d) at the lower center frequencies may be due to the extremely narrow signal bandwidths of these conditions (7 or 14 Hz). The fluctuations in the signal then are very slow, resulting in an unstable center of gravity, hence a poor estimate of the IWAIF. Figure 7(b) shows the ratio ΔI_{WAIF} (58 Hz)/ ΔI_{WAIF} (1 ST) for the two subjects. With noise bands, this ratio is very close to unity for conditions at higher center frequencies.

In conclusion, the results seem to indicate that the IWAIF model does well at narrow bandwidths (up to 1 ST), but only if the signal's bandwidth is large enough to ensure a stable estimate of the IWAIF.

B. Models acting on the temporal envelope

Several investigators have found that with narrow-band stimuli a change in the temporal envelope caused by a change in the spectral profile can be a valid cue (Richards, 1992; Berg *et al.*, 1992; Kidd, Jr. *et al.*, 1993; Green *et al.*, 1992). The PSE model proposed by Green *et al.* (1992) seems to be particularly successful. In this model differences in the power spectrum of the (temporal) envelope (PSE) between the two stimulus alternatives are used to predict threshold behavior.

One special property of the two-tone complexes used by Versfeld and Houtsma (1995) was that the temporal envelope was identical for both stimulus alternatives. This means that no envelope cues could be used. There is reason to believe that, with the noise bands of the present experiment, envelope cues play a very minor role. First of all, our subjects reported that they utilized pitch as a cue, not "roughness" or "smoothness" (verbal attributes that are typical for describing differences in the temporal envelope). Second, previous research (Green *et al.*, 1992) has shown that changes in the temporal envelope are mainly utilized when the spectral change is symmetric, that is, when it does not produce a shift in the "center of gravity" of the power spectrum, hence causing pitch cues to be weak or even absent. (An example of such a symmetric change is an increment in amplitude in the middle component of a linearly spaced multitone complex.) The present noise bands do not satisfy this condition. Third, it can be shown that the power spectrum of the temporal envelope of two noise bands with opposite spectral slopes are identical, but only if the spectral slope is linear on a linear-amplitude, linear-frequency scale. With the present stimuli this is approximately true for shallow spectral slopes. This does not mean that there cannot be any envelope cues, but indicates that the PSE model is bound to fail, since

it will predict no differences in PSE between the two stimulus alternatives. Last, the temporal envelope for signal pairs is identical not only for $\Delta L = 0$, but also for very large values of ΔL , since the signals then reduce to a sinusoid with frequency equal to either f_1 or f_2 : the temporal envelope is flat in both cases.

Yet, to verify a possible effect of change in the temporal envelope the PSE was calculated for the signals of the present experiment, and relations were determined between ΔL and ΔPSE (as described by Kidd *et al.*, 1993) for each of the experimental conditions. The calculations mostly resulted in capricious functions, indicating that the differences in PSE between the two stimulus alternatives were very small and probably due to discretisation errors. More importantly, ΔPSE was not constant at threshold.

In retrospect, we could have avoided temporal-envelope cues by using noise bands with a linear spectral slope on a linear-frequency and linear-amplitude scale. In that case, it can be shown that the power spectrum of the envelope remains unaltered when changing from a positive to a negative spectral slope. However, at the time we thought a slope in dB/oct to be perceptually more relevant.

C. Multichannel models

In multichannel models as developed by Plomp (1976), Durlach *et al.* (1986), and Ito (1990), a broadband signal is filtered by a set of (nonoverlapping) bandpass filters. The amount of activity is measured in each band. Thus the output of the model is a crude spectral representation of the signal. The frequency bands are usually identified with critical bands. Thus, if the entire signal falls well within a critical band (as is the case with most of the stimuli from experiments II and III), the model is reduced to a single-channel model, which can only register differences in overall level. Since the overall level was varied randomly between and within trials, a single-channel model will fail to describe the data for narrow-band signals. It therefore seems sensible to apply a multichannel model only to those results of experiment I that were obtained with broadband signals. Though the theoretical background of the multichannel model is fairly straightforward, a precise quantitative implementation is difficult since assumptions have to be made about, for instance, the auditory filter shape. Also, only few data points from experiment I can be used to fit the model. Nevertheless, some qualitative statements can be made. If the output of the different critical bands can be compared (correlated) with one another [which is an essential feature in the model of Durlach *et al.* (1986)], thresholds are not or only slightly influenced by a roving overall level, and the listener is able to achieve the low thresholds of experiment I. If the output of the different bands could be perfectly compared, one would expect an ever-decreasing threshold with increasing bandwidth, since widening of the bandwidth, while keeping the level difference fixed, results in a large spectral change. (In the log-amplitude, log-frequency domain, the spectral difference between two noise bands with opposite slopes is a straight line with a slope twice the imposed spectral slope, hence the difference is linearly related to the bandwidth.) The results, however, show that the opposite is true. This

