ABSTRACT

Title of thesis: SYNTACTIC PROCESSING AND WORD LEARNING WITH A DEGRADED AUDITORY SIGNAL

Isabel A. Martin, Master of Arts, 2017

Thesis directed by: Professor Yi Ting Huang
Department of Hearing and Speech Sciences

The current study examined real-time processing and word learning in children receiving a degraded audio signal, similar to the signal children with cochlear implants hear. Using noise-vocoded stimuli, this study assessed whether increased uncertainty in the audio signal alters the developmental strategies available for word learning via syntactic cues. Normal-hearing children receiving a degraded signal were found to be able to differentiate between active and passive sentences nearly as well as those hearing natural speech. However, they had the most difficulty when correct interpretation of a sentence required revision of initial misinterpretations. This pattern is similar to that found with natural speech. While further testing is needed to confirm these effects, the current evidence suggests that a degraded signal may make revision even harder than it is in natural speech. This provides important information about language learning with a cochlear implant, with implications for intervention strategies.

SYNTACTIC PROCESSING AND WORD LEARNING WITH A DEGRADED AUDITORY SIGNAL

by

Isabel A. Martin

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Master of Arts
2017

Advisory Committee:
Professor Yi Ting Huang, Chair
Professor Matthew Goupell
Professor Jared Novick

ACKNOWLEDGMENTS

I would like to first thank Dr. Yi Ting Huang, my committee chair, for guiding me through the steps of this process and for her infectious excitement about the subject matter. I would also like to thank Dr. Matthew Goupell and Dr. Jared Novick for their participation on my committee and for being open to recruiting friends and neighbors to participate in the study. Thanks also to Dr. Goupell for creating the noise-vocoded versions of the stimuli for this project. This study would not have been possible without the cooperation of the staff, parents, and children at Bethesda Montessori School, Celebration Christian Academy, Child's Garden Center, Crestview Montessori School, Greenwood School, Paint Branch Montessori School, and the UMD Center for Young Children, as well as Kathy Dow-Burger who helped to recruit community participants. Lastly, thank you to Language and Cognition Lab members Elinora Leonard, for being a fantastic partner in school recruitment, Zoe Ovans, for help with experimental design, and Allison Granados and Allesondra Sanchez, for help with data collection.

TABLE OF CONTENTS

Acknowledgements ii
Table of Contents iii
List of Figures iv
1. Introduction 1
    1.1 The role of executive functioning in language development and implications for learning with CIs 3
    1.2 Possible effects of degraded speech on word learning 5
    1.3 The utility of vocoded speech 7
    1.4 The current study 10
2. Methods 12
    2.1 Participants 12
    2.2 Procedure & Materials 13
        2.2.1 Noise-vocoded speech training task 14
        2.2.2 Word-learning task 17
    2.3 Coding 20
3. Results 20
    3.1 Mouse click: Accuracy 21
    3.2 Mouse click: Agent preference 25
    3.3 Real-time comprehension 27
4. Discussion 30
    4.1 Does an acoustically degraded signal make all sentence processing harder? 30
    4.2 Does an acoustically degraded signal increase the agent-first bias? 32
    4.3 Does an acoustically degraded signal only affect processing demands when syntactic revision is necessary? 33
    4.4 Further discussion 34
5. Conclusion 37
Appendix A 38
Appendix B 39
References 42

LIST OF FIGURES

Figure 1. Sample trial for Noise-vocoded speech training task 16
Figure 2. Sample trial for Word-learning task 18
Figure 3. In the vocoded speech condition, accuracy of click actions across all sentence constructions 22
Figure 4. Proportion of correct actions following sentence completion 24
Figure 5. Agent preference shown as the proportion of agent actions following sentence completion... 27
Figure 7. Natural speech condition. Fixation preference score.................................. 28
Figure 8. Natural vocoded condition. Fixation preference score 29

1. Introduction

Each year, approximately two of every 1,000 babies born in the United States are born with hearing loss, and approximately 90% of those children are born to hearing parents (National Institutes of Health, 2016). An important early decision these parents must make is what, if any, amplification technologies they will implement for their child and what modality(s) they will use for their child's communication and language development. Since cochlear implants (CIs) were first approved by the U.S. Food and Drug Administration in 1984, their use by children has grown steadily, and the technology has advanced considerably (NIDCD). It is estimated that now 90-95% of children who are born with severe-to-profound hearing loss receive one or two CIs (Komesaroff, 2007), and as of December 2012, approximately 38,000 CIs have been implanted in children in the United States (National Institutes of Health, 2016).

CIs give children access to sound and spoken language that they would likely not have without them, but the auditory signal received through CIs does not sound like natural speech. Incoming sound waves are converted into electrical currents to be delivered by the CIs, and through this process most of the pitch information is lost, resulting in a degraded auditory signal. Despite receiving this degraded input, children are able to extract semantic and syntactic meaning from their linguistic environment, and they acquire receptive and expressive language in much the same sequence as normal-hearing (NH) children (Duchesne, 2016). CIs lead to better and quicker spoken language improvement than would be predicted from rates of pre-implantation learning (Niparko et al., 2010). For the most part, children with CIs function well in mainstream classrooms, and appear to catch up to their NH peers (Ali & O'Connell, 2007).

However, while numerous studies have examined speech production/perception and language/literacy skills of children with CIs, the results are overall mixed and are highly variable both between individuals and across language domains (Nicholas & Geers, 2007; Nicholas & Geers, 2008; Peterson, Pisoni, Miyamoto, 2010; Sarant, Blamey, Dowell, Clark, & Gibson, 2001; Svirsky, Stallings, Lento, Ying, & Leonard, 2002; Svirsky, Teoh, & Neuberger, 2004; Szagun & Schramm, 2016). Even if CI users' overall language scores are within the average range, they are often in the low end of that range and significantly lower than NH peers in areas including structural language (e.g., vocabulary, morphology, and syntax) and metalinguistic skills (e.g., phonemic awareness, elision, auditory memory, idioms, lexical ambiguity) (Schorr, Roth, & Fox, 2008). Moreover, beyond standardized language tests, children with CIs also perform below their NH peers in academic areas such as math and reading (Sarant, Harris, & Bennet, 2015). Together, these patterns suggest that while standardized assessments may reveal the current state of children's language acquisition, questions remain about how these abilities emerged in the first place and what impacts they will have on future development.

The current study will address these questions by examining how a degraded auditory signal impacts the real-time mechanisms of language processing in children. Due to the spectrally degraded signal they receive through their devices, children with CIs may be at a language-learning disadvantage. Caregiver speech comes quickly and continuously over time. Since children use this input to learn vocabulary and sentence structures of their native language, they must develop strategies and skills to process language in real time. Understanding how the properties of an acoustic signal affect the kinds of strategies children employ when learning language may have important implications for methods of targeted instruction for children using CIs. NH children can incidentally pick up on most of the structures they need to understand, and use most syntactic forms without specific instruction. Children with CIs may seem like they are adept at using a variety of syntactic forms in casual conversation, but more direct targeting may be needed to teach them to use similar strategies as their NH peers and to be more efficient language learners. In the remainder of this introduction, I will explore the relationship between executive functioning skills and language development, particularly in individuals with CIs. Next, I will describe the use of syntactic bootstrapping for word learning in NH children and how that may be affected by a degraded speech signal. I will also briefly review literature on the relevance of using noise-vocoded stimuli to simulate the input received through CIs.

1.1 The role of executive functioning in language development and implications for learning with CIs

Language development in NH children is a complex process that takes place over multiple years. Still-developing executive functioning (EF) skills, such as memory, attention, sequential processing, novel problem solving, and conceptual learning, may have an impact on language learning. One late-developing aspect of EF that may be particularly important for language development is inhibition (Mazuka, Jincho, & Onishi, 2009). Up to around the ages of four to eight years, NH children have difficulty with inhibition in linguistic and non-linguistic tasks (Frye, Zelazo, & Palfai, 1995; Müller, Zelazo, Hood, Leone, & Rohrer, 2004; Jones, Rothbart, & Posner, 2003; Snedeker & Yuan, 2008). For example, when presented with the sentence *Put the frog on the napkin* *into the box*, and a display of objects including a napkin, a box, and two frogs – one on a napkin and one standing alone – both children and adults initially look at the empty napkin when they first hear the prepositional phrase “on the napkin.” But adults reinterpret the sentence when they hear the second prepositional phrase “into the box,” whereas children largely do not. Children fail to revise their initial interpretations 60% of the time, and instead place the stand-alone frog on the napkin and then in the box (Trueswell, Sekerina, Hill, & Logrip, 1999). As they get older, children’s abilities become more adult-like. Eight- and eleven-year-old children are able to inhibit initial sentence interpretation on the same task and perform more like adults, more often putting the frog that is on the napkin into the box (Weighall, 2008).

Understanding the interactions of language development and EF in NH children is an important foundation for looking at EF in CI users and allows for comparisons between the two groups. In CI users, difficulties with speech perception may have cascaded impacts on language processing and EF. Kronenberger and colleagues (2014) found that children with CIs were two to five times more likely to have problems with at least one area of EF. CI users may also have reduced attention to speech (Houston & Bergeson, 2014) and shorter working memory spans (Pisoni & Cleary, 2003). Since EF skills interact with language in a bidirectional manner (Castellanos, Pisoni, Kronenberger, & Beer, 2016), access to auditory input may be an important factor to develop EF, in addition to EF aiding in language learning. Children with hearing loss using total communication have poorer digit span performance than those in oral-only environments, presumably because those using only spoken language had more auditory information from an earlier age (Pisoni & Cleary, 2003). It may be that early auditory deprivation contributes to these changes, delaying both the development of language and other domain-general skills.

1.2 Possible effects of degraded speech on word learning

Difficulties with perceiving and processing language may create specific challenges for word learning. Children with CIs have been found to have a similar, or even faster, rate of lexical development as their NH hearing-age-matched peers in the months directly following implantation (Bollard, Popp, Chute, & Parisier, 1999; Svirsky, Chute, Green, Bollard, & Miyamoto, 2000). This corresponds to a similar period of rapid vocabulary development in NH toddlers, particularly for children who received their implants early. In fact, lexical development has been found to correspond to age of implantation (Schorr, Roth, & Fox, 2008). There is some evidence that as many as five years post-implantation, lexical-semantic skills are a strength compared to syntactic and morphologic skills for children with CIs (Young & Killen, 2002). But even though they seem to develop their vocabulary quickly, children with CIs show longer word recognition latencies (Grieco-Calub, Saffran, & Litovsky, 2009), and syntactic skills lag behind NH peers (Young & Killen, 2002). When lexical and grammatical development overlap, difficulties could result for the child with CIs.

