Within-ear and across-ear interference in a cocktail-party listening task

Douglas S. Brungart[a)]
Air Force Research Laboratory, Wright–Patterson Air Force Base, Ohio 45433-7901

Brian D. Simpson
Veridian, Dayton, Ohio

(Received 2 January 2002; revised 5 August 2002; accepted 5 August 2002)

Although many researchers have shown that listeners are able to selectively attend to a target speech signal when a masking talker is present in the same ear as the target speech or when a masking talker is present in a different ear than the target speech, little is known about selective auditory attention in tasks with a target talker in one ear and independent masking talkers in both ears at the same time. In this series of experiments, listeners were asked to respond to a target speech signal spoken by one of two competing talkers in their right (target) ear while ignoring a simultaneous masking sound in their left (unattended) ear. When the masking sound in the unattended ear was noise, listeners were able to segregate the competing talkers in the target ear nearly as well as they could with no sound in the unattended ear. When the masking sound in the unattended ear was speech, however, speech segregation in the target ear was substantially worse than with no sound in the unattended ear. When the masking sound in the unattended ear was time-reversed speech, speech segregation was degraded only when the target speech was presented at a lower level than the masking speech in the target ear. These results show that within-ear and across-ear speech segregation are closely related processes that cannot be performed simultaneously when the interfering sound in the unattended ear is qualitatively similar to speech. © *2002 Acoustical Society of America.*
[DOI: 10.1121/1.1512703]

PACS numbers: 43.66.Pn, 43.66.Rq, 43.71.Gv [LRB]

I. INTRODUCTION

One of the classic listening tasks in the study of auditory attention is the ‘‘cocktail-party’’ task, in which a listener is asked to extract information from a target speech signal that is masked by one or more simultaneous interfering talkers. Over the past 50 years, researchers have examined a number of different variations of the cocktail-party task [see Bronkhorst (2000) or Ericson and McKinley (1997) for recent reviews of this literature]. One common implementation of the cocktail-party experiment is the monaural listening configuration illustrated in the left panel of Fig. 1. In this configuration, the speech waveform from the target talker (T) is mixed together electronically with the speech waveform from the masking talker (M) and the combined signal is presented to the listener via headphones.

Previous experiments have shown that two distinct kinds of masking contribute to interference in the monaural cocktail-party task (Kidd *et al.*, 1998; Freyman *et al.*, 2001, 1999; Brungart, 2001b). ‘‘Energetic masking’’ occurs when the competing speech signals overlap in time and frequency in such a way that the listener is unable to detect some of the acoustic information contained in the target speech. ‘‘Informational masking’’ occurs when the competing speech signals are similar and the listener is unable to segregate the acoustically detectable elements of the target speech from the acoustically detectable elements of the masking speech. Despite the effects of these two kinds of masking, listeners are generally able to perform well in monaural speech segregation tasks with two competing talkers. This segregation is apparently achieved by taking advantage of differences in the characteristics of the competing voices ($F0$, vocal tract length, prosody, overall level, etc.) (Brungart, 2001b; Darwin and Hukin, 2000; Bregman, 1994; Brokx and Nooteboom, 1982) and by exploiting differences in the envelopes of the two speech signals by listening to the target speech ‘‘in the gaps’’ of the envelope of the masking speech (Festen and Plomp, 1990; Bronkhorst and Plomp, 1992).

A second common implementation of the cocktail-party listening task is the ‘‘dichotic’’ listening configuration illustrated in the middle panel of Fig. 1. In this configuration, stereo headphones are used to present the target talker T and the masking talker M to different ears. Because each ear receives an unaltered speech signal, the effects of energetic masking are negligible in the dichotic listening configuration. The effects of informational masking are also greatly reduced because differences in the apparent spatial locations of the talkers can be used to help segregate the competing speech signals (Freyman *et al.*, 2001). Because the effects of energetic and informational masking are greatly reduced in the dichotic listening configuration, performance in the dichotic cocktail-party listening task is generally much better than performance in the monaural cocktail-party listening task. Previous experiments have shown that, under most stimulus conditions, listeners in the dichotic cocktail-party task are able to attend to the signal in the target ear without any measurable interference from masking sounds in the unattended ear. Cherry (1953) found that a listener’s ability to

---

[a)] Electronic mail: douglas.brungart@he.wpafb.af.mil

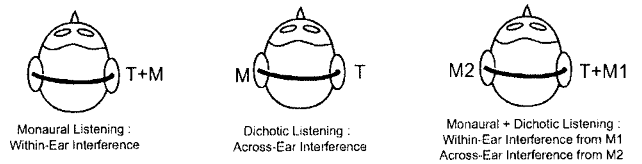

FIG. 1. Three configurations of the cocktail-party effect that isolate the effects of within-ear and across-ear interference. The left panel shows a monaural version of the cocktail-party task that produces only within-ear interference between the target talker T and the masking talker M. The middle panel shows a dichotic version of the cocktail-party task that produces only across-ear interference between talkers T and M. The right panel shows a hybrid version of the cocktail-party task that produces within-ear interference between T and M1 and across-ear interference between T and M2.

shadow an ongoing speech signal presented to one ear was unaffected by the presence of unrelated speech in the unattended ear. Moreover, other researchers have shown that the ability to selectively attend to a single ear extends to the case where multiple talkers are presented in the unattended ear (Drullman and Bronkhorst, 2000) and to dichotic tone detection tasks with a target tone in one ear and a random-frequency ‘‘informational’’ masker in the unattended ear (Neff, 1995; Kidd *et al.*, 1995). There are, however, a few situations where across-ear interference does occur in dichotic listening. When the semantic content of the speech signal in the unattended ear is surprising and unexpected, such as an unexpected occurrence of the listener’s first name (Moray, 1959; Wood and Cowan, 1995; Conway *et al.*, 2001), or related in some way to the signal in the target ear, such as a mid-sentence swap between the signals in the target and unattended ears (Triesman, 1960), errors often occur in the target-ear listening task. The dependence of across-ear interference on the semantic content of the interfering speech suggests that listeners perform some semantic processing on the acoustic signal in the unattended ear. This processing allows listeners to recall general physical characteristics of the speech signal in the unattended ear (such as the sex of the talker) after the completion of a dichotic speech segregation task (Cherry, 1953), but, under most circumstances, it does not result in any appreciable amount of across-ear interference.

One cocktail-party listening configuration that has thus far received relatively little attention is the hybrid configuration shown in the right panel of Fig. 1. In this configuration, a target talker (T) is presented to the listener’s right ear, a masking talker (M1) is presented in the same ear as the target speech, and a second masking talker (M2) is presented in the ear opposite the target speech. This allows a direct examination of any possible interactions between the informational and energetic ‘‘within-ear’’ interference that occurs from a masking talker in the same ear as the target speech and the primarily informational ‘‘across-ear’’ interference that occurs from a masking talker in the ear opposite the target talker. The remainder of this paper describes a series of experiments that were conducted with this hybrid monaural-dichotic cocktail-party listening task.

