ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Recognition of Sentences With Complex Syntax in Speech Babble by Adolescents With Normal Hearing or Cochlear Implants

Susan Nittrouer[a] and **Joanna H. Lowenstein**[a]

[a] Department of Speech, Language, and Hearing Sciences, University of Florida, Gainesville

ARTICLE INFO

Article History:
Received July 13, 2022
Revision received October 17, 2022
Accepted November 22, 2022

Editor-in-Chief: Peggy B. Nelson
Editor: Rachael Frush Holt

https://doi.org/10.1044/2022_JSLHR-22-00407

ABSTRACT

Purpose: General language abilities of children with cochlear implants have been thoroughly investigated, especially at young ages, but far less is known about how well they process language in real-world settings, especially in higher grades. This study addressed this gap in knowledge by examining recognition of sentences with complex syntactic structures in backgrounds of speech babble by adolescents with cochlear implants, and peers with normal hearing.
Design: Two experiments were conducted. First, new materials were developed using young adults with normal hearing as the normative sample, creating a corpus of sentences with controlled, but complex syntactic structures presented in three kinds of babble that varied in voice gender and number of talkers. Second, recognition by adolescents with normal hearing or cochlear implants was examined for these new materials and for sentence materials used with these adolescents at younger ages. Analyses addressed three objectives: (1) to assess the stability of speech recognition across a multiyear age range, (2) to evaluate speech recognition of sentences with complex syntax in babble, and (3) to explore how bottom-up and top-down mechanisms account for performance under these conditions.
Results: Results showed: (1) Recognition was stable across the ages of 10–14 years for both groups. (2) Adolescents with normal hearing performed similarly to young adults with normal hearing, showing effects of syntactic complexity and background babble; adolescents with cochlear implants showed poorer recognition overall, and diminished effects of both factors. (3) Top-down language and working memory primarily explained recognition for adolescents with normal hearing, but the bottom-up process of perceptual organization primarily explained recognition for adolescents with cochlear implants.
Conclusions: Comprehension of language in real-world settings relies on different mechanisms for adolescents with cochlear implants than for adolescents with normal hearing. A novel finding was that perceptual organization is a critical factor.
Supplemental Material: https://doi.org/10.23641/asha.21965228

Children born with severe-to-profound hearing loss have made remarkable gains in their abilities to acquire spoken language. With early intervention and cochlear implants (CIs), it is now routine for these children to perform within the typical range on standardized tests of language proficiency, with the "typical range" defined as better than 1 *SD* below the normative mean (e.g., Boons et al., 2013: Bradham et al., 2018; Geers et al., 2016; Rudge et al., 2022). These outcomes leave many clinicians and educators feeling secure in the belief that children with hearing loss who received CIs and appropriate early intervention will fare as well as their peers with normal hearing (NH) in the school setting, requiring little, if any, accommodations.

That perspective, however, discounts the very real challenges that remain for these children. Assessments with standardized test instruments take place in quiet settings, one-on-one with the clinician. Children usually have

Correspondence to Susan Nittrouer: snittrouer@phhp.ufl.edu. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

as long as they need to answer items on those assessments. But in the classroom, gym, cafeteria, or any other school setting, many people occupy the space, with lots of those people talking simultaneously. In addition, communication is ongoing; pause durations are not scaled to match the processing requirements of adolescents with hearing loss. The study described here was designed to assess these challenges in real-world, real-time communication for adolescents who use CIs. Specifically, the abilities of adolescents with CIs to recognize complex language structures in backgrounds consisting of speech babble were examined. With that broad goal in mind, three specific objectives were formulated:

(1) to assess the stability of speech recognition in challenging listening environments across the late-elementary and middle school years for adolescents with NH or with CIs;
(2) to evaluate speech recognition in babble backgrounds as a function of voice gender and number of talkers generating the babble, for sentences with complex syntax; and
(3) to explore potential mechanisms accounting for performance under these conditions.

It was anticipated that the outcomes of these analyses would inform the design of interventions aimed at mechanisms underlying the processing of complex language in real-world environments by school-age children and adolescents with CIs.

Recognizing Speech in Challenging Listening Environments

Research into the mechanisms underlying speech recognition in challenging listening environments has been ongoing for nearly as long as speech perception itself has been studied (e.g., Cherry, 1953; G. A. Miller, 1947). As currently understood, this process can be described as relying on properties of the sensory input, including target signal and competing background, as well as on characteristics of the listener, including the linguistic and cognitive functioning of the individual. First, the listener must apply principles of auditory scene analysis (Bregman, 1990; Darwin, 2008; Lorenzi et al., 2006; Tejani et al., 2017) to segregate the target signal from background, competing sounds. Effective segregation of the target signal occurs when the listener can take advantage of fluctuations in both the temporal and spectral envelopes of the background in order to recover bits of that target, a process typically termed *glimpsing* or *dip listening* (e.g., Alcántara et al., 2004; Howard-Jones & Rosen, 1993; Meddis & Hewitt, 1992; Rosen et al., 2013). The output of this step is an underspecified representation of the intended signal, requiring the listener to utilize cognitive and language resources to reconstruct the original message (Pichora-Fuller et al., 1995). These resources are taxed to a much greater extent under conditions of competing backgrounds than when the signal is fully accessible, as happens when a listener with NH is communicating in a quiet environment. The manner in which these cognitive and language resources are applied, to what extent, and under what circumstances have been the focus of ample investigation ever since G. A. Miller et al. (1951) observed that at a constant signal-to-noise ratio, words are recognized correctly more often if they are in sentence contexts, rather than in isolation: Even when "bottom-up" factors remained constant, recognition varied depending on context. Since that early demonstration of "top-down" effects, it has consistently been observed that the more predictable the linguistic context, the better a speech signal is recognized in a competing background (e.g., Boothroyd & Nittrouer, 1988; Buss et al., 2019; Kalikow et al., 1977; Nittrouer & Boothroyd, 1990; Smits & Zekveld, 2021). When it comes to cognitive factors, verbal working memory has most often been explored in relation to speech recognition in competing backgrounds. A stronger working memory capacity facilitates better recognition under those conditions, a finding that has largely been observed in older listeners (e.g., Akeroyd, 2008; Ben-David et al., 2019; Gordon-Salant & Cole, 2016; Humes, 2007; Schneider et al., 2010), although it has also been reported for children (MacCutcheon et al., 2019; McCreery et al., 2017, 2019; Sullivan et al., 2015). The idea that working memory capacity would be associated with speech-in-noise recognition is intuitively appealing, because it seems reasonable that an advantage would be accrued by listeners who can store more bits of the recovered target in a memory buffer for integration across time (Lad et al., 2020). One specific model of how working memory capacity contributes to speech recognition under difficult listening conditions specifies that other cognitive resources are taxed to a lesser degree when working memory better supports the process; that is, there is greater ease of language understanding (e.g., Rönnberg et al., 2013).

Listeners With Hearing Loss

Individuals with hearing loss encounter substantially greater challenges recognizing speech in competing backgrounds than do individuals with NH (Plomp, 1978). Nonetheless, the mechanisms underlying recognition for these listeners are not yet completely understood, at least partly because these mechanisms have often been assumed to be similar for listeners with NH and listeners with hearing loss. If so, mechanisms accounting for speech recognition in challenging conditions for listeners with hearing loss could be inferred from findings obtained from listeners with NH. But listeners with hearing loss have deficits in auditory functioning that go beyond the loss of sensitivity. Recognition is poorer for listeners with hearing loss than for those with NH when signals are presented at the same signal-to-noise ratio, even when recognition scores in quiet are similar across groups (Bernstein & Grant, 2009; Dubno et al., 1984). Because recognition is selectively poorer in noise for listeners with hearing loss—all else being equal—the mechanisms explaining recognition in noise are likely different for listeners with NH versus hearing loss. In particular, listeners with hearing loss have diminished spectral and temporal resolution (e.g., Bacon & Viemeister, 1985; Henry & Turner, 2003; Horn et al., 2017; Zheng et al., 2017), which can negatively impact speech recognition for target signals embedded in the speech of other talkers because it diminishes listeners' abilities to engage in glimpsing. Furthermore, explanatory mechanisms vary in the extent of involvement across recognition probabilities (Boothroyd & Nittrouer, 1988; Lewis et al., 2021); because these recognition probabilities are lower for listeners with hearing loss than for listeners with NH those mechanisms must differ for that reason alone. Consequently, it is essential that speech recognition in challenging conditions be explored specifically for listeners with hearing loss, rather than being inferred from studies of listeners with NH, with or without simulated hearing loss.

With that said, it remains the case that mechanisms underlying speech recognition in competing backgrounds might be the same for listeners with NH and for adults who lost their hearing after childhood. For these listeners, typical strategies for processing speech in noise would have developed before the onset of hearing loss. Evidence for this proposal comes from Litovsky et al. (2010) who reported that adults with late-onset hearing loss who received CIs demonstrated appropriate binaural processing abilities, but those with congenital hearing loss did not. In general, it could be expected that listeners with congenital hearing loss lack opportunities to acquire typical auditory processing strategies. On the other hand, it could simply be that development of strategies for listening to speech in noise is delayed as a consequence of delayed access to auditory input. This idea follows from the notion of "hearing age," which pegs the start of auditory learning to the initiation of amplification (e.g., Hayes et al., 2009; Holmes et al., 2018; Rudge et al., 2022). According to this idea, development of auditory functions proceeds at typical rates once amplification is received, albeit with milestones reached at older chronological ages. If this is indeed the case, it would mollify concern regarding children with hearing loss, especially those with CIs, being able to handle noisy school environments because it would mean they would eventually "catch up" in their abilities to do so. If this catching up was accomplished by adolescence, skill at recognizing speech in competing backgrounds would reach equivalent status at roughly the same time that the language of learning became most complex.

One study that investigated the possibility that children with hearing loss might catch up to their peers with NH in their abilities to recognize speech in competing backgrounds was conducted by Walker et al. (2019). In that study, children with hearing loss who used hearing aids were tested between first and fourth grade, along with peers with NH. Bamford–Kowal–Bench (Bench et al., 1979) sentences designed for use with children were presented in a multitalker babble. These sentences have simple syntactic structures. The dependent measure was the threshold in dB signal-to-babble ratio for recognition of roughly half of the key words in the sentences. Both groups of children showed improvements of a little more than 2 dB across the 3-year span. For children with NH, the majority of that improvement occurred early in the study and then slowed; for children with hearing loss, this improvement was relatively linear across the grades. These findings suggest different possibilities regarding developmental changes in speech-in-noise recognition after age 10 years for children with NH and for those with hearing loss. For children with NH, it may be that they have reached an asymptote in their speech-in-noise recognition abilities by fourth grade. For children with hearing loss, it may be that they continue to improve and catch up in subsequent years. These possibilities were tested in the study reported here by examining speech-in-noise performance for children with NH and for those with hearing loss from fourth to eighth grade.

The Current Study

This study was designed to examine speech recognition in the presence of other talkers for adolescents with CIs, and peers with NH, for the kind of complex language they are likely to encounter in middle and high school settings. The first specific objective of the study was to assess the stability of recognition abilities in challenging conditions across grades in late elementary and middle school (10–14 years), so at ages older than the range included in the Walker et al. (2019) study.

The second objective of this study was to assess how well these adolescents could recognize speech in challenging listening environments with the kind of complex language structures they are likely to encounter in middle and high school. Speech babble served as the competing background for testing. Target sentences were produced by a male talker, while voice gender and number of talkers were manipulated to create the babble. It could be predicted that listeners would be most successful at segregating the target signal from background babble when that babble was derived from a single female talker.

Significant differences exist between the speech of male and female adults in both fundamental frequency (so in harmonic structure) and vocal-tract length (so in formant structure). Accordingly, listeners are better able to segregate the speech of a target talker from the background speech of another talker when the talkers differ in sex—or as typically termed, *voice gender* (Skuk & Schweinberger, 2014). For listeners with NH this effect is observed (Darwin et al., 2003) starting at a young age (Nagels et al., 2021). For babble derived from a single male talker, the process should be more difficult. Under these conditions, it not only becomes harder to separate the signal components—because of similarities in fundamental frequency and vocal-tract length—it is harder to determine which detected components should be allocated to each source. Finally, the chore of segregating the target signal from background babble should be most difficult when that babble is the product of several talkers, because in these conditions it is maximally difficult to perceptually disentangle the signal components of each talker.

Sentence materials were designed with carefully controlled, complex syntactic structures. There is a long history of examining recognition of speech in noise using sentence materials. Several sets of sentences have been developed and standardized for use in the testing of speech-in-noise abilities (e.g., Kalikow et al., 1977; Nilsson et al., 1994; Spahr et al., 2012). Nonetheless, few of these materials have specifically controlled for syntactic structure, beyond sentence length. Where children are concerned, sentence materials are frequently selected explicitly to be simple in structure (and so short) so that sensory factors can be investigated alone, separately from language abilities (Walker et al., 2019). For this reason, new materials were developed for testing in this study, with sentence structures designed to examine how complex syntax may interact with background babble. All sentences consisted of two clauses; the way these clauses were combined was balanced across materials. Much as is the case for background noise, syntactic complexity increases the cognitive load of processing speech (e.g., Carroll & Ruigendijk, 2013; Wingfield et al., 2003). Complex syntax increases demands on working memory as sentence length increases (Lyxell & Rönnberg, 1993; Poll et al., 2016; Rönnberg et al., 2013). Effects on recognition in competing backgrounds, however, are not consistent across syntactic types, suggesting that it is not only sentence length that affects cognitive load. Center-embedded clauses with object-first structures appear to be most resistant to background noise, presumably because prosodic patterns are distinct for these sentences (Carroll & Ruigendijk, 2013). Of course, a listener must be able to recognize intonation—as instantiated by fundamental frequency—in order to take advantage of this distinction. Adolescents with hearing loss who use CIs may not be able to do this especially well, a prediction based on outcomes from studies showing that children with CIs demonstrate difficulty recognizing intonation in the speech signal (Peng et al., 2008; Richter & Chatterjee, 2021). Furthermore, sentences with center-embedded clauses are especially complex (Just & Carpenter, 1992; Riches et al., 2010; Smith et al., 1989), which could hinder recognition for children with poor syntactic knowledge. Thus, sentences with complex syntactic structures constrain recognition especially well, as long as the listener has adequate syntactic knowledge.

