ISH 2003. Do not distribute. To appear in Pressnitzer, D., de Cheveigné, A., McAdams, S., and Collet, L. (Eds). *Auditory signal processing: Physiology, psychoacoustics, and models.*

Discrimination of different temporal envelope structures of diotic and dichotic target signals within diotic wide-band noise

Steven van de Par[1], Armin Kohlrausch[1,2], Jeroen Breebaart[1], and Martin McKinney[1]

[1] Philips Research Laboratory, Eindhoven, The Netherlands{Steven.van.de.Par, Armin.Kohlrausch, Jeroen.Breebaart, Martin.McKinney}@philips.com
[2] Eindhoven University of Technology

1 Introduction

Binaural listening leads to considerable improvements in detection thresholds in masked listening conditions compared to monaural listening. When an interaurally out-of-phase tonal signal is masked by an in-phase noise masker, detection thresholds can be as much as 25 dB lower than when detecting an in-phase tone within the same noise masker provided that the masker is narrow-band (e.g. Zurek and Durlach, 1987).

When high-level auditory processing of speech for example is considered, detectability of target signals may not be the most relevant perceptual parameter. The extent to which specific properties (e.g. temporal envelope) of the target signal are audible may be more relevant for high-level auditory processing. This notion motivated our investigation of the discriminability of pairs of target signals with different temporal envelope structures that are partially masked by noise.

Studies on discriminability on the basis of frequency differences (Henning, 1973) and interaural time differences (Cohen, 1981; Stern, Slocum, and Phillips, 1983) have been done. Henning showed that for a low signal-to-noise ratio, frequency discrimination between sinusoidal signals in the presence of white diotic noise improved when the sinusoids were presented interaurally out-of-phase rather than in-phase. In these experiments, detection thresholds and frequency discrimination thresholds were both lower in the dichotic condition. It was concluded that apparently the information needed to discriminate frequency is preserved beyond the first stages of binaural interaction.

Cohen (1981) found that midline interaural time delay jnds for a tone were smallest when presented in diotic noise, increased with interaurally uncorrelated noise, but were largest when presented in interaurally out-of-phase noise. In contrast to Henning's findings, in these experiments ITD discrimination thresholds were highest in conditions where detection thresholds were lowest.

In the present study we explore to what extent and under which conditions listeners have a binaural advantage in the processing of temporal information of target signals. For this purpose we measured the discriminability between bandlimited noise and harmonic-complex-tone targets presented in diotic noise. The targets were presented either interaurally in-phase or out-of-phase. In addition, we measured detection thresholds for the noise target and the harmonic-complex-tone target.

2 Experiment I: the effect of target-signal bandwidth

2.1 Method and stimuli

Within each interval of a discrimination task, a 300 ms target was presented, temporally centred in a 400 ms masker. The masker was a 2-kHz low-pass noise that was interaurally in-phase with an SPL of 65 dB. Two types of targets were used: bandpass noise (BPN) with a flat spectral envelope; and a harmonic tone complex (HTC) with the same band-pass characteristic as the BPN, component spacings of 20 Hz, and a sinusoidal phase spectrum. In narrow-band conditions, the targets were 100 Hz wide and centred at 1 kHz, the first of five components of the HTC started at 960 Hz. In broad-band conditions, the targets were 1500 Hz wide and also centred at 1 kHz. The first of 75 components of the HTC started at 260 Hz. Both targets were either presented interaurally in-phase (N_0S_0) or out-of-phase (N_0S_π) depending on the measurement condition. Both masker and target had raised-cosine on and off-set ramps with 30 ms duration.

In the discrimination task, three intervals were presented on each trial. Two intervals contained the BPN target plus masker, while one interval contained the HTC target plus masker. In all intervals the target levels were the same. The subjects' task was to indicate which interval contained the HTC target. Target levels were controlled with a 2-down 1-up adaptive tracking procedure. At the start of each run, levels were adjusted with 8 dB steps and stepsizes were halved after each second reversal until a minimum step-size of 1 dB was reached after which another 8 reversals were measured. The median of the last 8 reversals was used as the measured threshold of that run. For each condition at least four thresholds were measured.

In addition to the discrimination task, subjects also performed a *detection* task, where only one of the three intervals contained a target. These detection conditions served as reference conditions to the corresponding discrimination conditions.

