The Dissertation Committee for Hoyoung Yi Certifies that this is the approved version of the following dissertation:

SPEECH PRODUCTION OUTCOMES IN YOUNG CHILDREN WITH EARLY IDENTIFIED HEARING LOSS

Committee:

Courtney Byrd, Supervisor

Barbara L. Davis, Co-Supervisor

Craig A. Champlin

Rajka Smiljanic

Julia Campbell

Madhu Sundarrajan

SPEECH PRODUCTION OUTCOMES IN YOUNG CHILDREN WITH EARLY IDENTIFIED HEARING LOSS

by

Hoyoung Yi

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2018

Acknowledgements

I would like to express my deepest gratitude to my mentor, Barbara Davis, for her patient guidance on this doctoral journey. Her support, mentorship, dedication, and belief in me has led to this moment. I am also indebted to her for the emotional support that she gave me and my family throughout my time at UT. She has continuously supported me even after her retirement. I cannot thank her enough for her exceptional care to all the things of my life. I was fortunate to have Courtney Byrd. She offered insightful comments and encouragement, and questions which helped me to be prepared for further discussions with other committee members. My sincere thanks also go to Craig Champlin for giving me his knowledgeable comments and advice. His thoughtful and logical comments on data analysis reminded me to acknowledge fundamental research questions and hypotheses. I was able to complete my experiment because he generously shared his lab space with me. I am grateful to Rajka Smiljanic for providing me the opportunity to join her team and access the laboratory and research facilities outside of CSD area. Collaboration with her incited me to widen my research into different areas, especially listener perspectives for speech production. Whenever I visited her office with my random questions, she was always willing to give her expertise. I thank Julia Campbell for her brilliant advice in regard to the study of audiology. I am grateful to Madhu Sundarrajan for her expertise in both clinical and academic areas. Most importantly, I thank her for her friendship and support. Without her generous consideration, I would not have completed my PhD and clinical practice at the same time.

I am truly grateful to the participating children, their parents and adult listeners who volunteered their time to help me to complete this project. Thanks to all the undergraduate research assistants who helped with data entry. Thanks especially, Dara, Maggie, Karinne, Suzanne, and Darby for their careful work in transcription reliability.

I would like to special thank Jie Yang and Suzanne van der Feest for their friendship and support. They have always sweetly encouraged me to complete this project whenever I express my concerns and worries about my progress. They have shared their expertise in child speech and language, and showed how successfully work as a young scholar in this department. To the many others from the UT community and beyond who have provided encouragement and advice: Thank you.

I always be grateful to my parents and in-laws for their endless support and belief in me. I know they have prayed for their children who are now 7,000 miles away from them. I also thank my sister, Hojeong for being a close friend of mother and father in spite of my absence at home. I want to thank my husband, Woonyoung. I could not have completed this project without his support and love. I dedicate this dissertation to him and my lovely daughter Jueun.

SPEECH PRODUCTION OUTCOMES IN YOUNG CHILDREN WITH EARLY IDENTIFIED HEARING LOSS

Hoyoung Yi, Ph.D.

The University of Texas at Austin, 2018

Supervisor: Courtney Byrd

Co-Supervisor: Barbara L. Davis

Relevant acoustic and linguistic information is not accessible to young children with a damaged or absent sense of hearing. While infant hearing screening enables early instrumentation and intervention for children with hearing loss, they have shown a substantial amount of variability in speech production outcomes. Comprehensive analysis of speech outcomes using contemporary speech and language measures is pivotal to specifying the diversity of speech characteristics in addition to understanding the relationship of auditory input system and production output system capacities.

Speech characteristics in spontaneous speech output of children with hearing loss were examined by analyzing associations across multiple speech and language measures, relationships of speech accuracy with input factors of phonotactic probability and neighborhood density, and speech intelligibility as perceived by adult listeners. Spontaneous connected speech samples were collected for twenty-one children with bilateral cochlear implants or hearing aids. Results revealed a largely consistent chronological age-related developmental pattern across the speech measures and mean length of utterance (MLU). Most of the children demonstrated a delay in acquiring later developing consonants and more complex word forms. Based on individual variance in performance across analysis measures, environmental factors contributing to individual differences in their speech production include chronological age at receiving instrumentation and familiarity of the speech sampling context. Phonotactic probability and neighborhood density showed positive relationships with consonant accuracy in CVC word forms produced. Vowel accuracy in CVC word forms was negatively related to neighborhood density in this cohort. Intelligibility outcomes were highly associated with speech measures, suggesting the importance of a whole-word analysis approach. These results emphasize the importance of analyzing functional daily communication interactions in children fit early with HA or CI instrumentation. Findings support use of measures that can reveal the sources of reduced intelligibility.

Table of Contents

List of Tables

List of Figures

1. Introduction

Children with hearing loss do not always develop speech characteristics typically as their speech acquisition is influenced by auditory access to listening experiences beginning before birth. Children identified early with hearing loss can, however, receive hearing devices (e.g. cochlear implants, hearing aids) and interventions as early as they can. This early instrumentation has been found to positively impact achievement of age-appropriate speech and language skills, speech intelligibility, and general oral communication skills (Ching et al., 2017; Connor, Craig, Raudenbush, Heavner & Zwolan, 2006; Geers, 2004; Nicholas & Geers, 2007; Tomblin, Ambrose, Walker, & Moeller, 2014; Yoshinaga-Itano, Sedey Coutler, & Mehl, 1998; Yoshinaga-Itano, 2013). Despite early implantation and amplification, considerable individual differences in speech development are observed across children identified early with hearing loss and fit with hearing instrumentation (Kral, Kronenberger, Pisoni, & O'Donoghue, 2016; Moeller & Tomblin, 2015; Pisoni, 2010; Spencer, 2004). Comprehensive approaches to evaluating and assessing spoken language abilities in children with hearing loss have important implications for understanding individual differences in their acquisition of spoken language abilities. This information is key to development of clinical speech and language assessment and intervention tools with maximal efficacy.

Speech production and speech perception are founded in a dynamic system across the early period of development. In this period, children's abilities are dynamically changing based on growth and maturity of the biological systems underlying perception, production, and neural aspects of development that impact children's cognitive capacities. Growth in these capacities is also founded in socially mediated input from children's environment based on functional daily communication interactions (see Davis & Bedore, 2013, for an overview).

The capacity of the auditory system to receive and transmit information to the central nervous system is a critical dimension of development affecting maturing speech production capacities in young children (Sininger, Grimes, & Christensen, 2010). This ability to utilize perception and production abilities is compromised by lack of auditory capacities for hearing vocalizations from the environment in infants with hearing loss. These infants may likely have an insufficiently diverse set of experiences with both perception and production aspects of development, leading to difficulties with acquisition of fully intelligible oral speech and language capacities as well as instantiation of fully functional neural-cognitive capacities.

The auditory system, particularly relative to speech perception capacities, is shaped by access to relevant acoustic input about language in early life. Hearing infants have auditory experience during not only the first year of life but also in the prenatal period. These experiences during the pre-natal period have been shown to affect auditory learning during the postnatal period (Moon & Fifer, 2000; Lecanuet & Schaal, 1996). At birth, hearing infants show a preference for listening to their own mother's voice (DeCasper & Fifer, 1980), a skill that must be developed via in-utero exposure to the acoustic properties of her voice. Before children can demonstrate speech production capacities for language-relevant vocal output, their ability to discriminate speech sounds has been demonstrated in classic studies using differential sucking rate or conditioned head turning to familiar versus novel sounds (Eimas, Siqueland, Jusczyk, & Vigorito et al, 1971; Werker & Tees, 1984; Kuhl & Meltzoff, 1996).

In response to research information about the importance of early auditory access, American Speech-Language-Hearing Association (ASHA, 2004) reported that Cochlear Implant (CI) surgery has been approved by the Food and Drug Administration (FDA) in the U.S. for children with profound sensory neural hearing loss from 12 months of age. Infants can be fit with hearing aids (HA) by 4 weeks of age. CI's are surgically implanted electronic devices that provide direct electrical stimulation to the auditory nerve in the inner ear. The purpose of the HA is to amplify input signal sounds for the infant.

Early identification of children with hearing loss through newborn hearing screening and subsequent early application of hearing instruments via HA or CI has mitigated some detrimental effects of early hearing loss. For children with hearing loss, many studies have demonstrated significant benefits of hearing aids and cochlear implants in facilitating auditory speech perception and development of spoken language (e.g. Yoshinaga-Itano et al., 1998). The level of speech perception ability provided by hearing instruments to children with significant levels of hearing loss can potentially support typical rates of spoken language development.

In an early study, Yoshinaga-Itano et al. (1998) compared the receptive and expressive language abilities of 72 children whose hearing losses were identified by 6 months of age (earlier-identified) and 78 children whose hearing losses were identified after the age of 6 months (later-identified). Earlier-identified children with hearing loss demonstrated expressive and receptive language score on the *Minnesota Child Development Inventory* (MCDI, 1974) better than later-identified children with hearing loss. Through the early hearing screening, infants were identified early in life allowing for very early intervention and amplification or implantation resulting in improved outcomes in speech and language for children who are deaf and hard of hearing.

Contemporary practices including both broadening of early identification of hearing loss and participation in early intervention have resulted in considerable benefits for speech development and improved speech intelligibility (Ertmer & Goffman, 2011; Montag, AuBuchon, Pisoni, & Kronenberger, 2014). However, children with hearing loss have shown a substantial amount of variability in speech outcomes measured by standard tests, variables from spontaneous speech output, and speech intelligibility (e.g., Ertmer & Goffman, 2011; Montag et al., 2014; Pisoni, 2000; von Hapsburg & Davis, 2006; Warner-Czyz & Davis, 2008). For instance, Ertmer and Goffman (2011) reported variability scores of six young Cochlear Implants (CI) recipients with 2 years of device experience compared with age matched hearing peers. Children with CI who were between 35 and 61 months old produced speech targets with less accuracy and higher variability than age matched hearing children.

Individual differences in performance of children with hearing loss who have received amplification or implantation have been observed based on analysis of a wide range of behavioral output (e.g., Montag et al., 2014; Pisoni, 2000). Some children with hearing loss exhibit well developed age appropriate speech and language abilities but others do not. Cognitive processes such as perception, attention, learning, and memory in addition to neurocognitive factors are important processing capacities that help to explain the large individual differences observed among young users of cochlear implants Bergeson, Pisoni, & Davis, 2003; Kral et al., 2016). Individual differences in communication skills of young children identified early with hearing loss have been explored through analysis of multiple environmental factors including early identification and duration of hearing device and communication modality (Connor et al., 2006; Geers, Mitchell, Warner-Czyz et al., 2017; Kirk, Miyamoto, Ying, Perdew, & Zuganelis, 2002). The individual diversity in speech and language capacities for children with hearing loss has also been examined by using various concurrent speech and language measures for their spontaneous speech outcome.

Previous research on children with hearing loss has focused largely on speech outcomes or language abilities respectively (Ertmer & Goffman, 2011; Moeller et al., 2007; Nicholas & Geers, 2007; Szagun, 2001, Warner-Czyz & Davis, 2008). The current study focuses on measurement tools designed to characterize speech development for young children with hearing loss as well as considering relationships between speech and language abilities based on various assessment tools. In addition, the relationships of speech capacities to the role of input factors (word characteristics) and measurements of speech intelligibility judged by adult listeners were examined to obtain a more comprehensive picture. Both child capacities and environmental factors related to instantiation of aspects of phonology in input (e.g., phonotactic probability) and listener judgements of the intelligibility of children's spontaneous speech forms were measured in the same child to give a comprehensive view.

Thirty-two spontaneous speech samples were analyzed from 21 children with early identified hearing loss. Their speech production patterns in recognizable lexical targets occurring within functional speech and language sampling procedures were measured in terms of phoneme inventory and accuracy, as well as word- and utterance-level phonological variables. Language measures of mean length of utterance and type-token ratios were also analyzed to consider the potential for interaction between speech and language capacities as well as to include language in addition to speech variables into this comprehensive developmental picture of children identified with hearing loss early. The potential importance of phonological and vocabulary variables available in input were incorporated via analysis of neighborhood density and phonotactic probability metrics. The interaction of speech measures with input characteristics was also considered in the context of speech intelligibility measures reflecting listener judgements to consider how these factors interact in understanding speech development profiles observed. The following hypotheses were tested to further specify relationships among variables tested as they relate to understanding sources of variability in speech outcomes in young children identified early with hearing loss.

Aim 1. Examine speech characteristics (e.g., percentage of consonant correctness, word complexity) of children identified early with hearing loss.

Hypothesis: Children identified early with hearing loss will demonstrate delayed/disordered speech characteristics based on various speech measures. Children with hearing loss will show a significant discrepancy in the pace of acquisition of later emerging consonants, vowels, and word/syllable shapes relative to their chronological and to their 'hearing' age (i.e., their age since amplification).

Aim 2. Examine oral language associated characteristics (e.g., Mean Length of Utterance, MLU) of children identified early with hearing loss.

Hypothesis: Children identified early with hearing loss will demonstrate delayed language development. Language development profiles will correlate with speech output profiles.

Aim 3. Examine vocabulary input factors including neighborhood density, phonological probability, and word frequency for children identified early with hearing loss and investigate relationships of input related outcomes relative to other speech and language measures observed.

Hypothesis: Input related characteristics will be observed for children with significant hearing loss. Children with early identified hearing loss will show the effects of phonological and lexical regularities correlated with their level of speech and language skills.

Aim 4. Examine the relationship(s) of intelligibility to speech and language measures in children identified early with hearing loss.

Hypothesis: Various measures of speech and language skills obtained will predict listener judgements of speech intelligibility for children with hearing loss. In children with hearing loss, speech intelligibility judgements by young adult listeners will be correlated with the children's level of performance on speech measures, language measures, and input-related measures.

2. Literature Review

2.1 Speech characteristics of children with hearing loss (Aim 1).

The goal of assessing speech characteristics of children with hearing loss is to determine how children with early identified hearing loss compare to hearing peers and how speech measures may relate to language as well as to environmental measures of their speech development. Accurate, reliable, and sensitive measures of speech production are necessary to characterize speech outcomes and to plan for appropriate early interventions for children with early identified hearing loss who may have concomitant speech delay/disorder. Comprehensive speech measures of speech production have been mostly reported based on phoneme inventory and accuracy across measures including formal articulation tests, elicited single words and sentences, and conversational speech samples. The speech measures in this study focus on use of spontaneous speech samples to gain a picture of children's functional speech capacities in daily communication settings where they are both initiating and responding to oral communications.

Numerous previous researchers have reported individual differences on speech and language development for children with hearing loss. Their diversity of speech and language outcomes have been examined with various environmental factors such as age of identification, duration of hearing device use, and communication modality (e.g. Connor et al., 2006; Geers et al., 2017). Connor et al. (2006) examined early CI implantation benefits in addition to the advantage provided by longer implant use. They analyzed a group of early implanted children (before age 2 years and 5 months) and focused on consonant-production accuracy and vocabulary outcomes. Their results revealed a benefit for both speech accuracy and vocabulary outcomes in children with longer experience using CI. They also showed a burst of growth immediately after implantation for the group. For the current study, children identified early in life who also received early implantation or hearing aids may not have a long period between birth and instrumentation fitting. In these children, the use of chronological age may be sufficient for considering their speech and language developmental age.

