Binaural release from informational masking in a speech identification task

Frederick J. Gallun, Christine R. Mason, and Gerald Kidd

Citation: The Journal of the Acoustical Society of America **118**, 1614 (2005); doi: 10.1121/1.1984876
View online: https://doi.org/10.1121/1.1984876
View Table of Contents: https://asa.scitation.org/toc/jas/118/3
Published by the Acoustical Society of America

Binaural release from informational masking in a speech identification task[a)]

Frederick J. Gallun,[b)] Christine R. Mason, and Gerald Kidd, Jr.
Communication Disorders and Hearing Research Center, Sargent College, Boston University, 635 Commonwealth Avenue, Boston, Massachusetts, 02215

(Received 10 August 2004; revised 1 June 2005; accepted 2 June 2005)

Binaural release from informational masking (IM) was examined in a speech identification task. Target and masker sentences were processed into mutually exclusive frequency bands, thus limiting energetic masking (EM), and presented over headphones. In a baseline condition, both were presented monotically to the same ear ($T_m M_m$). Despite minimal frequency overlap between target and masker, the presence of the masker resulted in reduced performance, or IM. Presenting the target monotically and the masker diotically ($T_m M_0$) resulted in a release from IM. Release was also obtained by imposing interaural differences in level (ILDs) and in time (ITDs) on the maskers ($T_m M_{\mathrm{ILD}}, T_m M_{\mathrm{ITD}}$). Any masker with a perceived lateral position that differed from that of a truly monaural stimulus resulted in a similar amount of release from IM relative to $T_m M_m$. For binaural targets and maskers ($T_0 M_{\mathrm{ILD}}, T_0 M_{\mathrm{ITD}}$), release was seen whenever ITDs or ILDs differed between target and masker. These results suggest that binaural cues can be very effective in reducing IM. Because mechanisms based on differences in perceived location make predictions that are similar to those of nonlocation-based binaural mechanisms, a variant of the equalization-cancellation model is also considered. [DOI: 10.1121/1.1984876]

PACS number(s): 43.66.Dc, 43.66.Pn, 43.66.Lj, 43.66.Ba [AK] Pages: 1614–1625

I. INTRODUCTION

Kidd *et al.* (2005) reported that with stimuli in which the spectral overlap of target and masker was minimized and the binaural parameters fluctuated randomly, a monaural target was identified more easily with a binaural than a monaural masker. This study attempted to systematically examine if perceived location, which was uncontrolled in that study, could have allowed listeners to distinguish between the auditory stimuli. Situations in which performance is governed by the presence of masker energy in the frequency regions occupied by the target (i.e., "energetic masking" or EM) make it difficult to differentiate the role of perceived location from the effects of nonlocation-based mechanisms that result in binaural release from masking (Colburn and Durlach, 1978, pp. 498–501). Using stimuli that fall into distinct auditory filters reduces the likelihood that observed effects can be explained by mechanisms that improve signal-to-noise ratio within a filter. Stimuli that meet these criteria include the spectrally processed speech described in Arbogast *et al.* (2002) and used in Kidd *et al.* (2005). In those studies, the target and masker were sentences comprised of sets of very narrow, mutually exclusive frequency bands, with which target identification performance was substantially reduced relative to the performance obtained with a noise-masker processed to have the same long-term spectrum as the speech. This additional masking has been attributed to "informational" masking (IM).

For the stimuli used in Arbogast *et al.* (2002) and Kidd *et al.* (2005), IM was due largely to listener uncertainty about which sentence was the target. Unlike with a noise masker, the listener was confronted with two very similar sentences, only one of which was the target. Although the talkers speaking the sentences were always different for masker and target, the narrowband processing disrupted fundamental frequency and intonation information. Both of these factors are known to provide important cues that listeners use to segregate natural competing speech (Darwin and Hukin, 2004). The result of the processing, then, was that listeners were forced to rely upon the temporal correlations between bands and timbral differences to group the bands successfully. Faced with this difficulty, listeners were motivated to make use of any additional information available—such as perceived location—to determine which bands corresponded to the target and which to the masker.

In Kidd *et al.* (2005), listeners were asked to report the keywords associated with a target sentence in the presence of a masker sentence processed into mutually exclusive frequency bands (called "different-band speech" or DBS). When both sentences were presented to the same ear, performance declined significantly relative to when no masker was present. The surprising result that motivated the experiments described here occurred when a second noise masker occupying the same bands as the speech masker (but different bands from the target) was added in the opposite (contralateral) ear. The addition of this "different-band noise" (DBN) masker in the contralateral ear *reduced* the amount of observed masking significantly by *increasing* identification performance. This result was unexpected because it is at odds with the assumption that noise in one ear does not energeti-

[a)]Parts of this research were presented at the 147th meeting of the Acoustical Society of America [J. Acoust. Soc. Am. **115**, 2457(A) (2004)].
[b)]Author to whom correspondence should be addressed; electronic mail: gallun@bu.edu

cally mask information in the opposite ear. Nonetheless, these results resembled those obtained when the noise was added to the same ear as the masking speech. Thus the noise "masked the masker" even though it was in the opposite ear. The reason for this is almost certainly that the two maskers, speech in the target ear and noise in the nontarget ear, shared the same narrow frequency bands. This resulted in binaural cues that were fluctuating randomly within each masker band and were uncorrelated between bands. It seems likely that some aspect of this situation resulted in either decreased intelligibility or—the possibility examined in this study—increased separation for the masking speech.

In general, it has been found that adding noise to the contralateral ear has very little effect on speech recognition performance in the ipsilateral ear. Brungart and Simpson (2002), for example, reported that performance was not changed for an ipsilateral speech identification task when a broadband noise was added to the opposite ear. Similarly, Brungart *et al.* (2005) found that when no speech masker was present, adding different-band noise to the contralateral ear produced performance statistically identical to that obtained when no masker was present. In the cases where adding a contralateral masker does have an impact, the finding is that adding a speech masker to the contralateral ear *increases* the amount of masking caused by an ipsilateral speech masker (e.g., Brungart and Simpson, 2002; Brungart *et al.*, 2005). The main difference between broadband and narrowband maskers is that a broadband masker has the same effect on both target and masker, but a narrowband masker has the ability to affect only one. This may explain why a release from masking has not been seen previously when broadband maskers have been added to the contralateral ear.

Kidd *et al.* (2005) suggested that the interaction of the narrowband DBS and DBN stimuli presented to opposite ears may have created a binaural image that allowed listeners to distinguish monaural-only target bands from masker bands on the basis of differences in perceived location. In addition, the combination of DBS and DBN maskers across ears may have reduced the intelligibility of the speech masker by some form of binaural interaction (for example, binaural summation) thus reducing its ability to compete with the target. The purpose of the experiments described in this paper is to examine whether perceived lateralization derived from narrowband binaural interactions is capable of explaining the finding reported by Kidd *et al.* (2005). Two main difficulties with evaluating the role of lateralization in the original study were that (1) the DBN masker may have been interfering with the intelligibility of the speech masker (second possible explanation above) and (2) it is not clear exactly what perceived lateralizations were associated with those stimuli. Because the DBN and DBS maskers were uncorrelated, the interaural differences that were created would have differed among the bands. Furthermore, the main finding occurred for cases where the target was monaural and the masker was either monaural or was binaural, leaving open the possible complication of a monaural versus binaural effect regardless of any differences that occurred in lateral position. The current experiments were intended to evaluate the above-offered lateral position explanation for the Kidd *et al.* findings while minimizing or controlling for these potentially confounding variables.

II. GENERAL METHODS

The experiments described in this paper were carried out largely simultaneously with the data collected for the experiments described in Kidd *et al.* (2005). The perceived lateralization study (presented as experiment 1) was conducted several months later. The three listeners were the same and the design of the experiments was influenced by the fact that these conditions were introduced as part of a larger group of conditions already in progress.

