Perception & Psychophysics
2007, 69 (7), 1140-1151

The effect of amplitude modulation on intelligibility of time-varying sinusoidal speech in children and adults

Dawna E. Lewis
Boys Town National Research Hospital, Omaha, Nebraska

and

Thomas D. Carrell
University of Nebraska, Lincoln, Nebraska

Although researchers are currently studying auditory object formation in adults, little is known about the development of this phenomenon in children. Amplitude modulation has been suggested as one of the characteristics of the speech signal that allows auditory grouping. In this experiment, we evaluated children (4 to 13 years of age) and adults to examine whether children's ability to use amplitude modulation (AM) in perception of time-varying sinusoidal (TVS) sentences is different from that of adults, and whether there are developmental changes. We evaluated performance on recognition of TVS sentences (unmodulated, amplitude-comodulated at 25, 50, 100, and 200 Hz, and amplitude-modulated using conflicting frequencies). Overall, the youngest children performed more poorly than did older children and adults. However, difference scores, defined as the percentage of phonemes correct in a given modulation condition minus the percentage correct for the unmodulated condition, showed no significant effects of age. Unlike the findings of previous studies (Carrell & Opie, 1992), these results support the ability of modulation with conflicting frequencies to improve intelligibility. The present study provides evidence that children and adults receive the same benefits (or decrements) from amplitude modulation.

Speech perception in natural environments is a complicated task that requires processing at many different levels. Before word recognition can occur, the listener must be able to separate the signal of interest from background noise. Auditory object formation is the ability to separate one sound source from a larger set of sound sources (Moore, 2003). Many of the processes that contribute to the formation of auditory objects may also be responsible for our ability to group sounds into a coherent speech signal. The formation of auditory objects includes both simultaneous and sequential grouping cues such as common fundamental frequencies (Assmann & Summerfield, 1987, 1990), simultaneity of onsets and offsets (Darwin, 1981, 1984), harmonicity and comodulation (Grose & Hall, 1992; Hall & Grose, 1990; Hall, Haggard, & Fernandes, 1984), continuity of pitch (Darwin & Bethell-Fox, 1977), interaural time differences (Darwin & Hukin, 1999), and temporal order (Bregman, 1990). In the perception of speech, phonetic and semantic relations and listener expectations can also provide information useful for grouping speech sounds into an auditory object (Miller, Heise, & Lichten, 1951). It has been suggested that this ability to perceive sounds as separate auditory objects may facilitate perception of speech in noise, as well as the ability to follow one person's voice in a group of talkers (see Bregman, 1990).

Because speech is a complex signal that contains multiple cues for perception (Denes, 1955; Repp, 1982), the listener may be able to rely on alternate cues in adverse listening conditions (e.g., reverberation, noise). Although this robust nature of speech is of great benefit to listeners, it also makes it difficult to evaluate the contribution of specific cues to perception. Time-varying sinusoidal (TVS) speech (Barker & Cooke, 1999; Carrell & Opie, 1992; Remez, Rubin, Pisoni, & Carrell, 1981) has been used as an impoverished form of speech that does not contain many of the grouping cues known to aid speech perception. Unlike real speech, TVS speech consists of three to four constant-amplitude, time-varying sinusoids that follow the center frequencies of formants of naturally spoken utterances. TVS speech does not contain the fundamental frequencies, harmonic structure, formant frequency transitions, or short-term spectral cues found in natural speech. As a result, acoustic cues thought to promote auditory object formation may be systematically investigated using TVS speech, by adding cues singly or in groups.

Carrell and Opie (1992) used TVS speech to determine the contribution of amplitude comodulation to auditory

D. E. Lewis, lewis@boystown.org

grouping in speech perception for adults. They compared recognition scores for unmodulated TVS (UTVS) sentences with sentences in which the three sinusoids were amplitude-modulated simultaneously (AMTVS) at 100 Hz. The results revealed improved intelligibility for the AMTVS sentences, supporting the hypothesis that modulation served as a mechanism for grouping the three sinusoids into a single auditory object, thereby increasing intelligibility. These findings were further supported in a later study by Barker and Cooke (1999).

To determine whether the improvement in recognition for AMTVS over UTVS speech was the result of the fact that amplitude-modulating each of the tones at the same frequency served as a grouping cue or whether the modulation created a signal that simply sounded more natural or speech-like, Carrell and Opie (1992) created a new set of sentences in which each of the tones was modulated at a different rate. Despite conflicting modulation, these stimuli had sound quality and spectral smearing characteristics similar to those of the comodulated stimuli. The results revealed higher intelligibility scores for the AMTVS sentences in comparison with the conflicting amplitude-modulated (CAMTVS) sentences. The authors concluded that it was amplitude *comodulation* that was critical for improving recognition.

Carrell and Opie (1992) also attempted to examine the underlying cause for improvements in recognition seen with amplitude comodulation by asking whether the comodulation effect seen in their experiments was due to a grouping mechanism similar to the one that underlies comodulation masking release (CMR). Previously, Hall et al. (1984) showed that a tone presented at subthreshold levels and centered in an amplitude-modulated narrowband noise can be audible. However, this effect occurs only if another band of noise that is amplitude-modulated at the same rate and phase is added at a different frequency. One interpretation of this phenomenon (CMR) is that the noise bands are grouped together by their common amplitude modulation, making the subthreshold tone more salient.

To assess the validity of a CMR-based underlying mechanism, Carrell and Opie (1992) compared TVS sentence recognition using no modulation and amplitude-modulation rates of 50, 100, and 200 Hz. Because CMR is stronger at low-modulation frequencies (Eddins & Wright, 1994; Hall, Cokely, & Grose, 1988), Carrell and Opie hypothesized that improved intelligibility for modulation rates of 50 and 100 Hz, but not for 200 Hz, would support a CMR-based explanation. The results indicated better recognition for sentences modulated at 50 and 100 Hz than for those modulated at 200 Hz or not modulated at all. Thus, they concluded that their results were consistent with a grouping mechanism similar to the mechanism underlying CMR.

C. C. Dunn (2003), however, reported that adult listeners demonstrated poorer intelligibility scores for TVS sentences modulated at 10 Hz than for sentences modulated at 100 Hz. These findings were in contrast to results for a CMR task in which these same listeners demonstrated greater release from masking for noises modulated at 10 Hz than for those modulated at 100 Hz. Dunn concluded that the improvements in speech intelligibility for the higher modulation rate may have resulted from the fact that the higher modulation rate mimics the glottal source in natural speech. Given the latter finding, it would be helpful to conduct additional investigations that extend the amplitude modulation to a lower rate than that used by Carrell and Opie (1992).

The ability of amplitude modulation between the rates of 50 and 100 Hz to increase the intelligibility of acoustically sparse TVS sentences is now well established in adult listeners. In addition, Perri (1997) demonstrated that speech masked by amplitude-modulated noise was more intelligible than speech masked by unmodulated noise. This held true despite the fact that the noise maskers were equilibrated for root-mean square (RMS) loudness. This pattern of results is consistent with the notion that amplitude modulation aids auditory grouping. Such a capability appears fundamental to auditory processing. For example, a listener must be able to segregate the simultaneous sounds into their component sources early in the process of speech perception.

If, as proposed above, auditory object formation is a fundamental component of speech perception, it is important to know how this capability develops. Although much is known about acoustic phonetics and more linguistic aspects of speech perception in children, much less is known about the process of extracting a desired speech signal from undesired noise (which may also be speech).