Speech outcomes for children with hearing loss can be potentially affected by communication modality (Dunn, Walker, Oleson et al., 2014; Geers et al., 2017; Kirk et al., 2002). Kirk et al. (2002) studied 106 children with CI to examine the effects of age at implantation on their communication skills. The children were divided into three groups based on mean age at implantation. The authors analyzed effects of communication modality in spoken word recognition. Children who used oral communication (OC) achieved significantly higher word recognition scores than the children who used total communication (TC) after controlling for length of device use and age at implantation. However, there was no significant effect of communication modality for their language development scores on language test results using their preferred communication mode.

Dunn and colleagues (2014) considered communication modality in examining the long-term effect of age at implantation on speech perception, language, and reading performance in 83 children from age 5 years to 13 years with CI. Their results revealed that children who attended educational programs emphasizing Oral Communication (OC) performed better on speech perception testing than children who used Total Communication (TC). Relative to expressive language, the OC group communication showed better scores than TC users even though the group difference was not statistically significant. One recent study also added the influence of degree of sign language exposure in speech recognition, spoken language, and reading skills for 97 children from ages 5 years to 11 years with CIs (Geers et al., 2017). Children with long-term sign exposure produced significantly less intelligible speech than those children in the short-term exposure or the no sign exposure groups. Children without early sign exposure also achieved significantly better language scores than both sign-exposed groups.

2.1.1 Phoneme inventory and accuracy

Understanding of speech acquisition in young children has relied on understanding both the inventory of phonemes they have developed (regardless of accuracy) and children's accuracy in deploying sounds in their inventory of spoken word forms. Moeller and colleagues (Moeller et al., 2007) studied young children with hearing loss longitudinally. They examined syllable complexity and consonant inventories from speech samples for children with hearing loss and compared them to age matched hearing peers' vocalizations. The authors evaluated twelve early identified children with bilateral sensory neural hearing loss from 10 to 24 months of age and compared them to 21 hearing peers over a period of 14 months. Speech outcomes for the children were analyzed based on 30-minute mother-child interactions in a laboratory playroom setting. The *Goldman-Fristoe*

Test of Articulation-2nd edition (GFTA-2, Goldman & Fristoe, 2000) was also administered individually to each child in its standard format at 36 months of age. The consistent production of canonical syllables defined as rhythmic consonant-vowel syllables was observed to emerge in output inventories later in the children with hearing loss than in their hearing age matches. The children with hearing loss produced smaller inventories of consonants and less complex syllable shapes relative to hearing peers. The analysis of consonants by place of articulation revealed that bilabial sounds increased at similar rates in both groups and converged by 22 to 24 months of age. In terms of manner of articulation, the two groups demonstrated similar developmental patterns for non-fricative sounds, but delayed fricative production was observed in the output of children with hearing loss. The gap between the two groups increased with age for fricative and affricate sounds. The authors also reported relationships between early syllable structure characteristics at age of 2 years and GFTA-2 scores at age of 3 years. The results showed a negative correlation between limited production of syllables with consonants and GFTA-2 scores indicating children who used a high number of syllables without consonants at 2 years of age had lower GFTA-2 scores at age 3. The results also demonstrated a positive correlation between using CV syllables at 10 to 12 months and GFTA-2 cores. That indicates children with a higher number of CV syllables at 10 to 12 months show higher GFTA-2 scores.

Warner-Czyz and Davis (2008) explored consonant and vowel inventories and phonological accuracy and error patterns in early words of four children with cochlear implants (CI) compared to hearing peers matched by vocal development age (i.e., meaningful word onset). Speech samples for the four children (mean chronological age at study onset: 24 months) with profound hearing loss were collected monthly for six months following the onset of meaningful words. Phonetic accuracy analyses were based on two dimensions: ‘horizontal’ (consonant place and vowel front-back characteristics) and ‘vertical’ (consonant manner, vowel height characteristics). Both the children with CI and hearing peers showed similar acquisition patterns evidenced by most frequent occurrence of labial and stop consonants. However, consonant productions in the hearing group were more accurate than in the group of children with CI. Vowel productions in both groups were observed to have similar segmental accuracy patterns. Both groups improved their level of consonant accuracy over time, but the hearing participants performed significantly more accurately than the children with CI in both early and later sessions. In terms of error patterns, level of accuracy for fricatives shifted from incorrect productions to correct productions in the hearing children and from omissions to partially correct productions in the children with CI.

Ertmer and Goffman (2011) assessed speech production accuracy and variability scores of six young CI recipients between 35 and 61 months old with 2 years of device experience compared with age matched hearing peers using words from the *First Words Speech Test* (FWST; Ertmer, 1999). The FWST consists of four sets of words based on the ages at which consonants are correctly produced by 50% of typically developing children. Ertmer and Goffman’s results indicated lower accuracy of speech production for consonants and vowels in children with CI than hearing peers. The children with CI produced speech targets with less accuracy and higher variability for the imitation tasks of the words on the FWST in a comparison with their hearing peers. The children with CI showed relatively high accuracy for the early-acquired phonemes (e.g. stop consonants) than for later emerging consonants (e.g. fricatives). They also showed higher variability for later emerging consonants than for early emerging consonants. Vowel accuracy was also higher in hearing children than in CI recipients. The authors concluded that two years of CI use may not be sufficient for children with hearing impairment to match the pace of typical development of spoken language abilities relative to hearing peers.

Eriks-Brophy and Whittingham (2013) compared speech errors between twenty-five preschool children with hearing loss at ages 36, 48, and 60 months and thirty-five hearing children. The children with hearing loss wore bilateral hearing aids (10) or unilateral cochlear implants (15). The authors used standardized tests; *The Goldman-Fristoe Test of Articulation*-2nd edition (GFTA-2, Goldman & Fristoe, 2000) and *Khan-Lewis Phonological Analysis*-2nd edition (KLPA-2, Khan & Lewis, 2002). Results from the GFTA-2 showed significant differences in the performance of the children with hearing loss and the hearing children. However, 72% of the children with hearing loss had a standard score at or above the level expected for their chronological age at the time of their last assessment. Only four consonants (i.e., /s/, /v/, /dʒ/, and blends with /w/) appeared to fall outside of the typical developmental range for phoneme acquisition at 5 years of age. The results from the KLPA-2 showed that children with hearing loss used more reduction and voicing processes as well as showing greater variability in performance on sounds tested as compared to the hearing children, highlighting the issue of variability as well as occurrence of non-developmental error patterns. Non-developmental patterns (e.g. stridency deletion) which were unscored by the KLPA-2 were also observed in the children with hearing loss. However, phonological patterns in the children with hearing loss resembled those of their peers, and 17 out of 25 (68%) children with hearing loss had standard score at or above 85 on the KLPA-2.

Flipsen and Parker (2008) examined phonological process use in longitudinal conversational speech samples from six young children with cochlear implants (CI) at ages ranging from 45 to 95 months. The analysis of phonological processes is based on rule-governed simplifications and systematic speech errors relative to adult speech patterns (Ingram, 1976). The index of phonological process has provided a more comprehensive description of the systematic patterns in children's acquisition of phonology in comparison to focusing on individual sound errors (Roberts Burchinal, & Footo, 1990). Flipsen & Parker (2008) examined forty speech samples and analyzed to identify phonological processes. Results demonstrated both developmental (i.e. final consonant deletion, cluster reduction, devoicing of stops, stopping, fronting, liquid simplification, and gliding) and non-developmental (i.e. initial consonant deletion, glottal stop substitutions, backing, vowel substitutions, vowel neutralizations, and diphthong simplification) patterns for children with CI. However, developmental patterns tended to occur with greater frequency than non-developmental patterns in the cohort of CI children studied.

Overall, these researchers have assessed speech outcomes for children with hearing loss based on inventories and phonological accuracy of consonants and vowels compared to their hearing peers. In order to obtain a more comprehensive picture of a child's phonological system, more speech measures need to be evaluated to assess the efficacy of cochlear implants or hearing aids in detail. The more comprehensive approach with multiple speech measures in addition to language measures provides professional guidance for speech and language intervention for children with speech disorder or delay associated with hearing loss. However, the results from the typical speech analysis only report quantitative differences in speech patterns between children with hearing loss and hearing peers. Several contemporary measures are available that might support a more finely grained understanding of speech production output patterns for planning intervention.

2.1.2 Speech characteristics in whole word approaches

Whole word measures provide another way to evaluate speech outcome patterns focused at the word level in phonological acquisition. The assessment of phonology in the range of sound classes, and syllable and word structures is recommended especially for young children in the early stages of phonological development, rather than analyzing the accuracy of individual phonemes (Stoel-Gammon, 2010; Ingram and Ingram, 2001, Vihman & Croft, 2007). Ingram & Ingram (2001) proposed a whole-word approach to assessment based on the hypothesis that children are word-oriented instead of segment-oriented in early acquisition (Ferguson & Farwell, 1975).

Ingram (2002) developed a measure that evaluates whole words instead of segments, the Phonological Mean Length of Utterance (PMLU). PMLU rewards the child not only for producing correct consonants, but also for producing all the segments that compose the target words. To examine how well the child matches the target words attempted, Proportion of Whole-word Proximity (PWP) was also introduced by Ingram (2002). PWP is a ratio between child PMLU and target PMLU. Another whole word measure is the Word Complexity Measure (WCM). This measure was developed by Stoel-Gammon (2010). The WCM highlights the presence of phonological characteristics focusing on the child's inventory of word, syllable and sound class parameters.

Several studies have applied whole word measures to characterize phonological development in hearing children (Burros & Goldstein, 2010; Cummings & Karson, 2015; Watson & Terrell, 2012). However, few studies have reported PMLU and PWP for children with hearing loss (e.g., Schauwers, Gillis, & Govaerts., 2008). Schauwers and colleagues compared PMLU and PWP between Dutch children with hearing loss and Dutch hearing children aged from 2 years 1 month and 2 years 6 months based on their spontaneous speech output. Results revealed that hearing children showed higher PMLU (6.3) and PWP (0.8) than children with cochlear implants, PMLU (4.5), and PWP (0.7). The current study will include the whole word measure to characterize speech production for children with hearing loss as well as in correlating with other types of measures gathered to help predict which measures are the most likely to indicate children in need of further support to acquire age appropriate speech output.

More recently, a linear relation between consonant correctness (i.e. Percent Consonant Correct, PCC, Shriberg et al., 1997) and word complexity (i.e. PWP) has been proposed to distinguish speech delay from speech disorder in hearing children (Cumming & Larson 2015; Ingram, 2015). Using the linear relationship between PCC and PWP scores from typically developing children, a child's PCC score was found to predict the range of possible PWP scores. Preliminary data with typically developing 2-year old children revealed that the relationship between PCC and PWP is influenced by word complexity measures such as word length and syllable complexity, especially consonant clusters (Ingram, 2015). Greater differences between PWP and PCC were found in clusters over singletons as well as for multi-syllable words over monosyllable words. On the other hand, a preliminary data analysis of children with speech sound disorders identified two distinct patterns. Children in a 'Linear' group demonstrated a linear increase in word production accuracy similar to typically developing peers. Alternatively, children in a 'Nonlinear' group produced more complex words with greater accuracy than some of their less complex words. Ingram (2015) identified children who showed linearity between PCC and PWP as 'speech delayed' whereas children in Nonlinear group were characterized as 'speech disordered'. This construct shows some promise for early identification of hearing impaired children who may need additional support to acquire age appropriate speech intelligibility.

Anderson and Cohen (2012) compared WCM scores between typically developing children and children with profiles of speech delay or disorder. Ninty-seven typically developing children between three to five years of the age and age-matched 8 children speech delay or disorder were compared in their WCM scores. Results revealed a striking difference between children with typically developing speech and children identified with speech delay or disorder. The range of mean WCM ratio scores for the typically developing group was 0.84 to 0.91 whereas the eight children with delayed or disordered speech development showed WCM ratio scores from 0.42 to 0.68.

To date, there is no study of hearing impaired children using the WCM. Results of previous research indicate that children with hearing loss demonstrate more difficulty than hearing children in developing later emerging sounds such as fricatives (Moeller et al., 2007; Warner-Czyz and Davis, 2008; Flipsen & Parker, 2008; Eriks-Brophy et al., 2013). Therefore, the study prediction is that children with hearing loss will score significantly lower on the WCM than typically developing children, thus identifying a potential measurement tool for this population. In addition, with analysis of the relationships between PCC-R and the parameters of the WCM, we will examine which later emerging sound can explain most consonant errors in children with hearing loss.

Whole word measures are another method to further examine speech characteristics of children with hearing loss. However, few studies have reported PMLU / PWP or WCM for children with hearing loss. Use of the whole word approach and examining relationships between PCC-R and whole word measure outcomes for children with hearing loss can help to determine the need for clinical support for them in acquiring intelligible speech outcomes.

2.2. Language characteristics of children with hearing loss (Aim2).

It is also of value to consider potential relationships and the prediction value of language based measures of development in young children with hearing loss to understand whether these measures correlate with speech production outcomes and should be incorporated routinely in assessment and intervention protocols. Two types of measures in broad use are Mean Length of Utterance (MLU, Brown, 1973) and Type Token Ratio (TTR, Templin, 1957). These measures assess morphological and syntactic complexity (MLU) and vocabulary use and size (TTR).

Previous studies have revealed that early identified children with CI or HA show delay in their morpho-syntactic development based on MLU. However, they also show improvements in production of longer and more complex sentences over time. Some children catch up with their hearing peers based on measures of morpho-syntax ability (Faes, Gills & Gills, 2015; Flipseon, 2014; Koehlinger, van Home, & Moeller, 2013; Nicholas & Geers, 2007; Szagun, 2000 & 2001).

Nicholas and Geers (2007) investigated the benefits of earlier and longer CI use in seventy-six children with hearing impairment at ages 3.5 and 4.5 years. Their longitudinal study used multiple measures (e.g., a standardized test for language development; the *Preschool Language Scale—Third Edition*, Zimmerman, Steiner, & Pond, 1992, & MLU) to assess language abilities to compare language abilities within CI groups. Spontaneous language sampling sessions were conducted twice for the CI group at the age of 3.5 years and 4.5 years to enable analysis of multiple language measures including MLU to measure syntactic development. Results showed that MLU's for children with CI increased with longer implant experience. Children who were implanted at 12 months of age caught up with their hearing peers by age 4 years and 6 months. However, MLU showed very little change between the 3.5 and 4.5 year old groups in this study. MLU may be more sensitive to language growth for broader age ranges for young children.

Flipsen and Kangas (2014) examined MLU values in a longitudinal data analysis of 6 children with CI (45 samples) and in a cross- sectional dataset of 4 children with CI. Across all 49 samples chronological age ranged from 3 years 9 months to 10 years 6 months. Results showed increased MLU values over time even though the children with CI demonstrated delayed onset of age appropriate MLUs.

Children with hearing aids also showed delayed development in their morpho-syntactic abilities as measured by MLU. Koehlinger et al. (2013) measured MLU for 145 children with hearing loss and 40 hearing children at the age of 3 and 6 years. The authors concluded that MLU's for children with hearing loss lagged behind with their hearing peers, but some of children with hearing loss in their cohort performed comparably to their hearing peers.

In other languages, children with hearing loss also have shown consistent results that MLU is lower in comparison to typically developing hearing peers but increases with longer CI use (German: Szagun, 2000 & 2001, Dutch: Faes et al., 2015). Szagun's (2001) investigation of language abilities in 22 German-speaking children with cochlear implants (CI) included MLU. Spontaneous speech samples were collected over 27-36 months during the one-word stage. MLUs derived from spontaneous speech samples and productive vocabulary by parent report from questionnaires were longitudinally analyzed as outcome measures. Age at implantation ranged from 14 to 46 months, and children's values with CI were compared to hearing peers matched at the study outset for vocabulary level. Progress on MLU for children with CI was slower than hearing peers. The CI group showed considerable individual differences in MLU. However, even though the CI group showed smaller vocabulary size and lower MLU values than the hearing group, their language abilities improved over time.

