

# Hearing speech against spatially separate competing speech versus competing noise

WILLIAM NOBLE and STEPHEN PERRETT

*University of New England, Armidale, New South Wales, Australia*

Listeners had the task of following a target speech signal heard against two competitors either located at the same spatial position as the target or displaced symmetrically to locations flanking it. When speech was the competitor, there was a significantly higher separation effect (maintained intelligibility with reduced target sound level), as compared with either steady-state or fluctuating noises. Increasing the extent of spatial separation slightly increased the effect, and a substantial contribution of interaural time differences was observed. When same- and opposite-sex voices were used, a hypothesis that the similarity between target and competing speech would explain the role for spatial separation was partly supported. High- and low-pass filtering showed that both parts of an acoustically similar competing signal contribute to the phenomenon. We conclude that, in parsing the auditory array, attention to spatial cues is heightened when the components of the array are confusable on other acoustic grounds.

When a broadband noise source is spatially separated from a speech signal, it is easier to listen to the speech, as compared with when both noise and speech are coming from the same spatial position (Hirsh, 1950). In the case in which a speech signal is coming from straight ahead ( $0^\circ$  azimuth) and a single noise source is displaced horizontally from it, the beneficial effect of separation occurs largely because the noise is reduced in overall loudness, owing to the shadowing effect of the head (Bronkhorst & Plomp, 1988). Although there is an increase in the loudness of the displaced noise in the ear that lies on the same side as it, this is less than the reduction in level at the shadowed ear. The effect is more noticeable for the higher frequency components of a broadband noise, since there is greater shadowing for sounds whose wavelength is less than the head's diameter (Masterton, Heffner, & Ravizza, 1969).

When two competing noise sources are symmetrically displaced from a speech signal (at  $0^\circ$ ) in opposite directions on the horizontal plane, spatial separation effects are less noticeable. This is because a head-shadow effect on one side will be counterbalanced by contrariwise shadowing on the other (Bronkhorst & Plomp, 1992). Bronkhorst and Plomp (1992), using independently fluctuating speech-shaped noises (Festen & Plomp, 1990), found some separation benefit even in conditions of symmetrical displacement. They observed that, in this condition, as compared

with nonspatially separate presentation, a listener can take advantage of momentary variations in the signal-to-noise ratio at a particular ear.

A benefit from separation has also been observed in the case of symmetrically displaced steady-state speech-shaped noises (Ter-Horst, Byrne, & Noble, 1993). Under these conditions, it might be expected that no separation effect will occur; the two noises should fuse so as to be identical with that of noise centered at  $0^\circ$ .

Listeners routinely have the task of trying to hear speech signals in competing conditions. Very often, the competing sound is itself speech that is spatially separated from the target signal. Spoken speech varies substantially through time in terms of level and spectral content. There may be more opportunities for a listener to have auditory "glimpses" of a signal that is to be attended to against such a background, as compared with the relatively greater consistency of random noise.

This point bears on the likely explanation for the separation benefit observed by Ter-Horst et al. (1993). In the separation conditions of that study, the same noise was presented at positions symmetrically displaced relative to  $0^\circ$  azimuth. Thus, the separate noises were completely correlated and would exhibit momentary additive and subtractive effects, recurring at particular frequencies across the audible spectrum, at any position other than  $0^\circ$ , because of systematic phase differences in the two wavefronts at positions occupied by the listener's ears. The resulting peaks and troughs in the spectrum of the signal should allow a listener to hear the speech at  $0^\circ$  better than when the noise is broadcast only from that same position. There is a measurable acoustic difference associated with the use of correlated versus uncorrelated noise under conditions of spatial unity versus separation, as will be explained in note 1, referred to in the Competing Signals subsection of Experiment 1.

---

The research reported here was supported by Grant A79700708 from the Australian Research Council. We thank Dean Davidson for help with software development, David Heap and Frank Niebling for help in apparatus design and construction, Don Hine for advice on statistical analysis, and Malcolm A. Perrett for DSP advice. Finally, we thank the reviewers for helpful criticism of the original version of the paper. Correspondence concerning this article should be addressed to W. Noble, School of Psychology, University of New England, Armidale, NSW 2351, Australia (e-mail: wnoble@metz.une.edu.au).

Symmetrically separate, uncorrelated noise will exhibit no consistent phase-linked effects, because the wavefronts are structurally unrelated to each other. Hence, it may be expected that no equivalent separation benefit will be found in such a case. The first experiment in the series reported here was designed to compare the effect on speech intelligibility of correlated versus uncorrelated noises symmetrically separated from the speech, in free-field listening conditions. In addition, we investigated the effect of speech as a competitor signal, as compared with noise. The measure of interest throughout was the effect on intelligibility of a target speech signal, positioned at 0° azimuth near the horizontal plane, when competing signals were presented from that same location (called the *mixed* condition) or from two flanking positions displaced at azimuths 30° to the left and the right of 0° (called the *separate* condition).

In the conditions of our experimental arrangement, speech was observed to offer a greater symmetrical separation benefit than did either form of broadband noise, and in subsequent experiments, we undertook examination of the factors that might account for this. Following a series of comparisons with noises that had a number of different temporal and spectral structures, attention was given to similarity versus dissimilarity of target and competing speech as such. When female versus male voices were used as target and competitor signals, partial support was gained for the hypothesis that a spatial effect arises or subsides depending on the similarity versus the dissimilarity of the signals employed. Simultaneous interaural time differences were identified as a significant factor enabling the speech separation advantage to occur, and in the final experiment, it was observed that both lower and higher frequency parts of the signal contribute to the effect.

The results can be interpreted in the light of recent reports examining the contribution of spatial cues to the segregation and identification of elements in the *auditory scene* (Bregman, 1990). Spatial cues play a role, but only in conditions in which other bases for segregation are less salient. Such a proposition leads to further reflection on the place of spatial hearing in the task of attending to the audible environment. If it is correct to identify the primary role of spatial hearing as serving a dynamic, orienting function, this may help to explain the inconstancy of its contribution to the perceptual organization of stationary auditory scenes.

## EXPERIMENT 1

### Correlated Versus Uncorrelated Noise

#### Method

##### Apparatus and Materials

Testing was conducted in a sound-isolated, semi-anechoic chamber with a 3 m × 3 m floor area, and 3.5 m in height. The listener sat in a height- and back-adjustable chair, with the head motionless, using a chinrest. In addition, a constant front-facing orientation was maintained during testing by the listener aiming at a light-emitting diode (LED) spot at 0° azimuth and elevation, using a laser pointer

attached to a lightweight adjustable head harness. Signals were presented either in combination from a center loudspeaker or separately from three loudspeakers (5-cm diameter midrange tweeters), depending on the test condition. The loudspeakers were arrayed 30° apart in an arc so as to subtend the same distance (45 cm) from the center of the listener's head when it was in position on the chinrest. The center loudspeaker of the three was at the straight-ahead position on the median vertical plane, with the whole arc 15° below the horizontal plane.

All the signals were inverse filtered to remove differences owing to individual loudspeaker transfer functions, using 400th-order digital filters, with 20 dB of cut, loaded into a PD1 (Tucker Davis). The digital sampling rate was 47 kHz. In the first two experiments, all the signals were bandpass filtered to attenuate energy below 500 Hz and above 18 kHz, in order to eliminate transfer function differences greater than 20 dB occurring below 300 Hz and above 18 kHz and filter inaccuracy below 500 Hz. In the third and fourth experiments, competitor signals were 4,096th-order inverse filtered, off line, enabling accurate inverse filtering to 300 Hz. In the fifth, sixth, and seventh experiments, all the signals were handled this way, enabling inverse filtering of both target and masker signals between 300 Hz and 18 kHz.

