The ability to listen with independent ears[a]

Frederick J. Gallun,[b] Christine R. Mason, and Gerald Kidd, Jr.
Department of Speech, Language and Hearing Sciences, and Hearing Research Center, Boston University, 635 Commonwealth Avenue, Boston, Massachusetts 02215

(Received 9 February 2007; revised 30 July 2007; accepted 6 August 2007)

In three experiments, listeners identified speech processed into narrow bands and presented to the right ("target") ear. The ability of listeners to ignore (or even use) conflicting contralateral stimulation was examined by presenting various maskers to the target ear ("ipsilateral") and nontarget ear ("contralateral"). Theoretically, an absence of contralateral interference would imply selectively attending to only the target ear; the presence of interference from the contralateral stimulus would imply that listeners were unable to treat the stimuli at the two ears independently; and improved performance in the presence of informative contralateral stimulation would imply that listeners can process the signals at both ears and keep them separate rather than combining them. Experiments showed evidence of the ability to selectively process (or respond to) only the target ear in some, but not all, conditions. No evidence was found for improved performance due to contralateral stimulation. The pattern of interference found across experiments supports an interaction of stimulus-based factors (auditory grouping) and task-based factors (demand for processing resources) and suggests that listeners may not always be able to listen to the "better" ear even when it would be beneficial to do so.
[DOI: 10.1121/1.2780143]

PACS number(s): 43.66.Pn, 43.66.Ba, 43.66.Dc [RLF] Pages: 2814–2825

I. INTRODUCTION

A number of influential models of binaural processing contain the assumption that listeners have access to both a monaural and a binaural mode of processing incoming auditory stimulation (e.g., Colburn and Durlach, 1978). In a recent example, Breebaart *et al.* (2001) modeled a wide variety of signal detection experiments by explicitly including both a binaural and a monaural processing path and allowing the model to choose the pathway that leads to the highest signal-to-noise ratio (SNR). While this makes conceptual sense from a bioengineering perspective and can explain a substantial amount of the data on binaural release from masking (see also Zurek, 1993), the evidence does not entirely support the assumption that listeners voluntarily switch between these two fundamentally different modes of processing. Strong evidence in favor of the availability of a monaural listening strategy comes from work on selective listening (e.g., Broadbent, 1958; Cherry, 1953; Moray, 1970; Treisman, 1964, 1969; Wood and Cowan, 1995) in which listeners often show substantial success in reporting the stimuli presented to one ear and ignoring competing stimuli presented to the opposite ear. On the other hand, several more recent studies (e.g., Heller and Trahiotis, 1995; Brungart and Simpson, 2002; 2004; Brungart *et al.*, 2005; Kidd *et al.*, 2003; Gallun *et al.*, 2005; Shub, 2006) have presented evidence suggesting that listeners cannot simply choose the preferred ear and ignore the other.

These recent examples of "contralateral interference" are also problematic because most models of binaural release from masking (e.g., Zurek, 1993) suggest that interaural differences in level (ILD) result in binaural release through the action of the "better-ear" effect, by which it is meant that listeners selectively attend to the ear with the better SNR. In realistic listening situations, this ability to choose a monaural better-ear listening strategy is fundamentally confounded with the use of perceived differences in location generated by ILDs. Whenever the SNR differs at the two ears there is by definition an ILD in the target, the masker, or both. Most models assume that it is the better SNR at one of the ears that drives performance. However, Gallun *et al.* (2005) showed that listeners can exploit target or masker ILDs in order to improve performance even in the absence of a better-ear effect. In that study, the better-ear hypothesis was excluded by manipulations that involved a monaural target and a fixed-level masker at the target ear. Results showed that increasing the masker level at the nontarget ear actually led to improved performance. These findings suggest that at least some of the results that have previously been attributed to better-ear listening may be the result of either (1) perceived differences in location (or some other perceptual cue) generated by the presence of ILDs or (2) a cancellation process similar to that proposed by Durlach (1960, 1963) in his equalization and cancellation (EC) model of binaural unmasking.

Another example of a situation where a change in performance occurred with no change in the SNR at the target ear is found in the results of Kidd *et al.* (2005) in which

[a] Portions of this research were presented at the 2006 Midwinter Meeting of the Association for Research in Otolaryngology.

[b] Current address: National Center for Rehabilitative Auditory Research, Portland VA Medical Center, 3710 SW US Veterans Hospital Road (NCRAR), Portland, OR 97239. Electronic mail: Frederick.Gallun@va.gov

listeners benefited from the addition of a second masker to the nontarget ear. Targets were narrow-band speech stimuli (described in the following) presented monaurally and the first masker was narrow-band speech in nonoverlapping frequency ranges ("different-band speech" or DBS; also described in the following) presented to the target ("ipsilateral") ear. When a second masker consisting of noise matched in frequency to the speech masker ("different-band noise" or DBN; described in the following) was presented to the contralateral ear, performance was usually better than when no contralateral masker was presented. This result presents problems for a better-ear explanation because there was a change in performance even though the stimulus at the "better" ear (the one that contained the target) had not changed. Three distinct hypotheses were initially proposed, all of which focused on how information in nontarget frequency regions at the nontarget ear could have been used to improve performance.

The first potential explanation for the Kidd *et al.* (2005) result is based on "cueing." If listeners are able to process the two ears independently, they might be able to use the information at the nontarget ear to tell them which frequencies to ignore at the target ear. This strategy, which is similar to what Wang and Brown (2006) and Brungart *et al.* (2006) would call an "ideal time-frequency mask," does not imply a binaural combination of information in the usual sense of the term. This type of independent processing has also been referred to in the past as a "contralateral cue" (e.g., Sorkin, 1965; Taylor and Forbes, 1969; Koehnke and Besing, 1992), although in that literature the task was always detection of an exact copy of the signal presented to the contralateral ear. As the data presented in the following show no evidence of this strategy, it will not be discussed further.

The next two "binaural combination" hypotheses are based on the assumption that listeners must choose when separating signals to use the frequency dimension or the dimension of spatial position (in this case, ear of presentation). This need to choose between strategies is similar to an explanation given by Brungart and Simpson (2007) for why speech from a different-sex talker (which resulted in less interference than a same-sex talker when presented alone) caused more interference than a same-sex talker when presented with a contralateral masker. For the stimuli presented in Kidd *et al.* (2005), the obvious strategy is to separate by frequency, since this dimension distinguishes the target from both maskers, which should then result in a binaural combination of the two maskers. The question to be answered is why such a strategy would lead to improved performance over the case where there was only a single speech masker. The first hypothesis is that if the listener separates target from masker on the basis of frequency, then the result is a monaural target composed of one collection of frequencies and a binaural masker composed of another set of nonoverlapping frequencies. If so, then this binaural difference would have acted as did the binaural differences in Gallun *et al.* (2005), allowing listeners to perceptually segregate the target from the masker and/or to use an EC-type operation to improve intelligibility of the target.

The second binaural combination hypothesis is that when the frequency-aligned speech and noise maskers were combined, it was as if they had been presented monaurally and that the resulting combined masker was simply less intelligible due to energetic rather than informational masking. Evidence against this second binaural combination hypothesis comes from the finding that performance was better in the condition where the speech and noise maskers were in opposite ears than in a control condition in which both maskers were presented monaurally to the target ear. While reduced intelligibility may certainly play a role, the difference between the masking exerted by the contralateral and monaural combinations suggests that intelligibility cannot explain the entire effect.

