Published in final edited form as:
Int J Pediatr Otorhinolaryngol. 2021 February ; 141: 110514. doi:10.1016/j.ijporl.2020.110514.

When Language Outgrows Them: Comprehension of Ambiguous Sentences by Children with Normal Hearing or Hearing Loss

Susan Nittrouer, **Joanna H. Lowenstein**
Department of Speech, Language, and Hearing Sciences, University of Florida

Abstract

Background: In spite of recent gains in language development made by children with hearing loss (HL) as a result of improved auditory prostheses and earlier starts to intervention, these children continue to struggle academically at higher grade levels. We hypothesize that one reason for these incongruent outcomes for language and academics may be that the language demands of school escalate as grade level increases, outstripping the language abilities of children with HL. We tested that hypothesis by examining a higher level skill that is essential for success with academic language, the ability to access multiple interpretations for a sentence.

Method: 122 children participated at the end of middle school: 56 with normal hearing (NH), 15 with moderate HL who used hearing aids (HAs), and 51 with severe-to-profound HL who used cochlear implants (CIs). Children's abilities to provide more than one interpretation for an ambiguous sentence were assessed. These sentences were ambiguous due either to words having multiple meanings or to syntactic structure that could evoke more than one interpretation. Potential predictors of those abilities were evaluated, including expressive vocabulary, comprehension of syntactic structures, grammaticality judgements, forward digit span, and several audiologic factors.

Results: Children with NH performed best, children with CIs performed poorest, and children with HAs performed intermediately to those groups. Children in all groups achieved higher scores on the multiple meanings than on the syntactic structure items. The variables that were associated with performance varied across groups. Audiologic factors did not explain any variability in performance on the ambiguous sentences task for children with HL.

Conclusions: The kind of linguistic flexibility needed to consider more than one interpretation for sentences lacking immediate, real-world context is essential to processing academic language. Children with HL – especially those with severe-to-profound HL who required CIs – showed deficits in this skill, which could contribute to their ongoing academic struggles. Continued language support is needed for these children to allow them to acquire the higher level language skills necessary for success through all of their years in school.

Address correspondence to: Susan Nittrouer, Department of Speech, Language, and Hearing Sciences, University of Florida, PO Box 100074, Gainesville, FL 32610. Phone: 352-273-5303, snittrouer@phhp.ufl.edu.

Keywords

children; cochlear implants; language development; hearing loss; academic achievement

1. Introduction

There is no question that prospects for children born with hearing loss (HL) to develop functional spoken language have improved dramatically over the past thirty years. Universal newborn hearing screening means these children are routinely diagnosed and enrolled in interventions at younger ages than ever before. Advances in the design of hearing aids (HAs) allow for better speech recognition and noise reduction for children with sufficient residual hearing to use these devices, and cochlear implants (CIs) provide children with severe-to-profound HL auditory stimulation that far surpasses anything available previously. Studies assessing language skills in children with HL who received appropriate treatment early in life report that at the end of preschool these children obtain mean standard scores between 85 and 100 on commonly administered vocabulary and language instruments [1–5]. These scores place most of these children in the range of what is considered “normal language abilities,” which leaves many professionals feeling confident that they are sending children with HL off to mainstream academic environments ready to face whatever further language learning or academic challenges come their way.

Unfortunately, research conducted with these children during the school years suggests such optimism may not always be warranted. Although some studies report that children with HL continue, on average, to perform just one standard deviation (SD) below the mean of well-matched peers with normal hearing (NH) (i.e., standard scores of 85) on tasks measuring vocabulary or morphosyntax through elementary school [6–10], others have found that children with HL show more severe deficits [11–14]. In particular, children with HL appear to encounter the most difficulty with more complex language functions at higher grades [15]. Moreover, measures of academic achievement by children with HL reveal that any gains made in language due to improved auditory prostheses and earlier interventions are not translating into improvements in academic performance. High school students with HL have been found to perform more than one SD below the means of students with NH on measures of reading comprehension, science, mathematics, and social studies [16;17]. In fact, children with HL are reading no better today than they were a generation ago, before cochlear implants were available as a treatment option [18]. Students with HL graduate from college at half the rate of NH peers, and these proportions have not changed since the middle of the 20th century [19]. Overall, these studies show depressed academic performance for children with HL, even when standard language measures suggest skill levels closer to those of peers with NH.

1.1. Academic language and more complex language skills

This seeming contradiction between the language abilities and academic performance of children with HL may reflect differences in the language used in everyday activities versus that used in school. These differences take a variety of forms, including vocabulary that is used only in the academic setting [20–23]. This vocabulary may be specific to the topic being studied (i.e., math or science) or be more generalized (i.e., instructional vocabulary). Much of the vocabulary of instruction is idiosyncratic to the school setting. It describes complex concepts, uses abstraction, and requires higher-order thinking [24;25]. Even when instructional language uses everyday vocabulary, it may do so in ways that differ from how it is used outside of school [26]. This increase in lexical complexity ramps up gradually as grade level gets higher [27]. Similarly, there are syntactic constructions that are used more frequently and in a requisite manner in academic settings; these typically involve long sentences with multiple embedded clauses [23;24]. Furthermore, academic language is decontextualized [28], meaning one cannot rely on personal experiences to aid comprehension. It is abstract and informationally dense, meaning more information is conveyed per utterance [29;30]. These differences in everyday and academic language may account for the discrepancies in outcomes for commonly used language measures and academic performance observed for children with HL. Almost without exception the standard instruments used to assess language abilities in children consist of items that can be considered part of everyday language.

Many complex sentences can have multiple interpretations, and without immediate context to specify the appropriate interpretation students need to have the linguistic knowledge and flexibility to remedy this inherent ambiguity. These multiple interpretations can arise from either the vocabulary used or the syntactic structure. Where vocabulary is concerned, variability in meaning of individual words across topics leads to ambiguity. The Coxhead Academic Word List [21] is a corpus of words found in a variety of academic subject areas. Hyland and Tse [31] explored the distribution of words in the Coxhead list across academic domains, and found that many words on that list have different meanings in different fields of study. For example, the word *issue* is often used to mean “to flow out” in science, but it is used to mean “the topic or problem of focus” in other subject areas. In another example, the word *attribute* can refer to a feature or a quality of a person or thing, but it can also mean “to ascribe to,” as in being the cause of something.

Where syntax is concerned, it is usually the lack of immediate, real-world context that can create the ambiguity. The sentence *The boy, the man’s son, and the girl walked to school together* is ambiguous in surface form, but that ambiguity would be easily resolved with real-world information regarding how many people were in the group walking to school; it would not require exquisite knowledge of how subordinate clauses operate. On the other hand, such knowledge could help resolve ambiguity in the sentence *The ossicles, the smallest bones in the body, and the tympanic membrane comprise the middle ear cavity* if the reader or listener did not have prior knowledge that the ossicles are indeed the smallest bones in the body. Overall, the ability to recognize the multiple possible interpretations of an utterance is the first step in understanding language in academic settings when that language is ambiguous. Without that ability students can be left behind in lectures or reading assignments, trying to apply the only interpretations they can derive to utterances where they do not fit.

1.2. The development of ambiguity recognition

The ability to recognize that a lexical item or a syntactic structure might support more than one interpretation is a skill that develops across childhood. Children as young as three years of age can point to pictures to demonstrate that they know a word such as *bat* can have more than one meaning, and that those meanings can represent different categories, such as an animal or an object [32]. However, the ability to provide two separate interpretations for a sentence that contains a lexical or syntactic ambiguity does not emerge until later in childhood [33]. Shultz and Pilon [34] looked at the comprehension of sentences with lexical ambiguity (i.e., multiple word meanings) or ambiguous syntactic structures in children ages 6, 9, 12 and 15 years. The sentence *He often goes to the bank* is an example of lexical ambiguity, because *bank* can mean either a financial institution or the side of a river. The sentence *He laughed at the school* is an example of an ambiguous syntactic structure, because either the school is the source of amusement or the boy is laughing while physically at school. Children were asked to restate each sentence in their own words (i.e., paraphrase), and were asked if there were any other interpretations to the sentence if they did not spontaneously provide a second one. The ability to provide both interpretations of the sentences involving lexical ambiguity started to emerge between the ages of 6 and 9, but children in that age range were unable to provide two interpretations of sentences involving syntactic ambiguity. In fact, recognition of syntactic ambiguity was not seen until the age of 12, and was still emerging at the age of 15. Similarly, Cairns et al. [33] found that 7-year-old children were able to recognize lexical ambiguity, but had much more difficulty recognizing syntactic ambiguity.

