Journal of Experimental Psychology:
Human Perception and Performance
2004, Vol. 30, No. 6, 1077–1091

0096-1523/04/$12.00 DOI: 10.1037/0096-1523.30.6.1077

Does the Information Content of an Irrelevant Source Differentially Affect Spoken Word Recognition in Younger and Older Adults?

Liang Li
Peking University and University of Toronto at Mississauga

Meredyth Daneman, James G. Qi, and
Bruce A. Schneider
University of Toronto at Mississauga

To determine whether older adults find it difficult to inhibit the processing of irrelevant speech, the authors asked younger and older adults to listen to and repeat meaningless sentences (e.g., "A rose could paint a fish") when the perceived location of the masker (speech or noise) but not the target was manipulated. Separating the perceived location (but not the physical location) of the masker from the target speech produced a much larger improvement in performance when the masker was *informational* (2 people talking) than when the masker was noise. However, the size of this effect was the same for younger and older adults, suggesting that cognitive-level interference from an irrelevant source was no worse for older adults than it was for younger adults.

This study used a paradigm developed by Freyman, Helfer, McCall, and Clifton (1999) that allowed us to bypass age differences in sensory processing so that we could directly investigate a popular cognitive explanation for the speech processing difficulties of older adults. First, we review the literature, which demonstrates the challenges involved in isolating sensory and cognitive determinants of age-related declines in speech comprehension, and then we describe how we used Freyman et al.'s informational-masking paradigm to assess the relative contributions of losses in inhibitory control (Hasher & Zacks, 1988; Lustig & Hasher, 2001) and age-related declines in hearing (Schneider, Daneman, Murphy, & Kwong See, 2000; Tun, O'Kane, & Wingfield, 2002) to the difficulties that older adults experience when listening to speech.

Sensory Versus Cognitive Factors in Speech Processing

Older adults often report that they have difficulty understanding speech in everyday conversational settings, especially when the environment is noisy and when there is more than one person speaking at a time (Committee on Hearing, Bioacoustics and Biomechanics, 1988; Hamilton–Wentworth District Health Council, 1988). Because they frequently find themselves in situations like this (e.g., family gatherings, mall conversations), older adults are prone to frustration and anxiety, and they may avoid or be excluded from social interactions.

Recent research has provided evidence to substantiate older adults' self-reports about their speech understanding difficulties. In general, the research has shown that older adults with normal or near-normal hearing may have no difficulty perceiving speech in quiet listening conditions, but they do have considerable difficulty when there are interfering stimuli or when they are tested in reverberant environments (Cheesman, Hepburn, Armitage, & Marshall, 1995; Dubno, Dirks, & Morgan, 1984; Duquesnoy, 1983; Frisina & Frisina, 1997; Gelfand, Ross, & Miller, 1988; Gordon-Salant & Fitzgibbons, 1995; Helfer, 1992; Nabelek & Robinson, 1982; Pichora-Fuller, Schneider, & Daneman, 1995; Stuart & Phillips, 1996; Tun & Wingfield, 1999; for reviews, see Pichora-Fuller, 1997; Schneider, Daneman, & Pichora-Fuller, 2002). However, the root of the age-related difficulties is not readily apparent.

Age-related difficulties in understanding speech could arise from several different sources. First, peripheral auditory deterioration (threshold elevations, losses in temporal synchrony, broadening of auditory filters) could degrade the signal available for linguistic and cognitive processing (e.g., Duquesnoy, 1983; Humes & Christopherson, 1991; Humes, Coughlin, & Talley, 1996; Humes & Roberts, 1990; for reviews, see Schneider, 1997; Schneider et al., 2002; Schneider & Pichora-Fuller, 2000, 2001; Willott, 1991). Declines in sensory functioning would place a greater processing load on the linguistic and cognitive systems of older adults. For example, older listeners might need to redeploy some of their limited cognitive resources to recover misheard words and/or phrases from the social and linguistic contexts in which the words and phrases were produced (Pichora-Fuller et al., 1995; Schneider et al., 2002). Hence, they would have fewer resources

---

Liang Li, Department of Psychology, Speech and Hearing Research Center, National Key Laboratory on Machine Perception, Peking University, Beijing, China, and Department of Psychology, Centre for Research on Biological Communication Systems, University of Toronto at Mississauga, Mississauga, Ontario, Canada; Meredyth Daneman, James G. Qi, and Bruce A. Schneider, Department of Psychology, Centre for Research on Biological Communication Systems, University of Toronto at Mississauga.

This work was supported by the Natural Sciences and Engineering Research Council of Canada Grant RGPIN 9952-02 and Canadian Institutes of Health Research Grants MT-15359 and MGC-42665. We thank Jane W. Carey and Alham Chelehmalzadeah for their assistance in data acquisition and illusion construction. We are grateful to Richard Freyman for providing the nonsense sentences and to Kathy Pichora-Fuller for her comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Liang Li, Department of Psychology, University of Toronto at Mississauga, 3359 North Mississauga Road, Mississauga, Ontario L5L 1C6, Canada. E-mail: liang@psych.utoronto.ca

available to integrate this information with past input and world knowledge, to store it in memory, and to formulate intelligent responses. The result would be an apparent loss of comprehension in such situations.

However, comprehension difficulties could be due to age-related changes in cognitive processing. One such candidate is age-related slowing (Cerella, 1990; Lindenberger & Baltes, 1994; Salthouse, 1985, 1993, 1996). According to this theory, slowing in brain functioning is thought to reduce the speed with which various cognitive operations can be executed and to produce age differences in performance on any task that requires a large number of cognitive operations in a short time. Thus, age-related cognitive slowing could account for why older adults might find it difficult to follow a conversation when the rate of speech is fast and when there are multiple speakers.

Another possible cognitive source for the age-related language comprehension difficulties is that older adults may find it difficult to inhibit the processing of irrelevant stimuli. It has been proposed that normal aging is associated with reduced inhibitory mechanisms for suppressing the activation of goal-irrelevant information (Hasher & Zacks, 1988; Hasher, Zacks, & May, 1999), allowing interfering signals to intrude into working memory (Daneman & Carpenter, 1980). The irrelevant signals squander central resources and disrupt the cognitive processing of goal-relevant information. Tun and Wingfield (1999) and Tun et al. (2002) have argued that older adults may find hearing in noisy backgrounds to be difficult not only because of auditory declines but also because they cannot inhibit the processing of irrelevant speech efficiently.

A prominent feature of the failure-of-inhibition theory is that the greater the similarity between target and distractors, the more difficult it becomes to inhibit the processing of irrelevant stimuli. In the context of listening comprehension, age differences should be larger when the distracting or masking stimulus is speech than when the masking stimulus is noise (Lustig & Hasher, 2001). The results of a number of studies support the notion that older adults may be less able to inhibit the processing of irrelevant stimuli when the task is listening to speech in a distracting setting. Duquesnoy (1983) reported that elderly listeners with hearing loss are unable to make full use of the spatial separation between target sentences and interfering sources. Dubno, Ahlstrom, and Horwitz (2002) showed that younger listeners have better sentence recognition in noise and higher spatial-separation benefits than do elderly listeners with normal hearing. However, the results from other studies suggest that elderly listeners with clinically normal hearing can make use of the spatial separation between a target sentence and a masker to improve speech recognition just as well as younger listeners (e.g., Gelfand et al., 1988). Given the inconsistent findings, it is not clear whether older listeners have greater difficulty inhibiting the processing of irrelevant signals when listening to speech than do their younger counterparts.

One of the reasons for the lack of consistency in the research findings is that irrelevant speech may not only introduce cognitive difficulties into the listening situation—it may peripherally mask the targeted speech as well. It may be particularly difficult to evaluate the relative contribution of cognitive-level effects to speech-processing declines when the presentation of irrelevant material simultaneously acts as a peripheral mask, because in such situations, both sensory (information-degradation hypothesis; Pichora-Fuller et al., 1995; Schneider et al., 2002) and cognitive factors (e.g., deficit in inhibitory control, speed of processing) could be responsible for the speech-processing declines.

How then does one determine the relative contributions of sensory-level effects and cognitive-level effects to the difficulties that older adults experience when listening to a conversation in a background of competing conversations? One way is to adjust the listening situation to make it equally difficult for younger and older adults to hear individual words when there is no contextual support to aid in the identification of these words. This is usually accomplished by finding the signal-to-noise ratio for each individual listener that produces equivalent word-recognition scores across the group. Meaningful material is then presented at these individually determined signal-to-noise ratios. If the information-degradation hypothesis is correct, then age-related differences in comprehension should be minimized. If, however, cognitive deficits are primarily responsible for age-related differences in comprehension, then these age-related comprehension differences should persist even when it is equally difficult for all participants to hear individual words. Support for the information-degradation hypothesis has been provided by several studies that have equated for perceptual difficulty in this manner (Murphy, Craik, Li, & Schneider, 2000; Schneider et al., 2000).