indicates, in terms of the model, that the capability of across-channel comparison is not perfect. It is possible that the ability to compare the output of different channels decreases as bands become more remote [a phenomenon similar to that observed with two-tone complexes (Versfeld, 1993)]. Also, it may be that not all, but only a limited number of channels are monitored. This possibility is discussed below.

VI. DISCUSSION

Noise bands with spectral slopes that change in sign can be discriminated while a roving intensity level is present. Changes are best perceived when the signal's bandwidth is about 3 ST. The presence of such a minimum suggests that probably more than one discrimination mechanism exists. Subjects report that, with the present signals, changes in the spectral shape of narrow-band stimuli are detected by pitch shifts, whereas spectral-shape differences in broadband stimuli are discriminated by comparison of timbre (dullness versus sharpness). Details of the noise sample seem to have little influence on threshold.

Although mere differences in the temporal envelope may provide the listener with a cue, it has been made plausible that for the current set of conditions these cues are unreliable and probably even absent.

Calculations with the EWAIF and the IWAIF model indicate, first of all, that these models cannot account for the present results if the relative bandwidth of the noise bands exceeds the value of 1 ST. In other words: the signals need to be well within the critical bandwidth. The EWAIF model as proposed by Feth (1974) can account for all data if the signal's bandwidth is 1 ST or less, but the IWAIF model fails to account with conditions at lower center frequencies. A possible explanation for this failure is that the auditory system has difficulties in estimating the center of gravity (i.e., calculating IWAIF) with these slowly fluctuating signals, since they are only 7 and 14 Hz wide at center frequencies of 125 and 250 Hz, respectively. This also explains the slight increase in $\Delta IWAIF$ for the 0.5-ST results in experiment I (29-Hz bandwidth). Multichannel EWAIF or IWAIF models (where EWAIF or IWAIF is calculated after the signal has been filtered into separate frequency channels) may account for signals with bandwidths that are larger than 1 ST (Anantharaman *et al.*, 1991).

In conclusion, the EWAIF model can account for the present results if the signal is well within the critical bandwidth and its bandwidth does not exceed 1 ST. This restriction holds also for the IWAIF model. Additionally, the IWAIF model can account for the data only if the signal's bandwidth is larger than about 20–30 Hz.

Threshold shifts due to different phase relations (experiment II) could not be explained by the EWAIF or the IWAIF model. The ANOVAs showed that the different noise samples gave a just-significant effect, so phase probably does play only a minor role in discrimination. This conclusion is supported by Dai *et al.* (1996), who showed that their pitch matches to narrow-band complex signals could be accounted for by the (phase-independent) IWAIF model much better than by the EWAIF model. Of course, one can never rule out the possibility that the signals are filtered in the auditory

system [both with respect to the amplitude (Berg *et al.*, 1992) and phase], such that the EWAIF or IWAIF of the resulting signal indeed can explain the obtained thresholds.

The choice of stimuli in the present experiments was based on earlier experiments with two-tone complexes (Versfeld and Houtsma, 1995). In addition to the results of the present experiments, Figs. 2, 4, 6, and 7 also display the (calculations based on) results obtained with two-tone complexes. Threshold behavior for the two stimulus types show differences with respect to absolute values (thresholds are always higher for noise bands than for two-tone complexes), and the position of the minimum (about 3 ST for noise bands and 1 ST for two-tone complexes). On the other hand, the between-subject differences for noise bands and two-tone complexes (cf. Fig. 4) are strikingly similar. In the following the relationship between the two stimulus types is discussed.