Szagun (2000) also studied the efficacy of TTR with young children with hearing loss. She collected spontaneous speech samples for 10 young German-speaking children with CI. She used TTR to measure each child's vocabulary quality. The results of the type/token ratio for the children with CI at the chronological age of 1 year 6 months indicated good distribution of vocabulary types relative to hearing children. She emphasized that children with CI may have better vocabulary size due to their high quantity and quality of parent input in comparison to hearing children in the initial stage of language acquisition.

In a longitudinal case study (Ertmer, Strong, & Sadagopan, 2003) examined the emergence of a wide range of oral language a deaf child with CI. The authors measured TTR for the child at 19, 20, 30, 36, and 42 months. They compared the child's TTR values to values for hearing children from Templin (1957). Results indicated that the child produced significantly fewer different words and used them more infrequently than children in the normative study of hearing children by Templin (1957).

Some studies demonstrate association between phonological and morphological production patterns in children with hearing loss (Bow, Blamey, Paatsch, & Sarant, 2004; Moeller, McCleary, Putamen et al., 2010; Koehlinger, Van Horne, & Moeller, 2013). Children with hearing loss are likely to be less accurate in producing phonological forms especially final fricative consonants (e.g. /s/ and /z/) that are implicated in English morphology. Children's patterns of lower performance on word final position articulations can affect isolating morphological production difficulties from speech sound production difficulties. Koehlinger et al. (2013) evaluated grammatical development for 3 and 6 year old children with hearing loss by using their spontaneous speech production. Their results showed that children with hearing loss had a shorter MLU than hearing children at both 3 and 6 years of age. In addition to the main finding of the study, the authors also revealed that articulation scores correlated well with morphology production abilities. Their results indicate that articulation skills may play a role in the degree of accuracy observed and should be considered as a potential factor in grammatical abilities in clinical settings.

There are many studies that examine language skills for children with hearing loss. However, few studies focus on both speech and language outcomes and the relationships of language skills to speech abilities in the same children with hearing loss. Based on previous research, the current study hypothesizes that the children with hearing loss will demonstrate delayed language development based on studies implementing currently available measures. This study also predicts that children's morphosyntactic skills are associated to speech outcomes based on the four speech measures evaluated.

2.3 VOCABULARY INPUT FACTORS ON SPEECH CHARACTERISTICS IN CHILDREN WITH HEARING LOSS (AIM 3).

Despite positive language outcomes for many early identified children with hearing loss (e.g., Nicholas & Geers, 2007), some children may have difficulty in storing representations of phonological and lexical forms (Han, Storkel, Lee, & Yoshigana-Itano, 2015; Schwartz, Steinman, Mystal, & Houston, 2013). Based on their lack of auditory access to the speech input signal in daily functional communication situations, children with significant hearing impairment may be less effective in detecting patterns in language input, and in reproducing sound patterns that match what they can access auditorily in salient words.

The specific characteristics of words such as phonotactic probability and neighborhood density are the variables analyzed in psycholinguistic research to understand the structure and organization of lexical representations. Phonotactic probability is one aspect of language processing, defined as the likelihood of occurrence of a sound sequence in a given language (Vitevitch & Luce, 1999). Phoneme sequences that are highly likely to co-occur in the language are referred to having high phonotactic probability, whereas phoneme sequences with a low likelihood of co-occurrence are labeled to low phonotactic probability. For example, /k/ and /ka/ in "car" occurs more frequently in English than the /p/ and/ or /pu/ in "push.

Various studies of children developing typically have established that phonotactic probability influences word learning (e.g., Storkel, 2001 & 2003), and accuracy of speech production (Zamuner, 2009; Edwards, Beckman & Munson, 2004). Storkel (2001) studied phonotactic probability effects on learning novel words for thirty-four typically developing hearing children (from ages 3 years 2 months to 6 years 3 months). Their results revealed that children learn novel nouns (non-words) with high phonotactic probability more easily than non-words with low phonotactic probability. Her subsequent study (Storkel, 2003) also revealed consistent results with common sound sequences being learned more rapidly than rare sound sequences in verb-learning (non-words) for 34 typically developing hearing children (from ages 3 years 2 months to 6 years 4 months) with the same word learning paradigm as in her previous study.

Edwards at al. (2004) examined phonotactic probability effects on speech production accuracy and fluency in non-word repetition tasks for 104 typically developing hearing children. They used two and three-syllable non-words for the repetition tasks and analyzed segmental accuracy, substitution errors, and reaction time on the tasks. Their results showed that children repeated consonants and vowels more accurately in common sound sequences than low frequency sequences. In addition, substitution errors on low frequency sequences (low phonotactic probability) resulted in higher frequency sequences, and children produced shorter durations in the high-frequency sequences.

Zamuner (2009) also used a non-word repetition paradigm for thirty-two younger Dutch learning children from ages 2 years 2 months to 2 years 8 months. Her results indicated that young children also produced words with high phonotactic probability more accurately than words with low phonotactic probability. These results of previous studies indicate that children's speech sound representations highly linked to the contexts the inputs occur in words in the lexicon (e.g. Edwards et al., 2004).

Another aspect of lexical processing, neighborhood density, is defined by the number of words that differ from a given word by one phoneme. Words with dense neighborhoods contain many neighbors (e.g., "tool": cool, pool, rule, wool etc.) while words with sparse neighborhoods include few neighbors (Vitevitch & Luce, 1999). Neighborhood density influences speech production in typically developing children (Newman & German, 2002; Sosa & Stoel-Gammon, 2012), and word learning (Heisler & Goffman, 2016; Storkel & Lee, 2011). Newman & German (2002) examined 267 typically developing hearing children and 53 hearing children with word-finding difficulties. The age range of the children was from 7 to 12 years. They used open-ended sentences to elicit target words ranging from low to high frequency of occurrence. Children with and without word-finding difficulties produced words in sparse neighborhoods more accurately than words in dense neighborhoods. This finding implies that one source of naming failure for both typically developing children and children with word finding difficulties is competition among similar sounding words.

Storkel and Lee (2011) examined the influence of neighborhood density on lexical acquisition in one of experiments they conducted. Storkel and Lee used CVC novel words to define sparse and dense neighborhoods. The referent identification tasks were completed by typically developing children and compared their performance in the baseline to following training. They found a low neighborhood density advantage for word learning in thirty-one 4 years old. Children learned words with sparse neighbors more accurately than words with more neighbors. Heisler & Goffman (2016) showed that sixteen typically developing hearing children (mean of age: 4 years and 5 months) produced phonemes in words with high density neighborhoods less accurately. Thus, the same competitive inhibition effects of words with denser neighborhood were observed in both word production and word learning.

Most of studies to investigate the influence of phonological and lexical factors in children developing typically have employed carefully controlled experimental tasks (e.g., Storkel & Lee, 2011). However, experimental paradigms may have limits relative to validly reflecting the natural process of word production and phonological acquisition. Sosa & Stoel-Gammon (2012) collected spontaneous speech production in free play settings for 15 typically developing hearing children between the ages of 2 years and 2 years 5 months. They elicited five productions of monosyllabic target word during various activities such as playing, reading and naming. They found words with dense neighborhoods had higher proximity (Whole Word Proximity; PWP, Ingram, 2002), but no phonotactic probability advantages in PWP.

Relative to language capacities, a majority of studies of young children with hearing impairment to date have focused on broad language abilities based on standardized tests or descriptive language tests that are not normed (e.g., Nicholas & Geers, 2007; Eriks-Brophy et al., 2013). Even though standardized tests give basic information about speech and language development relative to normative data, these tests may overlook critical details about speech and language representations that potentially impact test performance and functional communication outcomes in young children with hearing loss. It is important, in addition, to explore word characteristics relative to the various levels of performance on speech and language measures described above for children with hearing loss (e.g., PCC-R, Shriberg, Austin, Lewis, McSweeney, & Wilson, 1997). Children who show persistence of delay in speech and language development after receiving hearing instrumentation may have different representations relative to perceptual processing of input language as accessed by analyses of phonotactic probability and neighborhood density. These types of differences should optimally be included in assessment and intervention protocols if they contribute to understanding children with hearing impairment's processing and subsequent storage of words in input.

Schwartz and colleagues (2013) predicted that children with CI may overuse a strategy of phonotactic probability based on their preliminary data. They introduced their 5-year research project and reported preliminary data across their proposed experiments for children with CI and hearing peers. For example, they conduced semantic and phonological priming tasks with 13 children with CI and 27 hearing children. Their preliminary data showed greater inhibition to words following shared onset real words and to non-words following high PP non-words. Children with CI tended to judge non-words with high PP as real words. This indicates a potential strategy of overusing phonotactic probability as an indicator of real words.

A wide range of individual differences in performance of children with hearing loss who have received amplification or implantation (e.g. Yoshinaga-Itano, Baca, & Sedey, 2010) may potentially be explained by the way children with CI or HA use lexical characteristics to learn new words before and/or after implantation. Lack of hearing experiences can impact organization of lexical representations and contribute to delayed or disordered speech and language development (Han et al., 2015).

Han et al. (2015) investigated word characteristics relative to vocabulary size for children with hearing loss. They examined effects of phonotactic probability, neighborhood density, word length, and word frequency on words known by 14 children with CI in a subset of data from a longitudinal study of hearing loss by Yoshinaga-Itano et. al. (2010). They analyzed longitudinal outcomes from two normed vocabulary tests: the *MacArthur Communicative Development Inventories* (M-CDI, Fenson et al., 2006) words and the *Expressive One-Word Picture Vocabulary Test-3* (EOWPVT-3, Brownell, 2000). The word characteristics they analyzed were based on the target words instead of the actual word as the child produced it. Based on the results of the EOWPVT-3, the 14 children with CI were divided into three groups. These groups included 'Gap Closer' (5 children: improved performance from the age of 3 to the age of 7), 'Age Equivalent' (5 children: maintained an age equivalent or higher performance between 2 and 7 years of age), and 'Delayed' (4 children: borderline-to-delayed language performance at 3 years of age and performance below the Age Equivalent group at 7 years of age). The authors found no significant effects of phonotactic probability or word length, contrasting with previous findings from typically developing children (e.g. Storkel & Lee, 2011). However, neighborhood density showed a robust effect across groups. Children with CIs were found to be more likely to learn words with dense neighborhoods. The neighborhood advantage is consistent with the word production in typically developing children (e.g. Storkel, 2001 & 2003).

To date abundant studies have examined input characteristics and lexical representation in typically developing children with intact hearing acuity. High phonotactic probability and sparse neighborhood density advantages have been revealed relative to speech production and word learning for these children in previous studies. However, most data were collected in controlled experimental conditions. These experimental studies used lexical and phonological values dichotomized into high and low phonotactic probability and neighborhood density. Experimental designs dividing dependent measures into high and low values may disregard the whole spectrum of phonological and lexical representation of speech and language that children experience in functional daily input. Few studies have examined details about speech outcomes relative to input word characteristics beyond children’s performance on standardized tests. In addition, few studies focus on input factors related to speech and language outcomes in children with hearing loss. One important extension in this area of research is the application of these form characteristics to more naturalistically collected speech samples, such as spontaneous speech or elicited probes. Although this preliminary investigation only included CVC word forms due to a temporally unavailable online calculator (Storkel, personal communication, July 2017), it is possible to provide a continuous measure of phonotactic probability and neighborhood density to examine the degree of association between these form characteristics and speech patterns.

2.4 RELATIONSHIP OF INTELLIGIBILITY WITH SPEECH MEASURES IN CHILDREN WITH HEARING LOSS (AIM 4).

Intelligibility refers to listener understanding of speech in oral language based input (Hodge & Whitehill, 2010). It is a functional indicator of oral communication competence reflecting the talker's ability to convert language to a physical signal and the listener's ability to perceive and decode the signal to recover the meaning of the talker's message (Hodge & Whitehill, 2010). Measures of intelligibility have been used to help determine the need for intervention and provide an index of the severity of a speech disorder. Previous studies have indicated that children with normal hearing achieve adult-like intelligibility by the age of 4 years (Weiss, 1982; Gordon-Brannan &8 Hodson, 2000; Chin, Tsai, & Gao, 2003, Flipsen, 2008). Based on unfamiliar listeners who write down the words they hear a child say via audio recording of the child's conversations with an adult, the expected range of intelligibility is 26% to 50% by age 2 years, 71%-80% by age 3 years, 81%-90% by age 3 years 6 months, and 90%-100% by age 4 years (Weiss, 1982). Gordon-Brannan & Hodson (2000) recommend that speech language intervention is indicated for 4 year old hearing children with speech intelligibility scores of 66% or less (2SD below the mean) by unfamiliar listeners who respond the percentage of words orthographically transcribed correctly from a connected speech sample.

The improvement in speech intelligibility of children measured from before and after CI or HA device fitting provides indirect evidence of the efficacy of sensory aids (i.e., whether the children receive full benefit from their sensory aids). Previous studies of the intelligibility of children who use CIs or HAs indicates that they have showed improvement in their speech intelligibility after CI or HA fitting (Chin et al., 2003; Tobey, Geers, Brenner, Altuna, & Gabbert, 2003; Peng, Spencer, & Tomblin, 2004; Ertmer, 2007).

Chin, Tsai and Gao (2003) administered the Beginners' Intelligibility Test (BIT; Osberger, Robbins, Todd, & Riley, 1994) for both 51 children with CIs (age at time of testing ranged from 2 years 8 months to 10 years 8 months) and 47 children with normal hearing (age at time of testing ranged from 2 years 6 months to 6 years 9 months). The BIT requires that children repeat short and meaningful English sentences with relative objects and pictures. Intelligibility is measured by having small groups of normal-hearing adults transcribe the children's speech. For the group of children with normal hearing, the authors reported a result consistent with previous studies for the development of speech intelligibility by chronological age. Between the age of 2 and 4 years, their participants showed rapid increases in connected speech intelligibility that have been associated with increasing chronological age in hearing children. At the age of 4 years, the hearing children demonstrated adult-like intelligibility by achieving a range of 87% -100% correct on the BIT. At the age of 4 years, the children with CIs had a mean intelligibility score of 17 %, but by the age of 7 years, their intelligibility had increased to only 62%. The authors concluded that children with CI were significantly less intelligible than hearing peers based on these intelligibility data.

Castellanos, Kronenberger, Beer et al. (2014) studied 58 children with CI. Their mean age at initial testing was 4 years on the BIT and 13 years at long-term follow-up visits. The authors proposed that that there may be a relationship between preschool speech-language skills (e.g. speech intelligibility) and later speech-language outcomes (e.g., standardized language scores). They examined speech intelligibility with the BIT (Osberger et al., 1994) and vocabulary skills using *Peabody Picture Vocabulary Test 3rd and 4th editions* (PPVT-3, PPVT-4; Dunn & Dunn, 1997 & 2007). The PPVT is a one-word receptive vocabulary test, which requires children to choose one of four pictures matching a spoken word. They also used digit span tasks (short-term memory: forward digit span, short-term memory & executive function: backward digit span) and *Clinical Evaluation of Language Fundamentals Fourth Edition* (CELF-4; Semel, Wiig, & Secord, 2003) for evaluating the children's performance at the later testing. Results revealed that speech intelligibility (BIT) outcomes when children with CI were preschool age were strongly related to long-term speech and language outcomes, significantly predicting language (PPVT-4 and CELF-4) and forward short-term verbal memory capacity (Digit span forward). Overall, results indicated that a child's speech intelligibility score may identify and help to plan interventions for very young CI use users who may be at long-term risk for speech and language outcomes.