A second possible way to distinguish between sensory- and cognitive-level effects on comprehension is to manipulate the content of the competing signal. The hypothesis of a deficit in inhibitory control predicts that the difficulty in inhibiting irrelevant alternatives should increase as the similarity between the target and the irrelevant alternatives increases and that the rate of increase in difficulty should be greater in older adults (Lustig & Hasher, 2001). Thus, the interference in attending to the target talker should be greater when the irrelevant stimulus is another talker than when the irrelevant stimulus is speech-spectrum noise. If older adults are suffering from a deficit in inhibitory processes, then the decrease in performance that they experience when going from random noise as background to a talker as background should be greater than the decrease experienced by younger adults.

Of course, to rule out any contribution of age-related sensory degradation in this situation, one would also need to show that peripheral masking effects could not account for any differential age effects observed with the switch from a nonspeech noise background to a talker as background. For example, peripheral masking could be more severe in older adults than in younger adults when the masker is speech than when the masker is speech-spectrum noise. The envelope of a speech waveform has peaks corresponding to voiced segments and troughs corresponding to unvoiced segments and pauses between phrases and sentences. Troughs in the masker provide a brief opportunity for the listener to process the speech signal in the absence of a peripheral mask. Younger adults may be better able to profit from these troughs than are older adults, either because they are less subject than older adults to forward and backward masking or because temporal processing abilities are reduced with age. Hence, if older adults were more severely affected than younger adults by the switch from a noise masker to a speech masker, one could not be sure whether this was due to peripheral masking factors or to an inability, at the linguistic and cognitive level, to inhibit irrelevant alternatives evoked by the speech masker. To distinguish between a sensory and a cognitive interpretation in situations such as these, one needs to be able to either (a) equate younger and older listeners with respect to sensory factors (something that is extremely difficult to do) or (b) find a way to bypass age-related changes in sensory processing.

Recently, Freyman et al. (1999) found a clever and convenient way of bypassing differences in peripheral sensory processing when studying the differential effects of noise versus speech as a masker. They accomplished this by manipulating the perceived spatial separation between the target speech and a masker without introducing any significant changes in peripheral auditory cues (see the Spatial Separation and Peripheral Masking section below). Because the peripheral cues did not change with perceived location, any age differences in word recognition due to the type of masker and its perceived spatial position would presumably reflect age-related differences in cognitive processing. To see how Freyman et al. (1999) accomplished this, it is necessary to examine how peripheral acoustic cues associated with spatial separation could lead to a release from masking.

Spatial Separation and Peripheral Masking

Spatially separating the target and masker improves recognition of the target (Cherry, 1953; Hirsh, 1950; for a review, see Zurek, 1993). For example, thresholds for detecting targets in a noise background are much lower when target and masker are spatially separated (Arbogast, Mason, & Kidd, 2002; Dubno et al., 2002; Duquesnoy, 1983; Freyman et al., 1999; Gelfand et al., 1988). However, acoustical factors primarily account for this spatial-separation advantage. First, when, for example, the masker is to the left and the target is directly ahead, the listener's head reduces the intensity of the masker in the right ear, especially in the high-frequency region, thereby improving the signal-to-noise ratio in the right ear relative to when both masker and signal are frontally located. In other words, the sound shadow cast by the head improves the signal-to-noise ratio. Second, when the masker is to the left, the spatial separation between the ears introduces an interaural time difference for the masker but not for a frontally presented target. Interaural time differences between signal and masker (especially in the low-frequency region) allow the auditory system to unmask the signal (Bronkhorst & Plomp, 1988; Zurek, 1993). Hence, any age-related differences in sensitivity to signal-to-noise ratio or to interaural time differences (Grose, 1996; Pichora-Fuller & Schneider, 1992, 1998) would lead to a smaller spatial-separation effect in older adults than in younger adults.

There are ways, however, to reduce the contributions of these acoustic cues to the spatial-separation effect. As Freyman et al. (1999) and Koehnke and Besing (1996) have pointed out, presenting the target and masker in a highly reverberant environment significantly reduces the head-shadow advantage and obscures interaural time differences, thereby significantly reducing the spatial-separation effect. To further reduce the contribution of peripheral factors—such as signal-to-noise ratio and interaural time differences—to the spatial location effect, Freyman et al. (1999) manipulated the apparent location of a signal via the precedence effect.

When a sound source is produced in a reverberant environment, listeners not only receive the direct wavefront of the sound source but also numerous time-delayed reflections from walls, ceilings, and other surfaces. If the time delay between the arrival of the direct wave and each of the reflected waves is sufficiently short (e.g., 1–5 ms), listeners typically perceive the direct wave and its reflections as belonging to a single auditory event located at or near the point of origin of the direct wavefront. This phenomenon is generally known as the *precedence effect* (Wallach, Newman, & Rosenzweig, 1949; for reviews, see Blauert, 1997; Li & Yue, 2002; Litovsky, Colburn, Yost, & Guzman, 1999; Zurek, 1980).

In Freyman et al.'s (1999) experiment, a loudspeaker directly in front of listeners and a loudspeaker in the right hemifield delivered both target stimuli (nonsense sentences) and masking stimuli. For target sentences, the frontal loudspeaker always led the right speaker by 4 ms, a delay that was short enough that the leading and lagging signals were perceived as fused. Hence, the target sentence was perceived as having a frontal location (T_F). For the masking stimuli, the frontal loudspeaker either led or lagged behind the right speaker by 4 ms. When the frontal loudspeaker led the right loudspeaker, the masker was perceived as having a frontal location (M_F). Thus, when both masker and target were played over both loudspeakers with the same lag time, both were perceived as emanating from the front (T_FM_F). However, when the frontal loudspeaker presenting the masker lagged behind the right loudspeaker, the masker was perceived as coming from the right. This created a condition in which the perceived location of the target was to the front, but the perceived location of the masker was to the right (T_FM_R). In other words, a spatial separation between target and masker was perceived even though both target and masker were being physically presented from both loudspeakers.

An interesting feature of creating a perceived spatial separation in this manner is that, unlike in conditions in which the two sources are physically separated, there is no significant signal-to-noise ratio advantage at either ear, and there are no significant interaural time differences to support a spatial-separation effect. To see why this is the case, note that because the target does not change between conditions, neither the sound pressure level of the target in either ear nor the interaural time differences associated with the target change between T_FM_F and $T_FM_{R.}$ Moreover, although the perceived location of the masker shifts depending on which loudspeaker is leading the other, the sound pressure level and spectral characteristics of the masker at each ear remain essentially the same. There are a number of reasons for this. First, the degree of attenuation due to head shadowing remains the same for both M_F and M_R because the masker is being presented from the central and right loudspeakers in both cases. Hence, head shadowing will not produce a difference in signal-to-noise ratio between Conditions M_F and M_R. Second, the spectral profiles of the masker in each ear, which depend on the head-shadowing effect and the time delay between sounds from the left and right loudspeakers, do not differ in a significant way. Finally, the interaural cues associated with the masker do not differ substantially. Essentially, the only feature that changes across Conditions M_F and M_R is whether the frontal loudspeaker leads (M_F) or lags behind (M_R) the right loudspeaker for the masking stimulus (for further discussion, see the Appendix).

When Freyman et al. (1999) manipulated the perceived location of the masker, they found a larger advantage (4–9 dB) for spatial separation when the masker was one person talking than when the masker was speech-spectrum noise ($<$1 dB). Because the acoustic conditions (signal-to-noise ratio, interaural cues, etc.) did not change substantially with a change from speech to noise masking, Freyman et al. argued that the greater spatial-separation effect for the speech masker could not be attributed to differences in the acoustic conditions prevailing at the level of the ear for the two types of maskers. In other words, the release from masking arising from the perceived spatial separation of a speech target from a speech masker was not due to peripheral (i.e., energetic) factors. Hence, Freyman et al. referred to this effect as a release from informational masking and assumed that it occurs at a level remote from the auditory periphery—that is, at a cognitive level.

If the release from an informational masker occurs at a cognitive level, what cognitive mechanism or mechanisms could be responsible for its occurrence? A feature of many models of spoken word recognition is that the presentation of spoken words leads to automatic and obligatory activation of word meanings (e.g., Marslen-Wilsen, 1990; McClelland & Elman, 1986). This implies that when both target and masker are speech, the masker as well as the target could be initiating activity in the linguistic and cognitive systems responsible for speech processing. Hence, when the task is to report only the target words, it may become necessary to suppress activity in pathways activated by the masker. In other words, it becomes advantageous to inhibit these irrelevant alternatives (Hasher & Zacks, 1988; Hasher et al., 1999; Lustig & Hasher, 2001).