Numerous studies have shown differences between children and adults on a variety of speech-perception tasks in noise and reverberation (e.g., Fallon, Trehub, & Schneider, 2000; Hall, Grose, Buss, & Dev, 2002; C. E. Johnson, 2000) and in the presence of reduced spectral cues (Eisenberg, Shannon, Martinez, Wygonski, & Boothroyd, 2000). C. E. Johnson investigated the effects of reverberation and noise on the ability of listeners (ages 6–7, 10–11, 14–15 years, and adults) to identify vowels and consonants in nonsense words presented at a variety of sensation levels. The results indicated that children's maximum performance varied by listening condition and age, with some conditions not achieving adult-like performance until the late teenage years.

Using spectrally reduced speech, Eisenberg et al. (2000) evaluated speech recognition for children (5–7 and 10–12 years old) and adults using phonemes, words, and sentences. The results revealed that the youngest group required greater spectral resolution to perform at the same level as older children and adults. The younger children were also less able to use context to recognize words in sentences. Numerous studies have also shown that children differ from adults in their use of temporal and contextual cues in speech perception (Elliott, 1986; Elliott, Busse, Partridge, Rupert, & DeGraaff, 1986; Nittrouer & Boothroyd, 1990).

Theories regarding the underlying causes of the above-noted developmental differences vary. Explanations include peripheral factors such as auditory sensitivity (Schneider, Trehub, Morrongiello, & Thorpe, 1986, 1989), central auditory processes (Allen & Wightman, 1994, 1995), attention and memory (Buss, Hall, Grose,

& Dev, 1999; Hill, Hartley, Glasberg, Moore, & Moore, 2004; Wightman, Callahan, Lutfi, Kistler, & Oh, 2003), language experience (Eisenberg et al., 2000), poor information extraction (Grose, Hall, & Dev, 1997), and task difficulty (Fallon et al., 2000).

It is clear that there are a number of cues that affect auditory grouping. Research has suggested that amplitude comodulation can serve as one auditory grouping cue, improving intelligibility and holding acoustically independent tones in TVS speech together (Barker & Cooke, 1999; Carrell & Opie, 1992). It is also clear that many aspects of auditory perception are developmental and may reach adult-like levels at widely disparate ages, ranging from infancy to late adolescence.

Limited research on auditory grouping has been conducted with infants and young children. These studies have shown that very young infants demonstrate sequential auditory organization, suggesting that this ability may be present at birth (e.g., Demany, 1982; McAdams & Bertoncini, 1997; Trainor & Adams, 2000). However, the developmental time course of other aspects of auditory grouping remains unclear. To the extent that auditory grouping abilities may facilitate the perception of speech under typical or adverse listening conditions, this study examined the developmental time course of one characteristic of the speech signal that may allow auditory grouping.

The primary goal of the present investigation was to determine the extent to which children between the ages of 4 and 13 years benefit from amplitude modulation (AM) in the perception of TVS speech. Given expected improvements in recognition for modulated TVS speech, a secondary goal was to confirm the findings of Carrell and Opie (1992), indicating that comodulation can account for improvements in intelligibility, and to extend those findings to children. To answer this question, we included a condition using stimuli that have similar sound quality and spectral smearing characteristics as amplitude-comodulated speech, but are not comodulated. Finally, in the present experiment, we examined whether a CMR-based explanation can account for improvements in perception, by using four different comodulation rates, one of which was below the range of fundamental frequencies for natural speech. We hypothesized that children would be less able than adults to benefit from amplitude modulation, and that their ability to use this information would improve with age. Further, we expected that children would demonstrate better recognition for stimuli that are amplitude-comodulated, and that, as suggested by C. C. Dunn (2003), recognition would be poorest for the sentences with the lowest modulation rates.

METHOD

We conducted a pilot study to evaluate the abilities of children to recognize TVS speech and to investigate the procedural modifications necessary to obtain valid and reliable data from young children. The results revealed that children from 4 to 13 years of age were able to perform the required tasks, and that amplitude comodulation has, at least generally, similar effects on children and adults. However, results also suggested that we should use vocabulary and complexity of test sentences appropriate for young children, and that extended practice would be necessary.

Subjects

A total of 180 subjects (divided into six groups) participated in this study. The adult group consisted of 30 subjects between 19 and 40 years of age. The 150 children were subdivided into five age groups (4–5, 6–7, 8–9, 10–11, 12–13 years), with 30 subjects per group. Equal numbers of males and females were included in each age group. All of the subjects had normal hearing (≤20 dB HL for the octave frequencies 250–8000 Hz) and were native speakers of English. The subjects were recruited from Communication Disorders courses at the University of Nebraska, Lincoln, and from the Human Research Subjects Core at Boys Town National Research Hospital. The adult subjects received course credit or were paid for their participation. The children were paid for their participation and were rewarded with a small toy and book after completion of the experiment.

Stimuli

To ensure that the vocabulary was appropriate for the age groups used in this study, we compiled 60 sentences from ones used in previous investigations (C. D. Johnson & Owen, as cited in C. D. Johnson, Benson, & Seaton, 1997; Kenworthy, Klee, & Tharpe, 1990; Stelmachowicz, Hoover, Lewis, Kortekaas, & Pittman, 2000). Naturally produced tokens of four-word sentences (e.g., "I wrecked my bike," "Let me have it") were spoken by an adult male talker. The sentences were recorded at a sampling rate of 44.1 kHz, with an amplitude resolution of 16 bits, using a head-mounted microphone (Crown CM320) and a Sony A7 DAT recorder. The recordings were digitally transferred to a computer, and each sentence was saved as a separate sound file.

To create UTVS sentences from the original utterances, a spectrogram of each sentence was produced (SpeechStation 2, Sensimetrics, 1998). The center frequencies for each of the first four formants were estimated by visually tracing the center of each formant across time and recording the frequency and a gross estimate of the amplitude. The resulting values were entered into a computer program that constructs waveforms made up of independent sinusoidal waves (Tone16, Tice & Carrell, 1997). Peak amplitudes for the tones representing the four formants were set to 90-, 87-, 84-, and 81-dB RMS (re: 16 bits = 96-dB peak).[1] UTVS sentences were generated with a sampling rate of 10 kHz.

AMTVS sentences were created by amplitude-comodulating each of the UTVS sentences at 25, 50, 100, or 200 Hz. Conflicting amplitude-modulated TVS (CAMTVS) sentences were created by synthesizing each of the four time-varying sinusoids separately. Each sinusoid was then amplitude-modulated at a different frequency (97, 79, 113, 89 Hz), selected to replicate those used by Carrell and Opie (1992). For the CAMTVS sentences, the frequency of each tone was a prime number, in order to avoid the same lowest common denominator across the four tones. The resulting sound files were mixed (SoundForge, SonicFoundry, 1998) to create a CAMTVS version of each sentence. The modulating signal for both AMTVS and CAMTVS sentences was a triangular wave with an 80% duty cycle. The sentences were equilibrated at 70-dB RMS (re: 16 bits = 96-dB peak). Sample spectrograms of the first 600 msec of the sentence "My tooth is loose" are shown in Figure 1 for natural speech and all six experimental conditions.