The third goal of this study was to examine mechanisms underlying the recognition of speech in babble for adolescents with CIs. Several possible mechanisms were included in this investigation. One mechanism examined was spectral resolution, which should be related to how well these listeners could segregate the target signal from the background babble. This mechanism fits under the rubric of bottom-up mechanisms. This ability to segregate spectral components of complex signals has been shown to be related to speech-in-noise perception for adults with acquired hearing loss who receive CIs (Hong & Turner, 2006), so was a reasonable choice for investigation with these adolescents with congenital hearing loss and CIs.

The second kind of mechanism examined involved language knowledge and skill related to vocabulary, syntactic structures, and phonological sensitivity. All these factors support better speech recognition in challenging backgrounds by increasing predictability, and this set constitutes top-down mechanisms. Verbal working memory was another top-down process examined as a potential facilitating factor for speech recognition in babble by adolescents with CIs, but care was taken to use a measure requiring more cognitive processing than digit span. Work with adults has shown that simple span metrics are not good predictors of speech recognition, because they do not tax cognitive and linguistic resources sufficiently (Rönnberg et al., 2013). In particular, once sentence length extends beyond a typical word span (usually around seven) syntactic constraints become critical to facilitating recognition.

Finally, another less-frequently studied mechanism was examined: perceptual organization. When faced with a background of speech babble, the first chore for the listener is to segregate as much of the target signal as possible from that babble. Then, the listener must be able to combine these signal portions in order to recover the perceptual object, which in this case is the target speech. This ability to organize sensory information for the purpose of achieving recognition of the original object defines perceptual organization. Its operation is essential when the signal is degraded in some way, and thus is typically investigated under conditions of signal degradation (Remez et al., 1994; Shafiro et al., 2018). Listeners with hearing loss who use CIs would be expected to have poorer spectral resolution than listeners with NH (Horn et al., 2017; Kirby et al., 2015; Landsberger et al., 2018; Nittrouer et al., 2021), which means they have sparser signal information to piece together, a factor that makes it difficult to compare abilities to organize the structure that is recovered across listener groups. To "level the playing field," Nittrouer, Kuess, and Lowenstein (2015) presented to children with NH and those with CIs sentences processed as sinewave speech: that is, the center frequencies of the first three formants were replaced with sinewaves. Overall, this process should have provided signals of equivalent spectral distinctiveness to children in these two groups. This approach revealed that the children with CIs performed as well as the children with NH, with similar within-group variability. As a result, it was concluded that children with CIs generally have adequate abilities to integrate spectral components of the signal structure they manage to segregate from the background noise, and organize that structure in order to derive a linguistically meaningful percept. In this current study, we examined the extent to which this skill accounts for speech-in-babble recognition for adolescents with CIs. Of special relevance is the fact that the children in that 2015 experiment were largely the same individuals as in the study reported here, involving adolescents who were just completing eighth grade.

Preliminary Study

This study was undertaken to establish the efficacy of the sentence materials that would be used with adolescents in the main experiment. The two principal questions addressed in the preliminary study were whether the complex sentences we developed would show differences in recognition probabilities across syntactic structures and whether the babble conditions would show differences across voice gender and number of talkers used to generate the background babble. This preliminary study was conducted with young adults with normal speech, language, and hearing to assess effects in this population, which may be viewed as the standard for performance.

Method

Participants

Twenty-five young adults between the ages of 19 and 35 years (M = 22 years) participated in this study. All were university students at the time of testing, and had no histories of speech, language, or hearing disorders. All were native speakers of American English, and all passed hearing screenings of the octave frequencies from 0.25 to 8.0 kHz presented at 20 dB hearing level to each ear separately.

Equipment

For creation of the sentence materials, recordings were made in a soundproof booth, directly onto a hard drive using an AKG C535 EB microphone, a Shure M268 amplifier, and a Creative Laboratories soundcard. Hearing screenings were performed with a Welch Allyn RM262 audiometer and TDH-39 headphones. Perceptual testing took place in a soundproof booth. All stimuli were presented through a computer, with a Creative Laboratories soundcard. A Roland MA-12C powered speaker was used and was placed 1 m in front of the listener at 0° azimuth. Testing was video- and audio-recorded using a Sony HDR-XR550V video recorder, and listeners wore Sony FM transmitters to ensure good sound quality on the recordings. Receivers for these FM systems connected to the camera.

Stimuli

Five sets of sentences were presented in this testing. Complex sentences were presented in three different speech maskers (i.e., babble), comprising three sets, and two sets of simple sentences were presented in speech-shaped noise.

Complex Sentences in Speech Babble

Seventy-eight sentences were generated: 25 sentences in each of three categories to be used for testing, and one sentence in each of those categories to be used for practice. All sentences consisted of two clauses. For two categories of sentences, one clause was always subordinate to the other. For the third category, the two clauses were equal, joined by the conjunction *and*. Words used in these sentences were selected to have age of acquisition estimates of younger than 8 years (Kuperman et al., 2012), and all sentences had between eight and 11 words in them. Stimuli were recorded at 44.1 kHz and 16-bit digitization. The sentences can be found in the Appendix.

Subject–object (SO) sentences. These sentences were structured such that the subject of the main clause was the object of the subordinate clause (e.g., *The snow that the children shoveled blew across the field.*). The subordinate clause (*that the children shoveled*) is always center embedded, and the subordinate clause is retrospective in the sense that the object of that clause (*that*) appears first in the embedded clause, before the subject–verb pair. This structure makes these sentences especially complex, but also highly constraining. Furthermore, these sentences have a distinct intonation pattern, highlighting that center-embedded clause. Based on these considerations it was predicted that sentences in this set would offer the strongest constraints, so would support recognition most strongly.

Other (OTH) sentences. This set of sentences comprised three structures, all presumed to provide linguistic constraints of equivalent magnitude (e.g., Smith et al., 1989): (1) Object–object sentences (e.g., *The secretary sat at the desk that was painted red.*), (2) object–subject sentences (e.g., *The lady made the sign that said "free kittens."*), and (3) subject–subject sentences (e.g., *The girl who cleaned the kitchen ate some lunch.*). Except for the subject–subject sentences, these sentences are right-branching, which is a simpler construction than sentences with center-embedded clauses. Even in the subject–subject sentences, however, construction is simpler—so less constraining—than in the SO sentences, because the subject–verb–object structure is maintained in both the main and subordinate clauses. All three kinds of sentences in this category have a less-distinct inflectional pattern than the SO sentences. These sentences were predicted to provide linguistic constraints intermediate to those provided by the SO sentences and by the third group, conjoined clauses.

Conjoined-clauses (CC) sentences. These sentences consisted of two clauses of equal standing, joined with the conjunction *and* (e.g., *The cat ate the mouse and played with yarn.*). These sentences were predicted to impose the weakest constraints on recognition, so they would result in the lowest scores.

At the start of testing with each subject, the 75 complex sentences to be used in testing were randomly assigned to each of the three kinds of speech babble, creating three sets of these sentence materials. The babble was derived from three male talkers and three female talkers producing the coordinate response measure (CRM) sentences (Bolia et al., 2000). These sentences have been used in the past to investigate effects on recognition of gender and number of talkers producing the babble (e.g., Humes et al., 2017). During testing, sentences of each of the three kinds of syntactic structures were evenly distributed across each of three babble conditions: a single male talker, a single female talker, and a six-talker babble combination. Percent-correct recognition generally declines as the number of talkers contributing to the babble increases up to six talkers, and remains stable with the inclusion of additional talkers (Simpson & Cooke, 2005). In the single talker conditions, sentences serving as babble were randomly selected from each of the three talkers in that gender group. CRM sentences were temporally reversed as a way of minimizing informational masking, so that effects could largely be attributed to energetic masking. Although informational masking constitutes a challenge that adolescents encounter in real-world settings, it would have introduced an additional variable that would likely have interacted with the variables of primary interest in this study (type of babble and syntactic complexity) in ways that would have made outcomes difficult to interpret. In particular, evidence shows that the amount of informational masking available is inversely related to the similarity between target and background signals (Durlach et al., 2003). Accordingly, informational masking would have differed across babble types, confounding any assessment of the abilities of listeners to take advantage of dips in the background signal.

Simple Sentences in Speech-Shaped Noise

These adults were presented with two sets of sentences that have been used in the past: (1) four-word sentences that are syntactically correct, but semantically anomalous (e.g., *Dumb shoes will sing*) and (2) five-word sentences that are syntactically correct and semantically informative (e.g., *The book tells a story*). Recognition scores for these sentences provided a standard against which to compare performance with the newly developed sentences that allowed an assessment of the generalizability of outcomes.

The four-word sentences closely followed the format first implemented by Boothroyd and Nittrouer (1988). These sentences are short, almost exclusively containing content words and always conforming to simple syntactic constructions. Sentences with these same structural properties have been used in the past (e.g., Loebach & Pisoni, 2008; Moberly & Reed, 2019; Stelmachowicz et al., 2000) to examine the effects of syntactic constraints separately from semantic constraints. The five-word sentences were culled from the larger Hearing-in-Noise Test (HINT) corpus (Nilsson et al., 1994). The specific sentences used here were selected to contain exactly five words and to be especially rich in semantic context. In addition to broad use with adult listeners, both sets of sentences have been used specifically with children, so are known to be within children's abilities to recognize (e.g., Eisenberg et al., 2000; Nittrouer & Boothroyd, 1990; Nittrouer, Kuess, & Lowenstein, 2015; Nittrouer & Lowenstein, 2010). Twenty-six of each kind of sentence were included: 25 for testing and one for practice. All sentences used are listed in the Appendix. Sentences were produced by the same male talker as the one who produced the complex sentences, during the same recording session. These sentences were processed such that the long-term average spectrum of each set of sentences (four- or five-word) was derived and used to shape noise for that set.

Testing Procedure

Procedures were approved by the institutional review board of the University of Florida. All sentence materials were presented at a −3 dB signal-to-noise (or signal-to-babble) ratio. This level is associated with response probabilities in a range around 50% correct (Boothroyd & Nittrouer, 1988; Gordon-Salant & Cole, 2016), minimizing the likelihood of obtaining ceiling or floor effects.

A MATLAB program was written to control stimulus presentation and keep track of order of presentation. The babble or speech-shaped noise was calibrated to an overall level of 68 dB sound pressure level before testing. At the start of testing, the order of presentation of the five sets of sentences was randomized, with the provision that the first, third, and last sets would be the complex sentences in babble and the second and fourth sets would be the four- and five-word sentences in speech-shaped noise. The order of presentation of sentences within each set was randomized for each subject individually.

At the start of testing with each set, the practice sentence was presented in quiet. The subject repeated the sentence. It was then presented in either the babble or noise background, and the subject repeated it again. During testing, sentences were presented only once in the babble or noise condition and the subject repeated the sentence. The software kept track of the order of sentence presentation. After testing in all babble and noise conditions, all sentences were presented in quiet for recognition.

Scoring Procedure

A MATLAB program facilitated scoring. The experimenter watched the testing video on one monitor, while the scoring software was open on the other monitor. The software displayed the words in each sentence with a box under each word, one sentence at a time, and the experimenter checked the boxes corresponding to words that were recognized correctly. The articles *a* and *the* were considered interchangeable, and *who* was considered an acceptable substitution for *whom*. The dependent measure was percent correct words.

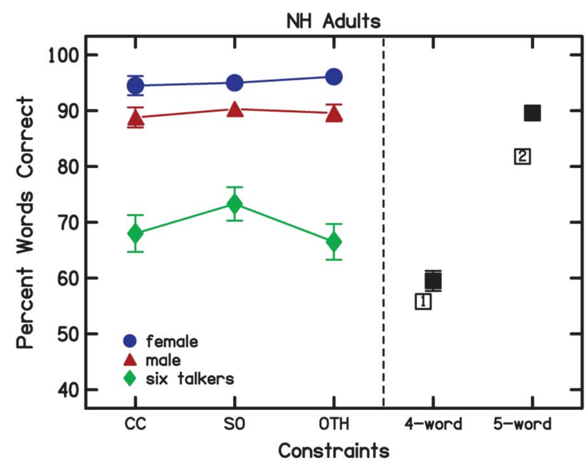
Two experimenters scored each subjects' responses independently and the software compared outcomes for each set of sentences, indicating exactly which words were discrepant. For the newly developed sentences, this meant scores were compared within each babble condition. If scores for a given subject showed agreement lower than 95% within a condition, the experimenters could go back and joint score. That did not occur with any of these subjects. When a mean reliability score was computed across the three babble conditions, this procedure provided an overall metric of interrater reliability for these newly developed sentences. Scores from Scorer 1 in each case were used for further analyses.

In addition, a split-halves comparison was performed for the newly developed sentences. In this procedure, percent correct word recognition scores for odd and even sentences presented in each babble condition were computed separately for each subject. Mean scores across the three babble conditions were then compared.

Results

Adults could recognize all sentences in quiet with 100% accuracy. Figure 1 shows mean percent correct word scores and standard errors for these 25 adults across each of the five sentence types for both the babble and noise conditions. In addition to these results, mean scores from two earlier studies using the 4- and 5-word sentence materials with adults are shown. This earlier mean score for four-word sentences is from Nittrouer, Tarr, et al. (2015) and comes from 40 adults between the ages of 18 and 38 years. The score for the five-word sentences is from Moberly et al. (2017), and is for 30 adults between the ages of 50 and 82 years. The older age of these listeners likely explains the slightly lower score for these sentences. Nonetheless, sufficient similarity is found between scores in each pair to support the proposal that methods are reproducible. Further analyses focused on the new sentences with complex syntactic structures.