Because it is not possible to distinguish among these various potential mechanisms on the basis of the data of Kidd *et al.* (2005) alone, the initial goal of this study was to consider some conditions in which identifying the optimal strategy becomes more difficult. The first experiment was carried out as an extension of the work presented in Kidd *et al.* (2005) and Gallun *et al.* (2005), involving the same listeners, the same equipment and response style, and the same types of stimuli. The new conditions tested contralateral presentation of noise matched in frequency to the target rather than to the speech masker. As the first experiment primarily provided evidence of contralateral interference, the second and third experiments were designed to further examine which aspects of the first experiment were responsible for this interference. The overall pattern of results suggests that the difficulties listeners experience can be attributed to (1) task-based factors involving a general limit on the number of source segregation strategies that can be employed in any given listening situation (demand for "processing resources"), and (2) stimulus-based factors involving a tendency for the auditory system to treat similar stimuli as if they were generated by a single source ("auditory grouping").

II. METHODS

A. Listeners

Seven listeners participated, six females and one male, all between the ages of 21 and 40 years and all with pure-tone thresholds within 15 dB of audiometric norms (ANSI 2004) at octave frequencies from 250 to 8000 Hz. All were paid listeners who were familiar with psychometric testing and with the type of processed sentences used. All had a minimum of three weeks experience participating 4–6 h per week in similar tasks. Each experiment involved a subset of three to four of these listeners and each took approximately four weeks of listening to complete.

B. Stimuli

Stimuli were sentences processed into narrow frequency bands as described in Arbogast *et al.* (2002). The original speech was taken from the coordinate response measure corpus (Bolia *et al.*, 2000) with the structure: "Ready [callsign] go to [color] [number] now," with eight callsigns, four colors (white, red, green, and blue) and eight numbers (1–8). To

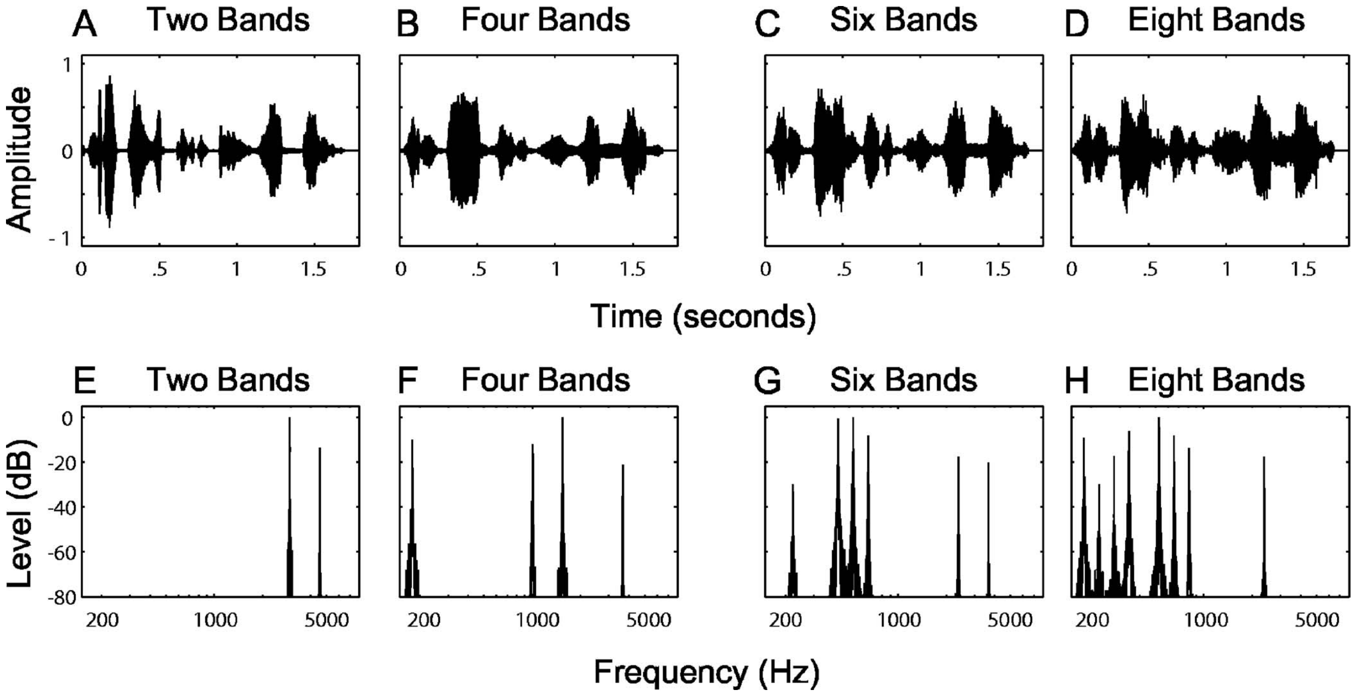

FIG. 1. Examples of wave forms and magnitude spectra for two-band, four-band, six-band, and eight-band target stimuli. Wave form amplitude is shown in arbitrary linear units (e.g., volts). Sound pressure level is displayed in decibels relative to the maximum level for any one band for each target. On each trial the particular bands to be used were chosen randomly from a possible 15. See the text for complete details of the stimulus generation.

restrict the frequency content, sentences were passed through a first-order high-pass Butterworth filter with a cut-off of 1200 Hz to roughly equate energy across the spectrum, after which 15, approximately one-third octave, fourth-order Butterworth filters evenly spaced on a logarithmic scale from 215 to 4891 Hz were used to divide the sentence into 15 narrow bands. Half-wave rectification and low-pass filtering at 50 Hz by a fourth-order Butterworth filter extracted the amplitude envelope within each band, reducing the speech wave forms to a set of 15 amplitude envelopes, each associated with one of the bands.

Processed speech targets and maskers were generated by randomly choosing two, four, six, or eight of the envelopes (depending on the experiment) and using them to modulate a set of pure tones with frequencies equal to the center frequencies of the chosen bands. Presenting the envelope-modulated tones together resulted in a sentence with restricted frequency content and reduced harmonic structure but with the amplitude variations over time that had occurred in those bands in the original sentence. Examples of the four types of processed speech are shown in Fig. 1. Note that on every trial new bands were randomly chosen to compose the target.

The types of processed speech and noise that were used as maskers appear in Fig. 2. In order to distinguish maskers that shared frequency content with the target ("same band") from those that did not ("different band"), the letters SB and DB will be used. For example, "different-band speech" (DBS) was composed only of bands not contained in the target speech and "different-band speech, reversed" (DBSr) was constructed by simply playing DBS wave forms backwards. In all three experiments, the talker, callsign, color, and number used for the masker sentence were all different from those used for the target sentence, but all of the talkers were male. In the first experiment, the target callsign was always "Baron," but in the second and third experiments the target callsign was randomized. To facilitate target sentence identification when the callsign was randomized, the callsign associated with the target sentence was indicated before each trial on the display in front of the listener.

Two types of noise maskers were used. Broadband noise (BBN) maskers were generated in the frequency domain and extended from 20 to 8000 Hz. Narrow-band noise maskers were generated by multiplying processed speech in the frequency domain by BBN. This resulted in noise maskers with the same number of bands as the processed sentences but none of the amplitude modulations necessary for interpreting them as speech. "Same-band noise" (SBN) had the same bands as the target speech, while "different-band noise" (DBN) did not. Four-band example wave forms and magnitude-frequency spectra for these five masker types appear in Fig. 2. Note that in order to ensure that the masker bands were the same or different from the target required a new draw of masker bands on each trial as well. A new draw of BBN was also generated on each trial. For all signals, level in dB SPL was calculated based on the duration of the entire signal.

