J Am Acad Audiol 18:5–16 (2007)

Directional Benefit in the Presence of Speech and Speechlike Maskers

Benjamin W.Y. Hornsby*
Todd A. Ricketts*

Abstract

Recent research suggests that omnidirectional hearing aids are relatively ineffective at improving speech understanding in everyday conversational speech settings when the background noise contains both energetic and informational masking components. Energetic masking refers to situations where the peripheral (or neural) activity of the target is less than that of the masker, thus making the target inaudible. In contrast, informational masking effects, in this paper, refer to additional masking effects that are not energetic in nature. The current study evaluated the benefits of directional technology in the presence of background noises that contained both energetic and informational masking components. Aided speech recognition (in both omnidirectional and directional modes) was assessed in the presence of three types of maskers (forward and reversed speech and speech-modulated noise) that varied in the amount of informational masking they were expected to produce. Study results showed significant directional benefit in all conditions. This finding suggests that in everyday conversational speech environments, directional technology is equally efficacious regardless of the magnitude of informational masking present in the background noise. In addition, study findings suggest that the semantic information present in the masking speech may play only a limited role in contributing to informational masking in everyday environments.

Key Words: Directional benefit, energetic masking, informational masking, hearing aids, hearing loss, speech intelligibility

Abbreviations: HINT = Hearing in Noise Test; REIG = real ear insertion gain; RT = Reverberation Time; SNHL = sensorineural hearing loss; SNR = signal-to-noise ratio

Sumario

Investigaciones recientes sugieren que los auxiliares auditivos omnidireccionales son relativamente ineficientes en cuanto a mejorar la comprensión del lenguaje en condiciones cotidianas de conversación, cuando el ruido de fondo contiene componentes de enmascaramiento tanto energéticos como de información. El enmascaramiento energético se refiere a situaciones donde la actividad periférica (o neural) de la señal meta es menor que aquella del enmascarador, haciendo la señal inaudible. En contraste, los efectos de enmascaramiento de información, se refieren a efectos de enmascaramiento adicional, que no son de naturaleza energética. El estudio actual evaluó los beneficios de la tecnología direccional en presencia de ruidos de fondo que contenían tanto componentes de enmascaramiento energético como de información. Se

---

*Department of Hearing and Speech Science, Vanderbilt Bill Wilkerson Center, Vanderbilt University

Benjamin Hornsby, Dan Maddox Hearing Aid Research Laboratory, Vanderbilt Bill Wilkerson Center, Room 8310 Medical Center East, South Tower, 1215 21st Ave. South, Nashville, TN 37232-8242; E-mail: ben.hornsby@vanderbilt.edu; Phone: 615-936-5132; Fax: 615-936-5013

This research was supported in part by Phonak, Inc., and the Dan Maddox Hearing Aid Research Foundation.

evaluó el reconocimiento amplificado del lenguaje (tanto en el modo omnidireccional como direccional) en presencia de tres tipos de ruidos enmascarantes (lenguaje anterógrado y reverso, y ruido de lenguaje modulado), con variaciones en la cantidad de enmascaramiento de información que debían producir. Los resultados del estudio mostraron un beneficio direccional significativo en todas las condiciones. Este hallazgo sugiere que en ambientes cotidianos de conversación, la tecnología direccional es igualmente eficiente, independientemente de la magnitud del enmascaramiento de información presente en el ruido de fondo. Además, los hallazgos del estudio sugieren que la información semántica presente en el lenguaje enmascarante puede jugar solamente un papel limitado, en ambientes cotidianos, en su contribución como enmascaramiento de información.

Palabras Clave: Beneficio direccional, enmascaramiento energético, enmascaramiento de información, auxiliares auditivos, pérdida auditiva, inteligibilidad del lenguaje

Abreviaturas: HINT = Prueba de Audición en Ruido; REIG = ganancia de inserción de oído real; RT = Tiempo de Reverberación; SNHL = hipoacusia sensorineural; SNR = tasa señal-ruido

One of the most common complaints of persons with hearing loss is difficulty understanding speech, particularly in the presence of a background noise (e.g., Kochkin, 2000). Substantial research has demonstrated that providing a well fit omnidirectional hearing aid can improve speech understanding and reduce the communication difficulties of persons with hearing loss (e.g., Mulrow et al, 1990; Humes et al, 1997; Larson et al, 2000, 2002). Despite these findings, many persons with hearing loss continue to report significant communication difficulties in everyday environments (Kochkin, 2000). The difficulty understanding speech in background noise is due largely to the masking effects of the background noise. In common everyday settings, background noise may originate from a number of different sources and may include competing talkers (speech) and/or nonspeech signals. When the background noise consists of speech, the masking effects can be due to both "energetic" and "informational" components.

The effects of "energetic" maskers have been well studied and are assumed to originate in the auditory periphery (i.e., cochlea or proximal portions of the auditory nerve). Energetic masking occurs when the excitation or neural response in a given frequency range, due to the target, is less than that produced by the background noise (e.g., Moore et al, 1997). The effects of energetic masking on threshold and speech understanding are, on average, quite predictable, at least for persons with normal hearing (e.g., French and Steinberg, 1947; ANSI S3.5, 1997).

In contrast, "informational masking" may be described as masking that is not "energetic" in nature, thus implying more "central" masking effects (e.g., Durlach et al, 2003). When the target and background noise consists of speech, informational masking may occur when there are high levels of uncertainty regarding the characteristics of the target stimulus and/or masker (Brungart and Simpson, 2004; Freyman et al, 2004). Informational masking may also occur when there are similarities in the temporal and/or semantic structure of the background competition and the speech target (e.g., Brungart, 2001; Brungart et al, 2001). In real-world settings, however, many factors come into play that can reduce the effects of informational masking on speech understanding. For example, informational masking is reduced when the target speech and maskers are spatially separated, as is common in everyday settings (Hawley et al, 1999, 2004; Arbogast et al, 2002, 2005). Likewise, differences in vocal characteristics and the context of the target and masking speech, as well as the presence of visual cues, would work together to reduce target/masker uncertainty and similarities in the temporal and semantic structure between the target and masking speech, thus reducing informational masking (Brungart, 2001; Brungart et al, 2001; Helfer and Freyman, 2005).