Jokes are often based on the kind of linguistic ambiguity being described. Hirsh-Pasek and colleagues [35] examined the comprehension of jokes specifically based on ambiguity in a group of children from 6 to 11 years of age. These jokes made use of both lexical and syntactic ambiguity. The 6- to 7-year-olds in the study found the task more difficult than the 10- to 11-year-olds. Across ages, they found that children made fewer errors for the items involving lexical ambiguity than for the items involving syntactic ambiguity. In another study, Spector [36] examined the comprehension of linguistic humor in typically developing and language-impaired high school students. Elements of this humor included ambiguous lexical items, as well as ambiguous syntactic structures. Here it was found that typical adolescents could describe the humorous elements more easily than the language-impaired adolescents, suggesting that linguistic ambiguity may present special challenges for children with language deficits.

Johnson, Ionson, and Torreiter [37] assessed the abilities of children between the ages of 7 and 10 years diagnosed with language-learning problems to provide multiple meanings for individual words. They made use of 20 words, each with at least four different meanings. Half the words were presented in sentences that provided context supporting one specific word meaning (four separate sentences per word), and half were presented in isolation. The children were asked to define the target word in each sentence, or to provide as many meanings as they could for the words in isolation. The authors found that these children were better at providing definitions when the words were presented in sentences rather than in isolation. Without sentence context, children could often provide a single meaning, but then had great difficulty identifying other meanings. This finding illustrates how strongly language comprehension depends on context, at least for children with language deficits.

Overall, these studies demonstrate that in typical development the ability to recognize ambiguity in linguistic structures that arises from words having multiple meanings emerges before the ability to detect ambiguity that involves syntactic structure. Furthermore, children and adolescents with language impairments have more difficulty coming up with multiple interpretations of ambiguous words, sentences, or jokes than children who are developing language typically. This deficit can be problematic, because it restricts language comprehension – in either oral or written presentation – to a single, sometimes incorrect interpretation.

1.3. Current study

The purpose of the current study was to assess the abilities of adolescents with HL and their peers with NH to provide different meanings of sentences constructed to be ambiguous, either due to multiple meanings of a word or to ambiguous syntactic structure. To do this, the Ambiguous Sentences subtest of the Comprehensive Assessment of Spoken Language (CASL) [38] was used. This is a standardized test that was designed for this purpose. In addition, we wanted to examine potential sources of variability in children's abilities to recognize alternative interpretations. It is possible that the ability to consider multiple interpretations is a distinct skill, independent of other language or cognitive abilities. Alternatively, a deficit in the ability to consider multiple interpretations could arise from another underlying language problem: for example, if a child simply has a weak vocabulary that child may not know more than one meaning for some words. Similarly, a child may not understand how some syntactic constructions function. To help decide between these alternative explanations for any findings that might result from this study, three other standardized language measures were administered and evaluated as potential predictor variables. First, a test of vocabulary knowledge (the Expressive One-Word Vocabulary Test, or EOWPVT [39]) was administered. Expressive vocabulary was examined, rather than receptive, because expressive vocabulary measures require children to retrieve the items from their lexicons, rather than simply to recognize a word that is presented and identify it in a group of pictures, which is the procedure for receptive vocabulary measures. Thus a higher level of vocabulary knowledge is being tested with expressive measures. In addition, a test of syntactic comprehension (the Sentence Comprehension of Syntax subtest of the CASL) and a test of knowledge about grammatical forms (the Grammaticality Judgment subtest of the CASL) were administered. This last subtest assesses a broader array of grammatical structures than the Sentence Comprehension of Syntax subtest, including noun number and verb tense. Thus the child must have knowledge of grammatical morphemes to do well on this subtest.

In addition to those language measures, the possibility was examined that children's working memory might help to explain their abilities to provide multiple interpretations for ambiguous utterances, because they need to be able to retain a whole sentence in memory long enough to consider multiple interpretations. For this purpose, forward digit span was measured.

2. Method

2.1. Listeners

In total 122 children participated in this study, all of whom had just completed eighth grade at the time of testing. All were participants in a longitudinal study [4], and most had been tested since they were infants. Fifty-six of these children had NH, meaning that thresholds for the octave frequencies between 0.25 kHz and 8.0 kHz were better than 20 dB hearing level in both ears. Fifteen of the children with HL used HAs, and the remaining 51 children with HL used CIs. Details of audiologic measures for the children with HL are presented in Table 1. Although not shown in this table, all children with HL received amplification and began intervention before turning 30 months of age. Twenty-five children with CIs wore a HA on the unimplanted ear for a period of a year or more at the time of receiving their first CIs (i.e., bimodal experience), and the other 26 children never wore a HA after receiving a first CI. Table 2 describes histories of auditory prostheses use for these children.

Children were well-matched on socioeconomic status (SES), which was assessed using an index that has been used before [40]. On this scale occupational status and highest educational level for each parent are ranked on scales from 1 to 8, from lowest to highest. It is based on the original methods of Hollingshead [41], but with occupations updated to reflect more modern jobs. These categories are shown in Appendix A. These scores are multiplied together for each parent separately, and the highest value obtained is used as the socioeconomic metric for the family. Scores of 30 and higher indicate that at least one parent had a four-year university degree or better, and a job commensurate with that level of education.

Non-verbal intelligence was also assessed, to ensure that any group differences would not reflect differences in general cognitive abilities. To assess non-verbal intelligence, the Leiter International Performance Scales – Revised [42] was given. This instrument assesses non-verbal reasoning, requires no verbal responses, and presents instructions through mime. The “Brief IQ” score was used, which consists of four subtests: figure-ground perception, form completion, sequencing abilities, and repeated patterns recognition.

All test materials were presented in audio-visual format. Nonetheless we assessed children’s audio-only speech recognition abilities to ensure that results could not be due to children not recognizing the test material. Because test materials were presented in quiet, that is how speech recognition was assessed. Five-word sentences derived from the Hearing in Noise Test [43] were used, but presented without noise. Twenty-five sentences were presented in audio-only format over a speaker 1 m from the child at 68 dB sound pressure level. These sentences had been recorded by a male speaker with a Midwest dialect. Participants repeated each sentence, and their responses were audio-video recorded for later scoring. One member of the laboratory staff scored these responses, and a second staff member scored 20% of the responses. Average word-by-word agreement between the scores of each staff member was .99, which was considered adequate reliability. Scores from the first staff member were used in reporting. Percent correct words was the measure derived from this assessment.

Table 3 displays means and SDs for each group for the demographic variables. The difference in mean age was significant, $F(2,119) = 4.507$, $p = .013$, indicating that the children with HL were, on average, a few months older than the children with NH. That was not considered a problem, because all children were at the same grade level. Differences in mean SES were not significant, nor were differences in the mean Brief IQ scores. Differences in mean scores on the speech recognition in quiet task were significant, $F(2,119) = 6.961$, $p = .001$, reflecting the slightly poorer performance of the children with HL. However, as the children with HL were able to repeat an average of 98–99% of words correctly, they should have had adequate recognition of speech in quiet. The addition of visual information in the test format used here would have only made recognition of the material presented that much more accessible.

2.2. Equipment

All testing was conducted in a soundproof booth. The materials for the CASL subtests (Ambiguous Sentences, Sentence Comprehension of Syntax, and Grammaticality Judgment) were video-recorded by a female speaker, and presented in audio-video format on a computer, rather than by live voice as is typically done in a clinical setting. This presentation mode ensured consistency of materials across subjects. All audio signals were presented with a Creative Labs Soundblaster soundcard and a Roland MA-12C powered speaker placed 1 m in front of the child at 0° azimuth. This system had a 44.1-kHz sampling rate, and 16-bit digitization. Video was presented on a widescreen monitor at a rate of 1,500-kilobits per second.

Digits in the digit span task was presented in audio-only format using the same soundcard and speaker as that used for the CASL subtests. The ability of each child to recognize each digit was checked before testing. Custom-written software controlled the presentation of the recorded digits. After the presentation of each list of digits, numerals appeared at the top of a touchscreen monitor and responses were collected by having children touch these numerals in the order recalled.