While the BIT broadly assesses speech production skills, it does not assess speech production skills in spontaneous conversational speech, which is the most socially valid level to measure intelligibility. For the BIT, the children were required to produce an imitative response to the examiner's spoken model based on naming objects and pictures. In this procedure, cognitive skills need to be activated, including verbal working memory that requires encoding, storage, and retrieval of phonological, lexical and syntactic information in addition to speech production skills. Speech production in spontaneous conversations occurs at different linguistic levels. Speech performance in words, sentences, conversation, and narratives can be included in intelligibility measurement outcomes. The current proposed study will focus on measurements of intelligibility in spontaneous speech production by young adult listeners listening to children with hearing loss. The goal is to understand the relationships of functional intelligibility in young children with other clinical measures that may help in planning assessment and intervention protocols.

Flipsen (2008) reviewed studies targeting intelligibility of spontaneous conversational speech produced by children with CI. The 10 studies he reviewed used several rating scales. For example, seven of the 10 studies used the Speech Intelligibility Rating Scale (SIR, Allen, Nikolopoulos, Dyar, & O'Donoghue, 2001). The SIR uses a scale from 1 to 5 to rate speech intelligibility. The articles were summarized to suggest that intelligibility outcomes are correlated with amount of implant use and the age of implantation. Results showed a clear trend for intelligibility to improve over time with faster progress for those who received their implants the earliest. Flipsen concluded that

CIs provided much better outcomes compared to older intervention approaches, at least relative to the intelligibility of spontaneous conversation. It should be noted that the SIR scaling procedure has several drawbacks (Ertmer, 2011). Each listener may have different subjective internal criteria when rating speech samples. For example, a 5-point scale can represent a continuum between the lowest and highest levels of intelligibility. A 5-point for intelligibility could have a different meaning for diverse listeners. One listener with very conservative approach may give a 5 point for intelligibility whereas the other listener may tend to give a 5 rating for relatively less intelligible speech.

In terms of relationships between intelligibility and speech accuracy, Chin and Kuhns (2014) examined relationships among speech intelligibility and speech characteristics in 10 children with CI between 8 years 5 months and 10 years 5 months. They used the BIT (Osberger et al., 1994) to obtain an intelligibility score and transcribed BIT sentences for speech errors (e.g. PCC: Schriberg et al., 1997). Their results showed no significant correlation between BIT and PCC scores (r = 0.461). Only the percent words omitted score showed a significant negative correlation with the BIT (r = -0.683). This outcome indicates that children who omitted more words in sentence repetition showed lower intelligibility scores based on the BIT.

Ertmer (2010) also studied the relationship speech intelligibility based on BIT scores and articulation scores from the GFTA-2 (Goldman & Fristoe, 2000) for 44 children with hearing loss ranging in age from 2 years 10 months to 15 years 5 months. Based on a child's actual production from the GFTA-2, total target correct, vowel correct, initial /

medial/ final consonant correct, consonant cluster correct, and initial consonant correct without clusters were examined to predict intelligibility. Ertmer's results revealed that the total target consonants score was the best single-variable predictor of intelligibility. However, the total target consonant score did not fully explain speech intelligibility (R^2 = .24: i.e. 24 % intelligibility was explained by the total target consonants score). Based on his findings, the author emphasized the importance of assessing connected speech intelligibility because word-based articulation test scores do not provide reliable estimates of connected speech intelligibility.

Summary and Rationale

To understand fully the reasons for differing outcomes in achievement of age appropriate speech and language skills in children with hearing loss, a comprehensive view of both input and output capacities may be extremely important. While there are many studies of speech output in these children with hearing loss, most focus only on phoneme characteristics, use of morphology, input word characteristics or intelligibility respectively rather than on a broad profile of these features within the same child. While formal tests and finely controlled experimental responses give some indication of the child's capacities, they do not show how the child functions in daily functional communication; what kinds of sound patterns, language capacities, vocabulary types and overall intelligibility to others he or she displays.

The current study examined relationships among various speech, language, and input related characteristics based on measures evaluated in the context of intelligibility outcomes for this cohort of children with early identified hearing loss. While these diverse factors have been tested in studies of each area of focus, no studies exist that test all these relevant areas in the same children and correlate them with one another. In addition, most previous studies used either experimental data or observational data only. This comprehensive approach is needed to understand the most critical aspects involved in development of children with hearing loss to support efficacy of assessment and intervention protocols.

3. Methods

3.1 RESEARCH DESIGN

The first two aims evaluated speech and language characteristics of children with hearing loss by using various speech and language measures to evaluate their spontaneous output in a functional communication setting. Overall effects of chronological age, hearing device and communication modality were examined by using mixed effects linear regression models for each speech or language measure. To consider the effect of chronological age on development of speech and language, children were descriptively compared related to normative values for their chronological age based on results from previous studies with typically developing hearing children for each speech and language measure. Relative to hearing devices, the children were divided into two groups: bilateral Cochlear Implants (CI) and Hearing Aids (HA). Participants were also classified according to their communication modality between Talking Only and Total Communication groups. Children in the Talking Only group use only oral communication. Children in the Total Communication group used both spoken language and some of American Sign language.

For aim 3, which was to investigate input characteristics including phonotactic probability and neighborhood density related to speech outcomes of children with hearing loss, only CVC word forms were selected for analysis due to a temporally unavailable online calculator (Storkel, personal communication, July 2017). All the variables of interest for the CVC words were derived from Storkel (2013). Mixed effects linear regression models were used to examine PP and ND on speech characteristics.

For study aim 4, an intelligibility test was conducted with young adult participants. Linear mixed logistic regressions were applied to correct and incorrect responses for words in the intelligibility study. To examine whether intelligibility judgements are related to values achieved on speech measures, four linear mixed logistic models were applied due to multicollinearity among the speech measures. To evaluate which speech measure was most related to the intelligibility judgments by adult listeners, exponential estimates for the speech measures across the four models were descriptively compared.

3.2 PARTICIPANTS

There were two types of participants as follows: children for the speech and language measures and young adults for the intelligibility tests. Eight out of 21 children participated in the study at least twice, thus the total number of speech samples analyzed was 32. Forty-six young adult listeners with normal hearing participated in the intelligibility test.

3.2.1 Children with hearing loss

The current study included twenty-one children with Hearing Loss (HL) (32 speech samples) from the total data collection (25 children, 37 speech samples) that occurred as a part of the aural habilitation summer camp at the University of Texas at Austin Speech and Hearing Center (UTSHC) between 2012 and 2016. The remaining children wore different types of hearing devices (e.g. bone conduction) or had no background information (e.g. communication modality). All the caregivers of the children gave written consent and they were paid ten dollars for each test that their child participated in. For the 13 children who participated in the cross-sectional data analysis, only one data collection was included per child. Eight children with longitudinal data participated in data collection at least twice across the period of data collection. Thirty-two speech samples were analyzed for the first three aims of the current study.

For the intelligibility experiment of study aim 4, thirty speech samples were selected for stimuli in the experiment. Only 30 (21 children) out of 38 speech samples (25 children) were available because of missing audio or video files or low sound quality. To be consistent in inclusion criteria of the participant for the first three study aims, 28 speech samples were selected for the analysis. One of the excluded speech samples was collected from a child who used bimodal sensory aids and the other speech sample was acquired from a child who did not have a part of parent questionnaires for the background information. Table 1 summarizes background information for the child participants with hearing loss.

Table 1. Characteristics of children with hearing loss.

Cross-sectional data set									
Child	Age[1]	Intelligibility	Sex	Sensory aid	Age of ID	Age at HA/CI	Modality	ST[3]	Hearing loss
OlSa	2:10	Yes	F	HA	At birth	6	Talking Only	1	--
DeBr	2:10	No	F	HA	At birth	9	Talking Only	3	--
RoSt	4:04	Yes	M	CI	After birth	20	Talking Only	2	--
MaRa	4:01	Yes	M	HA	At birth	6	Talking Only	1	Mild
GaSc	5:01	Yes	M	HA	After birth	48	Talking Only	--	--
TaWi	5:00	Yes	M	HA	After birth	33	TC[2]	2	Moderate
SuDi	5:03	Yes	F	CI	At birth	12	Talking Only	1	--
AbSy	5:07	Yes	F	HA	After birth	N/A	Talking Only	3	Moderate-Severe
EdMo	5:07	Yes	F	HA	After birth	30	Talking Only	1	Moderate
SeEr	6:01	No	M	CI	At birth	12	TC	2	Profound
AlBe	6:04	Yes	M	CI	At birth	12	TC	0	--
EvCa	7:04	Yes	F	CI	At birth	N/A	TC	2	--
CaBr	8:01	Yes	F	CI	After birth	24	TC	2	Profound
Longitudinal data set									
Child	Age[1]	Intelligibility	Sex	Sensory aid	Age of ID	Age at HA/CI	Modality	ST[3]	Hearing loss
CaFi1	2:03	Yes	M	CI	At birth	12	TC[2]	0	Profound
CaFi2	3:09	Yes							
CoCe1	2:08	Yes	M	CI	At birth	9	Talking Only	1	Profound
CoCe2	3:07	Yes							
MaCe1	2:08	Yes	M	CI	At birth	9	Talking Only	1	Profound
MaCe2	3:07	Yes							
NiBy1	2:03	No	M	CI	At birth	11	Talking Only	0	Profound
NiBy2	3:03	Yes							
NiBy3	4:02	Yes							
NiBy4	5:02	Yes							
TyMc1	4:01	Yes	M	HA	After birth	36	Talking Only	1	--
TyMc2	5:04	Yes							
SoAm1	4:11	Yes	F	HA	At birth	9	TC	1	Moderate-Severe
SoAm2	5:11	Yes							
MaTa2	5:01	No	F	CI	After birth	50	TC	2	Profound
MaTa3	7:01	Yes							
MaTa4	8:01	Yes							
StOd1	5:02	Yes	M	CI	After birth	36	TC	3	Profound
StOd2	6:01	Yes							

[1]Chronological Age
[2]Total Communication
[3]Speech Therapy: 0 – no, 1 – once a week, 2: twice a week, 3: more than twice a week

3.2.1.1. Data distribution of hearing device and communication modality

A distribution of the children by hearing device and communication modality was descriptively examined via contingency tables. Table 2 shows data distributions for all the 21 children. In the CI group, children who use Talking Only is less then children who use Total Communication. In the HA group, children who use Talking Only is more than children who use Total Communication. In a longitudinal data set (8 of 21 children with hearing loss), only two children wore HA. Table 3 shows the data distribution of the longitudinal data (8 children).

Table 2. Distribution of all participants (21 children) by hearing device and communication modality.

	CI	HA	Total
Talking Only	5 (41.7%)	7 (77.8%)	12 (57%)
Total Communication	7 (58.3%)	2 (22.2%)	9 (43%)
Total	12 (57%)	9 (43%)	21 (100%)

Table 3. Distribution of the longitudinal data for 8 children by hearing device and communication modality.

	CI	HA	Total
Talking Only	3 (50%)	1 (50%)	4 (50%)
Total Communication	3 (50%)	1 (50%)	4 (50%)
Total	6 (75%)	2 (25%)	8 (100%)

The age range of the 21 children in the overall sample (32 speech samples) was from 27 months old (2; 3) to 97 months old (8; 1). The mean of age for the CI group was

58 months and the mean of age for HA group was 57 months. For the total communication group, the mean of the children's age was 49 months. For the Talking Only group, the mean of the age for the children was 68 months. The chronological age of the Talking Only group was older than the total communication group. For the longitudinal data, all age information for multiple data collections was included. For example, TyMc participated in the study twice. His age was 49 months at the first data collection and 64 months at the second data collection. These two ages were included for the data description by hearing device and communication modality.

3.2.2 Adult Participants

Fifty-two young adult participants were recruited from lower division UT-Austin undergraduate classes. Participants completed questionnaires about their language background and basic vision and aural capacities. All listeners completed *The Language Experience and Proficiency Questionnaire (LEAP-Q,* Marian, Blumenfeld. & Kaushanskaya, 2007) prior to testing to identify whether they were monolingual English speakers with no second-language exposure from their main caregivers (i.e., parents) before the age of 12, English and other language bilinguals, or English late learners. Based on their response from the LEAP-Q, 32 participants only used English without any early second language experiences. Twenty participants had second language experiences before the age of 12. In this group, 16 participants were English dominant speakers even though they had early non-English language exposure. Four participants who were fluent non-English speakers were excluded for the data analysis.

The participants also performed tasks to measure their hearing threshold at 500 Hz, 1000 Hz, 2000 Hz, and 4000 via a portable audiometer with headphones. Two participants were tested but excluded from the analysis because their hearing thresholds were over 25dB. Finally, 46 participants were included for the data analysis. There were no difference between English speakers who had no early non-English exposure (47.8%) and English speakers who had early non-English experience before 12 years of the age (46.4%). All participants provided written informed consent.

3.3 DATA COLLECTION

3.3.1 Spontaneous speech and language sample

Data on vocalizations and connected speech output for the children with hearing impairment was collected in 20-30-minute sessions. Each child played with one or two graduate clinician level speech language pathologist(s) (SLP) in a quiet clinic room at the University of Texas at Austin Speech and Hearing Center (UTSHC). The computer for collecting audio samples was available in the room and a video camera was also built in. During the session, the child and clinicians sat on the floor. Developmentally appropriate toys available in the UTSHC were employed for collecting the connected speech and language samples for each child participant. Each speech sample was recorded by an Integrated Circuit (IC) recorder or built-in camera in the clinic room. There was one exceptional case for the data collection. One speech sample of a child was collected in his home while playing with his mother. The parent was instructed to record conversations between the child and herself with familiar toys for 20-30 minutes.

The twenty-one children used word combinations as their primary oral communication unit. All identified word tokens occurring in each child's data collection were selected for analysis. For example, when a child produced a speech sound "/bour/" by indicating or holding a toy "horse" while playing with an SLP in a session, the speech output "/bour/" was considered as a word indicating "horse". A primary transcriber trained in transcription of child speech transcribed all word tokens using broad phonetic transcription conventions. A second trained transcriber listened to recording files through audio or video and also used broad transcription. All transcribed data were entered for computer analysis using the *Phon* database program (Rose et al., 2006) with broad phonetic transcription. Suprasegmental features were not included for analyses of speech pattern outcomes.

Reliability of transcription was calculated by using a point-to-point method based on Davis, MacNeilage, and Matyear (2002). Reliability of transcription for the children was calculated by two transcribers who were undergraduate majors in communication sciences and disorders who had been trained in transcription of young children's speech. Using the audio or video file, the second transcriber randomly selected 20% of the sample from each child and coded sounds independently. Three speech samples of the thirty-two speech samples did not have audio or video files, thus these three samples were excluded for reliability. Reliability for consonants and vowels was computed separately. Reliability for individual consonants was 90.6% and for vowels was 90.8%.

3.3.2 Intelligibility experiment with adult participants

3.3.2.1 Materials

Thirty speech samples from 21 children were selected for the intelligibility experiment due to the lack of audio or video files and low quality of sound files. Fifteen utterances from each speech sample were selected as much as background noise of the sound files can be reduced. In addition, Mean Length of Utterance (MLU) of the selected utterances per child was matched to the original value of MLU for total utterances per child. The speech sample recordings were segmented into utterance-length files and equated for average root-mean-square amplitude. Six-hundred utterances were selected in total as the stimuli for the experiment. The utterances were randomly grouped into three sets (200 utterances per set). Each set included three to six utterances from each child's spontaneous speech sample. A set was divided into two sessions. One session consisted of one to two-word utterances and the other session consisted of three or more-word utterances. The following example sentences were included for the materials of the intelligibility tasks. For instance, "green", "a slinky", "my own bear", "it's a train", "the train is coming by" etc. were included. Monosyllable function words and be-verbs such as "a" and "is" were excluded for the data analysis. Two speech samples from 2 children with hearing loss were excluded because a child wore bimodal hearing devices and another child had no full version of the parent questionnaires.