Figure 1. Mean percent words recognized correctly across syntactic structures by young adults with normal hearing (with standard error bars). The CC, SO, and OTH sentences were presented in background babble. The four- and five-word sentences were presented in speech-shaped noise. CC = conjoined-clauses; SO = subject-object; OTH = other constructions; 4-word = four-word sentences; 5-word = five-word sentences; 1 = results from Nittrouer, Tarr, et al. (2015); 2: results from Moberly et al. (2017).



Reliability

The test of interrater reliability revealed a mean reliability metric of 99.0%, with the lowest reliability coefficient for any subject being 97.6%. This was considered acceptable reliability.

The split-halves analysis showed that mean correct word recognition across the 25 subjects was 83.4% correct (*SD* = 9.8% correct) for the odd-numbered sentences across the three babble conditions, and 86.3% correct (*SD* = 6.2% correct) for the even-numbered sentences. This difference was not significant and was viewed as representing good internal consistency.

Babble and Sentence-Structure Effects

Data screening revealed that scores across the nine Babble × Sentence-Structure Conditions were generally normally distributed; only the three female babble conditions displayed some negative skew to the distributions (i.e., < −1). Therefore, analysis of variance (ANOVA) was performed first on untransformed data and then on the transformed data, with babble condition and sentence structure as main effects. Results for both analyses are shown in Table 1. Both main effects were found to be significant in the analysis with the untransformed data, and the Babble × Structure interaction was also significant. When this analysis was performed with arcsine-transformed scores (i.e., arcsine(sqrt(prop))), the main effect of babble remained significant, but the main effect of structure was no longer significant. The interaction term, however, was still significant. These outcomes likely reflect the finding apparent in Figure 1 that syntactic structure had an effect on recognition probabilities only in the six-talker babble condition, when overall recognition was lowest.

To investigate outcomes across conditions more closely, paired-samples *t* tests were performed, using the transformed scores. First, mean scores were compared for the three babble conditions, using mean scores across syntactic constructions; this was considered appropriate because of the strong effect of babble condition for both the untransformed and transformed data. The analysis revealed that the score for every babble condition was different from every other babble condition, as shown in Table 2: Recognition was best for the female babble condition, followed by the male babble condition, and finally by the six-talker condition. These differences remained significant when a Bonferroni correction for multiple comparisons was applied (i.e., .05/3 = .017). Effect sizes, indexed by Cohen's *d*s, revealed moderately strong effects.

Because the main effect of syntactic structure was not significant when arcsine-transformed values were used, but there was a significant interaction term, the effects of syntactic structure were examined for each kind of babble separately. Paired-samples *t* tests were performed on transformed mean scores, but only scores from the six-talker condition showed any significant effects. For the six-talker babble condition, two of the three *t* tests were significant, as shown in Table 3: those comparing recognition probabilities for the SO structure to the other two kinds of structure. A Bonferroni correction for multiple comparisons was again applied, and both of these results remained significant.

Table 1. Results of analyses of variance performed with untransformed and arcsine-transformed scores from adults in the Preliminary Study.

Effect	*df*	*F*	*p*	η^2
Untransformed				
Babble	2, 48	92.35	< .001	.794
Syntactic structure	2, 48	3.95	.026	.141
Babble × Structure	4, 96	2.91	.025	.108
Arcsine transformed				
Babble	2, 48	114.17	< .001	.826
Syntactic structure	2, 48	1.57	.219	—
Babble × Structure	4, 96	2.71	.035	.101

Table 2. Outcomes of paired-samples *t* tests across babble conditions, using arcsine-transformed scores from adults.

Contrast	*df*	*t*	*p*	*d*
Female vs. male babble	24	5.83	< .001	1.04
Female vs. six-talker babble	24	14.90	< .001	2.58
Male vs. six-talker babble	24	8.71	< .001	1.82

Note. All *p* values are for two-tailed tests. All results remain significant with the application of a Bonferroni correction for multiple comparisons.

Discussion

This preliminary study was undertaken to develop new sentence materials and babble backgrounds for examining speech-in-babble recognition by adolescents with CIs, along with age-matched peers with NH. The sample of young adults with NH tested in this preliminary study were judged to be representative of this population, based on the fact that scores of these young adults for sentences provided in speech-shaped noise matched scores for the same materials from earlier studies. Items in the new sentence materials were found to have good internal consistency, and scoring methods were reliable. Recognition scores for the three babble conditions revealed expected differences: Recognition was best for the single-talker condition with babble from a female talker (i.e., opposite from the voice gender for target sentences), recognition was poorest for the six-talker babble, and recognition was intermediate for the single-talker condition with babble from a male talker (i.e., the same gender as the talker producing target sentences). When it came to the three kinds of syntactic constructions, differences in scores were not large, but that may reflect the fact that overall recognition scores were very high in two of the three babble

Table 3. Outcomes of paired-samples *t* tests across syntactic structures in the six-talker babble condition, using arcsine-transformed scores from adults.

Contrast	*df*	*t*	*p*	*d*
Subject-object vs. conjoined	24	2.88	.008	0.32
Subject-object vs. other	24	2.66	.014	0.38
Conjoined-clauses vs. other	24	0.52	.612	—

Note. All *p* values are for two-tailed tests. The comparisons involving subject-object constructions remain significant with a Bonferroni correction.

conditions. For the one condition with overall recognition scores farther from the ceiling, reasonable effects were observed among the three kinds of syntactic structures. Because recognition scores were expected to be lower overall for the adolescents to be tested in the main experiment, it was expected that effects of syntactic structure would be observed if these adolescents were indeed sensitive to them.

Main Experiment

This main experiment was conducted to enhance our understanding of the problems listeners with hearing loss face when listening to speech in competing backgrounds. The focus was on sentences with complex syntactic structures, defined as having two clauses; these were the sentences developed in the preliminary study. The competing background of primary interest was the speech of talkers other than the target talker, with a focus on the effects of talker gender and number of talkers. Finally, the listeners with hearing loss who were of interest in this report were adolescents who were born with severe-to-profound hearing loss and developed language through CIs. The syntactic constructions and babble conditions represent the environments these adolescents with hearing loss undoubtedly encounter on a regular basis as they navigate their school settings.

Method

Participants

Data were collected from 108 adolescents at the end of eighth grade. Fifty-six of these adolescents had NH (28 male, 28 female), defined as auditory thresholds for the octave frequencies from 0.25 to 8.0 kHz better than or equal to 20 dB hearing level. Fifty-two of these adolescents had congenital hearing loss and used CIs (23 male, 29 female). Table 4 displays demographic data, including means, medians, and standard deviations. Of these measures, only age at the time of testing showed a significant difference between groups, $t(106) = 2.96$, $p = .004$, indicating that the adolescents with CIs were slightly older overall than the adolescents with NH. This difference was not considered a problem for several reasons: the adolescents with CIs were the older group, which would presumably give them an advantage; the difference was small in magnitude (3 months); and all adolescents were at the same academic rank (just completing eighth grade).

All but 10 adolescents with NH had participated in a longitudinal study from the time they were toddlers; these additional 10 adolescents were subjects who participated in practice testing designed to ensure all procedures ran smoothly. Subjects in the longitudinal study traveled to the laboratory from across the country, so it was important for student testers to become comfortable with testing procedures before those subjects came to the laboratory. All 10 subjects with NH who were "practice participants" were tested at the end of the training period, when student testers were adequately skilled.

Table 4. Means, medians, and standard deviations for demographic and audiometric measures at eighth grade for adolescents with normal hearing (56) and adolescents with cochlear implants (52).

	NH			CI		
Variable	***M***	***Mdn***	***SD***	***M***	***Mdn***	***SD***
Age at time of testing (months)	173	173	6	176	176	5
Socioeconomic status	37	36	14	33	35	11
Nonverbal IQ	106	106	13	103	99	14
Age of identification				6	3	6
Age at first implant				26	15	29
Pre-implant better-ear PTA				100	100	16
Aided 4-frequency thresholds				20	20	5

Note. Age is given in months; socioeconomic status is on a 64-point scale; nonverbal IQ is a standard score; NH = normal hearing; CI = cochlear implant; pre-implant better-ear PTA = pure-tone average 3-frequency (0.5, 1.0, and 2.0 kHz) threshold in dB hearing level; aided 4-frequency thresholds = mean thresholds in dB hearing level for 4 frequencies (0.5, 1.0, 2.0, and 4.0 kHz).

To be included in the longitudinal study from the start, when they were toddlers, children had to have had no medical problems that would be expected to delay language acquisition, other than hearing loss in the case of children with CIs. English was the only language spoken to the children at home. Parents had NH or hearing that was readily corrected to normal levels with hearing aids. Intervention up to school age had to focus on spoken language, although it could include sign language as additional support. All parents confirmed that their goals for their children were that they could attend mainstream educational programs without the need for sign language interpreters, and all these children were in such settings from kindergarten until the time of this testing, at the end of eighth grade. Further information on criteria for participation and subject characteristics can be found in other publications (e.g., Nittrouer, 2010; Nittrouer et al., 2018).

Socioeconomic status (SES) was assessed using a two-factor scale on which occupation and highest educational attainment are ranked from 1 to 8, lowest to highest. These scores are multiplied together, and the product serves as the SES index. An SES index was obtained from each parent, and the highest value was used as the SES metric for the family (Nittrouer & Burton, 2005). SES scores in Table 4 indicate that the average participant had at least one parent with a 4-year university degree, and groups did not differ on overall SES. The lowest SES score for a child in each of the NH and CI groups was 12, which indicates a parent who likely works in a retail-sales job or other service profession. The highest SES scores were 64 for these groups, which represents a position as a medical doctor, chief executive officer of a large corporation, or equivalent. Thus, a wide SES range was represented.

Nonverbal cognitive abilities (i.e., nonverbal IQs) were assessed with the Leiter International Performance Scale–Revised (Roid & Miller, 2002). Standard scores for the "Brief IQ" are reported here. This composite score is derived from the individual scores of four subtests: figure-ground perception, form completion, sequencing abilities, and repeated patterns recognition. This set of measures is meant to capture the contributions of nonverbal cognition to fluid reasoning, including visual selective attention and memory. No adolescent had a score below 80, and scores given in Table 4 show that the groups were similar in performance.

General Procedure

Procedures were approved by the institutional review board of the University of Florida. All testing was conducted over the course of 2 days in sessions lasting no more than 1 hr each. Adolescents were given 1-hr breaks between test sessions. Four to six adolescents came to the laboratory on test days, and were tested with alternating schedules, such that when half were being tested the other half were taking a break.

Analysis #1: Examining Stability

The first analysis performed with data from these adolescents was intended to assess the stability of speech-in-noise recognition across the period of middle childhood. This was considered an important goal to address, because if the ability to recognize speech in competing backgrounds changes rapidly over the course of development, questions could be cast on the meaningfulness of outcomes for any single test age. In particular, it could be that adolescents with CIs are simply developmentally behind their peers with NH in acquisition of these skills, perhaps due to differences in what has been termed *hearing age*. That situation would suggest there is little to do in the way of intervention, because there would be a reasonable expectation that children with CIs would eventually "catch up" in their abilities to recognize speech in competing backgrounds. On the other hand, discovering that the problems adolescents with CIs face in recognizing speech in competing backgrounds is a chronic problem would suggest that appropriate interventions should be developed.

To address this question of stability in how well children and adolescents recognize speech in competing backgrounds, performance for the same stimuli across a 4-year span was examined. This longitudinal analysis was possible because most of the children tested at eighth grade had been tested previously; specifically, at fourth and sixth grades. Data collection at the younger ages was done under a protocol approved by the institutional review board of The Ohio State University.

Participants Across Ages

Forty-three of the adolescents with NH and 42 of the adolescents with CIs had been tested at fourth grade on the four- and five-word stimuli in speech-shaped noise. Only 19 of these children with NH, however, were tested at the same signal-to-noise ratio of −3 dB used at eighth grade; the other 24 had been tested at 0 dB, the signal-to-noise level used for all 42 children with CIs. At sixth grade, 29 of the adolescents with NH and 35 of the adolescents with CIs participated, and each group was tested at the signal-to-noise ratio used with their group at eighth grade: for the four- and five-word sentences, this was −3 dB for the adolescents with NH and 0 dB for the adolescents with CIs. The lower numbers at sixth grade were due to a lapse in funding for 1 year, during which time some subjects in the longitudinal study were unable to be tested.

Equipment

The same equipment as that used in the Preliminary Study was used to collect responses to the sentence materials in noise or in babble for these adolescents.

Sentence Materials

For this across-ages analysis, the four-word sentences with correct syntax, but anomalous semantics were used along with the five-word HINT sentences. The same sets of sentences were presented at both sixth and eighth grade. At fourth grade, there was some variation in the sentences used, but all sentences fit the criteria that define these two sets of stimuli. The reason for the variation was that both the four- and five-word sentences were presented in each of two conditions at fourth grade: in speech-shaped noise and as sinewave analogs. A complete list of all possible sentences used at fourth grade is available in Nittrouer, Kuess, and Lowenstein (2015). A random selection of sentences to use at sixth grade was made from those larger lists.

Procedure

Presentation of these sentences followed exactly the same procedures as those described in the Preliminary

Study section. Adolescents with NH were presented with these sentence materials at the −3 dB signal-to-noise ratio at all three test ages; adolescents with CIs were presented with these sentences at the 0 dB signal-to-noise ratio at all three test ages. Reliability was assessed as in the Preliminary Study, except that double scoring was only done on 20% of the data in this case (every fifth subject).

Results

Interrater reliability for scores of these adolescents at grades other than eighth grade has been reported elsewhere (e.g., Nittrouer, Kuess, & Lowenstein, 2015). For eighth grade, mean reliability for the four-word sentences ranged from 90 to 100, with a mean of 97.8 (*SD* = 2.4), and mean reliability for the five-word sentences ranged from 91 to 100, with a mean of 95.7 (*SD* = 2.4). This level of reliability was considered acceptable, and scores from the first scorer were used in further analysis.

