SELECTIVE AND DIVIDED ATTENTION: EXTRACTING INFORMATION FROM SIMULTANEOUS SOUND SOURCES

Barbara Shinn-Cunningham and Antje Ihlefeld

Boston University Hearing Research Center
Department of Cognitive and Neural Systems
677 Beacon St., Boston, MA 02215, USA
`<shinn, ihlefeld>@bu.edu`

ABSTRACT

The way in which sounds interact and interfere with each other (both acoustically and perceptually) has an important influence on how well an auditory display can convey information. While spatial separation of simultaneous sound sources has been shown to be very effective when a listener must report the content of one source and ignore another source (a condition known as *selective attention*), little is known about how spatial separation influences performance in *divided-attention* tasks, i.e., tasks in which the listener must report the content of more than one simultaneous source. This paper reports preliminary results from a pilot study investigating how perceived spatial separation of sources and consistency in source locations influences performance on selective- and divided-attention tasks. Results demonstrate that 1) in both selective- and divided-attention tasks, overall performance is generally better when sources are perceived at different locations than when they are perceived at the same location; 2) in both selective- and divided-attention tasks, randomly changing the perceived source locations from trial to trial tends to degrade performance compared to conditions where the source locations are fixed; and 3) both of the above effects are larger for selective-attention tasks than divided-attention tasks.

1. INTRODUCTION

In everyday environments, the signals reaching the ears of a listener are a mixture of acoustic energy from multiple, simultaneous sources. However, listeners are able to estimate the number of sound sources present in the environment, the spectro-temporal content of the sources, and their meaning, all with relative ease. Various cues, including differences in the pitch, timbre, and spatial location of the competing sources, are thought to aid listeners in separating and understanding simultaneous sound sources [1-13].

Many perceptual studies have examined how well listeners can extract the content of one sound source (the target) in the presence of competing sound sources (maskers), a situation requiring *selective attention* (e.g., see [4, 5, 14-19]). However, relatively few studies have considered how well listeners are able to understand the content of multiple, simultaneous sound sources (a situation requiring *divided attention*; e.g., see [20]). The current study directly compares performance in selective- and divided-attention tasks.

Past studies show that in selective-attention tasks, spatial separation of competing sources can improve target intelligibility, an effect known as "spatial unmasking." In conditions where the target and masker overlap in time and frequency (i.e., conditions where there is significant

energetic masking) and the masker is steady-state noise, a significant portion of this improvement with spatial separation can be attributed to differences in the interaural phase differences (IPDs) present in the target and masker signals reaching the ears. More specifically, spatial unmasking can be predicted simply by assuming that low-level binaural processing makes portions of a masked target more audible (e.g., see the modeling work in [3, 21]).

In conditions with significant *informational masking* (i.e., where both target and masker are audible and easily confused with one another), there appears to be a significant improvement in a listener's ability to attend and understand a target when the competing sources are *perceived* at different locations [16, 22-28]. In the current study, two competing sinusoidal-speech signals were generated to have minimal spectral overlap (producing essentially no energetic masking, similar to stimuli used in [27]). The sources were presented without fine-time IPD cues, but with interaural level differences (ILDs) and envelope interaural time differences appropriate for sources in different locations. The perceived spatial locations depended on the ILD and envelope ITD cues, even though were no fine-time IPD cues in the stimuli. In other words, the IPD cues thought responsible for spatial unmasking in conditions dominated by energetic masking were not present in the stimuli.

While spatial unmasking effects can be very large in studies of informational masking, the listeners in these studies invariably have *a priori* knowledge of where the competing sound sources are going to be located. Given that listeners appear to combat informational masking by focusing attention on the location of the source of interest [29], *a priori* knowledge of the source locations may be critical to elicit improvements in speech intelligibility with differences in perceived spatial separation. If the listener is not sure where to focus spatial attention, perceived spatial separation of competing sources may not aid performance. In the current study, the perceived locations of the target and masker were fixed in half of the experimental sessions and randomly alternated from trial to trial in the other half of the sessions to directly measure whether uncertainty about the perceived spatial locations of competing sources influences performance in selective- and divided-attention tasks.

2. METHODS

The current study quantifies the benefits of spatially separating a pair of simultaneous sound sources for 1) selective- and divided-attention tasks and 2) conditions in which source locations are fixed and in which source locations are chosen randomly from trial to trial.

We expected perceived spatial separation of the sources to have little influence on or even degrade performance in the divided attention task, where subjects should listen to

both sources at all times. We hypothesized that spatial separation would improve performance in a selective-attention task only when listeners had the correct expectation about where the target source would be located.

2.1. Subjects

Four normal-hearing college students, ages 23-26 (including two members of the Hearing Research Center familiar with the goals of the experiment, labeled S1 and S2) were recruited for the study. The two subjects who were not experienced listeners were paid for their participation.

2.2. Stimuli

Raw speech stimuli were taken from the Coordinated Response Measure corpus [30], which consists of sentences of the form "Ready <*call sign*>, go to <*color*> <*number*> now." In the corpus, the call sign is one of the set ["Baron," "Eagle," "Tiger," and "Arrow"]; the color is one of the set [white, red, blue, green]; and the number is one of the digits between one and eight. In the current study, only sentences spoken by the male talker 0 were used, and sentences with the number seven were excluded (as it is the only two-syllable digit and is therefore relatively easy to identify).