Presentation level was 68 dB sound pressure level for all tasks. All tasks, except for the digit span task, were audio-video recorded using a Sony HDR-XR550V video camera so that scoring could be done later. Children wore Sony FM transmitters in specially designed vests. The FM receivers provided direct-line input to the video camera to ensure good sound quality on the recordings.

2.3. Stimuli and Procedures

2.3.1 General Procedures—All procedures were approved by the local Institutional Review Board. Children came to the laboratory for two consecutive days of testing, in groups of two to six children. The tasks were administered in three individual test sessions of no more than one hour each, and children were given breaks between sessions. The measures analyzed are listed in Table 4.

2.3.2. Ambiguous sentences—Comprehension of ambiguous sentences was assessed using the Ambiguous Sentences subtest of the CASL [38]. In this task, participants were presented with a sentence that is ambiguous in interpretation, and they had to provide two possible interpretations for the sentence. The task consists of 43 test sentences. Two additional items were presented as practice before testing. The 43 test sentences are categorized by the test authors as ambiguous due to multiple word meanings (22 items) or due to syntactic structure (21 items) [44]. Those categorizations were used to assess whether children in this study were able to provide alternative sentence interpretations more readily for one kind of ambiguity than for the other. The prediction was that children would more readily recognize multiple interpretations when lexical ambiguity (i.e., multiple meanings) was involved, rather than syntactic ambiguity.

The published ceiling for this subtest is five consecutive incorrect items, meaning that testing should be discontinued when a child fails to provide two appropriate alternative interpretations for five items in a row. However, pilot testing was conducted with an additional group of children with NH without using a ceiling. That work revealed that some participants provided correct answers for items past the published ceiling of five incorrect answers, but not past the point of ten consecutive incorrect answers. Thus for this project an experimental ceiling rule of ten consecutive incorrect items was used in testing.

Four dependent measures derived from the Ambiguous Sentences task were computed for this study: (1) Standard scores derived using the published test ceiling of five consecutive errors; (2) Raw numbers of correct answers for the 43 items derived using the experimental ceiling of 10 consecutive errors; (3) Raw numbers of correct answers for just the 22 items that are ambiguous due to multiple word meanings, derived using the experimental ceiling of 10 consecutive errors; and (4) Raw numbers of correct answers for just the 21 items that are ambiguous due to syntactic structure, derived using the experimental ceiling of 10 consecutive errors.

2.3.3. Expressive vocabulary—Expressive vocabulary was assessed using the Expressive One-Word Picture Vocabulary Test – 4 [39]. In this task, children are shown a series of pictures and must label each one with a single word. Testing is discontinued after six consecutive errors. This test is normed for individuals from two years through adulthood, so there was no risk of reaching the limit of test items before a child made six consecutive errors.

2.3.4. Sentence comprehension of syntax—Comprehension of syntactic structure was assessed using the Sentence Comprehension of Syntax subtest of the CASL. In this task, children are presented with pairs of sentences that differ in syntactic structure. They must decide if the two sentences have the same meaning, and simply report *yes* or *no*. These pairs are presented in sets of two (i.e., pairs of pairs), where the first sentence of each pair in a set is the same; only the second sentence of each pair differs. There are a total of 21 of these two-pair test items, and the child must respond correctly to each pair in the item to get credit for that item. Testing is discontinued after five consecutive errors. Two practice pairs were presented before testing.

2.3.5. Grammaticality judgments—The Grammaticality Judgment subtest of the CASL was also administered. This task consists of the presentation of single sentences that may or may not be grammatically correct. The child must decide if the sentence is correct, and provide a corrected version that changes only a single word of the sentence if it is reported as being incorrect. This task differs from the Sentence Comprehension of Syntax subtest in that it tests knowledge of morphology, as well as of syntax. It provides a deeper assessment of language proficiency than does the Sentence Comprehension of Syntax subtest.

For items on the Grammaticality Judgment subtest that include incorrect sentences, children receive a point if they can identify the sentence as incorrect, and another point if they can appropriately correct the error. A point is given for identifying a correct sentence as such. There are 57 items in this test: 46 items that need to be corrected, and 11 correct items. The maximum possible raw score is 103. Testing is discontinued after errors on five consecutive items. Three practice sentences were presented before testing.

2.3.6. Forward digit span—To assess working memory, the Forward Digit Span test of the Wechsler Intelligence Scale for Children [45] was used. A computer program was used to present recorded digits. After the digits were presented, the digits appeared at the top of a touchscreen monitor, and the child's task was to tap them in the order recalled. Two practice sequences were presented before testing. The length of the longest digit sequence recalled was used in analyses.

2.4. Scoring and analyses

For the Ambiguous Sentences task, the experimenter working with the child scored each item as testing proceeded, but two members of the laboratory staff independently rescored all items later offline. These scores were compared on an item-by-item basis and any scoring discrepancies were resolved by the second author. The Sentence Comprehension of Syntax, Grammaticality Judgment, and Expressive Vocabulary tasks were scored by the experimenter, and then reviewed by a laboratory staff member who was not present at the time of testing. This staff member watched the video recording, and assessed whether the scoring by the experimenter at the time of testing was accurate. If there was a discrepancy, it was corrected according to the second observer, who had access to the video recording so could replay it as necessary. Few scores needed to be changed. Digit span was scored by the computer.

All data were entered into a data file by one staff member, and subsequently checked by a second staff member. Data were screened for normality of distribution and homogeneity of variance. Data for all measures were found to meet appropriate criteria, so no transformations were applied to most measures, and parametric statistics were used for all measures. The one case in which transformations were applied involved an analysis in which scores for the Ambiguous Sentences items with multiple meanings were compared to the items for syntactic structure. This is explained below.

Standard scores are reported for measures when available, but raw scores were used in statistical analyses because they are slightly more sensitive indices of performance across groups than are standard scores. There are several reasons for this difference. One reason is that two or more raw scores may be associated with a single standard score. Also the association between raw and standard scores changes according to age. Thus, two children who are similar in age and score the same on a standardized test may get different standard scores, if they are on different sides of an age boundary. Furthermore, a non-standard ceiling was used for the Ambiguous Sentences measures, and standard scores are not available in that case. Finally, for this study scores were separated according to whether they involved items representing multiple meanings or syntactic structure; again, that precluded the use of standard scores.

3. Results

3.1. Ambiguous sentences

The top row of Table 5 shows means, medians, and SDs of standard scores for the Ambiguous Sentences task obtained with the ceiling rule of five consecutive incorrect items. Mean performance of children with NH and children with HAs was near the normative mean on this task, while children with CIs performed more poorly. A one-way analysis of variance (ANOVA) was conducted with hearing group as the between-subjects factor and post hoc comparisons using Bonferroni corrections. The first row of Table 6 shows results of this analysis. The main effect of group was significant, and only the post hoc comparison of children with NH versus children with CIs was significant; the other two comparisons were not (NH versus HA and HA versus CI). The effect size for the mean difference between children with NH and those with CIs, given as Cohen’s *d*, is shown in the last column.

The second row of Table 5 shows means, medians, and SDs of raw scores for all items on the Ambiguous Sentences task using the higher ceiling of 10 consecutive items incorrect. The second row of Table 6 shows outcomes of a one-way ANOVA performed on these data. Appendix B describes how many children had higher raw scores on the task when this ceiling was implemented instead of the published ceiling of five incorrect items. Roughly half of all participants (56/122) were able to provide two meanings for at least one additional test item after having missed five items in a row. Overall, both of these first two analyses reveal that children with CIs were poorer than children with NH at providing multiple interpretations for sentences. When it comes to children with HAs, it was found that they performed neither statistically poorer than children with NH nor statistically better than children with CIs. Therefore it can only be concluded that the abilities of children with sufficient residual hearing to use HAs to provide alternative interpretations for ambiguous sentences fell intermediate to those of children with NH and those with CIs.

The third row of Table 5 shows means, medians, and SDs of raw scores for the 22 items from the Ambiguous Sentences task that involved multiple meanings. The pattern of scores looks very similar to the patterns in the first two rows, indicating that relative performance for the three groups on that portion of the task was similar to overall performance. A one-way ANOVA performed on these data confirmed that impression. The third row of Table 6 shows these results. In this analysis the main effect of group was significant, and again only the post hoc comparison of children with NH versus children with CIs was significant.