To select sentences for use in this study, 24 young adult listeners (ages 19–29 years) participated in an identification task. Following a brief familiarization, the subjects listened to the 60 original sentences. Half of the subjects heard 30 of the sentences as UTVS speech and 30 as AMTVS speech modulated at 100 Hz, whereas the other half heard the reverse. Scoring was based on the number of phonemes correct per sentence. The 30 sentences with AMTVS scores of >60% that produced the greatest average improvement for AMTVS over UTVS sentences were selected as experimental

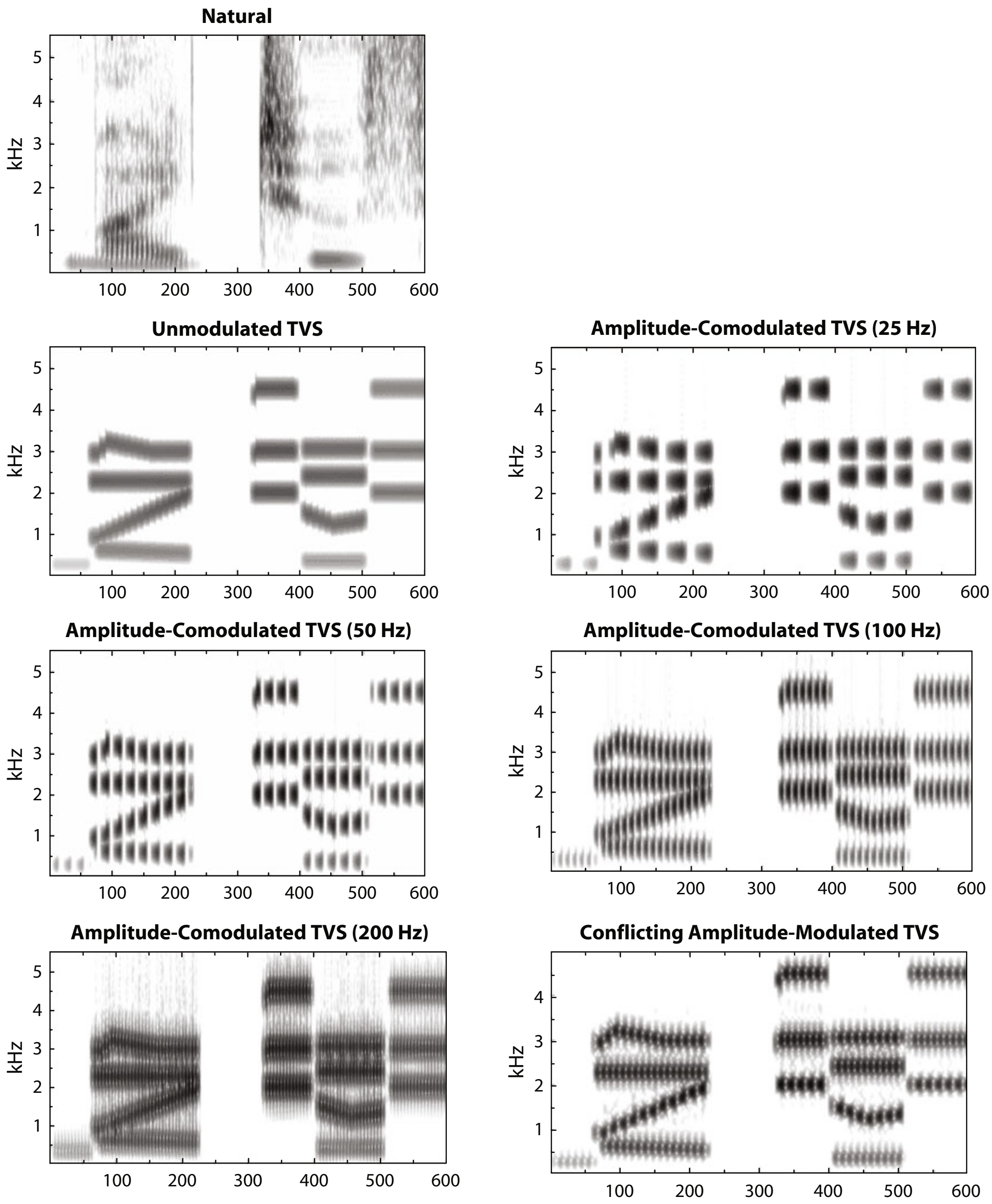

Figure 1. The first 600 msec of wide band spectrograms of the sentence "My tooth is loose" for natural speech and each of the six conditions (TVS = time-varying sinusoidal sentences).

sentences. The 6 sentences with the lowest AMTVS intelligibility scores were excluded, and the remaining 24 sentences were selected as practice sentences (see the Appendix for a list of sentences). The selection criteria were based on the need to use sentences that would be appropriate for the subjects (of all ages) being tested. On the basis of pilot data, as well as the results of previous research utilizing sinusoidal speech, we assumed that floor effects might limit analysis of results for the youngest children if the sentences were too difficult, whereas ceiling effects would not be likely, even for the adults.

Procedure

Before beginning the experiment, we assessed speech production using the Bankson–Bernthal Quick Screen of Phonology (Bankson & Bernthal, 1990) to identify any production errors that could influence scoring of verbal responses. Children with significant production errors were excluded from further testing. We assessed receptive vocabulary using the Peabody Picture Vocabulary Test (PPVT; L. M. Dunn & L. M. Dunn, 1997). To quantify developmental changes in speech recognition in noise, we also evaluated children using the Hearing in Noise Test for Children (HINT-C; Nilsson, Soli, & Gelnett, 1996).

Sentences were presented under computer control (MakeTape; Tice & Carrell, 1996). The sentences were output using a 16-bit TDT D/A converter at a sampling rate of 10 KHz, and low-pass filtered at 4300 Hz with a TDT PF1 programmable filter. Sentences were presented binaurally at a peak level of approximately 70 dB SPL via headphones (Sennheiser HD430).

We conducted a practice phase prior to beginning the experiment in which all of the listeners heard the 24 practice sentences. These sentences were used only for practice. Each listener heard four sentences in each of six conditions: UTVS sentences, AMTVS sentences modulated at 25, 50, 100, and 200 Hz, and CAMTVS sentences, in which each of the four time-varying sinusoids was modulated at a different rate. Each sentence was presented once, and the listener was encouraged to guess what he or she heard. Listeners were encouraged to respond with any sounds they heard, even if they did not hear complete words or sentences and even if what they heard did not make sense. Next, the listener heard the sentence spoken by the investigator, and then the processed sentence was presented two more times, with a 3.5-sec interstimulus interval (ISI) to help familiarize him/her with the stimuli. These procedures were used to ensure that subjects were comfortable listening to the sinusoidal speech prior to testing.

Because hearing a sentence even once can affect perception of subsequent presentations of that sentence, we used a block design for the sentences and then a modified Latin square design for the modulation conditions in both the practice and experimental phases. We selected this design to ensure that all of the sentences would be presented in each condition but that no subject would hear a sentence more than once. This design also ensured that all of the sentences were presented in at least one condition for each age group and that presentation order for conditions would vary systematically within each age group. Within each age group, we changed the block every 5 subjects. Thus, the first 5 subjects in a given age group heard the same sentences in the same order/condition. The next 5 subjects heard conditions in a different order, with sentences changed across condition, and so on across the 30 subjects per group.

During the experiment, each listener heard five sentences in each of the six conditions using the same presentation order specified for the practice sentences. On every trial, each listener heard a sentence three times, with a 3.5-sec ISI. During the experiment, listeners were not given any feedback regarding the sentences. After the third presentation, the adult listeners were given a 10-sec interval to write what they heard. Children were asked to repeat what they heard during this 10-sec interval, and the investigator transcribed the children's responses, asking for clarification as needed. Because the sentences were short, 10 sec was ample time for listeners of different ages to respond.[2]

Listeners' responses were scored phonemically by the first author. To enhance consistency in scoring, listeners were not penalized for additional phonemes or for extra words. In addition, if they wrote/said only a few words or phonemes, credit was given for those that were correct. This gave listeners the benefit of the doubt when it was unclear exactly where in the sentence they heard the word/phoneme.