Figure 2 shows Tukey box and whisker plots for percent correct word recognition of the four-word sentences by the adolescents in this study across the three ages tested: fourth, sixth, and eighth grade. Figure 3 shows similar Tukey box and whisker plots for percent correct word recognition of the five-word sentences. Paired-samples *t* tests were performed on scores for all possible comparisons across grades (fourth vs. sixth, fourth vs. eighth, and sixth vs. eighth) for both the four- and five-word sentences. None of these tests revealed significant results, indicating that these children performed essentially the same across these grades.

Figure 2. Tukey box and whisker plots for percent of words recognized correctly for children across three test ages (fourth, sixth, and eighth grades) for four-word sentences presented in speech-shaped noise. These plots display the median as the center line, the interquartile range as the box, and 1.5 × interquartile range as the whiskers. Scores outside of this range are displayed as individual points. NH = normal hearing; CI = cochlear implant.

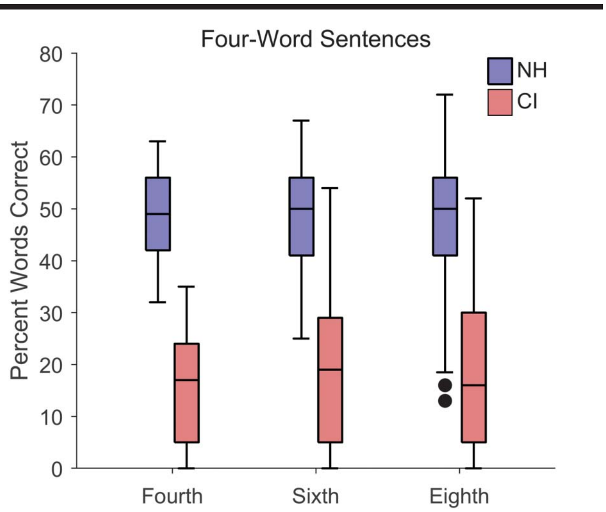
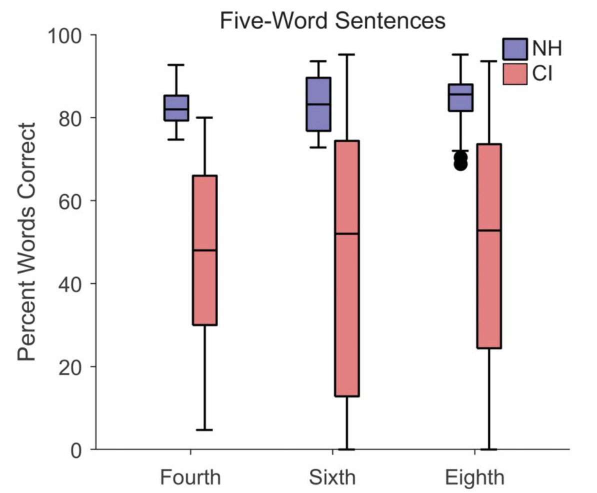


Figure 3. Tukey box and whisker plots for percent of words recognized correctly for children across three test ages (fourth, sixth, and eighth grades) for five-word sentences presented in speech-shaped noise. These plots display the median as the center line, the interquartile range as the box, and 1.5 × interquartile range as the whiskers. Scores outside of this range are displayed as individual points. NH = normal hearing; CI = cochlear implant.



Discussion

Overall, it was concluded that speech recognition in challenging backgrounds remains stable across a substantial period during middle childhood, for children with NH as well as those with hearing loss who use CIs. There was no significant change in recognition probabilities for either group of adolescents across this 4-year window. This finding suggests that the problems encountered by children and adolescents with CIs when it comes to recognizing speech in challenging backgrounds are chronic, and not just a case of delayed auditory development. Of particular interest to this study is the notion that whatever strategies these young people bring to the chore of recognizing speech under these difficult conditions remain constant across most of the years they spend in school, at least for upper elementary and middle school.

Analysis #2: Speech-in-Babble Performance

In this analysis, speech recognition was examined for the complex sentences in babble. The equipment, stimuli, and procedures were the same as those described in the Preliminary Study section.

Participants

The data of seven subjects had to be excluded from this analysis for the following reason. At the start of testing, adolescents with CIs were to be tested at a 0 dB signal-to-babble ratio, because in the earlier work using speech-shaped noise and sentences with simple syntax, they had been tested at a 0 dB signal-to-noise ratio. However, the first six adolescents with CIs who participated in this testing performed extremely poorly at this signal-to-babble ratio for the male and six-talker babble conditions, with means of just 12% correct word recognition in both conditions. These scores would not support an analysis of the mechanisms underlying speech recognition in babble. Consequently, the signal-to-babble ratio was changed to +3 dB for adolescents with CIs, and the data for these first six subjects was excluded from analysis. One adolescent with NH was mistakenly tested at the same signal-to-babble ratio as the adolescents with CIs. Thus, for this analysis, data were available from 55 adolescents with NH, all tested at a −3 dB signal-to-babble ratio, and from 46 adolescents with CIs, all tested at a +3 dB signal-to-babble ratio.

Results

When these sentences with complex syntactic structures were presented in quiet, recognition accuracy (and standard deviations) for the adolescents with NH was 96.7% (3.9%) for the SO structure, 97.6% (2.3%) for the OTH structure, and 97.4% (2.8%) for the CC structure. Accuracy of recognition in quiet for the adolescents with CIs was 87.8% (13.9%) for the SO structure, 89.5% (13.7%) for the OTH structure, and 90.4% (12.1%) for the CC structure. When paired-samples *t* tests were performed using arcsine-transformed scores to compare recognition across syntactic structures, none of the comparisons was significant for the adolescents with NH, even though it appears that recognition was slightly poorer for the SO structure. For the adolescents with CIs, recognition for the SO structure was significantly poorer than for the other two types of syntactic constructions: for OTH, $t(51) = -3.24$, $p = .002$, and for CC, $t(51) = -3.89$, $p < .001$.

Twenty percent of the responses for these complex sentences presented in background babble from each listener group were scored by two experimenters. Mean inter-rater reliability across the three types of babble and two listener groups ranged from 93 to 98.3, with a mean of 95.6 (SD = 1.8). This was considered acceptable reliability.

Figures 4 and 5 show mean percent correct word scores and standard errors for each type of syntactic structure in each kind of babble, for adolescents with NH and CIs, respectively. Mean scores and standard errors are also shown for the four- and five-word sentences. The *y*-axes differ for adolescents with NH and those with CIs, but both represent the same range of 60 percentage points.

Some differences in response patterns can be observed when scores from the adolescents with NH are compared with those from the adults with NH. Although scores for the single female babble condition are similar across the two groups of listeners, scores for the single male and six-

Figure 4. Mean percent words recognized correctly across syntactic structures by adolescents with normal hearing (with standard error bars). The CC, SO, and OTH sentences were presented in background babble. The four- and five-word sentences were presented in speech-shaped noise. Speech-in-noise and speech-in-babble ratios were −3 dB. CC = conjoined-clauses; SO = subject-object; OTH = other constructions; 4-word = four-word sentences; 5-word = five-word sentences.

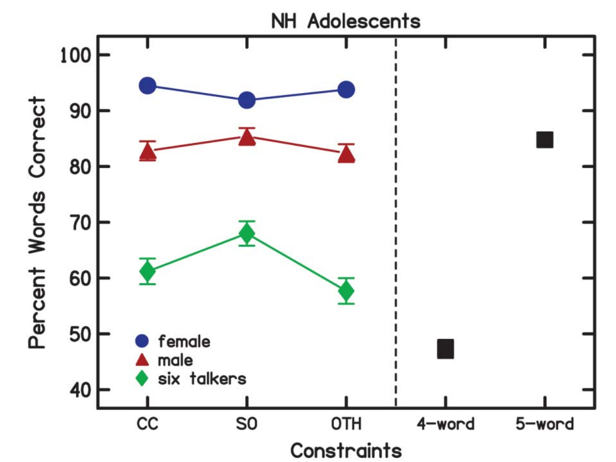


Figure 5. Mean percent words recognized correctly across syntactic structures by adolescents with cochlear implants (with standard error bars). The CC, SO, and OTH sentences were presented in background babble. The 4- and 5-word sentences were presented in speech-shaped noise. The speech-in-noise ratio was 0 dB and the speech-in-babble ratio was +3 dB. CC = conjoined-clauses; SO = subject-object; OTH = other constructions; 4-word = four-word sentences; 5-word = five-word sentences.

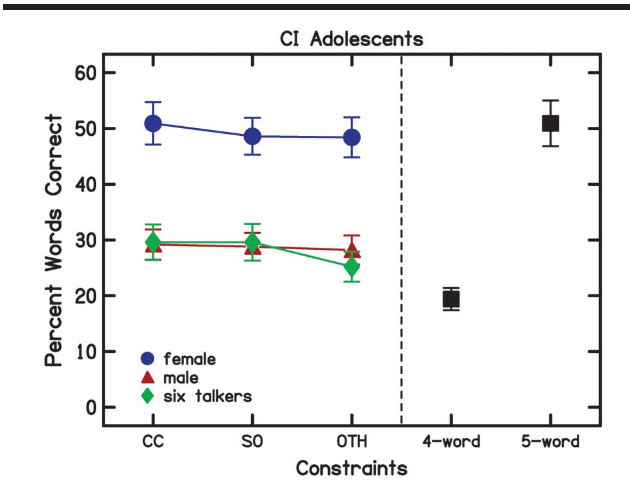


Table 5. Results of analyses of variance performed, using arcsine-transformed scores for adolescents with normal hearing or with cochlear implants separately.

Effect	*df*	*F*	*p*	η^2
Normal hearing				
Babble	2.108	201.08	< .001	.788
Syntactic structure	2.108	8.12	< .001	.131
Babble × Structure	4.216	8.92	< .001	.142
Cochlear implants				
Babble	2.90	105.03	< .001	.700
Syntactic structure	2.90	2.65	.076	—
Babble × Structure	4.180	1.06	.377	—

talker conditions are lower for the adolescents, compared with the adults. In addition, the effect of syntactic structure appears to be stronger for these adolescents than it was for the adults. Nonetheless, overall response patterns are reasonably similar across these listener groups.

Outcomes differ more substantially for the adolescents with CIs, beyond the simple difference in overall recognition probabilities. For one thing, the adolescents with CIs show no advantage for the babble condition of a single male talker over the six-talker condition. In fact, the advantage realized for the single female babble condition over the six-talker babble condition is diminished for these adolescents with CIs, compared with what is seen for the adolescents with NH: roughly 20 percentage points rather than 30. In addition, the adolescents with CIs show little, if any, effects of syntactic structure.

An ANOVA was performed on arcsine-transformed scores for the adolescents with NH and for those with CIs separately. Results are shown in Table 5. For the adolescents with NH, both main effects, as well as the interaction were found to be significant; for adolescents with CIs, only the main effect of babble condition was significant. To investigate these effects more closely, paired-samples *t* tests were performed, using the transformed scores. First, means were compared for the three babble conditions, for adolescents with NH and those with CIs separately. Results are displayed in Table 6 and show that adolescents with NH had significantly different scores across the three conditions. Recognition was highest for the female babble condition, followed by the male babble condition, and lastly, the six-talker condition. This pattern is the same as what was observed for the adults in the Preliminary Study. For the adolescents with CIs, recognition scores for the female babble condition differed from those for the other two conditions, but those two conditions showed similar recognition probabilities.

Finally, paired-samples *t* tests were performed on the transformed scores to examine the effects of syntactic structure, but only for the adolescents with NH. Adolescents in this group showed significant effects of syntactic structure, as well as a significant Babble × Syntactic-Structure interaction. Table 7 displays these outcomes. One significant difference was observed for each of the female and male babble conditions, both involving the SO construction. Two significant differences were observed for the six-talker babble condition. Recognition scores were found to be significantly higher for the SO construction than for either the CC or the OTH constructions. This finding matches outcomes from adults in the Preliminary Study, except that the effects, as indexed by Cohen's *d*s, were stronger.

Table 6. Outcomes of paired-samples *t* tests across babble conditions, using arcsine-transformed scores for adolescents at eighth grade.

Contrast	*df*	*t*	*p*	*d*
Normal hearing				
Female vs. male babble	54	8.80	< .001	1.36
Female vs. six-talker babble	54	19.80	< .001	3.16
Male vs. six-talker babble	54	10.85	< .001	1.66
Cochlear implants				
Female vs. male babble	45	11.13	< .001	0.99
Female vs. six-talker babble	45	14.70	< .001	0.98
Male vs. six-talker babble	45	1.05	.301	—

Note. All *p* values are for two-tailed tests. All three comparisons remain significant for the adolescents with normal hearing when a Bonferroni correction is applied; the two comparisons involving the female babble remain significant for the adolescents with cochlear implants when a Bonferroni correction is applied.

Discussion

This second analysis was conducted primarily to examine how adolescents with CIs perform with sentences possessing complex linguistic structures, embedded in speech babble. From the outset, it was apparent this sort of background is more challenging for these adolescents, because the first few

Table 7. Outcomes of paired-samples *t* tests across syntactic structures in the six-talker babble condition, using arcsine-transformed scores from adolescents with normal hearing.