The last row of Table 5 shows means, medians, and SDs of raw scores for the 21 items from the Ambiguous Sentences task that were determined to involve syntactic structure. It is apparent that children in all groups had more difficulty with these items than with the multiple-meaning items; on average they were able to answer correctly five fewer syntactic than multiple meaning items. A one-way ANOVA was performed on these data, and results are shown in the last row of Table 6. These results are similar to those seen for the other Ambiguous Sentences measures.

The results described above indicate that children with HL severe enough to require CIs had more difficulty providing alternative interpretations for ambiguous sentences than did children with NH; results for children with HAs were intermediate to and not significantly different from those of either of the other groups. An additional question that could be asked is whether there were group differences in how the two kinds of ambiguity were handled by children in the three groups. To examine potential group differences with the two types of Ambiguous Sentences test items, a two-way, repeated-measures ANOVA was performed, with type of ambiguity as the repeated measure and group as the between-subjects factor. Because there were slightly different numbers of items used for each type of ambiguity (22 for multiple meaning items and 21 for syntactic structure items), raw scores were converted to proportions by dividing the raw score by the total number of items for each type of ambiguity. Arcsine transformations were then performed on those proportions, and used in the analysis. The main effect of type of ambiguity was significant, $F(1,119) = 93.668$, $p < .001$, $\eta^2 = .440$, as was the main effect of group, $F(2,119) = 6.136$, $p = .003$, $\eta^2 = .093$. The interaction of Type of Ambiguity x Group, however, was not significant. These results indicate that children in all groups were better able to provide two alternative interpretations for sentences when the ambiguity was based on words having multiple meanings than on ambiguous syntactic structure. The effect of type of ambiguity was consistent across groups.

3.2 Potential predictor measures

3.2.1. Expressive vocabulary—The first row of Table 7 shows means, medians, and SDs for the standard scores on the expressive vocabulary measure. Appendix C provides means, medians, and SDs for raw scores. As can be seen in Table 7, mean scores for all groups were above the normative mean on this task. Nonetheless, there were differences among the groups. The first row of Table 8 provides results of a one-way ANOVA supporting that observation. Post hoc comparisons with Bonferroni corrections were computed for this vocabulary measure, as well as the other three measures shown in Table 7. As was seen for scores on the Ambiguous Sentences test, there was a significant difference in performance between children with NH and children with CIs, but no significant differences between children with HAs and either of the other two groups. Consequently it can be concluded that children with NH performed best, children with CIs performed poorest, and children with HAs performed intermediate to those two groups.

3.2.2. Sentence comprehension of syntax—The second row of Table 7 shows means, medians, and SDs for the standard scores on the Sentence Comprehension of Syntax subtest of the CASL. Appendix C provides means, medians and SDs for raw scores. Children with NH and children with HAs performed very similarly to each other, and both groups appear to have better performance than the children with CIs. The second row of Table 8 provides results of a one-way ANOVA supporting at least the first of those observations. The mean difference in scores between children with NH and those with CIs was found to be significant; however, the mean score for children with HAs was not found to be significantly different from those of children with NH or children with CIs.[1] Effect sizes for this measure, whether considered as η^2 or *d*, were the smallest of all the measures, indicating the least variability in performance across groups.

3.2.3. Grammaticality judgments—The third row of Table 7 shows means, medians, and SDs for the standard scores on the Grammaticality Judgment subtest of the CASL, and the third row of Table 8 provides results of a one-way ANOVA performed on these data. Appendix C provides means, medians, and SDs for raw scores. In this case, the mean score for children with NH was at the normative mean for the instrument, children with HAs performed somewhat – though not statistically – more poorly, and children with CIs demonstrated the poorest performance, with their mean score nearly one SD below the normative mean. In order to test this impression that this subtest provided more serious challenges to the children in this study, a two-way, repeated-measures ANOVA was performed on standard scores for the Sentence Comprehension of Syntax and the Grammaticality Judgment subtests. Unlike other analyses, standard scores were used in this case because these subtests were normed on the same samples of children, and in this case, having different numbers of items would confound results. There were significant main effects: subtest, $F(1,119) = 59.553$, $p < .001$, $\eta^2 = .334$; and group, $F(1,119) = 7.344$, $p = .001$, $\eta^2 = .110$. The Subtest x Group interaction was not significant, indicating that children in all three groups performed more poorly on the Grammaticality Judgment subtest than on the Sentence Comprehension of Syntax subtest. These outcomes support the suggestion that these children found the Grammaticality Judgment task to be more challenging than the Sentence Comprehension of Syntax task. And although mean standard scores for all groups were lower for this measure than for the vocabulary or sentence comprehension measures, the Cohen's *d* indexing the effect size of difference between children with NH and those with CIs was largest for this measure. That finding mirrors other findings observed for these same children at younger ages [46]. In that earlier work this trend was attributed to the fact that scores for the Grammaticality Judgment task loaded higher on a phonological factor than did scores for the Sentence Comprehension of Syntax task. That strong loading of Grammaticality Judgment scores on a phonological factor is likely due to the inclusion of multiple inflectional morphemes as targets in this subtest; these low-salient elements are essentially phonological components of the words they help form. Children with CIs have phonological deficits that are disproportionately large compared to any deficits observed for morphosyntactic abilities [46–48], a factor that likely explains why their performance on the Grammaticality Judgment subtest was lower when compared to children with NH than their performance on the Sentence Comprehension of Syntax subtest.

---

[1] The failure to find a significant difference between means of children with HAs and children with CIs when a difference of similar size between means of children with NH and children with CIs was found to be significant may be attributable to the small sample of children with HAs. The computed difference between means of children with NH and those with CIs is 1.80, with a 95% confidence interval of .508 to 3.10. The computed difference between means of children with HAs and those with CIs is 1.44, with a 95% confidence interval of −0.839 to 3.726. Thus, the confident interval is larger due to the smaller sample. This explanation may hold for other insignificant findings of means between children with HAs and those with CIs, but without a larger sample there is no way to know for sure. This may also explain why some differences between children with NH and children with HAs were not found to be significant.

3.2.4. Forward digit span—

The last row of Table 7 shows means, medians, and SDs for forward digit span, and the fourth row of Table 8 presents results of a one-way ANOVA for this measure. Mean digit span for both the groups of children with NH and children with HAs was almost 1 digit longer than mean digit span for children with CIs, and this difference was significant. For this measure, post hoc comparisons were significant for both NH versus CI and HA versus CI, confirming this observation. As with grammaticality judgments, measures of verbal working memory have been observed to load high on a phonological factor [46], and that fact likely explains why children with CIs performed more poorly on this task, relative to children in the other groups, than on the measures of vocabulary and sentence comprehension.

3.3. Predicting the ability to recognize sentence ambiguity

3.3.1. Audiologic factors—

The four measures described above were explored as potential predictors for the two types of sentence ambiguity (multiple meanings and syntactic structure) separately, in analyses described below. Factors related to the children's HL, however, were examined as potential predictors of the Ambiguous Sentences task as a whole. For children with HAs, Pearson product-moment correlation coefficients were calculated between Ambiguous Sentences raw scores and age of identification, as well as unaided and aided three-frequency PTA thresholds at the time of testing. None of these correlation coefficients was significant. For the children with CIs, Pearson product-moment correlation coefficients were calculated between Ambiguous Sentences raw scores and each of the following: age of identification, unaided, three-frequency PTA thresholds obtained just before implantation, aided three-frequency PTA thresholds obtained at the time of testing, and age at first CI. For children with two CIs, age of second CI was also considered. None of these correlation coefficients was significant, indicating that these audiologic factors did not explain any variance on this task.

As a final analysis, the effect of bimodal experience was considered for the children with CIs. Raw scores on the Ambiguous Sentences task (and SDs) were 16.7 (8.8) for the 25 children with at least one year of bimodal experience, and 11.4 (8.5) for the 26 children who did not have any bimodal experience. This difference was significant, $t(49) = 2.193$, $p = .033$, with Cohen's $d = 0.61$. Having a period of at least a year of bimodal experience was associated with higher scores on this task.

3.3.2. Predicting ambiguity arising from words with multiple meanings—

To examine the other skills that contributed to performance with the items from the Ambiguous Sentences task involving words with multiple meanings, Pearson product-moment correlation coefficients were computed between the raw scores for multiple meaning items on the Ambiguous Sentences task and each of the four potential predictor measures. These analyses were performed for each group separately, and are presented in Table 9. Correlation coefficients were generally higher for the language-related measures of expressive vocabulary, sentence comprehension, and grammaticality judgments than for forward digit span. Nonetheless, it is of some interest that even though these items involved words that have multiple meanings, for no group was vocabulary the measure with the highest correlation coefficient.