**Target speech signal.** The *target speech* was a continuous 45-min discourse (a topical book review) read in British English by a male talker (W.N.), recorded onto digital audio tape (DAT). When the signal was recorded, the talker used the ongoing dBA read-out from a sound level meter, to ensure an even normal voice level throughout. During testing, the signal was reproduced via the DAT recorder, set in play mode throughout the course of the trials. The 45-min discourse was long enough in almost every case to provide a nonrepeating target for the listener to follow (1 or 2 listeners needed longer than 45 min to complete the task in Experiment 2, the most time consuming in the series; the tape was begun afresh to accommodate this). The target speech signal was inverse filtered via the PD1 and was passed through a digitally controlled attenuator, an audio mixer, and an audio amplifier before being presented to the listener from the center loudspeaker. During each trial, the level of the target speech signal was under the listener's control via a hand-held cordless trackball that fed output to the digital attenuator via a PC.

**Competing signals.** The *competing speech* signals were produced from a 20-sec segment of discourse read by a second male talker, taken from a CD recording of a fictional work. The 20-sec segment was then broken into approximately equal parts (A and B). Both parts began and ended at natural breaks in the monologue. The parts were spliced together, with B following A for the left channel and A following B for the right, and were stored on a hard disk in two 16-bit digital data files. During a trial, each 20-sec file was loaded into a digital buffer of an AP2 (Tucker Davis) and was played as a continuous loop. As with all the other competing signals, these were streamed to the PD1 for loudspeaker inverse filtering, then converted to analogue signals before being amplified and switched to the appropriate loudspeakers (left and right). When played in a trial, the left signal was identical to the right, except that it was temporally displaced by about 10 sec.

The *correlated noise* was produced using a 1-sec segment of Gaussian noise digitally generated on the AP2. This was then duplicated to produce left and right signal data. These data were placed in a left and a right digital buffer and were played as continuous loops. The result was streamed to the PD1, where inverse filtering and speech-shaped filtering were applied. When played in a trial, the left signal was identical to the right. *Uncorrelated noise* used a 1-sec segment of Gaussian noise digitally generated on the AP2 and duplicated to produce left and right digital signal data. Each 1-sec segment of data was then cut into two equal parts (A and B), and the parts were spliced together, with B following A for the left signal and A following B for the right. These data were then buffered and

processed in the same way as the correlated noise. When played in a trial, the left signal was identical to the right, except that it was temporally displaced by 500 msec.

When competing signals were played from the center loudspeaker (mixed condition), they were generated in the same way as described above, except that they were inverse filtered according to the characteristics of the center loudspeaker, switched through an audio mixer along with the target speech signal before being amplified, then fed to the center loudspeaker.

All the competitor sounds were calibrated using a 0.5-in. Bruel & Kjaer microphone (Type 4176) attached to a B & K precision sound level meter (Type 2235), the microphone being located at the position to be occupied by the center of a listener's head. Whether being presented from the center loudspeaker or from the flanking loudspeakers, all the competitor signals were first individually measured, using the "slow" response and A-weighted circuit of the sound level meter, and were adjusted to achieve identical levels at the microphone position. The signals were then remeasured when presented simultaneously, with the whole system being adjusted to achieve a level of 60 dBA.<sup>1</sup> Calibration of the system was checked repeatedly throughout the experimental series and daily within the course of any experiment.

### Listeners and Procedure

Nine people took part in the first experiment (3 females, 6 males), recruited from the student and staff population of the university. They reported no deficiency in hearing in either ear, and all scored at or below the "low fence" on Coren and Hakstian's (1992) Hearing Screening Inventory, a criterion applied throughout the series of experiments. The Hearing Screening Inventory has a specificity (correct detection of normal hearing) of 93% and a sensitivity of 91% against the AMA audiometric standard of 25 dB at 0.5, 1, 2, and 4 kHz. The authors pilot-tested all the procedures and experiments reported here and also provided data for this and the subsequent experiments. No particular differences were observed in the nature of the results from these more experienced listeners and the majority recruited from beyond the laboratory.<sup>2</sup> The samples recruited to each new experiment were a mix of people with experience of at least one earlier experiment in the series and of people with no such experience. In all the experiments, the majority of the listeners were new recruits.

The listener's task was to use the trackball to adjust the target speech signal to a level that yielded nearly complete intelligibility (we have called this a *90% criterion*). The instruction was to adjust the speech to a point at which, if it were any softer, it would be difficult to follow, and if it were any louder, it would be effortless to follow. This criterion had been found, in pilot experimenting, to provide more repeatable determinations than, for example, a 50% intelligibility criterion. All the listeners were familiarized with the task before actual testing began.

The consistency and validity of the technique of adjusting a continuous speech signal to a nominated level of intelligibility has been demonstrated by Speaks, Parker, Harris, and Kuhl (1972). In that study, 3 normal hearing listeners each made 10 adjustments of a nonrepeating speech signal to each of 3 nominated percentage-of-intelligibility levels (25%, 50%, and 75%) in competing white noise: Standard deviations of the 10 adjustments ranged from 0.8 to 1.3 dB. Speaks et al. also observed that the listeners' adjustments to target intelligibility levels matched (to within 1–2 dB) their estimates of the intelligibility of a series of signals at various fixed signal-to-noise ratios. Furthermore, in a task in which listeners ( $N = 8$ ) estimated the intelligibility of sentence lists presented at a series of fixed signal-to-noise ratios, the correlation with measured intelligibility (percentage of words correctly repeated) at those same ratios was .93. Agreement was particularly close at the extremes of *nearly complete intelligibility* and *nearly complete unintelligibility*.

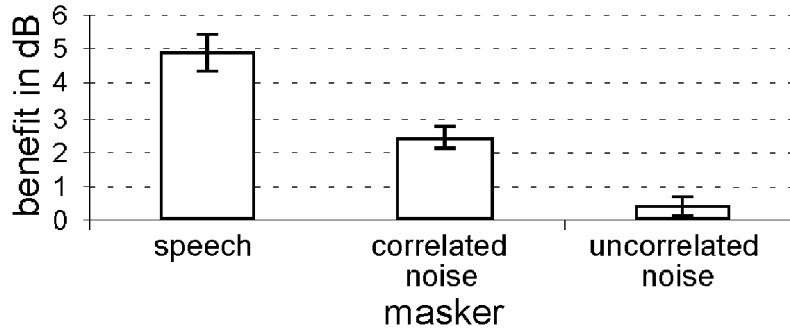
The test–retest reliability of the speech adjustment technique is also established. Walker and Byrne (1985) used a repeating 2.5-min segment of continuous speech, heard in speech spectrum noise at either 55 or 85 dBA. Testing was carried out on 3 successive days with 12 normal hearing listeners. Their task on each day was to make two adjustments of the speech to achieve 25%, 50%, and 75% intelligibility in both noise levels. The median standard error of measurement was 0.45 dB, and there was no difference in the stability of performance across different intelligibility or noise levels. In a second experiment in which four trials per condition were used, it was noted that, in noise conditions, no trend was observable across the trials. Finally, Kollmeier and Wesselkamp (1997) showed the equivalence of outcomes for adjusted and performance-based speech intelligibility testing, and Peissig and Kollmeier (1997) used the adjustment procedure in a speech-versus-speech experimental paradigm somewhat similar to aspects of the present experiments.

At the start of a trial in the present experimental series, the laser pointer light was illuminated, enabling targeting on the LED. The speech and competitor signals were initiated when the listener pressed a switch on the hand-held attenuator. The target was always initially at a clearly higher level than the competing signals. In the first experiment, there were five trials per condition, six conditions altogether (speech, uncorrelated noise, and correlated noise, each in mixed and separate conditions), with presentation order randomized across conditions. In subsequent experiments, there were four trials per condition. There was no time limit on any trial, but the listeners were signaled after 60 sec via the LED's changing from continuous to pulsed. A trial was terminated when the listener was satisfied that the *nearly complete intelligibility* criterion was being maintained and had again pressed the attenuator switch. Arithmetic average attenuator values over the trials per condition for each listener were used as the basis for calculating results.