In each trial, two different sentences were used as sources. One utterance always contained the call sign "Baron" and was designated the *target*. (Note that in selective attention tasks, the listeners were instructed to report the color and number of the sentence containing the call sign "Baron.") The second utterance (designated the *masker*) was chosen to contain one of the other three call signs, chosen randomly from trial to trial. (Note that the designation "target" and "masker" is essentially arbitrary in divided-attention tasks, where subjects had to report the content of both sources.) In all cases, the numbers and colors in the competing utterances were randomly chosen, although constrained to differ from each other within each trial.

In order to generate sources that were perceived at different spatial locations, the target and masker sentences were convolved with anechoic head-related impulse responses measured on a manikin head (KEMAR) for sources from either 0˚ (straight ahead) or 90˚ to the right of the listener (both at 0˚ elevation and 1 m distance; see [4, 31]). The amplitude envelopes of these left- and right-ear signals were used in subsequent processing to produce modulated-sinusoid speech in which the interaural level differences (ILDs) and the interaural time differences (ITDs) in the *envelopes* of the left- and right-ear signals were preserved within each frequency band, but the fine-time IPD was zero.

Each HRTF-processed speech signal was bandpass filtered into 15 non-overlapping frequency bands of 1/3 octave width, with center frequencies spaced evenly on a logarithmic scale between 215 and 4895 Hz (similar to the processing used in [27]). On each individual trial, eight of the 15 target bands were chosen randomly to construct the target. The envelope for each band was extracted using the Hilbert transform. The resulting envelopes (one for the left ear and one for the right ear) were used to amplitude modulate a sine-wave carrier whose frequency matched the center frequency of the passband. The resulting left- and right-ear modulated sinusoids had zero IPDs, but appropriate ILDs and envelope ITDs for a source at the corresponding location. The eight tones modulated by the left-ear envelopes were summed to form the left-ear target signal and the eight modulated right-ear tones were summed to form the right-ear target signal. To form the masker, six of the seven bands that did not overlap with the target bands

were randomly selected; their left- and right-ear amplitude envelopes were extracted and multiplied by the appropriate carrier tones; and the resulting modulated sinusoids summed to form the left- and right-ear masker signals. The resulting stimuli, although qualitatively unlike natural speech, are all 100% intelligible in quiet for a practiced listener (see [32] for a more complete description of this kind of degraded speech stimulus).

The resulting binaural target and masker signals were normalized to have the same RMS energy, then the levels were adjusted to set the desired target-to-masker energy ratio (TMR). The TMR was chosen randomly on each trial. Performance was measured at TMRs of –40, -30, -20, -10, and 0 dB for all subjects. For S3 and S4, a TMR of +10 dB was also used. The left and right target and masker signals were then added to generate the binaural stimulus for each trial.

2.3. Procedure

Each subject performed four experimental sessions, each lasting roughly one hour. In each session, one of four tasks was performed (*deterministic selective, randomized selective, deterministic divided,* and *randomized divided*, performed in this same order for all subjects). Note that this preliminary experiment proceeded from "easy" to "hard" conditions, in order, so that any learning effects should make *divided* better than *selective* attention results and *random* better than *deterministic* results. In *selective-attention* tasks, subjects were instructed to report the color and number of the target sentence (the sentence containing the call sign "Baron"). In the *divided-attention* tasks, subjects were instructed to report the colors and numbers from the target and masker sentences. In *deterministic* tasks, the target location was fixed throughout the session at 0˚. In the *randomized* tasks, the locations of the target and masker sentences were randomly chosen on each trial.

In each of the four sessions, subjects performed multiple blocks consisting of 50 trials. In half of the blocks, both target and masker were simulated at 0˚ (co-located sources). In the other half of the blocks, one source was simulated at 0º and one at 90˚ (spatially-separated sources). Co-located and spatially separated trial blocks were alternated throughout the course of each session. Within each session, subjects completed approximately 50 trials at each TMR and condition for a total of 500-600 trials per session.

In the co-located blocks of the two randomized task sessions, target and masker locations were fixed at 0˚ (not truly random); however, in the corresponding spatially separated blocks, the target and masker locations were selected at random (i.e., the target and masker could be at 0˚ and 90˚ or at 90˚ and 0º, respectively). Note that the subjects had to report the content of at least one source from 0˚ on almost all trials: only on half of the trials in the *randomized selective,* spatially separated conditions was the content of the source from 0˚ irrelevant (i.e., only on 6% of the trials). Similarly, listeners reported the content of a 90˚ source on only 31% of the trials over the course of the experiment. The goal was to cause subjects to build up an expectation that they should attend to sources at 0˚ (see Section 4).

In most conditions, the masker was kept at a constant, audible level and the target level adjusted to achieve the desired TMR. However, due to a technical error, the desired TMR was achieved in the *randomized divided* task by adjusting either the target or the masker level, chosen randomly on each trial (however, note that in this task subjects were asked to report the content of both sources).

At the end of each trial, subjects indicated the color(s) and number(s) of the sentences through a graphical user

interface. The computer controlling the experiment recorded their responses. No feedback was provided to the subjects.

2.4. Analysis

For all tasks, the percentage of correct trials was calculated as a function of the TMR. Trials were considered "correct" in the selective-attention tasks only if both the color and number of the target were correctly identified. In the divided-attention tasks, the colors and numbers of both target and masker sentences had to be correct for a trial to be "correct." Thus, given that there are four possible colors and seven possible numbers, chance performance in the selective-attention tasks is 1/4 * 1/7 or roughly 3.5%. In the divided-attention tasks, the chance of randomly guessing both colors is 1/6 and the chance of randomly guessing both numbers is 1/21, so chance performance is less than 1%.