All four measures were rather strongly correlated with these multiple-meaning items from this Ambiguous Sentences task, so stepwise regression analyses were done to determine which measures accounted for the most unique variance in raw scores for multiple meaning items. To address concern that multicollinearity could constrain the validity of the regression analysis, variance inflation factors were computed, but they did not reveal evidence of significant multicollinearity.

These stepwise regression analyses were performed on scores for each group separately, and outcomes are shown in Table 10. For both the children with NH and the children with CIs, the most unique variance was explained by scores on the Grammaticality Judgment subtest. This finding may reflect the fact that these two measures – Ambiguous Sentences and Grammaticality Judgments – require the most sophisticated linguistic skills of all measures included in this study. Therefore, children who score well on one of these measures may simply score well on the other, because they are indexing overall linguistic proficiency. Instead, the predictor variables that explain additional variance may be of more interest, providing a sense of what skills specifically explain abilities to provide multiple interpretations of ambiguous sentences. Here, the pattern for children with NH seems to fit better with the nature of the multiple meaning items than the pattern for the children with HL. For children with NH, vocabulary scores explained additional unique variance on these multiple meaning items. That is reasonable given that the task would seem to be assessing deep lexical knowledge. For children with CIs, however, scores on the Sentence Comprehension of Syntax task were the only ones to explain additional unique variance. Although the reason for this finding cannot be known without further study, it may reflect a different language processing strategy on the part of children with CIs. These children appear to have been attempting to use syntactic constraints to access alternative interpretations for these sentences when word meaning was uncertain. This suggested pattern for children with CIs is only more apparent for the children with HAs, for whom the only significant predictor variable involved scores on the Sentence Comprehension of Syntax task. Overall it appears that these children with HL were depending on their knowledge of syntax instead of their knowledge of vocabulary to support their comprehension of ambiguous sentences with multiple meanings.

3.3.3. Ambiguity arising from syntactic structure—To examine the language skills that contributed to performance with the items from the Ambiguous Sentences task that were determined to involve syntactic structure, Pearson product-moment correlation coefficients were computed between the raw scores for syntactic structure items on the Ambiguous Sentences task and each of the other measures. These analyses were performed for each group separately, and are presented in Table 11. As with the multiple meaning items, all of these correlation coefficients are rather high. Therefore, in order to determine which measures accounted for the most unique variance in raw scores for syntactic structure items, stepwise regressions were performed for each group separately, with sentence comprehension, grammaticality judgments, vocabulary, and digit span as the predictor variables. Variance inflation factors were again computed, and again failed to reveal any evidence of significant multicollinearity.

Outcomes for these stepwise regression analyses are shown in Table 12. Again, for both the children with NH and the children with CIs, the most unique variance was explained by scores on the Grammaticality Judgment subtest. For children with NH, the only measure that explained any additional unique variance was sentence comprehension. This finding makes sense, as this measure assesses syntactic comprehension. In this case, the measure of working memory – digit span – was the factor included in the second step of this analysis for children with CIs, indicating that it explained the most unique variance, after grammaticality judgments. This finding likely reflects the fact that children needed to retain the entire sentence in a memory buffer in order to analyze its syntactic structure, unlike the multiple meaning items where children only needed to retain a single word. However, it was the factor that was included in the third step for children with CIs that was surprising in this case. Instead of being sentence comprehension, as was seen for children with NH, it was vocabulary. And vocabulary was the only variable that explained unique variance in this analysis for children with HAs. In a trend that is exactly opposite to that observed for multiple meaning items, it appears that children with HL may have been attempting to use lexical constraints to derive alternative interpretations for these sentences with ambiguous syntactic structure.

4. Discussion

Prospects that children born with HL will develop functional spoken language have improved greatly with the advent of better auditory prostheses and earlier starts to intervention. During the early school years many of these children perform academically on par with their peers with NH. But academic performance of children with HL has been observed to decline as grade level increases, in some areas appearing even to eradicate the advantages afforded by better prostheses and earlier intervention. These grade-related declines have not been recognized until recently, so have not been well examined. The purpose of the current study was to explore one potential source of the academic challenges that children with HL are encountering at later grades.

The language of school grows increasingly disconnected from any immediate, real-world context as children progress through grade levels. It becomes more abstract, and students must be able to recover meaning based on the specialized content of the subject being taught. In course work at these higher grades, it is often the case that any individual sentence may have multiple potential interpretations, with the correct one for the immediate oral presentation (e.g., lecture) or written document (e.g., textbook) determined by the specific subject area. Accordingly, the student who is able to entertain only a single interpretation of an utterance may struggle in any attempt to follow a lecture or to comprehend text. On the other hand, the student who has the linguistic flexibility to consider alternative interpretations of an utterance – and then use content knowledge to settle on the right interpretation – will be able to keep up with the material being presented, either in lecture or written format.

The purpose of the study reported here was to assess the abilities of children with HL to provide multiple interpretations of a sentence that on its own is ambiguous. This skill is both more complex and more abstract than many of the vocabulary or morphosyntactic skills assessed in basic tests of language abilities. Results of this study revealed that children with HL who used CIs were indeed poorer than children with NH at providing multiple interpretations of these ambiguous sentences. Results for children with HL who used HAs, however, did not reveal statistically significant differences in this ability to provide multiple interpretations, between them and either of the other groups. Thus it can only be concluded that children with HAs performed intermediate to children in the other two groups on this task. Nonetheless, the deficit in the ability to provide multiple interpretations of ambiguous sentences that was so clearly observed in the children with CIs could help explain the reports of academic challenges observed for children with HL. Those reports either do not specify whether the children for whom data are being described had HAs or CIs (e.g., [16]) or only children with CIs were included in the study (e.g., [17]).

Looking across the three groups, one pattern that is revealed is that all children demonstrated more difficulty with sentences that were ambiguous due to syntactic structure, rather than due to any single word having multiple meanings. The magnitude of this ambiguity-related difference was similar across groups. This finding matches reports of other investigators (e.g., [33; 34])

Another outcome that was apparent in these data is the “inverted” pattern of underlying sources of variability found for different groups of children. Specifically, the explanations for within-group variability observed for children with NH were just what would be predicted: Variability in outcomes for the multiple meaning items was strongly dependent on vocabulary skills, and variability in outcomes for the syntactic structure items was strongly dependent on skill at comprehending syntactic structure. For children with HL, however, the pattern of relationship was reversed. This reversed pattern was most strongly observed for children with HAs, where the only predictor variable retained in stepwise regression for the multiple meaning items was children’s abilities to comprehend syntactic structure and the only predictor variable retained in stepwise regression for the syntactic structure items was vocabulary. For children with CIs, other predictor variables were found to explain additional unique variance, but the general pattern matched that found for the children with HAs. Thus, not only were the children with HL less capable of deriving multiple interpretations for an utterance, but the strategy they used for doing so was atypical. This finding could indicate additional difficulties for children with HL in their efforts to deal with the language of the classroom. These children could be putting more cognitive resources into unproductive efforts.

Another finding that was unique to the children with CIs was that for the syntactic structure items, working memory abilities helped to explain their scores. This outcome is likely related to the finding that children with CIs performed poorer than children in the other two groups on this measure.

The clinical implications of this study are clear: Children with HL need continued support for language learning as they progress through the school years, and some of that support must involve developing flexible comprehension strategies for complex language. Teachers and clinicians should incorporate training specifically in considering alternative interpretations for sentences that are ambiguous when presented on their own. Although one small component of what should be a broader intervention plan, this process could help these children acquire the needed flexibility for dealing with academic language.

Another component of these findings that should help direct future clinical efforts concerns the finding that both grammaticality judgments and working memory explained significant amounts of variance in scores for children with CIs. In past work both measures have been found to load highly on a phonological factor, indicating that these skills are themselves strongly affected by how well the children with CIs are able to recognize phonological structure in the speech signal. Consequently, any intervention strategies that would help children refine their sensitivity to phonological structure should improve both their awareness of grammatical structure and their working memory abilities. In turn, those heightened abilities could enhance their skills at recovering multiple interpretations for the sentences they hear or read in school.