Although the magnitude of informational masking effects may be reduced in real-world environments (e.g., settings where the background noise consists of spatially separated individual talkers) compared to some experimental conditions, recent research suggests that residual effects may remain substantial (Hornsby et al, 2006). Hornsby et al measured the signal-to-noise ratio, or SNR, needed for 50% sentence recognition, for persons with hearing loss in both the unaided and aided conditions, in the presence of multiple (two, four, or seven) individual talkers or multiple speech-modulated noises that were spatially distributed around the listener. Informational masking effects were estimated by comparing thresholds in the speech maskers to thresholds obtained using speech-modulated noises as maskers. The masking noises were shaped to match the long-term spectrum of the speech maskers and were modulated with the envelopes of the individual talkers. The magnitude of informational masking varied (between approximately 0.0 and 2.5 dB) based on the number of maskers and whether participants were listening unaided or aided.

In addition, the study results suggested that omnidirectional hearing aids were relatively ineffective in improving speech understanding when the masking noises contained an informational masking component (i.e., speech-masker conditions). Specifically, when the background noise consisted of actual talkers, improvements in speech understanding in the aided condition (quantified as the improvement in SNR) were small (range of 0.7–1.4 dB), and aided performance was not significantly different than unaided performance. In contrast, when the background noise consisted of modulated speech noises, the use of omnidirectional aids resulted in significant (range of 1.5–2.9 dB) improvements in speech understanding compared to unaided performance.

The finding that omnidirectional hearing aids are limited in their ability to improve speech understanding deficits associated with informational masking in everyday speech, although discouraging, is not totally unexpected. Hearing aids are designed, primarily, to restore audibility and are thus not expected to heavily impact the central factors (e.g., stimulus-masker uncertainty/similarity) that appear to be largely responsible for informational masking effects that occur in speech settings. This finding does, however, highlight the limitations of omnidirectional hearing aids in real-world noisy environments, particularly those that contain speech signals as the primary masker. Given these limitations, there is a clear need for additional research investigating the utility of additional technological and counseling strategies that may help reduce the negative effects of informational maskers on hearing aid wearers.

One option for improving speech understanding in the presence of speech maskers is the use of directional technology in hearing aids. The use of directional technology in hearing aids has been shown to improve speech understanding in a variety of noise backgrounds, at least in laboratory settings, compared to performance with well-fit omnidirectional hearing aids (see Ricketts and Dittberner, 2002, for a review). Despite this success in laboratory environments, the benefits of directional technology in real-world settings appear to be more modest (Cord et al, 2004) and may be related to the fact that the background noises present in real-world settings contain both informational and energetic masking components. The benefits of directional technology are due primarily to their ability to improve SNR, which for most noises is monotonically related to speech understanding. For highly informational maskers, however, the function relating SNR to speech understanding may vary substantially and may not always be monotonic (e.g., Dirks and Bower, 1969; Brungart, 2001). Thus, the benefits of directional technology may be limited or highly variable in the presence of maskers containing both informational and energetic masking components. The current study extends our previous work (Hornsby et al, 2006) by investigating the benefits in speech understanding provided by directional technology in the presence of speech and speechlike background noises containing both energetic and informational masking components.

PROCEDURES

Participants

Participants consisted of 14 adults (seven male and seven female) with symmetrical (≤20 dB difference between ears at any audiometric test frequency), mild-to-moderately severe flat or sloping sensorineural hearing losses (SNHL). SNHL was defined as having no significant air-bone gap (<10 dB) at any frequency tested and no history of conductive pathology. All participants' hearing thresholds were between 25 and 70 dBHL at 500Hz and between 50 and 85 dBHL at 3000 Hz, bilaterally (see Figure 1). Participants with hearing loss ranged in age from 23 to 83 years (median age: 74 years).

Test Setting

The test environment was the same as that used in our previous study (Hornsby et al, 2006). Briefly, a 3.2-meter square (2 meters high) sound-treated room, modified with reflective panels, served as the test environment. Frequency-specific reverberation times (RT_{60}; time required for 60 dB decay after signal offset) were measured at the position of the listener's head (without the listener present) using frequency-modulated tones. Measured RT_{60} values, at octave frequencies, were 485 msec (250 Hz), 440 msec (500 Hz), 400 msec (1000 Hz), 310 msec (2000 Hz), and 220 msec (4000 Hz).

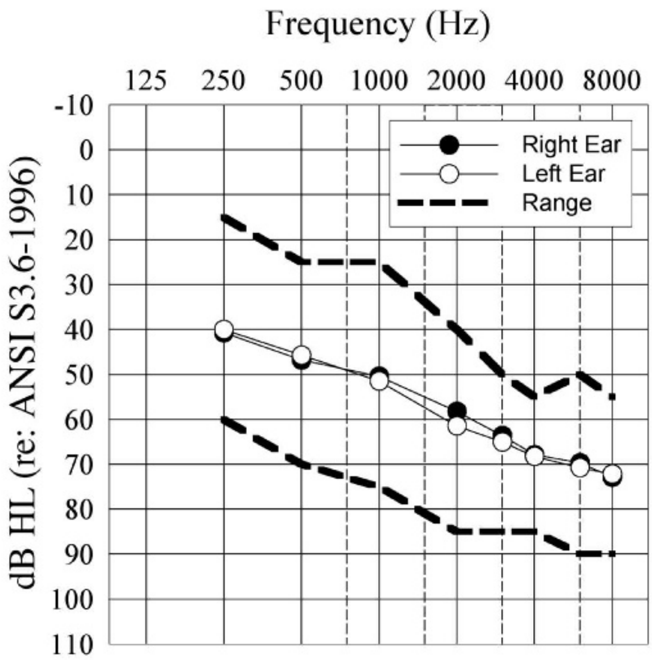

Figure 1. Average audiograms of study participants with hearing loss. Open and filled symbols are for the left and right ears, respectively. The range of hearing losses included in this study is shown by dashed lines.