A corollary of this theory is that it should be easier to inhibit these irrelevant alternatives the more dissimilar they are from the target stimulus (Tun et al., 2002; Tun & Wingfield, 1999). Hence, a competing coherent conversation between two individuals in the listener's own language should prove to be more difficult to suppress than a babble of voices in an unintelligible language or speech-spectrum noise. In addition, the greater the separation between the target voice and its competitors, the easier it should be to separate relevant information from competing information. Presumably, a perceived spatial separation between target and masker would make it easier to identify the irrelevant activation and, therefore, make it easier to suppress it (Bregman, 1990).

It is important to note that one would not necessarily expect a spatial-separation effect when the overlap in semantic content between target and masker was negligible. If the target were speech and the masker speech-spectrum noise, it is unlikely that the noise would activate connections and pathways in whatever circuits are responsible for word identification. Hence, one would not expect much of a cognitive benefit from separating masker and target, providing that the peripheral acoustic properties of the two types of masker were similar. However, if both target and masker were speech, the masker as well as the target could initiate activity in the linguistic and cognitive systems responsible for speech processing. Here, one would expect that any manipulation (e.g., spatial separation) that perceptually distinguishes target from masker (Bregman, 1990) might make it much easier to inhibit the activity elicited by the speech masker. In other words, one should find a greater effect of spatial separation the greater the similarity between masker and target. The results of Freyman et al. (1999) are consistent with the notion that perceived spatial separation provides a greater release from masking when the masker is speech than when it is speech-spectrum noise. Moreover, the Freyman et al. results provide a way of testing this hypothesis without the need to consider possible peripheral masking effects, because, as we have already shown, separating the perceived location of the masker from the perceived location of the target does not introduce, in any significant way, peripheral cues (signal-to-noise ratio increases, interaural differences) that could contribute to a release from peripheral masking. In other words, these results provide a way of bypassing age-related differences in peripheral mechanisms and directly testing cognitive processing.

In the present study, we used the fusion phenomenon of the precedence effect to induce perceived spatial separation of target sentences from maskers for both older and younger adults to test the hypothesis that older adults have a deficit in inhibitory processes. According to that hypothesis, older adults should have more difficulty inhibiting the irrelevant masker, particularly when the masker and target are both speech. Moreover, if older adults find it more difficult than younger adults to inhibit irrelevant alternatives, they might also derive less of a benefit from perceived spatial separation than do younger adults.

Experiment

Method

Participants. Twelve university students (19–22 years old) and 12 older adults recruited from the city of Mississauga, Ontario, Canada (63–75 years old) participated in the study. All participants had normal and balanced (less than 15-dB difference between the two ears) hearing thresholds in the speech range (250–3000 Hz). Their first language was English. Average hearing thresholds as a function of frequency for the two age groups are shown in Figure 1.

As Figure 1 shows, the thresholds of older adults over the speech range were generally higher than were those of younger adults. Beginning at about 3 kHz, however, threshold differences between younger and older adults started to increase. Hence, even though these older adults were considered to have good hearing, they were best characterized as being in the early stages of presbycusis and, therefore, were likely to be experiencing subclinical declines in a number of auditory functions.

Apparatus and materials. Listeners were seated in a chair placed in the center of an Industrial Acoustic Company (Bronx, NY) sound-attenuated chamber during testing. All acoustic stimuli were digitized at the sampling rate of 20 kHz using a 16-bit Tucker Davis Technologies (TDT; Gainesville, FL) System II hardware (DD1) and custom software. Stimuli were converted to analog forms using TDT DD1 digital-to-analog converters under the control of a Dell computer with a Pentium processor. The analog outputs were low-pass filtered at 10 kHz, attenuated by two programmable attenuators (PA4, for the left and right channels), amplified via a Technics power amplifier (SA-DX950), and then delivered from two balanced loudspeakers (40 watts; Electro-Medical Instrument [Mississauga, Ontario, Canada]). The loudspeakers were placed at angles of 45° to the left and right of the listener. The distance between a loudspeaker and the center of the listener's head was 1.03 m. The loudspeaker height was approximately ear level for a seated listener with average body height.

Target speech stimuli were 312 English nonsense sentences spoken by a female talker (Talker A). These sentences, which were developed by Helfer (1997) and used in experiments by Freyman et al. (1999), are syntactically correct but not semantically meaningful. In each target sentence (e.g., "His *inn* will *betray* the *foot*"; "The *goose* can *kick* a *street*"), there are three key words (italicized in the examples). Because these sentences are meaningless, listeners could not use contextual cues to identify the words. Target sentences were presented over both the right and the left loudspeakers, with the right speaker leading the left speaker by 3 ms. Thus, listeners perceived the target sentences as originating on the right side (see Figure 2). Because previous studies (e.g., Schneider, Pichora-Fuller, Kowalchuk, & Lamb, 1994) have failed to find age-related differences in the precedence effect, we are reasonably certain that the three delays used here (i.e., left leading right by 3 ms, simultaneous, and right leading left by 3 ms) had the same effect on perceived location in both younger and older adults.

There were two types of masking stimuli: noise and speech. The noise masking sound was steady speech-spectrum noise that was recorded from

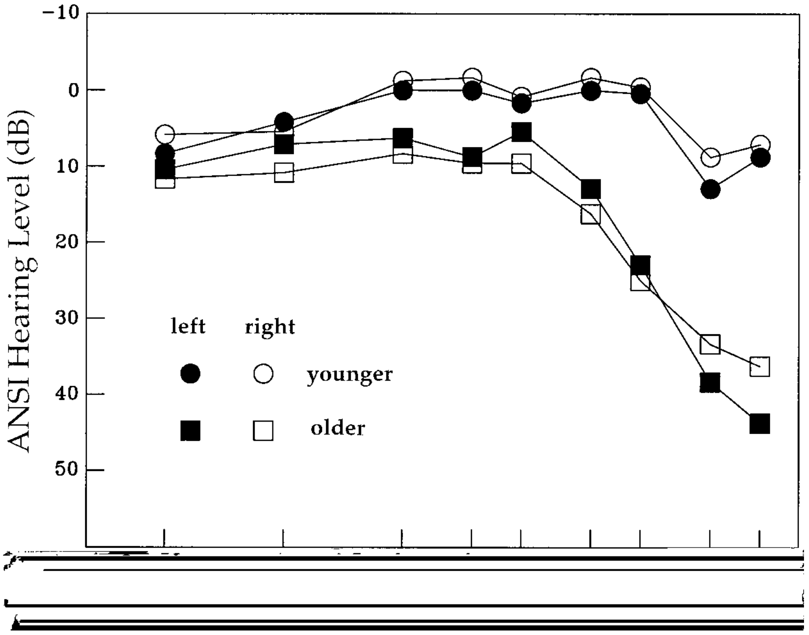

Figure 1. Average hearing thresholds in left and right ears for the two age groups. ANSI = American National Standards Institute.

an audiometer (Interacoustic [Assens, Denmark], Model AC5). The speech masker was a different set of linguistically correct but semantically meaningless sentences spoken by two female talkers, whose waveforms were mixed with equal root-mean-square levels from the two sources (see Freyman, Balakrishnan, & Helfer, 2001). These masker sentences were repeated in a continuous loop.

Targets and maskers were calibrated using a Brüel & Kjær (Copenhagen, Denmark) sound-level meter (Type 1616). A microphone was placed at the location usually occupied by the listener's head, and the reading was taken using the slow-norm scale. Measurements were conducted separately for each loudspeaker. During a session, the target sentences were presented at a level such that each loudspeaker, playing alone, would produce an average sound pressure of 60 dBA at the location corresponding to the center of the listener's head. The sound pressure level of the target remained constant throughout the experiment. The sound pressure levels of the masker were adjusted to produce four signal-to-noise ratios: −12, −8, −4, and 0 dB.

Procedure. There were 12 listeners in each of the two age groups. Six listeners in each group listened to sentences masked by noise in the first testing session and to a different set of sentences masked by speech in the second session. The remaining six listeners experienced these two sessions in the opposite order.

As shown in Figure 2, the masker was presented over the two loudspeakers using one of three delay times: (a) right leading left by 3 ms, (b) no lag between the loudspeakers, or (c) right lagging behind left by 3 ms. For right–left delays of +3, 0, and −3 ms, listeners heard the masker as originating from right, center, and left, respectively.

Twenty-four blocks of 13 sentences each were created for all possible combinations of the three masker delays, four signal-to-noise ratios, and two types of maskers (speech-spectrum noise and nonsense sentences). The order of presentation of the different perceived locations of the masker was completely counterbalanced across listeners, with each experiencing the four signal-to-noise ratios in a different random order. Hence, the type of masker, its level, and its location remained constant during each 13-trial block.