II. EXPERIMENT 1: A HYBRID MONAURAL-DICHOTIC COCKTAIL-PARTY TASK

A. Methods

The experiments described in this paper employed the coordinate response measure (CRM), a call-sign-based intelligibility test that has been shown to produce a substantial amount of informational masking in diotic listening tests with two or more simultaneous talkers (Brungart, 2001a, b). The CRM phrases were taken from the publicly available CRM speech corpus for multitalker communications research (Bolia *et al.*, 2000), which contains phrases of the form ‘‘Ready (call sign) go to (color) (number) now,’’ spoken by four male and four female talkers with all possible combinations of eight call signs (‘‘Arrow,’’ ‘‘Baron,’’ ‘‘Charlie,’’ ‘‘Eagle,’’ ‘‘Hopper,’’ ‘‘Laker,’’ ‘‘Ringo,’’ ‘‘Tiger’’); four colors (‘‘blue,’’ ‘‘green,’’ ‘‘red,’’ ‘‘white’’); and eight numbers (1–8).

In experiment 1, the signal presented to the right (target) ear always consisted of a mixture of two simultaneous phrases from the corpus: a target phrase, which was randomly selected from the phrases containing the call sign ‘‘Baron’’ and a masking phrase, which was randomly selected from all the phrases with a different call sign, color, and number than the target phrase. The level of the target phrase was scaled relative to the masking phrase to produce one of five different randomly selected signal-to-noise ratios (-8, -4, 0, 4, or 8 dB).

The signal presented to the left (unattended) ear consisted of one of three different masking sounds:

(1) Speech-shaped noise that was filtered to match the average long-term spectrum of all of the phrases in the CRM corpus (Brungart, 2001b) and presented at a rms level 20 dB higher than the rms level of the masking phrase in the target ear.[1]
(2) A randomly selected CRM phrase with a different call sign, color, and number than the phrases used in the target ear, presented at the same rms level as the masking phrase in the target ear.
(3) A randomly selected CRM phrase with a different call sign, color, and number than either of the phrases used in the target ear, presented at a rms level 15 dB lower than the rms level of the masking phrase in the target ear.

In addition to these three experimental conditions, two control conditions were tested. The first control condition was a purely monaural listening condition, with two competing talkers in the target ear and no signal in the unattended ear. The second control condition was a purely dichotic listening condition, with only the target talker in the target ear and a single masking talker in the unattended ear.

These five conditions were tested separately for two different target talkers: a male talker (talker 0 from the corpus) and a female talker (talker 5 from the corpus). In each case, the masking talkers were randomly selected from the remaining three talkers in the corpus who were the same sex as the target talker.[2] Thus, the talkers in any given stimulus presentation were always either all males or all females.

A total of eight paid volunteer listeners with normal

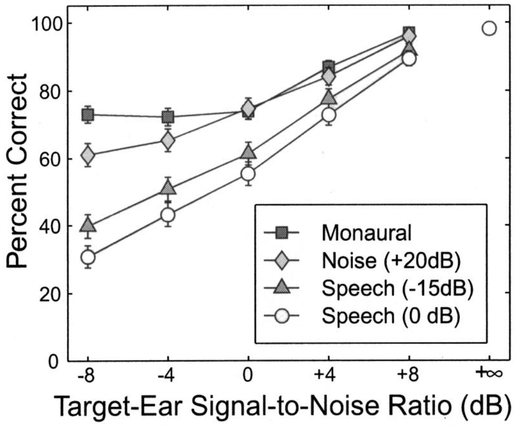

FIG. 2. Color and number identification performance in experiment 1, a dichotic listening experiment with one target talker and one masking talker in the target ear and an interfering speech or noise masker in the unattended ear. Each of the four curves shows performance for a different masking condition in the unattended ear (see legend), and the curves are all plotted as a function of signal-to-noise ratio (SNR) in the target ear. The error bars show the 95% confidence intervals for each data point (±1.96 standard errors).

hearing (three male, five female) participated in the experiment. All had previous experience in experiments using the CRM. These listeners were seated at a control computer in a sound-treated listening booth and they were instructed to listen in their right ear for the target phrase containing the call sign ''Baron'' and respond by selecting the color–number combination contained in the target phrase from a matrix of colored numbers displayed on the CRT of the control computer. They were also instructed to ignore any signals occurring in their left ear.

The listeners first participated in a block of 120 practice trials in which they heard only the male target talker's voice in the right ear and no masker in either ear. This allowed them to become familiar with the voice characteristics of the male target talker. Then they participated in two blocks of 120 trials for each stimulus condition in a randomly assigned order that was counterbalanced across the listeners with a latin square design. Finally, they participated in two blocks of 120 trials in the dichotic control condition. The procedure was then repeated using the female target talker. Thus, each of the eight listeners in the experiment participated in a total of 480 trials in each of the five experimental and control conditions tested in the experiment.

B. Results

The curves in Fig. 2 show the percentage of trials in which the listeners correctly identified both the color and the number in the target phrase as a function of the SNR in the target ear for each of the contralateral-ear masking configurations tested in experiment 1. In the dichotic control condition (shown by the open circle at a target-ear SNR of ∞) the listeners responded correctly in nearly 100% of the trials. This result is consistent with the results of other experiments that have shown listeners have no difficulty segregating competing speech signals that are presented to different ears (Cherry, 1953; Drullman and Bronkhorst, 2000). It also provides a performance baseline for a condition that involves only across-ear interference.

In the monaural condition, where there was no signal in the contralateral ear, performance decreased as the target-ear SNR decreased from 8 to 0 dB, but leveled off at SNR values less than 0 dB (filled squares in Fig. 2). This performance curve closely matches the results of an earlier diotic experiment that used the CRM corpus to measure the effects of SNR on two-talker speech segregation with same-sex talkers (Brungart, 2001b). The only difference is that performance at negative SNRs plateaued at roughly 70% correct responses in this experiment and at roughly 60% correct responses in the previous experiment. This difference probably occurred because the listeners were provided with *a priori* information about the target voice that they did not receive in the earlier experiment. The results of this monaural condition provide a performance baseline for a condition that involves only within-ear interference.

The addition of the +20 dB speech-shaped noise to the unattended ear had relatively little impact on overall performance (diamonds in Fig. 2). When the SNR in the target ear was 0 dB or higher, the noise had no effect on performance. When the SNR in the target ear was less than 0 dB, the noise produced only a slight (less than 10 percentage point) decrease in overall performance (relative to the monaural control condition). Thus it appears that even a relatively high-level contralateral noise masker produces only a small amount of across-ear interference in the two-talker target-ear segregation task.

When an interfering speech signal was added to the contralateral ear, however, performance was much worse than in the monaural control condition (open circles in Fig. 2). This reduction in performance was particularly large at negative target-ear SNR values: whereas performance in the monaural condition plateaued at negative target-ear SNR values, performance in the contralateral speech condition decreased monotonically at negative target-ear SNR values. This resulted in a net decrease in performance as large as 40 percentage points in the contralateral-speech condition when the target-ear SNR was −8 dB. When the target-ear SNR was greater than 0 dB, the contralateral speech masker produced a more modest 10 percentage point decrease in performance relative to the monaural control condition. The level of the contralateral speech signal had relatively little impact on overall performance: attenuating the masking talker in the contralateral ear by 15 dB improved performance by less than 10 percentage points across the range of SNRs tested (triangles in Fig. 2). Thus, it does not appear that the overall level of the signal in the unattended ear has much impact on the amount of across-ear interference it produces.

