The extent to which a position-based explanation accounts for binaural release from informational masking[a]

Frederick J. Gallun,[b] Nathaniel I. Durlach, H. Steven Colburn,
Barbara G. Shinn-Cunningham, Virginia Best, Christine R. Mason, and Gerald Kidd, Jr.
[illegible]

(Received 17 December 2007; revised 14 April 2008; accepted 15 April 2008)

Detection was measured for a 500 Hz tone masked by noise (an "energetic" masker) or sets of ten randomly drawn tones (an "informational" masker). Presenting the maskers diotically and the target tone with a variety of interaural differences (interaural amplitude ratios and/or interaural time delays) resulted in reduced detection thresholds relative to when the target was presented diotically ("binaural release from masking"). Thresholds observed when time and amplitude differences applied to the target were "reinforcing" (favored the same ear, resulting in a lateralized position for the target) were not significantly different from thresholds obtained when differences were "opposing" (favored opposite ears, resulting in a centered position for the target). This irrelevance of differences in the perceived location of the target is a classic result for energetic maskers but had not previously been shown for informational maskers. However, this parallellism between the patterns of binaural release for energetic and informational maskers was not accompanied by high correlations between the patterns for individual listeners, supporting the idea that the mechanisms for binaural release from energetic and informational masking are fundamentally different.

[DOI: 10.1121/1.2924127]

PACS number(s): 43.66.Ba, 43.66.Dc, 43.66.Nm [RLF] Pages: 439–449

I. INTRODUCTION

A. Informational and energetic masking

The term masking, as it is used in this study, refers to a decrease in the detectability of a target in the presence of an interferer. In "energetic" masking (EM), the interference can be associated with overlap of the target and the interferer acoustic energy or neural activity at a given place of excitation (i.e., the basilar membrane). In "informational" masking (IM), the overlap of excitation between the target and the masker at the auditory periphery is negligible, and the interference is assumed to take place more centrally in the auditory pathway. Obviously, these are two extreme examples and the reality is that the same masker can cause both EM and IM. Because the experiments described here were designed to examine the degree to which the same mechanisms can explain binaural release from these two quite different types of masking, artificial stimuli were constructed that would allow the two types of masking to be examined largely in isolation. It should be noted at the outset, however, that the results reported here may not generalize to masking in which the reduction in performance concerns the discriminability, intelligibility, or identifiability of the target, and the target is supra-threshold. In those cases, the tasks of the listener are different from the detection task described here. For this reason, the mechanisms underlying binaural release from masking may be different as well. Nonetheless, in the interest of starting with the most fundamental case and moving to more complex situations in a systematic manner, this study is concerned with detection in a two-interval forced-choice detection task.

When the listener's task is to detect the presence of a tone of a given frequency, the amount of EM can be estimated by the use of a model (such as estimating the energy passed by filters with widths set to the critical bandwidths specified by Moore and Glasberg, 1983), but the degree of IM is harder to determine. In the majority of cases, the presence of IM is indicated by a rise in threshold or a decrease in performance across two situations for which the EM is the same or even reduced. Durlach et al. (2003a) suggested that the two main sources of IM appear to be the target-masker similarity and stimulus uncertainty. One way of describing these situations is that a target that should be clearly audible is in some way confused with the masker or the masker distracts the listener from the target, resulting in the perception of no target or the misapprehension that the masker is the target. For this reason, the energetic model is insufficient to predict performance in the IM conditions. Furthermore, the amount of individual variability tends to be much greater for IM. This aspect of IM has been modeled through the addition of a filter width parameter and an internal noise parameter, both of which vary across listeners (e.g., Lutfi, 1993; Oh and Lutfi 1998; Durlach et al., 20058 0 TD(Tj/de5J/F5 1 Tf19.9

---

[a] Portions of this research were presented at the 2007 Midwinter Meeting of the Association for Research in Otolaryngology.

[b] Present address: National Center for Rehabilitative Auditory Research, Portland VA Medical Center, 3710 SW US Veterans Hospital Road (NCRAR), Portland, Oregon 97239. Tel.: 503-220-8262. Electronic mail: frederick.gallun@va.gov.

reported modeling results that relied upon negative frequency weightings, which is incompatible with a filter-based energetic model.

B. Binaural hearing

The goal of this study was to examine whether the dif-

As

lationships between the amounts of release obtained by individual listeners for the two masker types raise doubts about whether they truly share a common mechanism that is based on interaural differences rather than on perceived location.

