



Published in final edited form as:

Ear Hear. 2013 January ; 34(1): 3–14. doi:10.1097/AUD.0b013e31825e2841.

Release from Perceptual Masking for Children and Adults: Benefit of a Carrier Phrase

Angela Yarnell Bonino, M.S.,

Department of Allied Health Sciences, The University of North Carolina at Chapel Hill

Lori J. Leibold, Ph.D., and

Department of Allied Health Sciences, The University of North Carolina at Chapel Hill

Emily Buss, Ph.D

Department of Otolaryngology/Head and Neck Surgery, The University of North Carolina at Chapel Hill

Abstract

Objectives—The purpose of this study was to test the hypothesis that a carrier phrase can improve word recognition performance for both children and adults by providing an auditory grouping cue. It was hypothesized that the carrier phrase would benefit listeners under conditions where they have difficulty perceptually separating the target word from the competing background. To test this hypothesis, word recognition was examined for maskers that were believed to vary in their ability to create perceptual masking. In addition to determining the conditions under which a carrier-phrase benefit is obtained, age-related differences in both susceptibility to masking and carrier-phrase benefit were examined.

Design—Two experiments were conducted to characterize developmental effects in the ability to benefit from a carrier phrase (i.e., “say the word”) prior to the target word. Using an open-set task, word recognition performance was measured for three listener age groups: 5- to 7-year-old children, 8- to 10-year-old children, and adults (18 to 30 years). For all experiments, target words were presented in each of two carrier-phrase conditions: (1) carrier-present and (2) carrier-absent. Across experiments, word recognition performance was assessed in the presence of multi-talker babble (Experiment 1), two-talker speech (Experiment 2), or speech-shaped noise (Experiment 2).

Results—Children’s word recognition performance was generally poorer than adults’ for all three masker conditions. Differences between the two age groups of children were seen for both speech-shaped noise and multi-talker babble, with 5- to 7-year-olds performing more poorly than 8- to 10-year-olds. However, 5- to 7-year-olds and 8- to 10-year-olds performed similarly for the two-talker masker. Despite developmental effects in susceptibility to masking, both groups of children and adults showed a carrier-phrase benefit in multi-talker babble (Experiment 1) and in the two-talker masker (Experiment 2). The magnitude of the carrier-phrase benefit was similar for a given masker type across age groups, but the carrier-phrase benefit was greater in the presence of the two-talker masker than in multi-talker babble. Specifically, children’s average carrier-phrase benefit was 7.1% for multi-talker and 16.8% for the two-talker masker condition. No carrier-phrase benefit was observed for any age group in the presence of speech-shaped noise.

Correspondence concerning this article should be addressed to Angela Yarnell Bonino, Department of Allied Health Sciences, CB 7190, University of North Carolina, Chapel Hill, NC 27599. Fax: 919-966-0100. abonino@med.unc.edu.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conclusions—Effects of auditory masking on word recognition performance were greater for children than for adults. The time course of development for susceptibility to masking appears to be more prolonged for a two-talker speech masker than for multi-talker babble or speech-shaped noise. Unique to the current study, this work suggests that a carrier phrase can provide an effective auditory grouping cue for both children and adults under conditions expected to produce substantial perceptual masking.

Keywords

speech perception in noise; carrier phrase; auditory development; perceptual masking; auditory grouping

Introduction

Despite having a peripheral auditory system that is believed to provide the brain with an adequate representation of sound (reviewed by Werner & Leibold, 2004), children consistently require a more favorable signal-to-noise ratio (SNR) than adults for recognizing speech embedded in competing background sounds (e.g., Elliott et al., 1979; Fallon et al., 2000; Hall et al., 2002; Nittrouer & Boothroyd, 1990). For example, Nittrouer and Boothroyd (1990) observed that phoneme recognition in speech-shaped noise was similar between 4- to 6-year-olds tested at a +3 dB SNR and adults tested at a 0 dB SNR. Children also require a more favorable SNR for the identification of words in a competing speech background composed of two talkers (e.g., Hall et al., 2002) or multiple talkers (e.g., Elliott et al., 1979; Fallon et al., 2000). In light of children's increased susceptibility to masking, determining which acoustic cues improve performance in the presence of competing background sounds is a question of both clinical and theoretical importance. The objective of the current experiments was to determine if children's masked word recognition benefits from the inclusion of a carrier phrase (i.e., "say the word"). Specifically, carrier-phrase improvement was examined for several different competing backgrounds across three age groups: 5- to 7-year-olds, 8- to 10-year-olds, and adults. *Developmental Effects in Susceptibility to Masking*

This work was motivated, in part, by observations of pronounced child-adult differences as well as age-related changes across childhood for masking produced by competing speech (e.g., Fallon et al., 2000; Hall et al., 2002; Wightman & Kistler, 2005; Wightman et al., 2010; Wilson et al., 2010). Two speech maskers previously used to evaluate speech perception in children are multi-talker babble and two-talker speech. Multi-talker babble is often used in clinical measures, including the Hearing in Noise Test (HINT; Nilsson et al., 1994), the Quick Speech in-Noise Test (QuickSIN; Killion et al., 2004), the Bamford-Kowal-Bench Speech-in-Noise Test (BKB-SIN; Etym tic Research, 2005), and the Words in Noise (WIN; Wilson, 2003; Wilson et al., 2010; Wilson & McArdle, 2007). A typical multi-talker babble is comprised of six to 20 talkers, with individual talker streams considered to be unintelligible. Child-adult differences as well as differences across childhood have been reported for tasks using multi-talker babble (e.g., Elliott et al., 1979; Fallon et al., 2000; Paps0 & Blood, 1989; Wilson et al., 2010). For example, Fallon et al. (2000) reported that 5-, 9- and 11-year-old children required a more favorable SNR than adults when asked to point to a picture corresponding to the final word of a sentence presented in eight-talker babble. Compared to adults, children required an SNR advantage of 5, 3, and 2 dB for the group of 5-, 9-, and 11-year-olds, respectively. Larger developmental effects have been reported for maskers containing fewer talkers. For example, 5- to 10-year-olds tested by Hall et al. (2002) identified spondee words in the presence of a continuous two-talker masker using a four-alternative forced-choice picture pointing task. The average

child-adult difference was 7 dB for the two-talker masker. In contrast, the average child-adult difference was 3 dB for speech-shaped noise for the same listeners.

Speech maskers may be challenging for listeners of all ages, but particularly for children, because speech maskers are believed to produce substantial perceptual masking (e.g., Carhart et al., 1969) in addition to energetic masking (e.g., Fletcher, 1940). In this context, energetic masking refers to masking produced as the result of overlapping excitation patterns in the peripheral auditory system. In contrast, perceptual masking appears to be the result of limited or ineffective central auditory processing. The term “perceptual masking” was first used by Carhart et al. (1969) to refer to the elevated masked speech reception thresholds with a speech masker in comparison to speech-shaped noise. Specifically, Carhart et al. proposed that perceptual masking the masker, which may be particularly problematic when the signal and masker are similar (e.g., Brungart et al., 2001). The perceptual demand of speech maskers appears to be related to the number of talkers contained in the masker. For adults, this effect is most obvious for a two-talker masker, and is minimal once the masker contains as many as 10 talkers (Freyman et al., 2004). Note, however, that the effects of perceptual masking appear to persist in school-aged children for maskers containing more than 10 talkers, including a 12-talker (Elliott et al., 1979) and a 20-talker masker (Papso & Blood, 1989).

It has been suggested that perceptual masking is closely related to or identical to the phenomenon known as “informational masking” (e.g., Brungart et al., 2001; Freyman et al., 2004; Wightman & Kistler, 2005), which was originally observed for tonal stimuli. In the classic simultaneous informational masking paradigm (Neff & Green, 1987), the listener is asked to detect a pure-tone signal embedded in a masker complex composed of multiple pure-tones that are spectrally uncertain. The detrimental effects, referred to as informational masking, of multitonal maskers can be substantial for all age groups, even though experimental controls are employed to minimize energetic masking. Child-adult threshold differences in this paradigm can be as large as 50 dB (e.g., Oh et al., 2001). Moreover, improvements in susceptibility to masking have been observed with increasing age across childhood (e.g., Leibold & Neff, 2007; Oh et al., 2001; Wightman et al., 2003).

The underlying mechanism(s) responsible for children’s increased susceptibility to masking, for both tone detection and speech recognition, are not fully understood. Children’s difficulties cannot be fully explained by immaturity of the peripheral auditory system (reviewed by Werner & Leibold, 2004) or general inattentiveness (e.g., Viemeister & Schlauch, 1992; Wightman & Allen, 1992). For speech perception measures, it is possible that immature linguistic abilities affect performance (e.g., Fallon et al., 2002). Note, however, that children’s increased susceptibility to masking is not unique to speech stimuli. Indeed, mounting evidence suggests that children’s pronounced and prolonged susceptibility to masking under complex listening conditions reflects immature central auditory processes.

Sound Source Segregation

One possible explanation for children’s increased susceptibility to masking relative to adults is immature sound source segregation abilities. Sound source segregation is the process that listeners use to perceptually segregate the signal from the competing background sounds. Adults can use differences in stimulus features between the signal and masker to perform sound source segregation, including asynchronous onset, incoherence of dynamic stimulus properties, and differences in sound quality or timbre (reviewed by Bregman, 1993). Previous studies have shown that introducing cues thought to facilitate sound source segregation can assist adult listeners for tone detection in a random-frequency, multi-tonal complex (e.g., Kidd et al., 1994; Neff, 1995) or for speech recognition in a masker containing a small number of talkers (e.g., Brungart et al., 2001; Freyman et al., 2004;

Freyman et al., 1999). Note that only modest improvements in tone detection (e.g., Neff, 1995) or speech recognition (e.g., Freyman et al., 2004; Freyman et al., 1999) have been reported under conditions associated primarily with energetic masking, such as broadband or speech-shaped noise.