On each trial, the listener pressed the central button of a response box to start the masking sound, which was delivered by the two loudspeakers. About 1 s later, a single target sentence was automatically presented from the two loudspeakers. The masker was gated off with the target. The listener was instructed to repeat the target sentence as best as he or she could immediately after the stimuli ended. Tape recordings were made of all sessions, and each listener's performance was scored offline.

Results

A logistic psychometric function,

$$y = 1/[1 + e^{-\sigma(x-\mu)}],$$

was fit to each listener's data using the Levenberg–Marquardt method (Wolfram, 1991), where y is the probability of correctly identifying a key word, x is the signal-to-noise ratio, μ is the signal-to-noise ratio corresponding to 50% correct identification (the threshold value), and σ determines the slope of the psychometric function. Figure 3 shows percentages of correct identification of key words as a function of signal-to-noise ratio for younger and older listeners when the masker was speech. Figure 4 plots the data for the condition in which the masker was speech-spectrum noise in the same way. Psychometric functions were fit to the percentage of correctly identified items (out of 39: 3 words in 13 targets) at each of four signal-to-noise ratios. Figures 3 and 4 show that the percentage of words correctly identified increased with signal-to-noise ratio in all conditions and that the psychometric functions provide a reasonably good description of the results. Hence, we explored the effects of age, type of masker, and perceived location of the masker on the two parameters of the psychometric function: (a) the threshold (the value of signal-to-noise ratio corresponding to 50% correct) and (b) the slope (which reflected how rapidly performance increased with signal-to-noise ratio).

A 2 (masker) × 2 (age) × 3 (perceived location) analysis of

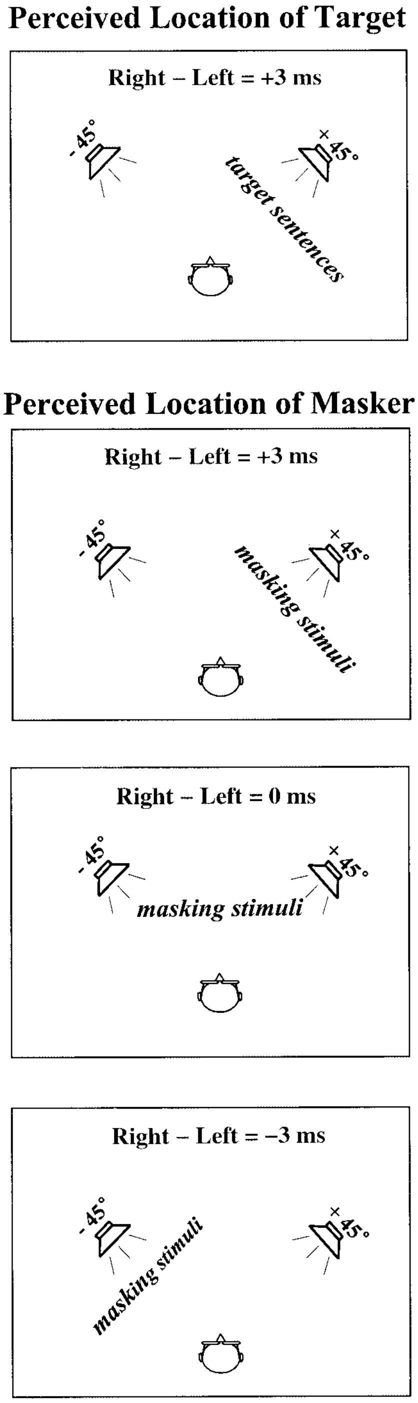

Figure 2. Diagrams showing the perceived locations of target speech and masking stimuli in the different experimental conditions. In all conditions, the perceived location of the target was on the right because the right loudspeaker led the left loudspeaker by 3 ms. There were three perceived locations for the masking stimuli: (a) right (right − left = +3 ms), (b) central (right − left = 0 ms), and (c) left (right − left = −3 ms).

variance (ANOVA) on individual thresholds found a main effect of age, *F*(1, 22) = 22.049, *MSE* = 14.715, $p < .001$. On average, older adults required a higher signal-to-noise ratio for 50% accuracy. However, none of the two- or three-way interactions with age were significant: Age × Masker, *F*(1, 22) = 1.073, *MSE* = 2.796, *p* = .311; Age × Perceived Location, *F*(2, 44) = .050, *MSE* = 4.329, *p* = .95; Age × Masker × Location, *F*(2, 44) = .045, *MSE* = 2.438, *p* = .956. Hence, the threshold difference between younger and older adults did not change with the type of masker or with the perceived location of the masker.

There was a significant main effect of masker type on thresholds, *F*(1, 22) = 25.210, *MSE* = 2.796, $p < .001$, a significant main effect of perceived location, *F*(2, 44) = 39.990, *MSE* = 4.329, $p < .001$, and a significant Masker × Perceived Location interaction, *F*(2, 44) = 14.477, *MSE* = 2.438, $p < .001$. For the noise masker, there appeared to be a small release from masking when the perceived location of the masker differed from that of the target. However, the release from masking was much larger for both age groups when the masker was speech. Finally, for both the noise and the speech masker, it did not seem to make any difference whether the perceived location of the masker was in the central position or further to the left as long as its perceived location was different from that of the target.

Although age had a significant effect on thresholds, an equivalent ANOVA on the individual slopes found that the slope of the psychometric function did not change with age, *F*(1, 22) = 0.772, *MSE* = 0.022, *p* = .389. Moreover, none of the two- or three-way interactions with age were significant: Age × Masker, *F*(1, 22) = 0.470, *MSE* = 0.010, *p* = .501; Age × Location, *F*(2, 44) = 0.269, *MSE* = 0.012, *p* = .765; Age × Masker × Location, *F*(2, 44) = 0.085, *MSE* = 0.015, *p* = .918. Hence, the slopes of younger and older adults were the same in all conditions.

The type of masker, however, did significantly affect the slope of the psychometric function, *F*(1, 22) = 31.172, *MSE* = .010, $p < .001$, as did the location of the masker, *F*(2, 44) = 4.780, *MSE* = 0.012, *p* = .013, but there was no Masker Type × Location interaction, *F*(2, 44) = 1.360, *MSE* = 0.015, *p* = .267. The slopes of the psychometric functions were steeper for noise maskers. In addition, slopes were steeper for both noise and speech maskers when the perceived location of the masker was the same as that of the target.

This pattern of results can be seen more clearly when the data are averaged across listeners. Figure 5 shows the average psychometric functions for younger and older listeners when the perceived location of the speech masker was on the left, center, and right. Note that the psychometric functions for younger and older listeners are equivalent in all conditions if the psychometric functions of the older listeners are shifted to the left by 2.8 dB (as is the case in Figure 5). In other words, the only way in which younger and older listeners differed was that older listeners needed a higher signal-to-noise ratio than did younger listeners to achieve the younger listeners' levels of performance. Figure 5 also indicates that the target was much easier to detect when the perceived location of the masker differed from that of the target. Moreover, the slope of the psychometric function was steeper when target and masker were perceived to be coming from the same position. Finally, there is no evidence of any differences in performance for the two locations (left and center) when the perceived location of the masker differed from that of the target.

Figure 6 plots the equivalent data for the noise masker. The psychometric functions for the older listeners have been shifted to the left by the same amount as in Figure 5. Hence, independent of