A. Listeners

Three females, aged 22–25 years, participated in 2 h sessions three to four times each week. All were very familiar with the stimuli through several thousand trials of listening experience prior to the introduction of these conditions. All had audiometric thresholds of 20 dB HL or better in each ear for octave frequencies from 250 to 8000 Hz.

B. Stimuli

The stimuli resembled those of Arbogast *et al.* (2002). Sentences spoken by four male talkers from the coordinate response measure (CRM) corpus of Bolia *et al.* (2000) were used. Each sentence has the structure: "Ready [callsign] go to [color] [number] now." There are eight callsigns, four colors ("white," "red," "green," and "blue") and eight numbers (1–8). Each sentence wave form was digitally preprocessed to restrict the frequency content. Down-sampling from 40 to 20 kHz was followed by a high-frequency emphasis using a first-order high-pass butterworth filter with a cut-off of 1200 Hz. Each wave form was then passed through a bank of fifteen, approximately one-third octave, fourth-order butterworth filters. The filters were evenly spaced on a logarithmic scale from 215 to 4891 Hz with successive center frequencies spaced at a ratio of 1.25, making them approximately one-third of an octave apart as well. Half-wave rectification and low-pass filtering at 50 Hz by a fourth-order butterworth filter extracted the amplitude envelope within each band.

Thus, each wave form was reduced to a set of amplitude envelopes, each corresponding to one of the fifteen frequency bands. Processed sentences were generated by randomly choosing a subset of the fifteen envelopes and using that subset to modulate pure tones with frequencies equal to the center frequencies of the bands, amplitudes of one, and starting phases of zero. Digitally adding the envelope-modulated tones resulted in a sentence with restricted frequency energy and reduced harmonic structure but with temporal content that reflected the original sentence. The processed sentences (when composed of at least six bands and presented at an adequate level) were perfectly intelligible to listeners (see also Brungart *et al.*, 2005) with only a few minutes of exposure—although this is probably due in part to the closed-set nature of the task. After preprocessing, the sentences were up-sampled to 50 kHz, scaled to equal rms amplitude, and saved to disk for later presentation.

Processed sentences were presented via a Tucker-Davis Technologies (TDT) System II hardware array. The digital to analog conversion was at 50 kHz with 16 bits amplitude quantization and all signals were low-pass filtered at 7.5 kHz. The attenuation and signal routing as well as the relative timing of the signals were varied across conditions and are discussed in the following. Stimuli were presented to the listener over TDH-50 headphones in a double-walled sound-attenuating (IAC) booth.

III. EXPERIMENT 1: PERCEIVED LATERAL POSITION

In order to examine whether increasing masker lateralization can cause release from informational masking, it was first important to show that variations in interaural parameters resulted in changes in perceived lateralization for these listeners with these stimuli. Consequently, an experiment was conducted in which listeners were not required to identify the sentences, but simply to rate their lateral positions.

A. Methods

On all trials, the stimulus was a sentence from the CRM corpus processed into six narrow frequency bands using the method described earlier (Sec. II B). In the first experiment, the sentence was presented to the right ear at 50, 60, or 70 dB SPL and the sentence in the left ear was either (1) not present (monaural presentation), (2) presented at a different level from the sentence at the right ear [interaural differences in level (ILD) condition], or (3) presented at the same level as the sentence at the right ear but either advanced or delayed in time relative to the onset of the sentence in the right ear [interaural differences in time (ITD) condition].

For the ILD trials, the left-ear sentence was presented at levels that varied in 10 dB steps from 30 to 70 dB SPL. To calculate the ILD created for a given combination, the level at the left ear (L_L) is subtracted from the level at the right ear (L_R). A positive ILD denotes a situation in which the ILD favors the right ear (L_R exceeds L_L), while a "negative" ILD indicates that L_L exceeds L_R.

For the ITD trials, onset and ongoing ITDs were created by delaying the onset of one of the sentences (left or right) by either 300 or 600 μs. To calculate ITD, the onset time for the left-ear sentence (T_L) is subtracted from the masker onset time for the right-ear sentence (T_R). In cases where the right-ear sentence was delayed, T_L exceeded T_R and the ITD was negative. The ITD values used were 600, 300, 0, −300, and −600 μs.

On 80% of the trials, the speech was binaural. On the remaining trials, the speech was monaural at the right ear. Listeners reported the perceived lateralization of a sound source on a rating scale—a method similar to that of Sayers and Toole (1964). After listening to each sentence, listeners rated the location on a nine-point scale. A midline location was to be indicated with a rating of five, all the way to the left with a rating of one and all the way to the right with nine. No feedback was given and listeners were not informed of the relative likelihoods of monaural and binaural sentences. The precise distribution was random and each listener completed 15 blocks of 50 trials each. Due to the truly random design, the actual numbers of each sentence type were not equal, but this method did result in at least 15 ratings of each combination of interaural parameters for each listener. The monaural signals were rated at least 75 times by each listener.

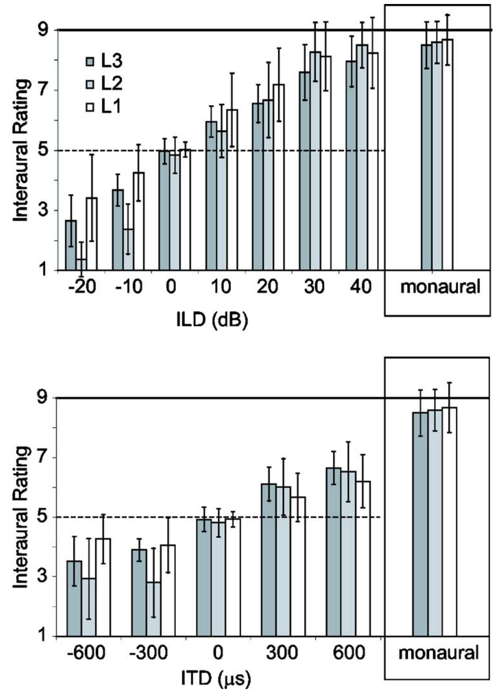

FIG. 1. Average ratings associated with the perceived interaural locations of the stimuli presented in experiment 1. The top panel indicates the ratings associated with ILDs and the bottom those associated with ITDs. Error bars indicate ± one standard deviation around the mean rating.

B. Results

The top panel of Fig. 1 shows the average ILD ratings and the bottom panel the average ITD ratings. The bars on the far right of each panel show the ratings for monaural stimuli. Both data sets show a monotonic increase in ratings that occurs for all subjects for both ILDs and ITDs. The effect of ILD on listener ratings was examined by a two-way ANOVA on masker level at the right and left ears. The effect of level at the right ear was significant [$F_{(2,4)}=51.097$, $p=0.001$] as was level at the left [$F_{(4,8)}=85.939$, $p<0.001$]. The lack of a significant interaction [$F_{(8,16)}=1.929$, $p=0.125$] shows that as the level at one ear increased, the effect of the relative level was not changed. This confirms the use of ILD as a measure of lateralization independent of level at either ear. A two-way ANOVA on masker level and ITD showed no significant effect of level [$F_{(2,4)}=0.544$, $p=0.588$], but a significant effect of ITD [$F_{(4,8)}=27.414$, $p<0.001$]. This confirms that although the range of ratings was smaller than for ILD, the various ITDs were in fact perceived as different.

Regardless of level at the left ear, ILDs between −20 and 20 dB were rated in a manner significantly different from monaural ($p<0.05$) but 30 and 40 dB were not ($p>0.10$). All of the mean ratings at the various ITDs were statistically different ($p<0.01$) from the monaural ratings. Thus, listeners rated all of the ITDs as producing perceptions that differ in laterality from a monaural stimulus, but only stimuli with ILDs of greater than 20 dB are rated as occupying a lateral position similar to that of a monaural stimulus. These results are in accordance with those of Hafter and Kimball (1980), who found that the ability to discriminate a large ILD from a truly monaural signal declines to chance once the ILD exceeds 45 dB. The second set of comparisons involving the ILDs and ITDs used in experiment 3 found that 10 and 20 dB ILDs were rated as different from 0 dB ($p<0.05$), but that while 600 μs was different from 0 μs ($p<0.05$), 300 μs was not ($p=0.07$).