### Results

The outcome of primary interest throughout the experimental series was the difference in performance when the listeners listened in spatially mixed versus spatially separate conditions, and these are the results we will present (we will also refer to aspects of raw attenuation data from which these difference values were derived). For any given competitor signal, the target speech was always adjusted from an identical initial level in both the mixed and the separate spatial conditions. If spatial separation makes it easier to hear the target speech, this would be reflected in adjustment of that target to a lower level in separate versus mixed conditions. Any such benefit is expressed as a positive value in the histograms displayed in this and in each of the ensuing Results sections. Figure 1 shows the effects of spatial separation of (1) speech from competing speech, (2) speech from competing correlated noise, and (3) speech from competing uncorrelated noise.

It is clear that correlated noise provides a small (2.4 dB) separation benefit and that uncorrelated noise provides practically none (0.4 dB). This contrasting outcome suggests that the result reported by Ter-Horst et al. (1993) was due to their use of correlated noises.<sup>3</sup> The most notable benefit in the present experiment, however, was from spatial separation of speech from speech (5.0 dB). An analysis of variance (ANOVA) across the benefit measure yielded [ $F(2,16) = 55.12, p < .0005$ , with the differences between all three signal types significant at the  $p < .05$



**Figure 1.** Average effects (and  $\pm 1$  SE bars) on the intelligibility of a target speech signal at  $0^\circ$  azimuth owing to symmetrical spatial separation of competing speech, competing correlated noise, and competing uncorrelated noise. Values represent the difference in attenuation needed to maintain 90% intelligibility in separate versus mixed conditions. Positive values indicate benefits from such separation (lower target signal level needed to maintain intelligibility).

level (Tukey's HSD test). The subsequent experiments were designed to examine further the particular effect of speech as a competing signal.

## EXPERIMENT 2 Same Versus Different Temporal and Spectral Patterns

The first experiment indicated that symmetrical spatial separation gave rise to a benefit in hearing target speech at  $0^\circ$  against two different competing speech signals  $30^\circ$  on either side of it. Two substantial acoustical variations in competing speech take the form of differences in temporal and spectral patterns. In the mixed condition, energy bursts from either competitor speech signal may coincide or interleave in time, whereas in the separate condition, interleaved energy will occur at different spatial loci. Similarly, the spectral content of the competing speech signals differs, and those differences, in the separate condition, will be specific to different locations.

### Method

#### Materials

The second experiment was designed to investigate the contribution of the temporal and spectral characteristics noted above to the spatial separation effect. It involved forms of noise that varied in independent versus identical ways in temporal or spectral patterns.

**Temporal variation.** A 500-msec sample of Gaussian noise was digitally generated on the AP2 and was duplicated to produce left and right signal data. The samples were cut into two equal segments (A and B), each windowed to apply 25-msec on/off ramps. Segments of silence of 250 msec were interspersed with the noise segments. For the *different temporal* condition, A was followed by silence, then by B, then by silence in the left channel; in the right channel, B was preceded by silence and then was followed by silence and then by A. For the *same temporal* condition, A, B, and the periods of intervening silence coincided with each other. The left and right signal data under both conditions were placed in digital buffers and were played as continuous loops. Filtering and calibration procedures were the same as those described for Experiment 1.

**Spectral variation.** Two different 250-msec samples of random noise were generated in Matlab and were filtered, using a 1,024th-order filter to create spectral notches at 1.5, 3, 6, and 12 kHz in one of the samples and either at those same frequencies or at 1, 2, 4, and

8 kHz in the other. For the *different spectra* condition, the 1.5- to 12-kHz pattern was assigned to the left channel, and the 1- to 8-kHz pattern was assigned to the right. For the *same spectra* condition, the 1.5- to 12-kHz pattern was used in both channels. The filtered data were saved as two separate 250-msec 16-bit data files on PC hard disk. During a trial, each 250-msec file was loaded into a digital buffer and was played as a continuous loop, one for the left channel and one for the right.

The target speech signal used for all the conditions was the same as that in Experiment 1. A competing speech condition was also included in Experiment 2, using the same competing signal as before.

#### Listeners

Ten listeners were recruited from the student and staff population of the university (3 females, 7 males), meeting the same hearing standard as before. Some of the listeners had taken part in the previous experiment; the majority were new recruits. No evident difference in response patterns was observed between more and less experienced listeners. There were 10 conditions in this experiment (same vs. different temporal patterns of competing noise, same vs. different spectral patterns, and competing speech, each in mixed and separate spatial conditions), with four trials per condition.

### Results

The effects of spatial separation (positive and negative differences between mixed and separate conditions) are shown in Figure 2. An ANOVA showed a significant main effect [ $F(4,36) = 19.85, p < .0005$ ] for type of competing signal. Post hoc testing (Tukey's HSD) confirmed the impression from the graphed results that the different temporal patterns offered a spatial separation advantage significantly greater than that from the same temporal pattern ( $p < .05$ ). By contrast, different versus identical spectral notches yielded similar outcomes, and there was no clear difference between either of the spectral conditions and the one temporal condition that showed separation benefit. Competing speech showed a beneficial separation effect significantly greater than that from any of the competing nonspeech signals ( $ps < .05$ ).

### Discussion

These results allow us to suggest that nonsimultaneous variation in the levels of competing signals explains some-

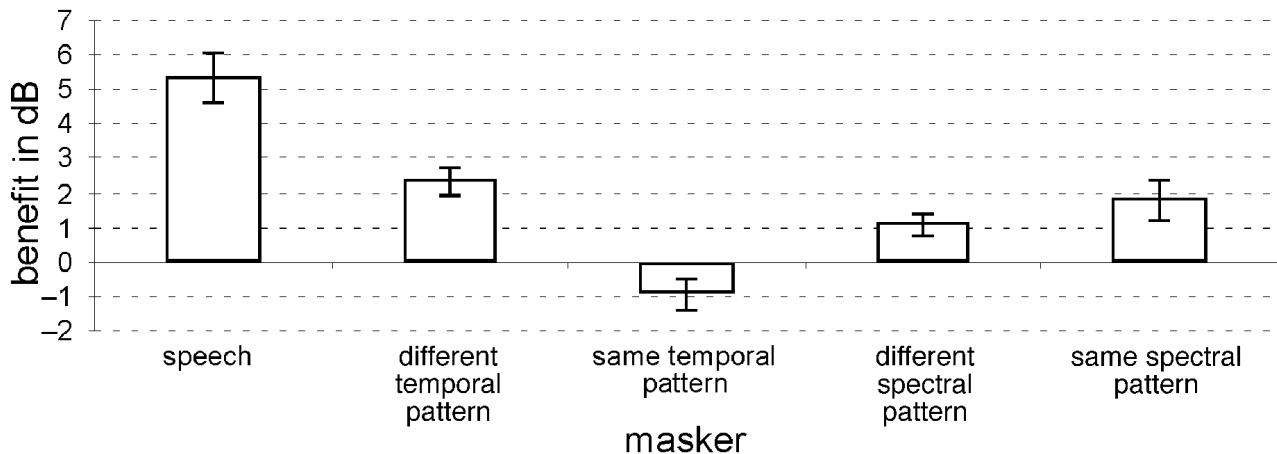


Figure 2. Effects of spatially separating target speech from competing speech, competing noise of different and same temporal structure, and competing noise of different and same spectral structure. Positive values indicate benefits from such separation (lower target signal level needed to maintain intelligibility).

thing of the spatial separation effect observed with speech as competitor. The contribution of spectral variation is less clear-cut: No obvious difference was observable between the two conditions in question. The slightly greater effect for the same spectra condition may be argued to be consistent with the contrast between correlated versus uncorrelated noise in Experiment 1. Notches occurring at the same spectral positions in two otherwise uncorrelated noises may be understood as analogous with notches arising at particular spectral positions owing to phase cancellation effects in two spatially separate but otherwise correlated noises.