In particular, it is well known that NH children use a variety of tools available to them from their language input to learn the meanings of words. One key strategy they employ is syntactic bootstrapping, in which they use the structure of a sentence to inform semantic mapping of new words (Landau & Gleitman, 1985; Gleitman, 1990; Gleitman & Gleitman, 1992). They can employ these strategies from a young age. NH children as young as two-years-old have been shown to understand the argument structure of verbs to differentiate between transitive and intransitive constructions, such as *Big Bird is flexing with Cookie Monster* and *Big Bird is flexing Cookie Monster* (Gleitman, 1990). And three- to five-year-old NH children can distinguish between novel mass and count nouns depending on the accompanying syntax (e.g., *a* blick vs. *some* blick; Brown, 1957). In order to make use of syntactic bootstrapping, a few conditions must be met. First, the child must be able to hear the syntactic cues (e.g., "…is flexing with…" is different from "…is flexing…"). Second, he must have an underlying ability to understand the relevant syntactic structure (e.g., "…is flexing with…" implies doing something together) and use those cues to infer word meaning (e.g., "…is flexing with…" must mean something you can do yourself and "…is flexing…" is something done to another person).

Interestingly, research has shown that even when these conditions are met and NH children can use syntactic bootstrapping with a particular structure, they are less reliable at using this strategy when they encounter uncertainty in their input. French-speaking, NH toddlers use prosody and syntax to correctly interpret right dislocations ("il mange, le lapin" *he eats, the rabbit*, meaning "the rabbit eats") when familiar words are used in the test sentences. Importantly, though, when presented with novel, nonsense verbs (e.g., "daser"), the children fall back on canonical mapping. They interpret "*he dased, the baby*" (baby as agent) as "*he dased the baby*" (baby as theme), despite prosodic cues to the correct interpretation (Dautriche et al., 2014). This demonstrates that children use different strategies for understanding sentences depending on the overall context. It seems that with a given entity there is only a small cognitive cost to store this information and delay assignment of a syntactic role. In fact, greater uncertainty, in the form of unknown words, syntactic errors, or implausible utterances, has been previously shown to increase canonical interpretations in multiple sentence structures in NH subjects (Ferreira, 2003; Levy, Bicknell, Slattery, & Rayner, 2009; Gibson, Bergen, & Piantadosi, 2013; Huang, Abadie, Arnold, & Hollister, 2016).

In summary, for NH children, syntactic bootstrapping is an important tool for learning language, but it is not a flawless strategy. Difficulties with auditory perception through CIs and processing differences linked to EF skills could compound these factors, making syntactic bootstrapping less available to CI users. Even in the best conditions, syntactic and morphological markers are often short and unstressed, and may be more difficult to perceive in acoustically degraded input. Challenges with EF skills may affect syntactic revision, which relies on short-term memory and inhibition, and impact CI users' ability to fast map using syntactic information. A degraded signal may also lead to general uncertainty of what has been said, adding to cognitive load. These factors may force children receiving degraded input to rely even more heavily on canonical word order.

1.3 The utility of vocoded speech

To determine the role that the uncertainty of the sound processed through CI technology may have on language learning, the current study explores children's ability to use syntactic bootstrapping with a degraded auditory signal. However, one challenge to investigating this question is that CI users have varying etiologies, experience with the device (i.e., duration of time since implanted), amount of language exposure pre- and post-implantation, and education. Even children who have been pre-lingually implanted differ along a number of factors that one would ideally control for, such as age at implantation, years of hearing experience, unilateral vs. bilateral implantation, and education history. Additionally, recruitment of CI users within a geographical area in a limited amount of time presents significant challenges for the feasibility of such an approach.

In order to combat these difficulties and isolate the effects of degraded input, researchers have turned to providing NH participants with noise-vocoded speech stimuli (Dorman, Loizou, Kemp, & Kirk 2000; Faulkner, Rosen, & Smith, 2000; Sheldon, Pichora-Fuller, & Schneider, 2008; Obleser, Meyer, & Friederici, 2011; van Heugten, Volkova, Trehub, & Schellenberg, 2013; Vongpaisal, Trehub, Schellenberg, & van Lieshout, 2012; Warner-Czyz, Houston, & Hynan, 2014). Noise-vocoded speech is a processed auditory speech signal that can be used to roughly simulate the input received through a CI. It is created by dividing the speech signal into separate analysis bands, determining the amplitude envelope for each analysis band, modulating a narrowband noise with the same frequency limits as the analysis band with the extracted amplitude envelope (i.e., create modulated synthesis bands), and recombining the synthesis bands. This creates a signal whose temporal envelope is similar to the original but whose spectral envelope has been degraded. NH adults have been shown to reliably comprehend noise-vocoded speech with training with as few as three to five channels (Loizou, Dorman, & Tu, 1999; Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995). Even NH children as young as two-years-old can extract meaning from noise-vocoded speech, though their performance is above chance only at eight channels or more (Newman & Chatterjee, 2013).

However, comprehending a degraded auditory signal does not come without its difficulties. The effort needed to understand noise-vocoded speech requires more attention than non-degraded signals and may tie up cognitive functioning in other domains as well. While NH subjects can recall natural speech even if they are not paying attention to it, degraded speech is processed much better if the subject actively attends to it (Wild et al., 2012; Huyck & Johnsrude, 2012). Similarly, neuroimaging studies demonstrate that processing noise-vocoded speech recruits high-order cortical areas throughout the left hemisphere, including areas important for memory and attention, rather than solely the auditory cortex (Obleser, Wise, Dresner, & Scott, 2007; Obleser, Wostmann, Hellbernd, Wilsch, & Maess, 2012). NH children also experience a decreased ability to discriminate speaker characteristics in vocoded speech when short-term memory demands are high (Roman, 2015).

Measures of pupil dilation that track listening effort indicate that even when speech intelligibility is high, degraded audio (in the form of vocoded speech or through a CI processor) increases cognitive effort exerted during listening tasks (Winn, 2016). Effort continues to increase as degradation worsens (Winn, Edwards, & Litovsky, 2015). Previous knowledge of content may decrease effort when listening to vocoded speech, seen in reduced cognitive activation when top-down information is available (Sohoglu, Peele, Carlyon, & Davis, 2012). Though, while known context does decrease effort when listening to vocoded speech, the effort reduction occurs later than it does when listening to natural speech (i.e., after a sentence rather than during the sentence) (Winn, 2016). These examples all point to effort being a significant factor affecting abilities to perceive and comprehend vocoded speech, particularly as part of an ongoing speech stream, where silent processing time after a sentence may be non-existent.

In the case of syntactic analysis, degraded input could be expected to make processing even more difficult. Obleser et al. (2011) found that brain activity was significantly affected by the interaction between syntactic complexity and quality of the audio signal in NH adult German speakers. As the experimenters increased syntactic complexity, more "upstream" areas were engaged in performing abstract processing. When the signal was also degraded, there was a greater demand for acoustic analysis, which diminished the upstream processing. In a study involving adult CI users, Hahne et al. (2012) found that syntactic repair and violation effects were more vulnerable than semantic integration. Their strategy for auditory comprehension appears to rely on using semantic compensation to make up for syntactic violations.

1.4 The current study

The current study explores the degree to which comprehending degraded audio affects incremental syntactic parsing or late-arriving syntactic revision in NH children. To address this question, this study follows prior work by Huang and Arnold (2016) on the effects of syntactic revision on word learning in active and passive sentences in NH children and adults, and extends it to investigate degraded input processing by NH children. In the original study by Huang & Arnold, NH adults and five-year-old children accurately interpreted active sentences like 1(a), choosing correct referents for the novel words (e.g., a large, scary-looking creature likely to be eating a seal). Both age groups also reliably chose the correct referent in active and passive sentences when the first noun phrase (NP) in the sentence was a known entity (2(a) and (b)).

(1) a. Active Novel NP1: The *blicket* will be quickly eating the seal.
b. Passive Novel NP1: The *blicket* will be quickly eaten by the seal.

(2) a. Active Novel NP2: The seal will be quickly eating the *blicket*.
b. Passive Novel NP2: The seal will be quickly eaten by the *blicket*.

Importantly, passive constructions with a novel first noun phrase (NP1) tell a different story. Eye tracking showed that when NH adults and children heard sentences such as 1(b), they first interpreted “blicket” as the agent of the sentence. Adults revised their misinterpretation after they heard the disambiguating passive syntactic cues and settled on the correct meaning of “blicket” as theme. Children, though, were not able to revise initial misinterpretations as the adults had. They exhibited an agent-first bias, assuming the first NP they encountered was the agent of the sentence, as is true in canonical subject-verb-object word order. In sentences with a novel word in the second NP position (NP2), children and adults were able to delay assigning a role to NP1 when it was a familiar word, thus holding off the agent-first bias and using syntactic bootstrapping to correctly interpret the passive sentence after the disambiguating passive markers.

Using vocoded stimuli with NH children, this experiment will illustrate how syntactic bootstrapping may differ in children with CIs. One possibility is that an acoustically degraded signal makes all sentence processing harder. Prior work has shown that comprehending vocoded speech is more difficult than comprehending natural speech (Eisenberg, Shannon, Martinez, Wygonski, & Boothroyd, 2000). If this holds true in the current study, accuracy of correct responses will be depressed for all sentence constructions compared to the natural speech condition in Huang and Arnold. Another possibility is that a degraded auditory signal affects overall processing, increasing the agent-first bias. This outcome is likely if the degraded signal contributes to overall uncertainty, which adds to the cognitive load so that children are unable to inhibit the agent-first bias in any construction and instead rely on canonical word order interpretation. If this is the case, children in the current study will interpret all NP1s as agents. They will perform better with actives than passives in both sentences (1) and (2).

A final possibility is that a degraded auditory signal affects processing demands only when syntactic revision is necessary. This is plausible because increasing syntactic complexity has been shown to diminish the amount of top-down processing used in sentence comprehension (Obleser et al., 2011) and syntactic repair has been found to be vulnerable in adult CI users (Hahne et al., 2012). If this is true of the current study, the children should perform better with actives than passives in sentence (1), but show little difference in sentence (2). Their accuracy may be lower than the natural speech condition in all four sentence constructions because of the difficulty of processing vocoded speech. However, there will be a greater difference between performance on active and passive novel-NP1 sentences (1(a) and (b)) than there was with natural speech input.