These results clearly show that within-ear and across-ear speech segregation are not independent processes. Listeners are extremely good at segregating a target speech signal from an interfering talker in the opposite ear. Listeners are also relatively good at segregating a target speech signal from an interfering talker in the same ear. But listeners have a great deal of difficulty segregating a target speech signal from an interfering talker in the same ear when an interfering talker is simultaneously presented to the opposite ear.

There are at least two possible ways to view this interaction between within-ear interference and across-ear interference. One possibility is that the presence of the masking talker in the same ear as the target speech degrades the listener's ability to ignore the interfering signal in the unattended ear. The other possibility is that the presence of the masking speech signal in the unattended ear degrades the listener's ability to segregate the two talkers in the target ear. By looking at the distribution of incorrect responses in the experiment, it is possible to distinguish between these two possibilities. Figure 3 shows how the color and number responses were distributed at each target-ear SNR in the experimental condition with the 0-dB masking talker in the unattended ear. The responses are divided into four categories: (1) responses that matched the color or number in the target phrase; (2) responses that matched the color or number in the masking phrase presented in the target ear; (3) responses that matched the color or number spoken in the masking phrase presented in the unattended ear; and (4) responses that did not match any of the colors or numbers presented in the stimulus. These results show that an overwhelming majority of the incorrect responses contained color and number coordinates that were presented in the target ear. Only a small portion contained the color–number coordinates presented in the unattended ear. This result suggests that listeners' performance was degraded in the contralateral masking condition because the presence of the masking talker in the unattended ear interfered with their ability to segregate the two talkers in the target ear, and not because the presence of the masking talker in the target ear impaired their ability to ignore the signal in the unattended ear.

The distribution of errors in Fig. 3 further suggests that this inability to segregate the talkers in the target ear was due to an increase in within-ear informational masking rather than an increase in within-ear energetic masking. The vast majority of the incorrect responses included the color or number words present in the target-ear masking phrase, and almost none included a color or number that was not spoken by any of the talkers in the stimulus. This result indicates that the errors occurred because the listeners were unable to distinguish between the target and masking talkers in the target ear (informational masking), and not because they were unable to detect the acoustic elements of the two speech signals in the target ear (energetic masking). Thus it appears that the decrease in performance that occurred in the contralateral speech-masking conditions of experiment 1 occurred primarily because of a marked increase in the informational masking component of the within-ear interference in the target ear.

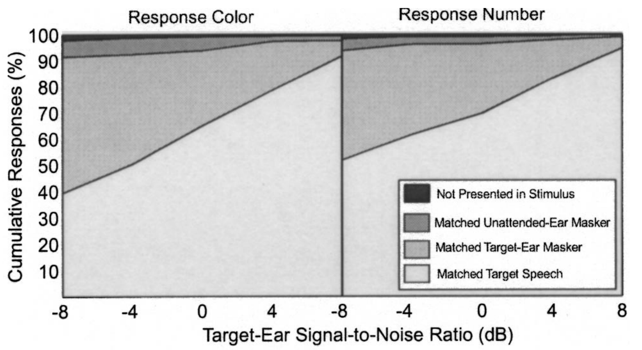

FIG. 3. Cumulative distributions of the responses in experiment 1 in the stimulus condition with two talkers in the target ear and a 0 dB masking talker in the unattended ear. The responses are divided according to their relation with the color and number words used in the target and masking phrases of the stimulus. The results are shown separately for the color responses (left panel) and the number responses (right panel) to simplify the display of data from situations where the color response and the number response did not match the same talker. See text for details.

III. EXPERIMENT 2: THE IMPACT OF ACROSS-EAR INTERFERENCE ON ENERGETIC MASKING IN THE TARGET EAR

The results of the first experiment show that the presence of a speech signal in the unattended ear produces a substantial increase in the amount of informational masking in the two-talker within-ear segregation task in the target ear. However, because previous experiments have shown that two-talker within-ear segregation with the CRM is almost completely dominated by informational masking (Brungart, 2001b), it is difficult to determine from these results whether there is also an interaction between the presence of an interfering talker in the unattended ear and the amount of energetic masking that occurs in the target ear. In order to isolate the effects of speech in the unattended ear on the energetic portion of within-ear interference in the target ear, a second experiment was conducted in which the masking talker in the target ear was replaced by a speech-shaped noise masker.

A. Methods

The procedure used in the second experiment was the same as the procedure used in the first experiment, except that the CRM masking phrase in the right (target) ear was replaced with speech-shaped noise that was filtered to match the average long-term spectrum of all of the phrases in the CRM corpus (Brungart, 2001b). This noise was scaled relative to the rms power of the target speech to produce one of five different target-ear SNR values (−16, −12, −8, −4, or 0 dB). Only two of the five unattended-ear masking conditions in experiment 1 were reproduced in experiment 2: the monaural control condition with no signal in the unattended ear and the 0-dB contralateral speech condition with a CRM phrase in the unattended ear.

Eight paid volunteer listeners participated in the experiment, four of whom were also participants in the first experiment. Each of the listeners first participated in a total of four blocks of 120 trials: one block with the male talker and one block with the female talker in each of the two unattended-ear masking conditions. A preliminary analysis of the data from these four blocks indicated that it would be useful to collect additional data at a higher target-ear SNRs value, so each listener was asked to participate in four additional blocks of 48 trials with the target-ear SNR fixed at +4 dB. Thus, each of the eight listeners participated in a total of 672 trials.

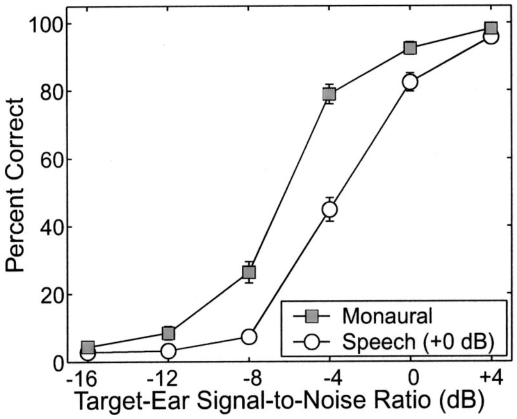

FIG. 4. Color and number identification performance in experiment 2, a dichotic listening experiment with a target talker and an interfering noise masker in the target ear and either an interfering speech signal (open circles) or no signal (filled squares) in the unattended ear. The curves are plotted as a function of SNR in the target ear. The error bars show the 95% confidence intervals for each data point (±1.96 standard errors).