Inspection of the raw attenuation data showed that, in the same temporal pattern condition, the listeners needed much lower target signal levels (whether in mixed or separate conditions) to maintain intelligibility, as compared with performance with other competing signals. The task is evidently easier when there are equal periods of silence interspersed with bursts of noise. There were considerable differences across individual listeners' responses in the same temporal pattern condition. Performance levels with signals mixed versus separate were nonetheless highly correlated ( $r = .96$ ), suggesting that the two conditions were effectively identical.

We concluded that nonsynchronous fluctuations in sound level, coupled with some order of spectral difference, accounts for some of the effect of spatially separating competitor and target speech. Nonetheless, there remained a particular effect for speech, which led us to an experimental test using supposedly closer approximations to that signal.

### EXPERIMENT 3 Approximations to Speech

#### Method

##### Materials

To explore the temporal/spectral issue, using signals more akin to the acoustic patterns of speech, two forms of competing noise were

constructed in which spectral characteristics either were or were not included within a speech-shaped fluctuating envelope. The noise that retained spectral variations we have called *blurred speech*. It was generated by obtaining the sinusoidal components of overlapping 350-msec sections of the competing speech signals (described earlier). The sections were captured at approximately 90-msec intervals throughout the duration of the 20-sec speech signal. The phase of each of the sinusoidal components in each section was randomized, and the resulting sections were overlapped at approximately 90-msec intervals, with the sound level of each being the same as the original signal. Two separate blurred speech noise streams were generated according to the same principles described in relation to competing speech presentation.

The second competing noise, *fluctuating noise*, was produced in a way similar to that for the blurred speech signal, but instead of the sinusoidal components being randomized, each 350-msec section was filled with white Gaussian noise. The signal was then filtered to give it an overall spectrum identical to the competing speech signal. The essential difference between these noises was that the spectrum of the fluctuating noise signal was relatively constant, whereas that of the blurred speech signal changed over time. Power spectra and envelopes for these two competitors are shown in Figure 3.

Given the consistently stronger spatial separation effect for speech as a competitor, as compared with nonspeech sounds, it was appropriate to observe the effect of a close acoustic approximation to speech, with meaning removed. Thus, in addition to the above noise conditions and a normally running speech competitor, the competitor speech loops were also presented reversed.

##### Listeners

There were 11 listeners in this experiment (5 females, 6 males), meeting the same hearing standard as that in the previous experiments. As before, the majority of the listeners were new to the experimental task. There were eight conditions: blurred speech, fluctuating noise, reversed speech, and competing (forward) speech, each in mixed versus separate conditions, with four trials per condition.

##### Results

The effects of separation in each condition are shown in Figure 4. It is clear, first of all, that the two speech competitors (forward and reverse) yielded broadly similar outcomes, as did the two forms of noise, and that both speech

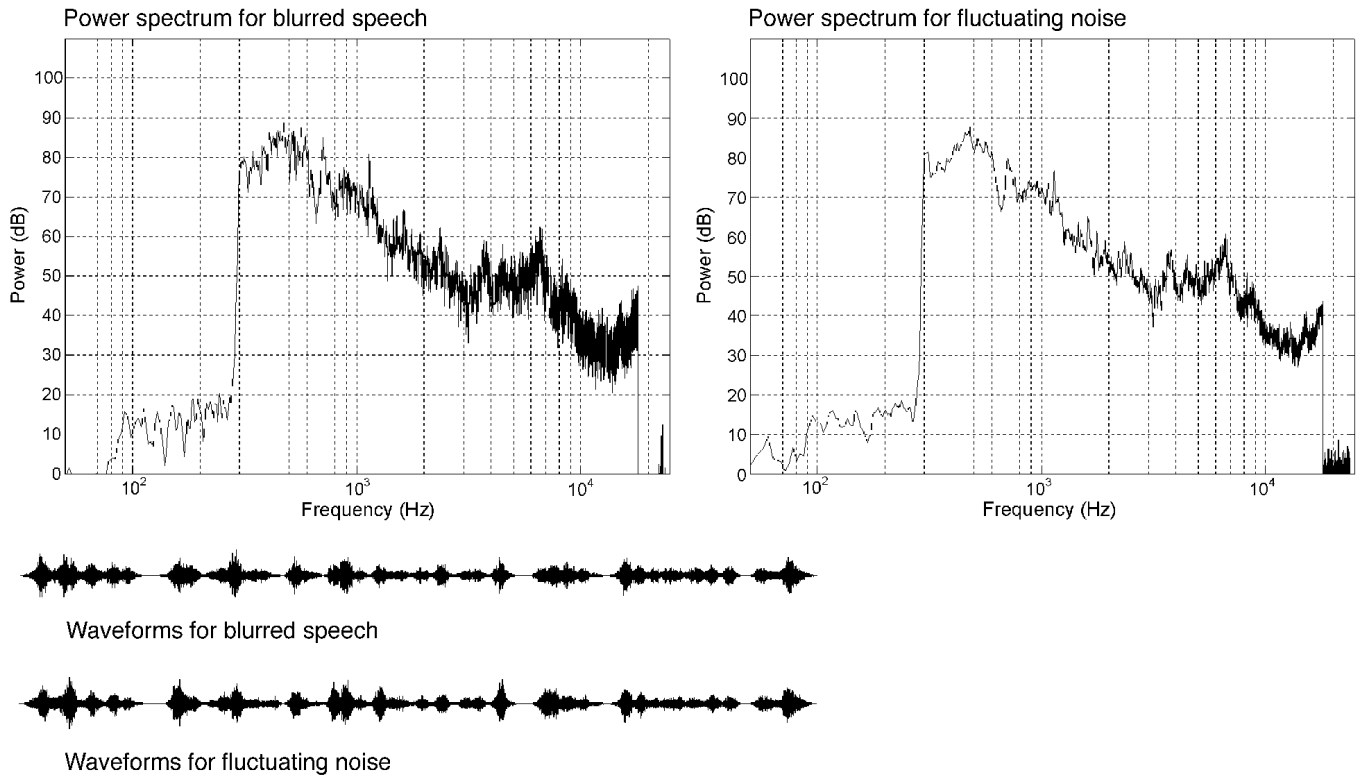


Figure 3. Power spectra and envelopes for the two competitor signals used in Experiment 3: blurred speech and fluctuating noise. The images depict sounds after loudspeaker inverse filtering has been applied.

conditions were distinct from both noise conditions. A one-way ANOVA revealed a significant main effect [ $F(3,30) = 22.9, p < .0005$ ]; post hoc testing using Tukey's HSD showed that the differences between both of the speech and both of the noise competitors were significant at the  $p < .05$  level but that the differences between forward and reverse speech and between blurred speech and fluctuating noise were not. We noted from raw attenuation data that in the mixed condition, listeners set the target speech 1.5 dB higher, on average, when forward-running speech was the competitor, as compared with reverse speech (0.5 dB higher in the spatially separate conditions). As a result, there was a slightly greater (1 dB) benefit from spa-

tial separation of competing forward speech, as compared with reversed.