2. Methods

2.1 Participants

Participants were 47 pre-K and Kindergarten children recruited from the greater Washington, DC, Metro area through private schools, the University of Maryland Infant & Child Studies Database, and the local community. Six subjects were excluded because of inability to achieve eye-tracking fixations following the vocoded speech training tasks. One subject was excluded because of inadequate performance on the training tasks. Mean age of the remaining 40 subjects was 5;4 (SD = 0;4, range = 4;10 to 5;11). The mean age of the children tested by Huang and Arnold in the natural speech condition was 5;5 (SD=0;3, range=5;0 to 5;11). This age group was originally chosen because children at this age produce passives but still have some difficulty comprehending them, particularly when revision is necessary. Choosing participants for the current study who are within the same age range allowed for comparison between the current data and that from the natural speech condition in Huang and Arnold. Participants had hearing within normal limits, according to parental report.

2.2 Procedure & Materials

The procedure for the current study largely follows that of the word-learning task from Huang and Arnold. One important change was adding a vocoded speech training task preceding the presentation of the word-learning task. Since vocoded speech sounds unlike a natural speech recording, a training task ensured that the children had some previous exposure to this signal before hearing the sentences in the critical trials. The current study is interested in processing strategies when the words are understood, not whether or not the children can perceive the words in the first place. Including a training task ensured that the children were able to decipher most of the words and any effects were a result of different processing strategies.

All audio stimuli were recorded in a noise-reducing sound booth by an adult female speaker using a clear tone at a slow, natural pace. Stimuli for the word-learning task were the same recordings used in Huang and Arnold. Vocoded speech training stimuli were recorded for this experiment. The natural speech recordings were passed through second-order Butterworth bandpass filters, using forward-backward filtering, into eight channels. Corner frequencies were logarithmically spaced between 200 and 8000 Hz. For each channel, the envelope was extracted using the Hilbert envelope transform, with a third-order low-pass Butterworth filter at 400 Hz. The envelopes were then used to modulate a narrowband noise and combined into a single waveform. The vocoded stimulus had the same overall energy as the natural speech stimulus.

2.2.1 Noise-vocoded speech training task

Children were told that they would be listening to a "robot," and would first hear some sentences so they could get used to how the robot's voice sounded. The initial familiarization phase of this task consisted of short phrases and active sentences (see Appendix A). Each sentence was presented three times: first the noise-vocoded recording, then the natural speech recording, then the noise-vocoded recording again. This pattern of presentation provides lexical feedback for the vocoded speech, which may enhance understanding of words when training with vocoded sentences (Davis, Johnsrude, Hervais-Adelman, Taylor, & McGettigan, 2005). During audio presentation of the sentences, a drawing depicting each sentence was shown on a computer screen to further reinforce the meaning and encourage attentiveness. The familiar nouns from the word-learning task were used as part of the training task in order to make them maximally familiar to participants in the critical trials. No nonsense words were used in these sentences. Twelve sentences were grouped into six pairs of roughly equal grammatical complexity.

Before beginning audio playback on each sentence of the familiarization phase, the experimenter drew the child's attention to the picture and discussed it together to give the child context for what the vocoded sentence would be. In a typical exchange, the experimenter would ask the child what animal or person he saw on the screen (e.g., a mouse, a girl) and what they were doing (e.g., eating cheese, catching a ball). Each participant heard the first sentence in the list (see Appendix A). The experimenter then dynamically determined whether the child should hear the next sentence in that pair for more exposure or could move on to the first item of the next pair of sentences. This was decided based on the child's spontaneous reactions to the vocoded speech and comments about difficulty understanding, as well as answers to direct questions such as "Is it hard to understand the robot?" The sentences were presented in the same order to each child, with approximately two seconds between the natural speech and vocoded speech versions of each sentence.

The children were then presented with a test phase, comprised of an additional set of six pairs of sentences accompanied by images. The noise-vocoded recording of a sentence played immediately on presentation of an image, with no experimenter dialogue. The child was asked to repeat the sentence verbatim, after which a natural speech recording of the same sentence was played as an accuracy check. The child's progression through the sentences was similar to that in the familiarization phase described above. If the child repeated the sentence incorrectly, the second sentence in a pair was played. If the child repeated the sentence correctly, the experimenter moved on to the next pair. Some repetition mistakes were not considered significant enough to necessitate the second sentence in a pair, such as "the" for "a." Mean number of test sentences presented to each child was 7 out of a possible 12 sentences (SD = .99, range = 6 to 10). The whole vocoded speech training task took approximately ten minutes, and children appeared to be comfortable listening to and understanding the vocoded recordings by the end of the task.

Display	Sentence
	<u>Familiarization phase</u> *"The boy is in the garden."* VS *"The boy is in the garden."* NS *"The boy is in the garden."* VS
• • •	x 6-12 training sentences
	<u>Test phase</u> *"The polar bear is standing on the ice."* VS (child's repetition) *"The polar bear is standing on the ice."* NS
• • •	x 6-12 test sentences

Figure 1. Sample trial for Noise-vocoded speech training task. Natural speech (NS); Vocoded speech (VS).

Prior to testing with children, pilot testing with eight NH adults was conducted to determine if the vocoded speech signal should include 8, 16, or 32 channels. The adults had little difficulty comprehending the eight-channel noise-vocoded speech during training trials, so it was determined that eight-channel vocoding was sufficient for this task. This is consistent with prior research showing that speech perception in adult CI users improves up to 7 to 10 electrodes (similar to the bands in vocoded speech), but there is little benefit to increasing the number of electrodes beyond 10 (Friesen, Shannon, Baskent, & Wang, 2001). Additionally, during pilot testing, adults were able to comprehend the vocoded speech after few training trials, without needing the full 24 sentences listed in Appendix A. This led to the dynamic nature of the training task described here. A novel-word familiarization task was also originally proposed to prevent the children from mishearing novel, noise-vocoded words as known words in their lexicon. However, pilot testing revealed no confusion upon hearing the novel words for the first time in vocoded sentences, so this task was omitted from the experimental design.

2.2.2 Word-learning task

Participants were seated in front of a computer connected to an EyeLink 1000 desktop eye tracker (SR Research, Mississauga, Ontario, Canada). The experimenter explained that the "robot" would tell them about what they saw on the screen and that sometimes the robot would use silly words to describe pictures they had never seen before. Children were directed to listen to the sentences and follow the instructions to pick an object on the screen.

Prior to presentation of the test sentence, a familiarization phase showed an animation of drawings of a familiar item, a likely agent, and a likely theme interacting with each other, such as a large, scary creature chasing a seal and a seal chasing a small creature. This familiarization phase was developed and used by Huang and Arnold, and reinforces the likely-agent and likely-theme roles of the unknown items. The vocoded instruction "Look at the [familiar noun]" also emphasizes the familiarity of the known entity in each trial. The child was then presented with vocoded recordings of a set of 12 sentences, randomly alternated between active and passive constructions. The vocoded instruction "Click on the [novel noun]" followed each critical sentence and the child chose one of the pictures on the computer screen with a mouse click.

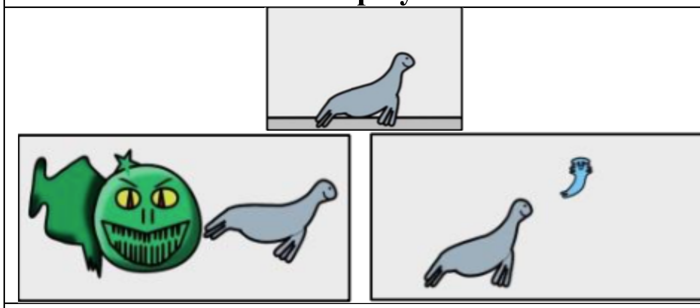
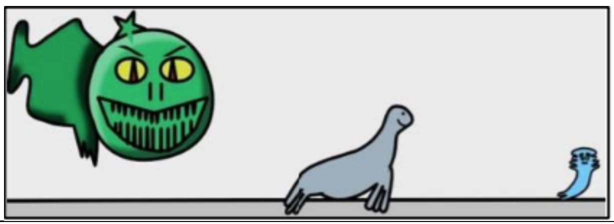
Display	Sentence
	Familiarization phase *"Look at the seal!"* VS *"Look at these!"* VS
	Test phase *"The blicket will be quickly eating the seal."* VS *"Click on the blicket."* VS
• • •	x 12 critical trials and 6 filler trials

Figure 2. Sample trial for Word-learning task. Natural speech (NS); Vocoded speech (VS).

The full set of sentences for the word-learning task was identical to that used by Huang and Arnold (see Appendix B). Each sentence included a novel nonsense noun and a familiar noun. Sentences were created using 12 words from the ARC non-word database (Rastle, Harrington, & Coltheart, 2002). The critical trials follow a 2 x 2 design, with the first factor being the sentence construction – active or passive – and the second factor being novel word position – first or second noun phrase. Four sentences were created with each word: Active Novel-NP1, Passive Novel-NP1, Active Novel-NP2, and Passive Novel-NP2 (see (3) below). Half of the children were randomly assigned to the novel-NP1 condition and the other half to the novel-NP2 condition, as suggested by data from Huang et al. (2013) that indicates children show interference when the novel NP alternates position between trials. Sentence construction was varied between trials for each participant. Auxiliary verbs and adverbs were inserted into each sentence to create a period of ambiguity before the disambiguating syntactic cue was heard.

(3) a. Active Novel-NP1: The *blicket* will be quickly eating the seal.
b. Passive Novel-NP1: The *blicket* will be quickly eaten by the seal.
c. Active Novel-NP2: The seal will be quickly eating the *blicket*.
d. Passive Novel-NP2: The seal will be quickly eaten by the *blicket*.

From this total 48 sentences, four sets of 12 sentences each were assembled, with six active and six passive sentences in each, and no repeated novel words within a set. Each set of sentences also included six filler trials to mask the critical sentence manipulations. Filler sentences used the same structure as the active test sentences, but included familiar objects only (e.g., The sheep will be slowly eating the grass).

The pilot testing with NH adults also ensured that the syntactic markers distinguishing between active and passive sentences could be perceived in eight-channel noise-vocoded speech recordings. Since NH adults accurately comprehended the natural speech recordings in Huang and Arnold, the pilot stimuli included only static pictures of the familiar object, likely agent, and likely theme. Because familiarization animations were not included, some items were ambiguous. All adults selected the distractor rather than target object when presented with the stimuli including the novel word "coopa" (see Appendix B). Those hearing the passive sentences including the novel word "daylon" also responded incorrectly. Both of these patterns reflected ambiguities in the visual stimuli, according to participant report. These ambiguities were minimized by the object familiarization animations that preceded critical trials with child participants. All other pilot responses demonstrated that the NH adults could perceive the syntactic markers, resulting in correct selection of the target objects.