Contrast	*df*	*t*	*p*	*d*
Female babble				
Subject–object vs. conjoined	54	2.79	.007	−0.38
Subject–object vs. other	54	1.66	.102	—
Conjoined-clauses vs. other	54	0.69	.494	—
Male babble				
Subject–object vs. conjoined	54	2.00	.051	—
Subject–object vs. other	54	2.51	.015	0.27
Conjoined-clauses vs. other	54	0.43	.669	—
Six-talker babble				
Subject–object vs. conjoined	54	3.76	< .001	0.43
Subject–object vs. other	54	6.32	< .001	0.59
Conjoined-clauses vs. other	54	1.93	.058	—

Note. All *p* values are for two-tailed tests. All four significant comparisons remain significant when a Bonferroni correction is applied.

adolescents tested were unable to recognize any substantial portion of words at the signal-to-babble ratio at which they have been tested when speech-shaped noise was used as the competing background. Furthermore, these adolescents with CIs showed no advantage for the single-talker babble condition over the six-talker babble, when the single talker was the same gender as the talker producing the target signal. It is tempting to attribute this outcome entirely to poor spectral resolution, which is presumed to be poorer for listeners with CIs; the third analysis addressed this assumption. Finally, the adolescents with CIs showed no effects of the different syntactic constructions used to generate the sentences for this experiment, a finding that could indicate a lack of access to the fundamental frequency (arising from poor spectral resolution) or poor syntactic knowledge. Both these possibilities were also explored in the next analysis.

Outcomes for the adolescents with NH served as a benchmark against which to compare the performance of the adolescents with CIs. The performance of these adolescents with NH was similar to that of the adults in the Preliminary Study, although scores were not quite as high. Nonetheless, these adolescents showed strong effects of the babble conditions, with scores for the single-male babble condition better than those for the six-talker condition. Moderate effects of syntactic structure were observed, and these effects were strongest when overall recognition probabilities were at lower values overall (i.e., the six-talker babble condition). This finding replicates what Boothroyd and Nittrouer (1988) reported, that top-down linguistic constraints had stronger effects when recognition probabilities were lower. In that experiment, however, those effects could clearly be attributed to the syntactic (and semantic) structure of the sentences. In the current experiment, at least some of the advantage observed for the SO construction might be attributable to a more distinct intonation pattern. The next analysis could help sort that out.

Analysis #3: Underlying Mechanisms

In this third analysis, the mechanisms likely to underlie speech recognition in competing backgrounds were explored, with an emphasis on the explanatory mechanisms for adolescents with CIs. In line with factors that have previously been suggested as underlying speech recognition under such conditions, language-based measures (vocabulary, syntactic knowledge, and phonological sensitivity) and verbal working memory using a task more complex than digit span were included in these analyses. These four measures represent what is termed *top-down* contributions to speech recognition in challenging environments. In addition, two *bottom-up* factors were included as predictor variables. First, spectral resolution was examined by obtaining thresholds for spectral modulation depth detection. Second, a measure of perceptual organization was included as a predictor variable. Specifically, we incorporated measures of sinewave speech recognition obtained from these adolescents at a younger age.

Participants

In the first part of this analysis, scores for the five predictor variables were compared across the groups of adolescents with NH and with CIs, to get a sense of the magnitude of differences between groups. For this procedure, scores from all 56 adolescents with NH and all 52 adolescents with CIs were included to provide as representative samples as possible. Next, the relationships of these predictor variables to recognition of the complex sentences in background babble were examined for the 55 adolescents with NH and 46 adolescents with CIs whose data were included in the second analysis, described above.

Equipment

The same equipment was used for presentation of audio materials as already described, and responses for all predictor measures were audio-video recorded, as already described, with the exception of the verbal working memory and spectral modulation tasks. For those tasks, responses were recorded by the MATLAB routines controlling the experiment.

Predictor Variables

Six predictor variables were used in this third analysis. These measures are listed here, and more detail regarding each one can be found in Supplemental Material S1.

Vocabulary. An expressive vocabulary measure was used: the Expressive One-Word Picture Vocabulary Test (Martin & Brownell, 2011). This is a standardized test instrument, so standard scores were used in analysis.

Syntactic comprehension. The Sentence Comprehension of Syntax from the Comprehensive Assessment of Spoken Language was used (Carrow-Woolfolk, 1999). This is a standardized test instrument, so standard scores were used in analysis. Stimuli for this measure were video-recorded by a laboratory staff member and presented on a computer monitor to all adolescents. This procedure ensured that all subjects had precisely the same presentation of stimulus materials: They were presented with the same recorded instructions and all test sentences were presented with the same inflectional patterns.

Phonological sensitivity. The Final Consonant Choice task was used. This task has been used extensively in this laboratory, and has been found to have high sensitivity (Nittrouer et al., 2012). The task assesses subjects' abilities to recognize similarity in phonological structure at the ends of words. The dependent measure is the percentage of items correctly answered.

Verbal working memory. An immediate serial recall task was used to assess verbal working memory, because it is more complex than a digit span task. This task has been used extensively in this laboratory, and methods have been shown to be reliable (Nittrouer & Miller, 1999). The dependent measure is the percentage of words recalled in the correct serial position across 10 lists.

Spectral resolution. Spectral modulation depth detection thresholds were obtained as a way of assessing spectral resolution. The modulation rate of these stimuli was held constant at 0.5 cycles per octave. This modulation rate was selected because modulation at this rate is roughly the rate at which formants are represented in speech. An adaptive procedure was used to estimate the 70.7% threshold on the psychometric function, and this threshold served as the dependent measure (Nittrouer et al., 2021).

Perceptual organization. It is well recognized that a central challenge for listeners with CIs involves organizing the extremely degraded signal they receive in order to recover a linguistically meaningful object. This process is made all the more difficult when the speech signal is embedded in a background of noise or speech babble, because the listener must first recover as much of the target signal as possible. To disentangle the perceptual organization abilities of these subjects from their abilities to segregate the target signal from the background noise, they were presented with sinewave speech signals during testing in fourth grade. Percent correct word recognition for these sinewave sentences were used as the measure of perceptual organization abilities for these adolescents. No differences were observed at fourth grade between children with NH and those with CIs (Nittrouer, Kuess, & Lowenstein, 2015).

Results

Overall performance. Performance on all six of the predictor variables was compared between adolescents with NH and those with CIs. Table 8 displays means, medians, and SDs for these measures, for both groups. Table 9 displays the results of independent-samples *t* tests conducted to compare scores between the two groups. For all scores that represent percentages, arcsine-transformed scores were computed and used. These analyses provide a sense of how these adolescents with CIs were faring in general when it came to language abilities.

Results shown on Table 9 reveal that for four of the six predictor variables, the adolescents with CIs performed, on average, a little more than 0.5 *SD* below the mean of adolescents with NH. The two exceptions to this pattern were phonological sensitivity, where the adolescents with CIs performed more than 1.0 *SD* below the mean of the adolescents with NH, and perceptual organization, where the adolescents with CIs performed on par with the adolescents with NH. This last finding is important, because it indicates that these adolescents with CIs were, on average, just as competent as the adolescents with NH when it came to perceptually organizing available signal components. This fact makes perceptual organization a critical variable to include in the analyses of mechanisms underlying speech recognition in challenging conditions for listeners with CIs. Although the adolescents with CIs possess the same proficiency with perceptual organization of degraded signals as

Table 9. Outcomes of independent-samples *t* tests performed on predictor variables, comparing scores for adolescents with normal hearing (56) and adolescents with cochlear implants (52).

Predictor variable	*df*	*t*	*p*	*d*
Vocabulary	106	3.54	< .001	0.68
Syntactic comprehension	106	2.72	.008	0.52
Phonological sensitivity	106	6.57	< .001	1.25
Verbal working memory	106	3.62	< .001	0.71
Spectral resolution	106	–2.77	.007	–0.53
Perceptual organization	106	0.354	.724	—

Note. All *p* values are for two-tailed tests. Arcsine-transformed values were used, when appropriate for percentage scores. Cohen's *d*s represent effect sizes for differences between scores of adolescents with NH and adolescents with CIs.

Table 8. Means, medians, and standard deviations for predictor variables at eighth grade for adolescents with normal hearing (56) and adolescents with cochlear implants (52).

	NH			CI		
Measure	*M*	*Mdn*	*SD*	*M*	*Mdn*	*SD*
Vocabulary (SS)	113	114	16	101	101	18
Syntactic comprehension (SS)	107	106	13	99	101	13
Phonological sensitivity (%)	89.6	91.7	6.7	72.6	75.0	18.3
Verbal working memory (%)	71.5	70.8	17.1	59.1	58.3	17.9
Spectral resolution (dB)	7.5	7.2	2.8	9.8	8.8	5.3
Perceptual organization (%)	31.2	32.0	17.1	28.7	30.5	12.8

Note. NH = normal hearing; CI = cochlear implant; SS = standard score (*M* = 100; *SD* = 15); % = percent correct; dB = threshold in dB.

adolescents with NH, this skill might explain significantly more variability in speech-in-babble recognition for the adolescents with CIs, because they recover less of the signal detail from the babble background; that is, they have a more degraded signal available to them.

Audiometric measures. Before the contributions of the predictor variables to speech recognition were examined, the contributions of the audiometric measures were assessed for the adolescents with CIs. These measures included (1) age of first CI; (2) pre-implant, better-ear 3-frequency pure-tone average thresholds; and (3) aided 4-frequency thresholds. Pearson product–moment correlation coefficients (hereafter, simply correlation coefficients) were computed for each of these three factors and arcsine-transformed scores for each of the three babble conditions. The only one of these audiometric measures that was close to demonstrating a significant relationship was the aided threshold. The correlation coefficients for this measure were −.241 (p = .086) for six-talker babble, −.291 (p = .036) for the female babble, and −.304 (p = .029) for the male babble. These were deemed to be weak relationships, so were not considered further.

Relationships between speech-in-babble recognition and predictor variables. For this analysis, the mean recognition score across the three babble conditions was computed for each adolescent and used as the dependent variable. Correlation coefficients between each predictor variable and these recognition scores were computed. Scores from the measure of nonverbal IQ were also included in these correlational analyses, because the nonverbal assessment tool used here assesses visual attention, which has been found to account for speech recognition in challenging conditions (McCreery et al., 2020; Moberly & Reed, 2019). These correlation coefficients are shown in Table 10 for adolescents with NH and adolescents with CIs separately. Correlation coefficients between each predictor variable and speech recognition scores for each babble condition individually can be found in Supplemental Material S2. There it can be seen that the patterns observed for mean scores across babble conditions are observed for the individual conditions as well. Supplemental Material S3 displays the correlation coefficients among all of the predictor variables for each group separately.

Looking at these relationships first for adolescents with NH, it appears that the primary mechanisms underlying speech-in-babble scores for this group involve top-down factors: nonverbal IQ, vocabulary, syntactic comprehension, phonological sensitivity, and verbal working memory. These variables, however, may be overlapping in the variability they explain. Consequently, a stepwise multiple regression was conducted using all these variables as predictors and mean speech recognition scores across the three babble conditions as the dependent measure. Table 11 shows the outcomes of this analysis, revealing that only verbal working memory and syntactic comprehension explained significant amounts of unique variance. Overall, these results indicate that for these adolescents with NH how well they performed in their (generally good) abilities to recognize speech in background babble could be explained by top-down effects. This finding matches predictions from most models of speech recognition in challenging environments.

Looking next at these relationships for the adolescents with CIs, a very different pattern emerges. Although some variability in scores can be explained by phonological sensitivity and spectral resolution, it is clear that the largest portion is explained by how well these adolescents were able to perceptually organize degraded signal structure. Indeed, when a stepwise multiple regression was performed with these three predictor variables, only perceptual organization was found to explain a significant amount of unique variability.

Table 10. Pearson product–moment correlation coefficients between mean recognition scores across the three babble conditions and each of the six predictor variables for adolescents with normal hearing (55) and adolescents with cochlear implants (46).

	NH		CI	
Predictor variable	***r***	***p***	***r***	***p***
Nonverbal IQ	**.336**	**.011**	−.051	.719
Vocabulary	**.318**	**.018**	.167	.268
Syntactic comprehension	**.455**	**< .001**	.116	.444
Phonological sensitivity	**.381**	**.004**	**.418**	**.004**
Verbal working memory	**.463**	**< .001**	.186	.217
Spectral resolution	−.024	.860	**−.385**	**.008**
Perceptual organization	.201	.203	**.666**	**< .001**

Note. Significant coefficients are bolded. NH = normal hearing; CI = cochlear implant.

Discussion

This third analysis was undertaken to examine mechanisms underlying the abilities of adolescents—primarily those with hearing loss who use CIs—to recognize speech in challenging backgrounds. Listeners with hearing loss encounter difficulty recognizing speech in the presence of

Table 11. Stepwise regression outcomes for mean speech recognition scores across the three babble conditions for adolescents with normal hearing. Predictor variables were syntactic comprehension, phonological sensitivity, and verbal working memory.

Step no.	Significant predictor variables	Standardized coefficients	R^2
1	Verbal working memory	β = .463, p < .001	.215
2	Verbal working memory	β = .316, p = .023	.283
	Syntactic comprehension	β = .300, p = .030	

noise or competing talkers more severe than what listeners with NH experience. Typically, variability in recognition scores across listeners with hearing loss under such circumstances is attributed to language or cognitive processes. In many studies reaching these conclusions, however, only measures of language and cognitive functions are obtained, so there is no way to compare the magnitude of these effects with that of other factors; in particular, factors associated with perceptual processes. Previous experiments have often been conducted with listeners with NH, and findings extrapolated to listeners with hearing loss. In the current analysis, care was taken to include two measures of what may be termed *bottom-up, perceptual factors,* and outcomes were obtained from listeners with NH and those with CIs so underlying mechanisms could be compared.

When it comes to these adolescents with CIs, results of this third analysis reveal that variability in speech recognition was most strongly explained by one bottom-up, perceptual factor: how well the adolescent was able to organize degraded signals in order to obtain a linguistically relevant object. One top-down factor, phonological sensitivity, explained a significant amount of variability, as did the other bottom-up factor examined, spectral resolution. When contributions of these predictor variables were examined in aggregate, however, only the ability to perceptually organize degraded signals was found to explain a significant amount of unique variability. This finding differed from the primary outcome for adolescents with NH, whose recognition scores were found to be explained primarily by language and cognitive factors of the nature described in earlier studies for children with NH (e.g., McCreery et al., 2020). Of course, overall recognition probabilities were quite different across these two groups, which likely helps to explain differences in underlying mechanisms for the adolescents with NH and those with CIs. Moreover, where adults with acquired hearing loss who use CIs are concerned, there is evidence that top-down processes are harnessed to a greater extent in speech recognition by patients with better spectro-temporal resolution, so when the bottom-up information is more readily available (Moberly et al., 2021). In the current study, the adolescents with NH had better spectral resolution than the adolescents with CIs, which fits with outcomes of Moberly et al.