C. Procedure

All of the stimuli were stored on a computer and played through Tucker-Davis Technology (TDT System II) 16 bit digital-to-analog converters at a rate of 50 kHz, then low-pass filtered at 7.5 kHz. For experiment one, target and masker levels were controlled by independent programmable attenuators (TDT PA4), whereas for experiments two and

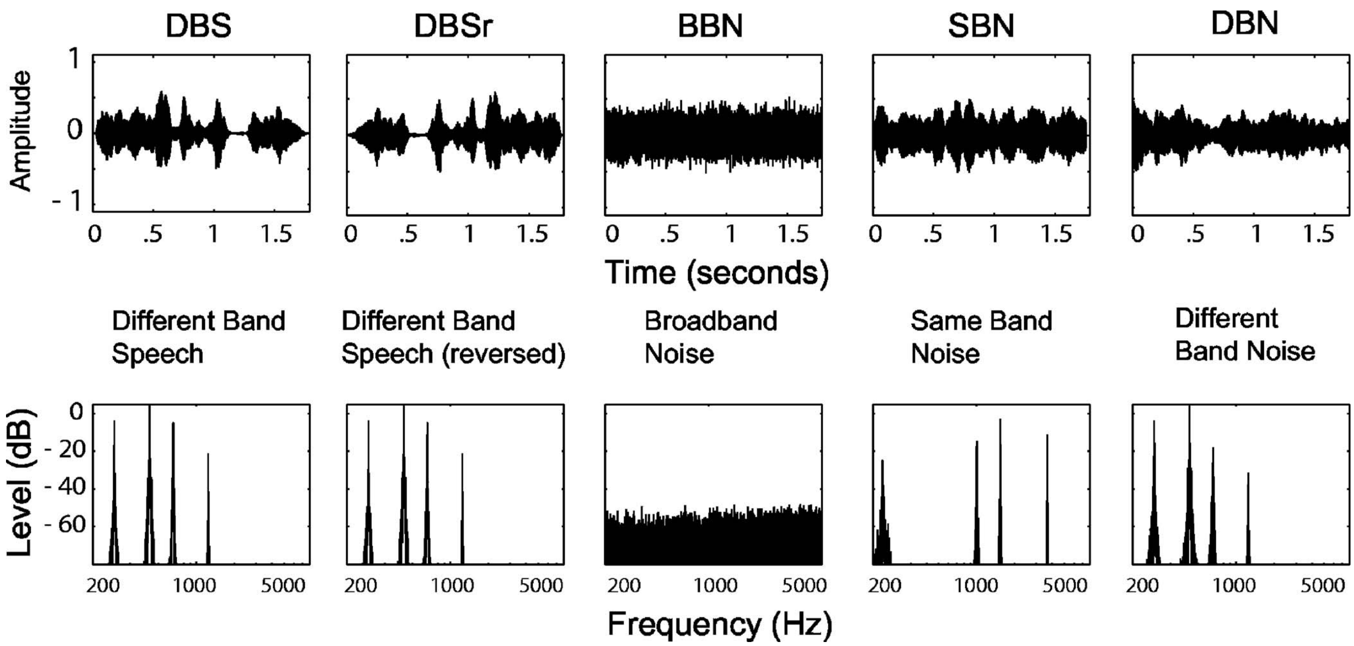

FIG. 2. Examples of wave forms and magnitude spectra of the various maskers used in the three experiments. Each is matched to the four-band target shown in Fig. 1. Wave form amplitude is shown in arbitrary linear units. Sound pressure level is displayed in decibels relative to the maximum level for the four-band target shown in Fig. 1. See the text for complete details of the stimulus generation.

three, relative target and masker levels were set digitally and overall level was controlled by two independent programmable attenuators. Listeners were seated in individual double-walled IAC booths. The stimuli were presented through matched and calibrated TDH-50 earphones. The task of the listener was to identify the color and number from the sentence having the callsign "Baron." Both the color and the number had to be reported accurately for a listener to be correct on each trial. Chance performance was thus about 3% (four colors by eight numbers). Responses and after-trial feedback were given via a handheld device with an LCD screen and buttons corresponding to the possible colors and numbers.

III. EXPERIMENT ONE: CONTRALATERAL SAME-BAND NOISE

The purpose of the first experiment was to examine the degree to which listener performance in a speech-on-speech masking task was affected by the presence of contralaterally presented SBN. Evidence for the influence of an independent masker presented contralaterally came from the improvements found by Kidd *et al.* (2005) when different-band speech (DBS) was presented ipsilaterally and the contralateral masker was DBN. One hypothesis proposed for those improvements was that listeners combined the DBN and DBS into a single binaural signal and made use of the binaural differences generated by a monaural target and a binaural masker. Such a hypothesis would predict that similar improvements should occur for a contralateral SBN noise masker because now the target would be binaural and the DBN masker would be monaural. An additional, but not mutually independent, hypothesis was that the DBN also reduced the intelligibility of the DBS speech masker in the results of Kidd *et al.* (2005). Were this additional hypothesis correct, it would imply that performance should be *reduced* by adding SBN (rather than improved) since now the *target* intelligibility would decrease. The first experiment was primarily designed to give an initial answer as to whether performance improved or declined.

A. Design

The narrow-band processing described in Sec. II was used to generate eight-band targets and spectrally matched eight-band SBN maskers as well as six-band DBS maskers. On each trial, the target and DBS masker were presented to the right ear while the SBN masker was presented to the left ear. The target was presented at 60 dB SPL on all trials. Three levels of DBS were used: 50, 60, and 70 dB SPL. Six levels of SBN were used: 30, 40, 50, 60, and 70 dB SPL and a no-noise control. Within each block of 50 trials, all masker levels were varied randomly from trial to trial and sorted for analysis afterwards. Each listener participated in a minimum of 150 trials for each combination of levels in each condition. Three listeners (L1, L2, and L3) participated immediately after participating in the conditions described in Kidd *et al.* (2005) and Gallun *et al.* (2005).

B. Results

Average performance for the three listeners appears in Fig. 3, with error bars indicating ±1 s.d. across listeners. The large error bars indicate the large differences in the performance of individual subjects that are typical in tasks involving maskers that are separated from the targets in frequency (e.g., Durlach *et al.*, 2005). These same subjects produced similarly divergent levels of performance in the conditions reported in Kidd *et al.* (2005) and Gallun *et al.* (2005), but as with those data sets the patterns of performance were consis-

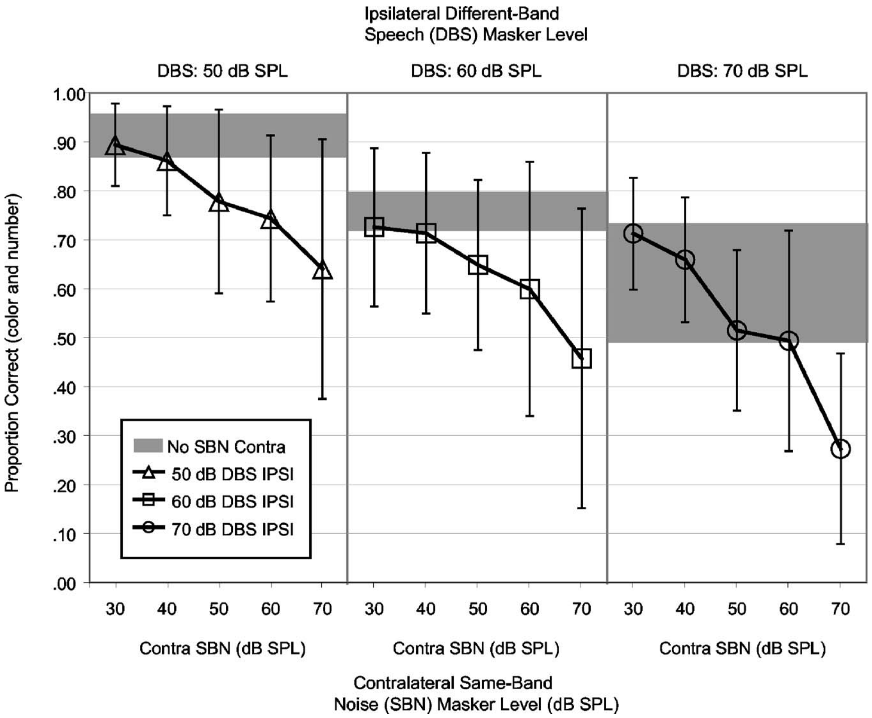

FIG. 3. Average performance for three listeners in experiment one. The listener's task was to identify the color and number keywords contained in an eight-band processed sentence presented to the right ear at a level of 60 dB SPL. Two maskers were always present: an *ipsilateral* (right-ear) different-band sentence (DBS) masker and a *contralateral* (left-ear) same-band noise (SBN) masker. The three panels show performance as a function of contralateral noise level for an ipsilateral sentence masker presented at a level of 50, 60, or 70 dB SPL. The shaded area in each panel indicates ±1 s.d. deviation around mean performance across listeners with no contralateral masker present. Error bars on the symbols show ±1 s.d. across listeners.

tent across listeners. The shaded areas represent the region of performance obtained when no contralateral masker was presented (mean ±1 s.d.).