Hearing Aids

The behind-the-ear version of the Phonak Claro™ (211dAZ) was used as the test instrument. This is a digital, 20-channel, low-threshold, fast-acting compression hearing aid. This hearing aid's digital noise reduction circuitry, referred to by the manufacturer as "fine scale noise cancellation," was disabled for all testing. The devices were fit bilaterally in both omnidirectional and adaptive directional modes using the manufacturer's fast-acting compression algorithm, which they refer to as "digital perception processing" (fast-acting DPP™). Using this algorithm, nominal compression thresholds across subjects ranged from approximately 30 to 50 dB SPL, and nominal attack and release times were 5 and 90 milliseconds respectively. Further details related to compression processing were not deemed of interest since past research has shown that directional benefit is generally unaffected by the presence of low-threshold compression and the compression parameters selected (Ricketts et al, 2001).

Given the relatively discrete competing noise source positions used in this experiment (i.e., single interferers from multiple loudspeakers), the adaptive, rather than fixed, directional mode was chosen for use in this study. It is important to note that "automatic" switching between directional and omnidirectional modes was not used in this study. That is, the aids were programmed to function in either omnidirectional or adaptive directional mode with the specific mode chosen by the experimenter (i.e., the aid could not "automatically" switch between programs). In adaptive directional mode, the two omnidirectional microphones both constantly function in a directional array. The directional sensitivity pattern is then altered by automatically varying the physical directional properties (by varying internal delay) until an attenuation pattern resulting in the lowest output is obtained. This process of shifting directional sensitivity patterns requires only several milliseconds to complete. The adaptive directional circuitry is limited so that directional parameters that result in nulls in the front hemisphere are avoided. In this way important sounds that arrive from the front hemisphere are not inadvertently, and undesirably, attenuated (Checkley and Kuehnel, 2000).

In theory, the adaptive mode should allow the device to optimize the directional characteristics for a given speaker configuration and therefore allow for assessment of optimal, but realistic, directional benefit in the presence of discrete, stationary, competing maskers.

Participants with hearing loss were fit bilaterally. The hearing aid fittings, based on the NAL-RP prescriptive method (Byrne and Dillon, 1986), were verified using the composite noise test signal (0° azimuth) on the Frye Systems Fonix 6500 real ear analyzer. For a 65 dB input, average measured real ear insertion gain (REIG) values were within 4 dB of prescribed gain values at test frequencies of 250 to 4000 Hz and within 6 dB at 6000 Hz. The average difference between omnidirectional and directional REIG was less than approximately 2 dB at frequencies between 250 and 6000 Hz (see Figure 2).

Speech Understanding Test Materials and Masking Stimuli

Speech understanding in noise was assessed using a modified version of the Hearing in Noise Test (HINT; Nilsson et al, 1994). The modifications to the HINT procedure were all related to the presentation and type of competing noise used as described below. The HINT is an adaptive procedure used to determine the SNR necessary to achieve 50% correct sentence recognition. Using this test the sentence level is adaptively varied in the presence of a constant level background noise to determine threshold. During speech testing the background noise was turned on approximately one second prior to the presentation of the first sentence and was on continuously during testing (i.e., there were no silent periods between sentences). Each experimental condition was evaluated using two ten-sentence lists. Experimental condition and list order were randomly assigned to each subject using a Latin square design. In addition, the participants were tested in both omnidirectional and directional modes with the presentation order counterbalanced.

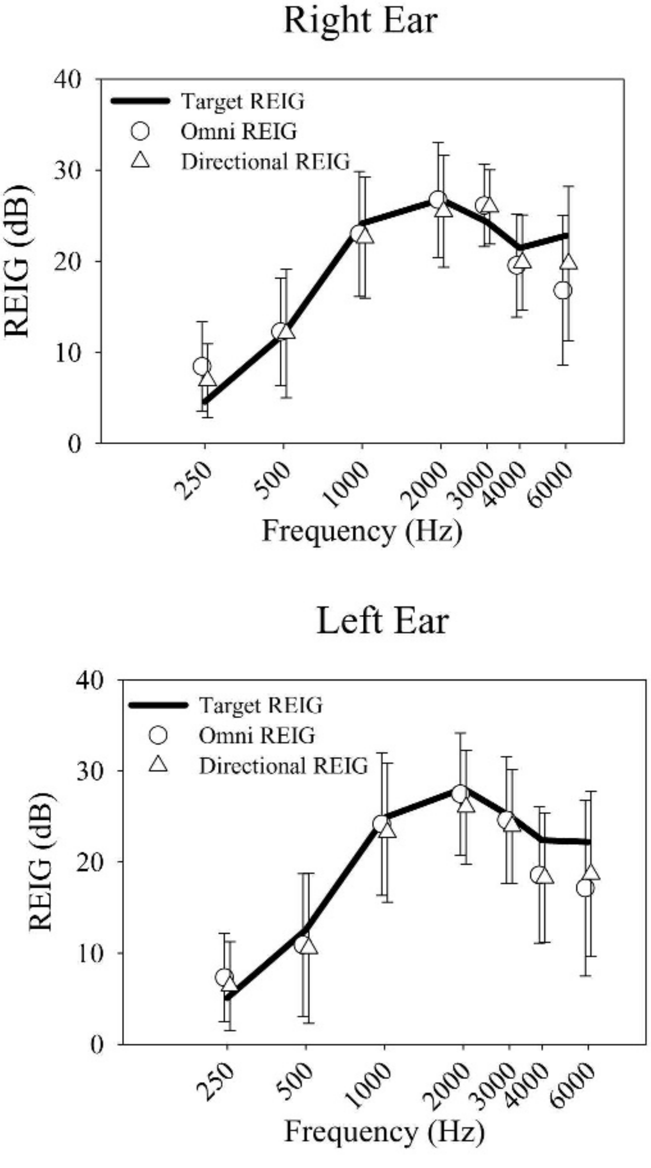

Figure 2. Average measured and target real ear insertion gain (REIG in dB) for the right and left ears in omnidirectional (circles) and directional (triangles) modes. Error bars represent one standard deviation around the mean. For clarity, symbols are slightly offset on the frequency axis.