3. RESULTS

3.1. Selective Attention

Figure 1 plots results for the selective-attention tasks as a function of TMR. The top left panel shows the across-subject mean and standard error. Results for individual subjects are shown in the remaining four panels. The same basic layout is used in all subsequent figures.

In all selective-attention tasks, color and number identification increases with TMR. Although not shown, when subjects misidentify the target color or number, they generally select the corresponding value from the masker utterance. This effect is often observed in selective-attention

studies using the CRM corpus (e.g., see [33]) and is thought to reflect the similarity of target and masker sentences (spectro-temporal, prosodic, linguistic, etc.). Performance is best for the spatially separated target and masker when the target spatial location is fixed (solid lines with triangles in Fig. 1). Performance is slightly worse when the sources are spatially separated, but the target location varies from trial to trial (dashed lines with triangles). When the target and masker are co-located at 0˚, listeners have more difficulty identifying the target color and number (solid and dashed lines with circles).

Figure 2 looks in detail at spatially separated, randomized results from the selective attention task. In Fig. 2, results for the spatially separated, deterministic condition are repeated from Fig. 1 (solid black line with triangles). The results for the spatially separated, randomized condition (dashed black line with triangles in Fig. 1) is broken down based on the random location from which the target was presented. Performance on trials in which the target came from straight ahead are shown in Fig. 2 as dotted black lines; performance on trials in which the target came from 90˚ to the right of the listener are shown by dotted gray lines.