As can be seen in Fig. 3, the overall trend is that raising either DBS level (successive panels) or SBN level (successive points within a panel) resulted in decreased performance. A repeated-measures analysis of variance (ANOVA) was performed on the percent correct data and DBS level (three values) and SBN level (six values, which includes no SBN masker) were entered as independent factors. The effect of DBS level was statistically significant [$F(2,4)=65.85$, $p<0.001$] as was the effect of SBN level [$F(5,10)=8.95$, $p<0.002$]. The interaction did not reach statistical significance ($p=0.065$).

C. Discussion

The observed decline in performance with increasing contralateral stimulation provides strong evidence that listeners were unable to simply listen to the ear with the higher target-to-masker ratio, as essentially all binaural models would predict. If the matched frequency content of the target and contralateral masker did cause an obligatory binaural combination ("auditory grouping"), then the resulting signal was so degraded by the combination of speech and noise that the improvement due to the binaural differences between target and DBS masker was offset by the reduced intelligibility of the target.

An alternative interpretation, similar to that suggested by Brungart and Simpson (2007) for a finding with unprocessed speech, is that listeners were unable to allocate sufficient processing resources to apply both an ear-based listening strategy and a frequency-based strategy simultaneously and were forced to choose one or the other. This implies that if the ipsilateral different-band speech masker had not been present, then listeners would have been able to segregate the signals at the two ears and performance would not have suffered. It is this hypothesis that is tested in experiment two.

One potential contaminating factor is that of sound conduction through and around the head. At the higher monaural stimulation levels for the contralateral masker there is the chance that the contralateral masker was in fact acting as an ipsilateral masker due to "cross-hearing." If interaural attenuation were only 40 dB, for example, the 50, 60, and 70 dB contralateral same-band noise maskers would have been equivalent to 10, 20, and 30 dB ipsilateral same-band noise maskers. This possibility will also be investigated further in experiment two.

IV. EXPERIMENT TWO: VARIATIONS IN NOISE BAND LOCATION, NUMBER, AND SPECTRUM

Based on the results of the first experiment, it was hypothesized that listeners were unable to process the information at the two ears independently due to the processing demand of also segregating the ipsilateral target and masker by frequency. The implication of this hypothesis is that since the listeners chose to perform the frequency segregation, the frequencies that were similar at the two ears were combined binaurally (despite their dissimilarity in the temporal domain) resulting in a "target," which was actually a mixture of target and SBN masker. It is not difficult to see why this would have led to reduced performance. One possible interpretation of the mixed processing is that a single binaural image was created in which the speech and the noise envelopes were combined, resulting in a less intelligible signal. Another interpretation is that listeners were "distracted" by the noise, although that interpretation seems unlikely given the improvement in performance found by Kidd *et al.* (2005) when the contralateral noise bands corresponded to the ipsilateral speech masker bands. The present experiment is intended to determine whether the apparently obligatory combination of corresponding bands across ears depends on the presence of an ipsilateral segregation task. While the alternative explanation, that the contralateral interference effect only depends on a high degree of similarity between the stimuli in the two ears, is possible, it seems unlikely given the temporal dissimilarity between the processed speech and the same-band noise.

In order to examine this issue more carefully, the ipsilateral DBS masker was removed, and two manipulations of the target and remaining noise masker were introduced. First, the number of bands in the target was varied by presenting two, four, six, or eight bands of speech and, second, either a long-term spectrally-matched noise (SBN) or a BBN was used as a contralateral masker. These two manipulations examined the effects of increasing both the temporal and the spectral differences between the signals at the two ears. Conceptually, these additional differences were regarded as information suggesting that the target and contralateral masker should be treated as independent sources. Under this framework, for every additional frequency band added to the target and to the masking noise, the listener is given additional evidence supporting the conclusion that the amplitude fluctuations in the target and the noise are independent. Thus, the greatest effects of binaural combination should be seen with the fewest numbers of bands. Similarly, since increasing the bandwidth of the noise increases the range of modulation frequencies present in the noise [the high-frequency cut-off is equal to the bandwidth, see Ewert and Dau 2000], the dissimilarity of the modulation spectra at the two ears should be much greater for BBN than for narrow-band noise (SBN).

In order to compare the energetic masking effects of SBN and BBN directly, control conditions were included in which both maskers were presented ipsilaterally instead of contralaterally. These control conditions were also used to provide data relevant to the issue (mentioned in the discussion of experiment one) of possible ipsilateral masking caused by sound conduction through and around the head from the contralateral masker. For this reason, lower noise levels were used ipsilaterally than contralaterally, although the average levels of the broadband and narrow-band noise were equated as described in the following.

A. Design

The narrow-band processing described in Sec. II was used to generate two, four, six, and eight band targets (shown in Fig. 2) and spectrally matched noise (SBN) as well as independent BBN. On each trial, the target was presented to the right ear at an overall level of 50 dB SPL. On each block of 50 trials, the location (ear of presentation), level, and frequency content (SBN or BBN) of the masking noise was kept constant, but the order in which the blocks was run was mixed and the listener was not told in advance precisely which combination to expect. Four new listeners (L4, L5, L6, and L7) participated. None had the experience of the listeners in experiment one, but all had experience with psychophysical testing and had participated in at least one other experiment. In the conditions in which the noise was presented to the nontarget ear (contralateral noise presentation), the noise could be either BBN or SBN (with the number of bands matched to that of the target) and the possible noise levels were 30, 40, 50, 60, or 70 dB SPL. In the conditions in which the noise was presented to the target ear ("ipsilateral" noise presentation), the noise could have the same frequency configurations, but the possible noise levels were 10, 20, 30, 40, and 50 dB SPL. These levels represent 20 dB of simulated interaural attenuation, which is an overly conservative estimate (for these headphones, which are used in clinical audiometry, the conservative estimate of interaural attenuation commonly used is 40 dB). Each listener completed a minimum of 150 trials for each noise level in each condition. The contralateral noise conditions were completed before the conditions in which the noise was presented to the target ear. For comparison, each listener also completed a set of trials in which no noise was presented ("target-alone" condition).

B. Results

The average percent correct data are shown in Fig. 4. The shaded areas represent the region of performance obtained when no masker was present (mean ±1 s.d.). The main effect of number of bands was analyzed through a repeated-measures ANOVA performed on the proportion correct data obtained in the target-alone condition. The effect of number of bands was significant [$F(3,9)=114.327$, $p<0.001$] and paired t-tests showed that performance was the same for the six-band and eight-band targets ($p=0.58$) but that the two-band and four-band targets differed from each other ($p<0.001$) and from the six-band and eight-band targets ($p<0.02$ or lower in all cases).