General Discussion

The primary purpose of the study reported here was to examine the speech-in-babble recognition abilities of adolescents who have congenital hearing loss severe enough to require CIs, when listening to sentences with complex syntax. The motivation for this focus was that children encounter ever-increasing complexity in language as they progress through school, and listening environments become more challenging, as well. Yet, much of the research done on speech recognition in challenging environments for children with CIs use sentences with simple syntax, and few have looked at the effects of background babble, especially the effects of various kinds of babble. We wished to evaluate how students in higher grades are performing in these real-world conditions that they encounter in school. Three specific objectives were addressed. The first objective was to examine the stability of recognition in challenging environments across a several-year age range for adolescents with NH or CIs. The second objective was to examine the abilities of these adolescents to benefit from variation in background babble and syntactic structures to recognize speech under these circumstances. Finally, the third objective was to explore the mechanisms that underlie variability in these adolescents' abilities to recognize speech under these circumstances.

Before any of those objectives could be addressed, appropriate stimuli needed to be developed. That was accomplished in a preliminary study involving young adults with NH.

Regarding the first objective, results revealed that the speech-in-noise recognition of both adolescents with NH and adolescents with CIs remained stable across a 4-year span from late elementary school through middle school. For the adolescents with NH, this finding matches previous reports regarding thresholds for word recognition (Corbin et al., 2016) and word detection (Bonino et al., 2021) in speech-shaped noise; those thresholds reach mature levels in children with NH at 10 years of age and remain stable thereafter. The significance of the result found here is that there was no evidence that the adolescents with CIs were simply delayed in the development of abilities to recognize speech in challenging environments.

Regarding the second objective, the adolescents with CIs performed much more poorly than the adolescents with NH, in spite of the fact that they were listening at a signal-to-babble ratio that was 6 dB more favorable than that used with the adolescents with NH. There was no overlap in scores between the two groups, in spite of the difference in signal-to-babble ratios. Adolescents with CIs were most successful at segregating the target signal from background babble when the babble was derived from a single talker of the opposite gender than that of target signal: This was when the target signal was produced by a male talker and the babble was derived from a female talker. Unlike adolescents with NH, however, these adolescents with CIs obtained no benefit when babble derived from a single talker, if that talker was the same gender as the target signal. In this latter case, recognition scores were, on average, identical to those obtained with six-talker babble. It is tempting to attribute the poor speech-in-babble outcomes for these adolescents entirely to the poor spectral resolution that CI processing imposes on signals. If that were entirely the source of the problem, however, it would be expected that the adolescents with CIs who demonstrated spectral resolution in the range of adolescents with NH would have been capable of processing speech in background babble with similar skill to the adolescents with NH. But this was not what was found. Even the adolescents with CIs who showed good spectral resolution, which could be defined as better than the median score of the adolescents with NH, had poorer overall recognition—even at the more favorable signal-to-babble ratio—and showed no advantage for the single male talker babble. The extreme disparity in overall abilities between the two groups and differences in outcomes across the various babble conditions indicates there must be factors, other than those explored here, that explain the extremely poor speech recognition abilities of the adolescents with CIs when listening to speech in a background of babble.

These adolescents with CIs obtained no benefit from differences in syntactic structure across the sentence materials. Both the adults and the adolescents with NH showed advantages for the SO construction, although it remains unclear whether that was due to the more constraining nature of that construction or a distinct inflectional pattern. However, it is notable that this syntactic effect only appeared for recognition scores near the middle of the cumulative normal distribution for these listeners with NH, suggesting they had no need for the contributions of top-down syntactic structure when adequate bottom-up information was available. The adolescents with CIs showed no such advantage for the SO construction, even though most of their recognition scores were in that middle range. It is tempting to attribute this lack of effect to poorer syntactic knowledge on the part of the adolescents with CIs: although their scores on the Syntactic Comprehension task were within the normal range according to the normative sample, those scores were still significantly below the scores of the adolescents with NH. Nonetheless, the lack of any correlation between Syntactic Comprehension and speech recognition scores for the adolescents with CIs indicates that extent of syntactic knowledge did not account for any variability in their speech recognition scores. Thus, there is no reason to suspect that better comprehension scores would have resulted in better speech recognition scores for these adolescents with CIs.

Regarding the third objective, outcomes for the adolescents with NH matched what would be predicted by previous investigations: their abilities to recognize speech in babble were largely explained by the top-down factors of working memory and syntactic knowledge. Outcomes for the adolescents with CIs told a different story. The abilities of these listeners to recognize these complex sentences in a background of babble were largely explained by their skill at perceptually organizing a degraded signal to recover a linguistically relevant object. This finding suggests a source of effect that has not previously been identified, but which can be explored in future investigations.

Clinical Implications

The results of this study provide information regarding how intervention strategies should—and should not—be designed in order to help children and adolescents with hearing loss enhance their speech recognition abilities, not only in challenging environments but in quiet as well. Auditory training regimens have proliferated over the years, even as the value of auditory training in general has been questioned (Boothroyd, 2010; Henshaw & Ferguson, 2013). This skepticism arises largely because "Evidence of dramatic general improvements in speech perception by hearing-impaired listeners after training is still lacking…" (p. 281, J. D. Miller et al., 2015). Approaches have differed across auditory training programs, with some focused on the learning of discrete phonemic contrasts (e.g., Fu & Galvin, 2007; J. D. Miller et al., 2008; Woods et al., 2015) and others focused on improving recognition of sentence-length materials (e.g., Sweetow & Sabes, 2006). One finding of this current study was that the ability to organize sensory components recovered from degraded signals contributes to speech recognition in challenging conditions. This is likely a skill that can be modified with training, so future studies should examine the best methods for doing so.

Another implication of this study is that more appropriate diagnostic markers remain to be discovered to support mapping of CIs. Recently, efforts have focused on spectral or spectrotemporal structure as a possible diagnostic marker of how well a patient's implants are mapped (Gifford et al., 2018; Horn et al., 2017; Landsberger et al., 2019; Won et al., 2007). This study, however, suggests these adolescents with CIs did not demonstrate large deficits in spectral resolution. And although these scores predicted speech recognition in babble to some extent, it was not the strongest predictor. Consequently, other potential sources of the extreme deficits in speech recognition exhibited by these listeners need to be examined. It may not be so much signal quality that determines speech recognition abilities, but rather how well the listener is able to organize the available signal fragments that matters.

Limitations and Future Directions

A possible limitation of this study is that the scores for perceptual organization were obtained at a younger age than the speech-in-babble scores, although that seems unlikely. The abilities of these adolescents to recognize sentences in noise did not change over the course of 4 years. And although there were group differences in speech-in-noise and speech-in-babble abilities, recognition scores for the sinewave sentences were similar across listener groups at that younger age. There is no reason to suspect that the adolescents with NH would have improved in those abilities, but not the adolescents with CIs. Thus these perceptual organization abilities were likely stable. Nonetheless, a future study should investigate these abilities at the same time that speech-in-noise (or speech-in-babble) recognition is examined, using the kind of complex sentences used here.

Another possible limitation is that sentences in the OTH category consisted of three syntactic structures. This design was implemented because it was not possible to present a sufficient number of sentences with each individual kind of structure, and there was no reason a priori to suspect differences in recognition scores across these three kinds of structure, as there was for the other two kinds of structure examined (SO and CC). But future work should examine the abilities of listeners—especially children and adolescents—to utilize each of the syntactic structures in that OTH category in service to speech recognition, controlling for intonation. That more careful assessment would help sort out whether the adolescents with CIs failed to show effects of syntactic structure because they failed to use top-down processes with these complex constructions, or they could not recognize differences in intonation patterns. Nonetheless, this study was a first step in addressing these questions.

Attentional control was inferred from the nonverbal IQ scores, rather than being separately assessed. This is a limitation in this study because recent findings by others have revealed that a listener's ability to inhibit attention to irrelevant sensory input is strongly associated with speech recognition in challenging conditions (McCreery et al., 2020; Moberly & Reed, 2019). Although a major finding of the current study was that perceptual organization of recovered signal structure is essential to speech recognition in noise, the task of organizing recovered signal components is likely affected by the listener's ability to attend to those components while inhibiting attention to other components. Future investigations should explicitly examine the role of attentional control and its relationship to perceptual organization.

Finally, only one measure of spectral resolution was obtained: depth detection. That measure was obtained at just one modulation rate (0.5 cycles per octave), and no measure of temporal processing abilities was obtained. Future investigations should expand the search for potential peripheral processes as mechanisms underlying speech-in-babble recognition.

Summary

Children born with severe-to-profound hearing loss have demonstrated remarkable gains in their language abilities since the advent of CIs, with many of these children scoring within the normal range on standardized instruments. But these measures are collected under ideal circumstances, which differ greatly from the conditions under which most of their interactions occur, especially academic activities. The goal of this study was to examine how adolescents with CIs function in realistic learning environments, when there are other people talking in the background and sentences can be complex in structure. Results revealed that these circumstances are much more challenging for children with CIs, putting them at an enormous disadvantage compared with their peers with NH. The primary factor explaining variability across adolescents with CIs in their abilities to recognize speech in these circumstances was how well they could perceptually organize the degraded signal structure recovered from background babble. This information should inform the design of interventions aimed at mechanisms underlying the processing of complex language in real-world environments by school-age children and adolescents with CIs, to help them optimize their abilities. At the same time, efforts must continue to identify and address the source of the extreme deficit faced by these children and adolescents in their overall abilities to recognize speech in challenging backgrounds.

Data Availability Statement

Data are available upon request from the first author.

Acknowledgments

This work was supported by Grants R01 DC006237 and R01 DC015992 from the National Institute on Deafness and Other Communication Disorders, the National Institutes of Health, to Susan Nittrouer. We thank the College of Public Health and Health Professions at the University of Florida for start-up funding for the preliminary study. The authors especially thank Donal G. Sinex for design and development of the speech-in-babble sentences. In addition, we thank the many parents and children who were part of our EDCHL family for so many years, along with staff members Taylor LeCorgne, Popy Shell, Amanda Caldwell-Tarr, Jill Twersky, Caitlin Rice, Eric Tarr, and Jamie Kuess, as well as student assistants Meganne Muir, Danielle Diaz, Angela Duan, Alexis Tolentino, Joanna Lubin, Lindsay Arena, Chandler Burkhart, Olivia

Coker, Shay Fitzgerald, Karen Worman, and Heather Smythe. Finally, we thank our faculty collaborator Aaron Moberly for lending his voice to this project.

References

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology, 47*(Suppl. 2), S53–S71. https://doi.org/10.1080/14992020802301142

Alcántara, J. I., Weisblatt, E. J., Moore, B. C., & Bolton, P. F. (2004). Speech-in-noise perception in high-functioning individuals with autism or Asperger's syndrome. *The Journal of Child Psychology and Psychiatry, 45*(6), 1107–1114. https://doi.org/10.1111/j.1469-7610.2004.t01-1-00303.x

Bacon, S. P., & Viemeister, N. F. (1985). Temporal modulation transfer functions in normal-hearing and hearing-impaired listeners. *Audiology, 24*(2), 117–134. https://doi.org/10.1121/1.2126933

Bench, J., Kowal, A., & Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *British Journal of Audiology, 13*(3), 108–112. https://doi.org/10.3109/03005367909078884

Ben-David, B. M., Gal-Rosenblum, S., van Lieshout, P. H. H. M., & Shakuf, V. (2019). Age-related differences in the perception of emotion in spoken language: The relative roles of prosody and semantics. *Journal of Speech, Language, and Hearing Research, 62*(4S), 1188–1202. https://doi.org/10.1044/2018_JSLHR-H-ASCC7-18-0166

Bernstein, J. G., & Grant, K. W. (2009). Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America, 125*(5), 3358–3372. https://doi.org/10.1121/1.3110132

Bolia, R. S., Nelson, W. T., Ericson, M. A., & Simpson, B. D. (2000). A speech corpus for multitalker communications research. *The Journal of the Acoustical Society of America, 107*(2), 1065–1066. https://doi.org/10.1121/1.428288

Bonino, A. Y., Ramsey, M. E., Pancoast, E. S., & Vance, E. A. (2021). Development of masked speech detection thresholds in 2- to 15-year-old children: Speech-shaped noise and two-talker speech maskers. *Ear and Hearing, 42*(6), 1712–1726. https://doi.org/10.1097/AUD.0000000000001062

Boons, T., De Raeve, L., Langereis, M., Peeraer, L., Wouters, J., & van Wieringen, A. (2013). Expressive vocabulary, morphology, syntax and narrative skills in profoundly deaf children after early cochlear implantation. *Research in Developmental Disabilities, 34*(6), 2008–2022. https://doi.org/10.1016/j.ridd.2013.03.003

Boothroyd, A. (2010). Adapting to changed hearing: The potential role of formal training. *Journal of the American Academy of Audiology, 21*(9), 601–611. https://doi.org/10.3766/jaaa.21.9.6

Boothroyd, A., & Nittrouer, S. (1988). Mathematical treatment of context effects in phoneme and word recognition. *The Journal of the Acoustical Society of America, 84*(1), 101–114. https://doi.org/10.1121/1.396976

Bradham, T. S., Fonnesbeck, C., Toll, A., & Hecht, B. F. (2018). The listening and spoken language data repository: Design and project overview. *Language, Speech, and Hearing Services in Schools, 49*(1), 108–120. https://doi.org/10.1044/2017_LSHSS-16-0087

Bregman, A. S. (1990). *Auditory scene analysis*. MIT Press. https://doi.org/10.7551/mitpress/1486.001.0001