Results in Fig. 2 show that the reason that overall performance is worse in the spatially separated random condition than in the deterministic condition is because listeners perform poorly when the target comes from the right (gray dashed lines) compared to when the target is straight ahead (black dashed lines). In fact, performance is comparable for cases where the target is always in front of the listener throughout the block of trials (solid black line) and when the target is randomly selected to come from straight ahead (dashed black lines).
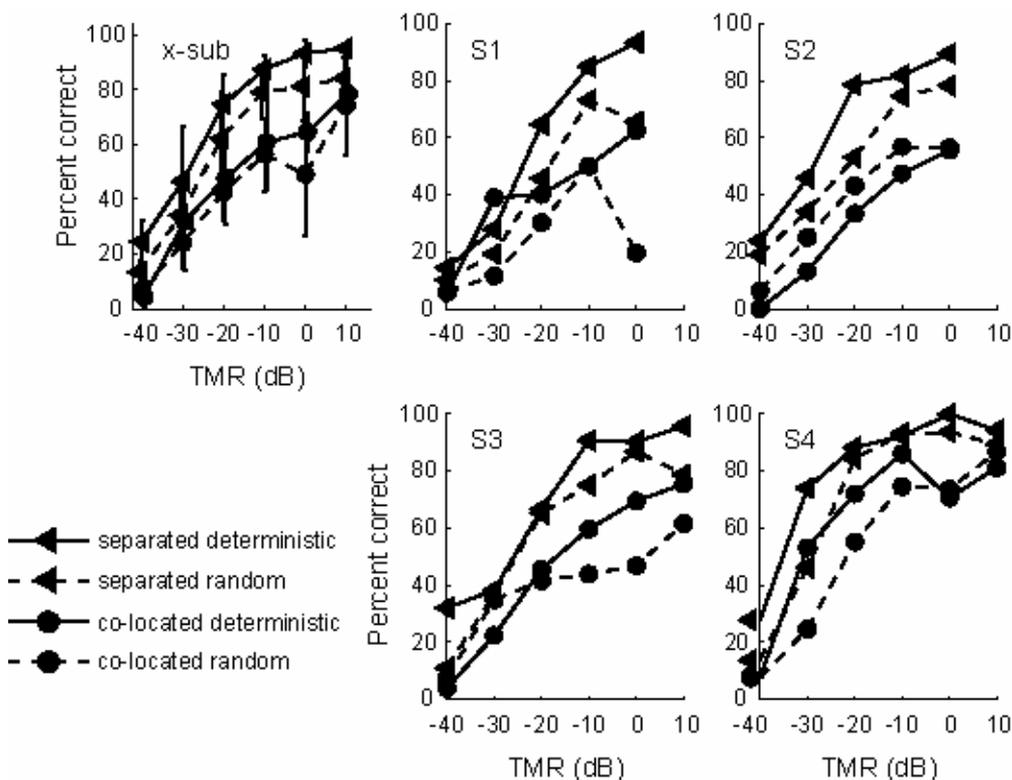
Figure 1. *Performance in selective-attention tasks as a function of TMR. The top-left panel plots the across-subject mean and standard error. The remaining panels show results for individual subjects. In each panel, results are shown for deterministic and randomized source locations (solid and dashed lines, respectively) and spatially separated and co-located sources (triangles and circles, respectively).*
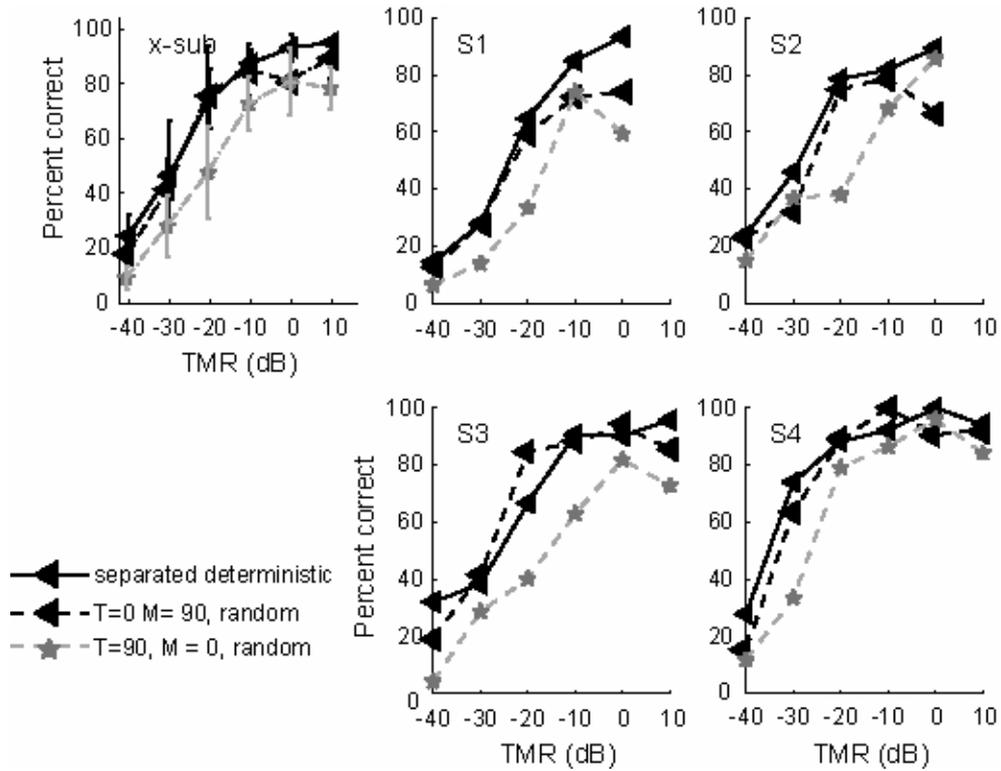
Figure 2. *Performance in randomized selective-attention tasks in which sources are spatially separated broken down by target and masker location as a function of TMR, laid out as in Fig. 1. In each panel, overall results for the spatially separated deterministic task (solid lines) are repeated from Fig. 1. Dotted lines show results in the spatially separated randomized task broken down according to whether the target was straight ahead (black) or at 90˚ (gray).*
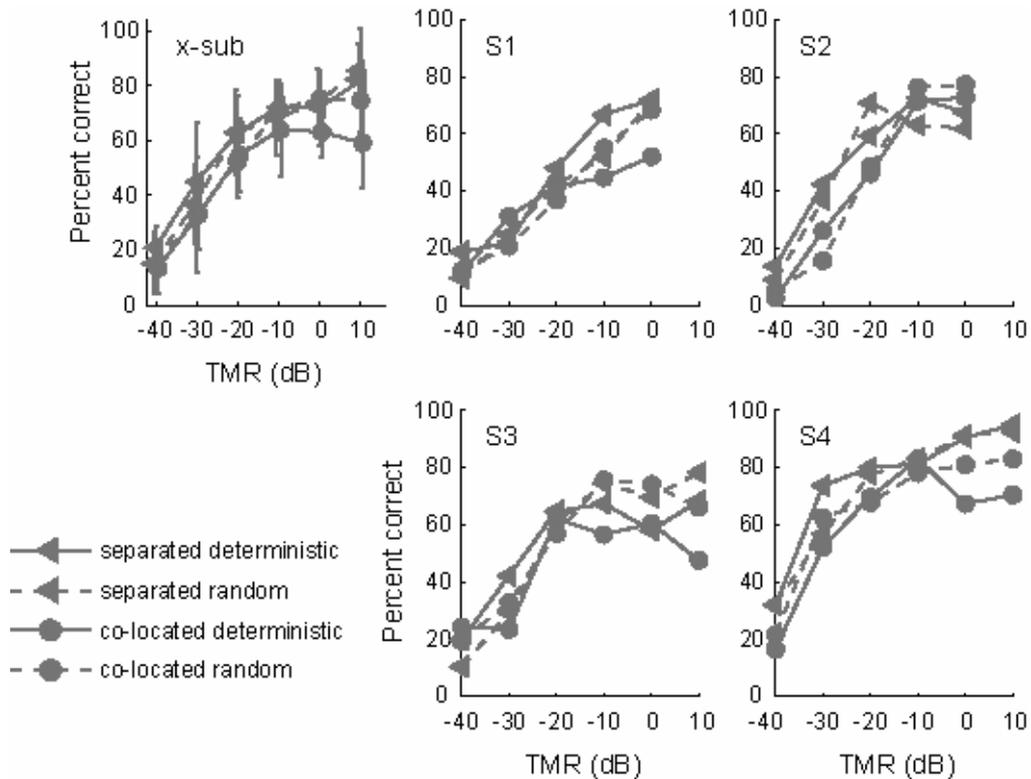
Figure 3. *Performance in divided-attention tasks as a function of TMR, laid out as in Figure 1. In each panel, results are shown for deterministic and randomized source locations (solid and dashed lines, respectively) and spatially separated and co-located sources (triangles and circles, respectively).*

3.2. Divided Attention

Figure 3 shows results from the divided-attention tasks. As in selective-attention tasks, performance improves with TMR. In fact, although it is not shown in Fig. 3, performance is near 100% for all divided-attention trials when scored only on the ability to identify the color and number of the louder sentence. Results across the four conditions are more similar for the divided-attention tasks than for the selective-attention tasks (the influences of spatial separation and spatial randomization are smaller in Fig. 3 than in Fig. 1).