Because levels were different for the contralateral and ipsilateral maskers, two sets of repeated-measures ANOVAs were performed. Analyses of simple main effects were performed on the individual condition of interest as a general ANOVA showed that there were significant interactions between all the factors. The most important result can be seen by comparing performance as a function of contralateral SBN level for the two-band and eight-band targets. While there is a significant reduction in performance for the two-band target [$F(4,12)=32.93$, $p<0.001$], there is no significant reduction for the eight-band target ($p=0.207$). The four- and six-band targets fall intermediate between these two extremes [$F(4,12)=7.94$, $p<0.01$ and $F(4,12)=3.57$, $p<0.04$]. On the other hand, the effect of level was not significant for the contralateral BBN maskers, regardless of the number of bands in the targets with which they were paired. For both the BBN and SBN presented to the target ear (ipsilaterally), the effect of noise level was significant for all of the numbers of bands ($p<0.01$).

Because so many of the performance values were near ceiling, a four-way repeated-measures ANOVA was performed on the original scores and on a transformed version of the scores, using the Rationalized Arcsine transform (Studebaker, 1985). In order to equate for the different noise levels used, the noise levels were rank-ordered before being entered into the analysis. The results mirrored those for the original scores. Nonetheless, while those conditions that generated scores that were immeasurably high can certainly be

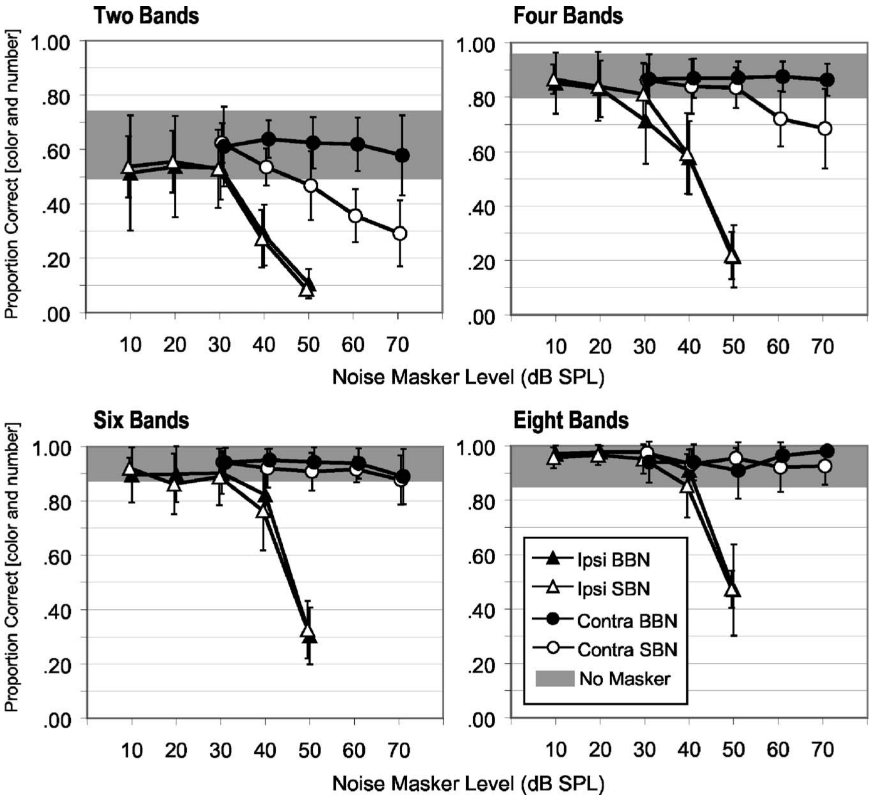

FIG. 4. Average keyword identification performance for four listeners in experiment two as a function of masking noise level, where maskers could be broadband noise (BBN; closed symbols) or same-band noise (SBN; open symbols) and presented to the same ear as the target (ipsilaterally; triangles) or the opposite ear (contralaterally; circles). Targets were presented to the right ear at a level of 50 dB SPL and the four panels show average subject performance for two-, four-, six-, and eight-band targets. The shaded area in each panel indicates ±1 s. d. around mean performance across listeners for that number of target bands with no masker present. Error bars on the symbols show ±1 s. d. across listeners.

regarded as immune to contralateral interference with these stimuli and levels, any comparisons across such conditions should be regarded as specific to these parameters. For example, the fact that no BBN maskers resulted in contralateral interference in this experiment should not be taken to mean that BBN could not potentially exert a contralateral masking effect under different circumstances. Similar care should be taken in interpreting the nonsignificant differences between six- and eight-band stimuli.

C. Discussion

The most important finding obtained from experiment two is that there was an interaction between the number of bands in the target and the amount of contralateral masking that is observed, but only when the long-term spectrum of the masker was matched to that of the target. This has implications for both of the questions that motivated the second experiment. Foremost, this result supports the interpretation that as the temporal similarity between target and masker decreases with increasing numbers of bands, the ability to segregate target from masker increases. This can be considered as support for an auditory grouping explanation of contralateral interference, which refers to the idea that the auditory system is sensitive to the amount of information available indicating that the target and masker are from independent sources. This interpretation is also consistent with the absence of any effect of presenting the broadband noise to the contralateral ear. One possible difficulty with this conclusion is that as the number of bands increases, there is an increase in the number of distinct spectral regions in which both ears are being stimulated. This similarity in spectral information at the two ears could be an indication that the two ears are being stimulated by the same sound source, in which case more bands should lead to increased grouping across the ears. Future work should examine the relative contributions of temporal and spectral similarity in conditions where there is ambiguity as to whether or not various signals should be combined or segregated. It is possible that temporal similarity (or dissimilarity) is simply a more powerful grouping cue than spectral similarity.

In terms of the influence of the ipsilateral speech masker in experiment one, the fact that so little contralateral masking was obtained with the six- and eight-band targets in the second experiment supports the interpretation that listeners are more capable of attending to the ear with the best signal-to-noise ratio if there is only one masker present. Such an interaction of task demands and stimulus configurations suggests that in order to understand the ability of listeners to process sounds independently at the two ears, it is necessary to consider carefully both the task demands and the nature of the stimuli being presented. For example, while BBN clearly can be an effective masker when presented ipsilaterally (it was just as effective as the SBN when presented to the target ear), when presented to the nontarget ear, there was no evidence of contralateral masking for even the two-band target. It would be interesting to determine whether or not this result holds in the presence of an ipsilateral masker that was spectrotemporally dissimilar to the noise as well as dissimilar to the target.

The data from experiment two also show that the results of the first experiment cannot be dismissed as an example of cross-ear masking. There are three reasons for this. The first is that the BBN and SBN stimuli produced indistinguishable masking patterns when presented to the target ear but significantly different patterns of masking when presented contralaterally. In particular, notice the difference in the slopes of the masking functions for ipsilateral and contralateral presentation. This suggests that fundamentally different mechanisms of masking are active. The second reason is that when presented ipsilaterally, there is no evidence of masking at levels below 30 dB SPL, which is the greatest conceivable level that even a very conservative estimate of interaural attenuation could provide. The third reason that acoustic crossover is ruled out by these results is that while there is clear evidence of ipsilateral masking for the six-band and eight-band targets, there is no evidence of contralateral masking at levels that produce contralateral masking in the two-band and four-band stimuli. These three findings provide clear evidence that the contralateral masking observed in experiments one and two was a result of specific interactions between the SBN masker and the target that were not facilitated by acoustic crossover but by binaural processing at higher auditory centers.

A final point regarding the results of experiment two concerns the difference between the masking exerted by the contralaterally presented BBN and SBN. Since the two maskers were equated for overall level, rather than level within the target band (as is shown clearly in Fig. 2), it is possible to argue that the different results for the contralateral maskers resulted from different amounts of energy within the target bands. This suggestion is incompatible with the fact that (as can be seen in all panels of Fig. 4) when the BBN and SBN were presented ipsilaterally they resulted in nearly identical performance (and there was substantial masking, so ceiling effects cannot be thought to have influenced the data). This suggests that the difference in effectiveness in the contralateral presentation was due not to within-channel energy but rather to spectro-temporal similarity to the target.