2.2 Results and discussion

The results showed good agreement between the five subjects, therefore we present here only the mean results in Fig. 1.

For the narrowband (100 Hz) N_0S_0 condition, detection thresholds for BPN (left open circle) and HTC (left open square) are nearly equal, while discrimination thresholds (left open diamond) are about 10 dB higher. This indicates that in the

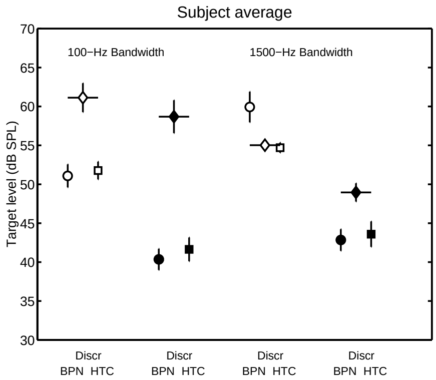

Fig. 1. Average discrimination thresholds (diamonds) and detection thresholds for BPN (circles) and HTC (squares) are shown for N_0S_0 conditions (open symbols) and N_0S_π conditions (filled symbols). The left half of the panel shows thresholds for 100-Hz wide targets, the right half for 1500-Hz wide targets. Vertical lines indicate averaged standard deviations based on individual subject standard deviations.

presence of the masking noise, targets need to be clearly audible in order to discriminate between the spectro-temporal structures of the BPN and the HTC.

The filled symbols on the left side of the panel show the same measurement conditions for an N_0S_π condition. Detection thresholds decrease by about 10 dB revealing the expected binaural advantage for a target at 1 kHz. The discrimination thresholds, however, are only about 3 dB lower than in the N_0S_0 condition. Apparently, the binaural cues present in the N_0S_π condition do not contribute much to the discriminability between BPN and HTC.

The same conditions presented above were repeated with wide-band (1500 Hz) targets and results are shown at the right side of the panel. For the N_0S_0 condition, detection thresholds are clearly lower for the HTC than for the BPN. It is possible that the temporal structure of the HTC present over a wide range of auditory filters contributes to the improved detectability, although it is not clear what across-channel process is responsible for the improvement.

In this same condition, the discrimination threshold is close to the HTC detection threshold. This can be understood by considering that at low target levels, in the discrimination task, the subject can only hear the HTC and not the BPN and thus is effectively performing a detection task.

For the N_0S_π condition we see again a clear binaural advantage for the detection thresholds. The masking level differences are 17 dB and 10 dB for the BPN and HTC, respectively. In this condition there is no indication for a difference in detection thresholds for the two types of targets. For the discrimination task there is an improvement of about 6 dB compared to the N_0S_0 condition. In the binaural condition, discrimination is now possible at a level where, with monaural listening, neither of the targets would be audible.

Comparing the narrow-band and the wide-band thresholds (for both N_0S_0 and N_0S_π) there is a clear reduction in the discrimination thresholds for the wide band conditions, while detection thresholds increase. This result seems to suggest that across spectrum integration contributes more to the discrimination task than to the detection task, at least for the dichotic condition.

3 Experiment II: the effect of centre frequency

In the previous experiment it was found that discrimination thresholds were lower for larger target bandwidths, while detection thresholds showed the opposite trend. These differences may be due to the target bandwidth but could also be related to frequency specific sensitivity. Here we performed separate measurements with narrowband targets at low and high frequencies in order to differentiate between the contributions of low-, mid-, and high-frequency auditory channels.

3.1 Method and stimuli

This experiment is similar to Experiment I except that the centre frequencies of the targets are 280-Hz and 1650-Hz with bandwidths of 60 Hz and 200 Hz, respectively. For the 280-Hz center frequency, the first of 3 components started at 260 Hz, for the 1650-Hz center frequency, the first of 10 components started at 1560 Hz. Three of the subjects from the first experiment participated in this experiment.

3.2 Results and discussion

The mean results of Experiment II are shown in Fig. 2 together with the mean narrowband results for the same subjects from Experiment I.