3.3.2.2 Procedures

All listening tasks were conducted in a sound proof booth located in the UT-Austin Speech Production Lab CMA 2.228. Prior to the listening task, all adult participants were asked to sign a consent form and fill out a detailed language background questionnaire (*The Language Experience and Proficiency Questionnaire*, LEAP-Q; Marian et al., 2007). Hearing screening was also performed using a portable audiometer via headphones in a quiet room. Hearing sensitivity of <25 dBHL at 1000 Hz, 2000 Hz, and 4000 Hz in both ears was required to take part in the experiment. The participants then listened to words and sentences in a quiet condition in the sound booth. The tasks were conducted through a PC via stimulus presentation software (e.g. 'PsychoPy', Peirce, 2007). All participants wore headphones and sound was presented at comfortable listening loudness levels equivalent to a naturally occurring conversation (i.e., approximately 60 dB). Each participant only listened to the 200 utterances from one of the three sets.

A set was divided into two sessions. One session only included one or two-word utterances and the other session included three or more-word utterances. The order of sessions was counterbalanced. The stimuli were randomly presented within a session. Each stimulus was presented only once. The participant was required to listen to each sentence and type as much as they could onto a computer. Each trial was presented only once, but participants could take as much time as they wished to write down each sentence.

3.4 DATA ANALYSIS

3.4.1 Speech characteristics of children with hearing loss (Aim 1)

Speech outcomes based on inventories and phonological accuracy of consonants and vowels as well as whole word phonological measures were analyzed. In order to obtain a more comprehensive picture of each child's phonological system, relevant measurements of speech outcomes to describe particular speech patterns for young children were analyzed for segmental accuracy and word-level variables.

In a speech sample, all unintelligible and unidentified words were excluded for this data analysis. Children's speech outcomes varied from the word-level to the sentence level. At a sentence level, monosyllable function words were pronounced various ways such as /ænd/, /ən/, or /n̩/ for "and". Thus, monosyllable function words were not included in the data analysis. Appendix 1 shows a list of excluded mono-syllable function words. Be-verbs (e.g. be, been, being, is, am, are) were also excluded. All speech measures were calculated per speech sample in order to examine speech characteristics between CI and HA group as well as the Talking Only and Total Communication groups.

To describe speech characteristics of children with hearing loss based on PCC-R, PVC-R, PWP, and WCM-ratio, children's type of sensory aids and communication modality, a linear mixed model was applied by using the lme4 package in R for each speech measure. In each model, a numerical value for a speech measure for the 32 speech samples was the dependent variable. Fixed effects included chronological age, hearing device and communication modality, as defined earlier. Due to repeated data collections for the 8 children in the longitudinal sample, participant (child) was a random intercept. In addition, the model with a random effect (child) improved the fit of the model for each speech measure. Multiple observations within a participant is a violation of the assumption of sample independence (Bewick, Cheek, & Ball, 2003). In this model, the intercept represents reference condition indicating a value of the speech measure for a child at the mean of chronological age in months, with cochlear implants, and in the Talking Only group.

3.4.1.1 Percentage of Consonant Correct-Revised (PCC-R)

The most widely used segment-oriented measure for phonological patterns is the Proportion of Consonants Correct (PCC) from conversational speech developed by Shriberg and Kwiatkowski (1982). PCC has been frequently cited in previous studies to compare the accuracy of the child's consonant production to the consonants in the target word (Moeller et al., 2010, Tobey et al, 2007). PCC-R (Shriberg et al., 1997) is a measure of consonant phoneme accuracy. Omissions and substitutions arescored as errors and consonant distortions are scored as correct. The values for consonant correct of these children with hearing loss were compared to the normative values by chronological age for hearing children developing typically (Campbell, Janosky, & Adelson, 2007). According to the Texas Speech-Language-Hearing Association (TSHA) Implementation Guidelines (2010), PCC-R includes the first two repetitions of a word unless the word is pronounced with different consonants. For example, when a child produces /baba/, /baba/, /baba/, /bai/, & /ba/ for "bye", the first two /baba/ and /bai/ are included for computing PCC-R. The current study followed the procedure from TSHA Implementation Guidelines (2010).

3.4.1.2 Percentage of Vowel Correct-Revised (PVC-R)

PVC-R (Shriberg et al., 1997) is a measure of vowel phoneme accuracy. Like PCC-R, omissions and substitutions are scored as errors and vowel distortions are scored as correct. PVC-R also does not include target vowels in the third repetition of a word unless articulation changes. When a child pronounces a vowel as his or her dialect (e.g. /bɪldiŋz/ for /bɪldɪŋz/ "buildings" using a Texas dialect), the child's vowel was scored as a correct vowel.

3.4.1.3 The Proportion of Whole-Word Proximity (PWP)

Ingram (2002) introduced the Phonological Mean Length of Utterance (PMLU) as a whole-word measure to account for phonological complexity. The PMLU provides information about the number of speech sounds in the words children produce, along with the number of correctly produced consonants. To measure degree of accuracy, Ingram (2002) also developed the Proportion of Whole-word Proximity (PWP). PWP is a measure that divides the child's PMLU score by the PMLU score of the words attempted. PWP measures how well the child matches the target words attempted. Each word receives points for the number of segments and the number of consonants correct for the PMLU score. The current study focuses on the PWP to assess how successful a child is at achieving whole-word production accuracy. The PWP is a measure for the child's target words calculated by dividing by the PMLU of the target words into the child's PMLU. For example, if a child produces /gin/ for /grin/ "green", the actual PMLU is 5, and the target PMLU is 7. PWP for this example is 0.7 (5/7). Ingram (2002) defined the procedures for the calculation of the PMLU (see Table 4), and the current study followed those procedures.

Table 4. Procedures for calculation of the Phonological Mean Length of Utterance (PMLU; Ingram, 2002).

1. *Sample-Size Rule*: Select at least 25 words, and preferably 50 words for analysis, depending on sample size. If the sample is larger than 50 words, choose a selection of words that cover the entire sample, e.g., every other word in a sample of 100 words. 2. *Lexical-Class Rule*: Count words (e.g., common nouns, verbs, adjectives, prepositions and adverbs) that are used in normal conversation between adults. This excludes child words, e.g., mommy, daddy, tata etc. Counting child words can inflate the PMLU if a child is a reduplicator. 3. *Compound Rule*: Do not count compound words as a single word unless they are spelled as a single word, e.g., 'cowboy' but not 'teddy bear', i.e. 'teddy bear' would be excluded from the count. This rule simplifies the decisions about what constitutes a word in the child's sample. 4. *Variability Rule*: Only count a single production of each word. If more than one occurs, then count the most frequent one. If there is none, then count the last one produced. Counting variable productions may distort the count if there is a highly variable single word. 5. *Production Rule*: Count 1 point for each consonant and vowel that occurs in the child's production. Syllabic consonants receive one point, e.g., syllabic 'l', 'r' and 'n'. (Some transcriptions may show these as two segments, i.e., a schwa plus consonant, e.g., 'bottle' [bɑdəl], but it should be counted as one consonantal segment.) Do not count more segments than are in the adult word. For example, a child who says 'foot' as *hwut+ has two consonants counted, not three. Otherwise, children who add segments will get higher scores despite making errors. 6. *Consonants Correct Rule*: Assign 1 additional point for each correct consonant. Correctness in vowels is not counted since vowel transcriptions are typically of low reliability. Syllabic consonants receive an additional point in the same way as nonsyllabic consonants. A child who applies liquid simplification, for example, will get 1 point for producing a vowel, e.g., 'bottle' *bɑdo], but 2 points if the syllabic consonant is correct. PWP: The following formula is used to determine the PWP for each word. (number of segments in child production + number of correct consonants) / (number of segments in target word + number of consonants in target word).

Ingram (2015) introduced the PCC-R - PWP intersect analysis to discriminate speech disorder from speech delay for young children. The procedure of the PCC-R – PWP intersect analysis is that children's words were divided into categories along two dimensions: 1) monosyllables versus multi-syllables, and 2) words with clusters versus words without clusters. PCC-R and PWP scores were calculated for these four categories. These two dimensions for word complexity generate a linear relationship between the two measures. In other words, this linear relationship between PCC-R and PWP is influenced by word complexity. Children with hearing loss were assigned to Linear or Nonlinear groups based on the linearity of relationship between PCC-R and PWP. The Linear group included children who showed monosyllable words with higher PCC-R and PWP scores than multi-syllable words, and higher scores of words without clusters than words with clusters. That is, consonant correctness correlated with word complexity. The simpler the phonotactics of the word, the higher the rate of consonant correctness. The Nonlinear group in this study included children who did not show any linear relation between consonant correctness and word complexity.

3.4.1.4 The Word Complexity Measure (WCM)

Stoel-Gammon (2010) developed the Word Complexity Measure (WCM) to evaluate phonetic complexity based on analysis in a single-word sample regarding three domains: word patterns, syllable structures and sound classes. WCM can be measured as a single number which can be used to quantitatively compare the child's forms with his/her target words that shows how much the child's productions are simplified by reporting a single number. However, this measure is not relational measures of accuracy or errors but independent analyses of two sets of data. So far, the WCM has been implemented in few studies including persons who stutter (Coalson et al., 2012) and aphasics (Haley & Jacks, 2013).

The WCM awards points for sounds and structures in the following areas: (1) more than two syllables (2) non-initial stress (3) word-final consonants (4) consonant clusters (5) a velar consonant (6) a liquid, a syllabic liquid, or a rhotic vowel (7) fricative or affricate (8) voiced fricative or affricates (in addition to the point received for #7). The WCM for all words each child produced were compared with the WCM for their word targets. Ratios of WCM for child productions to target WCM scores were calculated. All eight parameters of phonological complexity were categorized by word patterns, syllable patterns, and sound classes based on Stoel-Gammon (2010). For example, /grin/ for "green" gets rewarded 4 points for a velar consonant, a liquid, a consonant cluster, and a final consonant. If a child produces /gin/ for "green", his or her production for the word gets 2 points for a velar consonant and a final consonant. In this case, the ratio of WCM actual and target is 0.5 (2/4). The average of the WCM ratios for each child was used for analysis. The repetition rule was applied in the same manner as for PCC-R. Thus, the first two repetitions for a word were included when a child repeated the word more than twice in a speech sample.

To examine relationships between the 8 parameters of the WCM-Ratio and PCC-R, linear mixed models were applied for each parameter. Correlation analysis was not available due to multiple observations within a child for the children. A model included

PCC-R as a dependent variable, a parameter of the WCM-Ratio as a fixed effect, and child for a random effect. Thus, each model had its R-squared for a fixed effect and the R-squared was compared across the 8 parameters.

3.4.2 Language characteristics in children with hearing loss (Aim 2).

To characterize language capacity of children with hearing loss Mean Length of Utterance (MLU), Type Token Ratio (TTR), and the number of different word types were computed for the analysis. These three language scores were also differentiated by hearing device and communication modality. Three linear mixed models were applied by using the lme4 package in R for the language measures. In each model, a numerical value for the language measure (e.g. values of MLU across the speech samples) was the dependent variable. Fixed effects included chronological age, hearing device and communication modality and a random intercept of the participant (child). In this model, the intercept represents reference condition indicating a value of the language measure for a child at the mean of chronological age in months, with cochlear implants, and in a speaking only group.

3.4.2.1 Mean Length of Utterance (MLU)

Mean Length of Utterances (MLU, Brown, 1973) is one of the general tools used to assess grammatical abilities, including morphological and syntactical abilities in language development. MLU played a noticeable role in Brown's stage model of grammatical development and is a reference for developmental language status in hearing children. Rice, Redmond, amd Hoffman (2006) revealed robust reliability and validity for MLU as an index of clausal development by using 124 conversational samples from 39 children (5 years old) with SLI, 40 MLU-equivalent typically developing children (3 years old), and 45 age-equivalent controls. They concluded that MLU can be a reliable and valid index of general language development.

The MLU in morphemes (Brown, 1973) is a measure of morphological and syntactical skills. MLU score is calculated as the average length of utterance based on a count of the number of individual morphemes in each utterance. For example, if a child produces “riding horse”, the MLU of the utterance is 3. MLU played a seminal role in Brown’s stage model of grammatical development and has been used as a reference for developmental language status in children developing typically.

To examine relationships between speech measures and MLU, linear mixed models were applied for each speech measure because correlation analysis was not available due to repeated measures for the 8 children in the longitudinal data set. Speech measures were not included in a linear mixed model with MLU due to multicollinearity across the speech measures. A model included MLU as a dependent variable, a speech measure as a fixed effect, and child for a random effect. Thus, each model had its R-squared for a fixed effect and the R-squared was compared across the speech measures.

3.4.2.2 Type/Token Ratio (TTR)

Type/token Ratio (TTR, Templin, 1957) is a measure of quality of vocabulary at the lexical level. TTR was calculated by determining the number of different words produced and dividing that number by the total number of words produced in at least 50 utterances (see following formula). TTR = number of types / number of tokens. TTR provides evidence for the lexical richness of one's lexicon. When a child produces 50 tokens that include 30 different word types, his or her TTR will be 0.6 (30/50).

3.4.2.3 Number of different word types

The number of different word types in a spontaneous speech sample is another measure of quality of vocabulary use. Davis, van der Feest, & Yi (2017) used the cumulative number of different word types across longitudinal data sessions to indicate vocabulary growth in young typically developing children. The number of different words a child produced from their spontaneous speech sample was calculated to characterize the child's vocabulary size based on conversation with adults in a spontaneous play context.

3.4.3 Vocabulary input factors on speech outcomes in children with hearing loss (Aim 3).

The current study only included CVC words to examine the goal, of considering input related factors on speech outcomes, due to a temporally unavailable online calculator (Storkel, personal communication, July 2017). Instead of using the online calculator, the available CVC database in Storkel (2013) was used analysis in the current study. All the variables of interest for the CVC words were derived from Storkel (2013). The corpus wasobtained either from the Hoosier Mental Lexicon (HML; Nusbaum, Pisoni, & Davis, 1984) or from a dictionary that provides phonemic transcription of English words (Longman, 1993). The online calculator includes 4,832 words from child corpus of American English (Kolson, 1960; Moe, Hopkins, & Rush, 1982).

For this goal, all speech outcomes were calculated based on word-by-word comparisons, that are not based on original rules for the speech measures (e.g. PCC-R). In this analysis, instead of using PCC-R and PVC-R, consonant correctness and vowel correctness were labeled for speech outcomes respectively. The values of input characteristics for the PWP and WCM-Ratio are limited because various word lengths and a part of syllable patterns (e.g. consonant cluster) cannot be considered in CVC word forms. Thus, the PWP and WCM-Ratio were excluded for this input analysis.

PCC-R and PVC-R were calculated based on word-by-word comparisons. This procedure represents an adaptation of the original rules for calculating PCC-R/PVC-R. Instead of using PCC-R/PVC-R, consonant correctness and vowel correctness were labeled for the speech measures. Consonant correctness includes three types of scores (0, 0.5 &1). Zero, "0" indicates that both of two consonants do not match adult forms in a CVC word. "0.5" indicates that a CVC word includes one correct consonant of the two consonants. In this case, word positions (i.e. word initial and final position) were not considered to compute consonant correctness, consistent with original PCC-R calculation instructions. "1" means both of two consonants are exactly same as the adult forms in a CVC word. For calculating vowel correctness, a CVC word form includes two types of scores (i.e. 0, 1). Zero, "0" for vowel correctness indicates incorrect actual production in a vowel in a CVC word. "1" means correct actual child production for a vowel.