B. Results and discussion

Figure 4 shows the percentage of correct color and number identifications in experiment 2 as a function of the SNR in the target ear. In the condition with no interfering signal in the unattended ear (filled squares), the results were consistent with previous experiments that have used the CRM to measure performance as a function of SNR with a speech-shaped noise masker (Brungart, 2001a, b): identification performance was near 100% when the SNR was greater than 0 dB and dropped off rapidly as the SNR was reduced below 0 dB. When the masking speech was added to the unattended ear, overall performance decreased substantially: the performance curve was shifted to the right by roughly 2–4 dB. This result shows that the addition of a speech masker to the unattended ear produced a decrease in performance that was roughly equivalent to a 2–4-dB decrease in the SNR in that ear. Thus, it is either the case that the interfering talker in the unattended ear increased the amount of energetic masking in the target ear or that the noise in the target ear increased the amount of informational masking from the talker in the unattended ear. An analysis of the incorrect responses in the experiment provides at least partial support for the latter hypothesis. When the SNR in the target ear was −16 dB, 38% of the color responses and 25% of the number responses contained the color or number spoken by the masking talker in the unattended ear. This suggests that an increase in across-ear informational masking may have contributed to the overall decrease in performance in the contralateral-speech conditions of experiment 2. Because an interfering speech signal produces both informational and energetic masking, it is likely that energetic masking in the target ear also had some effect on the amount of informational masking caused by the contralateral speech masker in experiment 1. However, the fact that the vast majority of the incorrect responses in experiment 1 matched the color and number spoken by the masking talker in the target ear (see Fig. 3) suggests that informational masking from the interfering talker in the target ear played a larger role than energetic masking from the interfering talker in the unattended ear in that experiment.

IV. EXPERIMENT 3: ACROSS-EAR INTERFERENCE WITH A NOISE-MASKED SPEECH SIGNAL IN THE UNATTENDED EAR

The results of the first experiment show that a speech masker in the unattended ear produces substantially more across-ear interference in the hybrid monaural-dichotic cocktail-party task than a noise masker in the unattended ear. It therefore follows that the amount of across-ear interference may be reduced when the speech signal in the unattended ear is masked by noise. Experiment 3 was conducted to examine the change in across-ear interference that occurs as noise is added to the speech signal in the unattended ear.

A. Methods

The experimental procedure was generally similar to the procedures used in experiments 1 and 2. The signal presented to the listener's right (target) ear always consisted of two simultaneous phrases from the CRM corpus: a target phrase containing the call sign ''Baron,'' and a randomly selected masking phrase with a different call sign, color, and number than the target phrase. The rms level of the target signal was scaled relative to the masker to produce a fixed target-ear SNR value of −4 dB.[3] The signal presented to the listener's left (unattended) ear was a mixture of a randomly selected phrase from the CRM corpus and speech-shaped noise that was filtered to match the average long-term spectrum of all of the phrases in the CRM corpus (Brungart, 2001b). The contralateral speech signal was always presented at the same level as the masking speech in the target ear, and the level of the contralateral noise was adjusted relative to the speech signal to produce one of seven different SNR values in the unattended ear (−20, −12, −4, 4, 12, 20, and 28 dB). Note that this method of adding noise to a fixed-level speech masker caused the total energy in the unattended-ear stimulus to increase when the SNR in the contralateral ear decreased. When the SNR in the unattended ear was −20 dB, the combined speech and noise masker was approximately 20 dB more intense than the masking talker in the target ear (similar to the +20 dB contralateral noise condition of experiment 1). When the SNR in the unattended ear was +28 dB, the combined speech and noise masker in the unattended ear was presented at approximately the same level as the masking talker in the target ear (similar to the +0 dB contralateral speech condition of experiment 1).

The same eight listeners who participated in experiment 1 also participated in experiment 3. Each listener participated in a total of four blocks of 84 trials. Note that only the female target talker was tested in experiment 3: the target talker was always talker 5, and the masking talkers were randomly selected from the other female talkers in the corpus.

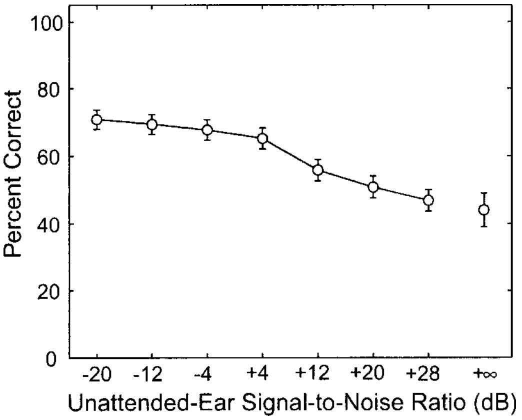

FIG. 5. Percentage of correct color and number responses in experiment 3, a dichotic listening experiment with one target talker and one masking talker in the target ear and a noise-masked speech signal in the unattended ear. The results are shown as a function of the SNR in the unattended ear; the SNR in the target ear was fixed at −4 dB. The point at +∞ shows the results from experiment 1 with a pure speech signal in the unattended ear. The error bars represent 95% confidence intervals (±1.96 standard errors).

B. Results and discussion

Figure 5 shows the percentage of correct color and number responses in experiment 3 as a function of the SNR in the unattended ear. When the SNR value was +28 dB and the signal in the unattended ear was primarily speech, performance was similar to the contralateral speech condition of experiment 1 (shown by the data point at +∞ dB); when the SNR value was −20 dB and the signal in the unattended ear was primarily noise, performance was similar to the contralateral noise masking condition of experiment 1. Between these two extremes, overall performance decreased monotonically with increasing SNR in the unattended ear. The only notable feature of this transition was the particularly steep decrease in performance between the SNR values of +4 and +12 dB.

The contralateral-ear SNR values in experiment 3 can also be used to estimate how intelligible the unattended-ear signal would be if listeners were instructed to attend to it rather than the target speech. The results from experiment 2 (Fig. 5) showed that performance with a speech-shaped noise masker in the target ear was near 100% when the target-ear SNR was +4 dB, and that it dropped to near 0% when the target-ear SNR was −12 dB. Thus all of the change in intelligibility caused by the addition of speech-shaped noise to a phrase from the CRM corpus occurs in the range of target-ear SNRs from −12 to +4 dB. In contrast, the results from experiment 3 (Fig. 5) show that less than a quarter of the 25 percentage point change in performance caused by the addition of noise to the CRM phrase in the unattended ear occurs in the range of SNRs from −12 to +4 dB. Viewed another way, the results show that a speech-shaped noise masker that had no effect on performance in experiment 2 when it was added to the CRM phrase in the target ear at an SNR of +4 dB produced a roughly 20 percentage point increase in performance in experiment 3 when it was added to the CRM phrase in the unattended ear at an SNR of +4 dB. Based on these results, it does not appear that the release from across-ear interference that occurs when the interfering speech signal in the unattended ear is masked by noise can be explained by a reduction in the intelligibility of the interfering speech. One possible alternative explanation is that the addition of noise to the unattended ear adds a distinguishing characteristic to the masking talker in that ear (‘‘noisiness’’) that makes it easier to segregate from the target speech. Another possibility is that the masking noise makes the interfering talker less ‘‘speechlike’’ and more ‘‘noiselike’’ in some other dimension that causes it to produce less across-ear interference than clean speech. Either of these alternatives could help explain why even very low levels of noise in the unattended ear (at SNRs of +12 dB or more) reduced the amount of across-ear interference in the hybrid monaural-dichotic cocktail-party task.