C. Trading ratios

One method of calculating the amount of ILD needed to produce the same lateralization as a given ITD (and vice versa) is by presenting a signal lateralized with one cue (ITD or ILD) and asking listeners to point to the location of the signal with an acoustic pointer that is moved by varying the other cue. The values found with this method tend to fall between 20 and 50 μs/dB [reviewed, along with methods that give divergent results, in Hafter and Jeffress (1968)]. The rating responses from the first experiment can also be used to establish a rough estimate of the "trading ratio" between interaural differences in level and in time. Since ratings of lateral position were assigned to all of the ITDs and ILDs using the same scale, the trading ratio was calculated by performing three linear regressions: first on the ratings of ILD and the ILDs presented, second on the ITD ratings and the ITDs presented, and third, between the predictions generated by the first two functions. This allowed an estimate of which ILD and ITD values result in the same rating of lateral position. For these listeners and these stimuli and equipment, the rating data produced a fairly large tradeoff between ITD and ILD: 47.6, 32.1, and 36 μs/dB for L1, L2 and L3, respectively. The mean was 38.6 μs/dB. In addition, the values associated with a rating of five (midline) were displaced from zero for the three listeners by 103.2, −67.3, and −139.7 μs, respectively. This corresponds to values of 2 to 4 dB in ILD, according to the trading ratios. While large, these values are not outside the expected range, especially given the fact that no ITDs were presented between 0 and 300 μs and no ILDs were presented between 0 and 10 dB. Given this very coarse sampling of interaural differences, the trading ratios obtained seem quite reasonable. Most significantly for interpreting the data from experiments 2 and 3, it can be seen from these values that the maximum ITDs employed, −600 and 600 μs, cover a range of roughly −20 to 20 dB in ILD.

IV. EXPERIMENT 2: IDENTIFICATION, TARGET MONAURAL

The second experiment was designed to examine the ways in which listeners make use of interaural differences in the masker to improve performance on an IM task where the target is monaural. In particular, the goal was to examine whether release from IM can be predicted from the ratings of lateral position obtained in experiment 1.

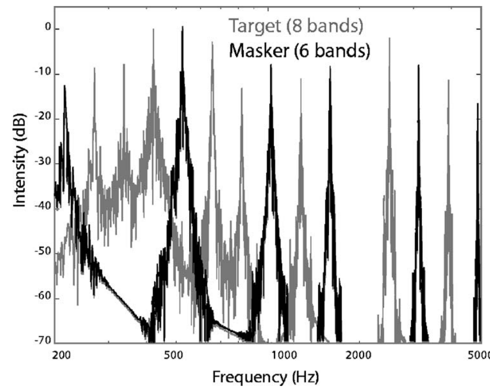

FIG. 2. Example of the magnitude spectra of target (grey) and masker (black).

A. Methods

The task was single-interval forced-choice closed-set speech identification using processed versions of the coordinate response measure (CRM) corpus of spoken sentences (Bolia *et al.*, 2000). Each trial involved presentation of two sentences: a "target" (identified by the presence of the callsign "Baron") and a "masker." Each unprocessed target sentence was first paired with a random masker sentence seven times, resulting in a total of 896 target-masker pairs. Target-masker pairs were constrained to differ on talker, callsign, color, and number. Each sentence was then processed into a collection of randomly chosen narrowband signals using the method described in Sec. II B. In order to ensure that the frequency bands comprising the target and the masker were mutually exclusive, eight target bands were randomly chosen and then six of the remaining seven bands were assigned to the masker. Target and masker-sentence lengths were not equalized. An example of the frequency content of a target and masker pair appears in Fig. 2. It can be seen that the target sentence (grey) and the masker sentence (black) each contain mutually exclusive bands and that the peaks are well separated. For most combinations of bands, the masker was at least 40 dB lower than the target at the center frequency of the target band and was always at least 20 dB lower.

On each trial, the listener was presented with a single target-masker pair and was instructed to report the color and number keywords associated with the target, designated by the presence of the callsign "Baron." In the masker sentence, any of the other callsigns could appear. The four talkers, four colors, and eight numbers resulted in 128 unique target sentences. After stimulus presentation, the listener made two responses, one indicating the color and the other the number associated with the target sentence. The first response (color) had four alternatives and the second response (number) had eight. Response was via a keyboard with the numbers and color names labeled on the keys. Feedback was given via an LCD screen.

Three conditions were tested. In the first condition, the monaural 60 dB SPL target was added to a monaural masker that could be at one of three levels: 50, 60, or 70 dB SPL. This resulted in three T/M ratios at the target ear ("ipsilateral" T/M values): 10, 0, and −10 dB, respectively. Adapting terminology from the binaural masking literature, this condition is referred to as $T_m M_m$: target monotic, masker monotic. The second condition, $T_m M_\alpha$, consisted of presenting the masker sentence to both the right (target or ipsilateral) and left (contralateral) ears with an interaural level difference (ILD). The target sentence was always presented monaurally at 60 dB SPL to the right ear and the masker at the ipsilateral ear was set to 50, 60, or 70 dB SPL. The ILDs were constructed in the same manner as for experiment 1. The range of ILDs was from −20 to 40 dB, although the range was smaller for each ipsilateral T/M due to the constraint that the masker in the contralateral ear would take on the same set of values (30–70 dB SPL) regardless of masker level at the target ear. In the third experimental condition, $T_m M_\tau$, masker ITD was investigated. A target was presented to the right ear at 60 dB SPL while the masker was presented to both ears with no ILD (the masker was presented to the two ears at 50, 60, or 70 dB for any given trial). ITDs were created in the same manner as for experiment 1. Again, the ITD values used were 600, 300, 0, −300, and −600 μs. The levels for each condition were presented randomly within a block of 50 trials. Blocks were generally repeated until a minimum of 100 trials had been obtained at every combination of levels for each listener.

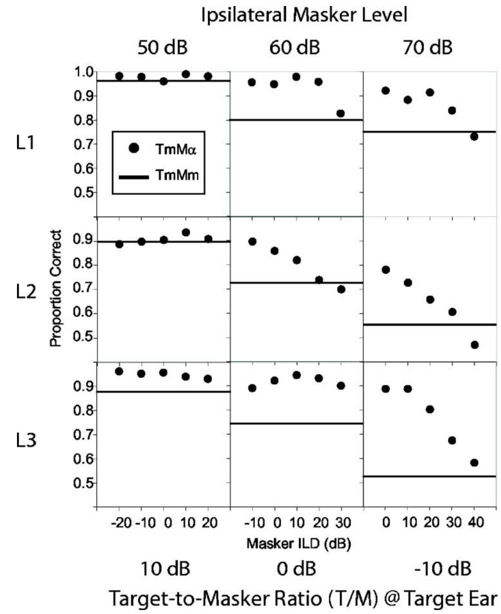

FIG. 3. Proportion correct in the masker ILD condition of experiment 2. Each row of panels contains the results for a single listener and each column contains those for a given T/M. The solid lines represent performance when the masker was presented only at the target ear and the symbols represent performance when the same masking sentence was also presented at the contralateral ear. The range of ILDs varies for the three T/M values because ILD was dependent upon the relationship between the masker levels at the two ears and the range of masker levels at the contralateral ear was always 30–70 dB (see the text).