RESULTS

The primary goal of this study was to examine the effect of amplitude modulation on children's perception of TVS speech. Figure 2 displays mean phoneme score as a function of age for each modulation condition. Perception was affected by modulation rate for all age groups. Most notably, the mean phoneme score across all ages was lowest for the 200-Hz modulation rate condition, followed by the unmodulated condition. At most ages, scores were best for the 50- and 100-Hz conditions. However, across all age groups, mean scores for the conflicting modulation condition were high and better than would have been predicted on the basis of Carrell and Opie's (1992) findings.

As seen in Figure 2, mean scores suggest that 4- to 5-year-olds performed more poorly than any other age group. We conducted a repeated measures ANOVA with modulation condition as the within-subjects variable and age as the between-subjects variable. We used a Greenhouse–Geisser correction to adjust the degrees of freedom due to a failure to meet the assumption of sphericity (Max & Onghena, 1999). The results revealed a main

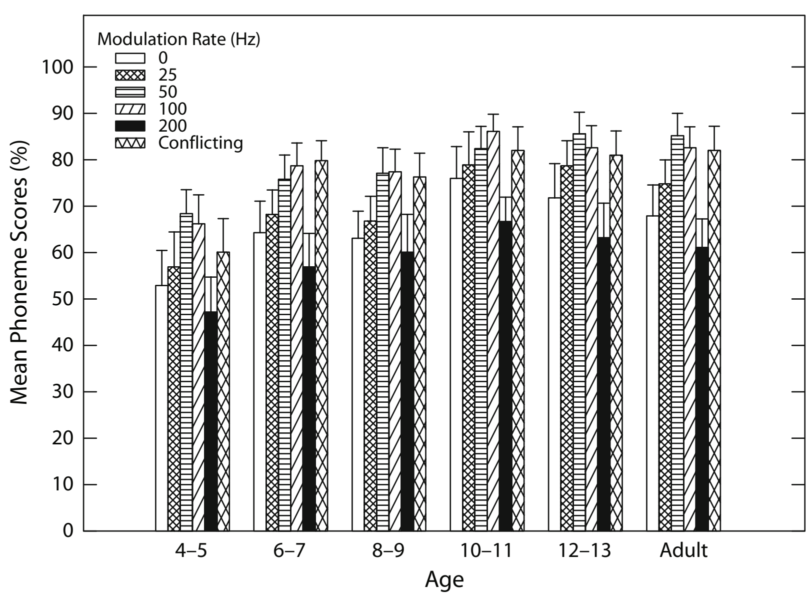

Figure 2. The mean phoneme scores across age groups as a function of modulation rate. The error bars represent 95% confidence intervals.

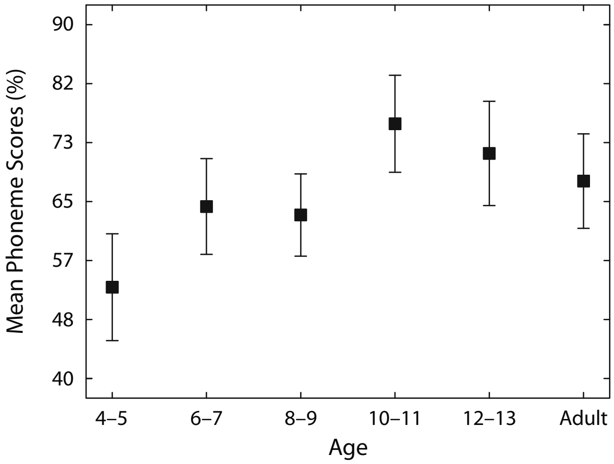

Figure 3. The mean phoneme scores for unmodulated time-varying sinusoidal (UTVS) sentences as a function of age. Error bars represent 95% confidence intervals.

effect of modulation condition [$F(4.2, 730.8) = 59.41$, $p < .001$; $\eta_p^2 = .26$] and age [$F(5,174) = 14.04$, $p < .001$; $\eta_p^2 = .29$], but no modulation condition × age interaction [$F(21, 730.8) = 0.742$, $p = .811$; $\eta_p^2 = .02$]. Bonferroni-adjusted post hoc tests revealed that 4- to 5-year-olds were significantly different from all other groups. In addition, differences were found between 8- to 9-year-olds and 10- to 11-year-olds.

To examine potential developmental effects on perception of UTVS (0 modulation) speech, we examined scores for these sentences separately. Figure 3 displays mean phoneme scores (and 95% confidence intervals) for UTVS sentences as a function of age group. Although mean scores for the 4- to 5-year-olds are lower than for any other group, the confidence intervals for individual groups revealed considerable overlap across age. We further investigated age differences using a one-way ANOVA, with age as the independent variable and phoneme score for UTVS sentences as the dependent variable. Although results indicated a significant effect of age [$F(5,174) = 5.21$, $p < .001$; $\eta_p^2 = .13$], post hoc Scheffé analyses revealed that the only significant differences were between the 4- to 5-year-old and 10- to 11-year-old children ($p < .001$) and between the 4- to 5-year-old and the 12- to 13-year-old children ($p < .015$).

To examine the effect of amplitude modulation on phoneme scores, we calculated difference scores between the scores for each modulation condition and the unmodulated (0 mod) condition. Figure 4 illustrates the modulation-based changes across age groups. Positive values indicate that subjects' phoneme scores improved in that modulation condition relative to the unmodulated condition. Negative values indicate that subjects' phoneme scores were better in the unmodulated condition. Notably, difference scores for the 200-Hz modulation condition (Diff 200) were negative for all age groups, indicating that subjects performed more poorly in this condition than they did in the unmodulated condition. All other difference scores were positive. Difference scores for the 25-Hz modulation condition (Diff 25) were also low across all age groups, suggesting that modulating TVS speech at a rate of 25 Hz provided only limited benefit for perception.

We conducted a repeated measures ANOVA with difference scores across the five modulation conditions as the within-subjects factor and age as the between-subjects

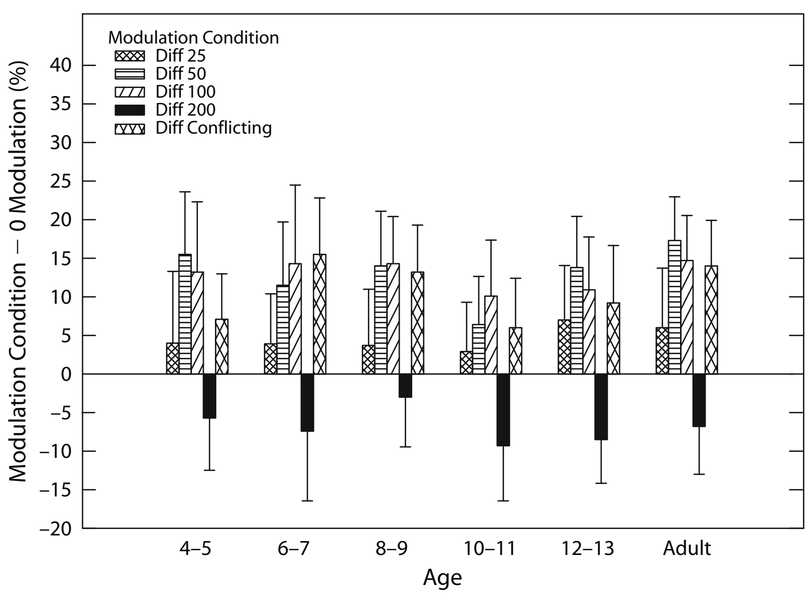

Figure 4. Modulation-based intelligibility changes across age. Each modulation condition represents the difference between the scores for that modulation condition and the unmodulated (0 mod) condition (e.g., "Diff 25" = scores for the 25-Hz modulation condition − scores for the unmodulated condition). Error bars represent 95% confidence intervals.

factor. We used a Greenhouse–Geisser correction to adjust the degrees of freedom due to a failure to meet the assumption of sphericity (Max & Onghena, 1999). The results revealed a main effect of difference scores [$F(3.4, 587.9) = 65.72$, $p < .001$; $\eta_p^2 = .27$] but no effect of age [$F(5,174) = 0.537$, $p = .748$; $\eta_p^2 = .02$], and no difference scores × age interaction [$F(16.9, 587.9) = 0.773$, $p = .725$; $\eta_p^2 = .02$]. Therefore, although the improvement or decrement in phoneme scores varied as a function of modulation rate, the degree to which performance deviated from the unmodulated condition was not significantly different across age groups.