Although the data are limited, researchers have begun to evaluate whether introducing sound source segregation cues, shown to benefit adults, can improve children's tone detection (Hall et al., 2005; Leibold & Bonino, 2009; Leibold & Neff, 2007; Wightman et al., 2003) or word recognition (Garadat & Litovsky, 2007; Litovsky, 2005; Wightman & Kistler, 2005; Wightman et al., 2006; Wightman et al., 2010). These studies show mixed results in the degree to which children benefit from segregation cues that have been shown to improve adults' performance. One potential sound source segregation cue that school-aged children appear to be able to use to improve their tone detection is spectro-temporal coherence (Hall et al., 2005; Leibold & Bonino, 2009). Using the "multiple bursts" paradigm developed by Kidd et al. (1994), Hall et al. (2005) asked children and adults to detect a train of 1000-Hz tone bursts embedded in a masker train composed of 60-ms bursts, where each masker burst contained two pure tones. The frequencies of these tones were either randomly selected for each masker burst (multiple-bursts different), or they were held constant across masker bursts within each interval (multiple-bursts same). Children and adults demonstrated a release from masking when the masker's frequency changed randomly across bursts compared to when frequency was fixed within an interval. One explanation for these findings is that in the multiple-burst different condition listeners were able to form separate auditory streams for the signal and the masker, because the signal was spectrally coherent across time in contrast to the dynamic masker.

The current study using speech stimuli was motivated by the multi-tonal studies described in the previous paragraph (Hall et al., 2005; Leibold and Bonino, 2009). Given the observation that both children and adults benefited from the provision of a coherent tonal signal in the multiple-burst different condition, the prediction for the current speech recognition study was that the introduction of a carrier phrase (i.e., "say the word") prior to target words would similarly improve performance. Based on previous data with tones, this benefit was expected to be larger for maskers that produce primarily perceptual (as opposed to energetic) masking. One complicating factor in this analogy between spectro-temporal coherence for tonal and speech stimuli is that additional linguistic-cognitive factors may be involved when the stimuli consist of speech. Both children and adults show improved recognition of target words when listeners are provided contextual information from preceding words in the sentence (e.g., Nittrouer & Boothroyd, 1990; Fallon et al., 2002). Another phenomenon observed in the speech perception literature is talker normalization. Talker normalization refers to the process used by listeners to recognize speech despite wide variation of the acoustic properties of a given word across speakers (e.g., Ladefoged & Broadbent, 1957; Mullennix et al., 1989). For example, Ryalls and Pisoni (1997) observed that preschool-aged children had better monosyllabic word recognition when words were presented by the same talker than when the talker was varied across trials. The benefit of having a fixed talker can be substantial in speech-shaped noise, improving performance by approximately 20% for both adults (Mullennix et al., 1989) and children (Ryalls & Pisoni, 1997). Based on limited studies that have examined related effects with adult listeners in speech maskers (e.g., Brungart et al., 2001; Helfer & Freyman, 2009; Kidd et al., 2008), it would not be surprising if greater developmental effects for cognitive-linguistic factors were observed in a two-talker masker than speech-shaped noise.

Another consideration in thinking about the analogy between the multiple-burst (tonal) and speech paradigms is that both the target and the masker speech are spectrally dynamic, in comparison to the fixed-frequency signals used in tonal studies of spectro-temporal

coherence. Natural speech is characterized by dynamic spectral and temporal properties. However, the anatomy of the vocal tract limits the particular spectral profiles that can be produced, as well as the speed of transition between these configurations. Therefore, speech may not provide the same opportunities to capitalize on spectro-temporal coherence as multi-tonal stimuli, but those cues may be available in a modified form. For example, knowledge of the spectro-temporal characteristics of a talker's speech stream could reinforce auditory segregation of that stream. Adult listeners are able to build up auditory streams for speech when there are pitch differences between speakers or based on coherent amplitude modulation across frequency in speech stimuli (reviewed by Carlyon, 2004). Furthermore, work by Kidd et al. (1998) and Durlach et al. (2003) suggests that adults are able to follow predictable frequency changes, such as a pattern that rises in frequency, in tonal signals presented in random-frequency, multi-tonal maskers. If listeners are able to build up an auditory stream over time by relying on the spectro-temporal properties of a target talker's voice, then using a carrier phrase might assist the listener in identifying the target word presented at the end of the phrase.

A carrier phrase is routinely provided during audiometric speech testing and many commercially-available recordings also include a carrier phrase. For normal-hearing adults tested in quiet, previous work suggests that a carrier phrase may provide a small benefit (e.g., Gladstone & Siegenthaler, 1971; Lynn & Brotman, 1981) or no improvement in performance (e.g., Martin et al., 1962). However, it is possible that a carrier phrase is beneficial in more challenging listening conditions. Lynn and Brotman (1981) tested adults' ability to identify monosyllabic words beginning with voiceless stop consonants (i.e., /p/, /t/, or /k/) in the presence of speech-shaped noise with and without the carrier phrase "you will say." The mean carrier-phrase benefit was 10% for adults in these conditions at a 0 dB SNR. Evidence of a possible carrier phrase benefit in multi-talker babble for children comes from work by Markham and Hazan (2004). In that study, monosyllabic word recognition was measured in 20-talker babble for each of two conditions: (1) a single monosyllabic word presented in isolation or (2) a carrier phrase followed by a series of three target words. Many of the children and adults tested performed near ceiling even when the words were presented in isolation. As a consequence, the carrier-phrase benefit was only examined for listeners whose performance was in the bottom quartile. Of those listeners, the difference between the two conditions was small (1.4%). Note, however, that interpreting the carrier-phrase benefit for this study is difficult, since most listeners were at ceiling, and three target words were presented after the carrier phrase. The limited data available in the literature on the benefit of a carrier phrase in noise are inconclusive, and to our knowledge there has not been any investigation of the benefit of a carrier phrase under conditions expected to produce substantial perceptual masking.

This paper presents two experiments that were conducted to examine carrier-phrase benefit in adults and school-aged children. These experiments examined the benefit associated with providing a carrier phrase in the presence of three different maskers that were expected to produce different amounts of perceptual masking. Experiment 1 tested the hypothesis that listeners would show improved monosyllabic word recognition with a carrier phrase in the presence of a multi-talker masker. This masker is commonly used in the clinic and was expected to produce moderate perceptual masking. Experiment 2 used a within-listener design to compare performance in low (speech-shaped noise) and high (two-talker speech) perceptual masking conditions to test the hypothesis that carrier-phrase benefit was related to auditory stream segregation. It was predicted that the carrier-phrase advantage would be larger for the two-talker masker than for the noise masker; such a result would be expected if the carrier phrase provided an auditory grouping cue. It was also hypothesized for both experiments that age-related differences in susceptibility to masking would be observed. Moreover, we expected to see larger-child adult differences for the two-talker masker.

Experiment 1

The first experiment examined child-adult differences in both susceptibility to masking and in the benefit provided by a carrier phrase for monosyllabic word recognition in the presence of continuous multi-talker babble. Although the individual talkers in multi-talker babble are considered to be unintelligible, this masker appears to result in perceptual masking (e.g., Elliott et al., 1979; Fallon et al., 2000; Paps0 & Blood, 1989; Wilson et al., 2010). If the provision of a carrier phrase spoken by the same talker as the target word provides an auditory grouping cue, listeners may show improved word recognition in multi-talker babble. Documenting age-related changes in susceptibility to multi-talker babble and carrier-phrase benefit are also clinically relevant, since multi-talker babble is often used in audiometric assessments, and few normative data are available.

Materials and Methods

Listeners—Eight 5- to 7-year-old children (mean=6:6 years:months (yrs:mos), SD=1:2 yrs:mos), eight 8- to 10-year-old children (mean=9:8 yrs:mos, SD=1:0 yrs:mos), and 10 adults (18 to 26 years of age, mean=21:0 yrs:mos, SD=2:3 yrs:mos) participated in this experiment. The rationale for testing a broad age range was because children appear to achieve adult-like performance at different ages based on the complexity of the masker (e.g., Hall et al., 2002; Leibold & Neff, 2007; Wightman et al., 2010). Whereas children older than about 7 years of age often perform like adults on speech tasks in speech-shaped noise (e.g., Nishi et al., 2010; Stelmachowicz et al., 2007), immature performance has been reported for adolescents for tasks using a speech masker (e.g., Wightman & Kistler, 2005; Wightman et al., 2010). All listeners had thresholds in quiet of 20 dB HL bilaterally for octave frequencies from 0.25 to 8 kHz (ANSI, 2004). Speech and language skills were not formally assessed. All listeners were native speakers of English and had no known history of chronic ear disease. Listeners were tested individually in a single-walled, sound-treated room for approximately one hour with regular breaks. This research was approved by the institutional review board at The University of North Carolina at Chapel Hill.