3.3. Effect of Spatial Separation

In order to quantify the effect of spatial separation in the various tasks, the difference in performance was calculated when target and masker were spatially separated and when they were co-located. Fig. 4 plots these differences for deterministic and spatially randomized conditions (diamond and square symbols, respectively) as well as selective- and divided-attention tasks (black and gray lines, respectively). Overall, performance is better in the spatially separated than the co-located conditions; values Fig. 4 are generally larger than zero. Furthermore, the benefit of spatial separation is larger in selective-attention than in divided-attention tasks.

3.4. Effect of Randomizing Location

Figure 5 plots the effect of randomizing the spatial locations of target and masker for the spatially separated trials; i.e., the difference in percent correct performance for deterministic conditions (in which the target is always presented from straight ahead and the masker is presented at 90˚) and randomized conditions (in which the target location is randomly chosen from trial to trial).

In selective-attention tasks (black lines in Fig 5), all subjects perform better when the target and masker locations are held fixed (target at 0º and masker at 90º), independent of the TMR. The difference between performance in the deterministic and randomized conditions is always positive for these conditions.

In the divided-attention tasks (gray lines in Fig. 5), the effect of randomizing the source locations varies from subject to subject. Although there appears to be a systematic effect of TMR on the effect of randomizing source location for S3, the effect of spatial randomization is less consistent for the remaining subjects. Averaged over subjects, the cost of spatial randomization in the divided-attention task is greatest for negative TMRs. When the "target" source is relatively quiet (low TMRs), the difference between the deterministic and randomized conditions is more consistently positive. In other words, performance is better when a relatively quiet source comes from an expected spatial location than when the quiet source location varies from trial to trial. Taken together with the finding that listeners are generally good at reporting the content of the louder source, these results suggest that attending to the location of the less salient source leads to improvements in the ability to understand this source in a divided-attention task.
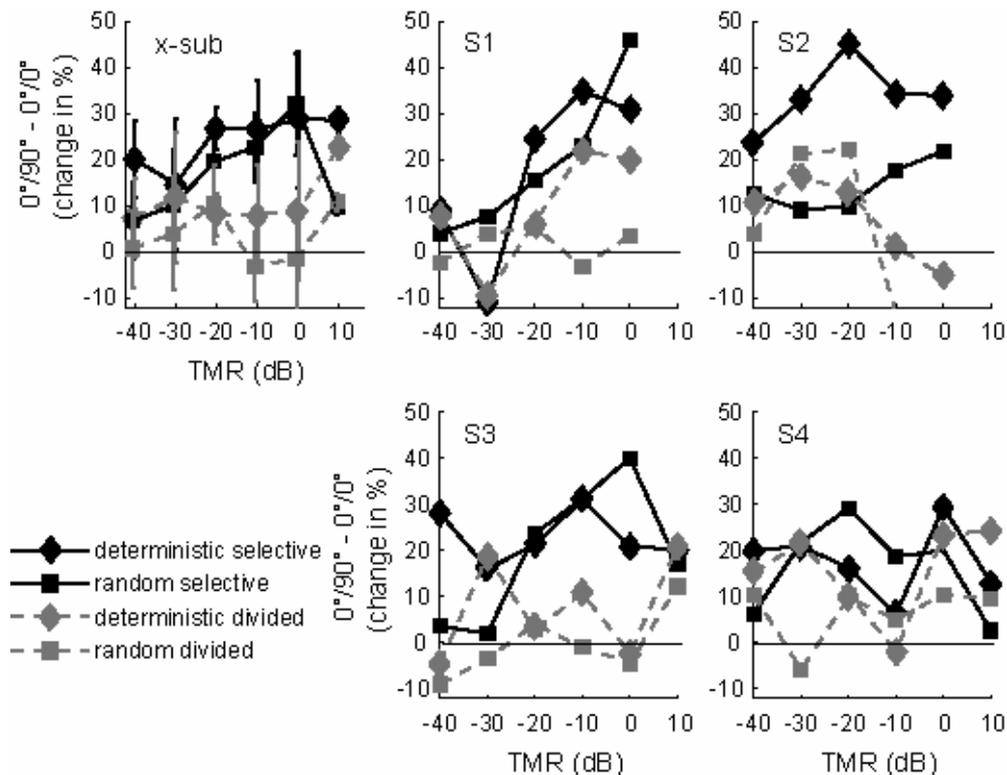
Figure 4. *Spatial gain (difference in percent correct performance for spatially separated and co-located sources), laid out as in Fig. 1. Within each panel results are shown for selective- and divided-attention tasks (black and gray lines, respectively) and for deterministic and randomized spatial presentation (solid and dashed lines, respectively).*
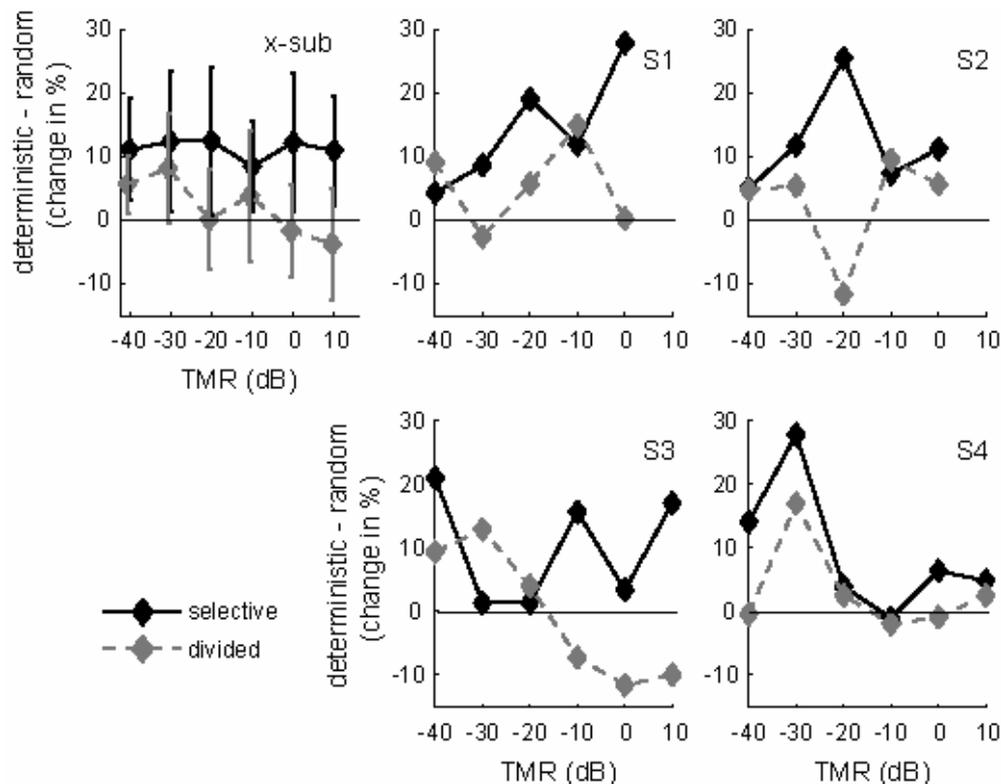
Figure 5. *Effects of spatial uncertainty (difference in percent correct performance for deterministic and randomized spatial presentation) for selective- and divided-attention tasks (solid black and dashed gray lines, respectively), laid out as in Fig. 1. Results are only shown for the cases in which the sources were spatially separated; in the co-located conditions, there was no consistent effect of spatial uncertainty.*