**Discussion**

It was somewhat harder, in the mixed condition, to follow the target speech signal when the competing signal was normal (forward) speech, as compared with reversed speech. The element of meaning appears to have added a small interfering effect to the task. Even though the competing signal was a 20-sec loop, the listeners could not avoid being distracted by the limited meaning that was available. However, spatial separation offered a benefit under both competitive speech conditions, suggesting that

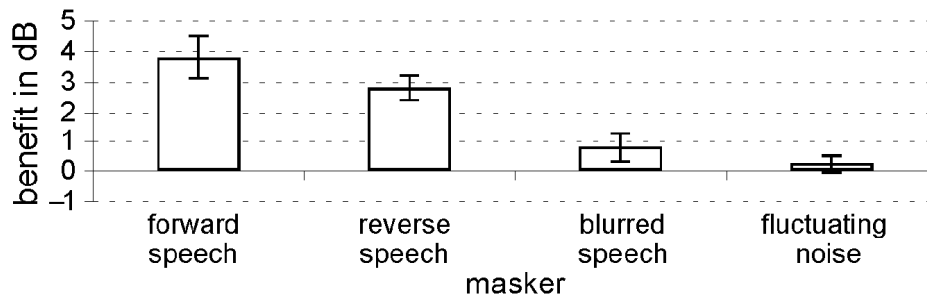


Figure 4. Effects of spatially separating target speech from competing speech, time-reversed competing speech, blurred speech, and fluctuating noise. Positive values indicate benefits from such separation (lower target signal needed to maintain intelligibility).

the listeners took particular advantage of *acoustic* features specific to speech when the sources were spatially separated. This may be understood as an instance of the phenomenon of *informational masking* (e.g., Kidd, Mason, Rohla, & Deliwala, 1998), in which structural similarity between target and competing signals adds to the interfering effect of sheer *energetic masking*. Spatial separation is especially effective in overcoming informational masking (Kidd et al., 1998).

The relative absence of an effect for either of the speech-like noises was unexpected. The design was based, in part, on the finding by Bronkhorst and Plomp (1992), who observed a benefit from symmetrical spatial separation of the fluctuating noise devised by Festen and Plomp (1990). Bronkhorst and Plomp (1992) used 90° of separation when two competing noise sources were played; it is probable that the greater spatial contrast enhances any beneficial effect. To check whether the present outcome was due to this characteristic, the next experiment was designed to investigate the speech versus noise contrast by using two levels of spatial separation. Since the noise compiled by Festen and Plomp was also somewhat different from those used here (see below), the opportunity was taken to employ the fluctuating noise stimuli they devised.<sup>4</sup>

#### EXPERIMENT 4 Comparing Different Extents of Spatial Separation

##### Method

##### Apparatus and Materials

Two further loudspeakers were placed at ±90° of the central loudspeaker, on the same plane as those at ±30° (i.e., 15° below the horizon), and with appropriate compensation made for their individual transfer functions. As before, the target speech signal was delivered to the central loudspeaker, with competing signals to that same source in the mixed condition or to either two loudspeakers at 30° to the left and right of the center, or two new locations at ±90° in the separate condition. The competing noise used in this experiment was wideband, separately modulated in lower and a higher frequency bands by multiplication with a male voice envelope (see Festen & Plomp, 1990, pp. 1726–1727).

##### Listeners

Five males and five females acted as listeners, most of them new to the experimental task and meeting the same hearing criterion as before. There were six conditions: speech versus speech at 0°, ±30°, and ±90° of separation and speech versus the fluctuating noise from Festen and Plomp (1990) at those same positions. There were four trials per condition.

##### Results and Discussion

Figure 5 shows the outcome of this experiment, expressed in the same way as for the earlier ones. A two-way ANOVA showed no interaction [ $F(1,9) = 0.70$ ,  $MS_e = 0.42$ ] but main effects both for angle of separation [ $F(1,9) = 9.01$ ,  $p = .015$ ] and for type of competitor [ $F(1,9) = 55.74$ ,  $p < .0005$ ]. Increasing the angle of spatial separation from 30° to 90° affected (slightly) the extent of the benefit observed. With speech as the competing signal, the angular effect was less than 1 dB. With fluctuating noise as the com-

petitor, the difference between the smaller and the larger separation angles was a little greater but was still modest in acoustical terms (1.2 dB). Nonetheless, part of the difference between the results in the previous experiment and those reported by Bronkhorst and Plomp (1992) may be explained by the difference in degree of spatial separation of the target from the competitor signals used in the two studies. We note that a feature of results obtained by Yost, Dye, and Sheft (1996) was a relatively greater benefit with 90° than with 30° of symmetrical separation for identifying three simultaneous target speech signals in stationary binaural listening conditions.

The fluctuating noise samples of Festen and Plomp (1990) provided an outcome similar to those observed in the previous experiment, using our own forms of speech-like noise. Bronkhorst and Plomp (1992) reported a more substantial effect (4.8 dB, as compared with 1.5 dB in this experiment) when using Festen and Plomp's fluctuating noise symmetrically separated at ±90°. The relatively smaller amount of low-frequency energy in our signals, as compared with theirs, may account for this difference. Bronkhorst and Plomp (1988) stated that the degree of improvement in speech intelligibility from head-induced interaural time delay operating on a single spatially displaced masker is predictable from the low-frequency (250–500 Hz) component of that delay. Interaural time delay seems to account for the effect we are observing, a point made clearer by the next experiment.

The foregoing experiment showed an effect for greater versus lesser degrees of symmetrical spatial separation. As spatial angle increased, interaural level and time differences increase also. Our design largely controlled for overall differences in level, but in the free-field conditions of these experiments, a complex set of changes would occur with different source locations, owing to the individual listener's head-related transfer function (Wightman & Kistler, 1993). Using headphone listening, it is possible to isolate the contribution of interaural time difference to the symmetrical separation effect. This was the aim of the next experiment.

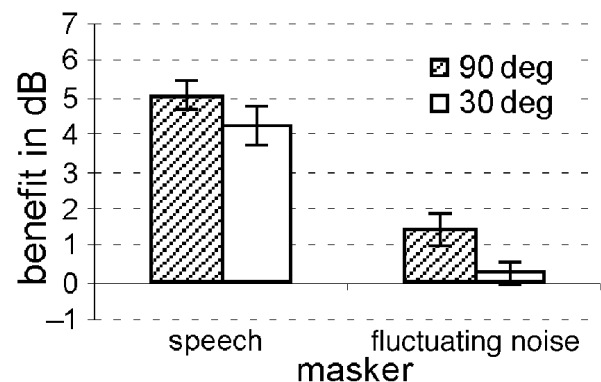


Figure 5. Effects of spatially separating target speech from competing speech and from the fluctuating noise from Festen and Plomp (1990) at two levels of spatial separation (±30° and ±90°). Positive values indicate benefits from such separation (lower target signal needed to maintain intelligibility).

Notwithstanding that outcome, the effect that remains consistent across the series of experiments is the substantial difference that occurs when interfering speech is symmetrically spatially separated from target speech, as compared with separation of speech and noise. We examined one possible basis for this effect in the experiment following the next one.

## EXPERIMENT 5

### The Contribution of Interaural Time Difference

#### Method

In this experiment, the AP2 was used to modify signals for headphone presentation. The left and right competitors were mixed without further processing and were delivered to both headphone channels, to provide the mixed conditions. For the separate conditions, a time delay of 300  $\mu$ sec was applied across the full spectrum of one competitor in one channel and of the other competitor in the other channel, to simulate interaural time differences occurring for one competitor 30° left and another 30° right. No other spatial cues were simulated. The target signal was routed through the digital attenuator (under the listener's control) before being split and mixed into each headphone channel. No inverse filtering was employed for headphone presentations. The signal generation used for loudspeaker presentation was identical to that in the previous experiment for 0° and 30° separation conditions. The target signal at 0° was the continuous male voice discourse used in the earlier experiments. Competing signals were the speech loops also used previously and the Festen and Plomp's (1990) noise samples, used in the previous experiment. As was mentioned in the description of the first experiment, the filtering procedure in this and the subsequent experiments allowed both the target and the competing signals in free field to be band-passed from 300 Hz to 18 kHz. For consistency, the same limits were placed on the signals presented over headphones.