V. EXPERIMENT THREE: CONTRALATERAL MASKING IN THE PRESENCE OF REVERSED SPEECH

In experiment two, the six- and eight-band targets suffered no measurable interference even when the contralateral noise was at the same levels as for the conditions in experiment one where substantial interference was observed. If the ability to hold the two ears separate depends on task demands, and specifically on the availability of cognitive processing resources, then it is possible the decreased interference resulted from the fact that there were only two tasks to be performed concurrently (interpret the speech target and keep the stimuli at the two ears separate) as opposed to the three tasks that had to be performed in experiment one (interpret the speech, separate the ears, and keep ipsilateral target and masker separate). Note that this argument could also be framed in terms of allocating attention, where attention (or processing resources) must be divided among three tasks in experiment one and between two tasks in experiment two. If a single limited resource is required for both the speech recognition and the source segregation tasks (as described in the model of Navon and Gopher, 1979) then there should be a distinct improvement when listeners go from three tasks to two.

This is similar to saying, as suggested by Brungart and Simpson (2007), that if listeners cannot apply multiple source-segregation strategies simultaneously, then introducing a masker will reduce performance proportional to the degree to which the second masker requires an orthogonal method of source segregation. By this argument, the contralateral interference conditions in experiment two allowed much improved performance over similar conditions in experiment one because now listeners were able to devote a greater proportion of their processing resources to segregating target from masker on the basis of ear of presentation.

In order to examine why there was more contralateral interference for the eight-band target in the first experiment than for the six- or eight-band targets in the second experiment, a subset of the stimuli from experiment two were combined with an additional ipsilateral masker. This manipulation was expected to increase task complexity by requiring listeners to perform both an ear-based segregation and a frequency-based segregation. In addition, the stimuli were chosen in a manner that would allow examination of the interactions between the number of tasks to be performed and the spectrotemporal similarity of the target and the masker without the confounding issue of confusions between two intelligible speech stimuli. Different-band speech that had been reversed in time (DBSr, see Fig. 2) was chosen as the ipsilateral masker and six-band speech was chosen as the target. DBSr resembles the DBS masker of experiment one (and thus the target) in terms of its distribution of long-term temporal fluctuations and it matches the DBS masker exactly in terms of long-term average spectral composition. It differs from DBS, however, in that it is unintelligible and as such is less likely to be confused with the target. Thus, the increase in the number of tasks is more controlled because while the listener still must distinguish between temporal variations that form target words and those that do not, the additional task of discarding intelligible words that are not the target is removed. This is worth considering because, indeed, over 90% of all responses in every condition in experiment one included at least one word that had been presented either as a target or a masker. If the entire cause of contralateral masking in experiment one was confusion between the target words and intelligible words from the DBS masker, then the results should resemble those of experiment two instead, and there should be no masking for a six-band target.

A. Design

The processed six-band speech targets from experiment two were presented to the right ear, always at a level of 50 dB SPL. Six-band DBSr maskers were also presented to the right ear at a level set individually for each listener such that performance with no contralateral stimuli was roughly 85% correct identifications of color and number. The same four listeners from experiment two participated, and the DBSr levels were fixed for the remainder of the experiment as follows: L4:35 dB; L5:50 dB; L6:30 dB; L7:50 dB SPL. Once these levels had been established from psychometric functions obtained in a pilot test, data collection was started in the contralateral masking conditions. Two con-

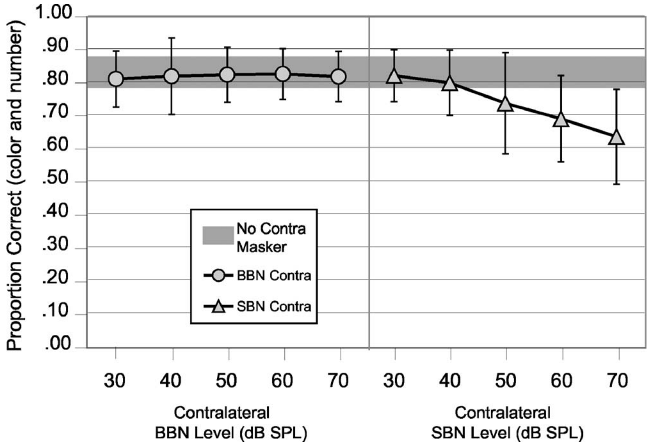

FIG. 5. Average keyword identification performance for four listeners in experiment three. The target was always an eight-band processed sentence presented to the right ear at a level of 50 dB SPL. Six-band different-band speech, reversed (DBSr) was also presented to the right ear (ipsilaterally) in all conditions. DBSr level was set individually for each listener in order to produce roughly equivalent performance when no masker was present (shaded area in each panel). The two panels represent performance as a function of contralateral masker level. Contralateral maskers were either broadband noise (BBN, left panel) or same-band noise (SBN, right panel). The shaded area in each panel indicates ±1 s.d. around mean performance across listeners with no contralateral masker present. Error bars on the symbols show ±1 s. d. across listeners.

tralateral conditions were tested, BBN and SBN (with six matched bands), each at five levels ranging from 30 to 70 dB SPL in 10 dB steps. Masker type and level was fixed for a block of 50 trials but the order of the blocks was randomized. All listeners completed 150 trials at each level for both conditions.

B. Results

Average results are plotted in Fig. 5. The shaded areas represent the region of performance obtained when no contralateral masker was presented (mean ±1 s.d.). A repeated-measures ANOVA was conducted on the data (including performance with no masker, resulting in six contralateral masker levels) and the results showed that there was an interaction between contralateral masker type (the two panels; BBN versus SBN) and contralateral masker level [$F(5,15)=8.92$, $p<0.001$]. Consequently, analyses of the effects of contralateral masker level were conducted separately for BBN and SBN. For BBN, the effect of contralateral masker level did not reach statistical significance [$F(5,15)=0.08$, $p=0.994$], while for SBN, the effect of level was significant [$F(5,15)=7.03$, $p<0.001$]. In order to determine which masker levels were responsible for the significant effect, an analysis was conducted examining the effect of masker type at each contralateral masker level. The results of that analysis indicated that the only significant difference was at a masker level of 70 dB [$F(1,3)=20.785$, $p<0.02$]. This statistical result was undoubtedly due to the wide variability in performance across listeners, as can be seen in Fig. 5.

C. Discussion

Despite the wide range of DBSr levels necessary to obtain sufficiently accurate proportion correct in the baseline task (a range of 20 dB) and the substantial variability across listeners, every listener showed a distinct downward trend in performance with increasing masker level for the SBN masker. This third example of contralateral masking demonstrates that even with an unintelligible masker that should not be easily mistaken for the target, the intelligibility of a monaurally presented six-band target is reduced by the presence of a frequency-matched noise masker presented to the opposite ear but not by a broadband noise. The results of experiment two argue strongly against a cross-ear masking explanation for such an effect. Similarly, the close match between the ipsilateral masking for BBN and SBN seen in experiment two argues against a difference in the within-band effectiveness of the two masker types. Currently, no computational models of binaural processing exist that can account for this pattern of results. While there is much that seems to be under voluntary control, these findings are quite difficult to reconcile with the view that the auditory system always makes use of the information that has the best signal-to-noise ratio. Were that the case, the complete indifference of the system to the stimulus at the left ear that occurred for the contralateral BBN masker would surely have been expected for the SBN masker as well. This was not the case.