5. Summary

Children born with HL in the recent past are leaving preschool programs with language skills that outpace anything imaginable in the days before newborn hearing screening and cochlear implants. These gains can give the impression that there is no more work left to do to help improve these children's language abilities. However, measures of academic achievement belie that impression. While children with HL may move through the early school years performing on par with their peers, academic achievement in later grades can reveal underlying weaknesses. The study reported here examined one potential source of those weaknesses by examining the hypothesis that these children may lack the linguistic flexibility needed to consider all possible interpretations of the language they hear or read in the school setting. Without that flexibility it can be difficult (i.e., require additional cognitive resources) to comprehend the material being presented. Here it was reported that children with HL, especially those with CIs, were impaired in this ability to recognize multiple interpretations for sentences.

Acknowledgments

This work was supported by Grant No. RO1 DC015992 from the National Institute on Deafness and Other Communication Disorders, the National Institutes of Health, to Susan Nittrouer. The authors thank Lauren Peterson, Olivia Coker, Chandler Burkhart, and Heather Smythe for their assistance with data collection and scoring.

Appendix A.

Educational and Occupational Ratings for Socio-Economic Status

These ratings were obtained by having 50 adults categorize each educational description or occupation. Mean ratings were used for this metric.

Educational Index

1.0=Completed elementary school

2.0=Completed junior high

2.5=Received General Education degree

3.0=Completed high school

3.5=Completed 1 or more years of technical/vocational school

4.0=Completed technical/vocational school

5.0=Completed 1 or more years of university/college

6.0=Bachelor's degree

6.5=Completed 1 or more years of graduate school

7.0=Master's degree

7.5=Course work completed for Ph.D., but no dissertation; Law degree without bar; Medical degree without internship completed

8.0=Ph.D.; Law degree with bar; Medical degree with internship completed

Occupational Index

1=maid, parking lot attendant, cafeteria worker, welfare recipient

2=fast food worker, meter reader, housekeeper, delivery man, garbage man, packer, housewife, bill collector, telemarketer, waiter/waitress (e.g., bars), butler, factory worker, taxi driver, telephone operator, assembly line worker, data entry, nanny, bartender, painter (e.g., house), dishwasher

3=daycare worker, construction worker, dispatcher, home appliance repairman, truck driver, bus driver, print room operator, gardener, machine operator, roofer, sales clerk, waiter/waitress (higher), brewer, camp counselor, dry cleaner, butcher, chef at a diner, exterminator, telephone company technician, mailman, car salesman, retail sales, military enlisted, post office clerks, welder, auto body repairman, bank teller/clerk, engraver, mechanic, beautician, service technician, janitor, carpet installer, brick mason, security guard, maintenance worker

4=barber, travel agent, proofreader, baker, plumber, insurance agent, farmer, florist, sales representative, court reporter, fast food manager, electrician, tailor, locksmith, jeweler, bookkeeper, undergraduate student, carpenter, corrections officer, piano teacher, loan officer, factory supervisor

5=advertising agent, actor/actress, construction foreman, librarian, interior decorating, real estate broker, missionary, funeral director, artist, laboratory technician, chef at a good restaurant, insurance adjustor, manufacturer, oral hygienist, musician, tavern owner, electrical contractor, L.P.N., public relations, social worker, executive assistant, office manager, radio/TV announcer, store manager (chain), executive secretary, personnel manager, accountant, contractor, graduate student, mortician, policeman, postmaster, fireman, medical technician, bank manager, firefighter

6=computer programmer, restaurant owner, store or small business owner, elementary school teacher, research assistant, book or magazine editor, optician, real estate developer, stock broker, high school teacher, military captain/lieutenant, chiropractor, registered nurse, military officer, lawyer, sheriff/police chief, clergyman, pharmacist, family therapist

7=mayor, symphony conductor, engineer, large business owner, school principal, architect, judge, psychologist, veterinarian, company president, university professor, dentist

8=university president, scientist, physician, surgeon

Appendix B.

Count of listeners for whom the raw score for the Ambiguous Sentences task improved when using the experimental ceiling of 10 consecutive incorrect items, rather than the published ceiling of five consecutive incorrect items, for each group separately. NH: Normal hearing. HA: Hearing aids. CI: Cochlear implants.

	NH N = 56	**HA N = 15**	**CI N = 51**
	N	N	N
No change	32	8	26
One item	9	4	7
2 to 5 items	6	2	12
6 to 10 items	7	0	6
More than 10 items	2	1	0

Appendix C

Means, medians, and SDs for three of the measures used as potential predictor variables, given as raw scores.

	NH			**HA**			**CI**		
	Mean	**Median**	**SD**	**Mean**	**Median**	**SD**	**Mean**	**Median**	**SD**
Expressive vocabulary	142	145	13.8	138	142	20.0	133	136	16.8
Sentence comprehension	17.6	18.0	2.6	17.3	28.0	3.2	15.8	17.0	4.1
Grammaticality judgments	68.7	68.0	9.3	64.0	65.0	15.4	58.2	59.0	14.8

References

[1]. Bradham TS, Fonnesbeck C, Toll A, Hecht BF. The Listening and Spoken Language Data Repository: Design and Project Overview. Lang Speech Hear Serv Sch 2018 1 9;49(1):108–20. [PubMed: 29222559]

[2]. Geers AE, Moog JS, Biedenstein J, Brenner C, Hayes H. Spoken language scores of children using cochlear implants compared to hearing age-mates at school entry. J Deaf Stud Deaf Educ 2009;14(3):371–85. [PubMed: 19155289]

[3]. Geers AE, Nicholas J, Tobey E, Davidson L. Persistent Language Delay Versus Late Language Emergence in Children With Early Cochlear Implantation. J Speech Lang Hear Res 2016 2;59(1):155–70. [PubMed: 26501740]

[4]. Nittrouer S Early development of children with hearing loss. San Diego, CA: Plural Publishing; 2010.

[5]. Tomblin JB, Oleson J, Ambrose SE, Walker EA, Moeller MP. Early Literacy Predictors and Second-Grade Outcomes in Children Who Are Hard of Hearing. Child Dev 2020 1;91(1):e179–e197. [PubMed: 30298910]

[6]. Boons T, De Raeve L, Langereis M, Peeraer L, Wouters J, van Wieringen A. Expressive vocabulary, morphology, syntax and narrative skills in profoundly deaf children after early cochlear implantation. Res Dev Disabil 2013 6;34(6):2008–22. [PubMed: 23584181]

[7]. Geers AE, Sedey AL. Language and verbal reasoning skills in adolescents with 10 or more years of cochlear implant experience. Ear Hear 2011 2;32(1 Suppl):39S–48S. [PubMed: 21832889]

[8]. Kronenberger WG, Xu H, Pisoni DB. Longitudinal Development of Executive Functioning and Spoken Language Skills in Preschool-Aged Children With Cochlear Implants. J Speech Lang Hear Res 2020 4 27;63(4):1128–47. [PubMed: 32204645]

[9]. Ruffin CV, Kronenberger WG, Colson BG, Henning SC, Pisoni DB. Long-term speech and language outcomes in prelingually deaf children, adolescents and young adults who received cochlear implants in childhood. Audiol Neurootol 2013;18(5):289–96. [PubMed: 23988907]

[10]. Schorr EA, Roth FP, Fox NA. A comparison of the speech and language skills of children with cochlear implants and children with normal hearing. Commun Disord Q 2008 8 12;29(4):195–210.