4. DISCUSSION

These experiments confirm that in informational masking tasks, spatial unmasking does not require fine-time IPD cues. This observation is consistent with the fact that in the presence of ordinary room reverberation (which interferes with IPD cues; see [31, 34]), spatial unmasking is relatively robust in the presence of informational maskers but is degraded in the presence of energetic maskers (see [3, 5, 35-37]). Perceived spatial separation of two sources improves a listener's ability to attend to a source of interest and increases the reliability with which listeners can identify the content of multiple competing sources. Despite the fact that in the current study, the competing sources did not overlap in their frequency content, perceived spatial separation of the competing sources lead to performance improvements in both selective- and divided-attention tasks. Furthermore, presenting sources from fixed locations yields better performance than randomly varying the source location in both selective- and divided-attention tasks.

4.1. Selective Attention

The main challenge in the selective attention tasks is to ignore the louder masker and report the content of the relatively quiet target. In the current experiment, listeners were generally better at extracting the content of the quieter target when the target was perceived as coming from in front of the listener. In fact, even though overall performance in the *random selective* task is worse than in the *deterministic selective* task, this overall degradation is due to poor performance on those random trials in which the target is to the side of the listener. For the trials in which the target happens to be in front of the listener, performance in the

randomized selective task is equivalent to performance in the *deterministic selective* task.

This result suggests that in the *random selective* condition, listeners adopt a strategy in which they always focus their attention towards 0°, even when this strategy yields poor performance when the target happens to come from the right of the listener. Given that such trials occur only 25% of the time during the *random selective* block of trials (and only 6% of the time, overall), such a strategy makes intuitive sense. Such an explanation suggests that in the *random selective* task, spatial uncertainty degrades performance not because listeners stop using focused spatial attention, but because they focus spatial attention towards the incorrect location on half the trials. It is also possible that listeners are simply better at hearing a target source perceived in front than to the side, or that listeners become particularly adept at listening in front, due to the frequency with which they were asked to report the content of a 0° source during this study. However, in the divided-attention tasks, listeners always understand the louder source, even when it is off to the side, suggesting that sources to the side are easy to understand when they are salient. We suspect that listeners build up an expectation that they should attend to the 0° direction throughout the experiment. Furthermore, it appears that this *a priori* expectation about what spatial location to attend affects the degree to which listeners benefit from spatial separation of target and masker sources. Further experiments are necessary to tease apart whether there are inherent spatial asymmetries in this task (e.g., perhaps listening to a target from straight ahead yields better performance than listening to a target to the side), investigate the role that practice plays in these results, and more fully explore how *a priori* expectations about spatial location influence spatial unmasking.

4.2. Divided Attention

In divided-attention tasks, listeners are extremely good at reporting the content of the louder source, independent of all other manipulations. As a result, any differences in overall performance in the divided attention conditions arise due to changes in the probability of correctly identifying the quieter source. While spatial separation improves a listener's ability to identify the quieter source's content, so does a consistent spatial configuration of the two competing sources. This result is inconsistent with the idea that in divided attention tasks listeners do not need knowledge about source direction because they simply have to report the content of all sources they hear. In particular, results suggest that listeners do not simply "listen everywhere" in the divided attention task, but deploy their attention to the location from which they expect the quiet source to appear.