2.3 Coding

The primary measures for statistical analysis focused on responses in the word-learning task. Responses in the vocoded speech training task were recorded, but performance on this task was not shown to have a significant effect on the word-learning task. Actions and eye movements were coded in the following ways.

Actions Mouse clicks following the test sentence and click instruction ("Click on the blicket") were recorded for all trials for 35 participants. Mouse clicks for an additional five participants were recorded by hand for a subset of trials (5 to 17 trials per subject, 4 to 11 critical trials per subject). Clicks were recorded as accurate if selection was of the target item and inaccurate if selection was of the distractor item or familiar object. When examining agent preference, clicks were converted to a binary variable where 0=likely agent and 1=likely theme or familiar object.

Eye movements Eye tracking data was recorded and coded for 35 participants. An additional five participants had no eye tracking data recorded due to eye tracker malfunction or early termination due to noncompliance. Eye movements were recorded in two millisecond intervals from the onset of the disambiguating cue to the offset of the sentence. Fixations were coded as looks to the familiar object, likely theme, likely agent, or looks away from these areas.

3. Results

Performance in the current experiment was analyzed for: (1) accuracy of mouse clicks comparing filler trials to critical trials and performance between critical sentence constructions; (2) agent preference of mouse clicks across conditions; and (3) real-time comprehension as measured by eye movements following the disambiguating cue (e.g.,

-ing for active sentences, *-ed by* for passive sentences). These analyses were further compared to those from the natural speech condition in Huang and Arnold.

3.1 Mouse click: Accuracy

To determine if participants could understand the vocoded stimuli and use this input to carry out task instructions, mean accuracy of filler trials was compared to chance using a two-tailed t-test. Chance was set at 50% because in critical trials there were two unfamiliar objects to choose from. While children did occasionally select the familiar object named in the sentence, these selections accounted for only 4.3% of actions. As shown in Figure 3, children were highly accurate at selecting the target object when all words in an active sentence were familiar (96%; $t(39) = 32.13$, $p < 0.001$). This result indicates that children were able to comprehend the vocoded speech and that any results from critical trials are due to interactions of task demands rather than inaccurate perception of the input alone. Accuracy on critical trials was lower, though still above chance (61%; $t(39) = 4.11$, $p < 0.001$). Comparison of performance on filler trials and critical trials suggests that there was a difference between comprehension of the noise-vocoded stimuli and action accuracy in a learning context ($F(1, 39) = 131.44$, $p < .001$).

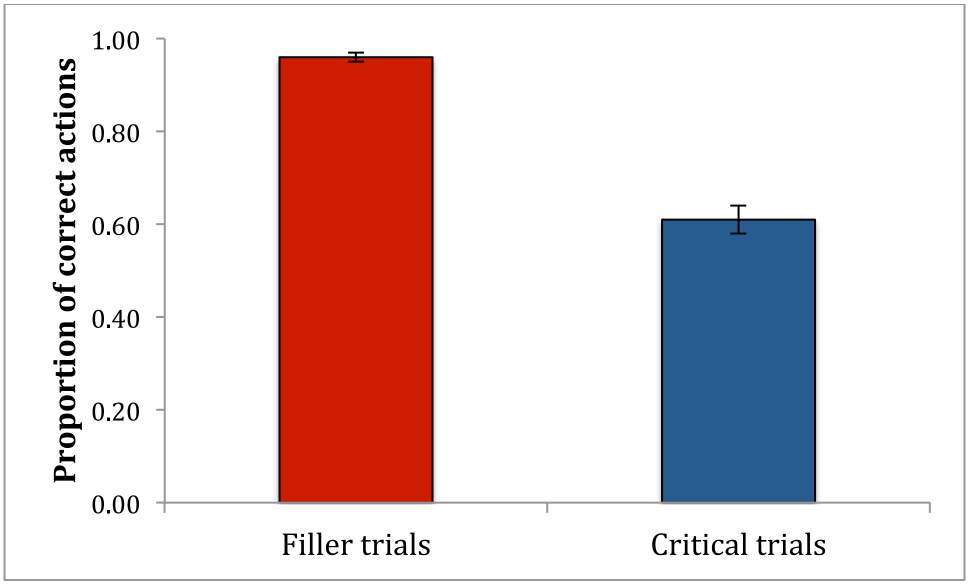


Figure 3. In the vocoded speech condition, accuracy of click actions across filler trials (all familiar words in active sentences) and critical trials (novel words in active and passive sentences).

Next, to determine interactions of sentence construction, novel word position, and speech type (natural vs. vocoded speech), accuracy in the critical trials was compared between conditions and between the current experiment and Huang and Arnold. Figure 4 shows accuracy data from the natural speech (Huang & Arnold, 2016) and vocoded speech conditions. Accuracy for each of four conditions was compared to chance (50%) using one-sample, two-tailed t-tests. In the natural speech condition, children's performance was above chance for NP1-active (86%; $t(19) = 9.65$, $p < .001$) and NP2-passive sentences (74%; $t(19) = 3.66$, $p = .002$). For NP2-active sentences, performance was above chance, although this effect was marginal (63%; $t(19) = 1.91$, $p = .07$). In contrast, accuracy was at chance for NP1-passive sentences (39%; $t(19) = -1.42$, $p = .17$). This demonstrates that when listening to natural speech, children are able to withhold the agent-first bias and use syntactic bootstrapping to determine word meaning when a known word is the first NP of a passive sentence. However, when they must revise initial misinterpretations (NP1-passive), they do not reliably learn the meanings of novel words.

Results from the vocoded speech show that accuracy was above chance for active sentences in the NP1 condition (79%, $t(19) = 7.51$, $p < .001$) and passive sentences in the NP2 condition (73%, $t(19) = 4.44$, $p < .001$). However, selection was at chance for actives in the NP2 condition (55%, $t(19) = .84$, $p = .40$) and below chance for passive sentences in the NP1 condition (33%; $t(19) = 3.27$, $p = .004$). Thus, similar to the natural speech condition, children exhibit an agent-first bias when interpreting passive sentences with a novel NP1, but they are able to inhibit this bias when NP1 is a familiar word. Additionally, accuracy in the vocoded speech condition was not significantly lower than in the natural speech condition. An ANOVA comparison showed no effect of speech type on accuracy across each of the four sentence permutations (all p's $>.14$). Word-learning performance in the NP1-passive condition for both speech conditions reveals children's agent-first bias, wherein they assign the agent role to the novel NP1 and do not reliably revise misinterpretations.

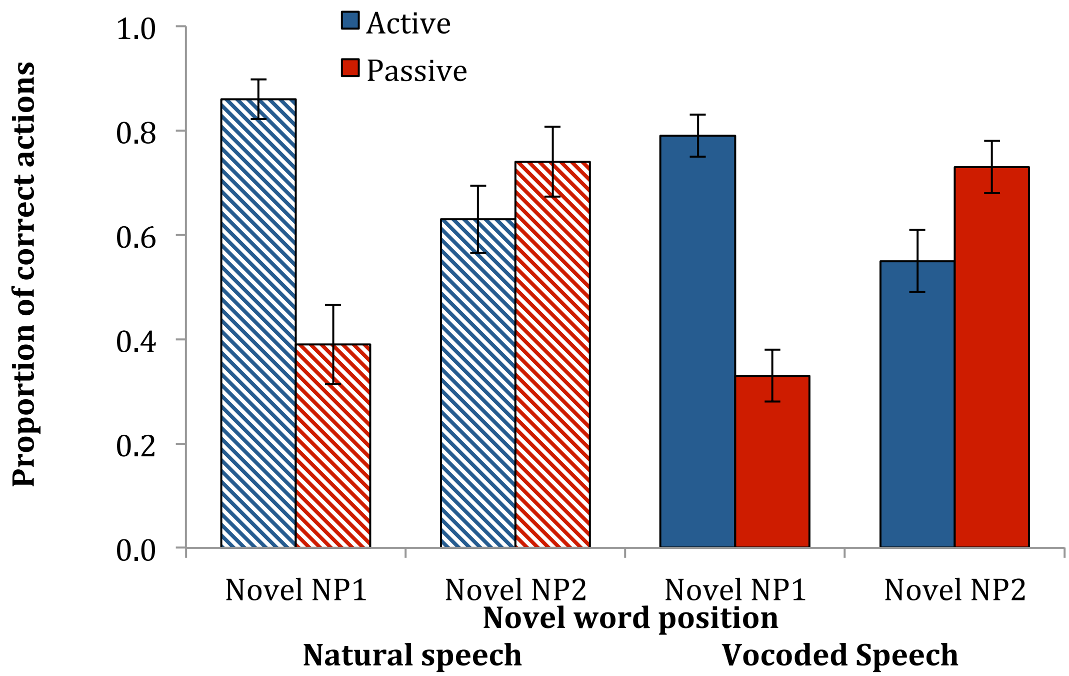


Figure 4. Proportion of correct actions following sentence completion in each novel word position.

However, chance performance in the NP2-active condition cannot be explained by the same mechanism, since no revision is necessary, and children are generally reliable interpreters of active sentences. Here, accuracy measures may be confounded by an agent preference. If children are disproportionately inclined to choose a likely agent across conditions (e.g., because it is the most interesting looking object in the scene), then their accuracy in the NP2-passive condition (in which the likely agent is the target object) may be inflated or their accuracy in the NP1-passive condition (in which the likely theme is the target object) may be deflated. Indeed, that seems to be the case across both speech conditions, as accuracy for NP2-passive sentences is surprisingly higher than for NP2-active sentences. Together, this suggests that in order to tease apart how children's behaviors change with respect to the linguistic cue (active or passive), we need to adopt a measure that is unbiased across condition.

3.2 Mouse click: Agent preference

Agent preference may instead be a more neutral measure to examine performance and differences in word learning when revision is necessary. It was calculated as the proportion of clicks on the likely agent for each construction type. Within each novel word position, differences in agent preference indicate the extent to which children can differentiate between active and passive sentences. In novel-NP1 sentences, accurate differentiation would result in an agent preference score approaching 1 for active sentences, when the target is the agent, and 0 for passive sentences, when the target is the theme. The opposite would be true for novel-NP2 sentences: agent preference scores approaching 0 for active sentences and 1 for passive sentences. Thus, if children always distinguished between active/passive constructions, the difference between the agent preference scores approaches a maximum of 1. In contrast, when children are less able to differentiate, the difference between the agent preference scores becomes smaller. For NP1-active sentences, in which the likely agent is the target object, the agent preference score should approach 1, since children interpret these sentences correctly. An agent-first bias in NP1-passive sentences would also cause the agent preference score to approach 1, because children initially assign the agent role to the novel NP and then fail to revise this bias. Importantly, as the agent preference score for NP1-passive sentences increases, the difference between agent preference in the two constructions decreases.