The second question addressed whether intelligibility improvements found with amplitude modulation were a result of amplitude comodulation or some other factor. As shown in Figure 4, difference scores for the 200-Hz modulation condition were much lower than those for the other conditions. Although all of the other conditions showed an improvement when TVS speech was amplitude-modulated, modulation at a rate of 200 Hz resulted in poorer speech perception across all age groups. This finding is not entirely unexpected, on the basis of the results of Carrell and Opie (1992). Examination of their mean results suggests that subjects performed more poorly in the 200-Hz modulation condition than in the unmodulated condition for three of the four sentences tested. Clearly, amplitude comodulation alone was not sufficient for improving sentence recognition. The rate of modulation was an important factor.

Because statistical analysis had shown an effect of modulation differences but not of age, we combined all age groups to examine the effect of modulation condition on difference scores. We conducted a repeated measures ANOVA with difference scores as the within-subjects factor and used a Greenhouse–Geisser correction to adjust the degrees of freedom due to a failure to meet the assumption of sphericity (Max & Onghena, 1999). The results revealed a significant effect of modulation rate [$F(3.4, 607.8) = 66.14$, $p < .001$; $\eta_p^2 = .27$]. Furthermore, post hoc Bonferroni tests revealed significant differences between the Diff 25 and Diff 200 conditions and all other conditions. There were no significant differences between any other conditions. Therefore, although the Carrell and Opie (1992) data suggest that amplitude comodulation is necessary to improve sentence recognition, results from the present study suggest that subjects perform as well with conflicting amplitude modulation as with amplitude comodulation at 100 Hz.

The final question in the present study examined whether a CMR-based explanation could explain improvements in perception. Figure 4 reveals that, whereas subjects do show improvement at the 25-Hz modulation rate, magnitude of the effect is less than that achieved at the 50- or 100-Hz modulation rates. Given that CMR demonstrates maximum effect at low modulation rates (Eddins & Wright, 1994; Hall et al., 1988), the trend seen here, taken by itself, would not support a CMR-like mechanism alone underlying the perceptual improvements.

Children were also tested using the PPVT and the HINT-C. Means and standard deviations for scores on the PPVT and HINT-C are shown in Table 1. The HINT-C scores represent the signal-to-noise ratio (SNR) corresponding to 50% correct, averaged across two lists. As expected, raw scores for the PPVT improved with increasing age. In addition, the SNR required for 50% performance on the HINT-C decreased with age. It should be noted that 5 children in the 4- to 5-year-old age range were unable to complete the HINT-C, either because of distractibility/fatigue by that point in the test session or because of difficulty with the background noise. This finding was not entirely unexpected, given that the HINT-C was designed for children 5 years old and older and that norms are available only for children as young as 6 years of age.

We further examined the relationship between development and children's perception of TVS speech through a correlation of the variables age, difference scores, raw score on the PPVT, and SNR on the HINT-C (Table 2). The negative correlation for the HINT-C occurs because, unlike with the PPVT, the SNR required to understand the sentences on this task decreased with age. As expected, the results revealed strong correlations among age, PPVT score, and SNR on the HINT-C. There was also a strong correlation among the difference scores for the TVS sentences. However, there were no statistically significant correlations among the age-related variables and the variables associated with perception of TVS speech. These results suggest that amplitude modulation independently affected perception and did not interact with other, age-related, variables.

DISCUSSION

The primary goal of the present study was to examine potential changes in children's ability to use amplitude modulation in the perception of TVS speech. In general, the youngest group (4–5 years old) performed more poorly than did older children and adults. However, when we examined the benefit resulting from the addition of amplitude modulation, there were no significant age effects. That is, amplitude-modulation-based performance changes (positive or negative) were similar across age groups. The absence of an age effect suggests that the ability to use this nonlinguistic auditory cue is established at least by 4 years of age. This was a surprising finding, given the many aspects of auditory perception that do show developmental trends across this age range (see, e.g., Hall, Buss, Grose, & Dev, 2004; Schneider et al., 1986, 1989). On the other hand, auditory grouping is an important pre-

Table 1
Mean Scores for the Peabody Picture Vocabulary Test (PPVT) and Hearing in Noise Test for Children (HINT-C)

	PPVT (Raw Score)		HINT-C	
Age (Years)	*M*	*SD*	*M*	*SD*
4–5	82.9	16.6	6.9	1.67
6–7	109.8	14.8	5.7	1.0
8–9	130.2	13.2	4.6	1.3
10–11	154.8	17.1	3.6	1.8
12–13	166.4	13.2	3.1	1.1

Table 2
Correlation Coefficients Between Age, Difference Scores From the Phoneme Perception Task, Performance on the Peabody Picture Vocabulary Test (PPVT), and Performance on the Hearing in Noise Test for Children (HINT-C)

	Age	Diff 25	Diff 50	Diff 100	Diff 200	Diff CFM	PPVT	HINT-C
Age	1.000	.055	.021	−.016	−.038	.022	.891*	−.680*
Diff 25		1.000	.567*	.594*	.250*	.433*	.030	−.040
Diff 50			1.000	.688*	.340*	.501*	−.088	.060
Diff 100				1.000	.548*	.669*	−.056	.051
Diff 200					1.000	.537*	−.051	.025
Diff CFM						1.000	−.038	.033
PPVT							1.000	−.624*
HINT-C								1.000

Note—"Diff *x*" indicates the difference score when sentences were modulated at *x* Hz. "Diff CFM" is the difference score for conflicting amplitude-modulated sentences. *Significant at the .01 level (two-tailed).

cursor for many higher level perceptual functions, so, in hindsight, this result should not have been unexpected. There are some auditory perceptual tasks for which young children demonstrate adult-like performance. For example, children's frequency-resolving abilities have been reported to be adult-like by 5 to 6 years of age (Allen, Wightman, Kistler, & Dolan, 1989; Hartley, Wright, Hogan, & Moore, 2000). Fallon, Trehub, and Schneider (2000, 2002) found that young children performed similarly to adults on speech perception in noise tasks when adult–child differences such as auditory sensitivity and cognitive difficulty of the task were taken into account. In the present study, we also tried to ensure that the instructions, procedures, and test materials would be appropriate for young children. As a result, nonsensory factors (such as language abilities) that might have affected results were minimized, providing a clearer picture of the effects of amplitude modulation across age.

In contrast to the results of Carrell and Opie (1992), listeners in the present study did not perform more poorly in the CAMTVS condition than in the AMTVS conditions. Carrell and Opie's hypothesis was that conflicting amplitude modulation would result in poorer intelligibility than would amplitude comodulation, because comodulation caused the sinusoids to be grouped as a single auditory object. Conversely, no difference in performance might suggest that amplitude modulation, in general, creates a signal that sounds more "speech-like." Thus, with such an outcome, it could be argued that an improved *speech-like quality*, rather than auditory grouping based on comodulation, contributes to improved performance. On the face of it, the results of the present study provide greater support for the latter explanation. That is, in this study, the improvements in performance when TVS sentences were amplitude-modulated may be related to how speech-like they sounded. Even when the sinusoids were *not comodulated*, the modulation rates used in this study (with the exception of 200 Hz) may have resulted in a signal that maintained enough of a speech-like quality to aid in perception.