There were 10 listeners (2 females, 8 males), most being new to the experimental arrangement, and all meeting the normal hearing criterion described earlier. There were eight conditions in this experiment: speech versus speech under free-field versus headphone listening (Sennheiser model HD 265 headphones), in both mixed and spatially separate conditions, and speech versus noise under the same conditions. There were four trials per condition.

#### Results

Figure 6 shows the outcome in terms of benefit from spatial separation of speech versus speech and speech versus noise under free-field and headphone listening. The marked contrast between speech and noise as competitors is evident in these results, and this remains the case when the only change was in interaural time differences between target and competing speech. A two-way ANOVA showed main effects for mode of listening [ $F(1,9) = 14.14, p = .004$ ] and for type of competitor [ $F(1,9) = 76, p < .0005$ ] and a nonsignificant interaction [ $F(1,9) = 0.08, MS_e = 0.27$ ]. The free-field listening mode was associated with a somewhat greater acoustical difference than headphone listening between mixed and separate conditions (5.1 vs. 3.7 dB). The free field enables the full effect of individual head-related transfer functions to operate. Nonetheless, the bulk of the effect of spatial separation of speech from speech may be attributed to time differences arising when signal sources are located independently in space.

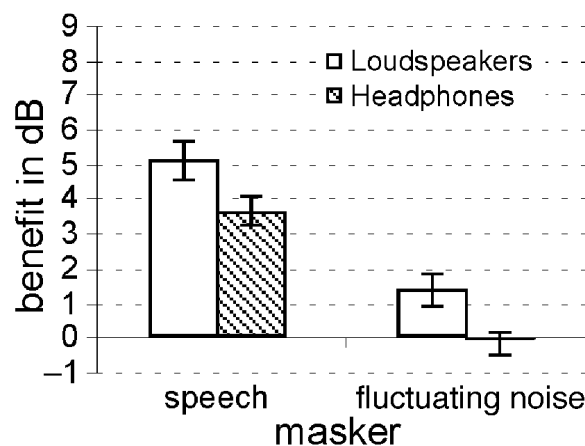


Figure 6. Effects of spatially separating target speech from competing speech and from the fluctuating noise from Festen and Plomp (1990) under free-field and headphone listening. Positive values indicate benefits from such separation (lower target signal needed to maintain intelligibility).

#### Discussion

There was a small separation effect for Festen and Plomp's (1990) fluctuating noise in free field, as compared with the same (30° separation) condition in the previous experiment. There was, concomitantly, a slightly higher average effect for the equivalent speech against speech conditions in this experiment, as compared with the previous one. These parallel outcomes may be due to differences in the composition of the samples tested. We might also expect some variation in outcomes for identical conditions across experiments when the task for a listener in any given experiment involves contrasts between those conditions and other ones that differ among experiments.

It is a feature of the experiments in which fluctuating noise was used (Experiments 3, 4, and 5) that, in the mixed condition, a lower target speech level was needed for 90% intelligibility, as compared with the experiments (1 and 2) in which more steady-state interference was used.<sup>5</sup> This outcome is consistent with Festen and Plomp's (1990) findings. Those authors also reported that their fluctuating noise functioned in the same way as speech, in terms of its masking effect (their observations being made under monaural headphone listening conditions). We had anticipated that fluctuating noise might act like speech in our experiments, involving symmetrical spatial separation of target from competitor signals, all heard binaurally in free field. Instead, we found that competing speech, whether normal or time reversed, provided a consistently different performance pattern from noise (Hawley, 2000, chap. 7, reports a similar speech–noise distinction). Further scrutiny of Festen and Plomp's results showed that when the competing speech signal was the same voice (reversed) as the target, it in fact acted similarly to steady-state noise. Only when the competing speech was different from the target (male vs. female) was the similarity with fluctuating noise observed.



We had not used the same voice for target and interference, but both voices were male talkers; a priori, these may be regarded as similar to, rather than different from, each other. The proposition emerging from the results in the series of experiments to this point was that fine-grain acoustic similarity between target and competition is a key to explaining the different outcomes from the use of speech versus nonspeech. In the next experiment, we examined the issue of greater and lesser similarity of target and competing speech by casting female and male voice targets against both female and male voice competition.

## EXPERIMENT 6

### Similarity of Target and Competing Speech

The question of similarity and difference between voices, in connection with masking effectiveness, has been the subject of research inquiry (e.g., Dirks & Bower, 1969). The point is well appreciated that while similarity may be simply achieved, such as by using the same voice for target and competition, what characterizes dissimilarity is harder to predict. In the design for the present experiment, we relied on the sex of the talkers as the marker of *dissimilarity*, recognizing that results could be influenced by any number of characteristics varying between any pair of voices (see Freyman, Helfer, McCall, & Clifton, 1999, p. 3586, on this point). The purpose of the experiment was to test the general hypothesis that symmetrical spatial separation of target from competitor voices will be affected differently depending on whether the signals are more or less similar to each other, while all being in the category of intelligible speech.

### Method

#### Materials

In approaching the question of similarity, we relied on stock recorded discourse, using narratives of published texts on CDs. This took us into uncharted territory as regards the male voice, but it was thought to be a fairer test of the hypothesis to start from positions equally unknown for either type of voice. The competing male voice loops were retained, new target speech discourses were selected (extracts from works of fiction) involving a female and a male reader, and a new pair of speech loops was produced from a second female-voice recording, along the same lines as those produced previously for the male competitor signals.

#### Listeners

There were 12 listeners in this experiment, 2 females and 10 males, recruited according to the same criteria as before. There were eight conditions (four trials per condition): male target versus male competitors, male target versus female competitors, female target versus female competitors, and female target versus male competitors, each in mixed and in spatially separate listening conditions.

### Results

Figure 7 shows the benefit from separation values for both the female and the male target speech against both female and male competitors. It is clear that symmetrical separation of male competing speech from the female tar-

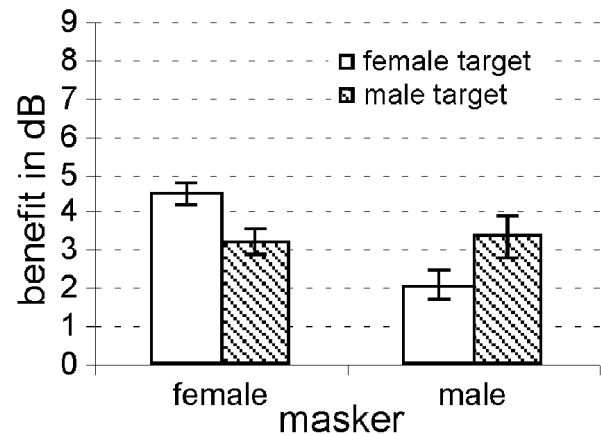


Figure 7. Effects of spatially separating female and male target speech from female or male competing speech. Positive values indicate benefits from such separation (lower target signal needed to maintain intelligibility).

get conferred little effect, whereas the female versus female condition yielded an outcome similar to those observed in the previous experiments involving male versus male. By contrast, no equivalent outcome was observed for the new male target, although there was a nonsignificant trend in the expected direction. A two-way ANOVA confirmed a significant target  $\times$  competitor interaction [ $F(1,11) = 13.61, p = .004$ ], with post hoc testing (Tukey's HSD) showing only the contrast between female target versus female competitors and female target versus male competitors as significant at the  $p < .05$  level.