Stimuli and conditions—Speech targets were Phonetically-Balanced Kindergarten words (PBK; Haskins, 1949). These words were selected based on kindergarteners' spoken vocabulary, and typically developing 5- to 7-year-olds can achieve 96 to 98% correct in quiet (Sanderson-Leepa & Rintelmann, 1976). The commercial recordings of the three PBK word lists (150 words in total) from Auditec of St. Louis were used. These three word lists have similar lexical properties, including measures of word frequency, lexical neighborhood frequency and density (Meyer & Pisoni, 1999). As is standard for many recordings of clinical word recognition (reviewed by Gelfland, 2009), each target PBK word was recorded by a male speaker with the carrier phrase "say the word" prior to the target word.

The two carrier-status conditions were created using the sound editing software Audacity (version 1.2.6). For the carrier-present condition, the carrier phrase "say the word" prior to each of the 150 target words was left intact. Thus, any potential coarticulation between the carrier phrase and the target word was preserved from the original recordings. The inter-stimulus interval was approximately 3.5 s for all carrier-present targets. The isolated words for the carrier-absent condition were obtained by replacing the carrier phrase prior to each target word with silence. Replacing the approximately 0.8-s carrier phrase with silence resulted in an inter-stimulus interval of approximately 4.3 s. This editing allowed for the PBK target word to be presented at the same time intervals across the two carrier-status conditions. After editing, all target words for both carrier-status conditions were verified independently by two experienced listeners to ensure that each word was intelligible and

free of audible distortion. Word lists were exported as WAV files, with a 16-bit resolution and 44.1 kHz sampling rate.

Target words were presented at 65 dB HL, computed based on the mean peak intensity level of the carrier phrase. Considerable variability in both peak and average RMS levels were observed across the 150 target PBK words. This variability appears to reflect the original recording procedure, where the speaker monitored a volume unit (VU) meter and attempted to peak the meter to 0 VU while saying the carrier phrase and then allowed the target word to “fall” naturally. For the 150 target words in isolation, the peak intensity spanned a range of 9.2 dB (SD=1.9 dB), and the RMS values spanned a range of 7.2 dB (SD=1.6 dB). Similar values for both mean peak and RMS levels were observed across the lists of 50 target words used in this experiment. Target words were not rescaled so as to preserve the level variability of the clinically-used Auditec recordings.

The masker was a continuous multi-talker babble. The multi-talker masker was created by summing 20 recordings of young adults reading different passages, and it is commercially available through Auditec of St. Louis. Different masker levels were used to test children and adults. Both groups of children were tested at an SNR of +10 dB, created by presenting the multi-talker babble at 55 dB HL. For adults, the multi-talker babble was presented at 60 dB HL, resulting in a +5 dB SNR. Based on extensive pilot data (at 0, +5 and +10 dB SNR), these SNRs were selected to roughly equate performance across children and adults, while avoiding ceiling and floor effects.

Listeners completed a list of 50 words for each carrier-status condition. Different target words were used for the two conditions. Prior to the start of each condition, 10 practice words were completed using words from List 2A of the Northwestern University Auditory Test No.6 (NU-6) recordings (Tillman & Carhart, 1966). Practice words did not duplicate any of the PBK target words and were spoken by the same male talker who produced the test words. During practice, listeners were encouraged to give responses and were reinstructed as needed. Following practice, each listener completed two word lists, with carrier-status fixed within a word list. Two of the three available word lists were randomly assigned a carrier-status condition and testing order across listeners. For a given list, the target words were presented in a fixed order, consistent with the commercially available lists.

Stimuli were presented via a two-channel audiometer (Grason-Stadler GSI 61; Eden Prairie, MN). The target words and continuous maskers were recorded on different compact disks (CDs) and were routed to independent channels of the audiometer through separate CD players. Both channels were calibrated prior to each testing session using a 1000-Hz calibration tone. The stimuli were presented to the left ear by a TDH-50P headphone.

Procedure—Listeners were seated inside a single-walled sound booth. An experimenter sat in front of the listener in the booth and was the primary coder. A secondary coder was located in the adjacent control room, monitoring the listener’s responses through the audiometer’s talk-back system. Both coders recorded the listener’s responses. In order for a target word to be scored as correct, the listener had to repeat the entire word correctly.

Assessing performance—Percent correct recognition performance was calculated based on all target words for each condition. Carrier-phrase benefit (or release from masking) was operationally defined as the difference in performance between the carrier-present and carrier-absent conditions. For all statistical analyses, percent correct scores were converted to rationalized arcsine units (Studebaker, 1985) to counteract non-uniformity of variance. Note, however, that performance is discussed as percent correct in both the text and the

figures. All statistical analyses were completed in SSPS (version 18.0) using a criterion of $\alpha=0.05$. The p -values for all post-hoc analyses incorporated Bonferroni adjustments.

Coder reliability—Coder reliability was computed for each test session.¹ Inter-rater reliability was calculated by examining the point-by-point agreement between the two coders. In order to be considered an agreement, both coders had to mark the response as correct, or both coders had to mark the response as incorrect.

Results

Coder reliability—Across conditions, words lists, and listener age groups, excellent average reliability (>90%) was observed between the two coders. Furthermore, coder reliability was never less than 80% for any listener. Due to the high reliability observed between coders, the scores used for statistical analysis and presentation of results are taken from the primary coder, located inside of the booth.

Group differences—Group average results are summarized in Figure 1. Mean percent correct is shown as a function of listener age group for the carrier-absent condition, represented by open squares, and for the carrier-present condition, represented by filled circles. Error bars indicate ± 1 standard error (S.E.) of the mean. Age-related changes in performance are suggested in Figure 1, with 5- to 7-year-olds having poorer word recognition scores than 8- to 10-year-olds. Despite age-related differences in performance, the average carrier-phrase benefit appears to be similar across listener age groups. The mean carrier-phrase benefit was 8.5% for 5- to 7-year-olds, 5.8% for 8- to 10-year olds, and 6.2% for adults.

A repeated measures analysis of variance (ANOVA) was used to test the trends observed in Figure 1. The analysis included the within-subjects factor of Carrier Status (carrier-present, carrier-absent) and the between-subjects factor of Age Group (5- to 7-year-olds, 8- to 10-year-olds, adults). Both of the main effects were significant: Carrier Status [$F(1,23)=20.87$, $p<0.001$] and Age Group [$F(2,23)=7.11$, $p=0.004$]. However, the Carrier Status \times Age Group interaction was not significant [$F(2,23)=0.28$, $p=0.76$]. To further examine the significant main effect of Age Group, post-hoc contrasts were performed. Averaging across the two carrier-phrase conditions, 5- to 7-year-old children performed significantly more poorly than 8- to 10-year-old children ($p=0.02$). Additionally, 5- to 7-year-old children at a +10 dB SNR performed more poorly than adults at a +5 dB SNR ($p=0.005$). However, at these SNRs, 8- to 10-year-olds' performance was equated to that of adults ($p=1$).

Individual differences—Recognition performance for individual listeners in the carrier-absent (open squares) and the carrier-present (filled circles) conditions is shown in Figure 2, plotted by listener age. The vertical lines indicate the carrier-phrase benefit for individual listeners. The carrier-phrase benefit ranged from -6 to 18%, -8 to 18% and 4 to 18% for individual 5- to 7-year-olds, 8- to 10-year-olds, and adults, respectively. In general, the individual data are consistent with the trends observed in the group data. All but three children (ages: 6:10, 9:5, and 10:11 yrs:mos) benefited from the inclusion of a carrier phrase. Interestingly, the two older children had the best performance (>70%) in the carrier-absent condition across all children. Seven adults benefited from the carrier phrase, and three adults did not; recall that adults were tested at a +5 dB SNR.

¹Using a testing procedure at a fixed-level allowed coder reliability to be compared after the completion of a testing session, rather than requiring the coders to agree before the next trial could be presented in an adaptive procedure.

Discussion

Age effects on susceptibility to multi-talker babble—Consistent with previous studies using multi-talker babble (e.g., Elliott et al., 1979; Fallon et al., 2000; Papso & Blood, 1989; Wilson et al., 2010), child-adult differences as well as age-related improvements during childhood in susceptibility to masking were observed. In this experiment, children were tested at a +10 dB SNR and adults at a +5 dB SNR. A 5-dB-SNR advantage resulted in similar performance for 8- to 10-year-old children and adults. However, 5- to 7-year-old children demonstrated significantly poorer average performance than either adults tested at a +5 dB SNR or 8- to 10-year-old children tested at a +10 dB SNR. Wilson et al. (2010) also reported age-related differences during childhood for identifying monosyllabic words in six-talker babble. For the age groups tested in that study (6 to 13 years), the 6-year-old group required a significantly higher SNR to achieve 50% accuracy than any other child age group. Specifically, the SNR at which children identified 50% of the words was 10 dB SNR for 6-year-olds and 6 dB SNR for 9-year-olds. This is roughly consistent with the age effect observed for children in the present experiment.

Carrier-phrase benefit—Despite age-related differences in susceptibility to masking, all three age groups showed a similar benefit with the inclusion of a carrier phrase. The average carrier-phrase benefit was 8.5%, 5.7%, and 6.2% for 5- to 7-year-olds, 8- to 10-year-olds, and adults, respectively. Testing the adults at a more challenging SNR than children may have allowed seven of the 10 adults to benefit from the provision of the carrier phrase. If adults had been tested at the same SNR as the children (+10 dB SNR), based on our pilot data, it is possible that many adults would have performed at or near ceiling in the carrier-absent condition. In order to maximize the observable carrier-phrase benefit, listeners' word recognition needed to be poor enough to permit improvement with the addition of the cue.