Buss, E., Lorenzi, C., Cabrera, L., Leibold, L. J., & Grose, J. H. (2019). Amplitude modulation detection and modulation masking in school-age children and adults. *The Journal of the Acoustical Society of America, 145*(4), 2565–2575. https://doi.org/10.1121/1.5098950

Carroll, R., & Ruigendijk, E. (2013). The effects of syntactic complexity on processing sentences in noise. *Journal of Psycholinguistic Research, 42*(2), 139–159. https://doi.org/10.1007/s10936-012-9213-7

Carrow-Woolfolk, E. (1999). *Comprehensive Assessment of Spoken Language (CASL)*. Pearson Assessments.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America, 25*(5), 975–979. https://doi.org/10.1121/1.1907229

Corbin, N. E., Bonino, A. Y., Buss, E., & Leibold, L. J. (2016). Development of open-set word recognition in children: Speech-shaped noise and two-talker speech maskers. *Ear and Hearing, 37*(1), 55–63. https://doi.org/10.1097/aud.0000000000000201

Darwin, C. J. (2008). Listening to speech in the presence of other sounds. *Philosophical Transactions of the Royal Society B: Biological Science, 363*(1493), 1011–1021. https://doi.org/10.1098/rstb.2007.2156

Darwin, C. J., Brungart, D. S., & Simpson, B. D. (2003). Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers. *The Journal of the Acoustical Society of America, 114*(5), 2913–2922. https://doi.org/10.1121/1.1616924

Dubno, J. R., Dirks, D. D., & Morgan, D. E. (1984). Effects of age and mild hearing loss on speech recognition in noise. *The Journal of the Acoustical Society of America, 76*(1), 87–96. https://doi.org/10.1121/1.391011

Durlach, N. I., Mason, C. R., Shinn-Cunningham, B. G., Arbogast, T. L., Colburn, H. S., & Kidd, G., Jr. (2003). Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity. *The Journal of the Acoustical Society of America, 114*(1), 368–379. https://doi.org/10.1121/1.1577562

Eisenberg, L. S., Shannon, R. V., Schaefer Martinez, A., Wygonski, J., & Boothroyd, A. (2000). Speech recognition with reduced spectral cues as a function of age. *The Journal of the Acoustical Society of America, 107*(5), 2704–2710. https://doi.org/10.1121/1.428656

Fu, Q. J., & Galvin, J. J., III (2007). Perceptual learning and auditory training in cochlear implant recipients. *Trends in Amplification, 11*(3), 193–205. https://10.1177/1084713807301379

Geers, A. E., Nicholas, J., Tobey, E., & Davidson, L. (2016). Persistent language delay versus late language emergence in children with early cochlear implantation. *Journal of Speech, Language, and Hearing Research, 59*(1), 155–170. https://doi.org/10.1044/2015_JSLHR-H-14-0173

Gifford, R. H., Noble, J. H., Camarata, S. M., Sunderhaus, L. W., Dwyer, R. T., Dawant, B. M., Dietrich, M. S., & Labadie, R. F. (2018). The relationship between spectral modulation detection and speech recognition: Adult versus pediatric cochlear implant recipients. *Trends in Hearing, 22*. https://doi.org/10.1177/2331216518771176

Gordon-Salant, S., & Cole, S. S. (2016). Effects of age and working memory capacity on speech recognition performance in noise among listeners with normal hearing. *Ear and Hearing, 37*(5), 593–602. https://doi.org/10.1097/AUD.0000000000000316

Hayes, H., Geers, A. E., Treiman, R., & Moog, J. S. (2009). Receptive vocabulary development in deaf children with cochlear implants: Achievement in an intensive auditory-oral educational setting. *Ear and Hearing, 30*(1), 128–135. https://doi.org/10.1097/aud.0b013e3181926524

Henry, B. A., & Turner, C. W. (2003). The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners. *The Journal of the Acoustical Society of America, 113*(5), 2861–2873. https://doi.org/10.1121/1.1561900

Henshaw, H., & Ferguson, M. A. (2013). Efficacy of individual computer-based auditory training for people with hearing loss: A systematic review of the evidence. *PLOS ONE, 8*(5), Article e62836. https://doi.org/10.1371/journal.pone.0062836

Holmes, E., Folkeard, P., Johnsrude, I. S., & Scollie, S. (2018). Semantic context improves speech intelligibility and reduces listening effort for listeners with hearing impairment. *International Journal of Audiology, 57*(7), 483–492. https://doi.org/10.1080/14992027.2018.1432901

Hong, R. S., & Turner, C. W. (2006). Pure-tone auditory stream segregation and speech perception in noise in cochlear implant recipients. *The Journal of the Acoustical Society of America, 120*(1), 360–374. https://doi.org/10.1121/1.2204450

Horn, D. L., Dudley, D. J., Dedhia, K., Nie, K., Drennan, W. R., Won, J. H., Rubinstein, J. T., & Werner, L. A. (2017). Effects of age and hearing mechanism on spectral resolution in normal hearing and cochlear-implanted listeners. *The Journal of the Acoustical Society of America, 141*(1), 613–623. https://doi.org/10.1121/1.4974203

Howard-Jones, P. A., & Rosen, S. (1993). Uncomodulated glimpsing in "checkerboard" noise. *The Journal of the Acoustical Society of America, 93*(5), 2915–2922. https://doi.org/10.1121/1.405811

Humes, L. E. (2007). The contributions of audibility and cognitive factors to the benefit provided by amplified speech to older adults. *Journal of the American Academy of Audiology, 18*(7), 590–603. https://doi.org/10.3766/jaaa.18.7.6

Humes, L. E., Kidd, G. R., & Fogerty, D. (2017). Exploring use of the coordinate response measure in a multitalker babble paradigm. *Journal of Speech, Language, and Hearing Research, 60*(3), 741–754. https://doi.org/10.1044/2016_JSLHR-H-16-0042

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122–149. https://doi.org/10.1037/0033-295X.99.1.122

Kalikow, D. N., Stevens, K. N., & Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *The Journal of the Acoustical Society of America, 61*(5), 1337–1351. https://doi.org/10.1121/1.381436

Kirby, B. J., Browning, J. M., Brennan, M. A., Spratford, M., & McCreery, R. W. (2015). Spectro-temporal modulation detection in children. *The Journal of the Acoustical Society of America, 138*(5), EL465–EL468. https://doi.org/10.1121/1.4935081

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44*(4), 978–990. https://doi.org/10.3758/s13428-012-0210-4

Lad, M., Holmes, E., Chu, A., & Griffiths, T. D. (2020). Speech-in-noise detection is related to auditory working memory precision for frequency. *Scientific Reports, 10*(1), 13997. https://doi.org/10.1038/s41598-020-70952-9

Landsberger, D. M., Padilla, M., Martinez, A. S., & Eisenberg, L. S. (2018). Spectral-temporal modulated ripple discrimination by children with cochlear implants. *Ear and Hearing, 39*(1), 60–68. https://doi.org/10.1097/AUD.0000000000000463

Landsberger, D. M., Stupak, N., Green, J., Tona, K., Padilla, M., Martinez, A. S., Eisenberg, L. S., & Waltzman, S. (2019). Temporal modulation detection in children and adults with cochlear implants: Initial results. *Otology & Neurotology, 40*(3), e311–e315. https://doi.org/10.1097/MAO.0000000000002122

Lewis, J. H., Castellanos, I., & Moberly, A. C. (2021). The impact of neurocognitive skills on recognition of spectrally degraded sentences. *Journal of the American Academy of Audiology, 32*(08), 528–536. https://doi.org/10.1055/s-0041-1732438

Litovsky, R. Y., Jones, G. L., Agrawal, S., & Van, H. R. (2010). Effect of age at onset of deafness on binaural sensitivity in electric hearing in humans. *The Journal of the Acoustical Society of America, 127*(1), 400–414. https://doi.org/10.1121/1.3257546

Loebach, J. L., & Pisoni, D. B. (2008). Perceptual learning of spectrally degraded speech and environmental sounds. *The Journal of the Acoustical Society of America, 123*(2), 1126–1139. https://doi.org/10.1121/1.2823453

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Science of the United States of America, 103*(49), 18866–18869. https://doi.org/10.1073/pnas.0607364103

Lyxell, B., & Rönnberg, J. (1993). The effects of background noise and working memory capacity on speechreading performance. *Scandinavian Audiology, 22*(2), 67–70. https://doi.org/10.3109/01050399309046021

MacCutcheon, D., Pausch, F., Fullgrabe, C., Eccles, R., van der Linde, J., Panebianco, C., Fels, J., & Ljung, R. (2019). The contribution of individual differences in memory span and language ability to spatial release from masking in young children. *Journal of Speech, Language, and Hearing Research, 62*(10), 3741–3751. https://doi.org/10.1044/2019_jslhr-s-19-0012

Martin, N., & Brownell, R. (2011). *Expressive One-Word Picture Vocabulary Test (EOWPVT-4)*. Academic Therapy Publications, Inc.

McCreery, R. W., Miller, M. K., Buss, E., & Leibold, L. J. (2020). Cognitive and linguistic contributions to masked speech recognition in children. *Journal of Speech, Language, and Hearing Research, 63*(10), 3525–3538. https://doi.org/10.1044/2020_jslhr-20-00030

McCreery, R. W., Spratford, M., Kirby, B., & Brennan, M. (2017). Individual differences in language and working memory affect children's speech recognition in noise. *International Journal of Audiology, 56*(5), 306–315. https://doi.org/10.1080/14992027.2016.1266703

McCreery, R. W., Walker, E. A., Spratford, M., Lewis, D., & Brennan, M. (2019). Auditory, cognitive, and linguistic factors predict speech recognition in adverse listening conditions for children with hearing loss. *Frontiers in Neuroscience, 13,* 1093. https://doi.org/10.3389/fnins.2019.01093

Meddis, R., & Hewitt, M. J. (1992). Modeling the identification of concurrent vowels with different fundamental frequencies. *The Journal of the Acoustical Society of America, 91*(1), 233–245. https://doi.org/10.1121/1.402767

Miller, G. A. (1947). The masking of speech. *Psychological Bulletin, 44*(2), 105–129. https://doi.org/10.1037/h0055960

Miller, G. A., Heise, G. A., & Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology, 41*(5), 329–335. https://doi.org/10.1037/h0062491

Miller, J. D., Watson, C. S., Dubno, J. R., & Leek, M. R. (2015). Evaluation of speech-perception training for hearing aid users: A multisite study in progress. *Seminars in Hearing, 36*(04), 273–283. https://doi.org/10.1055/s-0035-1564453

Miller, J. D., Watson, C. S., Kistler, D. J., Preminger, J. E., & Wark, D. J. (2008). Training listeners to identify the sounds of speech: II. Using SPATS software. *Hearing Journal, 61*(10), 29–33. https://doi.org/10.1097/01.HJ.0000341756.80813.e1

Moberly, A. C., Harris, M. S., Boyce, L., & Nittrouer, S. (2017). Speech recognition in adults with cochlear implants: The effects of working memory, phonological sensitivity, and aging. *Journal of Speech, Language, and Hearing Research, 60*(4), 1046–1061. https://doi.org/10.1044/2016_JSLHR-H-16-0119

Moberly, A. C., Lewis, J. H., Vasil, K. J., Ray, C., & Tamati, T. N. (2021). Bottom-up signal quality impacts the role of top-down cognitive-linguistic processing during speech recognition by adults with cochlear implants. *Otology &.Neurotology, 42,* S33–S41. https://doi.org/10.1097/mao.0000000000003377

Moberly, A. C., & Reed, J. (2019). Making sense of sentences: Top-down processing of speech by adult cochlear implant users. *Journal of Speech, Language, and Hearing Research, 62*(8), 2895–2905. https://doi.org/10.1044/2019_jslhr-h-18-0472

Nagels, L., Gaudrain, E., Vickers, D., Hendriks, P., & Baskent, D. (2021). School-age children benefit from voice gender cue differences for the perception of speech in competing speech. *The Journal of the Acoustical Society of America, 149*(5), 3328–3344. https://doi.org/10.1121/10.0004791

Nilsson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. *The Journal of the Acoustical Society of America, 95*(2), 1085–1099. https://doi.org/10.1121/1.408469

Nittrouer, S. (2010). *Early development of children with hearing loss*. Plural.