The strategy that listeners report employing in the divided attention task helps illuminate what is happening in these conditions. Subjects say that the more intense source is often so dominant in the acoustic signal that no special attention is required to "get it right." Listeners perform best in the divided-attention task when they can focus attention on the source that is relatively quiet and difficult to hear, register its content, and then report it; afterwards, they can easily recall the louder source and report its content. In order to properly focus attention on the quiet source, listener must learn what location to attend (something that they learn subconsciously when the source locations are fixed through a block). If listeners have the correct spatial expectation, focused spatial attention can help the listener hear out the quieter source. Even though the quieter source may be audible, it is less salient than the louder source. In effect, without focused spatial attention to mediate the competition between the quieter and louder sources, the louder source overwhelms the representation of the quiet source, making it difficult to hear the quiet source's content.

4.3. Informational Masking and Spatial Attention

In the current study, target and masker are constructed such that there are hardly any source attributes *other* than source location (such as timbre, prosody, fundamental frequency) that could reliably be used to determine which of the audible time-frequency events are from the target and which are from the masker. In other, more natural listening conditions, listeners may adopt other strategies, focusing attention on non-spatial attributes in order to hear out a source of interest. Thus, it is likely that the way in which spatial separation and spatial uncertainty affect performance on selective- and divided-attention tasks depends in part on the nature and similarity of the target and masker stimuli themselves (e.g., see [25, 28, 37]). When target and masker are similar in nearly all other possible dimensions, spatial attention appears to be very important for mediating competition between sources and reducing informational masking. However, in everyday situations, the importance of spatial attention may be less pronounced.

When designing auditory displays, all of the factors that can help to segregate one acoustic source from another and mediate perceptual competition between simultaneous sounds should be considered. Current results suggest that in cases where two independent sources of information are to be presented simultaneously to a listener, spatial acoustic cues can help a listener mediate competition between the information in the sources.

5. CONCLUSIONS

Building up a correct expectation for where competing sources will be located improves performance in both selective- and divided-attention tasks. By focusing attention on a source from a particular location, listeners are better able to hear out the content of the source, especially when this source is relatively quiet compared to another competing source in the environment. Attention (to spatial location or other source features) appears to be used to modulate the salience of competing sound sources in a manner very similar to the modulatory role that attention plays in other modalities, such as vision (e.g., see [38]).

In the presence of informational masking, spatial unmasking can occur when sources are simply *perceived* from different locations. The fact that in the current experiment, spatial separation provides a benefit in both selective- and divided-attention tasks shows that spatial unmasking can be important even in the absence of fine-time IPD cues that contribute to spatial unmasking through low-level binaural mechanisms [3, 21, 36, 37].

In the current study, listeners appear to use spatial attention to hear out the source of interest (the target in selective-attention tasks and the less salient source in divided-attention tasks). However, the stimuli used in the current experiment were designed to produce a large amount of informational masking; there are very few other source attributes that listeners could use to focus attention on the target and suppress competition from the masker or the more prominent, more intense competing source. Further experiments are necessary to determine whether spatial attention influences performance differently when competing sources differ from one another in more natural ways, and to fully explore how spatial expectations and spatial separation of competing sound sources influence performance in selective- and divided-attention tasks.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] W. A. Yost, "The cocktail party problem: Forty years later," in *Binaural and Spatial Hearing in Real and Virtual Environments*, R. Gilkey and T. Anderson, Eds. New York: Erlbaum, 1997, pp. 329-348.

[2] A. W. Bronkhorst, "The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions," *Acustica*, vol. 86, pp. 117-128, 2000.

[3] P. M. Zurek, R. L. Freyman, and U. Balakrishnan, "Auditory target detection in reverberation," *J Acoust Soc Am*, in press.

[4] S. Devore and B. G. Shinn-Cunningham, "Perceptual consequences of including reverberation in spatial auditory displays," Proc of International Conference on Auditory Displays, 2003.

[5] B. G. Shinn-Cunningham, "Speech intelligibility, spatial unmasking, and realism in reverberant spatial auditory displays," Proc of International Conference on Auditory Displays, Atlanta, GA, 2002.

[6] J. Bird and C. J. Darwin, "Effects of a difference in fundamental frequency in separating two sentences," Proc of 11th International Symposium on Hearing: Auditory Physiology and Perception, Grantham, UK, 1997.

[7] C. J. Darwin, "Auditory grouping and attention to speech," *Proceedings of the Institute of Acoustics*, vol. 23, pp. 165-172, 2001.

[8] A. de Cheveigne, "Vowel-specific effects in concurrent vowel identification," *J Acoust Soc Am*, vol. 106, pp. 327-40, 1999.

[9] A. de Cheveigne, H. Kawahara, M. Tsuzaki, and K. Aikawa, "Concurrent vowel identification. I. Effects of relative amplitude and F0 difference," *J. Acoust. Soc. Am.*, vol. 101, pp. 2839-2847, 1997.

[10] W. R. Drennan, S. Gatehouse, and C. Lever, "Perceptual segregation of competing speech sounds: the role of spatial location," *J Acoust Soc Am*, vol. 114, pp. 2178-89, 2003.

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

[12] C. J. Darwin, D. S. Brungart, and B. D. Simpson, "Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers," *J Acoust Soc Am*, submitted.

[13] M. Ebata, "Spatial unmasking and attention related to the cocktail party problem," *Acoustical Science and Technology*, vol. 24, pp. 208-219, 2003.

[14] M. L. Hawley, R. Y. Litovsky, and J. Culling, "The "cocktail party problem" with four types of maskers: Speech, time-reversed speech, speech-shaped noise, or modulated speech-shaped noise," Proc of MidWinter Meeting of the Association for Research in Otolaryngology, St. Petersberg Beach, FL, 2000.