3.4.3.1 Phonotactic probability

Phonotactic probability includes two measures: a) positional segment sum and b) biphone sum. Positional segment sum refers to how often a particular segment occurs in a certain position in a word (Storkel, 2004). Storkel (2004) reported that positional segment sum is computed by summing individual positional segment frequency for each segment in the CVC. Biphone frequency of a sound was defined as the co-occurrence probability of sounds within a word. The biphone sum is also computed by summing the frequency of twophoneme sequences (i.e., CV or VC) for the words instead of individual segments. Storkel provides the z scores of the positional segment sum and biphone sum in her supplemental materials (Storkel, 2013).

3.4.3.2 Neighborhood density

Neighborhood density is defined by the number of words that differ from a given word by one phoneme (Storkel, 2004; Storkel & Hoover, 2010; & Storkel, 2013). Storkel (2013) reported that neighborhood density is calculated by counting the number of words appearing in corpus that differ from the given CVC word by a one sound substitution, addition, or deletion in any word position. Z-values of neighborhood density were also computed and reported in her supplemental materials. Again, the current study used z-values of positional segment sum, biphone sum, and neighborhood density from Storkel (2013)'s materials.

3.4.4 Intelligibility tests for spontaneous speech outcomes for children with hearing loss perceived by adult participants (Aim 4).

Aim 4 of the study was to determine the best speech measures to predict intelligibility as judged by adult listeners. Twenty-eight among the 30 available speech samples were selected to analyze relationships between speech measures and intelligibility to be consistent with the inclusion criteria of the previous three aims for the current study. One of the excluded speech samples was collected from a child who used bimodal sensory aids and the other speech sample was acquired from a child whose family lost a part of their parent questionnaires.

Item identification measures for speech intelligibility were applied. To assess connected speech, listeners were presented with audio recordings of unfamiliar sentences and asked to write down the words they understood in each utterance. This identification measure is less subjective than the scaling procedures and resulted in increased reliability between listeners. It provides a functional intelligibility metric to consider correlations across types of speech measures that can support understanding of language acquisition in young children with hearing impairment (Ertmer, 2011). Because the intelligibility score was collected by word, all speech outcomes were calculated based on word-by-word comparisons, that are not based on original rules for the speech measures (e.g. PCC-R). If a child produced /ta/ for “truck”, the value for consonant correctness is 1/3 (33%). To be consistent with the Aim 3, instead of using PCC-R, PVC-R, PWP, and WCM consonant correctness, vowel correctness, word proximity, and word complexity were labeled for the speech intelligibility outcome respectively.

Prior to conducting inferential statistics, the overall findings of the intelligibility tests were described by a group-mean intelligibility score across children. The scores for the four speech measures per child (i.e., the original scores calculated for the speech measures) were used to describe relationships between speech measures and intelligibility scores. The goal was to determine the best speech measure to predict intelligibility, Mixed effects logistic regression models were applied with lme4 package (v1.1-14) in R (v3.4.2). More details of the regression models will be reported in the results section below.

4. Results

4.1 SPEECH CHARACTERISTICS OF CHILDREN WITH HEARING LOSS (AIM 1).

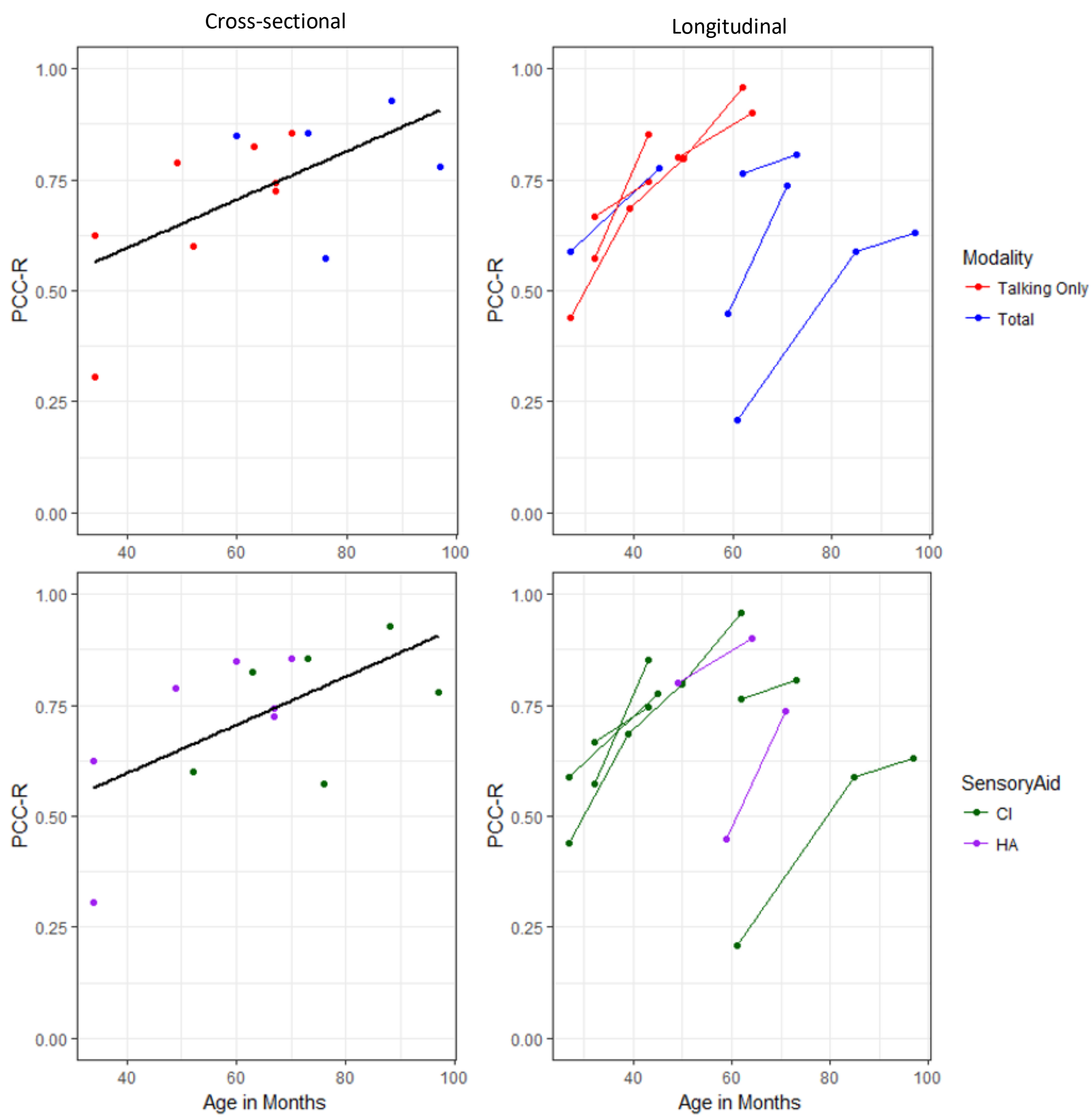
4.1.1 Percentage Consonant Correct-Revised (PCC-R)

For the thirty-two speech samples for 21 children with hearing loss, PCC-R was analyzed using a linear mixed regression with (v1.1-14) in R (v3.4.2). Fixed effects included chronological age, hearing device, and communication modality. Participant (child) was a random intercept. Model comparisons was performed using Akaike Information Criterion (AIC: Bozdogan, 2000). The fit of a model with the three main effects was improved by adding the random intercept of child ($p = 0.003$). The linear mixed model revealed that age is a significant predictor of PCC-R for these children with hearing loss. Controlling for both hearing device and communication modality, the older a child's chronological age, the higher score on PCC-R they achieve. Specifically, for two children who have the same types of hearing device and communication modality, a child who is 1 month older than the other child is predicted to have 0.01 point (1%) of PCC-R higher than the other child ($B = 0.011$, $t(10) = 7.489$, $p < .001$). In the longitudinal data, PCC-R improved over time for all the eight children. In the cross-sectional data, older children tended to show higher PCC-R values than younger children (See Figure 1).

Communication modality relates to the Talking Only versus Total Communication designation. Controlling for chronological age and hearing device, communication modality is a significant predictor of PCC-R ($B = -0.206$, $t(18) = -2.259$ $p = .037$). This result indicates that the Total Communication group showed a mean PCC-R score of 0.21

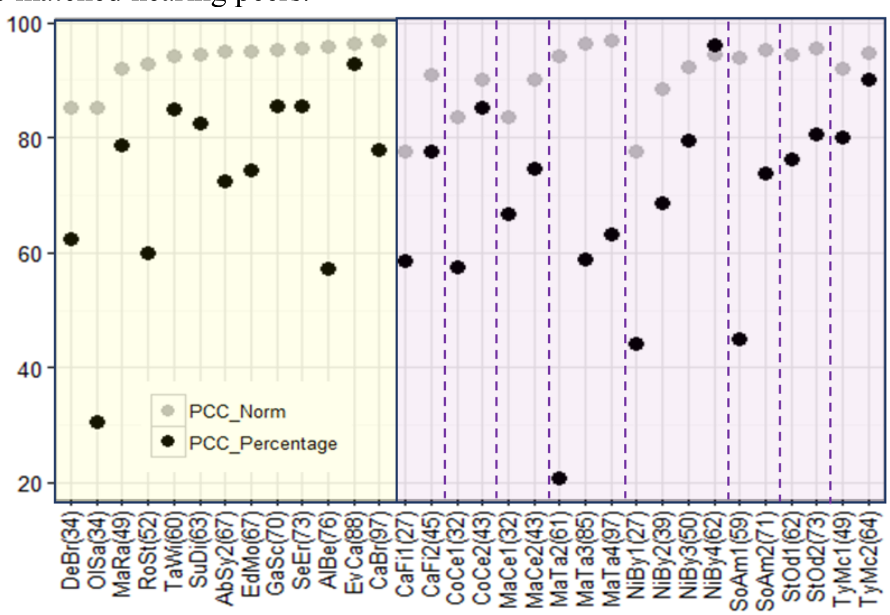
(21%) lower than the Talking Only group when the children were the same chronological age and used the same hearing device. The PCC-R for the two modality groups were approximately the same as each other without considering their chronological age. Figure 1 shows relationships between chronological age and PCC-R for the children in the cross-sectional (left) and longitudinal (right) data sets. Red dots indicate the Talking Only group. Blue dots denote children who use Total Communication. Within the current data sets, the Talking Only group was chronologically younger (mean: 49 months old) than the total communication group (mean: 70 months old). Even though children who use oral communication (i.e., talking only) are younger, they show higher scores on PCC-R especially in the longitudinal data set. When the chronological age effect is controlled, communication modality showed significant effects on PCC-R scores. When chronological age and communication modality are controlled, hearing device is not a significant predictor of PCC-R. B = - 0.020, t (18) = - 0.233, p = 0.818. The median PCC-R value of the CI group was 0.75 (75%), and of the HA group was 0.74 (74%).

Figure 1: PCC-R across chronological age in months. The left graphs include cross-sectional child group data and the right graphs include longitudinal child group data. Red colored dots are children who use Talking Only. Blue dots denote children who use total communication. Green dots show CI group and purple dots demonstrate HA group. The linear black line in the left graph is a "smoothed conditional mean" to show trends in the raw data for PCC-R across chronological ages. The eight lines in the right graph connect data within a child in the longitudinal data sets for the 8 children. PCC-R for the two groups of sensory aid users (CI vs. HA) and communication modalities employed.



PCC-R scores for the Children with hearing loss were also compared to normative data (Campbell et al., 2007). Figure 2 shows PCC-R scores for the cross-sectional (yellow) and the longitudinal (purple) data groups arranged by their chronological age within a data set. Overall, these children with hearing loss demonstrated lower PCC-R scores compared to age matched hearing children except on one data sample. Within the longitudinal data set, children demonstrated better performance on PCC-R when they were chronologically older, but none of the children achieved the normative score for hearing children except one child, NiBy.

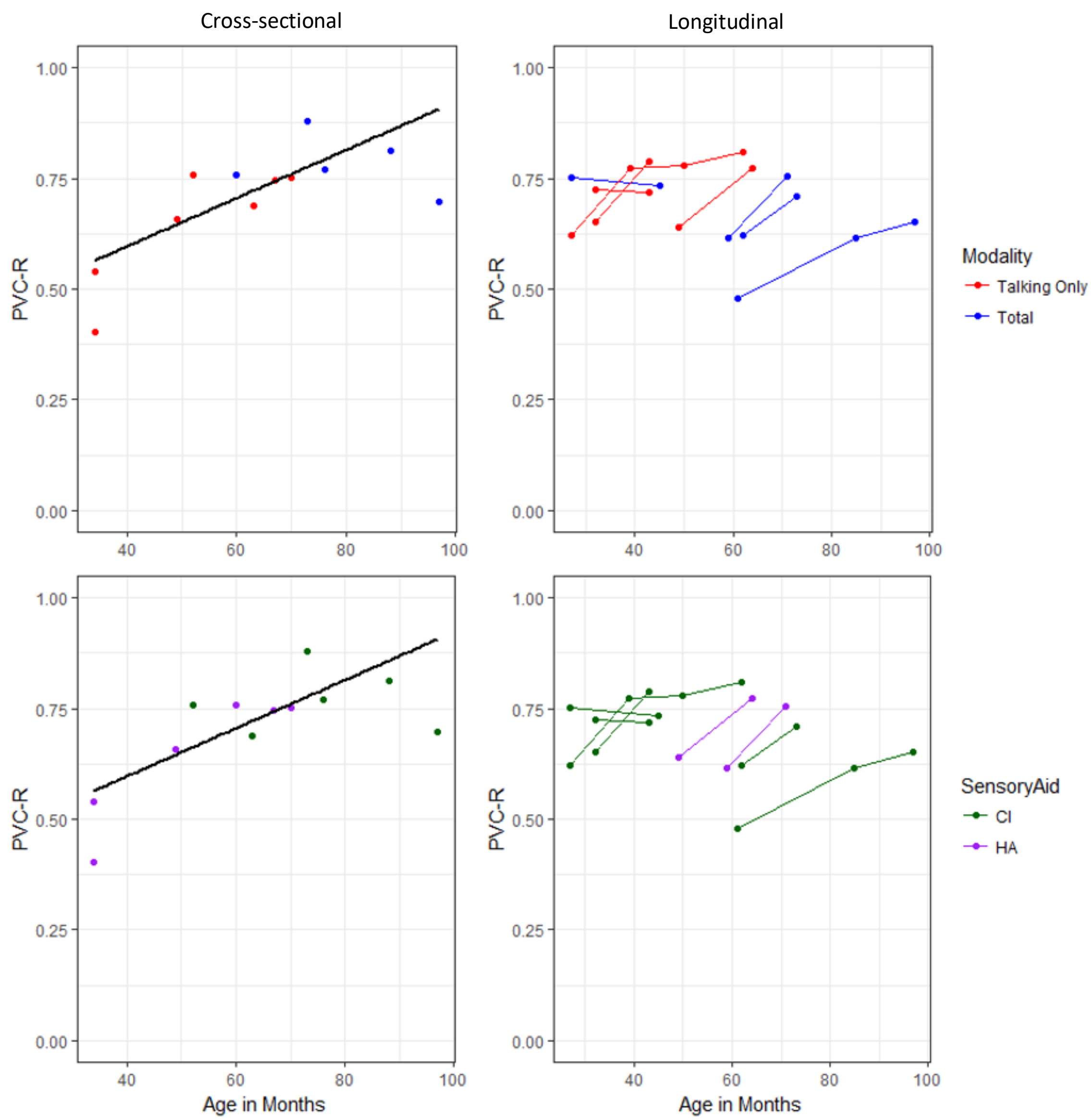
Figure 2. PCC-R ordered by chronological age in months. The left box (yellow) indicates children who had only one data sample (13 children, 13 samples). The right box (purple) shows the children who had more than 1 data sample (8 children, 19 samples). Dashed lines distinguish individuals within the longitudinal data group shown on the right. Black dots are scores for children with hearing loss PCC-R scores. Gray dots are PCC-R normative scores by age in months of chronological age matched hearing peers.