B. Results

The condition in which target and masker were both monaural, $T_m M_m$, was the same as that reported in Kidd *et al.* (2005), but additional data were collected in order to ensure that improvements were not due to learning effects. A comparison of the first half of the data with the second showed no significant effects. The data were combined to improve the estimates of baseline performance. This led to a greater number of trials for the monaural conditions (~400 points per listener). For each masker level in each condition, proportion correct was calculated for each listener by dividing the number of trials on which a listener was correct (correctly reported both the color and number of the target sentence) by the total number of trials at that masker level for that condition. ANOVAs were conducted on these "proportion correct" data.[1] The data presented as $T_m M_m$ are the same as those reported in Kidd *et al.* (2005). The individual data for all three listeners for the ILD condition ($T_m M_\alpha$) are plotted in the successive panels of Fig. 3. Each listener occupies one row of panels and the columns of panels indicate the level of the ipsilateral masker (50, 60, or 70 dB). The solid horizontal line in each panel indicates performance in the baseline condition $T_m M_m$. By comparing the solid lines across rows in Fig. 3 it can be seen that decreases in T/M in the target ear are associated with decreases in performance, despite the nonoverlapping target/masker bands. The solid lines are included in this and later figures as a monaural masker baseline reference: Any points that are higher than this baseline indicate release from masking due to the presence of a contralateral masker.

The first dichotic condition, $T_m M_\alpha$, involved a monaural target sentence and a binaural masker sentence. Figure 3 summarizes the effect of changing ILD, the level of which is shown in decibels (dB) along the abscissa for each panel. A two-way repeated-measures ANOVA found significant main effects of both ipsilateral masker level [$F_{(2,4)}=19.10$, $p<0.01$] and contralateral masker level [$F_{(4,8)}=24.27$, $p<0.01$]. The interaction was also significant [$F_{(8,16)}=5.50$, $p<0.01$]. These results show that for a monaural target, it is the ILD that matters, and not just the T/M in the target ear. As can be seen in Fig. 3, both increasing ipsilateral masker level and decreasing contralateral masker level (expressed as increasing ILD) were associated with decreasing performance. The interaction occurred because higher ipsilateral masker levels caused increasing ILD to have a larger impact on performance. Correlations were calculated between the performance values obtained with a 70 dB SPL ipsilateral masker and the mean ratings associated with ILD from experiment 1. The correlations were negative, indicating that for increasing ILDs, ratings increased and performance decreased. The values were −0.753, −0.946, and −0.962 for the listeners L1, L2, and L3. The variation in mean ratings accounted for 79% of the variance in performance (R^2 averaged across the three listeners).

The second dichotic condition, $T_m M_\tau$, explored the effect of masker ITD in the presence of a monaural target. The individual data for all three listeners are plotted in the successive panels of Fig. 4. As in Fig. 3, each listener occupies one row of panels and the columns indicate the level of the dichotic masker (50, 60, or 70 dB in both ears). Again, the solid line in each panel indicates the effect of a monaural masker at that level and points that fall above the line indicate an improvement relative to baseline. The symbols in this case indicate what occurs when the masker is binaural, with

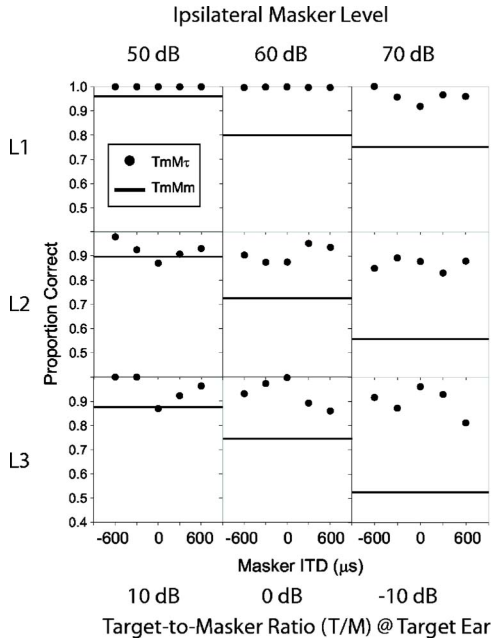

FIG. 4. Proportion correct for the masker ITD condition in experiment 2. Each row of panels contains the results for a single listener and each column contains those for a given T/M. Solid lines indicate performance when only the masker at the target ear was presented and symbols when a masker was presented at both ears at the same level. The ITD values indicate the interaural delay (in microseconds) for the masker. Positive ITDs indicate that the ipsilateral masker led.

the ITD values plotted along the abscissa. Each panel in Fig. 4 indicates that there was substantial release from masking for all of the ITD values used in the study and that, unlike with ILD, changes in masker ITD had essentially no systematic effect. For all three listeners at all three masker levels, the function is somewhat irregular and essentially independent of ITD. Although increasing ITD led to a shift in perceived masker location towards that of the monaural target (see Fig. 2), and the effect of ITD on perceived lateralization was significant in experiment 1, identification performance was largely unaffected. A two-way repeated-measures ANOVA conducted on T/M at the target ear and ITD found a reliable effect of T/M [$F_{(2,4)}$=61.93, $p<0.01$], but no effect of ITD [$F_{(4,8)}$=0.78, p=0.569]. Unlike T_mM_α, ceiling performance at a masker level of 50 dB SPL did not cause an interaction because performance was essentially independent of ITD at all levels. Another difference from T_mM_α is that correlations between performance and mean lateral position rating from experiment 1 were very small, accounting for only 4% of the variance in performance for the 70 dB SPL masker (average R^2 across listeners). To summarize, these results indicate that *all* ITDs produced a release from masking and that shifts toward and away from the location of the target (as shown by the results of experiment 1) were not enough to reliably change the amount of release.

By combining the results of the trading ratio analysis (described in Sec. III C) with the results from experiment 2, it can be seen that the amount of release from masking is dependent largely on the lateral position of the masker. Since the lateral positions associated with the ITD manipulations were restricted to values between −20 and 20 dB, it is not surprising that all the amount of release was quite large across all ITD conditions—release was also quite large for ILDs between −20 and 20 dB.

C. Masking release in dB

To compare our result more directly with those of other researchers, the averaged performance data from experiment 2 were used to plot three-point psychometric functions for each ITD and for a subset of the ILDs (0, 10, and 20 dB). These functions relate performance in the various conditions to the T/M at the target ear. The difficulty is that the slopes of the psychometric functions were much steeper for the monaural masking condition than for the binaural conditions. Unequal slopes result in T/M differences that depend heavily on which performance value is used for the comparison. The performance value associated with 2-down, 1-up adaptive tracking (70.7%) results in extremely large estimates of masking release (20–50 dB for the ILDs and 50–120 dB for the ITDs), suggesting that 70.7% would be an inappropriate value. A more reasonable value is 90%, because it requires little or no extrapolation of the functions—each data set contains a value near 90% correct. In order for a function to be included in this analysis, the correlation between predicted and actual performance values was required to exceed 0.85. "Release from masking" expressed in dB was calculated by subtracting the T/M that was associated with 90% performance in the monaural masker condition from the T/M associated with 90% correct in each binaural condition. When this analysis was performed for the ILD functions, 5.2 dB of release was estimated for the 0 dB ILD, 9.6 dB for the 10 dB ILD, and 10.9 for the 20 dB ILD. The other ILD functions ($-10, -20$, 30, and 40 dB ILDs) had fewer than three points and so accurate psychometric functions could not be estimated. For the ITD functions, the 300 μs ITD function only correlated 0.79 with the actual data and the 0 μs ITD function was nonmonotonic and so only correlated 0.10. Consequently, neither was included. The remaining functions (600, −300, and −600 μs ITDs) resulted in release estimates of 16.1, 21.7, and 24.4 dB, respectively. While large, these are the most conservative estimates possible given this data set.

D. Summary

Experiment 2 found that: (1) with monaural targets, the introduction of binaural maskers resulted in substantial release from masking relative to the monaural maskers (note the large number of points that fall above the horizontal lines in Figs. 3 and 4); (2) the correlations between ratings of lateral position and performance accounted for a large percentage of the variation in performance for the ILD manipulation (when the masker was at 70 dB SPL) but almost none of the variance for the ITD manipulation.