Other explanations of the present outcome exist. Although we designed this study to be similar to Carrell and Opie (1992), the methodologies we employed were somewhat different, and these differences may explain the differing outcomes of the two sets of findings. One difference was that the listeners heard many more sentence familiarization and test stimuli in the present work than in Carrell and Opie. The extra time spent learning to listen to these unusual sounds may have been sufficient to reduce the importance of amplitude comodulation and auditory grouping in sentence-identification tasks. A second difference was that the sentences in Carrell and Opie consisted primarily of vowels and semivowels, whereas the sentences in the present work contained a more representative sample of the phonemes of English. The effect of these differences might have been that the component tones of the sentences used by Carrell and Opie were more poorly grouped due to a lack of correlated frequency and amplitude changes, in comparison with the component tones in the present speech. Specifically, more stops, formant transitions, and fricative-to-voiced boundaries might have caused the present sentences to be better grouped, thereby reducing the importance of precise comodulation. A third difference was that four (rather than three) simultaneous tones were used to create the present sentences. This may have shifted the relative weighting of naturalness versus auditory grouping in sentence identification.

Any one of these factors, or all of them taken together, could have caused the listener to hear CAMTVS and AMTVS sentences as nearly equally intelligible, because the precise amplitude-comodulation of AMTVS speech may play a larger role in grouping when it is one of the only cues. Pseudoperiodicity and naturalness may have more of a role, because speech contains more alternative grouping cues. Moreover, substantial familiarization with these sentences may have allowed the listeners to depend less on auditory grouping mechanisms.

As has been found in earlier work, performance varies with amplitude-modulation rate, even when comodulation is employed. Results from Carrell and Opie (1992) and C. C. Dunn (2003) suggest two different interpretations for the role of CMR in the ability to use amplitude comodulation as a grouping cue. Improvements in performance in the 50-Hz modulation condition led Carrell and Opie to suggest an underlying mechanism similar to the one underlying CMR. C. C. Dunn, on the other hand, found that subjects performed more poorly when UTVS sentences were modulated at 10 Hz, which seems to pre-

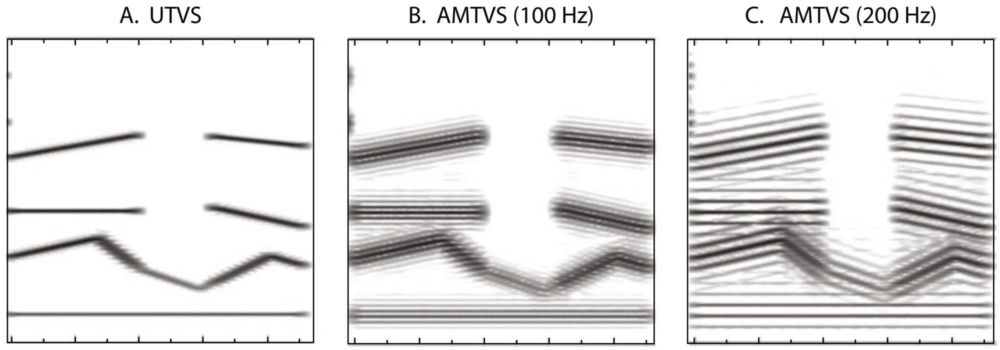

Figure 5. The same section of a narrow band spectrogram of a time-varying sinusoidal (TVS) sentence unmodulated (UTVS, panel A), modulated at 100 Hz (AMTVS, panel B) and at 200 Hz (AMTVS, panel C).

clude a CMR-like mechanism. In the present study, the 25-Hz modulation rate resulted in less improvement in phoneme scores than did the 50-Hz, 100-Hz, or conflicting modulation conditions. As in C. C. Dunn, the advantage of amplitude comodulation was less for lower modulation rates. C. C. Dunn theorized that a 100-Hz modulation rate would "mimic the glottal source in naturally spoken speech," whereas the 10-Hz modulation rate used in her study might have created additional periods of silence, apart from the natural silences in the original sentences, thus reducing intelligibility. It is possible that the 25-Hz modulation rate used in the present study may have interfered with the natural amplitude envelope based on the average rate of speech. These differing envelopes might reduce performance, relative to higher modulation rates (see also Miller & Licklider, 1950). During testing, a number of adults and older children commented about the "whistles and beeps" that accompanied the speech when they were listening to unmodulated TVS speech or speech modulated at 25 Hz. Several stated that the "added" sounds made the speech difficult to understand.

As in the Carrell and Opie (1992) study, the use of a 200-Hz modulation rate interfered with phonetic accuracy to the extent that subjects performed more poorly than with no modulation at all. Although 200 Hz is within the range of fundamental frequencies for female talkers (see Kent, 1997), modulating the sinusoids at 200 Hz may have resulted in sidebands that overlapped adjacent formant frequencies (see Figure 5).

Although we initiated the present work to investigate one aspect of the development of auditory grouping in children, the results go beyond this question. Prior to this study, it had been assumed that the tones in unmodulated TVS speech are acoustically independent. When something was done to group the tones, such as informing listeners that they were hearing speech (Remez et al., 1981), or using amplitude *comodulation* (Carrell & Opie, 1992), the brain would group the sounds and treat them as a single entity. In the case of conflicting modulation, however, separate tones would not be grouped. Data from Carrell and Opie support this theory. They found that phoneme perception was significantly better for the 100-Hz AMTVS sentences than for CAMTVS sentences. In the present study, phoneme perception for CAMTVS sentences was comparable to accuracy for sentences that were amplitude-comodulated at either 50 or 100 Hz. In Carrell and Opie, in which only a few sentences were presented, subjects may have been treating the sinusoids in the CAMTVS sentences as separate sounds. However, the present work indicates that this view needs to be refined. For example, with prior practice and the larger number of test sentences in the present study, subjects may have learned to group the sounds or rely less on auditory grouping.

Figure 6 shows a comparison between the mean phoneme scores in the present study and those of the adults from both Carrell and Opie (1992) and Barker and Cooke (1999). Scores for the unmodulated, 200-Hz-modulation, and conflicting-modulation conditions were higher in the present study. Scores were not higher for the 50- and 100-Hz modulation conditions, in which performance was generally high (>80%) and ceiling effects may have

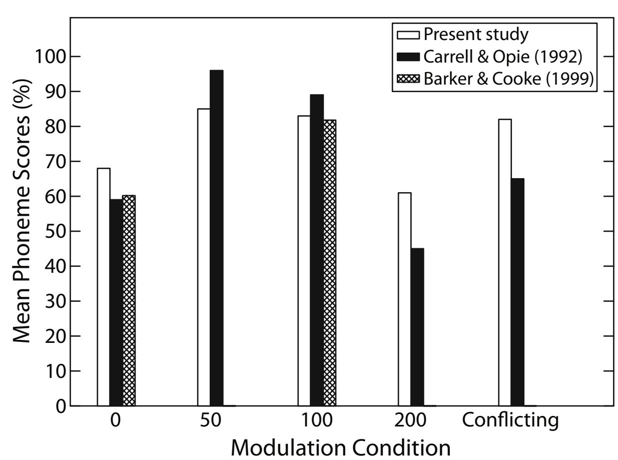

Figure 6. Mean phoneme scores as a function of modulation condition from the present study, Carrell and Opie (1992), and Barker and Cooke (1999). Not all conditions were tested in each study.

limited performance. Although experimental differences across the three studies limit direct comparisons, the higher scores for the present study raise interesting possibilities.