V. EXPERIMENT 4: ACROSS-EAR INTERFERENCE WITH MULTIPLE TALKERS IN THE UNATTENDED EAR

The results of experiment 3 show that the amount of across-ear interference in the hybrid monaural-dichotic cocktail-party task can be reduced by adding noise to the speech signal in the unattended ear. As mentioned previously, one possible explanation for this result is that the noise makes the interfering speech less ‘‘speechlike’’ and more ‘‘noiselike,’’ and thus easier to distinguish from the target speech. Another possible way to make a speech signal less ‘‘speechlike’’ and more ‘‘noiselike’’ is to increase the number of talkers in the signal; each additional talker fills in some of the ‘‘gaps’’ in the overall envelope of the signal. In the limit, a signal with an infinite number of simultaneous talkers with random onsets will be indistinguishable from speech-shaped noise (Bronkhorst and Plomp, 1992). Thus one might expect that a release from masking could be obtained by adding additional talkers to the unattended ear. A fourth experiment was conducted to test this hypothesis.

A. Methods

The experimental procedures used in experiment 4 were similar to those used in the earlier experiments. In most of the conditions, the speech signal in the right (target) ear consisted of a mixture of two CRM phrases: a target phrase containing the call sign ‘‘Baron,’’ and a randomly selected masking phrase that was scaled to produce a target-ear SNR of either +4 or −4 dB. The signal in the unattended ear in these conditions consisted of zero, one, two, three, or four randomly selected masking phrases from the CRM corpus, each presented at the same level as the masking speech in the target ear. These conditions are denoted by TM, TM-M, TM-MM, TM-MMM, and TM-MMMM, respectively. Two additional configurations were also tested. In the TMM condition, the signal in the target ear consisted of the target phrase mixed with two randomly selected CRM masking phrases. The masking phrases were combined and then scaled relative to the target phrase to make the overall target-ear SNR −4 dB. In the T-MM condition, only the target phrase was presented to the target ear and two randomly selected masking phrases were presented to the unattended ear at the same level.

Each of these conditions was tested in two different modes. In the standard onset mode, which was also used in experiments 1–3, all of the target and masking CRM phrases started simultaneously. In the random onset mode, each of the masking phrases started at a randomly selected point (uniformly distributed over the length of the utterance), played to the end of the waveform, and then wrapped around to play from the beginning of the waveform to the randomly selected starting point. This randomization varied the temporal positions of the call signs, colors, and numbers in the masking phrases without changing the overall lengths of the utterances. The conditions with random-onset maskers are denoted in the same way as the conditions with the standard onset maskers, but the M associated with the randomized masker is replaced by an R. Thus, the TM-RR condition consisted of a target phrase and a standard-onset masking phrase in the target ear and two random-onset masking phrases in the masking ear. The eight random-onset conditions are TM-R, TM-RR, TM-RRR, TM-RRRR, T-RR, TR, TRR, and TMR.

As in the first two experiments, experiment 4 was conducted separately with two different masking talkers: one male talker (talker 0) and one female talker (talker 5). Six listeners participated in the experiment, including five who also participated in experiment 1. In the first phase of the experiment, where the target-ear SNR was always fixed at −4 dB, each listener first participated in four blocks of 120 trials (two with a male talker, two with a female talker) in each of the seven standard-onset conditions (conducted in random order). They then participated in four blocks of 120 trials in each of the eight random-onset conditions. In the second phase of the experiment, where the target-ear SNR was always fixed at +4 dB, the six listeners each participated in four blocks of 120 trials (two with a male talker, two with a female talker) in five of the 15 listening configurations tested in the first phase of the experiment (TM, TM-M, TM-R, TM-MMMM, and TM-RRRR). Thus, each listener participated in a total of 8640 trials in the two phases of experiment 4.

B. Results and discussion

The left panel of Fig. 6 shows the percentages of correct color and number identifications in the conditions from the first phase of experiment 4 with two competing talkers in the target ear and a target-ear SNR of −4 dB. When no competing talkers were present in the unattended ear (the TM condition), the listeners correctly identified the color and number in approximately 75% of the trials. When the first talker was added to the unattended ear, the percentage of correct identifications dropped to about 50%. There was not, however, any additional degradation in performance with the addition of the second, third, or fourth masking talkers in the unattended ear. Randomizing the onsets of the talkers significantly improved performance in the conditions with three or four competing talkers in the unattended ear, but performance in these random-onset conditions was still much worse than in the TM condition with no masking talker in the unattended ear. These results suggest that the amount of across-ear interference in the hybrid monaural-dichotic listening task is relatively insensitive to the characteristics of the speech signal in the unattended ear: the amount of interference is roughly the same with a single standard-onset CRM phrase that starts at the same time as the target phrase as it is with a four-talker random-onset signal.

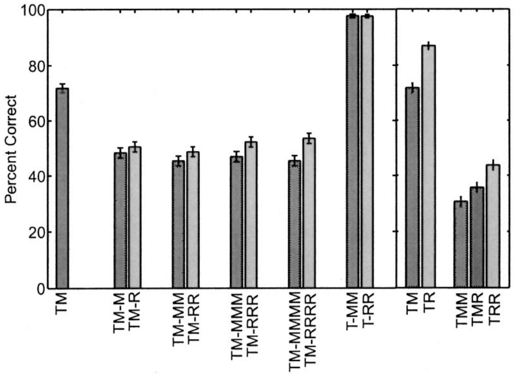

FIG. 6. Color and number identification performance in the first phase of experiment 4, a dichotic listening experiment with zero to four interfering talkers in the unattended ear, one or two interfering talkers in the target ear, and a target-ear SNR value of −4 dB. The conditions are labeled in the form TA-B, where T represents the target phrase, A represents the masking phrases in the same ear as the target phrase, and B represents the masking phrases in the unattended ear. In each case, an M stands for a standard-onset masker, and an R stands for a random-onset masker (see text for details). Note that the TM condition is plotted twice (once in the left panel and once in the right panel) to allow an easier comparison across conditions. The error bars represent 95% confidence intervals (±1.96 standard errors).

The right panel of Fig. 6 shows the effects that randomizing the onsets of the masking speech signals had on the amount of within-ear interference in the monaural cocktail-party task when there were no interfering talkers in the unattended ear and the target-ear SNR was fixed at −4 dB. In the two-talker monaural listening task (TM), randomizing the onset of the masking phrase (TR) produced a 15 percentage point increase in overall performance. In the three-talker monaural listening task, performance was worst when neither interfering talker was randomized (TMM), slightly better when one interfering talker was randomized (TMR), and best when both interfering talkers were randomized (TRR). Thus it appears that onset randomization substantially reduced the amount of within-ear interference in the multi-talker listening task. This was true even in the TRR condition, where a reduction in the number of opportunities to listen to the target talker in the "gaps" of the two-talker random-onset masker should have increased the amount of energetic masking in the stimulus. The most likely explanation for this result was a release in informational masking that occurred because the color and number coordinates in the masking phrases were much less likely to overlap with the color and number coordinates in the target phrase in the random-onset conditions of the experiment.