The multichannel model of Durlach *et al.* (1986) can only be applied to signals with a bandwidth that exceeds the critical band. Then, it can qualitatively account for the results of experiment I, but only if some kind of nonoptimal processing takes place. Figure 2 showed that for bandwidths beyond about 3 ST, threshold increases for both noise bands and two-tone complexes. Although these data points provide scarce evidence, it seems that *mainly two regions of the spectrum* are used in the decision process. This idea, already proposed by Bernstein *et al.* (1987), is in fact a very simple but interesting version of the multichannel model. With a roving intensity level the minimum number of channels which have to be observed in order to detect a change in the spectral shape is two. With two-tone complexes it is obvious that only two channels are involved. With large-bandwidth noise bands most information is present at the edges of the noise band. It is possible that only the information in the edge bands is used. Alternatively, the higher thresholds for noise bands, as well as the shift in the position of the minimum (3–6 ST for noise bands and 1 ST for two-tone complexes) suggests that, with noise bands, not the edges, but rather some intermediate bands are monitored. Unfortunately, the present data cannot give conclusive evidence for this hypothesis.

In the literature, Bernstein *et al.* (1987) reported several experiments where the threshold for detecting an increase in amplitude of only one component in a 21-component spectrum was compared with thresholds for detecting broadband spectral changes (e.g., flat versus tilted spectra). They found that thresholds obtained with broadband changes could not be predicted by the single-component thresholds unless it was assumed that only two regions of the spectrum were used in the discrimination process. In profile-analysis experiments it was found in general that complex changes give poorer thresholds than would be expected from optimum combination of thresholds obtained with single-component changes in the same multitone spectrum and (Green and Kidd, Jr., 1983; Green *et al.*, 1987; Richards *et al.*, 1989; Berg and Green, 1992).

Farrar *et al.* (1987) studied the discriminability of different speechlike noise spectra, embedded in long-term averaged speech noise. In an attempt to predict their results with a simple multichannel model they found rather large values

for the internal noise variance, indicating poor performance. Inspection of their Figure 6 indicates that only two regions are used, viz., those containing the spectral peaks. The same variance was assumed for all bands in their multichannel model. Had it been possible to let the internal noise variance vary across bands, they probably would have found a small variance in two bands and large variances in the other bands, indicating that only two regions of the spectra were used.

The ability of the auditory system to detect a peak or a notch in an otherwise flat spectrum has been studied to some extent in the literature (Moore *et al.*, 1989; Schacknow and Raab, 1976). A simple multichannel model predicts that threshold should decrease as the bandwidth of the peak or notch increases. The results, however, showed only a slight dependence or even none at all. This again suggests that only two regions of the spectrum are used. One region is then situated at the peak (notch), the other at the nonchanging part of the spectrum.

It seems that the discrimination of complex spectral changes can be explained qualitatively by assuming that only two regions are used. It has to be kept in mind, however, that all experiments reported are discrimination experiments where only two stimulus alternatives had to be compared. It is very well possible that, in experiments where, for instance, a set of spectra have to be discriminated or identified, observation of more than two spectral regions will be involved. This is true for example, for experiments where different timbres (Plomp, 1976) or vowels (Pols *et al.*, 1969) have to be judged for dissimilarity. In that case, the whole spectrum has to be monitored and not just the two regions containing most of the information that is relevant for discrimination.

ACKNOWLEDGMENTS

This work has been supported by the Netherlands Organization for the Advancement of Pure Research (NWO) through the PSYCHON foundation, Grant No. 560-262-034. Some of the results have been presented previously in Versfeld (1992). The author is very grateful to Adrian Houtsma for all his support, and is indebted to Dik Hermes and Armin Kohlrausch for their comments on earlier versions of this manuscript. Two anonymous reviewers are acknowledged for their very constructive remarks. The author once again thanks Theo de Jong for his indispensable technical assistance.

APPENDIX: DERIVATION OF Δ IWAIF FOR NOISE BANDS WITH OPPOSITE SPECTRAL SLOPES

The amplitude spectrum $L_p(f)$ of a noise band with a positive spectral slope, edge frequencies f_1 and f_2 , and levels L and $L + \Delta L$ at the two spectral edges [cf. Fig. 1(a)] is given by

$$L_p(f) = L + S \log_2 \left(\frac{f}{f_1} \right), \quad (\text{A1})$$

where S is the spectral slope expressed in dB/oct. Similarly, the amplitude spectrum $L_N(f)$ of a noise band with a negative spectral slope is given by

$$L_N(f) = L + \Delta L - S \log_2 \left(\frac{f}{f_1} \right). \quad (\text{A2})$$

Conversion of Eqs. (A1) and (A2) to a linear amplitude scale yields

$$A_p(f) = A_0 10^{L_p(f)/20} = A \left(\frac{f}{f_1} \right)^{m/2}, \quad (\text{A3})$$

$$A_N(f) = A_0 10^{L_N(f)/20} = A \alpha \left(\frac{f}{f_1} \right)^{-m/2}, \quad (\text{A4})$$

where A_0 is a reference amplitude, $A = A_0 10^{L/20}$, $\alpha = (A + \Delta A)/A$, and $m = S/(10 \log_{10}(2)) = 2 \log_e(\alpha)/\log_e(f_2/f_1)$.