[15] D. S. Brungart and B. D. Simpson, "Within-ear and across-ear interference in a cocktail-party listening task," *J. Acoust. Soc. Am.*, vol. 112, pp. 2985-2995, 2002.

[16] D. S. Brungart and B. D. Simpson, "The effects of spatial separation in distance on the informational and energetic masking of a nearby speech signal," *J. Acoust. Soc. Am.*, vol. 112, pp. 664-676, 2002.

[17] J. Peissig and B. Kollmeier, "Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners," *J. Acoust. Soc. Am.*, vol. 101, pp. 1660-1670, 1997.

[18] R. Drullman and A. W. Bronkhorst, "Multichannel speech intelligibility and talker recognition using monaural, binaural, and three-dimensional auditory presentation," *J. Acoust. Soc. Am.*, vol. 107, pp. 2224-2235, 2000.

[19] B. G. Shinn-Cunningham, J. Schickler, N. Kopco, and R. Litovsky, "Spatial unmasking of nearby speech sources in a simulated anechoic environment," *J Acoust Soc Am*, vol. 110, pp. 1118-29, 2001.

[20] W. A. Yost, J. Dye, Raymond H., and S. Sheft, "A simulated "cocktail party" with up to three sound sources," *Percept. Psychophys.*, vol. 58, pp. 1026-1036, 1996.

[21] P. M. Zurek, "Binaural advantages and directional effects in speech intelligibility," in *Acoustical Factors Affecting Hearing Aid Performance*, G. Studebaker and I. Hochberg, Eds. Boston, MA: College-Hill Press, 1993.

[22] R. L. Freyman, K. S. Helfer, D. D. McCall, and R. K. Clifton, "The role of perceived spatial separation in the unmasking of speech," *J. Acoust. Soc. Am.*, vol. 106, pp. 3578-3588, 1999.

[23] R. L. Freyman, U. Balakrishnan, and K. Helfer, "Spatial release from informational masking in speech recognition," *J. Acoust. Soc. Am.*, vol. 109, pp. 2112-2122, 2000.

[24] N. I. Durlach, C. R. Mason, G. Kidd, Jr., T. L. Arbogast, H. S. Colburn, and B. G. Shinn-Cunningham, "Note on informational masking," *J Acoust Soc Am*, vol. 113, pp. 2984-7, 2003.

[25] N. I. Durlach, C. R. Mason, B. G. Shinn-Cunningham, T. L. Arbogast, H. S. Colburn, and G. Kidd, Jr., "Informational masking: counteracting the effects of stimulus uncertainty by decreasing target-masker similarity," *J Acoust Soc Am*, vol. 114, pp. 368-79, 2003.

[26] T. L. Arbogast, C. R. Mason, and J. Kidd, Gerald, "The effect of spatial separation on informational and energetic masking of speech," *J. Acoust. Soc. Am.*, vol. 112, pp. 2086-2098, 2002.

[27] T. L. Arbogast, "The effect of spatial separation on informational and energetic masking of speech in normal-hearing and hearing-impaired listeners," in *Communication Disorders*. Boston: Boston University, 2003.

[28] G. Kidd, Jr., C. R. Mason, and T. L. Arbogast, "Similarity, uncertainty, and masking in the identification of nonspeech auditory patterns," *J Acoust Soc Am*, vol. 111, pp. 1367-76, 2002.

[29] T. L. Arbogast and J. Kidd, Gerald, "Evidence for spatial tuning in informational masking using the probe-signal method," *J. Acoust. Soc. Am.*, vol. 108, pp. 1803-1810, 2000.

[30] R. S. Bolia, W. T. Nelson, and M. A. Ericson, "A speech corpus for multitalker communications research," *J. Acoust. Soc. Am.*, vol. 107, pp. 1065-1066, 2000.

[31] B. G. Shinn-Cunningham, N. Kopco, and T. J. Martin, "Acoustic spatial cues contained in binaural room impulse responses from a classroom," *Journal of the Acoustical Society of America*, submitted.

[32] R. V. Shannon, F. G. Zeng, V. Kamath, J. Wygonski, and M. Ekelid, "Speech recognition with primarily temporal cues," *Science*, vol. 270, pp. 303-304, 1995.

[33] D. S. Brungart, "Informational and energetic masking effects in the perception of two simultaneous talkers," *J Acoust Soc Am*, vol. 109, pp. 1101-9, 2001.

[34] K. Kawakyu and B. G. Shinn-Cunningham, "How does the brain compute where multiple sounds are located?," Proc of Annual Forum of the Harvard-MIT Health Science and Technology Program, Boston, MA, 2002.

[35] J. Kidd, Gerald, C. R. Mason, T. L. Rohtla, and P. S. Deliwala, "Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns," *J. Acoust. Soc. Am.*, vol. 104, pp. 422-431, 1998.

[36] B. G. Shinn-Cunningham, "Acoustics and perception of sound in everyday environments," Proc of 3rd International Workshop on Spatial Media, Aizu-Wakamatsu, Japan, 2003.

[37] B. G. Shinn-Cunningham, "Spatial hearing advantages in everyday environments," Proc of ONR workshop on Attention, Perception, and Modeling for Complex Displays, Troy, NY, 2003.

[38] R. Desimone and J. Duncan, "Neural mechanisms of selective visual attention," *Annu Rev Neurosci*, vol. 18, pp. 193-222, 1995.