Figure 7 compares performance in each of the five listening conditions tested in the second phase of experiment 4 where the target-ear SNR was +4 dB (right panel) to performance in the same five conditions in the first phase of experiment 4 where the target-ear SNR was fixed at −4 dB (left panel). As would be expected, overall performance was substantially better at the higher target-ear SNR value. Be-

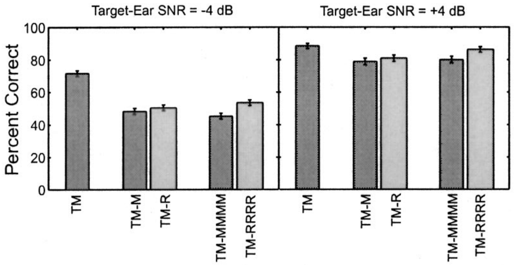

FIG. 7. Comparison of color and number identification performance in the first and second phases of experiment 4. The left panel shows performance in five of the listening configurations from the first phase of experiment 4 where the target-ear SNR was fixed at −4 dB. The right panel shows performance in the same listening configurations from the second phase of experiment 4 where the target-ear SNR was fixed at +4 dB. See text for details. The error bars represent 95% confidence intervals (±1.96 standard errors).

cause the listeners were performing well in all the conditions, the range of performance across the different listening conditions was smaller when the target-ear SNR was +4 dB than when it was −4 dB. There was, however, a significant performance improvement of approximately 7 percentage points between the TM-M and TM-RRRR conditions of the experiment ($p<0.001$, one-tailed *t*-test). Thus it appears that the qualitative differences between a single-talker standard-onset masker in the unattended ear and a four-talker random-onset masker in the unattended ear led to a decrease in across-ear interference in the TM-RRRR configuration even in the relatively easy conditions where the target-ear SNR was fixed at +4 dB.

Note, however, that the results of experiment 4 do not provide much evidence to support the hypothesis that the ''noiselike'' contralateral signal that results from the presence of multiple interfering talkers in the unattended ear will produce less across-ear interference than a ''speechlike'' contralateral signal generated from a single interfering talker in the unattended ear. At both positive and negative target-ear SNRs, performance with four talkers in the unattended ear was roughly the same as performance with a single talker in the unattended ear. Thus, although there are many reasons to believe a four-talker signal is more ''noiselike'' than a single-talker signal, four-talker speech is much more similar to single-talker speech than it is to speech-shaped noise in terms of the amount of across-channel interference it produces in a dichotic cocktail-party listening task. It appears that substantially more than four talkers are necessary to produce a multitalker speech signal that generates the same amount of across-ear interference as speech-shaped noise.

The results of experiments 2–4 also indicate that the amount of across-ear interference in the hybrid monaural-dichotic cocktail-party task cannot be predicted from the intelligibility of the individual CRM phrases in the unattended ear. Low-level noise that had no effect on the intelligibility of the CRM phrase when it was added to the signal in the target ear in experiment 2 was found to substantially improve performance when it was added to the unattended ear in experiment 3. Increasing the number of masking talkers has been shown to substantially reduce the intelligibility of the target phrase in a monaural CRM listening task (Brungart *et al.*, 2001), and randomizing the onsets of the masking speech signals produced large changes in the intelligibility of the target phrase in the monaural conditions of experiment 4, but these manipulations had little or no effect on overall performance when they were applied to the masking speech signals in the unattended ear. Thus, it appears that the amount of across-ear interference that occurs in the hybrid monaural-dichotic task cannot be predicted from the intelligibility of any one of the interfering speech signals in the unattended ear. Indeed, all signals that contain at least one speech signal and are free of noise seem to produce roughly the same amount of across-ear interference in the cocktail-party listening task.

VI. EXPERIMENT 5: ACROSS-EAR INTERFERENCE WITH A TIME-REVERSED TALKER IN THE UNATTENDED EAR

The results of experiment 4 show that the amount of across-ear interference is roughly constant for a wide range of different unattended-ear speech signals. A single-talker speech signal produces the same amount of interference as a four-talker speech signal, and a random-onset speech signal produces only slightly less interference than a standard-onset speech signal. These results suggest that the amount of across-ear interference caused by a speech signal in the unattended ear has more to do with its qualitative similarity to the target speech than with its semantic similarity to the target phrase. However, it is difficult to test this hypothesis directly with the phrases of the CRM corpus, which are all semantically similar. One manipulation that is capable of removing semantic content from a speech signal without eliminating its ''speechlike'' characteristics is time reversal, which produces a meaningless signal with the same temporal and spectral characteristics as normal speech (Duquesnoy, 1983; Hygge *et al.*, 1992). Experiment 5 was conducted to determine what effect time reversal has on the amount of across-ear interference caused by a speech signal in the unattended ear.

A. Methods

Experiment 5 was conducted concurrently with experiment 1 and used essentially the same methodology. That is, a target talker and masking talker were presented to the target ear, and a masking speech signal was presented in the unattended ear at the same level as the masking phrase in the target ear. The only difference was that the masking phrase in the unattended ear was played backward rather than forward. The same eight listeners who participated in experiment 1 also participated in experiment 5, and the four blocks of 120 trials collected from each listener were intermixed with the other experimental conditions collected in experiment 1.

B. Results

Figure 8 compares the results of experiment 5 to two of the experimental conditions from experiment 1: the monaural control condition with no signal in the unattended ear, and the 0-dB contralateral-speech condition with a nonreversed talker in the unattended ear. This comparison shows that the

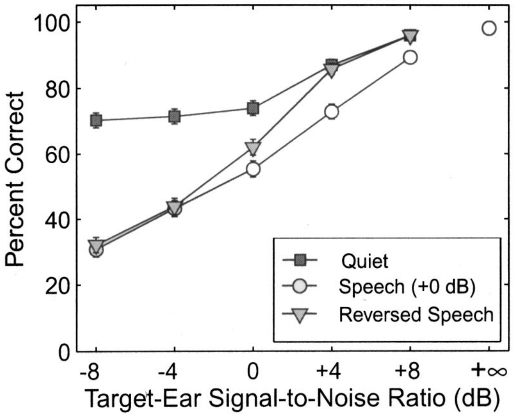

FIG. 8. Color and number identification performance in experiment 5, a dichotic listening experiment with one target talker and one masking talker in the target ear and a time-reversed speech signal in the unattended ear. The triangles show data in the condition with a time-reversed speech signal in the unattended ear. For comparison, results are also shown for the conditions of experiment 1 with no signal in the unattended ear (squares) and with an unattenuated speech signal in the unattended ear (circles). The error bars show the 95% confidence intervals for each data point (±1.96 standard errors).

amount of across-ear interference generated by the time-reversed speech in the unattended ear varied systematically with the SNR in the target ear: when the target-ear SNR was less than 0 dB, the time-reversed speech condition (triangles) was no different than the non-reversed condition from experiment 1 (circles); when the target-ear SNR was greater than 0 dB, the time-reversed condition was no different than the monaural condition of experiment 1 (squares); and when the target-ear SNR was 0 dB, performance with the time-reversed signal fell in between the other two conditions.

The strong relation between the target-ear SNR and the amount of across-ear interference caused by a time-reversed contralateral speech signal suggests 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. At positive target-ear SNRs, where the within-ear segregation process is relatively easy, listeners may be able to take advantage of the semantic differences between forward and reversed speech in the unattended ear. At negative target-ear SNRs, where the within-ear segregation process is relatively difficult, listeners do not appear to be able to take advantage of these semantic differences.