Agent preference scores can be seen in Figure 5. Agent preference in the natural speech condition was highest for conditions in which the likely agent was the target object – NP1-active (86%) and NP2-passive (74%). Agent preference was lower when the likely theme was the target object – NP1-passive (61%) and NP2-active (37%). Using a 2 x 2 ANOVA, comparison of agent preference scores across sentence construction (active vs. passive) within levels of novel word position (NP1 vs. NP2) shows significant effects of sentence construction for both novel-NP1 ($F(1, 19) = 7.61$, $p = .01$) and novel-NP2 constructions ($F(1, 19) = 18.77$, $p < .001$). As discussed above, a significant difference between agent preference scores within novel-word conditions indicates that children are able to differentiate active and passive sentence constructions in the natural speech condition.

Agent preference scores for the vocoded speech condition are again highest when the likely agent is the target object – NP1-active (79%) and NP2-passive (73%). Scores when the likely theme was the target object were 67% for NP1-passive sentences and 45% for NP2-active sentences. However, while the difference between active and passive constructions was significant for novel-NP2 sentences ($F(1, 19) = 14.58$, $p = .001$), it was not for novel-NP1 sentences ($F(1, 19) = 3.39$ $p = .08$). These results suggest that children in the vocoded speech condition do not distinguish between active and passive constructions when revision is necessary (NP1-passive), even though they can make this distinction when no revision is necessary. There is not a significant effect of speech type on agent preference, though results are trending toward significance ($F(1,39) = 1.81$, $p = .18$).

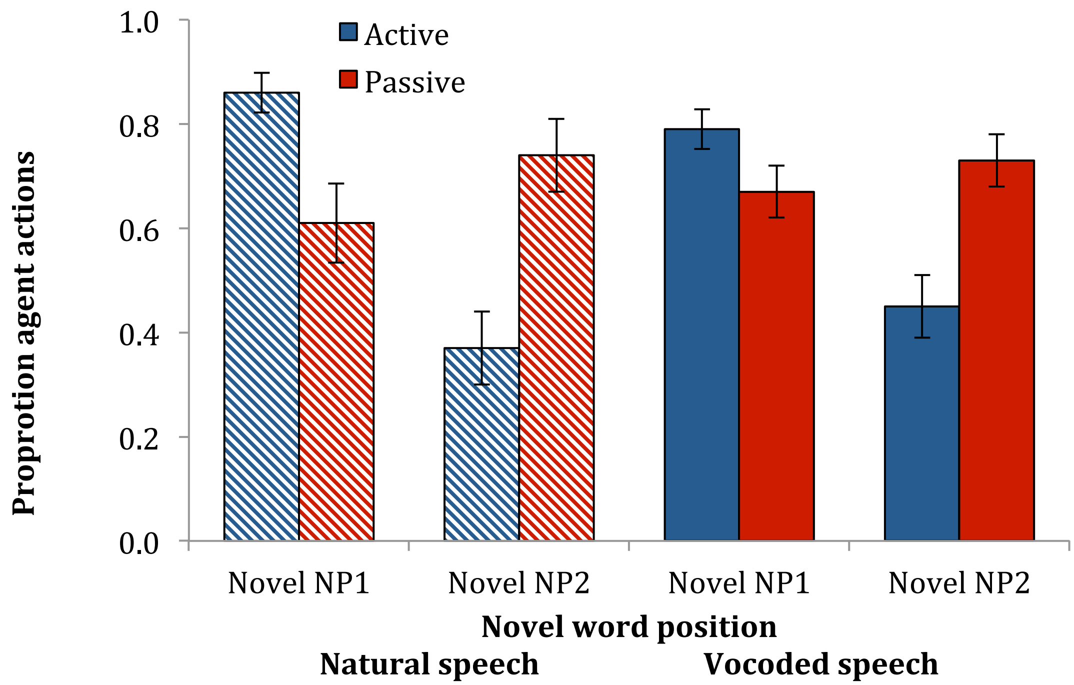


Figure 5. Agent preference shown as the proportion of agent actions following sentence completion in each novel word position.

3.3 Real-time comprehension

Analysis of eye tracking data was used to examine real-time comprehension and sensitivity to syntactic cues. Fixations were recorded from the onset of disambiguating syntactic cues (e.g., "-ing" in "eating" vs. "-en" in "eaten") until sentence offset, a period of approximately 1150ms. Time windows were shifted by 400ms to account for the time it takes children to generate saccadic eye movement (see Huang, et al., 2013). Figures 6 and 7 show fixation preferences by condition for natural and vocoded speech, respectively. Fixation preference was calculated by subtracting looks to the competitor from looks to the target for passive sentences and looks to the target from looks to the competitor for active sentences. This means that positive values indicate greater sensitivity to syntactic cues in passive sentences and negative values indicate greater sensitivity in active sentences.

If children are sensitive to syntactic cues in the sentences, we would expect to see the fixation preference score lines for active and passive constructions pull apart from each other within each novel word condition. This is true of novel-NP2 sentences in both speech conditions. The effect of sentence construction is significant in the natural speech condition ($F(1, 19) = 16.86$, $p = .001$) as well as the vocoded speech condition ($F(1, 17) = 14.24$, $p = .002$). However, these lines do not pull apart in the novel-NP1 sentences in either speech condition. The fixation preference scores for active sentences are negative, meaning children are correctly fixating on the target object. But they are also negative for passive sentences, meaning that children have not revised their agent-first bias after the disambiguating syntactic markers. Thus, they continue to look to the competitor object. Additionally, the effect of sentence construction in novel-NP1 sentences is not significant in either the natural speech condition ($F(1, 19) = .03$, $p = .88$) or the vocoded speech condition ($F(1, 16) = 1.38$, $p = .26$).

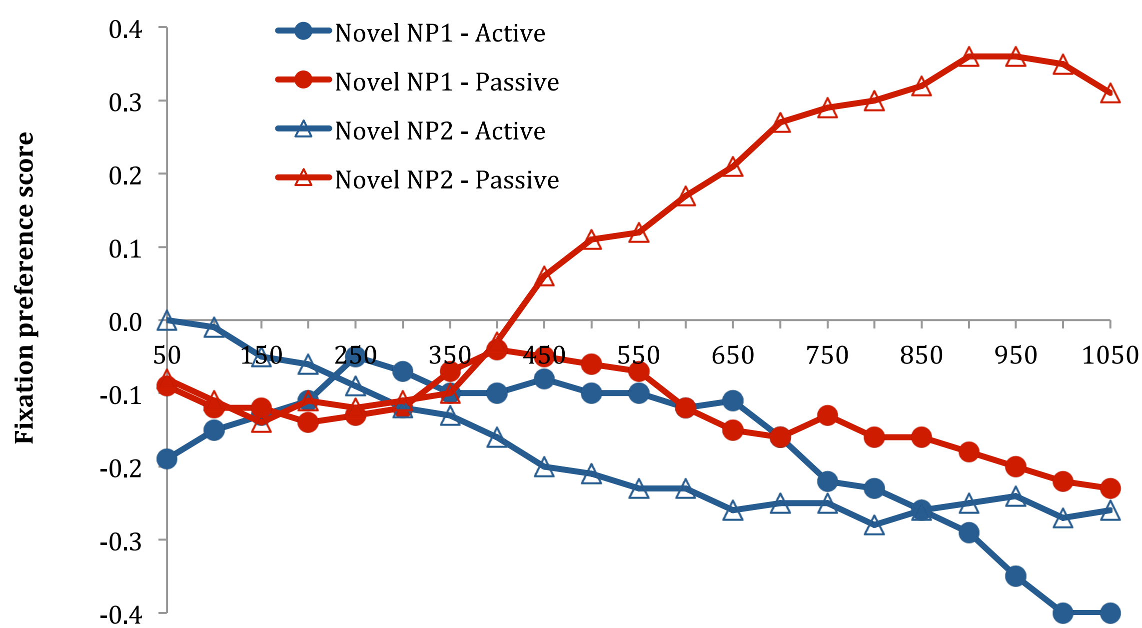


Figure 6. Natural speech condition. Fixation preference score. Correct fixations to the Target are indicated by positive scores in passive trials (in red) and negative scores in active trials (in blue).

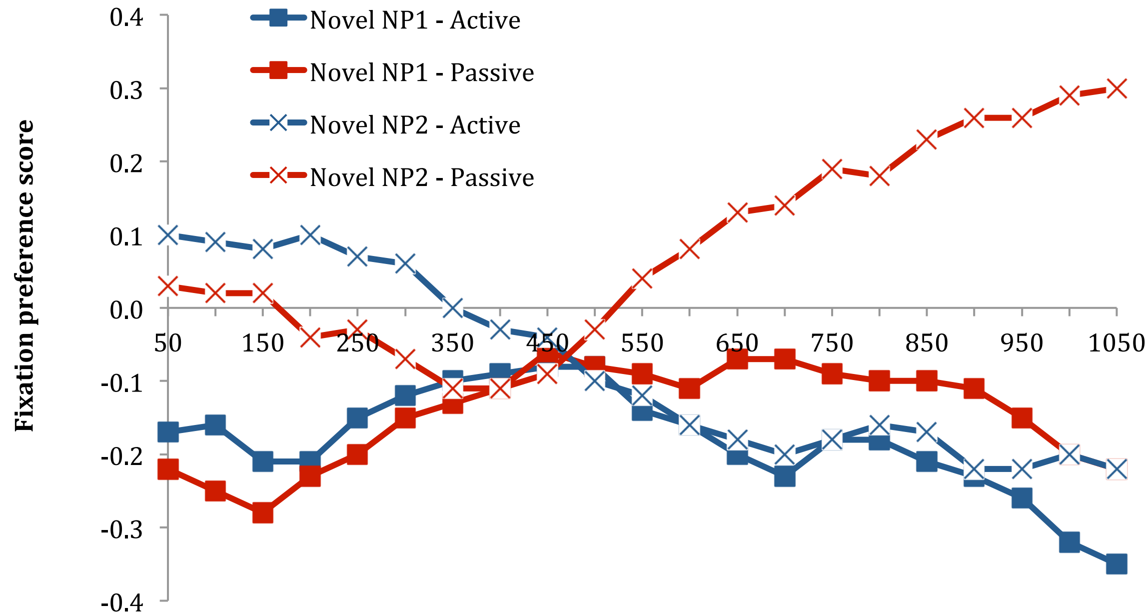


Figure 7. Vocoded speech condition. Fixation preference score. Correct fixations to the Target are indicated by positive scores in passive trials (in red) and negative scores in active trials (in blue).