II. METHODS

A. Listeners

Seven female listeners between the ages of 20 and 25 with audiometrically normal hearing were paid for their participation. L4 had considerable prior experience with psychophysical listening but very little with stimuli of this sort. None of the others had experience listening in psychoacoustical experiments. All were graduate students at Boston University in hearing-related disciplines (primarily speech and language pathology).

B. Stimuli

The target to be detected was a 250 ms, 500 Hz tone with 10 ms raised-cosine onsets and offsets. Noise maskers were generated digitally by creating a frequency vector with values spaced at 1 Hz intervals between 100 and 1000 Hz and associating each frequency value with a randomly chosen amplitude and phase value, drawn from rectangular distributions (thus resulting in random but not Gaussian noise). Signals were then converted to the time domain and normalized so that, after attenuation, the overall rms level was 60 dB sound pressure level (SPL) (spectrum level of 31.5 dB SPL). Multitone maskers were generated digitally by choosing from a linear distribution of ten frequencies that fell between 100 and 400 Hz and between 600 and 1000 Hz (leaving a 200 Hz wide “protected region” between 400 and 600 Hz). Each masker frequency was then associated with a random phase value drawn from a rectangular distribution and with an amplitude that was randomly varied within ±5 dB of an arbitrary starting amplitude, also from a rectangular distribution (in decibels). Ten new multitone masker frequencies were chosen randomly before each interval of each trial, always maintaining the 200 Hz protected region. Time-domain conversion and amplitude normalization assured that, after attenuation, the overall rms level of the multitone masker was 70 dB SPL. The maskers, like the targets, were 250 ms in duration with 10 ms raised-cosine-onsets and offsets. The difference in the rms levels of the noise and multitone maskers was initially the result of a programming error but fortuitously led to similar diotic target thresholds for both maskers.

ILD values were introduced into the target by reducing the level at the left or right ear by either 6 dB (“smaller differences”) or 12 dB (“larger differences”). ITD values were introduced by shifting the wave form by either 300 μs, which is equivalent to a phase delay of 54° (smaller differences) or 600 μs, which is equivalent to a phase delay of 108° (larger differences). The target and masker envelopes were applied after the phase shifts, ensuring that the onsets and offsets were synchronized, regardless of interaural diffive interaural configurations arranged in a random order. Each block consisted of either the noise or the multitone masker and either larger or smaller interaural differences. This resulted in average thresholds based on 40 reversals of the adaptive tracks for each masker type for each of the binaural conditions at each size of differences. Because the diotic conditions for the larger and smaller differences were identical, they were combined for analysis, allowing the base line measures of EM and IM to be more accurately measured (80 reversals rather than 40).

E. Calculation of the binaural masking level difference

Colburn and Durlach (1965) defined the BMLD as the ratio of the diotic threshold to the maximum of the levels at the two ears (at threshold) for a given "binaural" condition, expressed in decibels. Thus, the calculation simply involves subtracting the higher of the threshold target levels at the two ears in a given binaural condition (for example, 39 dB) from the threshold target level in the diotic condition (for example, 45 dB). Thus, in this example, the BMLD is 6 dB. For conditions where the BMLD is due entirely to the ITD, this is not problematic. For the ILD conditions, however, this may be a conservative estimate of the BMLD due to the fact that the loudness of a tone presented monaurally is less than that of the same tone presented binaurally (reviewed by Durlach and Colburn, 1978).

Consider the situation where a 12 dB ILD has been introduced by reducing the target level at the right ear by 12 dB but keeping the target level at the left ear the same. If the threshold at the ear with the higher level is unchanged, then the BMLD is 0 dB according to this calculation. If, on the other hand, the threshold is considered to be the level from which one ear is raised by 6 dB and the other lowered by 6 dB, then the BMLD is 6 dB. In addition, if the cue the listener is using is in some way related to the loudness of the target, then the calculation based on the maximum of the levels at the two ears fails to take into account that the listener has now detected a softer target in the ILD condition than in the diotic condition. Presumably, this ability reflects a binaural processing advantage, but the BMLD calculation shows none.