There are two conflicting explanations for these data. The first is that differences in lateral position *do* increase the ability of listeners to distinguish between target and masker bands, but that the lateral position associated with the monaural target was substantially different from the positions associated with all of the ITD values and the majority of the ILD values. This is supported by the finding that all of the lateral position ratings obtained for the ITDs in experiment 1 were statistically different from the monaural ratings. This was also true for many of the ILD values and the few that had similar location ratings to the monaural stimulus accounted for the only reduced performance found in experiment 2.

The second explanation is that differences in lateral position are not useful *per se*, but rather that the ILD values that caused reductions in performance resulted in percepts that were not perceived as binaural. This argument asserts that the majority of the maskers were perceived as binaural and that the very salient difference between a binaural masker and a monaural target was responsible for the improved performance at all of the ITD and the majority of the ILD values. Experiment 3 was designed to gather evidence on the specific question of whether a binaural/monaural difference is required to obtain the high levels of performance seen in experiment 2.

V. EXPERIMENT 3: IDENTIFICATION, TARGET DIOTIC

It is possible that the results from the second experiment depend heavily on the fact that the target was monaural and the maskers binaural. In this experiment, the target and masker were both binaural, allowing a test of whether differences in T/M, ITD, and ILD are all effective at causing a release from masking. This situation is much more similar to real-world listening conditions in which all targets and maskers are usually binaural. Experiment 3 serves as a test of the hypothesis that while a perceptible monaural/binaural difference is sufficient to produce maximal release from masking, it may not be necessary

A. Methods

The target and masker sentences were drawn from the same pool of processed sentences as were used in experiment 2. The task and response method were also identical. The primary difference in experiment 3 was that the target was no longer presented monaurally. In the first condition, T_0M_α, the target was presented diotically (same level at both ears with no interaural delay) at 60 dB and the masker was presented to both ears with an average level of 50, 60, or 70 dB across the two ears. There was no ITD on target or masker in the ILD condition, T_0M_α. Three ILDs were created at each nominal masker level: 0, 10, and 20 dB. For the 0 dB ILD, the masker sentence was at the same level in both ears. When a 10 or 20 dB ILD was imposed, however, the masker level at the left ear was raised and the level at the right ear was lowered. Thus, for an average masker level of 60 dB and a 20 dB ILD, the left ear would receive the masker sentence at 70 dB and the right ear would receive the same sentence at 50 dB. Hence the T/M was different at the two ears. All three average masker levels and all three ILDs were presented in random order on every block of ILD trials. Each listener participated in an average of 100 trials at each combination of average masker level and ILD.

For experiment 3, the target was presented to both ears, so ILDs had to be constructed in order to minimize the mismatch between the T/M at the two ears. To produce average T/M's of -10, 0, and 10 dB and ILDs of 0, 10, and 20 dB, on successive trials the masker at the left ear was presented at levels ranging between 45 and 70 dB (in 5 dB steps) and the masker at the right ear at levels between 55 and 80 dB, also in 5 dB steps. Thus, an average T/M of -10 dB and an ILD of 20 dB was created by presenting the left masker at 60 dB and the right masker at 80 dB. This also created a subset of conditions that allow ILDs to be examined with T/M held constant at one ear.

For the second condition, T_0M_τ, varying ITD, the target was presented as described earlier and the masker was presented to both ears with no ILD (both maskers were presented at 50, 60, or 70 dB SPL for any given trial), hence the same T/M of -10, 0, or 10 dB was present at both ears. The difference between the maskers was created by adding a delay of 300 or 600 μs to the beginning of the masker presented to the right ear. This meant that the masker sentence in the left ear started playing before the one in the right, producing both an onset and an ongoing ITD (as well as an offset ITD). The ITD values tested were 600, 300, and 0 μs relative to the target sentence. All three masker levels and all three ITDs were presented in random order on every block of ITD trials. Each listener participated in an average of 100 trials at each combination of masker level and ITD.

The results of experiment 1 suggested that maskers with these ILDs and ITDs were experienced at a variety of interaural locations and that maskers with ILDs of 10 and 20 dB were rated in a manner statistically different from those with 0 dB ILDs. Ratings of maskers with 600 μs ITDs were statistically different from ratings of those with 0 μs ITDs, but those with 300 μs ITDs were not. This difference seems to have been due to the increased variability of the ratings of masker with ITDs relative to those with ILDs and could perhaps have been overcome with the collection of more data.

B. Results

The masking data will be described as in the second experiment. "Proportion correct" refers to the proportion of the total trials at any one combination of parameters for which both the color and number of the target sentence were correctly reported by a given listener. Figure 5, the results of the T_0M_α condition, shows the effect on performance of introducing a difference in level in the speech masker presented to the two ears when the target sentence was presented diotically. The individual data for all three listeners are plotted in successive panels along each row and the columns of panels indicate the level of the masker. Masker ILD is plotted along the abscissa of each panel, with the 0 dB value corresponding to the case in which there was no interaural difference in masker level (T_0M_0). Performance in T_0M_0 (indicated by the dashed line as well as the leftmost point in each panel) serves as a baseline for the other two ILD values. It also serves as a comparison with T_mM_m (indicated by the solid line in each panel), both of which, it seems reasonable to assume, would lead to identical performance. As can be seen, this was generally the case. The exceptions are discussed in Sec. V C. The question that motivated this portion of the experiment was whether or not masker ILD

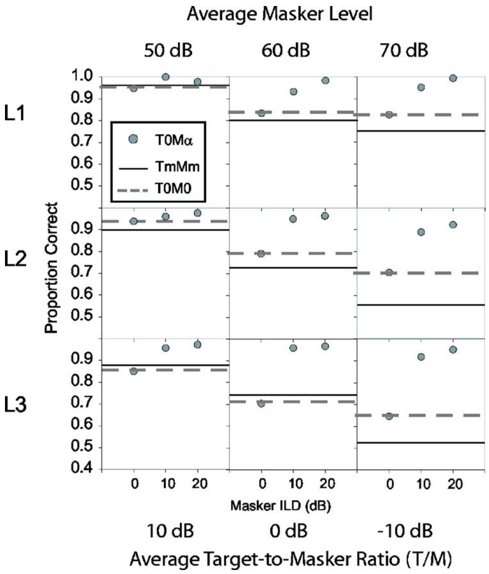

FIG. 5. Proportion correct as a function of masker ILD for the diotic targets in experiment 3. Each row of panels contains the results for a single listener and each column contains those for a given T/M. The solid lines show performance when both target and masker were monaural. The symbol at a masker ILD of 0 (as well as the dashed line) indicates that both target and masker were diotic. The remaining symbols indicate the result of lowering the level at one ear and raising it at the other to produce ILDs of 10 or 20 dB.