Nittrouer, S., & Boothroyd, A. (1990). Context effects in phoneme and word recognition by young children and older adults. *The Journal of the Acoustical Society of America, 87*(6), 2705–2715. https://doi.org/10.1121/1.399061

Nittrouer, S., & Burton, L. T. (2005). The role of early language experience in the development of speech perception and phonological processing abilities: Evidence from 5-year-olds with histories of otitis media with effusion and low socioeconomic status. *Journal of Communication Disorders, 38*(1), 29–63. https://doi.org/10.1016/j.jcomdis.2004.03.006

Nittrouer, S., Caldwell, A., Lowenstein, J. H., Tarr, E., & Holloman, C. (2012). Emergent literacy in kindergartners with cochlear implants. *Ear and Hearing, 33*(6), 683–697. https://doi.org/10.1097/aud.0b013e318258c98e

Nittrouer, S., Kuess, J., & Lowenstein, J. H. (2015). Speech perception of sine-wave signals by children with cochlear implants. *The Journal of the Acoustical Society of America, 137*(5), 2811–2822. https://doi.org/10.1121/1.4919316

Nittrouer, S., & Lowenstein, J. H. (2010). Learning to perceptually organize speech signals in native fashion. *The Journal of the Acoustical Society of America, 127*(3), 1624–1635. https://doi.org/10.1121/1.3298435

Nittrouer, S., Lowenstein, J. H., & Sinex, D. G. (2021). The contribution of spectral processing to the acquisition of phonological sensitivity by adolescent cochlear implant users and normal-hearing controls. *The Journal of the Acoustical Society of America, 150*(3), 2116–2130. https://doi.org/10.1121/10.0006416

Nittrouer, S., & Miller, M. E. (1999). The development of phonemic coding strategies for serial recall. *Applied Psycholinguistics, 20*(4), 563–588. https://doi.org/10.1017/S0142716499004051

Nittrouer, S., Muir, M., Tietgens, K., Moberly, A. C., & Lowenstein, J. H. (2018). Development of phonological, lexical, and syntactic abilities in children with cochlear implants across the elementary grades. *Journal of Speech, Language, and Hearing Research, 61*(10), 2561–2577. https://doi.org/10.1044/2018_jslhr-h-18-0047

Nittrouer, S., Tarr, E., Wucinich, T., Moberly, A. C., & Lowenstein, J. H. (2015). Measuring the effects of spectral smearing and enhancement on speech recognition in noise for adults and children. *The Journal of the Acoustical Society of America, 137*(4), 2004–2014. https://doi.org/10.1121/1.4916203

Peng, S. C., Tomblin, J. B., & Turner, C. W. (2008). Production and perception of speech intonation in pediatric cochlear implant recipients and individuals with normal hearing. *Ear and Hearing, 29*(3), 336–351. https://doi.org/10.1097/aud.0b013e318168d94d

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *The Journal of the Acoustical Society of America, 97*(1), 593–608. https://doi.org/10.1121/1.412282

Plomp, R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *The Journal of the Acoustical Society of America, 63*(2), 533–549. https://doi.org/10.1121/1.381753

Poll, G. H., Miller, C. A., & van Hell, J. G. (2016). Sentence repetition accuracy in adults with developmental language impairment: Interactions of participant capacities and sentence structures. *Journal of Speech, Language, and Hearing Research, 59*(2), 302–316. https://doi.org/10.1044/2015_jslhr-l-15-0020

Remez, R. E., Rubin, P. E., Berns, S. M., Pardo, J. S., & Lang, J. M. (1994). On the perceptual organization of speech. *Psychological Review, 101*(1), 129–156. https://doi.org/10.1037/0033-295x.101.1.129

Riches, N. G., Loucas, T., Baird, G., Charman, T., & Simonoff, E. (2010). Sentence repetition in adolescents with specific language impairments and autism: An investigation of complex syntax. *International Journal of Language & Communication Disorders, 45*(1), 47–60. https://doi.org/10.3109/13682820802647676

Richter, M. E., & Chatterjee, M. (2021). Weighting of prosodic and lexical-semantic cues for emotion identification in spectrally degraded speech and with cochlear implants. *Ear and Hearing, 42*(6), 1727–1740. https://doi.org/10.1097/aud.0000000000001057

Roid, G. H., & Miller, L. J. (2002). *Leiter International Performance Scale–Revised (Leiter-R)*. Stoelting Co.

Rönnberg, J., Lunner, T., Zekveld, A., Sorqvist, P., Danielsson, H., Lyxell, B., Dahlstrom, O., Signoret, C., Stenfelt, S., Pichora-Fuller, M. K., & Rudner, M. (2013). The Ease of Language Understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience, 7,* 31. https://doi.org/10.3389/fnsys.2013.00031

Rosen, S., Souza, P., Ekelund, C., & Majeed, A. A. (2013). Listening to speech in a background of other talkers: Effects of talker number and noise vocoding. *The Journal of the Acoustical Society of America, 133*(4), 2431–2443. https://doi.org/10.1121/1.4794379

Rudge, A. M., Coto, J., Oster, M. M., Brooks, B. M., Soman, U., Rufsvold, R., & Cejas, I. (2022). Vocabulary outcomes for 5-year-old children who are deaf or hard of hearing: Impact of age at enrollment in specialized early intervention. *The Journal of Deaf Studies and Deaf Education*. https://10.1093/deafed/enac009

Schneider, B. A., Pichora-Fuller, M. K., & Daneman, M. (2010). Effects of senescent changes in audition and cognition on spoken language comprehension. In S. Gordon-Salant, R. D. Frisina, A. N. Popper, & R. R. Fay (Eds.), *The aging auditory system* (pp. 167–210). Springer. https://doi.org/10.1007/978-1-4419-0993-0_7

Shafiro, V., Fogerty, D., Smith, K., & Sheft, S. (2018). Perceptual organization of interrupted speech and text. *Journal of Speech, Language, and Hearing Research, 61*(10), 2578–2588. https://doi.org/10.1044/2018_jslhr-h-17-0477

Simpson, S. A., & Cooke, M. (2005). Consonant identification in N-talker babble is a nonmonotonic function of N. *The Journal of the Acoustical Society of America, 118*(5), 2775–2778. https://doi.org/10.1121/1.2062650

Skuk, V. G., & Schweinberger, S. R. (2014). Influences of fundamental frequency, formant frequencies, aperiodicity, and spectrum level on the perception of voice gender. *Journal of Speech, Language, and Hearing Research, 57*(1), 285–296. https://doi.org/10.1044/1092-4388(2013/12-0314)

Smith, S. T., Macaruso, P., Shankweiler, D., & Crain, S. (1989). Syntactic comprehension in young poor readers. *Applied Psycholinguistics, 10*(4), 429–454. https://doi.org/10.1017/S0142716400009012

Smits, C., & Zekveld, A. A. (2021). Approaches to mathematical modeling of context effects in sentence recognition. *The Journal of the Acoustical Society of America, 149*(2), 1371−1383. https://doi.org/10.1121/10.0003580

Spahr, A. J., Dorman, M. F., Litvak, L. M., Van, W. S., Gifford, R. H., Loizou, P. C., Loiselle, L. M., Oakes, T., & Cook, S. (2012). Development and validation of the AzBio sentence lists. *Ear and Hearing, 33*(1), 112–117. https://10.1097/AUD.0b013e31822c2549

Stelmachowicz, P. G., Hoover, B. M., Lewis, D. E., Kortekaas, R. W., & Pittman, A. L. (2000). The relation between stimulus context, speech audibility, and perception for normal-hearing and hearing-impaired children. *Journal of Speech, Language, and Hearing Research, 43*(4), 902–914. https://doi.org/10.1044/jslhr.4304.902

Sullivan, J. R., Osman, H., & Schafer, E. C. (2015). The effect of noise on the relationship between auditory working memory and comprehension in school-age children. *Journal of Speech, Language, and Hearing Research, 58*(3), 1043–1051. https://doi.org/10.1044/2015_jslhr-h-14-0204

Sweetow, R. W., & Sabes, J. H. (2006). The need for and development of an adaptive Listening and Communication Enhancement (LACE) Program. *Journal of the American Academy of Audiology, 17*(8), 538–558. https://10.3766/jaaa.17.8.2

Tejani, V. D., Schvartz-Leyzac, K. C., & Chatterjee, M. (2017). Sequential stream segregation in normally-hearing and cochlear-implant listeners. *The Journal of the Acoustical Society of America, 141*(1), 50–64. https://doi.org/10.1121/1.4973516

Walker, E. A., Sapp, C., Oleson, J. J., & McCreery, R. W. (2019). Longitudinal speech recognition in noise in children: Effects of hearing status and vocabulary. *Frontiers in Psychology, 10,* 2421. https://doi.org/10.3389/fpsyg.2019.02421

Wingfield, A., Peelle, J. E., & Grossman, M. (2003). Speech rate and syntactic complexity as multiplicative factors in speech comprehension by young and older adults. *Aging Neuropsychology and Cognition, 10*(4), 310–322. https://doi.org/10.1076/anec.10.4.310.28974

Won, J. H., Drennan, W. R., & Rubinstein, J. T. (2007). Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. *Journal of the Association for Research in Otolaryngology, 8*(3), 384–392. https://doi.org/10.1007/s10162-007-0085-8

Woods, D. L., Doss, Z., Herron, T. J., Arbogast, T., Younus, M., Ettlinger, M., & Yund, E. W. (2015). Speech perception in older hearing impaired listeners: Benefits of perceptual training. *PLOS ONE, 10*(3), Article e0113965. https://doi.org/10.1371/journal.pone.0113965

Zheng, Y., Escabi, M., & Litovsky, R. Y. (2017). Spectro-temporal cues enhance modulation sensitivity in cochlear implant users. *Hearing Research, 351,* 45–54. https://doi.org/10.1016/j.heares.2017.05.009

Appendix (p. 1 of 3)

Sentences Used in Testing

Subject–Object Sentences

P. The dog that the bear chased wore a hat.
1. The picture that the housekeeper dusted fell off the shelf.
2. The soldier whom the president thanked walked up the stairs.
3. The cookie that the baby grabbed broke into pieces.
4. The student whom the principal scolded told his mother.
5. The computer that the technician rebooted caught on fire.
6. My aunt whom I see rarely came to visit.
7. The banana that I brought to school dropped from my bag.
8. The stars that the sailor saw came out at midnight.
9. The snow that the children shoveled blew across the field.
10. The baby that the mother held drank her bottle.
11. The girl that the man lifted held the basket.
12. The lady that the man touched held the blanket.
13. The bear that the lion followed ate the food.
14. The boy that the girl pushed hugged the cat.
15. The lamb that the goat chased bit the newspaper.
16. The doctor whom the nurse helped checked the baby.
17. The mailman whom the grandmother fed delivered the letter.
18. The sheep that the farmer cleaned drank the water.
19. The cowboy that the sheriff saw rode the train.
20. The teacher that the student greeted bought the book.
21. The plumber that the man hired fixed the pipe.
22. The dragon that the knight fought destroyed the village.
23. The policeman the woman called captured the robber.
24. The artist that the girl admired painted murals.
25. The cat that the mouse hated played with yarn.

Other Sentences

P.	**The puppy liked the boy who played with him most.**	**OS**
1.	The secretary sat at the desk that was painted red.	OO
2.	I saw the tree where the eagles built their nest.	OO
3.	The baby grabbed the cookie that the ants were on.	OO
4.	The driver went down the street that I live on.	OO
5.	The clown jumped in the car that the monkeys were pounding.	OO
6.	The dancers wore the shoes that the director brought.	OO
7.	The house was on the hill where lightning struck.	OO
8.	The baker showed the cake that the children had eaten.	OO
9.	The janitor picked up the apple that rolled across the floor.	OS
10.	The housekeeper dusted the picture that fell off the shelf.	OS
11.	The children shoveled the sand that was near the ocean.	OS
12.	The birds ate the worms that crawled on the driveway.	OS
13.	Sharks bite swimmers who float in the ocean.	OS
14.	Daddy hugged the child who made the birthday cake.	OS
15.	The singer thanked the crowd that came to see him.	OS
16.	The lady made the sign that said “free kittens.”	OS
17.	The girl who cleaned the kitchen ate some lunch.	SS
18.	The principal who thanked the student walked outside.	SS
19.	The man who opened the grocery store made the coffee.	SS
20.	The grass that grew in his yard got very wet.	SS
21.	The man that held a blanket touched the lady.	SS
22.	The lion that ate the food followed the bear.	SS
23.	The goat that bit a newspaper chased the lamb.	SS
24.	The nurse that checked the baby helped the doctor.	SS
25.	The grandmother that read the letter fed the mailman.	SS

Appendix (p. 2 of 3)

Sentences Used in Testing

Conjoined-Clauses Sentences

P. My aunt took a vacation and came to visit.
1. The carpenter hung the picture and fixed the door.
2. The president thanked the soldier and walked up the stairs.
3. Father drove way too fast and got a ticket.
4. The teacher gave out homework and went to lunch.
5. The technician rebooted the computer and drank some coffee.
6. The banana dropped from my bag and was quickly squashed.
7. The sailor saw the stars come out and went to bed.
8. The grass got very wet and turned bright green.
9. The mother held her baby and talked to the girl.
10. The bear chased the dog and wore a hat.
11. The man lifted the girl and held a basket.
12. The man touched the lady and held a blanket.
13. The lion followed the bear and ate the food.
14. The girl pushed the boy and hugged a cat.
15. The goat chased the lamb and bit a newspaper.
16. The grandmother fed the mailman and read the letter.
17. The farmer clipped the sheep and drank the water.
18. The sheriff pushed the cowboy and rode the train.
19. The student greeted the teacher and bought the book.
20. The pilot married the actress and bought the plane.
21. The man hired the plumber and went to work.
22. The knight fought the dragon and saved the village.
23. The woman called the policeman and hid under her bed.
24. The cat ate the mouse and played with yarn.
25. The artist painted murals and built some sculptures.

Four-Word Sentences

P. Cooks run in brooms.
1. Wide pens swim high.
2. Blocks can't run sharp.
3. Drive my throat late.
4. Stars find clean roof.
5. Tame beans test ice.
6. Green hands don't sink.
7. Bad dogs sail up.
8. Socks pack out ropes.
9. Cats get bad ground.
10. Sad cars want chills.
11. Hard corn feels mean.
12. Knees talk with mice.
13. Lend them less sleep.
14. Big apes grab sun.
15. Tin hats may laugh.
16. Soap takes on dogs.
17. Cars jump from fish.
18. Trucks drop sweet dust.
19. Let their flood hear.
20. Long kids stay back.
21. Thin books look soft.
22. Cups kill fat leaves.
23. Lead this coat home.
24. Four rats kick warm.
25. Soft rocks taste red.

Appendix (p. 3 of 3)

Sentences Used in Testing

Five-Word Sentences
P. The two farmers were talking.
1. The team is playing well.
2. They waited for an hour.
3. The silly boy is hiding.
4. The mailman shut the gate.
5. They knocked on the window.
6. He hung up his raincoat.
7. The apple pie was good.
8. The woman cleaned her house.
9. The painter uses a brush.
10. The ball bounced very high.
11. School got out early today.
12. Someone is crossing the road.
13. The man called the police.
14. He really scared his sister.
15. He found his brother hiding.
16. He wore his yellow shirt.
17. The young people are dancing.
18. They are coming for dinner.
19. They had a wonderful day.
20. He climbed up the ladder.
21. The sun melted the snow.
22. She argues with her sister.
23. The children helped their teacher.
24. The ball broke the window.
25. The family bought a house.