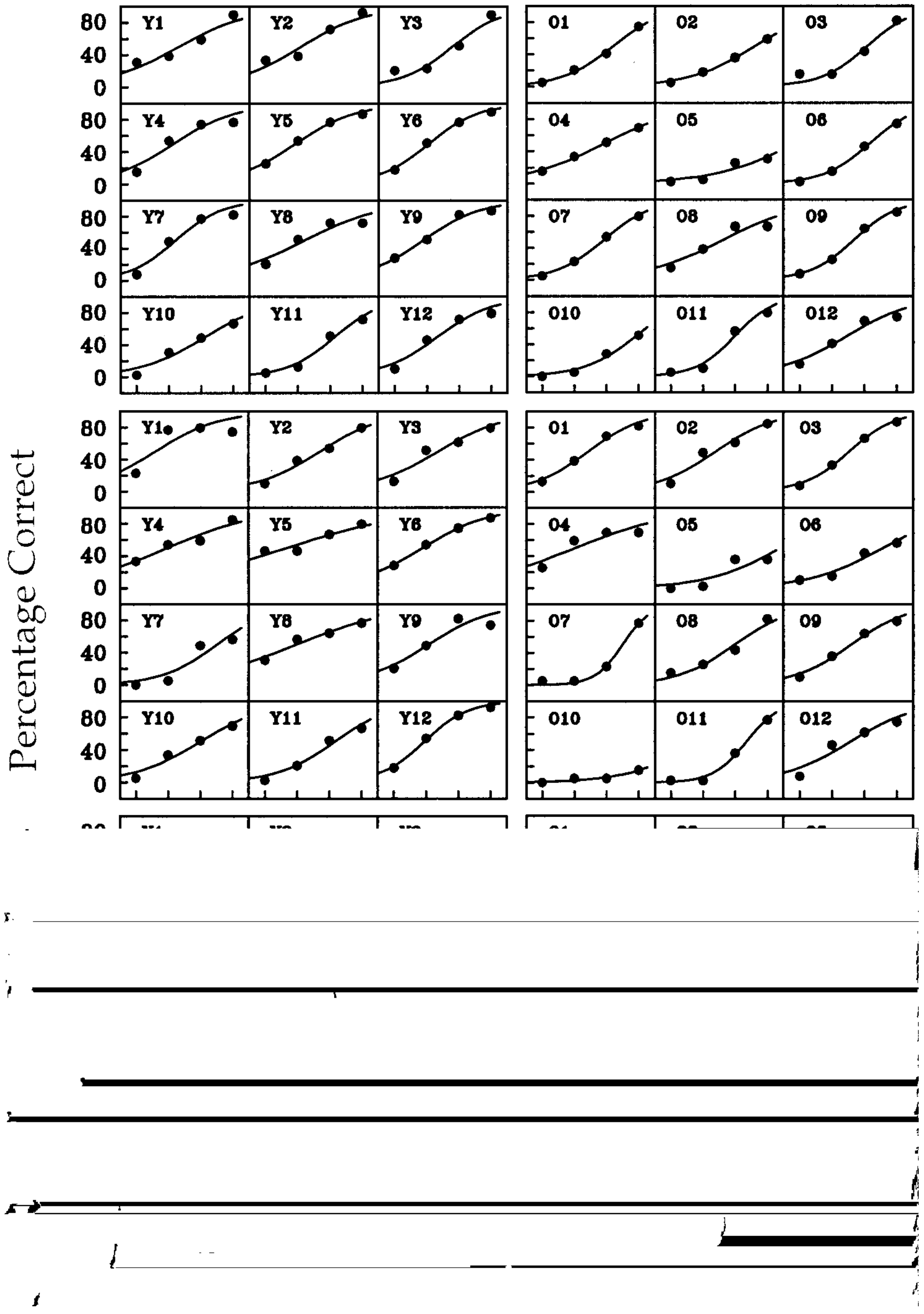

Figure 3. Percentages of correct word identification as a function of signal-to-noise ratio for the 12 younger (Y) and 12 older (O) listeners when the masker was speech. The top panels present the data for the condition in which the perceived location of the masker was on the left. The middle panels present the data for the condition in which the perceived location of the masker was in the center. The bottom panels present the data for the condition in which the perceived location of the masker was on the right.

the type of masker, older listeners needed the same increment in signal-to-noise ratio to perform equivalently to younger listeners. A comparison of the slopes in Figures 5 and 6 also indicates that the slopes are steeper for noise as a masker than for speech as a masker. Figure 6 indicates that there was a slight improvement in performance (lower thresholds) when the perceived location of the noise masker differed from that of the target. However, this improvement was much smaller than when the masker was speech (see Figure 5). Again, as was the case for the speech masker, slopes were steeper when target and masker shared the same

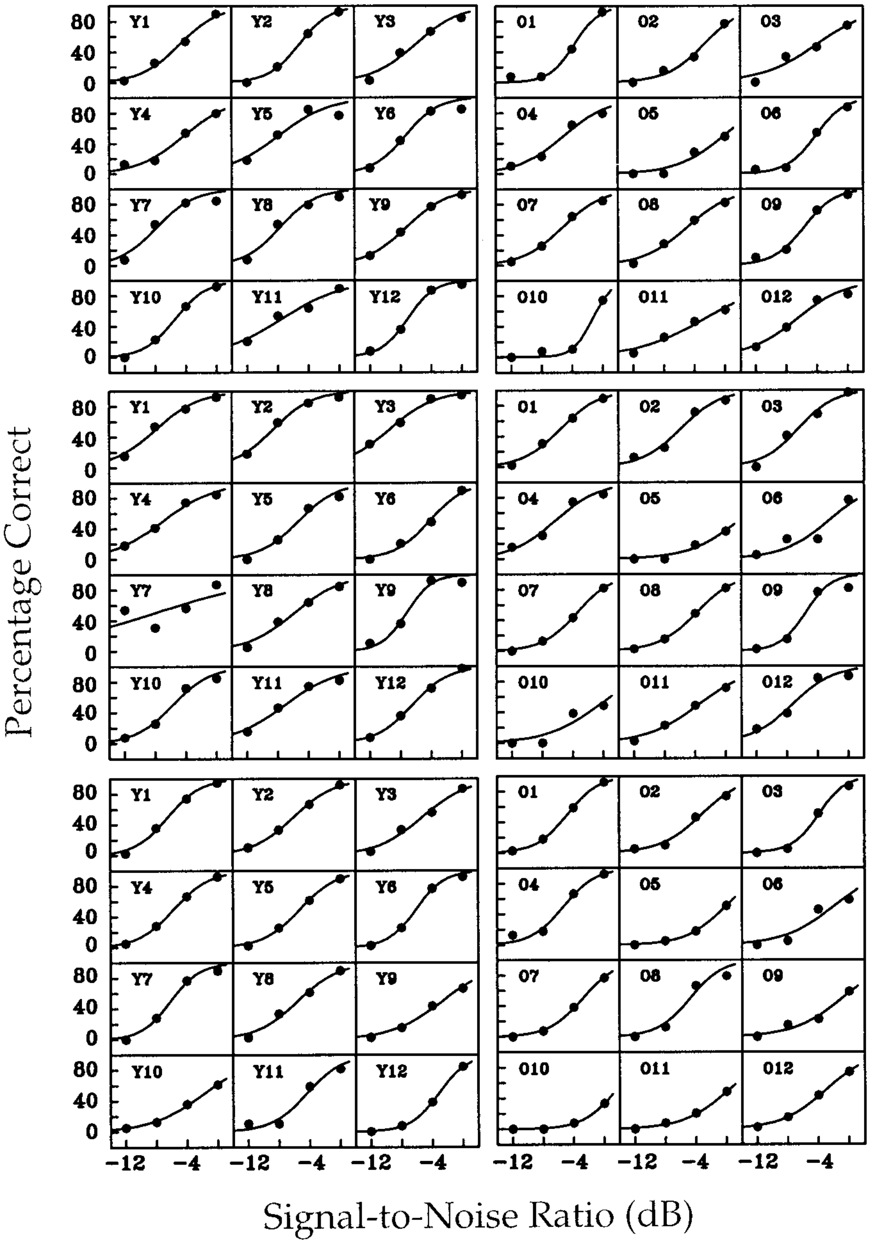

Figure 4. Percentages of correct word identification as a function of signal-to-noise ratio for the 12 younger (Y) and 12 older (O) listeners when the masker was noise. The top panels present the data for the condition in which the perceived location of the masker was on the left. The middle panels present the data for the condition in which the perceived location of the masker was in the center. The bottom panels present the data for the condition in which the perceived location of the masker was on the right.

perceived location, and there is no indication that performance differed between the perceived left and central locations.

Discussion

The overall picture that emerges from these data is quite simple. First, separating the perceived location of the target from that of the masker provides a greater release from masking when the masker is speech than when the masker is noise. Second, the slopes of the psychometric functions are steeper for a noise masker than for a speech masker. Third, for both noise and speech maskers, the slopes of the psychometric functions are steeper when there is no perceived separation between target and masker than when there is a perceived separation. Fourth, performance is equivalent for the two perceived locations of the masker that are different from that of the target. Finally, the only difference between younger and older listeners is that older listeners need a higher signal-to-noise ratio to reach the same performance level as that of younger listeners, regardless of the masker type or the perceived location of the masker.