The results of experiment three are also important because they support the mounting evidence that there can be substantial contralateral interference in speech identification even when the contralateral masker cannot be confused with the target. Evidence of such interference was also obtained by Brungart and Simpson (2002) with unprocessed reversed speech maskers. They concluded that "a listener's ability to ignore a masker in the contralateral ear is directly related to the difficulty of the within-ear segregation task in the target ear (p. 2993)." Their results (and those reported here) also agree with the conclusion reached by Brungart *et al.* (2005), who suggest that speech-on-speech masking is strongly influenced by "speech-like fluctuations in the spectral envelope" of the masker. This conclusion is supported by the analysis in Kidd *et al.* (2005) that showed that processed-speech maskers with modulation spectra matched to the processed-speech target were more effective than same-band noise maskers that had shallower dips in the envelope but a dissimilar modulation spectrum.

VI. GENERAL DISCUSSION

Taken together, the results of these three experiments provide strong evidence against any simple model of binaural processing in which listeners are always able to base decisions entirely on selecting the ear with the best signal-to-noise ratio. The most important influence on the ability to listen with independent ears is task-based demands on processing resources. In these experiments the task-based factor was the presence of multiple maskers that required more than one segregation strategy to distinguish the target from the maskers. Also important are stimulus-based auditory grouping factors, which can be thought of as the presence of information suggesting that target and masker arise from independent sources. Here, it was suggested that temporal similarity and spectral similarity are both important but that temporal dissimilarity can be more potent than spectral similarity when no additional task is required (although ceiling effects reduced the strength of this conclusion).

The evidence that listeners are not always capable of "turning off" an ear in order to avoid interference has important implications for the interpretation of results in which performance is improved by presenting different target-to-masker ratios at the two ears. In particular, it suggests that any model that assumes that listeners always maximize performance by comparing the output of a binaural and a monaural strategy will be wrong under some listening conditions. This problem has also been noted by Shub and his colleagues (Shub and Colburn, 2004; Shub *et al.*, 2005; Shub, 2006). What remains to be determined is the extent to which observing such nonoptimal performance depends on stimuli that represent unusual deviations from what listeners normally experience in real-world environments. For one very important group of listeners at least, those with cochlear implants, the sort of narrow-band processing and monaural presentation used in these experiments are actually quite close to what they experience. This is because the processing is based on simulations of the signal transformations that allow the cochlear implant to directly stimulate the auditory nerve with sound information (Shannon *et al.*, 1995; Arbogast *et al.*, 2002). Consequently, these results may contain important information about the sorts of interference that might occur for listeners with two cochlear implants. Even for normal hearing listeners, the factors that were hypothesized to have led to the use of the nonoptimal binaural listening mode are quite common. In particular, the correlation of frequency content across ears is going to be present for most binaural stimuli. It should also be remembered that Brungart and Simpson (2002, 2004), who used unprocessed speech, reported more interference from speech signals than from noise.

The results of these three experiments seem to indicate the presence of two significant factors that give rise to contralateral interference. The first involves the complexity of the listener's task, while the second is based on the auditory system's propensity to treat similar stimuli as if they were generated by a single source in the environment. Task complexity, in this context, refers to what Norman and Bobrow (1975), Navon and Gopher (1979), and Wickens (1984) have all called demands on task-specific processing resources. The suggestion is that the human perceptual system is composed of multiple types of processing resources and that whether or not interference is observed depends upon the degree to which multiple tasks require access to the same resources. In this conceptualization, processing resources refers to everything from input channels or feature analyzers to system processes like retrieval of representations from short-term or long-term memory. Certainly task complexity or multiple demands for processing resources is a factor that would be present at many levels in a real-world environment.

The second factor that seems to have influenced listener performance in these experiments is the spectrotemporal similarity between the targets and the maskers. The aspects of similarity that were varied by the choice of stimuli included spectral similarity (in the use of on- and off-frequency maskers and narrow-band and wideband maskers) and ongoing temporal envelope similarity (as affected by the use of speech or noise envelopes for the maskers and narrow-band or wideband noise maskers).

In experiment one, listeners were provided with an eight-band target, but a different-band ipsilateral speech masker and a contralateral same-band noise masker were also present. To be identified, the target still had to be segregated from the spectrally different bands presented ipsilaterally and the spectrally similar bands presented contralaterally. If the same resources were required for all three tasks, performance would have been reduced by the need to perform all three. Even if the speech identification task required different resources, there would still have been a dual-task to be performed with the two segregation tasks.

In experiment two, the same eight-band target was used (among others), but without the ipsilateral different-band speech masker. Now the same-band noise had no effect when presented contralaterally. The same result was obtained with the six-band target, but significant contralateral masking was seen for the two- and four-band targets. In addition, there was no contralateral masking for any of the targets in the case of the broadband masker. This release from masking due to increasing numbers of bands can be explained if it is first assumed that the demands on processing resources were reduced by removing the ipsilateral speech masker. This allowed near-baseline performance to be achieved with the six- and eight-band targets. Because the number of target bands was reduced, however, demands on resources were increased, as demonstrated by performance levels obtained with no masker present (grey bands in Fig. 4). For the two- and four-band targets, the masking observed in the presence of the narrow-band maskers may be evidence that listeners were less able to achieve the segregation of the signals at the two ears. If some of the resources necessary to perform the ear-based segregation were being used to understand the two- and four-band speech, then it makes sense that performance declined in the presence of the contralateral maskers. What still needs to be explained is the difference between the BBN and the SBN, with no masking whatsoever occurring for the BBN.

While all of the stimuli were similar in their temporal onsets, which would have encouraged grouping by common onset (cf., Bregman, 1990), it was only those maskers that were spectrally similar to the targets (SBN as opposed to BBN) that led to contralateral masking. If it takes processing resources to overcome the auditory system's tendency to combine similar stimuli into a single auditory object, it seems plausible that the decreased intelligibility of the two-band target (compare grey bands in Fig. 4) may have drawn resources from the segregation task and thus increased the degree to which the ipsilateral and contralateral bands were combined into a single binaural object. Consequently, listeners may have experienced a binaural percept as they did with the perfectly correlated DBS maskers presented binaurally to subjects by Gallun *et al.* (2005). Because listeners had fewer cues indicating that the SBN and the target were from different auditory sources than they did for BBN and the target, it is possible that a combined binaural signal was created for the two- and four-band targets, but only for the SBN stimuli.

The ongoing changes in envelope would have provided the strongest source of information indicating that this binaural combination should not occur and as this information was increased (by increasing the number of bands), so the amount of contralateral masking decreased. Clearly, more research is required to understand the interaction between the availability of processing resources and degree to which listeners erroneously combine targets and maskers into single objects, thus removing ongoing differences in pitch, location, and intensity between target and masker.

In experiment three, a six-band target was significantly less intelligible in the presence of a masker with speech-like modulations (DBSr) at the ipsilateral ear and a SBN at the contralateral ear. This decrease in performance did not occur with the same ipsilateral DBSr masker and BBN at the contralateral ear. This finding supports the idea that listener performance is influenced both by the processing load imposed, which explains why a six-band target suffered in experiment three but not experiment two, and by the spectrotemporal similarity between target and masker, which explains why the similar SBN hurt performance but the dissimilar BBN did not.

As a final note, it is worth examining the question of whether the large individual differences seen here (and in many other studies of informational masking) should be considered as evidence that different listeners were employing different strategies to solve the tasks with which they were presented. While this is probably an appropriate description of some studies (see, for example, Oxenham *et al.*, 2003), what characterized the use of different strategies in that case was a difference in the *pattern* of results, with some listeners performing similarly on two tasks and others performing quite differently. The listeners in these experiments, however, differed only in the level of performance that they achieved or, in the case of the third experiment, the level of masking noise that was needed to achieve a particular level of performance. Thus, there is evidence that some listeners were better at applying the segregation strategies than others, but not that the actual strategies differed among listeners. It should be noted, however, that these experiments were not designed with the goal of examining individual differences. Future work in this area would benefit from an emphasis on designing experiments and computational models that would allow differences in performance that can be obtained with a single strategy to be clearly differentiated from the use of different strategies by different listeners.