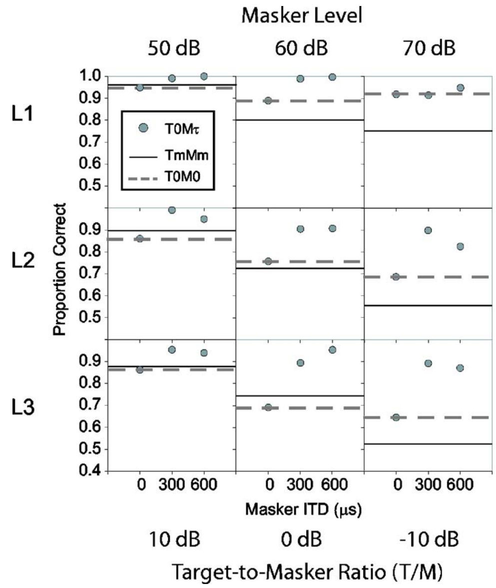

FIG. 6. Proportion correct for the ITD condition in experiment 3. Each row of panels contains the results for a single listener and each column contains those for a given T/M. The solid lines show performance when both target and masker were monaural. The symbol at a masker ITD of 0 (as well as the dashed line) indicates that both target and masker were diotic. The remaining symbols indicate the result of delaying the onset of the masker at one ear in order to produce onset and ongoing masker ITDs of 300 or 600 μs.

supports a release from masking when the target is also binaural. Focusing again on the condition where performance is least influenced by ceiling effects (average T/M of -10 dB) shows that average proportion correct performance in T_0M_0 was 0.73. Creating a 10 dB ILD in the masker increased performance to 0.92 and a 20 dB ILD increased it to 0.95. A repeated measures two-way ANOVA confirmed the effects of T/M [$F_{(2,4)}=15.14$, $p<0.05$] and ILD ($F_{(2,4)}=19.32$, $p<0.01$]. The interaction of the two factors was also reliable [$F_{(4,8)}=22.66$, $p<0.01$]. As in experiment 2, the interaction was due to the near ceiling performance found at large T/M values. Correlations between lateral position ratings and performance for the -10 dB T/M data were 0.998, 0.994, and 0.836 for L1, L2, and L3, respectively. Thus lateral position can account for 90% of the variance in performance (average R^2 across listeners).

Figure 6 shows the performance values obtained for each of the three listeners in the T_0M_τ condition, in which the target was diotic and various ITDs were imposed upon the masker. Each listener's data are confined to a single row and the columns contain all of the data associated with a given masker level (50, 60, or 70 dB). As with the 0 dB masker ILD, the 0 μs masker ITD, T_0M_0 (indicated by the dashed line as well as the leftmost point in each panel) serves as a baseline for the other two ITD values. A comparison between T_mM_m (indicated by the solid line in each panel) and T_0M_0 shows that, as in the ILD conditions, there are a number of panels in which diotic performance is better than monaural, primarily for the T/M of -10 dB (see the discussion in Sec. V C for more on this point). As with ILD, there was substantial release associated with the imposition of masker ITD, indicated by points that fall above the horizontal lines. In the majority of panels, the improvement for a 300 or 600 μs ITD is obvious relative to the 0 μs condition. Those not showing improvement typically have near ceiling performance in T_0M_0. A repeated-measures two-way ANOVA confirmed the effect of T/M [$F_{(2,4)}=15.81$, $p<0.05$] and ITD [$F_{(2,4)}=9.53$, $p<0.05$]. The interaction of the two factors was not reliable in this condition [$F_{(4,8)}=2.06$, $p=0.178$]. Correlations calculated between lateral position ratings and performance in the -10 dB T/M condition were 0.671, 0.654, and 0.954 for L1, L2, and L3, respectively, thus accounting for 60% of the variance in performance (average R^2 across listeners).

C. Masking release in dB

Using the same method as in experiment 2, an estimate of the amount of masking release in dB was obtained by fitting curves to the psychometric functions for the T_0M_0 conditions and then subtracting the T/M associated with 90% correct on those functions from the T/M associated with 90% correct on the functions associated with the dichotic conditions. For these functions, the correlations were required to exceed 0.85. Only the correlation for the 300 μs ITD function, which was nonmonotonic, failed this test. The amount of release estimated from this difference was 27.7 dB for the 10 dB ILD and 73.5 dB for the 20 dB ILD condition. The estimate of release associated with the 600 μs ITD condition was 21.6 dB. Clearly, these estimates are made unreliable by the large differences in the slope of the functions associated with diotic and dichotic maskers. Nonetheless, it is also the case that these values are no less than the values found by Arbogast *et al.* (2002) and they give an approximate idea of the size of the effects expressed in dB. For a more precise assessment, performance should be held constant while masker level is varied or, equivalently, complete psychometric functions should be obtained for all conditions.

D. Better-ear advantage

In comparing binaural to monaural masking, it is important to consider the "better-ear advantage" (Bronkhorst and Plomp, 1988). If a binaural masking situation is such that the T/M at one ear becomes better than it was in the monaural case, improvements in performance could be a simple result of this improvement in T/M rather than a result of binaural processing or a difference in perceived location. For experiment 2, this was not an issue because the T/M at the target ear was independent of binaural manipulations. For the ILD condition in experiment 3, however, the actual T/M at one of the ears was equal to the diotic T/M for that condition plus half of the ILD imposed. Thus, a 20 dB ILD created both an interaural difference and a change in T/M at one ear by 10 dB relative to the T/M in that ear for T_0M_0. In order to help separate the effect of binaural interactions from the effects of an improved T/M at the better ear, there are several relevant performance values that can be compared. If the T/M at the better ear was the sole determining factor, performance would be unaffected by changing the T/M at the contralateral ear. The average performance across listeners for the case where ILD was 0 dB and T/M at both ears was 0 dB was 0.78. When the better ear T/M was kept at 0 dB and the worse ear was lowered to −20 dB, performance increased to 0.95. This improvement of 0.17 cannot be due to a better-ear advantage because the T/M at the better ear remained the same. It cannot be due to a binaural effect based on a time difference because the ITD was 0 μs in both cases. Consequently, it is concluded that, at least for the stimuli used in this experiment, an ILD alone is capable of improving performance substantially and the effect is not mediated by the T/M at the better ear.

E. Summary

The performance obtained when target and masker were diotic was compared to that obtained when masker ILD and ITD were systematically varied. Substantial improvement was found for all listeners at nearly every combination of masker level and interaural difference. In general, less improvement was obtained when the masker level was 50 dB, probably due to ceiling effects. These results suggest that the release from masking found in experiment 2 was not due entirely to the monaural/binaural comparison between target and masker. Rather, it seems that interaural differences *can* support substantial release from IM. As with the first experiment, there was little difference between the release observed for a 10 and a 20 dB ILD or a 300 and a 600 μs ITD. This is in accordance with the trading ratios obtained in Sec. III C, which suggested that these listeners found these values of ILDs and ITDs to be roughly equivalent in terms of the lateral positions they produce.

VI. DISCUSSION

A. Lateral position and release from masking

The primary question driving these experiments was whether or not listeners can use differences in the perceived lateralization of target and masker to improve their performance on a speech identification task. The use of spectrally nonoverlapping stimuli allowed this question to be addressed in a configuration that substantially reduced the amount of energetic masking. This was important because, as discussed by Durlach and Colburn (1978), if the target and masker fall in the same auditory filters then performance that correlates with differences in perceived location is also the performance predicted by nonlocation-based binaural mechanisms. Despite the lack of EM, the amount of binaural release in this study was anticipated to be high based on similar work by Arbogast *et al.* (2002). That study showed that stimuli processed in this way resulted in substantial release from masking when target and masker each were presented from spatially separated loudspeakers compared to both from the same loudspeaker. In that study the stimuli were presented in a mildly reverberant chamber rather than over headphones, meaning that the ITD and ILD information occurred together naturally and the effect of better-ear T/M could not be ascertained. This study had the advantage that it could examine the effects of ITD and ILD in isolation while controlling for better-ear advantage.

The first experiment showed that ILD and ITD manipulations with these stimuli were sufficient to produce a wide range of lateral position ratings. Statistical tests showed that there were significant effects of ILD and ITD on the lateralization ratings obtained from the listeners. A trading ratio between ITD and ILD was estimated for these stimuli and the average value was 38.6 μs/dB. This suggests that the largest ITD values presented (±600 μs) were equivalent to ILD values of only 20 dB.

The results of experiment 2 showed that listener performance was related to whether or not the masker had a binaural configuration that had been reliably rated as different from the monaural target. The correlations between lateral position ratings and performance were high for the ILD conditions in experiment 2, but not for the ITD conditions (all of which supported good performance). The ambiguity of the relationship between release and perceived location raised the possibility that release might be based on the difference between a monaural target and a binaural masker. The results of the third experiment did not support this interpretation, however. When the target was diotic, the introduction of ILDs and ITDs still produced release from masking and the correlations between performance and lateral position accounted for at least 60% of the variance across all listeners and both manipulations.