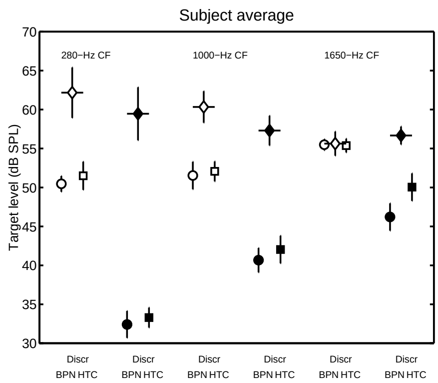

Fig. 2. Average discrimination thresholds (diamonds) and detection thresholds for BPN (circles) and HTC (squares) are shown for N_0S_0 conditions (open symbols) and N_0S_π conditions (filled symbols). The left, middle and right parts of the panel show thresholds for 280-Hz, 1-kHz and 1.65-kHz centre-frequency targets, respectively. Vertical lines indicate averaged standard deviations based on individual subject standard deviations.

For the N_0S_0 conditions (open circles and squares), detection thresholds increase somewhat with center frequency. This result is consistent with the idea that critical bandwidth increases towards high frequencies, which results in more masker energy per critical band. For the same conditions, discrimination thresholds (open diamonds) show a decrease at high frequencies, where they equal detection thresholds.

For the N_0S_π conditions (filled circles and squares), thresholds increase much more towards high-frequencies, in agreement with previous data that show a reduced binaural advantage for high frequencies (e.g. van de Par and Kohlrausch, 1999). The high sensitivity at the lowest frequency suggests that the wideband N_0S_π detection thresholds of experiment I may be based on listening only to the lowest frequency bands containing binaural cues.

Discrimination thresholds (filled diamonds) for N_0S_π show some decrease towards high center frequencies, although this trend differed slightly among subjects. Compared to the N_0S_0 discrimination thresholds, the binaural advantage is largest at low-frequencies. For none of the centre-frequencies, do the N_0S_π discrimination thresholds reach the same low threshold values that are reached in the wide band case of experiment I. Therefore, the low thresholds in the wideband condition cannot be attributed to a high sensitivity in binaural discrimination for one particular frequency range. Apparently, across-spectrum integration of binaural cues leads to a considerable advantage in the discrimination task.

4 Experiment III: the effect of harmonic-tone-complex phase spectrum and F0

From the previous experiments it was concluded that binaural discrimination performance for wideband conditions is better than can be expected based on the individual narrowband conditions. In this experiment we investigate specific properties of the temporal structure of HTC targets that could contribute to the discriminability in wideband conditions.

4.1 Method and stimuli

Two experiments were conducted, similar to the previous experiments, to investigate how changes in the temporal structure of the wide band (1500-Hz bandwidth) HTC target affect discrimination performance under N_0S_π conditions.

In the first set of experiments, the phase relations between the sinusoidal components were varied. In addition to the standard sinusoidal phase structure, also random phase, Schroeder-positive, and Schroeder-negative phase were used. For the latter two conditions the frequency sweep rate resulted in one sweep per period of the HTC across the spectrum of the complex. The random-phase condition was chosen to examine whether or not a target with reduced peakedness would increase discrimination thresholds. The targets with Schroeder phase were chosen to examine the effect of asynchrony. With the rather small separation between partials of 20-Hz, Schroeder-phase targets are peaked in an asynchronous way across auditory filters. If across-channel synchrony is important for discrimination performance, this stimulus is expected to show higher discrimination thresholds than the sinusoidal phase conditions.

In the second set of experiments the influence of frequency separation between partials (F0) was investigated. In addition to the 20-Hz separation, also 40- and 80-Hz separations were used.

In both experiments three subjects of the first experiment participated.

4.2 Results and discussion

The results of Experiment III are shown in Fig. 3. The left panel shows data for different phase conditions of the HTC target. Detection thresholds do not differ significantly for the different phase conditions. The discrimination thresholds, however, show a marked elevation for the random-phase condition, while the Schroeder-phase conditions show only slightly higher discrimination thresholds than the sine-phase condition. Because the random-phase stimuli are the only targets with a low peak-factor, these data suggest that the peakedness of the HTC target is the critical property that enables the low discrimination thresholds observed for wide-band targets. The Schroeder-phase stimuli contain periodic linear frequency sweeps. The data show that synchronicity of the waveform peaks across auditory filters is not needed for a good discrimination performance.