Consistent with previous studies that suggest a carrier-phrase benefit in noise (Lynn & Brotman, 1981; Markham & Hazan, 2004), most listeners obtained a carrier-phrase benefit in Experiment 1. This finding provides evidence that a carrier phrase improves speech recognition performance in the presence of multi-talker babble. However, the underlying mechanism responsible for the carrier-phrase benefit is not clear. One explanation is that the carrier phrase provides an auditory grouping cue, assisting listeners under conditions where the listener has difficulty separating the target word from the competing background sounds. If the carrier phrase is an auditory grouping cue, it is likely that the benefit seen in Experiment 1 is dependent on the use of a masker that produces sufficient perceptual masking. However, the multi-talker masker used in Experiment 1 contained 20-talkers. Freyman et al. (2004) have previously shown that the amount of perceptual masking for adults is greater for maskers containing fewer talkers. Thus, carrier-phrase benefit may be greater for a masker composed of a smaller set of talkers.

Experiment 2

The goal of the second experiment was to test the hypothesis that a carrier phrase provides an auditory grouping cue that helps the listener segregate the target from the masker under conditions of relatively high perceptual masking. In order to test this hypothesis, children and adults were asked to repeat monosyllabic words embedded in a continuous two-talker masker or speech-shaped noise. The two-talker masker was expected to produce both perceptual and energetic masking, whereas speech-shaped noise was expected to produce primarily energetic masking. Comparing performance across the two maskers was of interest because sound source segregation cues are likely to improve recognition in cases where performance is limited by perceptual masking, but not when performance is limited by energetic masking (e.g., Freyman et al., 1999). The expected finding for Experiment 2 was

that a substantial carrier-phrase benefit would only be observed in the two-talker masker condition.

Materials and Methods

Many of the methods were consistent across the two experiments. Details that are different from those of Experiment 1 are presented here.

Listeners—Nine 5- to 7-year-old children, nine 8- to 10-year-old children, and 16 adults participated in this experiment. None of these listeners had previously participated in Experiment 1. The mean ages were 6:6 yrs:mos (SD=0:9 yrs:mos) for the 5- to 7-year-old children and 9:7 yrs:mos (SD=0:9 yrs:mos) for the 8- to 10-year-olds. Three additional children were excluded: two due to experimenter error (ages: 5:4 and 10:3 yrs:mos) and one because no responses were given by this listener during practice (5:2 yrs:mos). Two groups of adults were tested. The first group of nine adults (18 to 28 years; mean=22:6 yrs:mos, SD=3:6 yrs:mos) was tested at the same SNRs as the children. The second group of seven adults (20 to 30 years; mean=23:4 yrs:mos, SD=3:3 yrs:mos) was tested at a harder SNR, to approximately equate performance with children. The same eligibility criteria were used as in Experiment 1. However, in this study one child (5:3 yrs:mos) was included who had two thresholds of 25 dB HL in the non-test ear.

Stimuli, conditions, and procedure—Each listener was tested in two separate masker conditions: (1) two-talker speech and (2) speech-shaped noise. The two-talker masker consisted of continuous, meaningful speech presented by two different male voices. The talkers were reading different passages from a series of fantasy novels for children. One sample was 4 min and 17 s, and the other was 7 min and 47 s in duration. Both samples were manually edited to remove silent pauses greater than approximately 300 ms and were scaled to equal RMS level. The 20-minute masker on the CD was the sum of the two streams, each composed of copies of the single sample concatenated head-to-tail. Each sample ended with a complete sentence, so there was no auditory or syntactic discontinuity when the samples were repeated. Listeners were also tested in the presence of speech-shaped noise. The spectrum of the noise masker was matched to the spectrum of the two-talker masker. This was achieved by passing Gaussian noise through an FIR filter with 2^{10} taps. A 20-minute sample of speech-shaped noise was recorded to a CD.

For all children and the first group of adults, the two-talker masker was presented at 55 dB HL (+10 dB SNR) and speech-shaped noise was presented at 60 dB HL (+5 dB SNR). For the second group of adults, tested at a harder SNR, both the two-talker and speech-shaped noise masker were presented at 65 dB HL (0 dB SNR). Stimuli were presented via an audiometer, as described for Experiment 1, to the right ear by an insert earphone (ER3A, Etym tic).

The same target words were used as in Experiment 1; however, they were concatenated into four lists of 35 words. The first 35 words from each of the three lists used in Experiment 1 were used to create three of the word lists for this experiment. To create the fourth list, 35 of the remaining words from each of the three 50-word lists were combined. The mean peak and RMS levels across the four word lists were comparable, with the range not exceeding 1 dB between lists. There were 10 PBK words that were not included in the four lists: these words served as practice. Practice words were presented in the context of a carrier phrase, and the masker was reduced by 5 dB. To allow for additional practice at the test SNR, the first 5 words of each test word list were not scored. Thus, performance was based on 30 target words for each condition. As in Experiment 1, target words were presented in a fixed

order for each list, whereas test order and the assignment of a given word list to a condition were randomized across listeners.

Results

Average reliability scores for all age groups, conditions, and maskers, were 90% or better. Reliability scores across the two coders were 80% or better for individual listeners, except for two children. These two children (ages: 6:5 and 7:2 yrs:mos) each had one condition where the coder agreement was only 76.7%. These cases may be a byproduct of young children having a higher degree of within-subject variability in speech production (e.g., Lee et al., 1999) and/or that the second experimenter was listening through the monitoring system of the audiometer in the adjacent control room. However, because of limited available data in the literature, it is not clear what the normal range of inter-rater reliability is for an open-set task that elicits children's speech productions while listening to a masker. As in Experiment 1, the scores generated by the primary coder inside of the booth were used for analyses.

Group differences—Mean results for Experiment 2 are summarized in Figure 3. The two panels show performance for speech-shaped noise (left) and the two-talker masker (right). For each panel, mean percent correct is shown for carrier-present (filled circle) and carrier-absent (open squares) conditions for each age group. Error bars indicate ± 1 S.E. of the mean. Note that the three leftmost points show data for children and adults tested at the same SNR, which was +10 dB SNR for the two-talker masker and +5 dB SNR for speech-shaped noise. Data from the group of adults tested at a 0 dB SNR are shown at the far right in each panel.

Age differences in susceptibility to masking—The first set of analyses examined developmental effects in susceptibility to masking. Performance was compared across the three age groups tested at the same SNR (+5 dB SNR for speech-shaped noise and +10 dB SNR for the two-talker masker) in the carrier-absent condition. Susceptibility to masking was not examined for the carrier-present condition because the interpretation of the child-adult difference would be confounded by ceiling effects for the adults. A repeated-measures ANOVA was conducted with the within-subjects variable of Masker (two-talker, speech-shaped noise) and the between-subjects variable of Age Group (5- to 7-year-olds, 8- to 10-year-olds, SNR-matched adults). Results indicated a non-significant main effect of Masker [$F(1,24)=1.24, p=0.277$], a significant main effect of Age Group [$F(2,24)=24.06, p<0.001$], and a significant Masker \times Age Group interaction [$F(2,24)=5.03, p=0.015$].

To further examine the significant Masker \times Age Group interaction, a pair of Univariate ANOVAs were performed, with one for each masker. It was confirmed that performance in the carrier-absent condition differed significantly across Age Group for both speech-shaped noise [$F(2,24)=9.91, p=0.001$] and the two-talker masker [$F(2,24)=17.29, p<0.001$]. Pairwise comparisons, based on the estimated marginal means with Bonferroni adjustments, were then conducted to examine performance across Age Group for each masker. For speech shaped noise, 5- to 7-year-olds performed more poorly than both 8- to 10-year-olds ($p=0.001$) and adults ($p=0.004$). No significant difference in performance was observed between 8- to 10-year-olds and adults in the noise masker ($p=1.0$). A different pattern of results was observed for the two-talker masker. For the two-talker masker, both 5- to 7-year-olds ($p<0.001$) and 8- to 10-year-olds ($p=0.006$) performed more poorly than the adult listeners in the carrier-absent condition. In contrast to speech-shaped noise, no significant difference in performance was seen between the two age groups of children ($p=0.07$). This finding suggests that the developmental time course of word recognition is more prolonged for the two-talker masker than for speech-shaped noise.

Carrier-phrase benefit—An important goal of this experiment was to compare carrier-phrase benefit between speech-shaped noise and the two-talker masker, and between the three listener age groups. In this analysis, children's data were compared to the group of adults who were tested at 0 dB SNR. The rationale for this approach was that many of the adults tested using a +10 dB SNR were at or near ceiling in the carrier-absent condition, potentially reducing the opportunity to observe a benefit of the carrier phrase. In contrast, the performance of adults tested at a 0-dB SNR was roughly equated to that of children.

A repeated-measures ANOVA was used to examine carrier-phrase benefit. For this analysis, carrier-phrase benefit was calculated as the difference in performance (in rationalized arcsine units) for the carrier-present and the carrier-absent conditions. The analyses included the within-subject factor of Masker (two-talker, speech-shaped noise) and the between-subject factor of Age Group (5- to 7-year-olds, 8- to 10-year-olds, adults at 0 dB SNR). The main effect of Masker was significant [$F(1,22)=32.81$, $p<0.001$], whereas neither the main effect of Age Group [$F(2,22)=1.75$, $p=0.2$] nor the Masker \times Age Group [$F(1,22)=3.09$, $p=0.07$] were significant. The significant main effect of Masker confirmed the trend, as seen in Figure 3, that all three listener age groups had substantial carrier-phrase benefit in the two-talker condition, but not for speech-shaped noise. The average carrier-phrase benefit in the two-talker masker condition was 10.2% for 5- to 7-year-olds, 21.5% for 8- to 10-year-olds and 22.9% for the adults (at 0 dB SNR). Despite the appearance of a developmental trend for carrier-phrase benefit in Figure 3, there are not statistically significant differences in the ability to benefit from the carrier phrase between the three listener age groups. Moreover, despite children being more susceptible to a two-talker masker than adults (SNR-matched), they are able to benefit from the carrier phrase to the same extent as adults (approximately performance-matched).