VII. GENERAL DISCUSSION

A. Comparison with results of previous experiments

At this point, we are aware of only one previous study that has systematically examined performance in hybrid monaural-dichotic cocktail party tasks similar to the ones examined in these experiments. In a recent study on cocktail party listening, Drullman and Bronkhorst (2000) examined a total of three different hybrid configurations: a TM-M condition with a target talker in one ear and one interfering talker in each ear, a TMM-M condition with two interfering talkers in the target ear and one in the unattended ear, and a TMM-MM condition with two interfering talkers in each ear. In general, the results Drullman and Bronkhorst reported for these hybrid listening configurations were much different than the results of the experiments reported in this paper. Specifically, Drullman and Bronkhorst found no difference in performance between the hybrid monaural-dichotic listening configurations and the corresponding monaural listening configurations with the same number of masking talkers in the target ear. In other words, they found that the addition of a masking talker in the unattended ear had no effect on the listener's ability to segregate competing talkers in the target ear.

The most likely explanation for this discrepancy with our results is a difference in the vocal characteristics of the masking voices used in the cocktail-party listening tasks. In our experiments, all of the configurations were tested with same-sex target and masking talkers. In the Drullman and Bronkhorst study, the TM configuration was tested with male target and masking talkers, but the TM-M configuration was tested with a male target talker, a male masking talker in the unattended ear, and a *female* masking talker in the target ear. Thus, in evaluating the effect of a masking talker in the unattended ear, our study compared performance with a same-sex masker in the target ear to performance with same-sex talkers in both ears, while the Drullman and Bronkhorst study compared performance with a same-sex masker in the target ear to performance with a same-sex masker in the unattended ear and a *different-sex* masker in the target ear. Previous studies have shown that monaural speech segregation is substantially easier with a different-sex masker than with a same-sex masker. Brungart (2001b), for example, found performance in the monaural CRM task improved from about 60% correct identifications to approximately 85% correct identifications when a same-sex masking talker was replaced with a different-sex masking talker at a 0-dB SNR. Thus, it is likely that Drullman and Bronkhorst did not find any degradation in performance when a masking talker was added to the unattended ear because the additional interference caused by that talker was offset by the improvement in performance that occurred when the same-sex masking talker in the target ear was replaced by a different-sex masking talker. The TMM-M and TMM-MM configurations tested by Drullman and Bronkhorst do not have any direct parallels in our experiments, so it is difficult to say whether we would have encountered the same results with our CRM task. However, the poor overall level of performance in these configurations (≈20% correct responses with words and near 0% correct with sentences) suggests that the amount of across-ear interference caused by adding an interfering talker to the unattended ear may be limited when more than one masking talker is present in the target ear.

B. A shared-resource model of within-ear and across-ear speech segregation

The results of these five experiments show that there are substantial interactions between the within-ear and across-ear segregation processes that occur in the hybrid monaural-dichotic cocktail-party task. Although these interactions are complicated, many of their important features are consistent with a shared-resource model of attention (Wickens, 1984, 1980; Hirst and Kalmar, 1987) where speech segregation ability is constrained by a limited pool of shared attentional resources that listeners must choose to allocate either to within-ear speech segregation or to across-ear speech segregation. Such a model could explain the main results of these experiments simply by assuming that the shared pool contains enough attentional resources to do either across-ear segregation or within-ear segregation, but not enough resources to perform both segregation tasks at the same time. A shared-resource model could also explain why the amount of interference caused by the masking sound in the unattended ear systematically increases with the difficulty of the selective attention task in the target ear: when the segregation task in the target ear is relatively easy (i.e., when the target-ear SNR is greater than 0 dB), relatively few attentional resources are required by the within-ear segregation task and enough resources are left over to effectively segregate the target from the signal in the unattended ear; when the segregation task in the target ear is more difficult (i.e., when the listener has to concentrate on the quieter of two talkers in the target ear), fewer resources are available for across-ear segregation and listeners are more susceptible to interference from a contralateral masking sound. This would explain why the presence of a forward or reversed talker in the unattended ear had a much larger effect on performance when the target-ear SNR was negative than when the target-ear SNR was positive.

A shared-resource model of attention could also explain why some kinds of masking sounds in the unattended ear (like speech and reversed speech) cause more across-ear interference than other kinds of sounds (like noise). In general, the results of these experiments indicate that listeners are much more susceptible to across-ear interference from sounds that are qualitatively similar in some way to speech (such as reversed speech or multitalker speech) than from sounds that are qualitatively different than speech (i.e., noise). A shared-resource model could explain this effect by assuming that more resources are required for across-ear segregation for a speechlike signal and that this additional resource requirement either leaves the listener with fewer resources available for the within-ear segregation task or with insufficient resources to perform the across-ear segregation task.

Unfortunately, substantially more research is needed before this shared-resource model of attention can progress beyond the conceptual stage. One major question that still needs to be addressed is what kind of properties determine the amount of across-ear interference caused by a contralateral masking sound. Time-reversed speech and multitalker speech are qualitatively much different (and much less intelligible) than normal speech, but they appear to generate almost as much across-ear interference as a normal speech signal. However, the results of experiment 3 show that a ‘‘noisy’’ speech signal produces much less across-ear interference than a time-reversed speech signal even at an SNR of +4 dB, where the intelligibility of the noisy speech would be near 100%. Only when the parameters that determine the amount of interference caused by a contralateral masker are better understood will it be possible to begin developing a quantitative model of within-ear and across-ear speech segregation.

VIII. SUMMARY AND CONCLUSIONS

This series of experiments has examined the interactions that occur between within-ear speech segregation and across-ear speech segregation in a hybrid monaural-dichotic cocktail-party task. The results have shown that listeners are not generally able to perform both of these segregation tasks simultaneously: they can segregate a target speech signal that is masked by a single interfering talker in the target ear, or one that is masked by an interfering speech signal in the unattended ear, but not one that is masked by interfering speech signals in both ears at the same time. In general, the amount of across-ear interference generated by a signal in the unattended ear depends on the SNR of the two talkers in the target ear. When the target-ear SNR is positive, only a speech signal in the unattended ear seems to produce any appreciable degradation in performance. When the target-ear SNR is negative, noise signals in the unattended ear produce a slight degradation in performance, and ‘‘speechlike’’ signals produce a dramatic degradation in performance. This degradation occurs even when the speech signal in the unattended ear is distorted by time reversal or onset randomization. Taken together, these results suggest that within-ear segregation and across-ear segregation are closely related processes that may draw from a single shared pool of attentional resources.

In conclusion, it seems appropriate to comment briefly on the relation between the results of these experiments and those of previous experiments in auditory attention. Many experiments have replicated Cherry’s original finding that listeners are able to easily segregate unrelated speech signals that are presented to different ears (Moray, 1959; Triesman, 1964; Egan *et al.*, 1954; Drullman and Bronkhorst, 2000), and some have extended this result to show that listeners can ignore more than one talker in the unattended ear (Drullman and Bronkhorst, 2000; Triesman, 1964). However, we know of no experiments that have shown that listeners are unable to ignore a speech signal in the unattended ear when more than one talker is presented in the target ear. Indeed, the only experiments that have found any effects of contralateral masking on speech perception in the target ear have used unattended ear signals that either contained key words that were highly relevant to the listener [such as the listener’s name (Moray, 1959)] or were contextually related to the target speech signal [such as a speech signal that was switched between the two ears in mid sentence (Triesman, 1960)]. While one might argue that the similar call sign, color, and number structures of the CRM phrases might cause context-related across-ear intrusions in the contralateral speech conditions of experiment 1, such a context-related argument could not explain why nearly all of the incorrect responses in experiment 1 contained the color or number spoken by the interfering talker in the target ear, or why the relatively unintelligible four-talker stimuli from experiment 4 and the completely unintelligible reversed speech stimuli from experiment 5 caused nearly as much across-ear interference at negative target-ear SNRs as a single-talker speech signal.