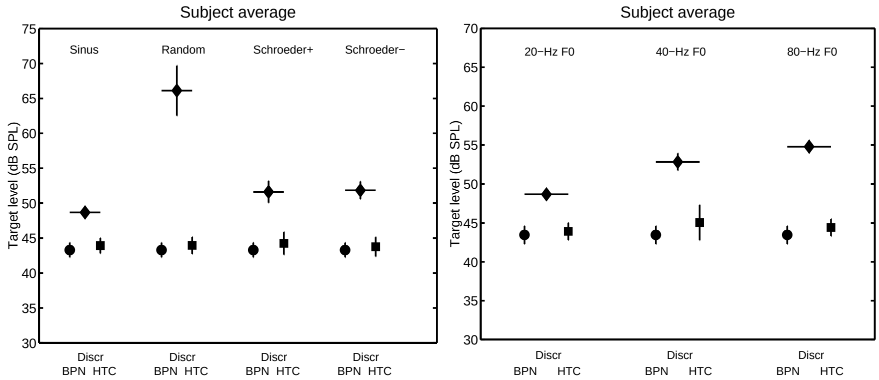

Fig. 3. Average discrimination thresholds (diamonds) and detection thresholds for BPN (circles) and HTC (squares) are shown for N_0S_π conditions. The clusters of data in the left panel show sinusoidal-phase, random-phase, Schroeder-positive, and Schroeder-negative phase conditions. The clusters of data in the right panel show thresholds for targets with F0s of 20-Hz, 40-Hz and 80-Hz, respectively. Vertical lines indicate averaged standard deviations based on individual subject standard deviations.

In the right panel of Fig. 3, N_0S_π data are shown as a function of frequency separation between partials (F0) of the sine-phase HTC. It is clear that detection thresholds are almost constant as a function of F0. Discrimination thresholds on the other hand show an increase at higher F0s where there is an increase in the rate of variation in binaural cues. A rate limitation in the binaural processing of dynamic interaural disparities, which would result in some kind of temporal averaging of these disparities, might explain the increase in discrimination thresholds because short-term changes in the interaural disparities may be the cues used for discriminating the two target types. With such an assumption, *detection* thresholds are not expected to change because the average amount of interaural disparities will not change depending on F0. Such a rate limitation is observed, e.g., in the study by Grantham and Wightman (1979) where the interaural correlation of a masking noise was varied sinusoidally over time and a short interaurally out-of-phase sinusoidal probe had to be detected. For correlation modulations faster than 4 Hz, probe position relative to the masker correlation phase did not influence the measured thresholds. This led the authors to define a minimum binaural integration time of about 44-243 ms. Such an integration time is just long enough to resolve the target with the lowest periodicity (20 Hz) in our experiments.

A second reason for the increase in discrimination thresholds with increasing F0 may be the reduction in peakedness of HTC targets because fewer sinusoidal components will fall within one auditory filter. Again, detection thresholds are not expected to be influenced by such an effect because peakedness is not a factor that seems to influence detection thresholds.

5 Conclusions

It appears that the across-spectrum integration of binaural and monaural information is much better for discrimination conditions than for detection conditions. For detection conditions, thresholds increase towards larger bandwidths while the opposite tends to be the case for the discrimination experiments. An important factor for the rather good discriminability between the two types of wide-band binaural targets seems to be the temporal peakedness of the targets. For random-phase HTC targets with rather weak peakedness, discrimination from noise targets was only possible at very high target levels.

References

Cohen, M.F. (1981) Interaural time discrimination in noise. J. Acoust. Soc. Am. 70, 1289-1293.

Henning, G.B. (1973) Effect of interaural phase on frequency and amplitude discrimination. J. Acoust. Soc. Am. 54, 1160-1178

van de Par, S., and Kohlrausch, A. (1999) Dependence of binaural masking level differences on center frequency, masker bandwidth, and interaural parameters. J. Acoust. Soc. Am. 106, 1940-1947.

Stern Jr., R.M., Slocum, J.E., and Phillips, M.S. (1983) Interaural time and amplitude discrimination in noise. J. Acoust. Soc. Am. 73, 1714-1722.

Grantham, D.W., and Wightman, F.L. (1979) Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation. J. Acoust. Soc. Am. 65, 1509-1517.

Zurek, P.M. and Durlach, N.I. (1987) Masker-bandwidth dependence in homo-phasic and antiphasic tone detection. J. Acoust. Soc. Am. 81, 459-464.