Individual differences in carrier-phrase benefit—In agreement with previous studies of perceptual masking (e.g., Wightman & Kistler, 2005; Wightman et al., 2006), large individual differences in performance were observed in the two-talker masker condition. Individual listeners' percent correct scores in the carrier-absent (open squares) and the carrier-present (filled circles) conditions for the two-talker masker are plotted in Figure 4 as a function of age. Adult data are those from the 0-dB SNR conditions. The vertical lines indicate the release from masking associated with the addition of a carrier phrase for each listener. Carrier-phrase benefit ranged from -3.3 to 33.3% for 5- to 7-year-olds, and from 0.0 to 53.3% for the 8- to 10-year-olds. All except four children showed carrier phrase benefit, ages: 5:3, 6:1, 7:2, and 10:11 yrs:mos. Note that the 10:11-year-old's performance in the carrier-absent condition was similar to SNR-matched adults', and this child may have been at ceiling. All seven adults tested at 0 dB SNR showed a carrier-phrase benefit, which ranged from 6.7 to 50% across listeners.

In contrast to the two-talker masker condition, some listeners performed better in the absence of a carrier phrase in the speech-shaped noise condition. Of the nine 8- to 10-year-olds tested: one child had a positive carrier-phrase benefit (9%), two children had 0% benefit, and six children had negative benefit scores (maximum score was -25%). Negative scores were also seen for some of the 5- to 7-year-olds and adults. It is unclear why some listeners would do better in the absence of a carrier phrase. However, in light of this finding not being statistically significant, observing negative values may be a result of variability of performance around 0%.

Discussion

Age effects in susceptibility to masking—Child-adult differences in susceptibility to masking were observed in the presence of speech-shaped noise. The difference in mean

performance for 5- to 7-year-olds and SNR-matched adults was 16.7% in the carrier-absent condition. In contrast, 8- to 10-year-olds performed similarly to adults. Thus, children's ability to recognize words in speech-shaped noise appears to reach mature levels during the age span of children tested for the current study. This time course of development is consistent with previous studies showing developmental differences in speech perception in the presence of filtered noise during childhood (e.g., Allen & Wightman, 1992; Hall et al., 2002; Nishi et al., 2010, Ryalls & Pisoni, 1997). Results from this experiment are also consistent with the previously reported developmental differences for PBK word recognition in the presence of speech-shaped noise (e.g., Lewis et al., 2010; Stelmachowicz et al., 2007). Specifically, Lewis et al. (2010) reported improved performance across the age range of 5 to 7 years, whereas Stelmachowicz et al. (2007) reported that PBK word recognition performance was stable for 7- to 14-year-old children with normal hearing. Together with results from the current experiment, these findings suggest that by the age of 8 to 10 years, children's word recognition in speech-shaped noise is adult-like.

In contrast to noise, more complex maskers believed to create informational or perceptual masking have been associated with prolonged immaturities for both tone detection (e.g., Leibold & Neff, 2007; Wightman et al., 2003) and speech recognition tasks (e.g., Hall et al., 2002; Wightman & Kistler, 2005). Previous work from Hall et al. (2002) suggests that there was no change in thresholds for recognizing spondees in a two-talker masker across the age span of 5 to 10 years. Consistent with the findings from Hall et al., results from Experiment 2 showed similar performance for 5- to 7-year-olds and 8- to 10-year-olds for an open-set word recognition task in a two-talker masker. Average percent correct in the carrier-absent condition was 47.5% for 5- to 7-year-olds and 62.2% for 8- to 10-year olds. Both groups of children were significantly poorer at recognizing words than the group of SNR-matched adults, who had an average performance of 81.9% correct in the two-talker masker condition. These findings indicate substantial child-adult differences; however, in contrast to results with speech-shaped noise, performance for the two-talker masker condition did not significantly improve between the two age groups of children.

Carrier-phrase benefit—For children and adults, substantial carrier-phrase benefit was observed for the two-talker masker (high perceptual masking), and no systematic carrier-phrase benefit was seen for speech-shaped noise (low perceptual masking). The average carrier-phrase benefit was 17.8% for the two-talker masker and -2% for speech-shaped noise, when performance was collapsed across the three listener age groups. This pattern of performance across the two maskers is consistent with the carrier phrase being an auditory grouping cue. Or in other words, the acoustic information provided by the carrier phrase improves word recognition under conditions where the listener has a difficult time separating the target word from the masker. Interestingly, children as young as 5- to 7-years of age were able to benefit from the carrier phrase as effectively as 8- to 10-year-old children and adults.

Conclusions

Two groups of children (5- to 7-year-olds and 8- to 10-year olds) and adults were tested on a word recognition task in three different competing maskers: speech-shaped noise, multi-talker babble, and a two-talker speech. As detailed in the discussion section of each experiment, significant child-adult differences were observed for all three maskers for 5- to 7-year-old children. In contrast, 8-10 year old children performed similarity to adults in speech-shaped noise, yet this age group had immature performance for the speech maskers. Consistent with previous studies (e.g., Hall et al., 2002; Wightman & Kistler, 2005; Wightman et al., 2010), these findings suggest that the developmental time course of masked word recognition is longer for more complex maskers. The second objective was to

determine if listeners benefited from a carrier phrase under conditions expected to produce perceptual masking.

Potential Mechanisms Responsible for the Carrier-Phrase Benefit

Unique to the current study, this work tested the hypothesis that a carrier phrase provides an auditory grouping cue in the presence of maskers expected to produce substantial perceptual masking. Consistent with this hypothesis, carrier-phrase benefit differed across the three maskers, which were believed to vary in the amount of perceptual masking they produced. Word recognition improved when a carrier phrase was provided for both speech masker conditions, with a trend for greater benefit with the two-talker masker than multi-talker babble. Across all child listeners, the average carrier-phrase benefit was 7.1% in multi-talker babble and 16.8% in the two-talker masker. In contrast to the speech maskers, listeners did not benefit from the provision of the carrier phrase in the speech-shaped noise condition. These findings are consistent with previous speech studies of adult listeners, which have reported differences in the amount of benefit from an auditory grouping cue based on the perceptual-demand of the masker (e.g., Freyman et al., 2004; Freyman et al., 1999).

Although the carrier phrase improves performance in the present experiments, it is unclear what feature(s) of the carrier phrase are responsible for this effect. As discussed in the introduction section, one possible explanation is that the carrier-phrase effect is related to cognitive-linguistic factors that assist listeners in speech comprehension. Although there are limited data under conditions of high perceptual masking, findings from Kidd et al. (2008) suggest that adults' word recognition in the presence of a single stream of competing speech is better when the target sentence is syntactically correct compared to when words are randomly combined to form a target sentence. In that study, energetic masking was prevented by presenting two sentences spoken in a sequential, interleaved-word format, with one sentence being the target and the other the masker. Results from Brungart et al. (2001) and Helfer and Freyman (2008) also indicate that adults have better target word identification in speech maskers when they are provided a prime of the listener's voice prior to the trial. Thus, it is plausible that linguistic-cognitive factors would play a role in speech perception in the context of speech maskers. However, it is unlikely that linguistic-cognitive effects contributed substantially to this study's findings, considering that other studies have found large effects of cognitive-linguistic factors in speech-shaped noise (e.g., Mullennix et al., 1989; Nittrouer & Boothroyd, 1990; Ryalls & Pisoni, 1997) and a significant carrier-phrase benefit was not observed for speech-shaped noise in Experiment 2.

The pattern of current results is consistent with the idea that the carrier phrase allows listeners to build up an auditory stream based on the common spectral fluctuations of a target talker's voice that occur over time. Evidence of listeners being able to use a spectro-temporally coherent signal to improve performance comes from informational masking studies using a "multi-burst different" paradigm with tonal stimuli (e.g., Hall et al., 2005; Kidd et al., 1994; Leibold & Bonino, 2009). In the multi-burst different condition, listeners are asked to detect a fixed-frequency signal embedded in a random-frequency, multi-tonal masker. Providing multiple bursts of the signal tone results in improved tone detection (Kidd et al., 1994; Leibold & Bonino, 2009). This manipulation increases the spectro-temporal coherence of the signal, assisting the listener in separating the spectrally-fixed target from the spectrally-uncertain masker stream. For the speech maskers in the current study, the observed carrier-phrase benefit might likewise reflect improved spectro-temporal coherence for the target word. Giving the listener additional exposure to the talker's common spectral fluctuations may facilitate separation of the signal from the competing speech background. It may also be that the additional time and information provided by the carrier phrase increases the listener's ability to attend to common properties of the target talker's voice. Support for this idea comes from work by Kidd et al. (2008), in which adults heard two sentences (one

the signal and one the masker) in an interleaved-word format as discussed above. Kidd et al. showed that performance improved for later occurring words in nonsense sentences when target words were presented by the same talker, suggesting that the listener is able to have benefit more from “linkage variables” (e.g., fixed talker, fixed perceived interaural location, and correct syntactic structure) as the number of words increases.