Speech materials were presented through a single loudspeaker (Tannoy System 600) located approximately 1.2 meters from the subject (0° azimuth) and positioned at approximately ear level.

Speech recognition was assessed in a four- and seven-masker configuration using actual talkers (forward speech), reversed speech, and modulated speech-shaped noises as maskers. Recent work suggests that the informational masking effects of individual talkers peaks as the number of individual talkers approaches eight (Simpson and Cooke, 2005). Using the four- and seven-masker configurations allows us to maximize informational masking effects while maintaining face validity in the experimental configuration (e.g., speech maskers surrounding a listener).

In each masker configuration, speech understanding was assessed using maskers that varied in the amount of informational masking they were expected to produce. The forward speech maskers consisted of male talkers reading a specific topic and were assumed to produce the highest amount of informational masking. The male talkers read test passages from the Connected Speech Test (CST; Cox et al, 1987, 1988). Each CST passage consists of approximately ten sentences describing a specific topic (e.g., lemons). The passages were derived from a children's educational reading source (see Cox et al, 1987, for details). The CST passages were chosen as masking materials because they provide a contextually rich dialogue, rather than single unrelated sentences, that may result in informational masking more consistent with that present in everyday communication settings. Recordings of the individual talkers reading the CST passages were made in an anechoic chamber and stored on a computer hard disk. All talkers were native speakers of American English. Offline digital filtering, using the FIR2 function and a 1000th order FIR filter implemented in Matlab™, was used to match the long-term rms spectra of each CST passage recorded by each talker to that of the long-term rms spectra of the HINT materials.

The modulated noise maskers used in this study were derived from uncorrelated segments of Gaussian noise that were first spectrally shaped, using the FIR2 function and a 1000th order FIR filter implemented in Matlab, to match the long-term rms spectrum of the HINT materials. The shaped noises were then modulated using the envelopes of the same single talkers reading the CST passages. The envelopes of the single-talker maskers were derived in Matlab by implementing half-wave rectification of a given single-talker passage, followed by low-pass filtering of the half-wave rectified signal, using a sixth-order butterworth filter with a 30 Hz low pass cutoff frequency. The envelope was then applied to the shaped noise providing a primarily energetic masker that retained the long-term spectral and temporal patterns of the single-talker maskers. The long-term spectrum of all individual-talker and modulated noise maskers closely approximated the long-term spectrum of the HINT sentences. The mean error, across one-third octave bands from 160 to 8000 Hz, between individual talkers, modulated noises and the HINT spectrum was 0.19 and -0.22 dB, respectively, with standard deviations less than 1 dB. The maximum error, in a given one-third octave band, for any individual talker or modulated noise was ~3 dB.

In addition, a third masker type consisting of reversed speech was added to the experimental conditions. Specifically, the forward speech maskers were reversed in the time domain, resulting in an unintelligible masking noise with similar spectral and temporal properties as the forward speech. It was expected that reversing the speech would create a masker with similar energetic masking properties compared to the speech but without semantic information that may contribute to target and masker similarity, thus reducing informational masking.

We estimated informational masking effects by comparing recognition performance in the forward speech-masker condition (more informational masking) to performance in the reversed speech and speech-modulated noise masker conditions (less informational masking). In summary, sentence recognition was assessed in a total of 12 conditions (two aid types [omni and directional] x two noise configurations [four and seven maskers] x three masker types [forward and reversed speech and speech-modulated noise]).

Each masker was presented from a separate loudspeaker that was spatially separated from the target speech, which was always located at a 0° azimuth. Although spatially separating the speech and noise provides a more realistic test configuration, it should be noted that this configuration was expected to reduce the amount of informational masking compared to that measured in some previous laboratory studies (Hawley et al, 1999, 2004; Arbogast et al, 2002, 2005). In the four-masker configuration, speakers were located symmetrically at azimuths of 45°, 135°, 225°, and 315°. In the seven-masker configuration, three additional speakers were placed at azimuths of 90°, 180°, and 270°. All maskers were presented from Definitive Technology BP-2X bipolar loudspeakers placed 1.2 meters from the listener. Informal testing confirmed that, although difficult, listeners with normal hearing could attend to the speech of individual speech maskers even when all seven were presented simultaneously.

In each test condition the combined level of the maskers (at the position of the listener's head with the listener absent), was a constant 65 dBA. The level of each individual masker was equated prior to adjusting the overall level of the combined maskers to 65 dBA. Calibration prior to each test session was performed to assure an overall level of 65 dBA (±1 dB; Larson Davis 814 sound level meter set to slow averaging) in all experimental conditions.