VII. SUMMARY

This series of three experiments provides further support for the conclusion that listeners are not always able to choose a monaural better-ear listening strategy when it would be helpful to do so.

(i) In experiment one, performance decreased relative to performance in the presence of a DBS masker presented ipsilaterally when a SBN masker matched in long-term spectrum to the target was presented contralaterally.

(ii) In experiment two, interference from both ipsilateral and contralateral SBN maskers was observed for two- and four-band targets but only ipsilateral interference occurred for six- and eight-band targets. Only ipsilateral interference was observed for BBN maskers, but it occurred for all numbers of target bands.

(iii) In experiment three, the presence of different-band speech presented ipsilaterally and reversed in time (DBSr) resulted in contralateral interference for SBN maskers but not BBN maskers when the target had six bands and the DBSr was presented at a level that resulted in a minimum amount of masking for all listeners.

These results support the conclusion that the findings of Kidd *et al.* (2005) in which release from DBS masking occurred when a DBN masker was presented contralaterally represented a true case of contralateral interference and not simply cross-ear masking. Nor are those results entirely attributable to the generation of a binaural masker image, although creation of such an image cannot be ruled out. The results of experiment two, in particular, suggest that listeners may indeed combine information across ears, especially when the spectrotemporal similarity of the information at the two ears is high and the processing demands of the task are also high. The results of experiment three provide additional support for the importance of both task-based factors such as processing demands and stimulus-based auditory grouping factors, such as spectrotemporal similarity, in determining whether or not contralateral interference will occur.

ACKNOWLEDGMENTS

The first experiment evolved out of a suggestion by Virginia Richards, to whom we are gratefully indebted. We also would like to thank Antje Ihlefeld and Douglas Brungart for comments on an earlier draft of this manuscript. We are also grateful for the assistance of our listeners as well as to Kelly Egan and Jackie Stachel for assistance with data collection. This work was supported by the Department of Veterans Affairs, Veterans Health Administration, Rehabilitation Research and Development Service through Associate Investigator Award No. C4855H to Frederick Gallun at the National Center for Rehabilitative Auditory Research as well as AFOSR Award No. FA9550-05-1-2005 and by Grant Nos. DC00100, DC04545, and DC04663 from NIH/NIDCD. F.G. was also supported by F32 DC006526 from NIDCD.

ANSI (**2004**). ANSI S3.6-2004, "American National Standard Specification for Audiometers," (American National Standards Institute, New York).

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.

Bregman, A. S. (**1990**). *Auditory Scene Analysis* (MIT, Cambridge, MA).

Broadbent, D. E. (**1958**). *Perception and Communication* (Pergamon, New York).

Brungart, D. S., Chang, P. S., Simpson, B. D., and Wang, D. (**2006**). "Isolating the energetic component of speech-on-speech masking with ideal time-frequency segregation," J. Acoust. Soc. Am. **120**, 4007–4018.

Brungart, D. S., and Simpson, B. D. (**2002**). "Within-ear and across-ear interference in a cocktail-party listening task," J. Acoust. Soc. Am. **112**, 2985–2995.

Brungart, D. S., and Simpson, B. D. (**2004**). "Within-ear and across-ear interference in a dichotic cocktail party listening task: Effects of masker uncertainty," J. Acoust. Soc. Am. **115**, 301–310.

Brungart, D., and Simpson, B. (**2007**). "The effect of target-masker similarity on across-ear interference in a dichotic cocktail-party listening task," J. Acoust. Soc. Am. **122**, 1724–1734.

Brungart, D. S., Simpson, B. D., Darwin, C. J., Arbogast, T. L., and Kidd, G., Jr. (**2005**). "Across-ear interference from parametrically-degraded synthetic speech signals in a dichotic cocktail-party listening task," J. Acoust. Soc. Am. **118**, 292–304.

Cherry, E. C. (**1953**). "Some experiments on the recognition of speech, with one and two ears," J. Acoust. Soc. Am. **25**, 975–979.

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

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., 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.

Ewert, S. D., and Dau, T. (**2000**). "Characterizing frequency selectivity for envelope fluctuations," J. Acoust. Soc. Am. **108**, 1181–1196.

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.

Heller, L. M., and Trahiotis, C. (**1995**). "Interference in detection of interaural delay in a sinusoidally amplitude-modulated tone produced by a second, spectrally remote sinusoidally amplitude-modulated tone," J. Acoust. Soc. Am. **97**, 1808–1816.

Kidd, G., Jr., Mason, C. R., Arbogast, T. L., Brungart, D. S., and Simpson, B. D. (**2003**). "Informational masking caused by contralateral stimulation," J. Acoust. Soc. Am. **113**, 1594–1603.

Kidd, G., Jr., Mason, C. R., and Gallun, F. J. (**2005**). "Combining energetic and informational masking for speech identification," J. Acoust. Soc. Am. **118**, 982–992.

Koehnke, J., and Besing, J. M. (**1992**). "Effects of roving level variation on monaural detection with a contralateral cue" J. Acoust. Soc. Am. **92**, 2625–2629.

Moray, N. (**1970**). "Introductory experiments in auditory time sharing: Detection of intensity and frequency increments," J. Acoust. Soc. Am. **47**, 1071–1073.

Navon, D., and Gopher, D. (**1979**). "On the economy of the human-processing system," Psychol. Rev. **86**, 214–255.

Norman, D. A., and Bobrow, D. G. (**1975**). "On data-limited and resource-limited processes," Cognit Psychol. **7**, 44–64.

Oxenham, A., Fligor, B. J., Mason, C. R., and Kidd, G., Jr. (**2003**). "Informational masking and musical training," J. Acoust. Soc. Am. **114**, 1543–1549.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (**1995**). "Speech recognition with primarily temporal cues," Science **270**, 303–304.

Shub, D. E. (**2006**). "Monaural perception under dichotic conditions," Ph.D. thesis, Health Sciences and Technology, MIT, Cambridge.

Shub, D. E., and Colburn, H. S. (**2004**). "Monaural intensity discrimination under dichotic conditions," J. Assoc. Res. Otolaryngol. (Abstracts).

Shub, D. E., Pogal-Sussman, T. E., and Colburn, H. S. (**2005**). "The effects of distractor frequency on monaural intensity discrimination under monotic and dichotic conditions," J. Assoc. Res. Otolaryngol. (Abstracts).

Sorkin, R. D. (**1965**). "Uncertain signal detection with simultaneous contralateral cues," J. Acoust. Soc. Am. **38**, 207–212.

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

Taylor, M. M., and Forbes, S. M. (**1969**). "Monaural detection with a contralateral cue (MDCC). I. Better than energy detector performance by human observers," J. Acoust. Soc. Am. **46**, 1519–1526.

Treisman, A. (**1964**). "Verbal cues, language and meaning in selective attention," Am. J. Psychol. **77**, 206–219.

Treisman, A. M. (**1969**). "Strategies and models of selective attention," Psychol. Rev. **76**, 282–299.

Wang, D. and Brown, G. J., eds. (**2006**). *Computational Auditory Scene Analysis: Principles, Algorithms and Applications* (IEEE Press/Wiley-Interscience, Hoboken, NJ).

Wickens, C. D. (**1984**). "Processing resources in attention," in *Varieties of attention*, edited by R. Parasuraman and D. R. Davies (Academic Press, New York), pp. 63–102.

Wood, N., and Cowan, N. (**1995**). "The cocktail party phenomenon revisited: Attention and memory in the classic selective listening procedure of Cherry (1953)," J. Exp. Psychol. Gen. **124**, 243–262.

Zurek, P. M. (**1993**). "Binaural advantages and directional effects in speech intelligibility," in *Acoustical Factors Affecting Hearing Aid Performance*, edited by Studebaker, G. and Hochberg, I. (College-Hill Press, Boston).