Another possible mechanism to consider is coarticulation. There may have been information about the identity of the target word present in the carrier phrase due to anticipatory coarticulation (Fowler, 2005), such that excising the carrier phrase removed information that could have facilitated recognition of the target. Additionally, it is also possible that coarticulation influences the acoustic representation of the target word itself. Previous research suggests that listeners are able to compensate for the acoustic changes in a target phoneme when the preceding phoneme is provided (e.g., Mann, 1980; Lotto & Holt, 2006). In the absence of the carrier phrase, however, acoustic modifications due to coarticulation may have had a detrimental effect on speech recognition. There are two lines of evidence that argue against coarticulation effects being responsible for the carrier-phrase benefit. First, in a previous study from our lab (Bonino & Leibold, 2007), PBK target words were presented monitored lived-voice by a female talker in the presence of the same multi-talker babble used in Experiment 1. In this study, target words in the carrier-absent condition did not have any effects of coarticulation from a carrier phrase. Consistent with the above findings using recorded target words, results from 18 children (4 to 10 years of age) and eight adults demonstrated substantial mean carrier-phrase benefit, ranging from 5 to 6% at a 0 dB SNR. The second line of evidence is that a carrier-phrase benefit was not seen in the speech-shaped noise condition. Given that previous studies of coarticulation have reported effects even in the absence of auditory masking (e.g., Mann, 1980), the absence of a carrier-phrase effect in the speech-shaped noise condition suggests a limited role of co-articulation.

The final feature of the carrier phrase that might result in improved word recognition is that the carrier phrase indicates when in time to listen for the target word. The effect of a cue indicating when in time to listen for a tone appears to be relatively small in both quiet and noise (e.g., Egan et al., 1961; Green & Weber, 1980). In contrast to these findings, recent work, also with non-speech stimuli, has suggested that knowing when in time to listen is an effective cue for conditions that produce substantial informational masking (e.g., Best et al., 2007; Bonino & Leibold, 2008). For example, Best et al. (2007) found that listeners could more accurately identify a target birdsong embedded in other birdsongs when provided with a cue indicating the time interval that contained the target. In a related study, Bonino and Leibold (2008) found that a cue indicating when in time to listen for a pure-tone signal was 7 dB greater in a random-frequency, two-tone masker than for broadband noise. It is likely that listeners also benefit from knowing when in time to listen for speech stimuli under conditions expected to produce perceptual masking, a hypothesis that we are currently investigating in the laboratory.

Clinical Implications

In addition to the theoretical implications of this work, these results may have relevance to clinical practice. There has been a recent interest in clinical audiology to capture “real-world listening” performance by assessing speech understanding in noise (e.g., McArdle & Wilson, 2008). However, many of the word recognition measures used in clinics to assess speech understanding in noise were normalized in quiet. For example, standard word recognition tests normalized in quiet, such as the NU-6 (Tillman & Carhart, 1966), PBK (Haskins, 1949), and CID W-22 (Hirsh et al., 1952), are now commercially available from Auditec with multi-talker babble on the second channel of the CD. Indeed there are few normative data for these tests in multi-talker babble, especially for children. One exception is the WIN Test, which uses the NU-6 words (Wilson, 2003; Wilson et al., 2010; Wilson &

McArdle, 2007). Results of the present experiments suggest that caution needs to be exercised as clinicians attempt to use speech in noise measures for the pediatric population. As demonstrated in Experiments 1 and 2, a prolonged developmental course is observed for complex listening conditions. Moreover, extensive variability has been observed within and across age groups of children with normal hearing. Until further data are collected from children with normal hearing, it is difficult to know what is “normal” on these tasks. Moreover, children with limited or abnormal listening experience may show a delayed or different developmental pattern.

In addition to age-related changes in performance, differences in testing procedures can influence results. Historically, the production of a carrier phrase was viewed as a convenient method for the audiologist to “calibrate” his/her voice during the monitored live-voice mode of presentation. Although the variability concerns associated with presenting speech audiometry measures in monitored-live voice mode have been well documented (e.g., Brandy, 1966), there are clinicians who continue to use it for certain situations (reviewed by Madell, 2008). For clinicians who test using live-voice procedures, the carrier phrase may be dropped in an effort to reduce the administration time. However, results of the present experiments suggest that a carrier phrase actually improves performance in complex listening conditions, unlike the limited improvement typically reported for listening in quiet (e.g., Martin et al., 1962). This finding further supports the need to use standard recording procedures and testing protocols in clinical practices.

Acknowledgments

This work was supported by the American Academy of Audiology/American Academy of Audiology Foundation Research and NIH NIDCD grants R03DC008389, R01DC011038 and F31DC010308. We thank the members of the Human Auditory Development Lab for data collection assistance, particularly Ashley Halbach, Jack Hitchens, and Caitlin Rawn. Thanks are also extended to Andrew Lotto for his helpful comments on a previous version of this manuscript.

Sources of financial support include the American Academy of Audiology/American Academy of Audiology Foundation Research and the National Institutes of Health (NIDCD)

REFERENCES

- Allen P, Wightman FL. Spectral pattern discrimination by children. *Journal of Speech and Hearing Research*. 1992; 35:222–233. [PubMed: 1735972]
- ANSI. *Methods for Manual Pure-tone Threshold Audiometry*. American National Standards Institute; New York: 2004. ANSI S3.21-2004
- Best V, Ozmeral EJ, Shinn-Cunningham BG. Visually-guided attention enhances target identification in a complex auditory scene. *Journal of the Association for Research in Otolaryngology*. 2007; 8(2): 294–304. [PubMed: 17453308]
- Bonino, AY.; Leibold, LJ. Poster presented at the Sound Foundation through Early Amplification. Chicago, IL: 2007. Children’s and adults’ performance for PBK words in noise: Effect of spectro-temporal coherence.
- Bonino AY, Leibold LJ. The effect of signal-temporal uncertainty on detection in bursts of noise or a random-frequency complex. *Journal of the Acoustical Society of America*. 2008; 124(5):EL321–EL327. [PubMed: 19045685]
- Brandy WT. Reliability of voice tests of speech discrimination. *Journal of Speech and Hearing Research*. 1966; 9:461–465.
- Bregman, AS. Auditory scene analysis: Hearing in complex environments. In: McAdams, S.; Bigand, E., editors. *Thinking in Sound: The Cognitive Psychology of Human Audition*. Oxford University Press; New York: 1993.

- Brungart DS, Simpson BD, Ericson MA, et al. Informational and energetic masking effects in the perception of multiple simultaneous talkers. *Journal of the Acoustical Society of America*. 2001; 110(5):2527–2538. [PubMed: 11757942]
- Carhart R, Tillman TW, Greetis ES. Perceptual masking in multiple sound backgrounds. *Journal of the Acoustical Society of America*. 1969; 45(3):694–703. [PubMed: 5776931]
- Carlyon RP. How the brain separates sounds. *Trends in Cognitive Sciences*. 2004; 8(10):465–471. [PubMed: 15450511]
- Durlach NI, Mason CR, Shinn-Cunningham BG, et al. Informational masking: Counteracting the effect of stimulus uncertainty by decreasing target-masker similarity. *Journal of the Acoustical Society of America*. 2003; 114(1):368–379. [PubMed: 12880048]
- Egan JP, Greenberg GZ, Schulman AI. Interval of time uncertainty in auditory detection. *Journal of the Acoustical Society of America*. 1961; 33(6):771–778.
- Elliott LL, Connors S, Kille E, et al. Children's understanding of monosyllabic nouns in quiet and in noise. *Journal of the Acoustical Society of America*. 1979; 66(1):12–21. [PubMed: 489827]
- Etymic Research. BKB-SIN Speech-in-Noise Test version 1.03 (Compact Disk). 2005.
- Fallon M, Treheh SE, Schneider BA. Children's use of semantic cues in degraded listening environments. *Journal of the Acoustical Society of America*. 2002; 111(5):2242–2249. [PubMed: 12051444]
- Fallon M, Trehub SE, Schneider BA. Children's perception of speech in multitalker babble. *Journal of the Acoustical Society of America*. 2000; 108(6):3023–3029. [PubMed: 11144594]
- Fletcher H. Auditory patterns. *Reviews of Modern Physics*. 1940; 12:47–65.
- Fowler CA. Parsing coarticulated speech in perception: Effects of coarticulation resistance. *Journal of Phonetics*. 2005; 33:199–213.
- Freyman RL, Balakrishnan U, Helfer KS. Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *Journal of the Acoustical Society of America*. 2004; 115(5):2246–2256. [PubMed: 15139635]
- Freyman RL, Helfer KS, McCall DD, et al. The role of perceived spatial separation in the unmasking of speech. *Journal of Acoustical Society of America*. 1999; 106(6):3578–3588.
- Garadat SN, Litovsky RY. Speech intelligibility in free field: Spatial unmasking in preschool children. *Journal of the Acoustical Society of America*. 2007; 121(2):1047–1055. [PubMed: 17348527]
- Gelfand, SA. *Essentials of Audiology*. Third ed. Thieme Medical Publishers, Inc.; New York: 2009.
- Gladstone VS, Siegenthaler BM. Carrier phrase and speech intelligibility test score. *The Journal of Auditory Research*. 1971; 11:101–103.
- Green DM, Weber DL. Detection of temporally uncertain signals. *Journal of the Acoustical Society of America*. 1980; 67(4):1304–1311. [PubMed: 7372917]
- Hall JW, Buss E, Grose JH. Informational masking release in children and adults. *Journal of the Acoustical Society of America*. 2005; 118(3):1605–1613. [PubMed: 16247871]
- Hall JW, Grose JH, Buss E, et al. Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. *Ear and Hearing*. 2002; 23:159–165. [PubMed: 11951851]
- Haskins, H. Unpublished master's thesis. Northwestern University; Evanston, IL: 1949. A phonetically balanced test of speech discrimination for children.
- Helfer KS, Freyman RL. Lexical and indexical cues in masking by competing speech. *Journal of the Acoustical Society of America*. 2009; 125(1):447–456. [PubMed: 19173430]
- Hirsh IJ, Davis H, Silverman SR, et al. Development of materials for speech audiometry. *Journal of Speech and Hearing Disorders*. 1952; 17:321–337. [PubMed: 13053556]
- Kidd G, Best V, Mason CR. Listening to every other word: Examining the strength of linkage variables in forming streams of speech. *Journal of the Acoustical Society of America*. 2008; 124(6):3793–3802. [PubMed: 19206805]
- Kidd G, Mason CR, Deliwal PS, et al. Reducing informational masking by sound segregation. *Journal of the Acoustical Society of America*. 1994; 95(6):3475–3480. [PubMed: 8046139]
- Kidd G, Mason CR, Rohtla TL, et al. Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns. *Journal of the Acoustical Society of America*. 1998; 104(1):422–431. [PubMed: 9670534]