The converse comparison can also be made by looking at the effect of novel word position within each construction type. There is not a significant difference between fixation preference scores for active sentences with either a novel-NP1 or a novel-NP2 in both the natural speech ($F(1, 39) = .28$, $p = .64$) and vocoded speech ($F(1, 34) = .09$, $p = .77$) conditions. This demonstrates that children correctly interpret active sentences in real time and tend to fixate on the target object regardless of the location of the novel word. Thus, if children could reliably use syntactic cues in passive sentences to learn the meaning of novel words, we would also expect the passive fixation preference scores to be similar between novel word positions. This is not the case, though. Fixation preference for the target object during passive sentences was significantly higher in the novel-NP2 condition than it was in the novel-NP1 condition for both speech types (natural speech: ($F(1, 39) = 12.57$, $p = .001$; vocoded speech: ($F(1, 34) = 11.04$, $p = .002$). This demonstrates that children were sensitive to the syntactic cues in the NP2-passive sentences, as shown by their looks to the target object, but were not sensitive to the same syntactic cues in the NP1-passive sentences.

4. Discussion

This study examined real-time processing and word learning in children receiving a degraded audio signal, and specifically assessed whether increased uncertainty in the audio signal alters the developmental strategies available for word learning. It is clear from their above-chance performance on NP1-active and NP2-passive sentence structures that children are able to use syntactic cues in a degraded speech signal to deduce word meaning. However, syntactic revision makes syntactic bootstrapping more difficult, as is revealed by below-chance accuracy for NP1-passive sentences. Additionally, preference for agent interpretations leads children astray when the target object is the likely theme, as in NP2-active sentences, resulting in at-chance accuracy. What, then, are the differences in word learning and syntactic bootstrapping with a clear audio signal versus a degraded signal, and what can the current data tell us about learning with a degraded audio signal, and ultimately, a CI? In the remainder of this section, I will address these questions through the lens of the three original possibilities suggested in the introduction, then explore some further-reaching assumptions and implications this information may have for teaching children with CIs.

4.1 Does an acoustically degraded signal make all sentence processing harder?

The broadest possible outcome presented above was that a degraded audio signal would make all sentence processing difficult. If that were the case, we would expect to see depressed performance across the board with no evidence of sensitivity to syntactic cues and even poor performance in filler trials where no word learning or revision were necessary. This was not in fact the case. Children were highly accurate in their responses to filler trials (96%), indicating that they understood the noise-vocoded speech and were able to follow directions presented in this format. Even when novel words were introduced in the critical trials, accuracy across sentence constructions in the vocoded speech condition was not significantly lower than the natural speech condition.

Children's high accuracy in the present task echoes that of other studies that have shown that even toddlers can understand eight-channel vocoded speech (Newman & Chatterjee, 2013). The sentence-level training protocol implemented here strengthens previous results that found this to be more effective than phoneme-based training (Stacey & Summerfield, 2008). Previous research has shown that adults and older children (10-12 years) are typically able to understand vocoded speech better and with fewer channels than younger children (5-7 years) are (Dorman et al., 2000; Eisenberg et al., 2000). While the current study did not compare children's accuracy to adults', the very high accuracy achieved here (96%) may be impacted by the enhanced context of the sentences used in both the training and filler sentences of the current experiment. The visual stimuli that accompanied the recorded sentences as well as the closed response set for filler trials may have increased comprehension and accuracy of responses compared to more open-set tasks lacking visual context in other studies.

Additionally, the children acclimated relatively quickly to the vocoded speech during the training tasks and were frequently able to repeat novel sentences and phrases after one presentation. Furthermore, there was no correlation between accuracy on training trials and agent preference actions. Accuracy on training trials was defined as percent of correct sentence repetitions based on number of training test sentences presented. Accuracy in the training phase approached significance in the NP2-passive condition ($r = .42$, $p = .06$). In this condition, no revision is necessary and the target object is the agent. With vocoded speech stimuli, a key component for correct interpretation is the ability to perceive the syntactic cues (e.g., eat*en by*). Training accuracy approaching significance means that increased ability to perceive (and repeat) vocoded speech may create an advantage in this condition, where competing demands (i.e., revision) are not present. Correlations for all other conditions were insignificant (all p's > .4), which suggests that individual differences in perceptual abilities in the vocoded condition do not aid in interpretation of active sentences or when cognitive demands are harder, as in cases where revision is necessary.

4.2 Does an acoustically degraded signal increase the agent-first bias?

Another outcome posited in the introduction was that noise-vocoded speech would increase the agent-first bias, making it difficult for children to inhibit the bias in novel-NP2 conditions. We know from Huang and Arnold (2016) that having a novel NP1 promotes an agent-first bias in natural speech, but that children are able to inhibit that bias with novel NP2s. If degraded input strengthens the agent-first bias across conditions, the patterns of the NP1 condition should be mirrored in the NP2 condition, as children would interpret all NP1s as the agent of the sentence. This would lead to decreased accuracy on all passive sentences and similar agent preference in actions for active and passive constructions within each novel word position condition (high agent preference for novel-NP1 sentences and low for novel-NP2 sentences). Looking at the current data, we see that this is not true. There is an apparent agent-first bias in the vocoded novel-NP1 condition. The agent preference scores are high and the effect of sentence construction on agent preference is not statistically significant ($p = .08$). Importantly, the pattern in the vocoded NP2 conditions does not match. Children are able to differentiate passives from actives, as is demonstrated by significantly lower agent preference scores for NP2-active sentences vs. NP2-passive sentences ($p = .001$). A similar effect is seen in the natural speech condition. Children differentiate between the sentence constructions and show a lower agent preference for NP2-active sentences than for NP2-passives.

Based on these results, we can deduce that even with degraded input, children are able to distinguish between active and passive constructions when no revision is necessary. They are still able to inhibit the agent-first bias when there is a familiar entity in the NP1 position and use late-emerging syntactic cues to determine novel word referents. It seems that degraded input does not lead to overall uncertainty across all conditions. This suggests that word learning in CI users may be affected not by overall processing differences, but rather more subtle or specific difficulties arising from the properties of degraded audio input.

4.3 Does an acoustically degraded signal only affect processing demands when syntactic revision is necessary?

Finally, we have seen that children have a harder time differentiating between active and passive constructions when the novel word is in the NP1 position than when it is in the NP2 position. But does the degraded nature of vocoded speech make processing even harder than natural speech? To examine this, we can compare agent preference scores between the two speech types. The effect of sentence construction on agent preference for natural speech NP1 sentences was significant ($p = .01$), but there was no significant effect for vocoded speech NP1 sentences ($p = .08$). The effect of speech type on agent preference scores was approaching significance ($p = .18$). This suggests that revision may be harder (i.e., it is harder to overcome an agent-first bias) when the auditory signal is degraded. This may be because of a cascaded effect of the degraded input on the steps required to revise. When listening to each passive sentence, the child must be able to perceive the late-occurring syntactic cue, use this cue to re-map the sentence to a passive structure, and then revise the agent-first bias and reassign the theme role to the NP1. If there is increased uncertainty in the first step of this process, the child will have difficulty in the subsequent steps. Alternatively, increased effort throughout all of these steps may lead to reduced revision and comprehension, particularly in a continuous speech stream.

Further research with a larger sample may help to illuminate these differences. If degraded input makes syntactic revision even harder than it is with natural speech, this could have important implications for how children with CIs learn word meanings and use processing strategies. It seems to not be the case that CIs have a blanket effect on comprehension strategies, but instead affect children's understanding only when syntactic demands are high.

4.4 Further discussion

Noise-vocoded speech simulates some aspects of CI sound processing, and previous research has shown correspondence between performance of NH listeners with vocoded speech to CI listeners. However, simulations with noise-vocoded speech may not tell the whole story of how children with CIs would perform with similar constructions in a real-world situation. NH four- and five-year-olds have had years of receiving non-degraded speech. They are on their way to comprehending passive sentences and using syntactic bootstrapping. It may be the case that the current subjects are able to rely on their previous experience with this strategy and implement it even in the face of a less clear signal. On the other hand, young CI users, who may never have had access to a clear auditory signal, may not have built up the same skills and strategies to the extent that their NH peers have. When given the same experimental task, their performance may be worse or their inability to use syntactic cues to decipher word meaning may have more to do with lack of exposure or practice with bootstrapping rather than increased processing demands.

On the other hand, children with CIs have years of practice listening to and extracting meaning from a degraded signal, so they may have been able to develop the same or similar bootstrapping strategies. However, we can conclude some effect of processing demands from the current study. Both the vocoded speech group of the current study and the natural speech group from Huang and Arnold (2016) have received similar previous exposure to the constructions used as stimulus items, yet when receiving a degraded audio signal, the group hearing vocoded speech was less able to employ strategies that the natural speech group could. Even though the same strategies are available to both groups, the current group is less able to access them.

Findings from the current study could have important implications for the methods of language acquisition in children with CIs. If increased processing is required to decode the auditory signal when syntactic demands are high, CI users may be less efficient language learners – an effect not due solely to increased difficulty in understanding spectrally degraded speech. Further investigation with a larger sample may help to determine how large these effects are. Additional studies could also be developed that would clarify the mechanisms at play that make processing more difficult. For example, these effects may be found to be due to differences in a combination of EF skills already found to be problematic in CI users (Kronenberger et al., 2014), or perhaps to working-memory differences seen in children with CIs (Pisoni & Cleary, 2003). Or maybe the difficulties with inhibition exhibited by NH children (Jones, Rothbart, & Posner, 2003; Snedeker & Yuan, 2008) are even more pronounced in children learning with CIs.

To combat language-learning differences clinically, explicit instruction of structures and strategies may be advantageous for language learners with CIs. For something like syntactic bootstrapping from passive sentences, children with CIs may benefit from increased exposure to the constructions and instruction to attend to the syntactic cues. NH children who are presumed to have decreased exposure to passive structures perform poorly on tasks where revision is necessary (Huang, Leech, & Rowe, 2017). There is some evidence that increasing NH children's exposure to complex forms can improve comprehension and production (Vasilyeva & Waterfall, 2011). These strategies develop naturally in NH children, but since passive constructions are already low frequency in caregiver speech (Stromswold, 2005), children with CIs may need them to be highlighted for them. Additionally, as Winn (2016) noted, cognitive effort was reduced by contextual information only after completion of a sentence. If effort were found to be the key factor in processing differences, children with CIs may benefit from significant extra time for processing following syntactically complex information.