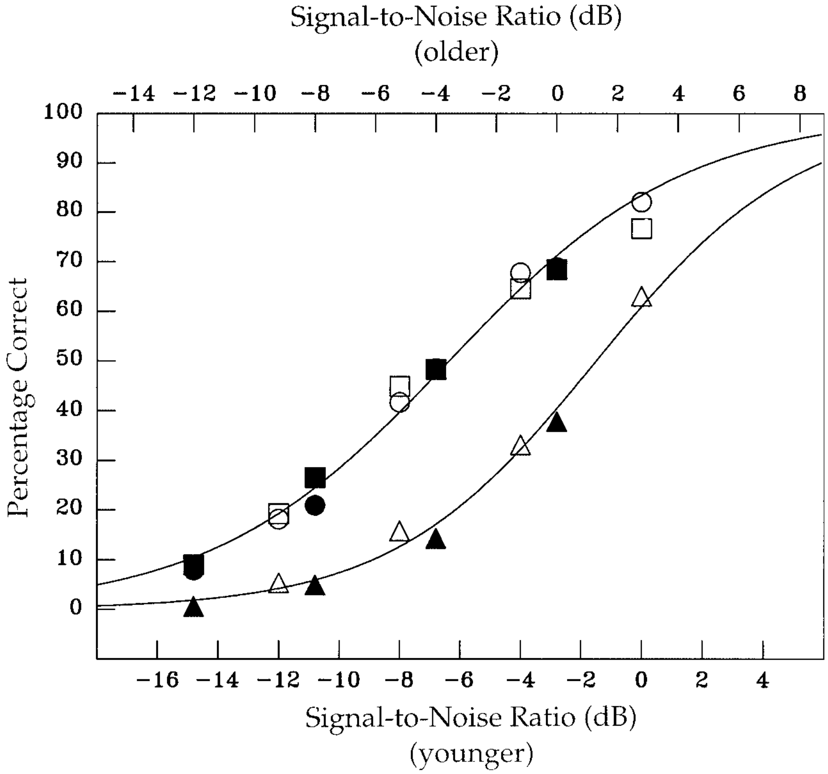

Figure 5. Mean percentages of correct responses as a function of signal-to-noise ratio for the three perceived positions of the speech masker: left (circles), center (squares), and right (triangles). Filled and unfilled symbols represent older and younger adults, respectively. Note that the signal-to-noise ratios for the older listeners have been shifted 2.8 dB to the left of those for the younger listeners. Consistent with the results of the analyses of variance, which found no difference in performance between left and center positions, a single psychometric function has been fit to the data for these two positions. A separate psychometric function was fit to the data for the right position.

The first two results essentially replicate those of Freyman et al. (1999), who found that when the masker was speech, thresholds for word recognition improved by 4–9 dB when the perceived location of the masker was shifted away from that of the target. In the present study, the average improvement in thresholds when the perceived location of the speech masker was shifted away from that of the target was 4.8 dB. When the masker was noise, Freyman et al. found a small release from masking (<1 dB) when the perceived location of the masker was shifted away from that of the target. In the present study, the average improvement in threshold when the perceived location of the noise masker was shifted away from that of the target was 1.7 dB.

The small release from a speech-spectrum noise masker that was observed could have been due to interaural time differences (in the low-frequency region) between the target speech and the noise masker. Freyman et al. (1999) found lower thresholds for the detection of low-frequency, ⅓-octave bands of noise in their condition in which the target was perceived frontally and the masker laterally than in the condition in which target and masker were both perceived frontally. Therefore, one might expect to find a release from energetic masking due to perceived spatial separation. Using a model based on the Articulation Index (Kryter, 1962), Freyman et al. (1999) estimated that the amount of release from energetic masking should be about 2.0 dB, which is close to the 1.7 dB found in this experiment.

The comparability of results across the two studies, despite three potentially significant differences in testing conditions, reinforces the argument that these unmasking effects cannot be attributed to peripheral acoustic factors. First, the Freyman et al. (1999) study was conducted in an anechoic environment, whereas ours was not. Hence, in our test situation, the signal at each ear consisted of the direct wavefronts from the two loudspeakers plus numerous, but highly attenuated, reflections from the walls, floor, and ceiling. In an anechoic environment, each ear receives only the direct wavefronts. If peripheral acoustic factors are responsible for the release of masking that is found when the perceived location of the speech masker is shifted away from that of the target, then one might expect to find differences between anechoic and echoic situations. Reverberant environments tend to minimize the head-shadowing effects and interaural time differences (see discussions in Freyman et al., 1999; Koehnke & Besing, 1996). Hence, if the Freyman et al. results were due to peripheral acoustic cues, the effectiveness of these cues should have been reduced in our situation, leading to a reduction in the size of the effect. That the effect sizes were comparable suggests that the contribution of peripheral acoustic factors to release from masking is negligible.

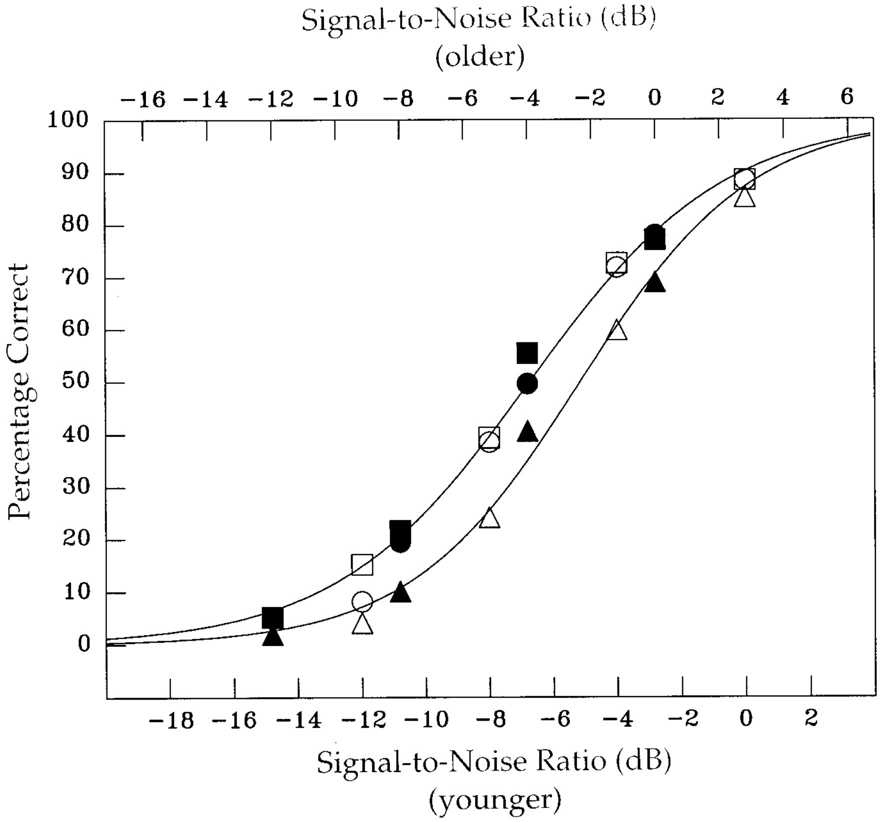

Figure 6. Mean percentages of correct responses as a function of signal-to-noise ratio for the three perceived positions of the noise masker: left (circles), center (squares), and right (triangles). Filled and unfilled symbols represent older and younger adults, respectively. Note that the signal-to-noise ratios for the older listeners have been shifted 2.8 dB to the left of those for the younger listeners. Consistent with the results of the analyses of variance, which found no difference in performance between left and center positions, a single psychometric function has been fit to the data for these two positions. A separate psychometric function was fit to the data for the right position.

The second difference between the two situations is that Freyman et al. (1999) had only one perceived separation condition between masker and target, whereas we had two. Specifically, in Freyman et al.'s spatial-separation condition, the right loudspeaker presenting masking led the center loudspeaker, so the masker was perceived on the right. When there is a lag, the delay in the time of arrival of the masker from the two loudspeakers introduces a modulation in the spectrum of the masker (comb filtering) that is absent when there is no delay. In our study, when the masker was on the left, we had a similar delay condition for spatial separation between the target (perceived on right) and the masker, and this produced comb filtering as well. However, we also had a spatial-separation condition in which the target was on the right and the masker was central (no lag between the loudspeakers). The peripheral acoustic signal when there is no lag is quite different from the signal when there is a lag (see the Appendix). Also, when there is no lag between the two loudspeakers, there are no interaural cues under anechoic conditions (assuming head symmetry; see the Appendix) that could contribute to the observed release from masking. If the release from masking is due, in part, to comb-filtering effects, then a smaller release from masking should have been observed when the perceived location of the masker was in front than was observed when the masker was perceived on the left. However, as Figures 5 and 6 show, both conditions provided an equivalent amount of release from masking. Hence, the results here strongly support the notion that the release from masking that occurs when the perceived location of the masker is shifted away from that of the target is not due to peripheral auditory processing of acoustic cues but, rather, to more central (i.e., cognitive-level) mechanisms.

A third difference that could have led to differences between our study and that of Freyman et al. (1999) is that we used two voices in the speech masker, whereas Freyman et al. used only one. The use of two talkers rather than one talker could have affected the degree of informational masking. Nevertheless, the results were similar across the two studies despite these differences.