B. Mechanisms based on perceived location

These results are concordant with those of Carhart *et al.* (1969), who showed that masker ITDs of 800 μs produced release from speech-on-speech masking with wideband stimuli. In that study, four wideband maskers, two of speech and two of modulated noise, were presented at opposing time delays with a diotic target. The introduction of these time delays produced release from masking similar to that obtained when all of the maskers had the same (parallel) time delays. This similarity of release poses problems for any mechanism that relies upon a single compensatory time delay for each frequency channel. Since a single delay would only compensate for one of the two sets of maskers, the authors argue that such a mechanism cannot completely explain the results. Instead, they suggest that release from masking based on signal location acts in parallel with binaural unmasking based on interaural compensation.

There is conflict between our results and those of Culling and Summerfield (1995), however. The stimuli in that study were four noise bursts with narrow bandwidths and different center frequencies that listeners were taught to identify based on their similarity to vowel spectra. If only two were presented, a recognizable vowel was perceived, but if all four were presented at the same level, vowel identification was at chance. Introducing a 390 μs ITD into one pair did not improve performance, but presenting one pair monaurally did cause improvement. This result is in conflict with those of experiment 3 in which a 300 μs ITD was effective in changing the amount of masking obtained relative to the T_0M_0 condition.

Darwin and Hukin (1999) proposed a solution to this apparent contradiction, based on a difference they had observed between segregating one time-varying signal from another (target-word identification) and segregating the steady portions of a complex stimulus (segregating a harmonic from a vowel). While the identification task showed an effect of ITD, the task of segregating a harmonic did not. They suggested that the harmonic was grouped with the other vowel components on the basis of grouping factors such as harmonicity and common-onset and that the ITD of the grouped object was assigned later by a weighted averaging of ITD across its frequency components. The target word, on the other hand, was grouped as a separate object from the remainder of the sentence and thus a difference in ITD between the target and the sentence was able to serve as a cue for tracking one object among others.

The arguments of Darwin and Hukin (1999) could be applied to this study if we assume that listeners used interaural differences to focus attention on the target instead of the masker. This would imply that both sentences were available for processing (consistent with reports of typical error patterns; e.g., Brungart and Simpson, 2002; Arbogast *et al.*, 2002), but that the listeners used interaural differences to track the information in the target and ignore the masker. Evidence that a successful grouping of the frequency components is possible when the target and masker are both presented monaurally is supported by the fact that listeners performed the task at above-chance levels in the T_mM_m condition. Cues that listeners could have been using monaurally included correlations between component envelopes (of which common-onset is an example) as well as "goodness-of-fit" tests on various band combinations, perhaps based on the fact that the target bands were constrained to produce the callsign "Baron" as the second word of the sentence. Once the bands have been successfully grouped, the main difficulty facing the listener is to track the proper sentence. For the monaural case, the only cue that identifies the target sentence is the set of bands that were associated with the word "Baron." Once the callsigns have been uttered, the listener must combat uncertainty about which sentence to track by focusing attention only on the target bands. When there is an interaural difference between target and masker, however, an additional cue is available for tracking the target sentence, an operation that Darwin and Hukin (1999) describe as selective attention.

C. Mechanisms based on cancellation

An alternative argument is that performance is improved by a binaural cancellation operation conceptually similar to the equalization-cancellation (E-C) mechanism (Durlach, 1963; also discussion of "Listener-Min" in Durlach *et al.*, 2003) that acts on the masker bands only, leaving the target bands unaffected. According to this explanation, initial grouping still occurs on the basis of monaural cues as described earlier, but once the masker bands are identified they are cancelled and thus removed. For the stimuli in experiment 2, all masker combinations were binaural and all targets were monaural, so if all the binaural bands could be equalized and cancelled the resulting stimulus would—ideally—contain only the target. While this usage of the E-C model is different from how it has been traditionally considered, it is in accordance with the basic "nulling" analogy from radar operations that was referenced from the earliest discussions of the model (e.g., Durlach, 1963). In most cases, the E-C model has been used to explain how binaural manipulations could lead to reductions in the effectiveness of a masker overlapping the frequencies of the target, a situation in which the sizes of differences between the interaural parameters of the target and masker correlate with the reduction in effectiveness. For these stimuli, neither the predictions nor the results follow this pattern. As long as the binaural masker bands can be successfully identified (and the mechanism can select the equalization parameters for just the masker bands), the masker may be "cancelled." This could explain why changes in the interaural parameters had so little effect on the amount of release obtained.

The only case where there seemed to be a relationship between ILD and binaural release was when the target was monaural and the ILD was so great that the ratings of lateral position were not significantly different from monaural. In this case, the system was unable to generate a release from masking. This limit is actually in agreement with data on release from masking with broadband binaural noise maskers and monaural tonal targets (reviewed in Durlach and Colburn, 1978). In that case, it has been found that although release does occur based on the noise ILD, the amount of release declines with increasing ILD. The limit seems to be similar to that seen in experiment 2—roughly 30 dB. If, as this suggests, the level-equalization simply has a limited range, then it is perhaps the case that the masker can only be completely equalized when the difference between the ears falls within that range. In fact, Breebaart *et al.* (2001) have modeled these data and suggest that, with noise maskers, large ILDs lead to a decorrelation at the two ears. The mechanism is based on peripheral nonlinearities and the suggestion is that the level adjustment alone is no longer sufficient to match the noise at the two ears.[2] If such a mechanism were to operate for stimuli such as those employed in this study, this would allow an E-C approach to explain the ILD data from experiment 2 as well.

Recent work by Akeroyd (2004) examined the central question of whether or not the band parameters can be selected independently, allowing (in our case) cancellation of only the masker. The stimuli were multiple-component signals in wideband noise that was interaurally 180° out of phase from each signal component. In the baseline condition, the noise had no interaural-phase difference (IPD) applied to it (and thus ITD was constant across frequency), while the signal components were each 180° out of phase at the two ears. Applying a constant equalization parameter to these signals would result in optimal cancellation of the noise. The test conditions involved noise that also had a constant IPD, still 180° out of phase from the signal components. In those cases, a single equalization parameter would not permit optimal cancellation of the noise. Nonetheless, the test conditions had thresholds indistinguishable from baseline performance, supporting a version of the E-C model in which the equalization and cancellation process selects equalization parameters independently in each band. This is also the conclusion reached by Edmonds and Culling (2005).

D. The diotic advantage

One aspect of the data that is not captured by either an E-C mechanism or an attentional explanation is the consistent finding that presenting target and masker to both ears results in better performance than presenting them to only one ear. Improvement due to diotic presentation is not the usual pattern of results, but it does occur (Langhans and Kohlrausch, 1992). Several of the factors that tend to result in a diotic improvement were present in this study: a cross-frequency analysis was required and the target had limited spectral overlap with the masker. It is possible that the two ears allow two independent samples to be taken of the stimuli, thus reducing the variability of the estimated signal parameter (as suggested by Zwicker and Henning, 1985 and by Schooneveldt and Moore, 1989). Additionally, as suggested by Langhans and Kohlrausch, interactions between the left and right hearing pathways could be responsible for an increased sensitivity in binaural but not monaural conditions. Another of their suggestions, that the headphones were not actually equivalent, was eliminated by a series of acoustical measurements. None of this explains why a difference should occur primarily for the -10 dB T/M (see Figs. 5 and 6).