Probably the biggest factor that differentiates the present experiment from all previous TVS studies is that listeners had extensive exposure to a variety of TVS sentences. Prior to testing, subjects practiced the task using 24 sentences representing each of the modulation conditions. Also, after hearing a sentence once, they were told what had been said and then heard that same sentence two more times. In previous studies, subjects heard only 1 or 2 sentences in order to demonstrate what they would hear during the experiment (e.g., Barker & Cooke, 1999; Carrell & Opie, 1992; Remez et al., 1981; Remez & Rubin, 1990; Rosner et al., 2003). Also, in the present experiment, subjects heard 30 test sentences, each repeated three times. In most previous studies, only a small number of sentences were used. For example, Carrell and Opie (1992) and Remez and Rubin (1990) used 4 sentences, and Remez et al. (1981) used only 1. An exception would be Barker and Cooke (1999), who used 20 experimental sentences. It is of interest to note that the degree of practice in the present study was minimal in comparison with that used in other experiments incorporating degraded speech stimuli (e.g., Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995). In fact, the duration of exposure in the present study was relatively short and was consistent with exposure in other nontypical listening environments (e.g., listening to a foreign talker).

In summary, although results revealed that phoneme scores were affected by modulation rate, no interaction between age and modulation rate was noted. Children and adults receive the same benefits (or decrements) from the addition of amplitude modulation to unmodulated TVS speech. However, the underlying causes for listeners' performance remain unclear. Given that scores did not improve at lower modulation rates, a CMR-based explanation is unlikely. In addition, improvements for conflicting amplitude-modulated sentences (re: no modulation) call into question a synchrony-based explanation. It is possible that the addition of amplitude modulation results in sounds that are more natural or speech-like, and it is this "naturalness" that causes the brain to understand the signal as speech. Further research is needed to investigate the underlying causes for listeners' improvements when TVS speech is amplitude-modulated. It is also possible that overall performance in the present study was positively influenced by exposure to all types of TVS speech, as well as by the use of verbal feedback during the practice phase. The potential benefit provided by exposure and feedback may have implications for methods used to teach listeners to understand unfamiliar or difficult-to-understand speech. Further study—examining the time course for learning to understand TVS sentences as well as the effects of different feedback methods—is needed to investigate this issue more closely. Finally, additional modulation frequencies should be examined in conflicting modulation conditions to investigate stimulus-based relations between modulated and unmodulated TVS speech and practice effects.

AUTHOR NOTE

This research was supported by NIH Grants F31 DC006582 and P30 DC04662. This article is based on research for a PhD dissertation at the University of Nebraska by the first author. Travel expenses to present these data at the American Auditory Society Meeting in Phoenix, AZ in March 2005 were provided by NIH Conference Grant 2 R13 DC006616. The authors thank Pat Stelmachowicz and three anonymous reviewers for comments on the manuscript. Correspondence concerning this article should be addressed to D. E. Lewis, Boys Town National Research Hospital, 555 North 30th Street, Omaha, NE 68131 (e-mail: lewis@boystown.org).

REFERENCES

Allen, P., & Wightman, F. (1994). Psychometric functions for children's detection of tones in noise. *Journal of Speech & Hearing Research*, **37**, 205-215.

Allen, P., & Wightman, F. (1995). Effects of signal and masker uncertainty on children's detection. *Journal of Speech & Hearing Research*, **38**, 503-511.

Allen, P., Wightman, F., Kistler, D., & Dolan, T. (1989). Frequency resolution in children. *Journal of Speech & Hearing Research*, **32**, 317-322.

Assmann, P. F., & Summerfield, Q. (1987). Perceptual segregation of concurrent vowels. *Journal of the Acoustical Society of America*, **82**(Suppl. 1), S120.

Assmann, P. F., & Summerfield, Q. (1990). Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies. *Journal of the Acoustical Society of America*, **88**, 680-697.

Bankson, N., & Bernthal, J. (1990). *Bankson–Bernthal Test of Phonology*. San Antonio: Special Press.

Barker, J., & Cooke, M. (1999). Is the sine-wave speech cocktail party worth attending? *Speech Communication*, **27**, 159-174.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.

Buss, E., Hall, J. W., III, Grose, J. H., & Dev, M. B. (1999). Development of adult-like performance in backward, simultaneous, and forward masking. *Journal of Speech, Language, & Hearing Research*, **42**, 844-849.

Carrell, T. D., & Opie, J. M. (1992). The effect of amplitude comodulation on auditory object formation in sentence perception. *Perception & Psychophysics*, **52**, 437-445.

Darwin, C. J. (1981). Perceptual grouping of speech components differing in fundamental frequency and onset time. *Quarterly Journal of Experimental Psychology*, **33A**, 185-207.

Darwin, C. J. (1984). Perceiving vowels in the presence of another sound: Constraints on formant perception. *Journal of the Acoustical Society of America*, **76**, 1636-1647.

Darwin, C. J., & Bethell-Fox, C. E. (1977). Pitch continuity and speech source attribution. *Journal of Experimental Psychology: Human Perception & Performance*, **3**, 665-672.

Darwin, C. J., & Hukin, R. W. (1999). Auditory objects of attention: The role of interaural time differences. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 617-629.

Demany, L. (1982). Auditory stream segregation in infancy. *Infant Behavior & Development*, **5**, 261-276.

Denes, P. (1955). Effect of duration on the perception of voicing. *Journal of the Acoustical Society of America*, **27**, 761-764.

Dunn, C. C. (2003). *Auditory temporal capabilities and sentence intelligibility in the elderly*. Unpublished doctoral dissertation, University of Nebraska, Lincoln.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test* (3rd ed.). Circle Pines, MN: American Guidance Service.

Eddins, D. A., & Wright, B. A. (1994). Comodulation masking release for single and multiple rates of envelope fluctuation. *Journal of the Acoustical Society of America*, **96**, 3432-3442.

Eisenberg, L. S., Shannon, R. V., Martinez, A. S., Wygonski, J., & Boothroyd, A. (2000). Speech recognition with reduced spectral cues as a function of age. *Journal of the Acoustical Society of America*, **107**, 2704-2710.

Elliott, L. L. (1986). Discrimination and response bias for CV syllables differing in voice onset time among children and adults. *Journal of the Acoustical Society of America*, **80**, 1250-1255.

Elliott, L. L., Busse, L. A., Partridge, R., Rupert, J., & DeGraaff, R. (1986). Adult and child discrimination of CV syllables differing in voicing onset time. *Child Development*, **57**, 628-635.

Fallon, M., Trehub, S. E., & Schneider, B. A. (2000). Children's perception of speech in multitalker babble. *Journal of the Acoustical Society of America*, **108**, 3023-3029.

Fallon, M., Trehub, S. E., & Schneider, B. A. (2002). Children's use of semantic cues in degraded listening environments. *Journal of the Acoustical Society of America*, **111**, 2242-2249.

Grose, J. H., & Hall, J. W., III (1992). Comodulation masking release for speech stimuli. *Journal of the Acoustical Society of America*, **91**, 1042-1050.

Grose, J. H., Hall, J. W., III, & Dev, M. B. (1997). MLD in children: Effects of signal and masker bandwidths. *Journal of Speech & Hearing Disorders*, **40**, 955-959.