For a noise band with a positive spectral slope, Eq. (3) in the main text can be written as (cf. Anantharaman *et al.*, 1993)

$$\text{IWAIF}_p = \frac{\int_{f_1}^{f_2} A_p^2(f) f df}{\int_{f_1}^{f_2} A_p^2(f) df}. \quad (\text{A5})$$

Substituting the expression for $A_p(f)$, given in Eq. (A3), into Eq. (A5) yields, after calculating the integral and doing some rearranging, an expression for IWAIF_p :

$$\text{IWAIF}_p = \frac{1+m}{2+m} \left[\frac{f_2^2 (f_2/f_1)^m - f_1^2}{f_2 (f_2/f_1)^m - f_1} \right]. \quad (\text{A6})$$

Similarly, IWAIF_N is given by

$$\text{IWAIF}_N = \frac{1-m}{2-m} \left[\frac{f_2^2 - f_1^2 (f_2/f_1)^m}{f_2 - f_1 (f_2/f_1)^m} \right]. \quad (\text{A7})$$

By using the equality $\alpha^2 = (f_2/f_1)^m$, the expression for $\Delta \text{IWAIF} = \text{IWAIF}_p - \text{IWAIF}_N$ can be written as

$$\Delta \text{IWAIF} = \frac{1+m}{2+m} \left[\frac{f_2^2 \alpha^2 - f_1^2}{f_2 \alpha^2 - f_1} \right] - \frac{1-m}{2-m} \left[\frac{f_2^2 - f_1^2 \alpha^2}{f_2 - f_1 \alpha^2} \right], \quad (\text{A8})$$

which is Eq. (4) in the main text.

- Anantharaman, J. N., Krishnamurthy, A. K., and Feth, L. L. (1991). "Auditory processing of complex signals using the multichannel EWAIF," *J. Acoust. Soc. Am.* **89**, 1938–1939.
- Anantharaman, J. N., Krishnamurthy, A. K., and Feth, L. L. (1993). "Intensity-weighted average of instantaneous frequency as a model for frequency discrimination," *J. Acoust. Soc. Am.* **94**, 723–729.
- Berg, B. G., and Green, D. M. (1992). "Discrimination of complex spectra: Spectral weights and performance efficiency," in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, Oxford), pp. 373–379.
- Berg, B. G., Nguyen, Q. T., and Green, D. M. (1992). "Discrimination of narrow-band spectra. I: Spectral weights and pitch cues," *J. Acoust. Soc. Am.* **92**, 1911–1918.
- Bernstein, L. R., and Green, D. M. (1987). "Detection of simple and complex spectral changes of spectral shape," *J. Acoust. Soc. Am.* **82**, 1587–1592.
- Dai, H. (1993). "On the pitch of two-tone complexes," *J. Acoust. Soc. Am.* **94**, 730–734.
- Dai, H., Nguyen, Q., Kidd, Jr., G., Feth, L. L., and Green, D. M. (1996). "Phase independence of pitch produced by narrow-band sounds," *J. Acoust. Soc. Am.* **100**, 2349–2351.
- Durlach, N. I., Braida, L. D., and Ito, Y. (1986). "Towards a model for discrimination of broadband signals," *J. Acoust. Soc. Am.* **80**, 63–72.