4.1.2 Percentage Vowel Correct-Revised (PVC-R)

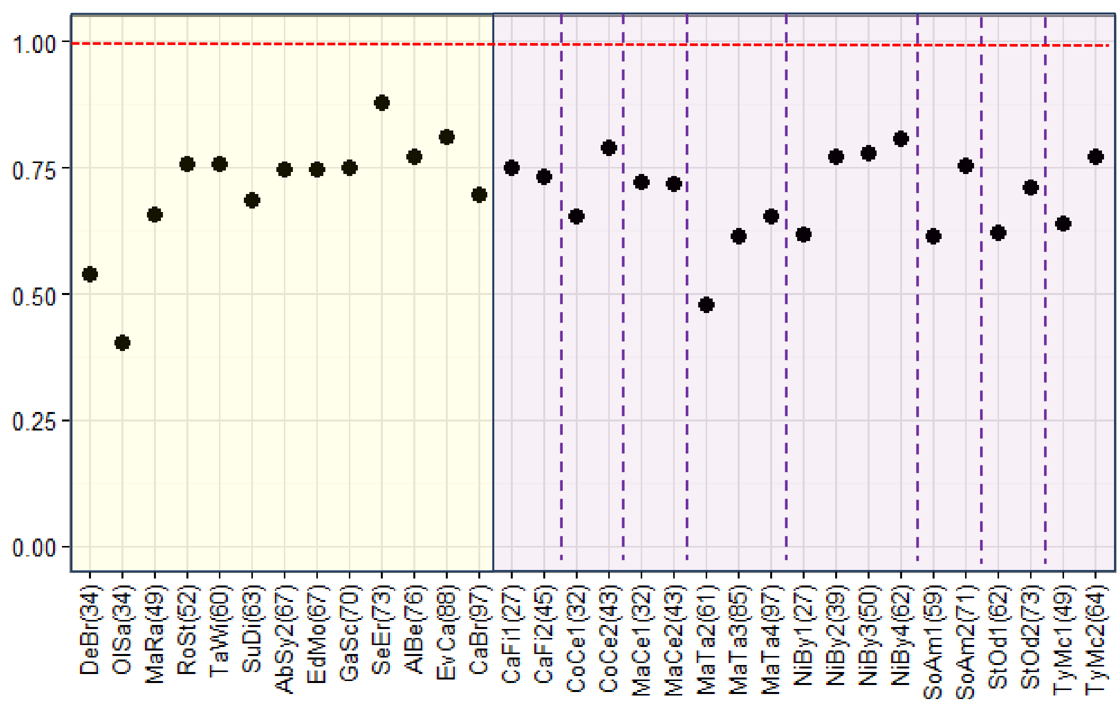
PVC-R was analyzed using a linear mixed regression including the three fixed effects of chronological age, hearing device, and communication modality and a random intercept of child. The fit of a model with the three main effects was improved by adding the random intercept of child (p = 0.003). The linear mixed model for the PVC-R revealed that age is a significant predictor of PVC-R for these children with hearing loss. Controlling for hearing device and communication modality, the older a child's chronological age, the higher score on PVC-R they achieve. Specifically, for two children who have the same types of hearing device and communication modality, a child who is 1 month older than the other child is predicted to have 0.004 point (0.4%) of PVC-R higher than the second child (B = 0.004, t (10) = 4.415, p = .001). Figure 3 shows relationships between chronological age and PVC-R for the children in the cross-sectional data set (left) and the longitudinal data set (right). Controlling for other predictors, neither fixed effects of hearing device or communication modality are significant (hearing device: B = - 0.058, t (18) = -1.233, p = 0.233, communication modality: B = -0.056, t (18) = - 1.103, p = 0.285).

Figure 3. PVC-R across chronological age in months. The left graphs include cross-sectional data and the right graphs include longitudinal data. Red colored dots are children who use talking only. Blue dots denote children who use total communication. Green colored dots are children who use CI. Purple dots denote children who use HA. The linear black line in the left graph is a “smoothed conditional mean” to show trends in the raw data, PVC-R across chronological age. The eight lines in the right graph connect data within a child in the longitudinal data set for those 8 children.



Accurate vowel production is achieved by 3 years of age (Pollock, 2002; Stoel-Gammon & Pollock, 2008, Templin, 1957). PVC-R scores for the children with hearing loss are displayed on Figure 4. None of these children with hearing loss exhibit fully developed vowel production accuracy. Within the longitudinal data set, the eight children demonstrated better performance on PVC-R as they got older, but none of the children achieved 100%.

Figure 4. PVC-R ordered by chronological age in months. The left box (yellow) indicates children who had only one data sample (13 children, 13 samples). The right box (purple) shows the children who had more than 1 data sample (8 children, 19 samples). Dashed vertical lines distinguish individuals within the longitudinal data group. Red horizontal line denotes 100% score of PVC-R.

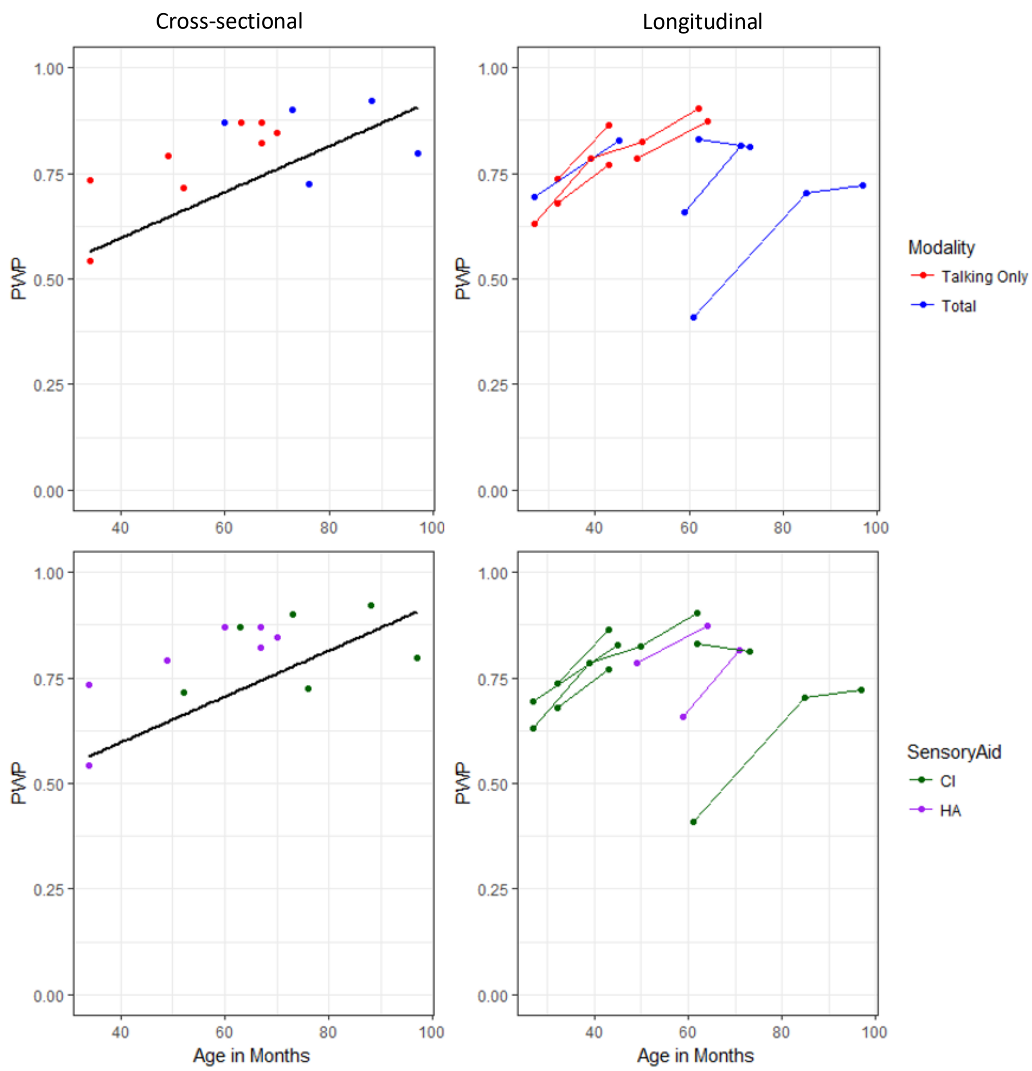


4.1.3 Proportion of Whole-word Proximity (PWP)

PWP was analyzed using a linear mixed regression including the three fixed effects of chronological age, hearing device, and communication modality and a random intercept of child. The fit of a model with the three main effects was improved by adding the random intercept of child (p = 0.002). The result of the linear mixed model showed that chronological age is a significant predictor of PWP for the children with hearing loss. When hearing device and communication modality are controlled, the older a child's chronological age, the higher score on PWP he/she achieves. Specifically, for two children who have the same types of hearing device and communication modality, a child who is 1 month older than the other child is predicted to have 0.006 point (0.6%) of PWP higher than the other child (B = 0.006, t (10) = 7.137, $p < .001$).

Figure 5 shows relationships between chronological age and PWP for the children in the cross-sectional (left) and longitudinal (right) data sets. Red dots indicate the Talking Only Group and blue dots denote children who use total communication. Green dots are children who use CI and purples dots are children who use HA. Controlling for chronological age and hearing device, communication modality is a significant predictor of PWP (B = - 0.118, t (18) = - 2.105, p = 0.050). This result indicates that a child who uses Total Communication showed a PWP score of 0.12 (12%) lower than the other child who used Talking Only when the children are the same age and wear the same hearing devices. Like PCC-R, the Talking Only group was younger (mean: 49 months old) than the total communication group (mean: 70 months old), even though the PWP scores for the two modality groups were approximately the same. This result indicates that when the age effect is controlled, communication modality showed a significant effect on PWP. When chronological age and communication modality are controlled, hearing device is not a significant predictor of PWP (B = - 0.003, t (18) = - 0.054, p = 0.958).

Figure 5. PWP across chronological age in months. The left graphs include cross-sectional child group data and the right graphs include longitudinal child group data. Red colored dots are children who use Talking Only. Blue colored dots denote children who use Total Communication. Green dots are children in the CI group and purple dots are children in the HA group. The linear black line in the left graph is a "smoothed conditional mean" to show trends in the raw data for PWP across chronological ages. The eight lines in the right graph connect data within a child in the longitudinal data sets for the 8 children.

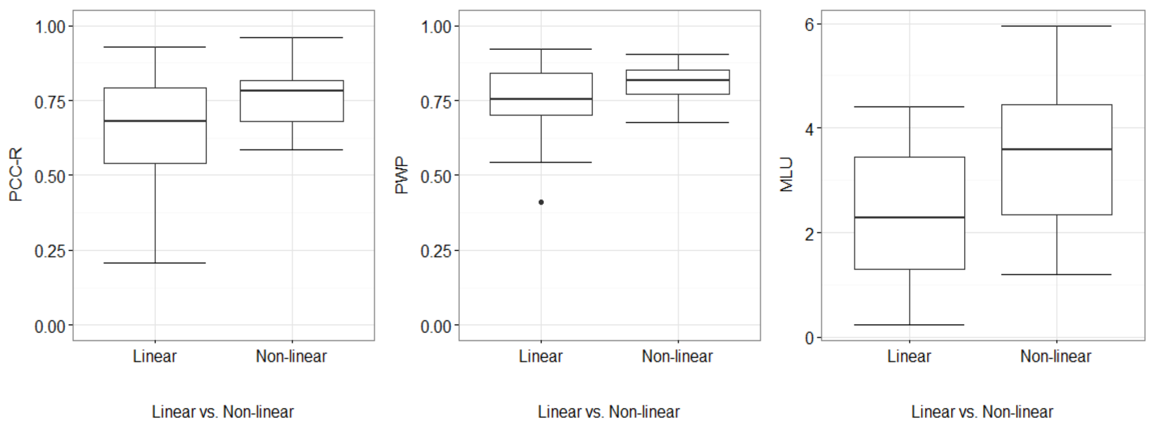


In terms of word complexity related to word length, these children did not show a trend toward better outcomes on PCC-R and PWP for monosyllable than multi-syllable words. Thus, only the dimension ‘consonant clusters’ (i.e., words with versus without clusters) was considered to examine the linearity of the PCC-R – PWP intersect. The 32 speech samples were divided into Linear and Non-Linear groups for the PCC-R - PWP intersect. The linear group included children who showed differences between PCC-R and PWP on simpler words (i.e. words without consonant clusters) smaller than on complex words (i.e. words with consonant clusters). Children in the Nonlinear group did not show any relationship between the PCC-R - PWP intersect and word complexity. The Linear group showed a linear increase in word production accuracy, in that less complex words (singletons) were produced with greater accuracy than more complex words (clusters). The Non-Linear group was unable to produce some consonants correctly regardless of word complexity. They produced more complex words with greater accuracy than some of the less complex words.

Figure 6 shows PCC-R, PWP and MLU scores between the Linear and Nonlinear groups. The Nonlinear group achieved PCC-R, PWP, and MLU scores greater than the linear group. To examine group difference in PCC-R, PWP, & MLU, three linear mixed models were applied by using the lme4 (v1.1-14) in R (v3.4.2). PCC-R, PWP or MLU was the dependent variable for each model. Fixed effects included the linear versus non-linear group. The participant (child) was a random intercept. The linear mixed models revealed that there were significant differences between the linear and non-linear groups on PCC-R

[B = 0.135, t (10) = 2.364, p = .040], and MLU [B = 1.174, t (10) = 2.664, p = 0.024]. However, the two groups did not show any significant difference for PWP [B = 0.065, t (10) = 1.769, p = 0.107].

Figure 6. PCC-R, PWP, and MLU for the linear and non-linear groups. The center line on each box plot indicates the median score, the edges of the box denote the 25th and 75th percentiles, and the whiskers denote maximum and minimum within 1.5 times the interquartile range.