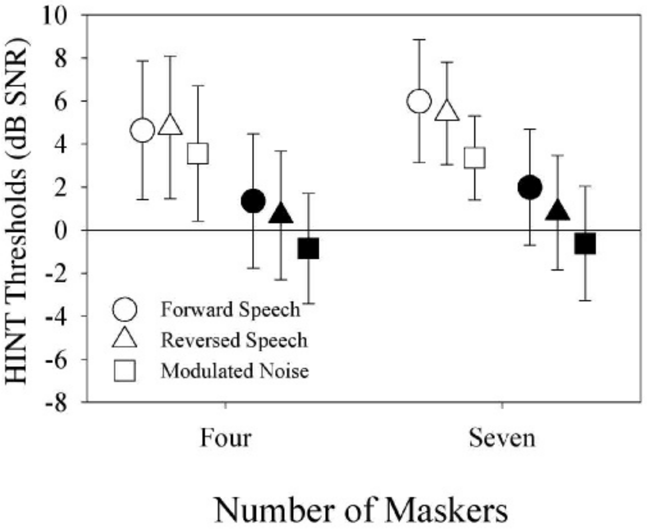

Figure 3. Thresholds (in dB SNR) required for 50% sentence recognition for aided participants with hearing loss as a function of number of maskers (four or seven). Circles, triangles, and squares represent performance in background noises of forward speech, reversed speech, and modulated noise, respectively. Performance in omnidirectional and directional modes is shown by the open and filled symbols, respectively. Error bars represent one standard deviation.

RESULTS

The current study examined the effects of type and number of maskers on the aided speech understanding of persons with hearing loss using both omnidirectional and directional hearing aid modes. Figure 3 shows HINT thresholds obtained in each masker type as a function of number of maskers in both omnidirectional and directional modes. Plotted in this fashion, a more negative threshold represents better understanding in noise.

Performance differences across conditions were examined using a three-factor, repeated-measures analysis of variance (ANOVA). The within subjects independent variables were microphone mode (omnidirectional or directional), masker type (forward speech, reversed speech, or speech-modulated noise), and number of maskers (four or seven); the dependent variable was the HINT score in each masker condition. A 0.05 level of significance was used in all analyses described here. Summary results of the primary analyses are shown in Table 1.

As expected, a significant main effect of microphone mode was observed with better performance occurring when listening with the microphones set to directional mode. When switching from omnidirectional to directional modes, performance improved, on average, between 3.3 dB and 4.6 dB across the test conditions. This directional benefit varied widely across individuals ranging from

Table 1. Results from a Three-Factor, Repeated Measures ANOVA

Overall ANOVA Results			
Effect	Df	F	p-level
Microphone	**1,13**	**580.8**	**<0.001**
Masker Type	**2,26**	**38.8**	**<0.001**
Num_Maskers	1,13	4.57	0.052
Microphone x Masker Type	2,26	1.47	0.248
Microphone x Num_Maskers	2,26	0.45	0.517
Masker Type x Num_Maskers	2,26	1.97	0.160
Microphone x Masker Type x Num_Maskers	2,26	0.87	0.433

Note: Significant effects with p-levels less than 0.05 are shown in boldface. Microphone: omnidirectional or directional; Masker Type: forward speech, reversed speech, or noise masker; Num_Maskers: number of maskers (i.e., four or seven).

a low of 0 dB to a maximum of 7.7 dB. No significant interaction effects between microphone mode and any other factor were observed, suggesting that, on average, significant directional benefit was present and similar in magnitude across all test conditions. Figure 4 shows directional benefit (omnidirectional-directional HINT thresholds) for the test conditions evaluated in this study.

In addition to the significant main effect of microphone mode, a significant main effect of masker type was also present. There were, however, no significant interaction effects. The lack of a significant interaction suggests the effect of masker type was similar in both omnidirectional and directional modes and in both the four- and seven-masker conditions. Follow-up testing was performed, using a series of three-factor ANOVAs, to compare performance differences between masker types. These planned comparisons showed that performance in the modulated noise background was significantly better ($df_{1,13}$; $p < 0.001$) than when the background noise consisted of forward or reversed speech. This additional masking, above that observed with the modulated noise, suggests that both the forward and reversed speech maskers contained an additional informational masking component that affected speech understanding.

In addition, there was no significant difference ($df_{1,13}$; $p = 0.078$) in performance between the forward and reversed speech background noise conditions. In other words, reversing the speech maskers provided no measurable release from informational masking in any of our test conditions. Figure 5 shows the magnitude of informational masking effects in the current study (the difference in HINT thresholds measured in speech and speech-modulated noise maskers) for the study participants in both omnidirectional and directional modes, as a function of number of maskers.

As seen in Figure 5, the magnitude of informational masking effects in the current study varied slightly (but not significantly) across microphone mode, type of masker (forward or reversed speech), and number of maskers. The lack of a significant interaction effect here may be due in part to relatively low statistical power. The low power, however, is due largely to the fact that relatively small differences are observed across conditions relative to the variability across study

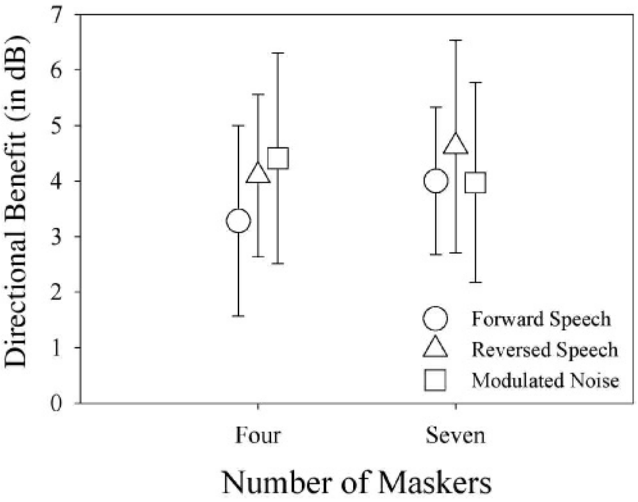

Figure 4. Directional benefit (omnidirectional thresholds–directional thresholds for 50% sentence recognition) in dB SNR as a function of number of maskers (four or seven) and type of masking noise (forward speech, circles; reversed speech, triangles; and speech modulated noise, squares). Error bars represent one standard deviation. Positive values reflect better performance in directional mode.