[11]. Antia SD, Lederberg AR, Easterbrooks S, Schick B, Branum-Martin L, Connor CM, Webb MY. Language and Reading Progress of Young Deaf and Hard-of-Hearing Children. J Deaf Stud Deaf Educ 2020 2 13;25(3):334–50. [PubMed: 32052022]

[12]. Fitzpatrick EM, Olds J, Gaboury I, McCrae R, Schramm D, Durieux-Smith A. Comparison of outcomes in children with hearing aids and cochlear implants. Cochlear Implants Int 2012 2;13(1):5–15. [PubMed: 22340747]

[13]. Johnson C, Goswami U. Phonological awareness, vocabulary, and reading in deaf children with cochlear implants. J Speech Lang Hear Res 2010 4;53(2):237–61. [PubMed: 20008682]

[14]. Tobey EA, Thal D, Niparko JK, Eisenberg LS, Quittner AL, Wang NY. Influence of implantation age on school-age language performance in pediatric cochlear implant users. Int J Audiol 2013 4;52(4):219–29. [PubMed: 23448124]

[15]. Kronenberger WG, Pisoni DB. Assessing Higher Order Language Processing in Long-Term Cochlear Implant Users. Am J Speech Lang Pathol 2019 11 19;28(4):1537–53. [PubMed: 31618055]

[16]. Marschark M, Shaver DM, Nagle KM, Newman LA. Predicting the Academic Achievement of Deaf and Hard-of-Hearing Students From Individual, Household, Communication, and Educational Factors. Except Child 2015 4;81(3):350–69. [PubMed: 26549890]

[17]. Sarant JZ, Harris DC, Bennet LA. Academic Outcomes for School-Aged Children With Severe-Profound Hearing Loss and Early Unilateral and Bilateral Cochlear Implants. J Speech Lang Hear Res 2015 6;58(3):1017–32. [PubMed: 25677804]

[18]. Harris M, Terlektsi E, Kyle FE. Literacy Outcomes for Primary School Children Who Are Deaf and Hard of Hearing: A Cohort Comparison Study. J Speech Lang Hear Res 2017 3 1;60(3):701–11. [PubMed: 28241207]

[19]. Garberoglio CL, Palmer JL, Cawthon S, Sales A. Deaf people and educational attainment in the United States: 2019. Washington, D.C.: U.S Department of Education, Office of Special Education Programs, National Deaf Center on Postsecondary Outcomes; 2019.

[20]. Baumann JF, Graves MF. What is academic vocabulary? J Adolesc Adult Lit 2010;54(1):4–12.

[21]. Coxhead A A new academic word list. TESOL Q 2000;34(2):213–38.

[22]. Dutro S, Morran C. Rethinking English language instruction: An architectural approach In: Garcia GG, editor. English learners: Reaching the highest level of English literacy.Newark, DE: International Reading Association; 2003 p. 227–58.

[23]. Snow CE. Academic language and the challenge of reading for learning about science. Science 2010 4 23;328(5977):450–2. [PubMed: 20413488]

[24]. Snow CE, Uccelli P. The Challenge of Academic Language In: Olson D, Torrance N, editors. The Cambridge Handbook of Literacy.New York: Cambridge University Press; 2009 p. 112–33.

[25]. Zwiers J Building academic language: Essential practices for content classrooms, grades 5–12. San Francisco, CA: Jossey-Bass; 2008.

[26]. Bailey AL, Butler FA, Stevens R, Lord C. Further specifying the language demands of school In: Bailey AL, editor. The Language Demands of School: Putting Academic English to the Test.New Haven, CT: Yale University Press; 2007 p. 103–56.

[27]. Gottlieb M, Ernst-Slavitt G. What is academic language? In: Gottlieb M, Ernst-Slavitt G, editors. Academic language in diverse classrooms: Definitions and contexts.Thousand Oaks, CA: SAGE Publications; 2014 p. 1–26.

[28]. Schleppegrell MJ. The language of schooling: A functional linguistics perspective. Mahwah, N.J.: Lawrence Erlbaum Associates, Inc.; 2004.

[29]. Nagy W, Townsend D. Words as tools: Learning academic vocabulary as language acquisition. Read Res Q 2012;47(1):91–108.

[30]. Uccelli P, Barr CD, Dobbs CL, Galloway EP, Meneses A, Sanchez E. Core academic language skills: An expanded operational construct and a novel instrument to chart school-relevant language proficiency in preadolescent and adolescent learners. Appl Psycholinguist 2015;36:1077–109.

[31]. Hyland K, Tse P. Is There an “Academic Vocabulary”? TESOL Q 2007;41(2):235–53.

[32]. Backscheider AG, Gelman SA. Children’s understanding of homonyms. J Child Lang 1995 2;22(1):107–27. [PubMed: 7759574]

[33]. Cairns HS, Waltzman D, Schlisselberg G. Detecting the Ambiguity of Sentences: Relationship to Early Reading Skill. Commun Disord Q 2004;25(2):68–78.

[34]. Shultz TR, Pilon R. Development of the Ability to Detect Linguistic Ambiguity. Child Dev 1973;44:728–33.

[35]. Hirsh-Pasek K, Gleitman LR, Gleitman H. What Did the Brain Say to the Mind? A Study of the Detection and Report of Ambiguity by Young Children In: Sinclair A, Jarvella RJ, Levelt WJM, editors. The Child’s Conception of Language.New York: Springer-Verlag; 1978 p. 97–132.

[36]. Spector CC. Linguistic humor comprehension of normal and language-impaired adolescents. J Speech Hear Disord 1990 August;55(3):533–41.

[37]. Johnson CJ, Ionson ME, Torreiter SM. Assessing Children’s Knowledge of Multiple Meaning Words. Am J Speech Lang Pathol 1997;6(1):77–86.

[38]. Carrow-Woolfolk E Comprehensive Assessment of Spoken Language (CASL). Bloomington, MN: Pearson Assessments; 1999.

[39]. Martin N, Brownell R. Expressive One-Word Picture Vocabulary Test (EOWPVT). 4th ed Novato, CA: Academic Therapy Publications, Inc.; 2011.

[40]. Nittrouer S, Burton LT. 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. J Commun Disord 2005 1;38(1):29–63. [PubMed: 15475013]

[41]. Hollingshead AB. The Two Factor Index of Social Position. New Haven: 1957.

[42]. Roid GH, Miller LJ. Leiter International Performance Scale – Revised (Leiter-R). Wood Dale, IL: Stoelting Co.; 2002.

[43]. Nilsson M, Soli SD, Sullivan JA. Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. J Acoust Soc Am 1994 2;95(2):1085–99. [PubMed: 8132902]

[44]. Carrow-Woolfolk E Comprehensive Assessment of Spoken Language-2 (CASL-2). Torrance, CA: WPS Publishing LLC; 2017.

[45]. Wechsler D Wechsler Intelligence Scale for Children. 3rd ed San Antonio, TX: The Psychological Corporation; 1991.

[46]. Nittrouer S, Muir M, Tietgens K, Moberly AC, Lowenstein JH. Development of phonological, lexical, and syntactic abilities in children with cochlear implants across the elementary grades. J Speech Lang Hear Res 2018 10;61(10):2561–77. [PubMed: 30242344]

[47]. Nittrouer S, Caldwell A, Holloman C. Measuring what matters: effectively predicting language and literacy in children with cochlear implants. Int J Pediatr Otorhinolaryngol 2012 8;76(8):1148–58. [PubMed: 22648088]

[48]. Nittrouer S, Caldwell-Tarr A. Language and literacy skills in children with cochlear implants: Past and present findings In: Young N, Kirk KI, editors. Pediatric Cochlear Implantation: Learning and the Brain.New York: Springer; 2016 p. 177–97.

Highlights

- Children with hearing loss often leave early intervention programs with language skills that appear close to those of children with normal hearing, but nonetheless encounter academic challenges at higher grade levels.
- This study tested the hypothesis that the language of school becomes more complex and disconnected from real-world context by examining the abilities of adolescents with normal hearing or with hearing loss to provide multiple interpretations for ambiguous sentences, a task akin to what is required in higher grades.
- Results supported the hypothesis, showing that although basic vocabulary and syntactic skills of the children with hearing loss were close to those of the children with normal hearing, their abilities to provide multiple interpretations for ambiguous sentences was deficient. This is an important skill for processing academic language.

Table 1.

Means, medians, and standard deviations (SD) for audiometric measures for children with hearing aids (HA) and children with cochlear implants (CI).

	HA N = 15			CI N = 51		
	Mean	**Median**	**SD**	**Mean**	**Median**	**SD**
Age of identification of hearing loss (months)	10	4	11	6	3	7
Unaided PTA threshold at the time of testing (HA) or pre-implant unaided PTA threshold (CI) (dB)	64	67	13	101	100	15
Aided PTA threshold at the time of testing	26	25	6	20	20	5
Age at first CI (months)				25	14	29
Age at second CI (months); N = 33				54	44	38

Note: PTA = pure tone average. PTA thresholds are given in dB hearing level for the three speech frequencies of 500, 1000, and 2000 Hz, for the better ear.

Table 2

Auditory prosthesis history for children with cochlear implants (CIs).