- Killion MC, Niquette PA, Gudmundsen GI, et al. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*. 2004; 116(4):2395–2405. [PubMed: 15532670]
- Ladefoged P, Broadbent DE. Information conveyed by vowels. *Journal of the Acoustical Society of America*. 1957; 29(1):98–104.
- Lee S, Potamianos A, Narayanan S. Acoustics of children's speech: Developmental changes of temporal and speech parameters. *Journal of the Acoustical Society of America*. 1999; 105(3): 1455–1468. [PubMed: 10089598]
- Leibold LJ, Bonino AY. Release from information masking in children: Effect of multiple signal bursts. *Journal of the Acoustical Society of America*. 2009; 125(4):2200–2208. [PubMed: 19354396]
- Leibold LJ, Neff DL. Effects of masker-spectral variability and masker fringes in children and adults. *Journal of the Acoustical Society of America*. 2007; 121(6):3666–3676. [PubMed: 17552718]
- Lewis D, Hoover B, Choi S, et al. Relationship between speech perception in noise and phonological awareness skills for children with normal hearing. *Ear and Hearing*. 2010; 31:761–768. [PubMed: 20562623]
- Litovsky RY. Speech intelligibility and spatial release from masking in young children. *Journal of the Acoustical Society of America*. 2005; 117(5):3091–3099. [PubMed: 15957777]
- Lotto AJ, Holt LL. Putting phonetic context effects into context: A commentary on Fowler (2006). *Perception & Psychophysics*. 2006; 68(2):178–183. [PubMed: 16773891]
- Lynn JM, Brotman SR. Perceptual significance of the CID W-22 carrier phrase. *Ear and Hearing*. 1981; 2(3):95–99. [PubMed: 7250563]
- Madell, JR. Evaluation of speech perception in infants and children. In: Madell, JR.; Flexer, C., editors. *Pediatric Audiology*. Thieme Medical Publishers, Inc.; New York: 2008. p. 89-105.
- Mann V. Influence of preceding liquid on stop-consonant perception. *Perception & Psychophysics*. 1980; 28(5):407–412. [PubMed: 7208250]
- Markham D, Hazan V. The effect of talker- and listener-related factors on intelligibility for a real-world, open-set perception test. *Journal of Speech, Language, and Hearing Research*. 2004; 47:725–737.
- Martin FN, Hawkins R, Bailey H. The nonessentiality of the carrier phrase in phonetically balanced (PB) word testing. *The Journal of Auditory Research*. 1962; 2:319–322.
- McArdle, R.; Wilson, RH. Selecting speech tests to measure auditory function. *The ASHA Leader*. Sep 02. 2008
- Meyer TA, Pisoni DB. Some computational analyses of the PBK test: Effects of frequency and lexical density on spoken word recognition. *Ear and Hearing*. 1999; 20(4):363–371. [PubMed: 10466571]
- Mullennix JW, Pisoni DB, Martin CS. Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America*. 1989; 85(1):365–378. [PubMed: 2921419]
- Neff DL. Signal properties that reduce masking by simultaneous, random-frequency maskers. *Journal of the Acoustical Society of America*. 1995; 98(4):1909–1920. [PubMed: 7593915]
- Neff DL, Green DM. Masking produced by spectral uncertainty with multicomponent maskers. *Perception & Psychophysics*. 1987; 41(5):409–415. [PubMed: 3601622]
- Nilsson M, Soli SD, Sullivan JA. Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and in noise. *Journal of Acoustical Society of America*. 1994; 95(2):1085–1099.
- Nishi K, Lewis DE, Hoover BM, et al. Children's recognition of American English consonants in noise. *Journal of the Acoustical Society of America*. 2010; 127(5):3177–3188. [PubMed: 21117766]
- Nittrouer S, Boothroyd A. Context effects in phoneme and word recognition by young children and older adults. *Journal of the Acoustical Society of America*. 1990; 87(6):2705–2715. [PubMed: 2373804]
- Oh EL, Wightman FL, Lutfi RA. Children's detection of pure-tone signals with random multitone maskers. *Journal of the Acoustical Society of America*. 2001; 109(6):2888–2895. [PubMed: 11425131]

- Papso CF, Blood IM. Word recognition skills of children and adults in background noise. *Ear and Hearing*. 1989; 10(4):235–236. [PubMed: 2776983]
- Ryalls BO, Pisoni DB. The effect of talker variability on word recognition in preschool children. *Developmental Psychology*. 1997; 33(3):441–452. [PubMed: 9149923]
- Sanderson-Leepa ME, Rintelmann WF. Articulation functions and test-retest performance of normal-hearing children on three speech discrimination tests: WIPI, PBK-50, and NU Auditory Test No. 6. *Journal of Speech and Hearing Disorders*. 1976; 41:503–519. [PubMed: 994481]
- Stelmachowicz PG, Lewis DE, Choi S, et al. Effect of stimulus bandwidth on auditory skills in normal-hearing and hearing-impaired children. *Ear and Hearing*. 2007; 28:483–494. [PubMed: 17609611]
- Studebaker GA. A “rationalized” arcsine transform. *Journal of Speech and Hearing Research*. 1985; 28:455–462. [PubMed: 4046587]
- Tillman, TW.; Carhart, R. An expanded test for speech discrimination utilizing CNC monosyllabic words: Northwestern University Auditory Test No. 6. U.S. Air Force; Brooks Air Force Base, TX: 1966.
- Viemeister, NF.; Schlauch, RS. Issues in infant psychoacoustics. In: Werner, LA.; Rubel, EW., editors. *Developmental Psychoacoustics*. American Psychological Association; Washington, D.C.: 1992. p. 191-210.
- Werner, LA.; Leibold, LJ. Ecological developmental psychoacoustics. In: Neuhoff, JG., editor. *Ecological Psychoacoustics*. Academic Press; New York: 2004. p. 191-217.
- Wightman, FL.; Allen, PA. Individual differences in auditory capability among preschool children. In: Werner, LA.; Rubel, EW., editors. *Developmental Psychoacoustics*. American Psychological Association; Washington, D.C.: 1992. p. 113-133.
- Wightman FL, Callahan MR, Lutfi RA, et al. Children’s detection of pure-tone signals: Informational masking with contralateral maskers. *Journal of the Acoustical Society of America*. 2003; 113(6): 3297–3305. [PubMed: 12822802]
- Wightman FL, Kistler DJ. Informational masking of speech in children: Effects of ipsilateral and contralateral distracters. *Journal of the Acoustical Society of America*. 2005; 118(5):3164–3176. [PubMed: 16334898]
- Wightman FL, Kistler DJ, Brungart DS. Informational masking of speech in children: Auditory-visual integration. *Journal of the Acoustical Society of America*. 2006; 119(6):3940–3949. [PubMed: 16838537]
- Wightman FL, Kistler DJ, O’Bryan A. Individual differences and age effects in dichotic informational masking paradigms. *Journal of the Acoustical Society of America*. 2010; 128(1):270–279. [PubMed: 20649222]
- Wilson RH. Development of a speech in multitalker babble paradigm to assess word-recognition performance. *Journal of the American Academy of Audiology*. 2003; 14:453–470. [PubMed: 14708835]
- Wilson RH, Farmer NM, Gandhi A, et al. Normative data for the Words-in-Noise Test for 6- to 12-year-old children. *Journal of Speech, Language, and Hearing Research*. 2010; 53:1111–1121.
- Wilson RH, McArdle R. Intra- and inter-session test, retest reliability of the Words-in-Noise (WIN) Test. *Journal of the American Academy Audiology*. 2007; 18:813–825.