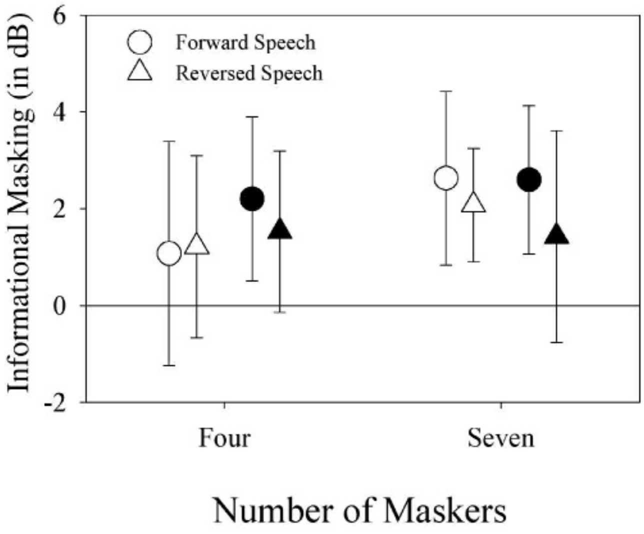

Figure 5. Informational masking effects, defined as the difference between the HINT thresholds in the forward or reversed speech masker minus the HINT thresholds in the modulated noise masker, as a function of number of maskers. Informational masking effects in omnidirectional and directional modes are shown by the open and filled symbols respectively. Results for forward (circles) and reversed (triangles) speech maskers are shown as separate symbols. Error bars represent one standard deviation.

participants. Informational masking effects (additional change in SNR required for 50% sentence recognition over that due to a modulated noise) ranged from approximately 1.0 to 2.6 dB across conditions.

DISCUSSION

Masker Type and Directional Benefit

The data shown in Figure 5 show that informational masking effects were similar in magnitude to those observed in the aided (omnidirectional) conditions of our previous study, which ranged from approximately 1.5 to 2.5 dB (Hornsby et al, 2006). It should be noted, however, that these values are relatively small compared to some studies designed to maximize informational masking effects.

For example, Arbogast et al (2002, 2005) highlighted informational masking while minimizing the effects of energetic masking by creating sine wave speech targets and maskers. The sine wave speech was created by modulating pure tones of a given frequency with the amplitude envelope of actual speech filtered at that center frequency. The authors minimized energetic masking by placing the sine waves used to generate the masking speech at center frequencies well removed from the center frequencies used to create the target speech. The result was intelligible target and masking speech that had minimal spectral overlap and thus minimal energetic masking. Using this method, Arbogast et al (2005) reported informational masking effects (the difference between speech recognition thresholds for sine wave speech obtained using modulated noise versus sine wave sentences as background maskers) of approximately 7.0 to 12.0 dB in their unaided participants with hearing loss. Thus, informational masking effects, under certain conditions, can be quite large (Hawley et al, 1999, 2004; Arbogast et al, 2002, 2005).

The results from the current study and those of Hornsby et al (2006), however, suggest that in everyday conversational speech settings (i.e., where the target speech is a single talker and the masking speech is made up of multiple, different, spatially separated talkers), informational masking effects, while present, may be relatively small. A primary focus of the current study was to explore the benefits of directional processing in the presence of background noises that were designed to vary in the amount of energetic and informational masking they were expected to produce. Despite the presence of a significant informational masking component, our study results revealed significant directional benefit that was independent of masker type. That is, the improvement in SNR provided by the directional technology resulted in comparable improvements in speech understanding regardless of whether the background noise consisted of actual speech, reversed speech, or speech-modulated noise. In addition, the magnitude of directional benefit (3.3 dB to 4.6 dB) in the current study was comparable to that seen in other studies where the background noise surrounded the listener (Ricketts and Dittberner, 2002). These findings are encouraging given previous results showing that omnidirectional hearing aids were relatively ineffective at improving speech understanding when the background noise consisted of spatially separated individual talkers as in the current study (Hornsby et al, 2006).

The finding of improved speech understanding with the use of directional technology in the presence of speech maskers is not totally unexpected given our attempt at simulating somewhat real-world conditions. In most cases, when measuring speech understanding in noise, a systematic increase in SNR (which is comparable to switching from omnidirectional to directional mode) results in a monotonic improvement in speech understanding, although the slopes of the performance functions may vary based on the target speech and masking materials used (e.g., Miller, 1947).

This monotonic decrease in performance as the SNR worsens can occur even in situations where informational masking is thought to play the prominent role (Brungart et al, 2001; Arbogast et al, 2002, 2005). In some cases, however, “plateaus” in the function relating speech recognition to SNR (i.e., a nonmonotonic function) have been observed, specifically when the test SNR is in the range of 0 to -10 dB (Egan et al, 1954; Dirks and Bower, 1969; Brungart, 2001). In some cases at these SNRs, performance remains stable despite a change in SNR. This “plateau” appears to be limited, however, to situations where substantial informational masking is present, such as when the masker is a single talker speaking with the same temporal characteristics as the target speech (Dirks and Bowers, 1969). In addition, the "plateau" often disappears and a monotonic function is again observed when more than a single talker is used as the masking noise (Brungart et al, 2001). Consequently, in real-world situations for which directional technology is capable of improving the SNR, improvements in speech understanding should be possible regardless of the type of background noise.

Informational Masking Effects in Reversed Speech

Another finding of interest in the current study was the relative lack of benefit obtained when reversing the forward speech maskers (see Figure 5). Our study results showed no significant difference in performance between the forward and reversed speech maskers in any condition. One interpretation of these results is that the semantic content of the forward speech masker resulted in no additional informational masking. This is consistent with previous research showing no difference in speech understanding when using forward and reversed speech maskers (Miller, 1947; Dirks and Bower, 1969; Duquesnoy, 1983; Festen and Plomp, 1990; Larsby and Arlinger, 1994).