On the other hand, the BMLD is intended to reflect the improvement obtained with two ears relative to performance with a single ear, for which it makes sense to examine changes in the level at the ear with the maximum target level. Using the maximum value calculation both allows comparisons with the results of previous studies and ensures that there is no [illegible] of binaural release from masking simply due to the method of calculating the differences. If the listener in the 12 dB ILD condition made responses based only on the signal at the left ear (which has the most intense target), then the results would be identical for all of the various binaural conditions. A measure based on differences between the most intense target levels presented would give the correct answer in that case, whereas a measure based on any other level would lead to overestimates of the binaural differences (

The BMLD values were also analyzed by performing two correlational analyses. In each case, each pair of values entered corresponded to the BMLDs for an individual listener in the same interaural condition. The first analysis correlated release from the noise masker with release from the multitone masker and was performed separately for larger and smaller interaural differences. The second analysis correlated release based on larger differences with release based on smaller differences and was performed separately for the noise and multitone maskers. The logic behind the first analysis was that perhaps, the nonsignificant effects of masker type were due to variability across listeners. The second analysis was done to determine whether or not the correlational analysis had sufficient power to show a significant difference where one was thought to exist.

The correlation between the BMLDs for the noise and multitone maskers for the smaller differences was nonsignificant ($=0.221$) as was also the case for the larger differences ($=0.223$). On the other hand, significant correlations (<0.01) were found between the BMLDs obtained with larger and smaller differences for both the noise masker ($=0.559$) and for the multitone masker ($=0.767$). Adding listener as a covariate had the effect of increasing the correlations slightly, but the level of significance did not change. These patterns of correlation show that while individual listeners were likely to have similar patterns of BMLDs across larger and smaller interaural differences for a given masker type, the pattern for each individual was not necessarily similar across masker types.

IV. DISCUSSION

These results cause difficulties for a purely position-based account of binaural release from IM because there was not even a trend toward greater BMLDs for reinforcing interaural differences as compared with opposing differences. The similarity between the patterns of masking release obtained with noise maskers and with multitone maskers sug-

starts by estimating the amount of energy falling in the critical-band filter centered on the target tone and then determining the changes in that filter width that would be necessary to produce the thresholds obtained in the experiment for the various listeners.

Using the same software that generated the experimental stimuli, 100 maskers of each type were generated and filtered with a range of filter widths. Figure 4 shows the effective masker level calculated for a range of filter widths and for both masker types. The mean energy through the critical band centered on the 500 Hz target frequency, which Moore and Glasberg (1983) estimated at 76.8 Hz, is marked by the dashed line. In accordance with the fact that the noise band included energy in the region between 400 and 600 Hz while the multitone maskers did not, the average energy falling in the critical band was greater for the noise (49 dB SPL) than ening and narrowing account for the performance of the HighIM listeners (L1, L5, L6, and L7) but not the LowIM listeners. This is consistent with the correlational analysis reported in Sec. IV, where it appears that the LowIM group was using the same mechanism for both masker types, but the HighIM group was not.

This analysis provides support for the hypothesis that at least some listeners were widening and narrowing the bandwidths of their effective filters in response to the maskers presented and the interaural differences imposed on the target. Given that Durlach - (2005) were able to capture much of their data with a band-widening analysis and that the CoRE model of Lutfi (1993) and Oh and Lutfi, (1998) also contains the concept of an effective auditory filter of variable bandwidth, such an approach is certainly worth considering.

Arbogast, T. L., Mason, C. R., and Kidd, G., Jr. (**2002**). “The effect of spatial separation on informational and energetic masking of speech,” J. Acoust. Soc. Am. **112**, 2086–2098

Best, V., Ozmeral, E., Gallun, F. J., Sen, K., and Shinn-Cunningham, B. G. (**2005**). “Spatial unmasking of birdsong in human listeners: Energetic and informational factors,” J. Acoust. Soc. Am. **118**, 3766–3773.

Brungart, D. S., Simpson, B. D., and Freyman, R. L. (**2005**). “Precedence-based speech segregation in a virtual auditory environment,” J. Acoust. Soc. Am. **118**, 3241–3251.

Colburn, H. S., and Durlach, N. I. (**1965**). “Time-intensity relations in binaural unmasking,” J. Acoust. Soc. Am. **38**, 93–103.

Colburn, H. S., and Durlach, N. I. (**1978**). “Models of binaural interaction,” in , edited by E. C. Carterette and M. P. Friedman (Academic, New York).

Darwin, C. J., and Hukin, R. W. (**1999**). “Auditory objects of attention: The role of interaural time differences,” J. Exp. Psychol. **25**, 617–629.