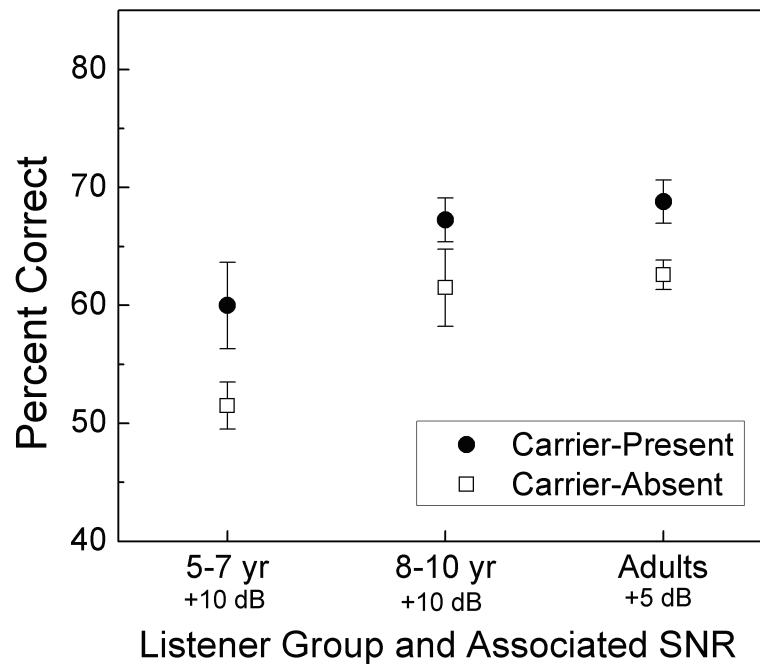


Figure 1.

Average performance scores in percent correct (± 1 SE) are presented for each of the three age groups (5- to 7-year-olds, 8- to 10-year-olds, and adults) in multi-talker babble. Carrier-status is indicated by symbol shape, with filled circles representing carrier-present and open squares representing carrier-absent. The SNR used for testing is indicated on the x-axis.

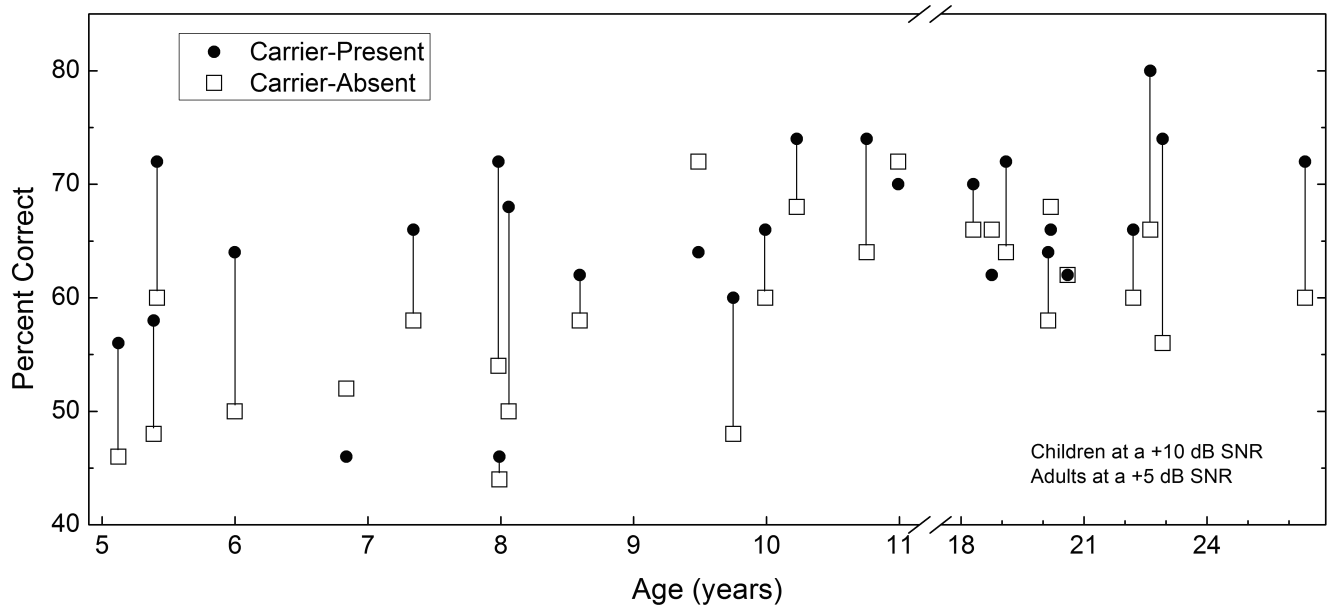


Figure 2.

Individual performance in multi-talker babble is shown as a function of listener age for listeners in the carrier-present (filled circle) and carrier-absent conditions (open square). The vertical line connecting each individual's data points indicates the amount of carrier-phrase benefit for each listener. The absence of a vertical line indicates no benefit or a reduction in percent correct associated with inclusion of a carrier phrase.

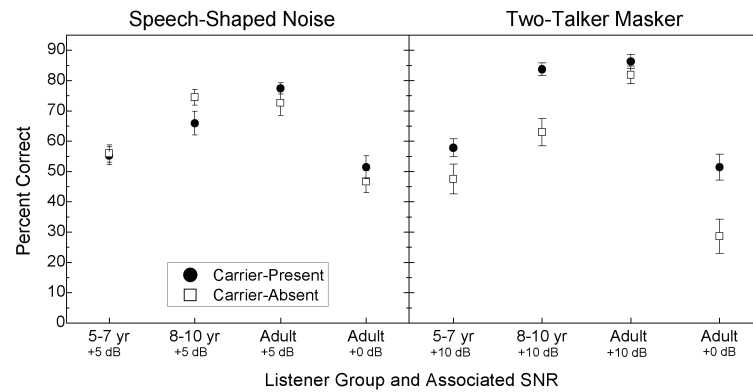


Figure 3. Average performance in percent correct across listeners (± 1 SE) is shown for Experiment 2. The panels separate performance for the two maskers, with speech-shaped noise on the left and the two-talker masker on the right. For each masker, performance is plotted for the two child groups (5- to 7-year-olds and 8- to 10-year-olds) and the two adult groups (SNR-matched and 0 dB SNR). The filled circle symbol represents performance in the carrier-present condition and the open square symbol indicates the carrier-absent condition. The SNR used for testing is indicated on the x-axis.

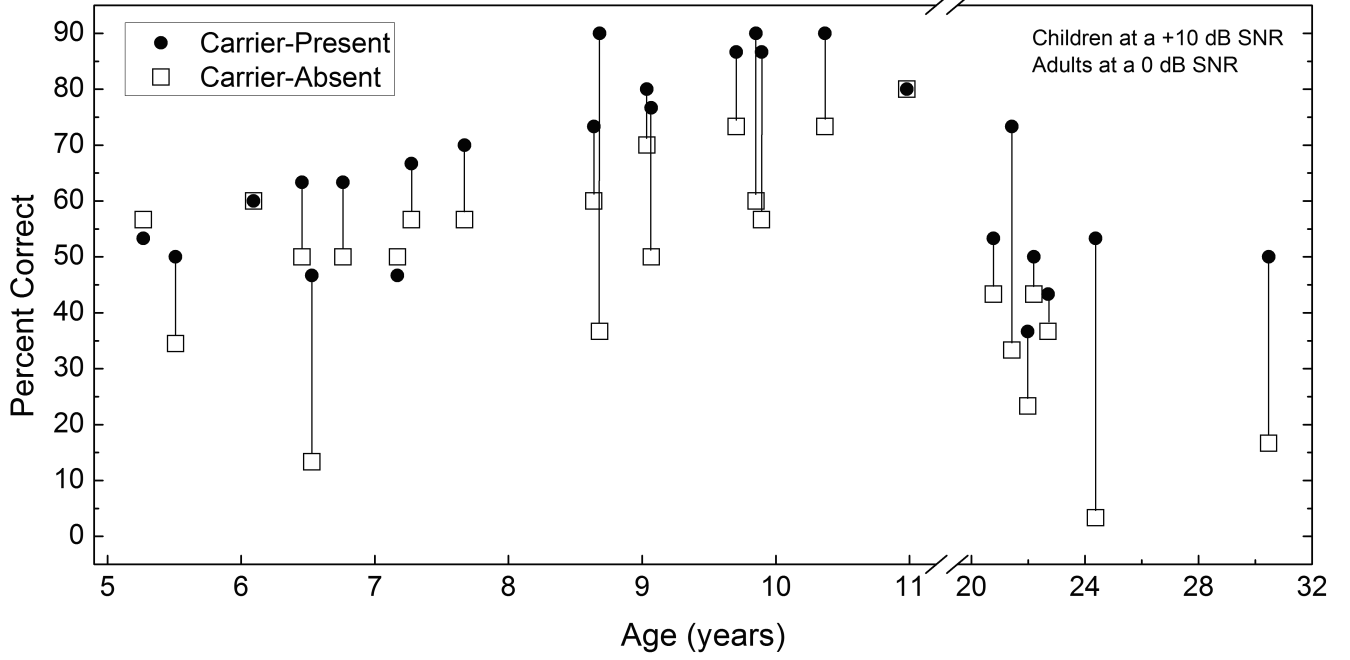


Figure 4. Individual performance in the two-talker masker is shown as a function of listener age for listeners in the carrier-present (filled circle) and carrier-absent conditions (open square). The vertical line between each individual's data points indicates the amount of carrier-phrase benefit. The absence of a vertical line indicates no benefit or a reduction in percent correct associated with inclusion of a carrier phrase. The adults shown on this figure are the group of adults tested at a 0 dB SNR, whereas children were tested at a +10 dB SNR.