Thus, although most of the classic models of auditory attention assume that listeners are able to extract some information from the signal in the unattended ear (Triesman, 1964; Deutsch and Deutsch, 1963), none would expect an irrelevant or unintelligible sound in the unattended ear to severely degrade a listener's ability to attend to a two-talker stimulus in the target ear. Future models of auditory attention will have to account for this result.

ACKNOWLEDGMENTS

The authors would like to thank Chris Darwin, Adelbert Bronkhorst, Gerald Kidd, Fred Wightman, and Doris Kistler for their helpful comments on the analysis of the data and on earlier versions of this manuscript. Portions of this work were supported by AFOSR Grant No. 01HE01COR.

---

[1] The +20 dB noise level was used because preliminary experiments showed that +0 dB noise had little effect in the unattended ear and we wanted to explore the upper limits of across-ear interference for a noise source. One concern of having such a strong contralateral noise masker is the possibility of acoustic crosstalk between the noise in the unattended ear and the signal in the target ear. In order to ensure that no crosstalk was occurring, we listened to unmasked CRM speech signals in the target ear that were as much as 50 dB quieter than the noise in the unattended ear and found that we were still able to respond correctly more than 90% of the time. Thus there did not appear to be any evidence that acoustic crosstalk was occurring with the +20 dB noise signal in the unattended ear.

[2] Because the selection was random, $\frac{1}{9}$ of the contralateral speech trials used the same masking talker in both ears. An analysis of the results revealed that these trials generated the same overall pattern of performance as the trials with different masking talkers in the two ears.

[3] The −4 dB value was chosen to ensure a reasonably large difference between performance in the contralateral-speech and contralateral-noise conditions of the experiment.

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

Bregman, A. S. (**1994**). *Auditory Scene Analysis* (MIT, Cambridge).

Brokx, J., and Nooteboom, S. (**1982**). "Intonation and the perceptual separation of simultaneous voices," J. Phonetics **10**, 23–36.

Bronkhorst, A. (**2000**). "The Cocktail Party Phenomenon: A Review of Research on Speech Intelligibility in Multiple-Talker Conditions," Acustica **86**, 117–128.

Bronkhorst, A., and Plomp, R. (**1992**). "Effects of multiple speechlike maskers on binaural speech recognition in normal and impaired listening," J. Acoust. Soc. Am. **92**, 3132–3139.

Brungart, D. (**2001a**). "Evaluation of speech intelligibility with the coordinate response measure," J. Acoust. Soc. Am. **109**, 2276–2279.

Brungart, D. (**2001b**). "Informational and energetic masking effects in the perception of two simultaneous talkers," J. Acoust. Soc. Am. **109**, 1101–1109.

Brungart, D., Simpson, B., Ericson, M., and Scott, K. (**2001**). "Informational and energetic masking effects in the perception of multiple simultaneous talkers," J. Acoust. Soc. Am. **110**, 2527–2538.

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

Conway, R. A., Cowan, N., and Bunting, M. (**2001**). "The cocktail party phenomenon revisited: The importance of working memory capacity," Psychonomic Bulletin and Review **8**, 331–335.

Darwin, C., and Hukin, R. (**2000**). "Effectiveness of spatial cues, prosody, and talker characteristics in selective attention," J. Acoust. Soc. Am. **107**, 970–977.

Deutsch, J., and Deutsch, D. (**1963**). "Attention: some theoretical considerations," Psychol. Rev. **70**, 80–90.

Drullman, R., and Bronkhorst, A. (**2000**). "Multichannel speech intelligibility and talker recognition using monaural, binaural, and three-dimensional auditory presentation," J. Acoust. Soc. Am. **107**, 2224–2235.

Duquesnoy, A. (**1983**). "Effect of a single interfering noise or speech source on the binaural sentence intelligibility of aged persons," J. Acoust. Soc. Am. **74**, 739–943.

Egan, J., Carterette, E., and Thwing, E. (**1954**). "Factors affecting multichannel listening," J. Acoust. Soc. Am. **26**, 774–782.

Ericson, M., and McKinley, R. (**1997**). "The intelligibility of multiple talkers spatially separated in noise," in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. R. Anderson (Erlbaum, Hillsdale, NJ), pp. 701–724.

Festen, J., and Plomp, R. (**1990**). "Effects of fluctuating noise and interfering speech on the speech reception threshold for impaired and normal hearing," J. Acoust. Soc. Am. **88**, 1725–1736.

Freyman, R., Balakrishnan, U., and Helfer, K. (**2001**). "Spatial release from informational masking in speech recognition," J. Acoust. Soc. Am. **109**, 2112–2122.

Freyman, R., Helfer, K., McCall, D., and Clifton, R. (**1999**). "The role of perceived spatial separation in the unmasking of speech," J. Acoust. Soc. Am. **106**, 3578–3587.

Hirst, W., and Kalmar, D. (**1987**). "Characterizing Attentional Resources," J. Exp. Psychol. Gen. **116**, 68–81.

Hygge, S., Ronnberg, J., Larsby, B., and Arlinger, S. (**1992**). "Normal-hearing and hearing-impaired subjects' ability to just follow conversation in competing speech, reversed speech, and noise backgrounds," J. Speech Hear. Res. **35**, 208–215.

Kidd, G. J., Mason, C., and Rohtla, T. (**1995**). "Binaural advantage for sound pattern identification," J. Acoust. Soc. Am. **98**, 1977–1986.

Kidd, G. J., Mason, C., Rohtla, T., and Deliwala, P. (**1998**). "Release from informational masking due to the spatial separation of sources in the identification of nonspeech auditory patterns," J. Acoust. Soc. Am. **104**, 422–431.

Moray, N. (**1959**). "Attention in dichotic listening: affective cues and the influence of instructions," Q. J. Exp. Psychol. **9**, 56–60.

Neff, D. (**1995**). "Signal properties that reduce masking by simultaneous random-frequency maskers," J. Acoust. Soc. Am. **98**, 1909–1921.

Triesman, A. (**1960**). "Contextual cues in selective listening," Q. J. Exp. Psychol. **12**, 242–248.

Triesman, A. (**1964**). "The effect of irrelevant material on the efficiency of dichotic listening," Am. J. Psychol. **77**, 533–546.

Wickens, C. D. (**1980**). "The structure of attentional processes," in *Attention and Performance VIII*, edited R. Nickerson (Erlbaum, Hillsdale, NJ), pp. 239–258.

Wickens, C. D. (**1984**). "The structure of attentional processes," in *Varieties of Attention*, edited by R. Parasuraman and D. Davies (Academic, Orlando).

Wood, N., and Cowan, N. (**1995**). "The Cocktail Party Phenomenon Revisited: Attention and Memory in the Classic Selective Listening Procedure of Cherry," J. Exp. Psychol. **124**, 243–262.