Domnitz, R. H., and Colburn, H. S. (**1976**). “Analysis of binaural detection models for dependence on interaural target parameters,” J. Acoust. Soc. Am. **59**, 598–601.

Durlach, N. I. (**1960**). “Note on the equalization and cancellation theory of binaural masking level differences,” J. Acoust. Soc. Am. **32**, 1075–1076.

Durlach, N. I. (**1963**). “Equalization and cancellation theory of binaural masking-level differences,” J. Acoust. Soc. Am. **35**, 1206–1218.

Durlach, N. I. (**1972**). “Binaural signal detection: Equalization and cancellation theory,” in , edited by J. V. Tobias (Academic, New York).

Durlach, N. I., and Colburn, H. S. (**1978**). “Binaural phenomena,” in , edited by E. C. Carterette and M. P. Friedman (Academic, New York).

Durlach, N. I., Mason, C. R., Gallun, F. J., Shinn-Cunningham, B., Colburn, H. S., and Kidd, G., Jr. (**2005**). “Informational masking for simultaneous nonspeech stimuli: Psychometric functions for fixed and randomly mixed maskers,” J. Acoust. Soc. Am. **118**, 2482–2497.

Durlach, N. I., Mason, C. R., Kidd, Jr., G., Arbogast, T. L., Colburn, H. S., and Shinn-Cunningham, B. G. (**2003a**). “Note on informational masking,” J. Acoust. Soc. Am. **113**, 2984–2987.

Durlach, N. I., Mason, C. R., Shinn-Cunningham, B. G., Arbogast, T. L., Colburn, H. S., and Kidd, G., Jr. (**2003b**). “Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity,” J. Acoust. Soc. Am. **114**, 368–379.

Edmonds, B. A., and Culling, J. F. (**2005**). “The role of head-related time and level cues in the unmasking of speech in noise and competing speech,” Acta. Acust. Acust. **91**, 546–553.

Freyman, R. L., Helfer, K. S., McCall, D. D., and Clifton, R. K. (**1999**). “The role of perceived spatial separation in the unmasking of speech,” J. Acoust. Soc. Am. **106**, 3578–3588.

Gallun, F. J., Mason, C. R., and Kidd, G. Jr. (**2005**). “Binaural release from informational masking in a speech identification task,” J. Acoust. Soc. Am. **118**, 1614–1625.

Goupell, M. J., and Hartmann, W. M. (**2006**). “Interaural fluctuations and the detection of interaural incoherence: Bandwidth effects,” J. Acoust. Soc. Am. **119**, 3971–3986.

Goupell, M. J., and Hartmann, W. M. (**2007a**). “Interaural fluctuations and the detection of interaural incoherence. II. Brief duration noises,” J. Acoust. Soc. Am. **121**, 2127–2136.

Goupell, M. J., and Hartmann, W. M. (**2007b**). “Interaural fluctuations and the detection of interaural incoherence. III. Narrowband experiments and binaural models,” J. Acoust. Soc. Am. **122**, 1029–1045.

Hafter, E. R. (**1971**). “Quantitative evaluation of a lateralization model of masking-level differences,” J. Acoust. Soc. Am. **55**, 1116–1122.

Hafter, E. R., Bourbon, W. T., Blocker, A. S., and Tucker, A. (**1969**). “A direct comparison between lateralization and detection under conditions of antiphasic masking,” J. Acoust. Soc. Am. **46**, 1452–1457.

Hafter, E. R., and Carrier, S. C. (**1970**). “Masking-level differences obtained with a pulsed tonal masker,” J. Acoust. Soc. Am. **47**, 1041–1047.

Hafter, E. R., and Carrier, S. C. (**1972**). “Binaural interaction in low-frequency stimuli: The inability to trade time and intensity completely,” J. Acoust. Soc. Am. **51**, 1852–1862.

Hafter, E. R., Carrier, S. C., and Stephan, F. K. (**1973**). “Direct comparison of lateralization and the MLD for monaural signals in gated noise,” J. Acoust. Soc. Am. **53**, 1553–1559.

Hirsh, I. J. (**1948**). “The influence of interaural phase on interaural summation and inhibition,” J. Acoust. Soc. Am. **20**, 536–544.

Jeffress, L. A., Blodgett, H. C., Sandel, T. T., and Wand. coe. coe. coe. coe. Soc. Am.