### Discussion

The hypothesis about similarity gains partial support from these results. It would be gratifying if the male target data had shown an effect equivalent to that for the female; we are satisfied, on the basis of the previous experiments, that such an outcome would be likely to be demonstrated with a different voice pairing. When we listened to the new male target voice, it had an apparently more accented manner, which may have reduced its confusability. It should be possible to identify particular pairs of voices that would maximize the similarity effect; our purpose was not that, but rather to test a more general hypothesis.

## EXPERIMENT 7

### The Contribution of Different Frequency Bands

The fifth experiment suggested that interaural time differences play a dominant role in the effect of symmetrical spatial separation of competing and target speech. The mechanism we surmise to be operating is one that engages the listener in attending to spatial differences between target and competitors by virtue of the interaural time difference patterns that are absent in the mixed condition but present, moment by moment, in the separate condition, thus specifying distinguishable locations for both competi-

tors. In the final experiment, we were interested in observing the relative contribution of low- and high-frequency parts of the signal to the overall separation effect.

### Method

Twelve listeners (3 females, 9 males) took part in this experiment, the sample being similar in composition to previous ones. We used the female target and the female competitor voices introduced in the previous experiment, but under low- and high-pass filtering conditions, as well as broadband (all-pass). Employment of these new speech signals would test the robustness of the previous finding. A cutoff at 2000 Hz was selected. This is in a frequency region below which interaural *fine structure* or carrier (phase) delay is understood to dominate and above which envelope-onset delay is prominent (Zurek, 1993). The 2000-Hz cutoff also allowed for approximately similar intelligibility of the target speech in the absence of competing sounds. There were six conditions: low-pass mixed and separate, high-pass mixed and separate, and all-pass mixed and separate. All listening was in the free-field conditions used in the earlier experiments.

### Results

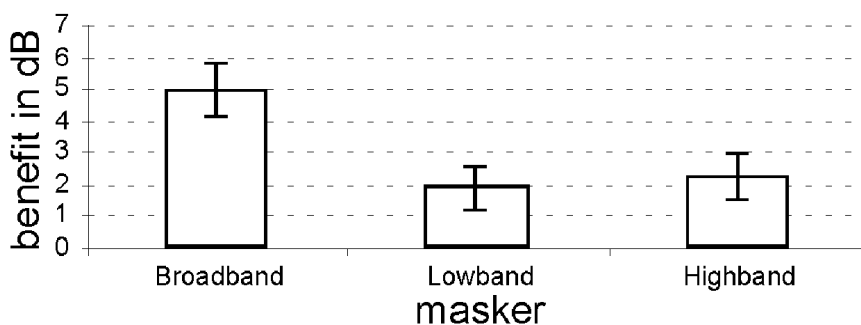
Figure 8 shows the separation effect under the three band-pass conditions. First, the broadband outcome confirmed the result observed in the same condition in the previous experiment. Second, it appears that low- and high-frequency regions of the signal, as presented in the conditions of our laboratory, contributed about equally to the effect. An ANOVA showed a significant main effect for type of signal [ $F(2,22) = 7.11, p = .004$ ], with the Tukey's HSD test showing significantly greater benefit under broadband than under either of the narrower band conditions ( $p < .05$ ) and no difference between the latter two conditions. These outcomes indicate that the listeners picked up both forms of interaural delay in the course of detecting the spatial difference between target and interfering signals. Such a conclusion should be tempered by the point that the signals were filtered to remove energy below 300 Hz. Given the observation from Bronkhorst and Plomp (1988), referred to in connection with the results from Experiment 4, regarding the contribution of time delays at frequencies between 250 and 500 Hz, we might expect dif-

ferent proportional contributions for low- versus high-frequency components for a signal that includes even more low-frequency energy.

## GENERAL DISCUSSION AND CONCLUSION

The starting point for the experiments reported here was the earlier finding of an advantage in hearing target speech at  $0^\circ$ , in free-field listening, under conditions of symmetrical horizontal separation of steady-state broadband noise, as compared with a mixed condition in which speech and noise were from the same ( $0^\circ$ ) locus. The first of the present experiments suggested that the earlier outcome may have been due to the use of correlated noise that generated phase cancellations at the listener's ears in the separated noise condition. In our first experiment, no separation benefit was observed with uncorrelated noise. By contrast with the results from correlated noise, competing speech showed a much clearer beneficial effect of symmetrical separation from target speech, motivating the succeeding experiments whose aim was to account for this outcome.

We found that symmetrically separated speech behaved differently from noise, whether fluctuating or steady state, the general observation being that competing speech more effectively interfered with the target signal in mixed conditions and/or that symmetrical spatial separation more effectively reduced the interference. Noise that alternated between flanking spatial positions provided an intermediate effect between speech and most other forms of noise, an outcome supporting the view that precise momentary cues to spatial distribution underpin some part of the separation effect. Furthermore, when interaural time differences were examined in isolation from other spatial influences (such as the individual head-related transfer function), a substantial part of the speech-versus-speech separation effect was retained. In a later experiment, we observed that both low- and high-frequency components of the signal contributed to that effect, suggesting that both carrier



**Figure 8.** Effects of spatially separating unfiltered (broadband) target speech from competing speech, lowband filtered target from competing lowband speech (2-kHz cutoff), and highband filtered target from competing highband speech. Positive values indicate benefits from such separation (lower target signal level needed to maintain intelligibility).

and envelope onset differences (Zurek, 1993) enabled the parsing of the auditory scene that was observed.

The relative absence of a separation effect for noise and the finding of Festen and Plomp (1990) that reversed speech using the same voice as the target behaves differently (in monaural conditions) from speech by an opposite-sex speaker led to an experiment that tested the effect of similar and dissimilar voices on the separation phenomenon. Partial support was gained for the hypothesis that when target and interfering voices are similar (female–female in this case), a spatial separation effect is observable, whereas the effect does not emerge when the voices are different (female–male). The relative absence of an effect for noise may thus be accounted for in terms of some order of dissimilarity between a target signal that is speech and interference that is not speech. This feature of our results may be further understood in the light of recent reports about the relative contribution of spatial cues to signal segregation and intelligibility, as well as about the relative masking effectiveness of different types of competing signals, particularly speech versus nonspeech.

There have been several studies in which the robustness of interaural difference cues in the organization of auditory attention has been investigated. Darwin and Hukin (2000) showed that when intrinsic cues for segregating voices, such as implied vocal tract length and prosodic continuity are ranged, sometimes singly or in combination, against cues specifying spatial differences between vocal components, the spatial cues are weakened or overridden. These findings are consistent with reports (e.g., Culling & Summerfield, 1995; Darwin & Hukin, 1999) that, as a perceptual grouping cue, signal components having interaural differences in common are not as effective as those having harmonics in common.

The present experimental results suggest, as a corollary, that when sources are distinguishable on other grounds, spatial difference is not attended to; on the other hand, when separate sources are more confusable on other grounds, locational cues feature prominently in how listeners deal with the input signal. Such a proposition can also be argued to follow from results reported by Freyman et al. (1999), who found that a perceived difference in the location of female target and female interfering speech benefited understanding of the target but that the same perceived locational difference had virtually no effect when the interfering signal was steady-state noise.

These experimental observations suggest that the salience of cues to spatial distinctiveness varies, depending on what other components form the structure of the auditory array. We speculate that such variability arises because spatial hearing may not be consistently critical with respect to the particular form of experimental task involved. This task, most typically (in our own and others' experiments), is one of trying to understand a speech signal directly in front in conditions of moderately laterally displaced interference.