Children with some bimodal experience (25)
Received a second Cl in the ear using the HA more than a year after the first Cl (17)
Discontinued use of the HA, without receiving a second Cl in that ear (5)
Continued use of the HA up through eighth grade (3)
Children with electric only experience (26)
Received bilateral CIs simultaneously (5)
Received a second Cl more than a year after receiving the first Cl (11)
Continued use of a single Cl up through eighth grade (10)

Note: Bimodal experience refers to the use of a cochlear implant (Cl) and hearing aid (HA) on different ears at the same time. Electric-only experience refers to the use solely of CIs, either single-sided or bilaterally.

Table 3.

Mean demographic variables (and standard deviations, *SD*s) for each group. NH: Normal hearing. HA: Hearing aids. CI: Cochlear implants.

	NH N = 56		**HA N = 15**		**CI N = 51**	
	M	***SD***	**M**	***SD***	**M**	***SD***
Age (years; months)	14;5	*0;6*	14;8	*0;5*	14;8	*0;5*
Socioeconomic status	37	*14*	32	*14*	33	*11*
Brief IQ	106	*13*	103	*9*	103	*14*
Speech recognition in quiet	99.8	*0.6*	98.9	*2.3*	98.4	*2.7*

Note: Brief IQ is from the Leiter International Performance Scale – Revised, and standard scores are shown. Socioeconomic status is on a 64-point scale. Speech recognition in quiet is percent words repeated correctly.

Table 4

Measures collected in study and used in analyses

Ambiguous Sentences, Comprehensive Assessment of Spoken Language (43 items)	
1	Standard scores for all items, using published ceiling of 5 consecutive incorrect items
2	Raw scores for all items, using experiment-specific ceiling of 10 consecutive incorrect items
3	Raw scores for ambiguous items due to multiple word meanings, using experiment-specific ceiling of 10 consecutive incorrect items (22 items)
4	Raw scores for ambiguous items due to syntactic structure, using experiment-specific ceiling of 10 consecutive incorrect items (21 items)
Potential predictor measures	
1	Expressive One-Word Picture Vocabulary Test
2	Sentence Comprehension of Syntax, Comprehensive Assessment of Spoken Language (21 items)
3	Grammaticality Judgment3, Comprehensive Assessment of Spoken Language (57 items)
4	Forward Digit Span, Wechsler Intelligence Scale for Children

Table 5

Means, medians, and SDs for the four measures of performance on the Ambiguous Sentences test.

	NH			HA			CI		
	Mean	**Median**	**SD**	**Mean**	**Median**	**SD**	**Mean**	**Median**	**SD**
Standard scores, all items	101	100	16.3	97	99	18.6	91	91	14.7
Raw scores, all items	21.0	23.0	9.6	18.1	17.0	10.6	14.0	13.0	9.0
Multiple-meanings items	13.2	14.0	4.8	11.8	13.0	5.9	9.6	11.0	5.1
Syntactic-structure items	7.8	7.5	5.8	6.3	4.0	6.4	4.5	2.0	5.0

Note: Standard scores were obtained using a ceiling rule of five consecutive incorrect items; all other scores were obtained using a ceiling rule of 10 consecutive incorrect items.

Table 6.

Outcomes of one-way analyses of variance performed on each of the measures from the Ambiguous Sentences test, with Bonferroni *p* values for post hoc comparisons between listeners with NH and CIs. NH: Normal hearing. HA: Hearing aids. CI: Cochlear implants.

	F	***p***	η^2	**NH vs CI**	**Cohen's *d***
Standard scores, all items	6.148	.003	.094	.002	.70
Raw scores, all items	7.239	.001	.108	.001	.75
Multiple-meanings items	7.097	.001	.107	.001	.73
Syntactic-structure items	4.773	.010	.074	.008	.61

Note: Degrees of freedom are 2,119 for all analyses. There were no significant differences between children with NH and those with HAs, or between children with HAs and those with CIs.

Table 7

Means, medians, and SDs for the four measures used as potential predictor variables. Standard scores are shown for expressive vocabulary, sentence comprehension, and grammaticality judgments. Span length is shown for forward digit span.

	NH			HA			CI		
	Mean	**Median**	**SD**	**Mean**	**Median**	**SD**	**Mean**	**Median**	**SD**
Expressive vocabulary	113	114	15.7	108	109	21.0	102	103	17.3
Sentence comprehension	107	106	12.9	106	108	15.2	100	101	16.1
Grammaticality judgments	100	99	11.3	94	97	16.8	88	89	15.8
Forward digit span	6.5	6.5	1.3	6.5	7.0	1.4	5.6	6.0	1.0

Table 8.

Outcomes of one-way analyses of variance performed on each of the other measures, with Bonferroni *p* values for post hoc comparisons between listeners with NH and CIs. NH: Normal hearing. HA: Hearing aids. CI: Cochlear implants.

	F	***p***	**η^2**	**NH vs CI**	**Cohen's *d***
Expressive vocabulary	4.590	.012	.072	.009	.61
Sentence comprehension	4.006	.021	.063	.019	.53
Grammaticality judgments	9.206	<.001	.134	<.001	.85
Forward digit span	10.281	<.001	.147	<.001	.85

Note: Degrees of freedom are 2,119 for all analyses. There were no significant differences between children with NH and those with HAs. The only significant difference in scores for children with HA and those with CIs was for digit span, $p = .015$ and Cohen's $d = .82$.

Table 9.

Pearson product-moment correlation coefficients between scores on the Ambiguous Sentences test – Multiple Meanings and each of the other measures, for each group separately. NH: Normal hearing. HA: Hearing aids. CI: Cochlear implants.

	NH		**HA**		**CI**	
	r	***p***	***r***	***p***	***r***	***p***
Expressive vocabulary	.564	<.001	.675	.006	.495	<.001
Sentence comprehension	.524	<.001	.843	<.001	.505	<.001
Grammaticality judgments	.618	<.001	.752	.001	.542	<.001
Forward digit span	.393	.003	.648	.009	.302	.031

Table 10.

Stepwise regression outcomes for scores on the Ambiguous Sentences test – Multiple Meanings, for each group separately. NH: Normal hearing. HA: Hearing aids. CI: Cochlear implants.

	NH	**HA**	**CI**
Step #1	GJ, $\beta = .618$, $p < .001$ $R^2 = .382$	SC, $\beta = .843$, $p < .001$ $R^2 = .711$	GJ, $\beta = .542$, $p < .001$ $R^2 = .294$
Step #2	GJ, $\beta = .439$, $p = .001$		GJ, $\beta = .376$, $p = .011$
	Voc, $\beta = .312$, $p = .016$ $R^2 = .447$		SC, $\beta = .289$, $p = .048$ $R^2 = .350$

Voc = Expressive vocabulary; SC = Sentence comprehension; GJ = Grammaticality judgments; DS = Forward digit span

Table 11.

Pearson product-moment correlations between scores on the Ambiguous Sentences test – Syntactic Structure and each of the other measures, for each group separately. NH: Normal hearing. HA: Hearing aids. CI: Cochlear implants.

	NH		**HA**		**CI**	
	r	***p***	***r***	***p***	***r***	***p***
Expressive vocabulary	.469	<.001	.760	.001	.636	<.001
Sentence comprehension	.570	<.001	.753	.001	.566	<.001
Grammaticality judgments	.591	<.001	.614	.015	.655	<.001
Forward digit span	.400	.002	.447	NS	.589	<.001

Note: NS = not significant; $p > .10$.

Table 12.

Stepwise regression outcomes for scores on the Ambiguous Sentences test – Syntactic Structure, for each group separately. NH: Normal hearing. HA: Hearing aids. CI: Cochlear implants.

	NH	HA	CI
Step #1	GJ, $\beta = .591$, $p < .001$ $R^2 = .349$	Voc, $\beta = .760$, $p = .001$ $R^2 = .577$	GJ, $\beta = .655$, $p < .001$ $R^2 = .429$
Step #2	GJ, $\beta = .392$, $p = .003$ SC, $\beta = .344$, $p = .009$ $R^2 = .427$		GJ, $\beta = .480$, $p < .001$ DS, $\beta = .346$, $p = .004$ $R^2 = .518$
Step #3			GJ, $\beta = .315$, $p = .018$ DS, $\beta = .291$, $p = .012$ Voc, $\beta = .313$, $p = .014$ $R^2 = .576$

Note: Voc = Expressive vocabulary; SC = Sentence comprehension; GJ = Grammaticality judgments; DS = Forward digit span