5. Conclusion

The current study looked at word-learning performance of NH children given a degraded input signal. The evidence shows that even with a degraded signal, children can extract syntactic cues from sentences to inform word-learning, and that performance is worse when revision of original misinterpretation is necessary, a pattern that is similar to that found for children hearing natural speech. Agent preference scores suggest that the greater uncertainty or cognitive load brought on by a degraded signal may make it even harder for children receiving degraded input to revise in these conditions. Children's performance in the current study demonstrates that there are not significant effects of speech condition across all construction types, but there may be in constructions in which revision is necessary. Relying on CI users' performance in conversational speech, where they may seem to be performing similar to NH peers, may not show these differences, as increased problems with revision would be hard to differentiate and could instead manifest in less efficient use of less observable word-learning strategies.

Appendix A

Vocoded Speech Training Task

Familiarization Phase

1 (a) A cat. (b) The dog runs.
2 (a) The fox walks. (b) The girl is catching.
3 (a) The boy is in the garden. (b) The seal is swimming slowly.
4 (a) The rocks are rolling quickly. (b) The frog is sitting on a log.
5 (a) The mouse is eating cheese. (b) The car is driving in the woods.
6 (a) The monkey is swinging in the tree. (b) The rabbit is racing the turtle.

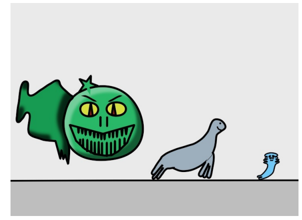
Test Phase

1 (a) A bush. (b) A giraffe.
2 (a) The sheep jumps. (b) The cow walks.
3 (a) The grass is long. (b) The penguin is swimming.
4 (a) The chicken sits on a next. (b) The elephant is taking a drink.
5 (a) The eagle flies very high. (b) The antelope is running in the field.
6 (a) The polar bear is standing on the ice. (b) The duck is hiding in the grass.

Appendix B

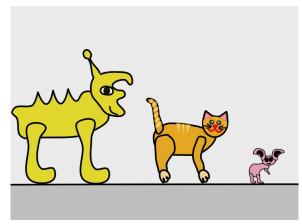
Novel word: "Blicket"

NP1 / Active: The blicket will be quickly eating the seal.
NP1 / Passive: The blicket will be quickly eaten by the seal.
NP2 / Active: The seal will be quickly eating the blicket.
NP2 / Passive: The seal will be quickly eaten by the blicket.



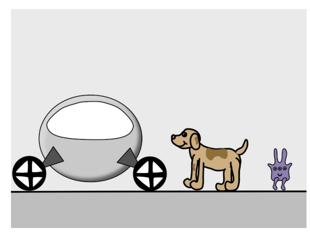
Novel word: "Nedoke"

NP1 / Active: The nedoke will be quickly scaring the cat.
NP1 / Passive: The nedoke will be quickly scared by the cat.
NP2 / Active: The cat will be quickly scaring the nedoke.
NP2 / Passive: The cat will be quickly scared by the nedoke.



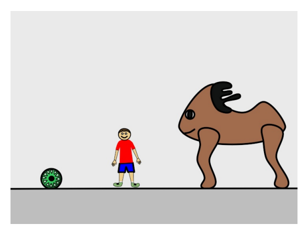
Novel word: "Coopa"

NP1 / Active: The coopa will be quickly chasing the dog.
NP1 / Passive: The coopa will be quickly chased by the dog.
NP2 / Active: The dog will be quickly chasing the coopa.
NP2 / Passive: The dog will be quickly chased by the coopa.



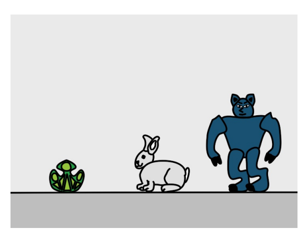
Novel word: "Hantil"

NP1 / Active: The hantil will be gently kicking the boy.
NP1 / Passive: The hantil will be gently kicked by the boy.
NP2 / Active: The boy will be gently kicking the hantil.
NP2 / Passive: The boy will be gently kicked by the hantil.



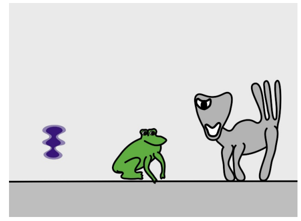
Novel word: "Leepo"

NP1 / Active: The leepo will be slowly eating the rabbit.
NP1 / Passive: The leepo will be slowly eaten by the rabbit.
NP2 / Active: The rabbit will be slowly eating the leepo.
NP2 / Passive: The rabbit will be slowly eaten by the leepo.



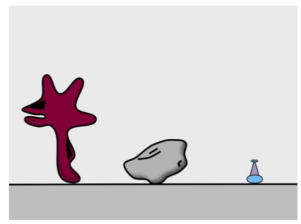
Novel word: “Daylon”

NP1 / Active: The daylon will be quietly catching the frog.
NP1 / Passive: The daylon will be quietly caught by the frog.
NP2 / Active: The frog will be quietly catching the daylon.
NP2 / Passive: The frog will be quietly caught by the daylon.



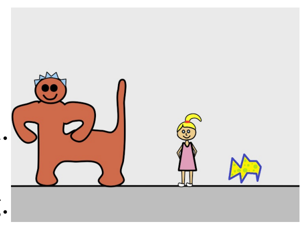
Novel word: “Tayvak”

NP1 / Active: The tayvak will be loudly smashing the rock.
NP1 / Passive: The tayvak will be loudly smashed by the rock.
NP2 / Active: The rock will be loudly smashing the tayvak.
NP2 / Passive: The rock will be loudly smashed by the tayvak.



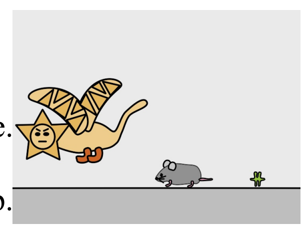
Novel word: “Chowvag”

NP1 / Active: The chowvag will be carefully lifting the girl.
NP1 / Passive: The chowvag will be carefully lifted up by the girl.
NP2 / Active: The girl will be carefully lifting the chowvag.
NP2 / Passive: The girl will be carefully lifted up by the chowvag.



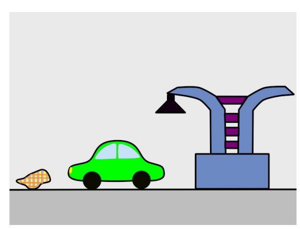
Novel word: “Vaychip”

NP1 / Active: The vaychip will be quickly grabbing the mouse.
NP1 / Passive: The vaychip will be quickly grabbed by the mouse.
NP2 / Active: The mouse will be quickly grabbing the vaychip.
NP2 / Passive: The mouse will be quickly grabbed by the vaychip.



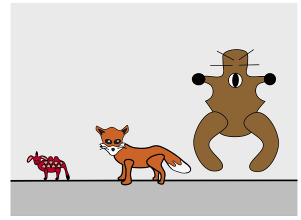
Novel word: “Noytoff”

NP1 / Active: The noytoff will be loudly squishing the car.
NP1 / Passive: The noytoff will be loudly squished by the car.
NP2 / Active: The car will be loudly squishing the noytoff.
NP2 / Passive: The car will be loudly squished by the noytoff.



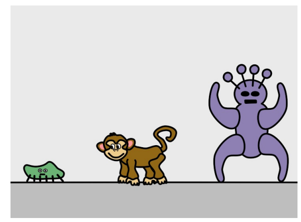
Novel word: "Bellwer"

NP1 / Active: The bellwer will be quickly chasing the fox.
NP1 / Passive: The bellwer will be quickly chased by the fox.
NP2 / Active: The fox will be quickly chasing the bellwer.
NP2 / Passive: The fox will be quickly chased by the bellwer.



Novel word: "Furpin"

NP1 / Active: The furpin will be quickly scaring the monkey.
NP1 / Passive: The furpin will be quickly scared by the monkey.
NP2 / Active: The monkey will be quickly scaring the furpin.
NP2 / Passive: The monkey will be quickly scared by the furpin.



References

Ali, W. & O'Connell, R. (2007). The effectiveness of early cochlear implantation for infants and young children with hearing loss. *NZHTA Technical Brief, 6*(5).

Bollard, P. M., Popp, A., Chute, P. M., & Parisier, S. C. (1999). Specific language growth in young children using the Clarion® Cochlear Implant. *Annals of Otology, Rhinology & Laryngology, 108*(4), 119-123.

Brown, R. (1957) Linguistic determinism and the part of speech. *Journal of Abnormal and Social Psychology, 55*, 1–5.

Castellanos, I., Pisoni, D. B., Kronenberger, W. G., & Beer, J. (2016). Neurocognitive function in deaf children with cochlear implants: early development and long-term outcomes. In Marschark, M. & Spencer, P. E. (Eds.), *The Oxford Handbook of Deaf Studies in Language* (pp. 264-275). New York, NY: Oxford University Press.

Dautriche, I., Cristia, A., Brusini, P., Yuan, S., Fisher, C., & Cristophe, A. (2014). Toddlers default to canonical surface-to-meaning mapping when learning verbs. *Child Development, 85*(3), 1168-1180.

Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., & McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences. *Journal of Experimental Psychology: General, 134*(2), 222-241.

Dorman, M. F., Loizou, P. C., Kemp, L. L., & Kirk, K. I. (2000). Word recognition by children listening to speech processed into a small number of channels: data from normal-hearing children and children with cochlear implants. *Ear and Hearing, 21*(6), 590-596.

Duchesne, L. (2016). Grammatical competence after early cochlear implantation. In Marschark, M. & Spencer, P. E. (Eds.), *The Oxford Handbook of Deaf Studies in Language* (pp. 113-131). New York, NY: Oxford University Press.

Eisenberg, L. S., Shannon, R. V., Martinez, A. S., Wygonski, J., & Boothroyd, A. (2000). Speech recognition with reduced spectral cues as a function of age. *Journal of the Acoustical Society of America, 107*(5), 2704-2710).

Faulkner, A., Rosen, S., & Smith, C. (2000). Effects of the salience of pitch and periodicity information on the intelligibility of four-channel vocoded speech: Implications for cochlear implants. *Journal of the Acoustical Society of America, 108*, 1877-1887.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47*, 164-203.

Friesen, L. M., Shannon, R. V., Baskent, D., & Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *The Journal of the Acoustical Society of America, 110*(2), 1150-1163.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10*, 483–527.

Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences, 110*, 8051-8056.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition, 1,* 3-55.

Gleitman, L. & Gleitman, H. (1992). A picture is worth a thousand words, but that's the problem: The role of syntax in vocabulary acquisition. *Current Directions in Psychological Science, 1*, 31–35.

Grieco-Calub, T., Saffran, J. R., & Litovsky, R. Y. (2009). Spoken word recognition in toddlers who use cochlear implants. *Journal of Speech Language and Hearing Research 52*(6), 1390–1400.