Other studies, however, have shown a measurable improvement in speech understanding with the reversal of a forward speech masker (Sperry et al, 1997; Freyman et al, 2001; Hawley et al, 2004; Summers and Molis, 2004). This effect tends to be largest, however, in conditions where substantial informational masking is thought to exist. For example, release from masking due to time reversal of a forward speech masker has been observed in conditions where there is no spatial separation between the target and masking speech and/or in conditions where substantial similarities exist in the spectro-temporal and/or semantic characteristics of the target and masking speech (e.g., Freyman et al, 2001; Hawley et al, 2004). Freyman et al (2001) observed a substantially larger improvement in speech understanding due to reversal of a two-talker speech masker when the target speech and maskers both came from a 0° azimuth (high informational masking condition). The release from masking due to time reversal was substantially smaller when the target speech and maskers were perceived (via use of the precedence effect) as spatially separated (resulting in less informational masking).

By creating a time-reversed copy of the forward speech masker, the underlying assumptions are that the temporal and spectral characteristics of the forward speech masker, and thus its energetic masking properties, are essentially unchanged while the informational masking component due to the semantic information in the masker is removed. However, a contrasting view was suggested by Rhebergen et al (2005). These authors suggested that the acoustic characteristics of reversed speech result in significantly higher amounts of energetic masking than would occur when forward speech was used as a masker. Specifically, they suggest the envelope of forward speech is dominated by the relatively abrupt onsets and gradual decays of plosive sounds. When the speech is reversed, the plosives provide an abrupt offset that may make the reversed speech a more effective forward masker; thus, we would expect *poorer* thresholds in the presence of a reversed speech masker if no additional factors were involved.

To examine this hypothesis, Rhebergen et al measured performance in the presence of a forward and reversed foreign language masker (Swedish) and reported *poorer* performance using the reversed foreign language masker. In this situation there was no usable semantic information in either masker, suggesting that the poorer performance was due to greater energetic masking effects in the reversed speech. It should be noted, however, that Dirks and Bower (1969) performed a similar experiment using a Latin masker and found no difference in performance in forward and reversed modes. Thus, the variation in forward masking effects in reversed and forward speech may vary as a function of language.

In the current study, although no difference between forward and reversed speech was observed, performance in the reversed speech background noise was significantly poorer than in the modulated noise. This could be due, as suggested by Rhebergen et al (2005), to an increase in energetic masking properties of the reversed speech (due to increased forward masking), compared to the speech-modulated noise, that was offset by a similar amount (but in the opposite direction) by a release from informational masking due to the time reversal of the speech maskers.

An alternative interpretation, if one assumes that no differences exist in the forward masking properties of forward and reversed American English speech, is that the similarities in the temporal and spectral characteristics of the target and masking (both forward and reversed) speech are the primary factors responsible for the "informational" masking observed in the current study, not difficulty disentangling the semantic information in the target from that in the masker. Dirks and Bower (1969) highlighted how similarities in the temporal structure of the target and masking speech could alter the masking characteristics of a background noise. While measuring synthetic sentence identification as a function of SNR in different background noises, they observed a "plateau" in performance as SNR decreased in certain test conditions. Specifically, speech understanding remained stable in cases where the same speaker delivered both the target and masking stimuli at approximately the same level (e.g., between 0 to -10 dB SNR).

This "plateau" effect is associated with informational masking, in that energetic masking effects are known to change with SNR, yet no appreciable change in performance is observed (Brungart, 2001). Dirks and Bower measured sentence recognition using both intelligible (English) and unintelligible (Latin) masking stimuli in both forward and reversed modes. In these conditions a "plateau" was most apparent when either the English or Latin maskers were presented in the forward mode. In an additional condition, a single subject was trained to identify reversed synthetic sentences as the target speech and listened to this reversed target speech in both forward and reversed (English and Latin) masking conditions. In this condition, the "plateau" was most apparent when the maskers were also played in reversed mode. Together these findings suggest that it is the similarities in the temporal properties of the target and masking speech, not the semantic information in the masker, which were responsible for this plateau in performance.

In the current study, there are substantial similarities in the short- and long-term temporal structure of the target and masking speech while substantial differences in semantic information exist. This is similar to situations that occur in everyday environments when listening to a single talker in a background noise of conversational speech from multiple talkers. In these cases, it is possible that the similarities in temporal structure of the target and masking stimuli, rather than the semantic properties of the masker, are largely responsible for any "informational masking" effects. Further research in this area is needed to identify the roles that these and other factors play in any informational masking effects that may be present in everyday conversational settings.

SUMMARY

In the current study, the use of directional technology provided a significant improvement in speech recognition for persons with hearing loss. This benefit was present, and comparable in magnitude, regardless of whether the background noise was primarily energetic (speech-modulated noise) or contained both energetic and informational masking components (forward and reversed speech). These findings offer additional support for the use of directional technology in everyday settings where the presence of background noise, whether speech or nonspeech, limits speech understanding.

REFERENCES

ANSI (American National Standards Institute). (1997) *ANSI S3.5-1997 American National Standard Methods for the Calculation of the Speech Intelligibility Index.* New York: American National Standards Institute.

Arbogast TL, Mason CR, Kidd G. (2002) The effect of spatial separation on informational and energetic masking of speech. *J Acoust Soc Am* 112(5):2086–2098.

Arbogast TL, Mason CR, Kidd G. (2005) The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am* 117(4):2169–2180.

Brungart DS. (2001) Informational and energetic masking effects in the perception of two simultaneous talkers. *J Acoust Soc Am* 109(3):1101–1109.

Brungart DS, Simpson BD. (2004) Within-ear and across-ear interference in a dichotic cocktail party listening task: effects of masker uncertainty. *J Acoust Soc Am* 115(1):301–310.