Using the precedence effect to shift the perceived location of the masker away from that of the target significantly alters the auditory scene (Bregman, 1990) without having much of an effect on peripheral factors. At a cognitive level, the ability to spatially separate a target from a masker might be expected to improve a listener's ability to process the information in the target without interference from the information contained in the masker. It follows that the advantage accrued by spatial separation should be greater the greater the similarity there is, at a cognitive level,

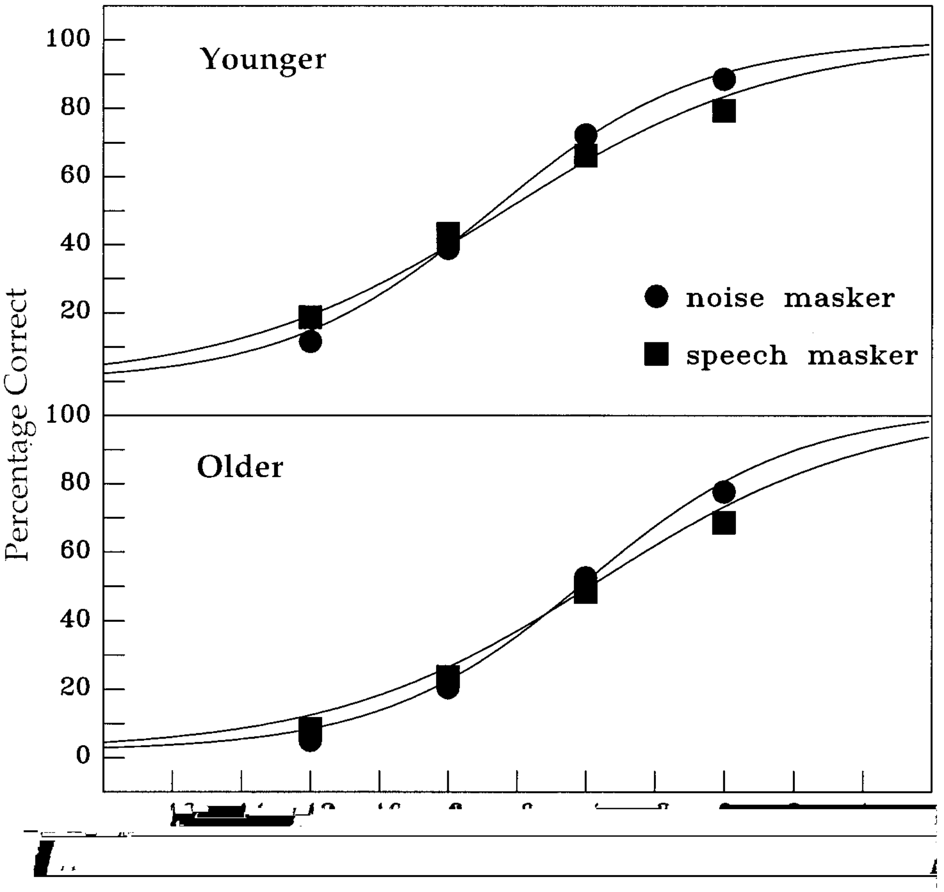

Figure 8. Mean percentages of correct responses as a function of signal-to-noise ratio when the perceived location of the masker differed from that of the target speech for both younger and older listeners. Data for the two conditions in which the location of the masker differed from that of the target (masker location left, masker location center) were pooled because performance in these conditions did not differ.

The present results suggest that when older adults are attending to a target speaker, they are no more susceptible than are younger adults to informational interference from other talkers. Of course, we need to be cautious about generalizing these results to more complex speech-recognition situations. Remember that the language task used here was simple word recognition. Because both the target sentences and the masking sentences were grammatically correct but meaningless utterances, listeners probably did not need to engage the full range of linguistic, semantic, and cognitive mechanisms that were available to them. It is possible that we might find that age differences in susceptibility to informational interference increase as the complexity of the listening task increases. For instance, if both target and masking speech consisted of meaningful sentences (e.g., "The man walked the dog"), both the degree of linguistic and semantic activation and the degree of similarity between target and masker sentences would be greater. Under such conditions, age-related differences in inhibitory control might become more apparent. This speculation awaits further experimentation. This qualification notwithstanding, the current findings strongly suggest that, at least in the case of simple word recognition, older adults experience greater difficulty in noisy situations than do younger adults because of age-related auditory declines (see also Schneider et al., 2000, 2002) and that there is no evidence to suggest that age-related changes at the cognitive level are contributing to these difficulties.

References

Arbogast, T. L., Mason, C. R., & Kidd, G. (2002). The effect of spatial separation on informational and energetic masking of speech. *Journal of the Acoustical Society of America, 112,* 2086–2098.

Blauert, J. (1997). *Spatial hearing.* Cambridge, MA: MIT Press.

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.

Cerella, J. (1990). Aging and information-processing rate. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 201–221). San Diego, CA: Academic Press.

Cheesman, M. F., Hepburn, D., Armitage, J. C., & Marshall, K. (1995). Comparison of growth of masking functions and speech discrimination abilities in younger and older adults. *Audiology, 34,* 321–333.

Cherry, E. C. (1953). Some experiments on the recognition of speech with one and two ears. *Journal of the Acoustical Society of America, 25,* 975–979.

Committee on Hearing, Bioacoustics and Biomechanics. (1988). Speech understanding and aging. *Journal of the Acoustical Society of America, 83,* 859–895.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466.

Dubno, J. R., Ahlstrom, J. B., & Horwitz, A. R. (2002). Spectral contributions to the benefit from spatial separation of speech and noise. *Journal of Speech, Language, and Hearing Research, 45,* 1297–1310.

Dubno, J. R., Dirks, D. D., & Morgan, D. E. (1984). Effects of age and mild hearing loss on speech recognition in noise. *Journal of the Acoustical Society of America, 76,* 87–96.

Duquesnoy, A. J. (1983). Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons. *Journal of the Acoustical Society of America, 74,* 739–743.

Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2001). Spatial release from informational masking in speech recognition. *Journal of the Acoustical Society of America, 109,* 2112–2122.

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.

Frisina, D. R., & Frisina, R. D. (1997). Speech recognition in noise and presbycusis: Relations to possible neural mechanisms. *Hearing Research, 106,* 95–104.

(1994). Gap detection and the precedence effect in young and old adults. *Journal of the Acoustical Society of America, 95,* 980–991.

Stuart, A., & Phillips, D. P. (1996). Word recognition in continuous and interrupted broadband noise by young normal-hearing, older normal-hearing, and presbyacusic listeners. *Ear & Hearing, 17,* 478–489.

Tun, P. A., O'Kane, G., & Wingfield, A. (2002). Distraction by competing speech in young and older adult listeners. *Psychology and Aging, 17,* 453–467.

Tun, P. A., & Wingfield, A. (1999). One voice too many: Adult age differences in language processing with different types of distracting sounds. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 54,* 317–327.

Wallach, H., Newman, E. B., & Rosenzweig, M. R. (1949). The precedence effect in sound localization. *American Journal of Psychology, 62,* 315–336.

Willott, J. F. (1991). *Aging and the auditory system: Anatomy, physiology, and psychophysics.* San Diego, CA: Singular.

Wolfram, S. (1991). *Mathematica: A system for doing mathematics by computer.* New York: Addison-Wesley.

Zurek, P. M. (1980). The precedence effect and its possible role in the avoidance of interaural ambiguities. *Journal of the Acoustical Society of America, 67,* 952–964.

Zurek, P. M. (1993). Binaural advantages and directional effects in speech intelligibility. In G. A. Studebaker & I. Hockberg (Eds.), *Acoustic factors affecting hearing aid performance* (pp. 255–276). Boston: Allyn & Bacon.

Appendix

Peripheral Cues Associated With Perceived Spatial Separation

The following is a more detailed explanation of why changing the perceived location of stimuli bypasses age-related differences in peripheral processing. Changing the physical location of a sound source produces acoustic changes in the sounds arriving at the eardrums from that source that could change the signal-to-noise ratio at each ear as well as the interaural timing relationships between the sounds at each ear. If these changes are favorable, they will lead to a release from energetic masking. For example, if a change in the physical location of a sound improves the signal-to-noise ratio, then the threshold for detecting the signal will be reduced. Moreover, a binaural listening advantage arises when the addition of a target to a masker changes the interaural correlation. If a change in the physical location of a source results in a larger change in interaural correlation when a signal is added to a masker, then the threshold for detecting the signal will be reduced. Hence, changing the physical location of a sound source can dramatically change the acoustic cues that can be used by peripheral auditory processes to unmask a signal. If peripheral auditory processes in older adults differ from those in younger adults (Schneider, 1997), the amount of unmasking produced by shifting the spatial location of a sound source could differ between younger and older listeners. Therefore, in the event that the masker is informational (e.g., semantically meaningful), it would be difficult to determine whether the age difference in the amount of masking is due to differences in peripheral auditory processes or to age differences in listeners' ability to inhibit the processing of irrelevant material.