Hahne, A., Wolf, A., Müller, J., Murbe, D., & Friederici, A. D. (2012). Sentence comprehension in proficient adult cochlear implant users: On the vulnerability of syntax. *Language and Cognitive Processes, 27*(7-8), 1192-1204.

Houston, D. M., & Bergeson, T. R. (2014). Hearing versus listening: Attention to speech and its role in language acquisition in deaf infants with cochlear implants. *Lingua. International Review of General Linguistics. Revue Internationale de Linguistique Generale*, *139*, 10–25.

Huang, Y., Abadie, L., Arnold, A., & Hollister, E. (2016). Novelty of discourse referents promotes heuristics in children's syntactic processing. Presented at the *29th Annual CUNY Conference on Human Sentence Processing.* Gainesville, FL.

Huang, Y. & Arnold, A. (2016). Word learning in linguistic context: Processing and memory effects. *Cognition, 156*, 71-87.

Huang, Y., Leech, K., & Rowe, M. L. (2017). Exploring socioeconomic differences in syntactic development through the lens of real-time processing. *Cognition, 159*, 61-75.

Huang, Y., Zheng, X., Meng, X., & Snedeker, J. (2013). Children's assignment of grammatical roles in the online processing of Mandarin passive sentences. *Journal of Memory and Language, 69*, 589-606.

Huyck, J. J. & Johnsrude, I. S. (2012). Rapid perceptual learning of noise-vocoded speech requires attention. *The Journal of the Acoustical Society of America, 131*(3), EL236-EL242.

Jones, L. B., Rothbart, M. K., & Posner, M. I. (2003). Development of executive attention in preschool children. *Developmental Science, 6*(5), 498–504.

Komesaroff, L. (Ed.) (2007). *Surgical consent bioethics and cochlear implantation.* Washington, DC: Gallaudet University Press.

Kronenberger, W.G., Beer, J., Castellanos, I., Pisoni, D.B., & Miyamoto, R.T. (2014). Neurocognitive risk in children with cochlear implants. *JAMA Otolaryngology - Head And Neck Surgery*, *140*(7), 608-615.

Landau, B. & Gleitman, L. (1985). *Language and experience: Evidence from the blind child*. Cambridge, MA: Harvard University Press.

Levy, R., Bicknell, K., Slattery, T., & Rayner, K. (2009). Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the National Academy of Sciences, 106*, 21086-21090.

Loizou, P. C., Dorman, M., & Tu, Z. (1999). On the number of channels needed to understand speech. *Journal of the Acoustical Society of America*, *106*(4 Pt. 1), 2097–2103.

Mazuka, R., Jincho, N., & Onishi, H. (2009). Development of executive control and language processing. Language and Linguistics Compass, 3, 59-89.

Müller, U., Zelazo, P. D., Hood, S., Leone, T., & Rohrer, L. (2004). Interference control in a new rule use task: Age-related changes, labeling, and attention. *Child Development, 75*(5), 1594–1609.

National Institutes of Health (2016, June 17). *Quick statistics about hearing*. Retrieved from https://www.nidcd.nih.gov/health/statistics/quick-statistics-hearing

National Institute on Deafness and Other Communication Disorders (NIDCD) (n.d.). *Cochlear implants fact sheet*. Retrieved from http://report.nih.gov/nihfactsheets/ViewFactSheet.aspx?csid=83

Newman, R. & Chatterjee, M. (2013). Toddlers' recognition of noise-vocoded speech. *The Journal of the Acoustical Society of America, 133*(1), 483-494.

Nicholas, J. G. & Geers, A. E. (2007). Will they catch up? The role of age at cochlear implantation in the spoken language development of children with severe to profound hearing loss. *Journal of Speech, Language, and Hearing Research, 50*, 1048–1062.

Nicholas, J. G. & Geers, A. E. (2008). Expected test scores for preschoolers with a cochlear implant who use spoken language. *American Journal of Speech-Language Pathology, 17*, 121–138.

Niparko, J. K., Tobey, E. A., Thal, D. J., Eisenberg, L. S., Wang, N., Quittner, A. L., & Fink, N. E. (2010). Spoken language development in children following cochlear implantation. *Journal of the American Medical Association, 303*(15), 1498-1506.

Obleser, J., Meyer, L., & Friederici, A. D. (2011). Dynamic assignment of neural resources in auditory comprehension of complex sentences. *Neuroimage, 56*(4), 2310-2320.

Obleser, J., Wise, R. J. S., Dresner, M. A., & Scott, S. K. (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *The Journal of Neuroscience, 27*(9), 2283-2289.

Obleser, J., Wostmann, M., Hellbernd, N., Wilsch, A., & Maess, B. (2012). Adverse listening conditions and memory load drive a common alpha oscillatory network. *The Journal of Neuroscience, 32*(36), 12376-12383.

Peterson, N. R., Pisoni, D. B., & Miyamoto, R. T. (2010) Cochlear implants and spoken language processing abilities: Review and assessment of the literature. *Restorative Neurology and Neuroscience, 28*, 237–250

Pisoni, D. B., & Cleary, M. (2003). Measures of working memory span and verbal rehearsal speed in deaf children after cochlear implantation. *Ear and Hearing*, *24*(1 Suppl), 106S–120S. http://doi.org/10.1097/01.AUD.0000051692.05140.8E

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC Nonword Database. *Quarterly Journal of Experimental Psychology, 55A*, 1339-1362.

Ridderinkhof, K. R., Band, G. P. H., & Logan, G. D. (1999). A study of adaptive behavior: effects of age and irrelevant information on the ability to inhibit one's actions. *Acta Psychologica, 101*, 315–37.

Roman, A. S. (2015). *Neurocognitive correlates of spectrally-degraded speech recognition in children* (Order No. 3684624). Available from ProQuest Dissertations & Theses Global. (1660541630).

Sarant, J. Z., Blamey, P. J., Dowell, R. C., Clark, G. M. & Gibson, W. P. (2001). Variation in speech perception scores among children with cochlear implants. *Ear and Hearing 22*, 18–28.

Sarant, J. Z., Harris, D. C., & Bennet, L. A. (2015). Academic outcomes for school-aged children with severe-profound hearing loss and early unilateral and bilateral cochlear implants. *Journal of Speech, Language, and Hearing Research, 58*. 1017-1032.

Schorr, E. A., Roth, F. P., & Fox, N. A. (2008). A comparison of the speech and language skills of children with cochlear implants and children with normal hearing. *Communication Disorders Quarterly, 29*(4), 195-210.

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

Sheldon, S., Pichora-Fuller, M. K., & Schneider, B. A. (2008). Priming and sentence context support listening to noise-vocoded speech by younger and older adults. *Journal of the Acoustical Society of America, 123*, 489-499.

Snedeker, J. & Yuan, S. (2008). Effects of prosodic and lexical constraints on parsing in young children (and adults). *Journal of Memory and Language, 58,* 574-608.

Sohoglu, E., Peelle, J.E., Carlyon, R.P., Davis, M.H. (2012). Predictive top-down integration of prior knowledge during speech perception. *Journal of Neuroscience, 32*(25), 8443-8453.

Stacey, P. C. & Summerfield, A. Q. (2008). Comparison of word-, sentence-, and phoneme-based training strategies in improving the perception of spectrally distorted speech. *Journal of Speech, Language, and Hearing Research, 51*, 526-538.

Stromswold, K. (2005). Why Children Understand and Misunderstand Sentences: An Eye-tracking Study of Passive Sentences. New Brunswick, NJ: Rutgers University.

Svirsky, M. A., Chute, P. M., Green, J., Bollard, P., & Miyamoto, R. T. (2000). Language development in children who are prelingually deaf who have used the SPEAK or CIS stimulation strategies since initial stimulation. *Volta Review, 102*(4), 199-213.

Svirsky, M. A., Stallings, L. M., Lento, C. L., Ying, E., & Leonard, L. B. (2002). Grammatical morphologic development in pediatric cochlear implant users may be affected by the perceptual prominence of the relevant markers. *Annals of Otology, Rhinology & Laryngology, 111*, 109-112.

Svirsky, M. A., Teoh, S., & Neuburger, H. (2004). Development of language and speech perception in congenitally, profoundly deaf children as a function of age at cochlear implantation. *Audiology & Neuro-otology, 9*, 224-233.

Szagun, G. & Schramm, S. A. (2016). Sources of variability in language development of children with cochlear implants: age at implantation, parental language, and early features of children's language construction. *Journal of Child Language, 43*, 505-536.

Trueswell, J.C., Sekerina, I., Hill, N.M., & Logrip, M.L. (1999). The kindergarten-path effect: studying on-line sentence processing in young children. *Cognition, 73,* 89-134.

van Heugten, M., Volkova, A., Trehub, S. E., & Schellenberg, E. G. (2013). Children's recognition of spectrally degraded cartoon voices. *Ear & Hearing, 35*(1), 118-125.

Vasilyeva, M., & Waterfall, H. (2011). Variability in Language Development: Relation to socioeconomic status and environmental input. In Neuman, S. B., & Dickinson, D. K. (Eds.), *Handbook of Early Literacy Research, Volume 3* (pp. 36-48). The Guilford Press: New York, NY.

Vongpaisal, T., Trehub, S. E., Schellenberg, E. G., & van Lieshout, P. (2012). Age-related changes in talker recognition with reduced spectral cues. *The Journal of the Acoustical Society of America, 131*(1), 501-508.

Warner-Czyz, A. D., Houston, D. M., & Hynan, L. S. (2014). Vowel discrimination by hearing infants as a function of number of spectral channels. *The Journal of the Acoustical Society of America, 135*(5), 3017-3024.

Weighall, A.R. (2008). The kindergarten path effect revisited: children’s use of context in processing structural ambiguities. *Journal of Experimental Child Psychology, 99,* 75-95.

Wild, C. J, Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., & Johnsrude, I. S. (2012). Effortful listening: the processing of degraded speech depends critically on attention. *The Journal of Neuroscience, 32*(40), 14010-14021.

Winn, M. (2016). Rapid release from listening effort resulting from semantic context, and effects of spectral degradation and cochlear implants. *Trends in Hearing, 20*, 1-17.

Winn, M. B., Edwards, J. R., & Litovsky, R. Y. (2015). The impact of auditory spectral resolution on listening effort revealed by pupil dilation. *Ear and Hearing, 36*(4), e153–e165.

Young, G. A. & Killen, D. H. (2002). Receptive and expressive language skills of children with five years of experience using a cochlear implant. *Annals of Otology, Rhinology, & Laryngology, 111*(9), 802-810.