- Farrar, C. L., Reed, C. M., Ito, Y., Durlach, N. I., Delhorne, L. A., Zurek, P. M., and Braida, L. D. (1987). "Spectral shape discrimination. I. Results from normal-hearing listeners for stationary broadband noises," *J. Acoust. Soc. Am.* **81**, 1085–1092.
- Feth, L. L. (1974). "Frequency discrimination of complex periodic tones," *Percept. Psychophys.* **15**, 375–379.
- Feth, L. L., and O'Malley, H. (1977). "Two-tone auditory spectral resolution," *J. Acoust. Soc. Am.* **62**, 940–947.
- Feth, L. L., and Stover, L. J. (1987). "Demodulation processes in auditory perception," in *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, NJ), pp. 76–86.
- Florentine, M., and Buus, S. (1981). "An excitation-pattern model for intensity discrimination," *J. Acoust. Soc. Am.* **70**, 1646–1654.
- Green, D. M. (1988). *Profile Analysis* (Oxford U.P., New York).
- Green, D. M., and Kidd, Jr., G. (1983). "Further studies of auditory profile analysis," *J. Acoust. Soc. Am.* **73**, 1260–1265.
- Green, D. M., Berg, B. G., Dai, H., Eddins, D. A., Onsan, Z., and Nguyen, Q. (1992). "Spectral shape discrimination of narrow-band sounds," *J. Acoust. Soc. Am.* **92**, 2586–2597.
- Green, D. M., Onsan, Z. A., and Forrest, T. G. (1987). "Frequency effects in profile analysis and detecting complex spectral changes," *J. Acoust. Soc. Am.* **81**, 692–699.
- Ito, Y. (1990). "Auditory discrimination of power spectra for roving two-tone stimuli," Ph.D. thesis, Massachusetts Institute of Technology.
- Kidd, Jr., G., Mason, C. R., Uchanski, R. M., Brantley, M. A., and Shah, P. (1991). "Evaluation of simple models of auditory profile analysis using random reference spectra," *J. Acoust. Soc. Am.* **90**, 1340–1354.
- Kidd, Jr., G., Uchanski, R. M., Mason, C. R., and Deliwala, P. S. (1993). "Discriminability of narrow-band sounds in the absence of level cues," *J. Acoust. Soc. Am.* **93**, 1028–1037.
- Li, X., and Pastore, R. E. (1995). "Perceptual constancy of a global spectral property: Spectral slope discrimination," *J. Acoust. Soc. Am.* **98**, 1956–1968.
- Moore, B. C. J., Oldfield, S. R., and Dooley, G. J. (1989). "Detection and discrimination of spectral peaks and notches at 1 and 8 kHz," *J. Acoust. Soc. Am.* **46**, 458–467.
- Plomp, R. (1976). *Aspects of Tone Sensation* (Academic, London).
- Pols, L. C. W., van der Kamp, L. J. Th., and Plomp, R. (1969). "Perceptual and physical space of vowel sounds," *J. Acoust. Soc. Am.* **46**, 458–467.
- Richards, V. M. (1992). "The detectability of a tone added to narrow bands of equal-energy noise," *J. Acoust. Soc. Am.* **91**, 3424–2435.
- Richards, V. M., and Nekrich, R. D. (1993). "The incorporation of level and level-invariant cues for the detection of a tone added to noise," *J. Acoust. Soc. Am.* **94**, 2560–2574.
- Richards, V. M., Onsan, Z. A., and Green, D. M. (1989). "Auditory profile analysis: Potential pitch cues," *Hearing Res.* **39**, 27–36.
- Schacknow, P. N., and Raab, D. H. (1976). "Noise-intensity discrimination: Effects of bandwidth conditions and mode of masker presentation," *J. Acoust. Soc. Am.* **60**, 893–905.
- Southworth, C., and Berg, B. G. (1995). "Multiple cues for the discrimination of narrow-band sounds," *J. Acoust. Soc. Am.* **98**, 2486–2492.
- Versfeld, N. J. (1992). "On the perception of spectral changes in noise bands," in *The Auditory Processing of Speech*, edited by M. E. H. Schouten (Mouton de Gruyter, Berlin), pp. 219–227.
- Versfeld, N. J. (1993). "Profile-analysis models applied to two-tone complexes," in *Contributions to Psychological Acoustics*, edited by A. Schick (BIS, Oldenburg), pp. 269–285.
- Versfeld, N. J., and Houtsma, A. J. M. (1991). "Perception of spectral changes in multi-tone complexes," *Q. J. Exp. Psychol.* **43A**, 459–479.
- Versfeld, N. J., and Houtsma, A. J. M. (1995). "Discrimination of changes in the spectral shape of two-tone complexes," *J. Acoust. Soc. Am.* **98**, 807–816.
- Versfeld, N. J., Dai, H., and Green, D. M. (1996). "The optimum decision rules for the oddity task," *Percept. Psychophys.* **58**, 10–21.
- von Bismarck, G. (1974). "Sharpness as an attribute of the timbre of steady sounds," *Acustica* **30**, 159–172.