Brungart DS, Simpson BD, Ericson MA, Scott KR. (2001) Informational and energetic masking effects in the perception of multiple simultaneous talkers. *J Acoust Soc Am* 110(5):2527–2538.

Byrne D, Dillon H. (1986) The National Acoustic Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid. *Ear Hear* 7(4):257–265.

Checkley P, Kuehnel V. (2000) Advantages of an adaptive multi-microphone system. *Hear Rev* 8(9):58–60, 74.

Cord MT, Surr RK, Walden BE, Dyrlund O. (2004) Relationship between laboratory measures of directional advantage and everyday success with directional microphone hearing aids. *J Am Acad Audiol* 15(5):353–364.

Cox RM, Alexander GC, Gilmore C. (1987) Development of the Connected Speech Test (CST). *Ear Hear* 8(5, Suppl.):119S–126S.

Cox RM, Alexander GC, Gilmore C, Pusakulich KM. (1988) Use of the Connected Speech Test (CST) with hearing-impaired listeners. *Ear Hear* 9(4):198–207.

Dirks D, Bower D. (1969) Masking effects of speech competing messages. *J Speech Hear Res* 12(2): 229–245.

Duquesnoy AJ. (1983) Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons. *J Acoust Soc Am* 74(3):739–743.

Durlach NI, Mason CR, Kidd G, Arbogast TL, Colburn HS, Shinn-Cunningham BG. (2003) Note on informational masking. *J Acoust Soc Am* 113(6):2984–2987.

Egan J, Carterette EC, Thwing E. (1954) Some factors affecting multi-channel listening. *J Acoust Soc Am* 26(5):774–782.

Festen JM, Plomp R. (1990) Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *Acoust Soc Am* 88(4):1725–1736.

French NR, Steinberg JC. (1947) Factors governing the intelligibility of speech sounds. *J Acoust Soc Am* 19:90–119.

Freyman RL, Balakrishnan U, Helfer KS. (2001) Spatial release from informational masking in speech recognition. *J Acoust Soc Am* 109(5, pt. 1):2112–2122.

Freyman RL, Balakrishnan U, Helfer KS. (2004) Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *J Acoust Soc Am* 115(5, pt. 1):2246–2256.

Hawley ML, Litovsky RY, Colburn HS. (1999) Speech intelligibility and localization in a multi-source environment. *J Acoust Soc Am* 105(6):3436–3448.

Hawley ML, Litovsky RY, Culling JF. (2004) The benefit of binaural hearing in a cocktail party: effect of location and type of interferer. *J Acoust Soc Am* 115(2):833–843.

Helfer KS, Freyman RL. (2005) The role of visual speech cues in reducing energetic and informational masking. *J Acoust Soc Am* 117(2):842–849.

Hornsby B, Ricketts T, Johnson E. (2006) The effects of speech and speech-like maskers on unaided and aided speech recognition in persons with hearing loss. *J Am Acad Audiol* 17(6):435–450.

Humes LE, Christensen LA, Bess FH, Hedley-Williams A. (1997) A comparison of the benefit provided by well-fit linear hearing aids and instruments with automatic reductions of low-frequency gain. *J Speech Lang Hear Res* 40(3):666–685.

Kochkin S. (2000) MarkeTrak V: "Why my hearing aids are in the drawer": the consumer's perspective. *Hear J* 53(2):34, 36, 39–41.

Larsby B, Arlinger S. (1994) Speech recognition and just-follow-conversation tasks for normal-hearing and hearing-impaired listeners with different maskers. *Audiology* 33(3):165–176.

Larson VD, Williams DW, Henderson WG, Luethke LE, Beck LB, Noffsinger D, Wilson RH, et al. (2000) Efficacy of 3 commonly used hearing aid circuits: a crossover trial. NIDCD/VA Hearing Aid Clinical Trial Group. *J Am Med Assoc* 284(14):1806–1813.

Larson VD, Williams DW, Henderson WG, Luethke LE, Beck LB, Noffsinger D, Bratt GW, et al. (2002) A multi-center, double blind clinical trial comparing benefit from three commonly used hearing aid circuits. *Ear Hear* 23(4):269–276.

Miller GA. (1947) The masking of speech. *Psychol Bull* 44:105–129.

Moore BCJ, Glasberg BR, Baer T. (1997) A model for the prediction of thresholds, loudness, and partial loudness. *J Audio Eng Soc* 45(4):224–240.

Mulrow CD, Aguilar C, Endicott JE, Tuley MR, Velez R, Charlip WS, Rhodes MC, et al. (1990) Quality-of-life changes and hearing impairment. A randomized trial. *Ann Intern Med* 113(3):188–194.

Nilsson M, Soli S, Sullivan JA. (1994) Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and in noise. *J Acoust Soc Am* 95:1085–1099.

Rhebergen KS, Versfeld NJ, Dreschler WA. (2005) Release from informational masking by time reversal of native and non-native interfering speech. *J Acoust Soc Am* 118(3):1274–1277.

Ricketts T, Dittberner AB. (2002) Directional amplification for improved signal-to-noise ratio: strategies, measurements, and limitations. In: Valente M, ed. *Hearing Aids: Standards, Options, and Limitations.* 2nd ed. New York: Thieme, 274–346.

Ricketts T, Lindley G, Henry P. (2001) Impact of compression and hearing aid style on directional hearing aid benefit and performance. *Ear Hear* 22(4):348–361.

Simpson S, Cooke M. (2005) Consonant identification in N-talker babble is a nonmonotonic function of N (L). *J Acoust Soc Am* 118(5):2775–2778.

Sperry JL, Wiley TL, Chial MR. (1997) Word recognition performance in various background competitors. *J Am Acad Audiol* 8(2):71–80.

Summers V, Molis MR. (2004) Speech recognition in fluctuating and continuous maskers: effects of hearing loss and presentation level. *J Speech Lang Hear* Res 47(2):245–256.