E. Better-ear listening

These conditions allowed us to test the hypothesis, initially proposed by Bronkhorst and Plomp (1988), that the effects of interaural level differences can be completely described as better-ear listening. While this describes their data quite well (and that of Hawley *et al.*, 2004), if the only effect of ILD is to change the T/M at the better ear, then for the ILD conditions in experiment 2 of this study there should have been no effect of ILD. This is because T/M at the better ear was held constant as ILD was manipulated. As is clear in Fig. 3, however, ILDs clearly lead to improvements despite the constant T/M at the "better" ear. Further support comes from Shinn-Cunningham *et al.* (2005), who showed that, with stimuli very similar to those used in this study, the improvements in performance obtained by separating the source locations of target and masker contained a large component related to differences in perceived location. In addition, these differences could be obtained by introducing variations in just ILD or just ITD. In fact, combining ILD and ITD information produced no greater release from masking.

The conclusion from the Shinn-Cunningham *et al.* (2005) study as well as from this one is that, for IM, the traditional models of spatial unmasking may encourage researchers to focus on the wrong mechanisms. Since there is relatively little masker energy in the target bands for these stimuli, the decorrelation caused by introducing a target and the effect of improved T/M at one ear may be fairly unimportant compared to traditional binaural conditions in which target and masker occupy the same frequency regions. This parallels the conclusion of Carhart *et al.* (1969) that binaural mechanisms that ignore perceived location may be missing an important part of the explanation for why interaural differences improve performance. It is important to also consider the differences in perceived location that serve to enhance the segregation of two very similar sentences, one of which is the target and the other of which is the masker.

F. Summary

This study found that for a monaural target masker ILDs and ITDs that favored the target ear were just as effective as cues that favored the opposite ear—so long as the lateral positions were rated as significantly different. This is in accord with the data that directly inspired this study, in which a binaural interaction with no long-term correlation between the ears still supported substantial release from IM. The results of experiment 2 can be summarized by the statement that when a sufficiently intense masker is perceived as occupying a lateral position different from that of a monaural stimulus, listeners somehow exploit the pairing of a monaural target with a binaural masker to overcome IM. The third experiment showed that the result is not entirely dependent on distinguishing a binaural masker from a monaural target, however. Substantial release due to the presence of ILDs and ITDs was obtained when the target and masker were both binaural. Two mechanisms were proposed to explain the data. According to the first mechanism, listeners use the perceptual differences arising from the binaural/monaural comparison and/or the differences in perceived location to maintain their attention on the target sentence and thus report the correct words. In the second mechanism, listeners actively suppress the activity caused by the masker by relying on an equalization-cancellation mechanism similar to that suggested by Durlach (1963).

ACKNOWLEDGMENTS

This work was supported by Grant Nos. DC00100, DC04545, DC04663, and F32 DC006526 from NIH/NIDCD and by the Boston University Hearing Research Center. The authors are grateful to Nathaniel Durlach for consultation on the E-C model, to Kelly Egan for her assistance in many aspects of this project, to the helpful comments of associate editor Armin Kohlrausch and two anonymous reviewers and, of course, to our listeners.

[1]It has been argued that a rationalized arcsine transform should be applied to proportion correct data prior to statistical analyses in order to normalize the data (see Studebaker, 1985). For this experiment and for the data in experiment 3, this transformation was performed and the statistical analyses were repeated. There were no differences in the results. In particular, the differences between conditions at T/M values of 10 dB, where performance approached ceiling values, were still not statistically reliable. As the results did not vary from those for the untransformed data, proportion correct is reported throughout for ease of interpretation.

[2]The applicability of the Breebaart *et al.* (2001) model to our data was suggested by Dr. Armin Kohlrausch, to whom we are grateful.

Akeroyd, M. A. (**2004**). "The across frequency independence of equalization of interaural time delay in the equalization-cancellation model of binaural unmasking," J. Acoust. Soc. Am. **116**, 1135–1148.

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.

Bolia, R. S., Nelson, W. T., Ericson, M. A. and Simpson, B. D. (**2000**). "A speech corpus for multitalker communications research," J. Acoust. Soc. Am. **107**, 1065–1066.

Breebaart, J., van der Par, S. and Kohlrausch, A. (**2001**). "Binaural processing model based on contralateral inhibition. I. Model structure." J. Acoust. Soc. Am. **110**, 1074–1088.

Bronkhorst, A. W. and Plomp, R. (**1988**). "The effect of head-induced interaural time and level differences on speech intelligibility in noise," J. Acoust. Soc. Am. **83**, 1508–1516.

Brungart, D. S. and Simpson, B. D. (**2002**). "The effects of spatial separation in distance on the informational and energetic masking of a nearby speech signal," J. Acoust. Soc. Am. **112**, 664–676.

Brungart, D. S., Simpson, B. D., Darwin, C. J., Arbogast, T. L., and Kidd Jr., G. (**2005**). "Across-ear interference from parametrically degraded synthetic speech signals in a dichotic cocktail-party listening task," J. Acoust. Soc. Am. **117**, 292–304.

Carhart, R., Tillman, T. W., and Greetis, E. (**1969**). "Release from multiple maskers: Effects of interaural time disparities," J. Acoust. Soc. Am. **45**, 411–418.

Colburn, H. S. and Durlach, N. I. (**1978**). "Models of binaural interaction" in *Hearing*, Handbook of Perception Vol. IV, edited by E. C. Carterette and M. P. Friedman (Academic, New York).

Culling, J. F. and Summerfield, Q. (**1995**). "Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay," J. Acoust. Soc. Am. **98**, 785–797.

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

Darwin, C. J. and Hukin, R. W. (**2004**). "Limits to the role of a common fundamental frequency in the fusion of two sounds with different spatial cues," J. Acoust. Soc. Am. **116**, 502–506.

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

Durlach, N. I. and Colburn, H. S. (**1978**). "Binaural phenomena" in *Hearing, Handbook of Perception* Vol. IV, edited by E. C. Carterette and M. P. Friedman (Academic, New York).

Durlach, N. I., Mason, C. R., Shinn-Cunningham, B. G., Arbogast, T. L., Colburn, H. S., and Kidd, G., Jr., (**2003**). "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 spatial unmasking of speech: Evidence for within-channel processing of interaural time delay," J. Acoust. Soc. Am. **117**, 3069–3078.

Hafter, E. R. and Jeffress, L. A. (**1968**). "Two-image lateralization of tones and clicks," J. Acoust. Soc. Am. **44**, 563–569.

Hafter, E. R. and Kimball, P. (**1980**). "The threshold for binaural interaction," J. Acoust. Soc. Am. **67**, 1823–1825.

Hawley, M. L., Litovsky, R. Y., and Culling, J. F. (**2004**). "The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer," J. Acoust. Soc. Am. **115**, 833–843.

Kidd Jr., G., Mason, C. R., and Gallun, F. J. (**2005**). "Combining informational and energetic masking for speech identification," J. Acoust. Soc. Am. (in press).

Langhans, A. and Kohlrausch, A. (**1992**). "Differences in auditory performance between monaural and diotic conditions. I. Masked thresholds in frozen noise," J. Acoust. Soc. Am. **91**, 3456–3470.

Sayers, B. McA. and Toole, F. E. (**1964**). "Acoustic-image lateralization judgments with binaural transients," J. Acoust. Soc. Am. **36**, 1199–1205.

Schooneveldt, G. P. and Moore, B. C. J. (**1989**). "Comodulation masking release for various monaural and binaural combinations of the signal, on-frequency, and flanking bands," J. Acoust. Soc. Am. **85**, 262–272.

Shinn-Cunningham, B. G., Ihelfeld, A., Satyavarta, S., and Larson, E. (**2005**). "Bottom-up and top-down influences on masking," Acta. Acust. Acust. (submitted).

Studebaker, G. A. (**1985**). "A 'rationalized' arcsine transform," J. Speech Hear. Res. **28**, 455–462.

Zwicker, E. and Henning, G. B. (**1985**). "The four factors leading to binaural masking-level differences," Hear. Res. **19**, 29–47.