However, when perceived spatial location is shifted using precedence, it can be shown that changing the perceived location of a source does not change the signal-to-noise ratio at the two ears or the interaural correlation in ways that could be used to provide a significant release from energetic masking. Hence, any changes in word recognition occasioned by a shift in the perceived location of a source would have to be due to other factors. Specifically, if a change in the spatial location of an informational masker were to lead to a greater reduction in masking in younger than in older adults, this age difference could not be attributed to age-related declines in the peripheral processes responsible for release from energetic masking. Conversely, if older and younger adults do not differ in their ability to inhibit the processing of irrelevant semantic information, then performance differences between older and younger adults should not change with shifts in the perceived spatial location of the masker. Below, we show why this would be the case.

Consider a situation in which there is a loudspeaker x degrees to the left of the listener and another loudspeaker x degrees to the listener's right. Suppose the distance from the left loudspeaker to the left ear is the same as the distance of the right loudspeaker to the right ear. Let $g(t)$ represent the output from the left loudspeaker. Because of the head-related transfer function, the output from the left loudspeaker arriving at the left ear is a linearly filtered version of $g(t)$. Let $H_{\mathrm{L,L}}$ represent the transformation that this linear filter imposes on the output from the left loudspeaker when it arrives at the left ear. Hence, the linearly filtered output from the left loudspeaker arriving at the left eardrum is $y_{\mathrm{L,L}}(t) = H_{\mathrm{L,L}}[g(t)]$, where the first subscript stands for the loudspeaker producing the sound, and the second stands for the ear being stimulated. Because this filter is linear and time-shift invariant, it follows that when the output from the left loudspeaker is delayed by δ seconds, $y_{\mathrm{L,L}}(t - \delta) = H_{\mathrm{L,L}}[g(t - \delta)]$. Note that because the distance from the left loudspeaker to the right ear is greater than the distance from the left loudspeaker to the left ear, the output, $g(t)$, from the left loudspeaker also arrives at the right ear τ s later. There it is filtered by the head-related transfer function for the right ear. Hence, the output from the left loudspeaker arriving at the right ear is $y_{\mathrm{L,R}}(t - \tau) = H_{\mathrm{L,R}}[g(t - \tau)]$, where the $H_{\mathrm{L,R}}$ represents linear filtering of the output from the left loudspeaker due to the head-related transfer function for the right ear. Similarly, when the same output is produced by the right loudspeaker, $y_{\mathrm{R,L}}(t - \tau) = H_{\mathrm{R,L}}[g(t - \tau)]$, and $y_{\mathrm{R,R}}(t) = H_{\mathrm{R,R}}[g(t)]$.

Now consider the situation in which the same output is presented over both loudspeakers in an anechoic environment, with the output from the left loudspeaker leading that from the right loudspeaker by δ s ($\mathrm{L} - \mathrm{R} = \delta$; perceived location of the sound is at the left). If it is assumed that the head is perfectly symmetrical, so that $H_{\mathrm{L,L}} = H_{\mathrm{R,R}}$, and $H_{\mathrm{R,L}} = H_{\mathrm{L,R}}$, then it follows that the left- and right-ear sounds are

$$y_{\mathrm{L,L}}(t) + y_{\mathrm{R,L}}(t - \tau - \delta)$$

and

$$y_{\mathrm{L,L}}(t - \delta) + y_{\mathrm{R,L}}(t - \tau),$$

respectively ($\mathrm{L} - \mathrm{R} = \delta$; perceived location is left). For the situation in which the left loudspeaker lags the right loudspeaker ($\mathrm{L} - \mathrm{R} = -\delta$; perceived location of the sound is at the right), the sounds at the right and left ear are

$$y_{\mathrm{L,L}}(t) + y_{\mathrm{R,L}}(t - \tau - \delta)$$

and

$$y_{\mathrm{L,L}}(t - \delta) + y_{\mathrm{R,L}}(t - \tau),$$

respectively ($\mathrm{L} - \mathrm{R} = -\delta$; perceived location is right). Finally, for the situation in which there is no lag between the left and right loudspeakers,

$$y_{\mathrm{L,L}}(t) + y_{\mathrm{R,L}}(t - \tau),$$

for the right ear, and

$$y_{\mathrm{L,L}}(t) + y_{\mathrm{R,L}}(t - \tau),$$

for the left ear ($\mathrm{L} - \mathrm{R} = 0$; perceived location is center).

Notice that the sound at the left ear when the left loudspeaker is leading is identical to the sound at the right ear when the right loudspeaker is leading, and the sound at the right ear when the left loudspeaker is leading is the same as the sound in the left ear when the right loudspeaker is leading. Hence, the interaural correlation when the left loudspeaker is leading is the same as the interaural correlation when the right loudspeaker is leading. Note also that if τ and δ were both 0, the interaural correlation would be 1.0 in all three conditions. Hence, in the current situation, interaural correlations of less than 1.0 are due solely to interaural timing differences between the sounds at the two eardrums. It follows that, under these conditions, the only way to change the interaural correlation is to change the timing relations between left- and right-ear sounds.

The addition of a target to a masker can change the timing relationships that exist between the two ears when only the masker is present. Hence, the interaural correlation could change when a target is added to a masker. However, when the masker and target are independent, it can be shown that adding the target to the masker changes the interaural correlation by the same amount in both conditions (left masker leading vs. right masker leading).

Now consider what happens when only the masker is played over the two loudspeakers. The equations above show that the interaural correlation for the condition in which the masker is perceived on the left ($\mathrm{L} - \mathrm{R} = \delta$) would be the same as the interaural correlation for the condition in which the masker is perceived on the right ($\mathrm{L} - \mathrm{R} = -\delta$). It can also be shown that when a target whose perceived location is on the right is added to the masker, the interaural correlation changes by the same amount when the masker is perceived on the left as it does when the masker is perceived on the right, provided that there is independence between target and masker. Hence, any difference in the amount of masking between these two conditions (left masker leading vs. left masker lagging) cannot be attributed to interaural timing differences between the two situations. The experimental results show that target recognition is higher when the masker is perceived to be coming from the left than when it is perceived to be coming from the right. Therefore, the release from masking that occurs when the perceived location of the masker is shifted from the left to the right cannot be due to differences in interaural timing relationships.

However, the same two conditions (left masker leading vs. left masker lagging) will produce differences in the long-term spectra of the maskers at the two ears because of comb-filtering effects. Comb filtering occurs when two correlated sounds, such as the linearly filtered maskers arriving at an ear from each loudspeaker, are added together. Specifically, a modulation will occur in the long-term spectrum of the summed sounds, with the rate of modulation in the spectrum depending on the pattern of timing differences between the two correlated sounds entering into the sum. Hence, the function relating power to frequency will have alternating peaks and troughs like the tines on a comb. When the timing relations between the two summed sounds in the right ear differ from the two summed sounds in the left ear (as they do for $\mathrm{L} - \mathrm{R} = \delta$ and $\mathrm{L} - \mathrm{R} = -\delta$), the pattern of modulation in the long term spectrum of the right ear will differ from that in the left ear. Because switching the masker presented over the two loudspeakers from left leading to left lagging switches the sounds produced by the masker from one ear to the other, comb-filtering effects could account for differences in the amount of masking between the condition in which the left masker was leading versus the condition in which the right masker was leading.

However, when there is no delay between the masking sounds produced by the left and right loudspeakers, comb filtering of the masker will be the same in both ears. Hence, a comparison of the condition in which the left masker is leading the right with the condition in which there is no delay between the left and right maskers can be used to determine the extent to which comb filtering of the masker may be affecting word-recognition accuracy. Because we found no differences in performance between these two conditions for either noise or speech maskers, we conclude that comb-filtering effects on performance in these conditions are negligible. However, the conclusion that comb-filtering effects are negligible in this situation must be qualified by noting that the change in interaural correlation that results from adding the target when there is no delay between the maskers is not the same as the change in interaural correlation that results from adding the target in the other two conditions in which there is a delay in masker.

The analyses above characterize what we would expect if the loudspeakers were symmetrically placed with reference to the head, the head was perfectly symmetrical, and the listening conditions were anechoic. Because we were testing in a sound-attenuating chamber, our test situation was not anechoic. Hence, reflections will affect the left- and right-ear sounds. Moreover, human heads are not symmetrical, and in our experiment, the head was not held in place, so there was no guarantee that the placement of loudspeakers was strictly symmetrical. All of these factors would add to the complexity of the signal. However, it is difficult to see how more or less random changes to the sounds at each ear (assuming that head movements are random) could account for the effects that we observed.

Received September 15, 2003
Revision received March 29, 2004
Accepted May 12, 2004 ■