Hall, J. W., III, Buss, E., Grose, J. H., & Dev, M. B. (2004). Developmental effects in the masking-level difference. *Journal of Speech, Language, & Hearing Research*, **47**, 13-20.

Hall, J. W., III, Cokely, J. A., & Grose, J. H. (1988). Combined monaural and binaural masking release. *Journal of the Acoustical Society of America*, **83**, 1839-1845.

Hall, J. W., III, & Grose, J. H. (1990). Comodulation masking release and auditory grouping. *Journal of the Acoustical Society of America*, **88**, 119-125.

Hall, J. W., III, Grose, J. H., Buss, E., & Dev, M. B. (2002). Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. *Ear & Hearing*, **23**, 159-165.

Hall, J. W., III, Haggard, M. P., & Fernandes, M. A. (1984). Detection in noise by spectro-temporal pattern analysis. *Journal of the Acoustical Society of America*, **76**, 50-56.

Hartley, D. E. H., Wright, B. A., Hogan, S. C., & Moore, D. R. (2000). Age-related improvements in auditory backward and simultaneous masking in 6- to 10-year-old children. *Journal of Speech, Language, & Hearing Research*, **43**, 1402-1415.

Hill, P. R., Hartley, D. E. H., Glasberg, B. R., Moore, B. C. J., & Moore, D. R. (2004). Auditory processing efficiency and temporal resolution in children and adults. *Journal of Speech, Language, & Hearing Research*, **47**, 1022-1029.

Johnson, C. D., Benson, P. V., & Seaton, J. B. (1997). *Educational audiology handbook*. San Diego: Singular.

Johnson, C. E. (2000). Children's phoneme identification in reverberation and noise. *Journal of Speech, Language, & Hearing Research*, **43**, 144-157.

Kent, R. T. (1997). *The speech sciences*. San Diego: Singular.

Kenworthy, O. T., Klee, T., & Tharpe, A. M. (1990). Speech recognition ability of children with unilateral sensorineural hearing loss as a function of amplification, speech stimuli and listening condition. *Ear & Hearing*, **11**, 264-270.

Max, L., & Onghena, P. (1999). Some issues in the statistical analysis of completely randomized and repeated measures designs for speech, language, and hearing research. *Journal of Speech, Language, & Hearing Research*, **42**, 261-270.

McAdams, S., & Bertoncini, J. (1997). Organization and discrimination of repeating sound sequences by newborn infants. *Journal of the Acoustical Society of America*, **102**, 2945-2953.

Miller, G. A., Heise, G. A., & Lichten, W. (1951). The intelligibility of speech as a function of the context of test materials. *Journal of Experimental Psychology*, **41**, 329-335.

Miller, G. A., & Licklider, J. C. R. (1950). The intelligibility of interrupted speech. *Journal of the Acoustical Society of America*, **22**, 167-174.

Moore, B. C. J. (2003). *An introduction to the psychology of hearing*. San Diego: Academic Press.

Nilsson, M. J., Soli, S. D., & Gelnett, D. J. (1996). *Development of the Hearing in Noise Test for Children (HINT-C)*. Los Angeles: House Ear Institute.

Nittrouer, S., & Boothroyd, A. (1990). Context effects in phoneme and word recognition by young children and older adults. *Journal of the Acoustical Society of America*, **87**, 2705-2715.

Perri, J. L. (1997). *Sentence intelligibility in amplitude-modulated noise by listeners with normal hearing*. Unpublished master's thesis, University of Nebraska, Lincoln.

Remez, R. E., & Rubin, P. E. (1990). On the perception of speech from time-varying acoustic information: Contributions of amplitude variation. *Perception & Psychophysics*, **48**, 313-325.

Remez, R. E., Rubin, P. E., Pisoni, D. B., & Carrell, T. D. (1981). Speech perception without traditional speech cues. *Science*, **212**, 947-950.

Repp, B. H. (1982). Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin*, **92**, 81-110.

Rosner, B. S., Talcott, J. B., Witton, C., Hogg, J. D., Richardson, A. J., Hansen, P. C., & Stein, J. F. (2003). The perception of "sine-wave speech" by adults with developmental dyslexia. *Journal of Speech, Language, & Hearing Research*, **46**, 68-79.

Schneider, B. A., Trehub, S. E., Morrongiello, G. A., & Thorpe, L. A. (1986). Auditory sensitivity in preschool children. *Journal of the Acoustical Society of America*, **79**, 447-452.

Schneider, B. A., Trehub, S. E., Morrongiello, G. A., & Thorpe, L. A. (1989). Developmental changes in masked thresholds. *Journal of the Acoustical Society of America*, **86**, 1733-1742.

Sensimetrics (1998). SpeechStation2 [Computer software]. Cambridge, MA: Author.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, **270**, 303-304.

SonicFoundry (1998). Sound Forge Version 4.0E [Computer software]. Madison, WI: Author.

Stelmachowicz, P. G., Hoover, B. M., Lewis, D. E., Kortekaas, R. W. L., & Pittman, A. L. (2000). The relation between stimulus context, speech audibility, and perception for normal-hearing and hearing-impaired children. *Journal of Speech, Language, & Hearing Research*, **43**, 902-914.

Tice, R., & Carrell, T. [D.] (1996). MakeTape (Version 1.2). Retrieved from hush.unl.edu/LabResources.html.

Tice, R., & Carrell, T. [D.] (1997). Tone16: Tone-analog waveform synthesizer. Retrieved from hush.unl.edu/LabResources.html.

Trainor, L. J., & Adams, B. (2000). Infants' and adults' use of duration and intensity cues in the segmentation of tone patterns. *Perception & Psychophysics*, **62**, 333-340.

Wightman, F. L., Callahan, M. R., Lutfi, R. A., Kistler, D. J., & Oh, E. (2003). Children's detection of pure-tone signals: Informational masking with contralateral maskers. *Journal of the Acoustical Society of America*, **113**, 3297-3305.

NOTES

1. To approximate the original formant amplitudes, amplitudes of the modulated tones could be lowered by 6 dB or set to 0. Therefore, each of the tones could be present at one of two levels or absent.

2. It is possible that procedures may have resulted in different error rates for children and adults. However, they are unlikely to have differentially affected individual subjects' performance across modulation conditions.

APPENDIX

Practice Sentences	Experimental Sentences	Deleted Sentences
Dark clouds bring rain	Boats sail at sea	Put your shoes on
You can't make me	Big cars are loud	Cats catch slow mice
The rain came down	Let me have it	Men wear long pants
The jar was full	Cows give sweet milk	Knock these blocks down
I don't want to	Birds like long worms	I don't feel good
Keep your hands off	He cut his finger	What a fun ride
Apes swing from trees	Sun melted the snow	
The dog came back	The glass bowl broke	
She found her purse	I know a song	
Have a nice day	The little baby sleeps	
The boy was asleep	The bus stopped suddenly	
The girls are reading	Cooks make hot food	
Pour me more tea	Go back to bed	
Turn the light off	Smart bears sleep late	
Do you know what	The floor looked clean	
Fresh bread smells great	I wrecked my bike	
The movie finished early	My tooth is loose	
Warm sun feels good	The truck carried fruit	
Where are we going	Bug bites will itch	
Toy trains move fast	That one is mine	
Can you see me	Dad buys new shirts	
Dump trucks fill holes	They wanted some potatoes	
It was my turn	Tell mom those jokes	
It's time for lunch	Tall men jump high	
	Lemons grow on trees	
	They took some food	
	Stay off the hill	
	Buy me that book	
	Most boys play ball	
	Blue planes fly far	

(Manuscript received March 14, 2006;
revision accepted for publication March 13, 2007.)