Spatial hearing acuity is highest in the frontal horizontal area and becomes coarser in regions lateral to the head (Middlebrooks & Green, 1991). It might be assumed from

this arrangement that frontal discrimination is the most critical role for spatial hearing. Yet, in humans, the frontal area is also under acute visual monitoring, so there are multiple and intersensory cues to the segregation of objects and events in this region (indeed, vision dominates, as is evidenced by the ventriloquism phenomenon; e.g., Warren, Welch, & McCarthy, 1981).

The critical function for spatial hearing may, instead, be understood as alerting an organism, especially one with frontally oriented vision, to the occurrence of objects and events outside the immediate field of view. The evidence from everyday observation and in the laboratory (Perrett & Noble, 1997) is that lateral signals are responded to by immediate rotational movement so as to face the source.<sup>6</sup> A broadly tuned lateral filter combined with a finely tuned frontal one may be just what is required to enable reorientation to occur efficiently. At the point at which the signal of interest is at 0° azimuth, spatial cues will serve to partition the target from otherwise acoustically confusable interference, even if it is close by, because of the spatial acuity of the system in that region. But spatial cues may be subsidiary when a frontal target is distinguishable from competitors on other acoustic grounds.

## REFERENCES

- BREGMAN, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- BRONKHORST, A. W., & PLOMP, R. (1988). The effect of head-induced interaural time and level differences on speech intelligibility in noise. *Journal of the Acoustical Society of America*, **83**, 1508-1516.
- BRONKHORST, A. W., & PLOMP, R. (1992). Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing. *Journal of the Acoustical Society of America*, **92**, 3132-3139.
- COREN, S., & HAKSTIAN, A. R. (1992). The development and cross-validation of a self-report inventory to assess pure-tone threshold sensitivity. *Journal of Speech & Hearing Research*, **35**, 921-928.
- CULLING, J. F., & SUMMERFIELD, Q. (1995). Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay. *Journal of the Acoustical Society of America*, **98**, 785-797.
- DARWIN, C. J., & HUKIN, R. W. (1999). Auditory objects of attention: The role of interaural time differences. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 617-629.
- DARWIN, C. J., & HUKIN, R. W. (2000). Effectiveness of spatial cues, prosody, and talker characteristics in selective attention. *Journal of the Acoustical Society of America*, **107**, 970-977.
- DIRKS, D. D., & BOWER, D. R. (1969). Masking effects of speech competing messages. *Journal of Speech & Hearing Research*, **12**, 229-245.
- FESTEN, J. M., & PLOMP, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *Journal of the Acoustical Society of America*, **88**, 1725-1736.
- FREYMAN, R. L., HELFER, K. S., MCCALL, D. D., & CLIFTON, R. K. (1999). The role of perceived spatial separation in the unmasking of speech. *Journal of the Acoustical Society of America*, **106**, 3578-3588.
- HAWLEY, M. L. (2000). *Speech intelligibility, localization and binaural hearing in listeners with normal and impaired hearing*. Unpublished doctoral dissertation, Boston University.
- HIRSH, I. J. (1950). The relation between localization and intelligibility. *Journal of the Acoustical Society of America*, **22**, 196-200.
- KIDD, G., MASON, C. R., ROHTLA, T. L., & DELIWALA, P. S. (1998). Release from masking due to spatial separation of sources in identification of nonspeech auditory patterns. *Journal of the Acoustical Society of America*, **104**, 422-431.

- KOLLMEIER, B., & WESSELKAMP, M. (1997). Development and evaluation of a German sentence test for objective and subjective speech intelligibility assessment. *Journal of the Acoustical Society of America*, **102**, 2412-2421.
- MASTERTON, B., HEFFNER, H., & RAVIZZA, R. (1969). The evolution of human hearing. *Journal of the Acoustical Society of America*, **45**, 966-985.
- MIDDLEBROOKS, J. C., & GREEN, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology*, **42**, 135-159.
- PEISSIG, J., & KOLLMEIER, B. (1997). Directivity of binaural noise reduction in spatial multiple-source arrangements for normal and impaired listeners. *Journal of the Acoustical Society of America*, **101**, 1660-1670.
- PERRETT, S., & NOBLE, W. (1997). The contribution of head motion cues to localization of low-pass noise. *Perception & Psychophysics*, **59**, 1018-1026.
- SPEAKS, C., PARKER, B., HARRIS, C., & KUHL, P. (1972). Intelligibility of connected discourse. *Journal of Speech & Hearing Research*, **15**, 590-602.
- TER-HORST, K., BYRNE, D., & NOBLE, W. (1993). Ability of hearing-impaired listeners to benefit from separation of speech and noise. *Australian Journal of Audiology*, **15**, 71-84.
- WALKER, G., & BYRNE, D. (1985). Reliability of speech intelligibility estimation for measuring speech reception threshold in quiet and in noise. *Australian Journal of Audiology*, **7**, 23-31.
- WARREN, D. H., WELCH, R. B., & MCCARTHY, T. J. (1981). The role of visual-auditory "compellingness" in the ventriloquism effect: Implications for transitivity among the spatial senses. *Perception & Psychophysics*, **30**, 557-564.
- WIGHTMAN, F. L., & KISTLER, D. J. (1993). Sound localization. In W. A. Yost, A. N. Popper, & R. R. Fay (Eds.), *Human psychophysics* (Springer series in auditory research, Vol. 3, pp. 155-192). New York: Springer-Verlag.
- YOST, W. A., DYE, R. H., JR., & SHEFT, S. (1996). A simulated "cocktail party" with up to three sound sources. *Perception & Psychophysics*, **58**, 1026-1036.
- ZUREK, P. M. (1993). A note on onset effects in binaural hearing. *Journal of the Acoustical Society of America*, **93**, 1200-1201.

#### NOTES

1. This was achievable for all the competitor signals, with the exception of the correlated noise employed in Experiment 1. In the case of correlated noise, two identical wavefronts are being added, which results

in a doubling of *sound pressure*, represented by a 6-dB increase in overall level when the two are presented simultaneously from the same position. When they are presented simultaneously from separate positions, the 6-dB increase is observable only at a point exactly midway between the two: At any other point (e.g., those occupied by a listener's ear canals), there is only a 3-dB increase, owing to the phase cancellation effect noted earlier. Adding two uncorrelated signals increases *sound intensity*, represented by a 3-dB increase that is observable at the midpoint and at the positions occupied by the listeners' ear canals, and whether or not the sounds are from the same locus or different ones. These acoustic differences very likely underlie the performance differences observed for the two sorts of noises in the first experiment.

2. In all but 2 conditions over the first six experiments (24 conditions in total), the difference between the authors' combined results and those of other listeners was less than 1 dB, and the direction of difference was negative in 12 of those conditions. In the seventh experiment, the differences, always positive, ranged from 1.1 to 2.1 dB. We are satisfied that the results obtained are not biased by inclusion of our own data.

3. We note that the different outcome for correlated versus uncorrelated noise was due largely to the lesser attenuation (higher target signal level) required under mixed conditions with correlated noise. This is consistent with the different acoustic effects, described in note 1, of combining identical versus nonidentical noises under conditions of spatial separation versus spatial unity.

4. Our thanks to J. M. Festen for making noise samples available to us based on Festen and Plomp's (1990) procedure.

5. The initial level of the target speech in any one of the experiments in question was determined by its audibility against the range of competing signals used in that particular experiment. That initial target speech level was set as *zero* for the experiment in question; its absolute value was not recorded. To compare attenuation levels across the five experiments, we calculated the differences in average attenuation values for comparable conditions in each (free-field target speech against forward-running competing speech) and applied the relevant difference as a correction factor to all free-field conditions within that experiment.

6. Yost et al. (1996, p. 1035) observed that in a condition in which signals heard via microphones fitted to a stationary manikin were not at 0°, listeners were frustrated by not being able to turn so as to zero in on the signal.

(Manuscript received February 21, 2001;  
revision accepted for publication February 22, 2002.)