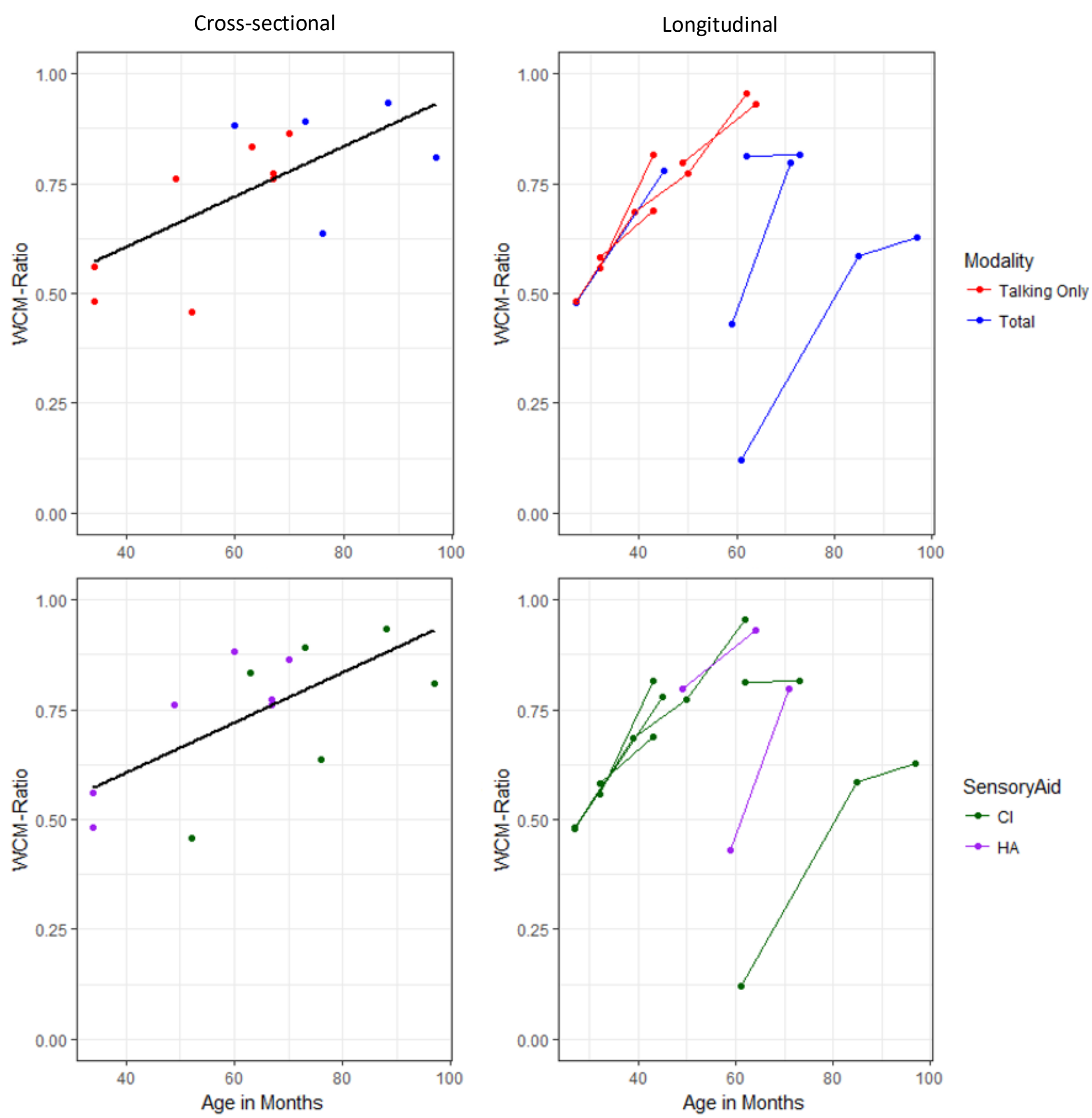
Within the longitudinal data set, 6 out of 8 children demonstrated PCC-R – PWP nonlinearity in their later speech samples. As those 6 children got older, they did not show any relation between word complexity (i.e. singleton versus clusters) and consonant correctness. However, 2 out of 8 children did demonstrate linearity. One child showed linearity in both speech samples. The other child first showed non-linearity when he was younger, but his later speech sample showed linearity. However, neither child produced any multi-syllable words with clusters in their later speech sample.

4.1.4 Word Complexity Measure (WCM)

Results of the linear mixed model revealed that age is a significant predictor of WCM-Ratio for these children with hearing loss. Controlling for hearing device and communication modality, the older a child's chronological age, the higher score on WCM-Ratio they achieve. Specifically, for two children who have the same types of hearing device and communication modality, a child who is 1 month older than the other child is predicted to have 0.011 point (1%) higher WCM-Ratio than the other child, B = 0.011, t (10) = 6.751, $p < .001$.

Figure 7 shows relationships between chronological age and the WCM-Ratio for these children in the cross-sectional data set (left) and the longitudinal data set (right). Red dots indicate children who use Talking Only and blue dots denote children who use Total Communication. Green dots are children with CI and purple dots denotes children with HA. Controlling for other predictors, both fixed effects of hearing device and communication modality are not significant factors in the regression for the WCM-Ratio (hearing device: B = 0.028, t (18) = 6., $p = 0.744$, communication modality: B = -0.176, t (18) = - 1.950, $p = 0.067$).

Figure 7.WCM-Ratio across chronological age in months. The left graph includes cross-sectional child group data and the right graph includes longitudinal child group data. Red colored dots are children who use Talking Only. Blue dots denote children who use Total Communication. Green dots denote children with CI and purple colored dots are children with HA. The linear black line in the left graph is a “smoothed conditional means” to show trends in the raw data for WCM-Ratio across chronological ages. The eight lines in the right graph connect data within a child in the longitudinal data sets for the 8 children.



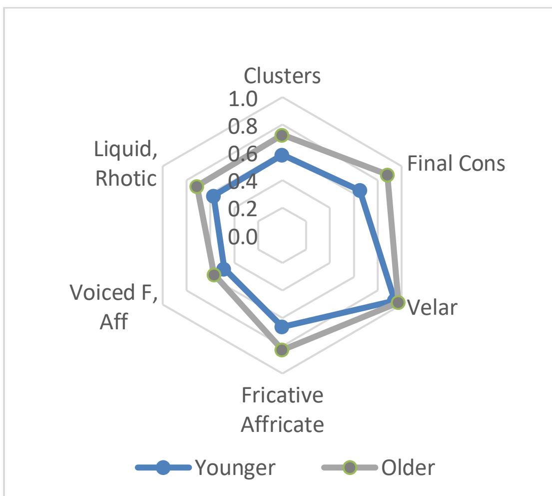
There are 8 parameters within the WCM-Ratio: (1) more than two syllables (2) non-initial stress (3) word-final consonants (4) consonant clusters (5) a velar consonant (6) a liquid, a syllabic liquid, or a rhotic vowel (7) fricative or affricate (8) voiced fricative or affricate. Table 5 shows the average WCM-Ratio scores of the 32 speech samples for these children with hearing loss. They showed the highest score on the WCM-Ratio for words with more than two syllables. However, only a few occurrences were observed in the WCM-Actuals and Targets across these children thus, the ratio was not reliable for comparing with the other parameters. The values of the parameter 'velar consonant' was higher than the other 6 parameters. Velar consonants were observed showing similar occurrences between WCM-Actual and Target for the children with hearing loss. Voiced fricative/affricates and consonant clusters were the lowest WCM-Ratios for these children. The children with hearing loss attempted to use words with voiced fricative/affricates, but actual productions of the children were much less observed in words as the children attempted them.

Table 5. Average values of the WCM-Ratio per 8 WCM parameters for the 32 speech samples and summary of the linear mixed models for PCC-R with the 8 parameters of WCM-Ratio and the Total WCM Ratio.

	Word Patterns		Syllable Patterns		Sound Patterns				
	> 2 syllables	Stress	Cluster	Final Cons	Velar	Fric, Affric	Voiced_Fric, Affric	Liq, Rhotic	Total Ratio
WCM-Ratio*	0.99	0.72	0.58	0.76	0.86	0.71	0.45	0.62	0.69
Estimate	0.13	0.20	0.73	0.59	0.50	0.66	0.297	0.548	0.852
R-squared	0.16	0.39	**0.82**	**0.72**	0.43	**0.70**	0.23	**0.72**	0.90
p-value	0.061	0.003	< 0.001	< 0.001	< 0.001	< 0.001	0.012	< 0.001	< 0.001

In the cross-sectional data set, the average values of the six parameters of WCM-Ratio for younger (n = 6, blue) and older (n = 7, silver) groups are shown in figure 8. The age range of the younger group is from 34 months to 63 months. The older group's age range is from 67 months to 97 months. The older group showed higher values on WCM-Ratio for all the parameters. However, the two groups showed almost the same scores on WCM-Ratio for the velar sounds (younger group: 0.94, older group: 0.97).

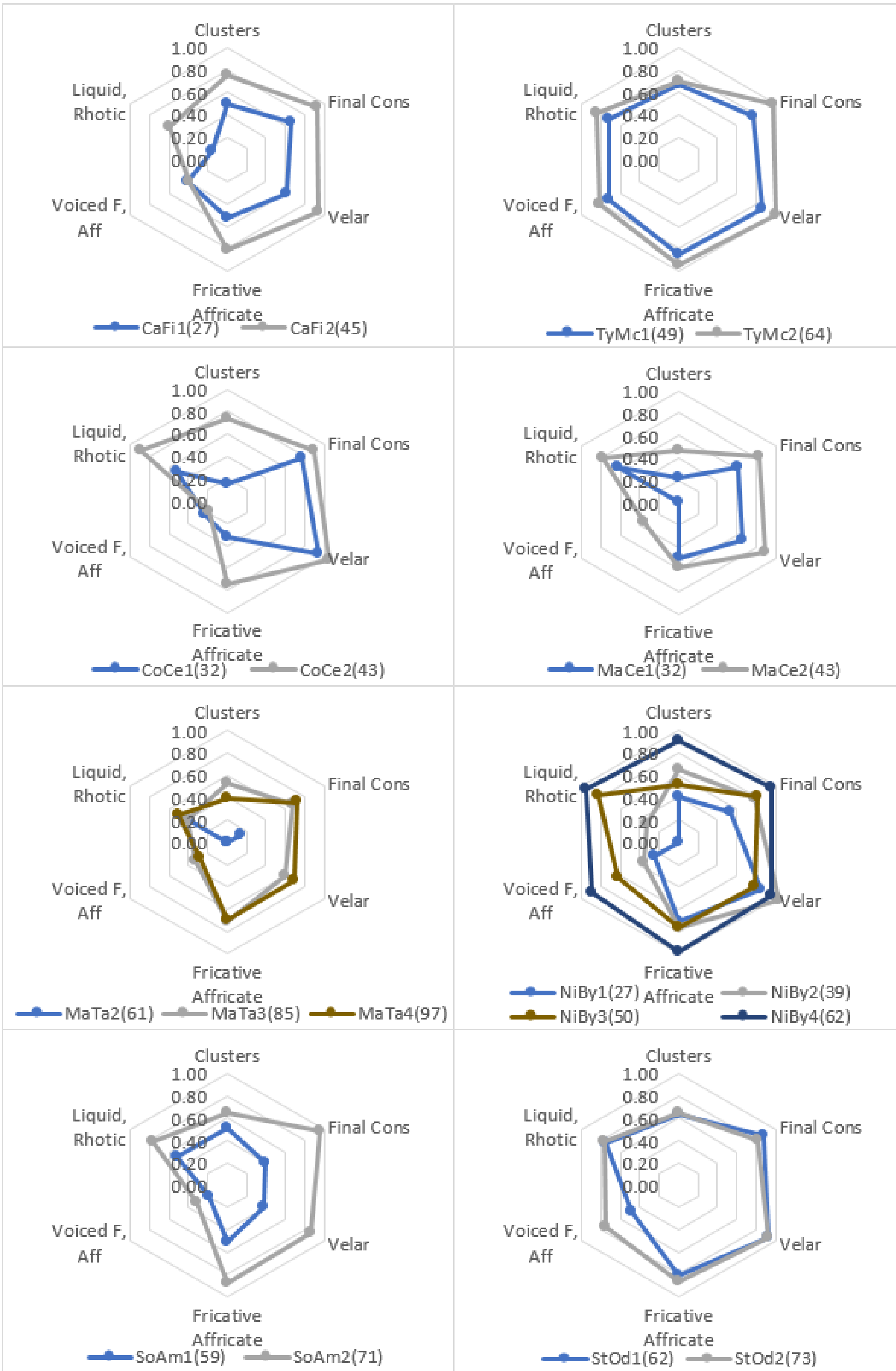
Figure 8. A radar chart of the WCM-Ratios for the 6 parameters of Syllable and Sound Patterns for the 13 speech samples in the cross-sectional data set. Blue line denotes younger children with hearing loss (n = 6) and silver line indicates older children with hearing loss (n = 7).



Within the longitudinal data set, the children demonstrated better performance on almost all 6 parameters of the WCM-Ratio when they were chronologically older (Figure 9). Two children showed higher WCM-Actual than WCM-Target values for velar sounds, thus the WCM-Ratio was more than 1.0 (CoCe2(43): 1.06 and NiBy2(39): 1.04). This is because both children showed a pattern of velar assimilation (e.g. / kɛkoʊ/ for “candle”).

However, NiBy did not produce velar assimilation in his later speech sample. When NiBy was 62 months old (5 years 2 months), his WCM-Actual was approximately the same as the WCM-Target even though his actual productions were not equal to his WCM-Targets across the 6 parameters. One other child, StOd, improved very little in use of consonant clusters, velars, fricative/affricates, and liquid//rhotics. He only made visually obvious progress in voiced fricative and affricate consonants (from 0.48 to 0.76).

Figure 9. Radar charts of the WCM-Ratios for the 6 parameters of Syllable and Sound Patterns in the longitudinal data set.



The results of linear mixed models with PCC-R and each parameter of the WCM-Ratio are reported in Table 5 above. Each model introduced its estimate, R-squared value and p-value for a fixed effect. R-square values for the linear mixed models across the 8 parameters of the WCM-Ratio were descriptively reported to evaluate relationships with PCC-R. The R-squared of a model with consonant cluster is 82%. This is the highest value among the 8 parameters of WCM-Ratios. Final consonants, liquid/rhotics, and fricative/affricates show high R-squared. Most consonant errors can be explained by deletion of sound(s) in consonant clusters, final consonant deletion, fricative/affricate and liquid/rhotic errors, among these 8 parameters for these children with hearing loss.

SUMMARY OF RESULTS FOR AIM 1.

Chronological age is a significant predictor for all speech measures; PCC-R, PVC-R, PWP, and WCM-Ratio. This outcome means that older children with hearing loss in this cohort achieve higher values than younger children for all speech measures. There is no significant difference between the HA and CI groups on any of these speech measures. However, for communication modality, PCC-R and PWP outcomes for the Talking Only and Total Communication groups are significantly different. The Talking Only group shows higher values of PCC-R and PWP than the Total Communication group. In contrast, PVC-R and WCM-Ratios do not show any significant difference between HA and CI groups related to the children's communication modality.

Compared to hearing children developing typically, these children with hearing loss showed lower PCC-R and PVC-R values. For the PCC-R – PWP intersect, these children with hearing loss can be divided into linear and Nonlinear groups. The Nonlinear group produced words with significantly greater PCC-R and MLU values than the linear group. This result indicates that even though the children in the Nonlinear group showed higher PCC-R and MLU values, their speech errors do not show any relationship between consonant correctness and word complexity. Because children in the Nonlinear group demonstrated PCC-R higher than the children in the linear group, it is hard to distinguish between speech delay and disorder based on the PCC-R-PWP analysis for this cohort.

Analysis of the relationship between the eight parameters of the WCM-Ratio and PCC-R, consonant clusters shows the greatest R-squared value. Final consonant, liquid/rhotic, and fricative/affricate also show high R-squared values. Word forms actually produced by the children who showed greater PCC-R tended to approximate more closely to the target forms. This was especially true for later acquired phonological features in children with typical hearing profiles such as consonant clusters, final consonants, fricatives, affricates, and liquids/rhotics.

4.2 LANGUAGE CHARACTERISTICS OF CHILDREN WITH HEARING LOSS (AIM 2).

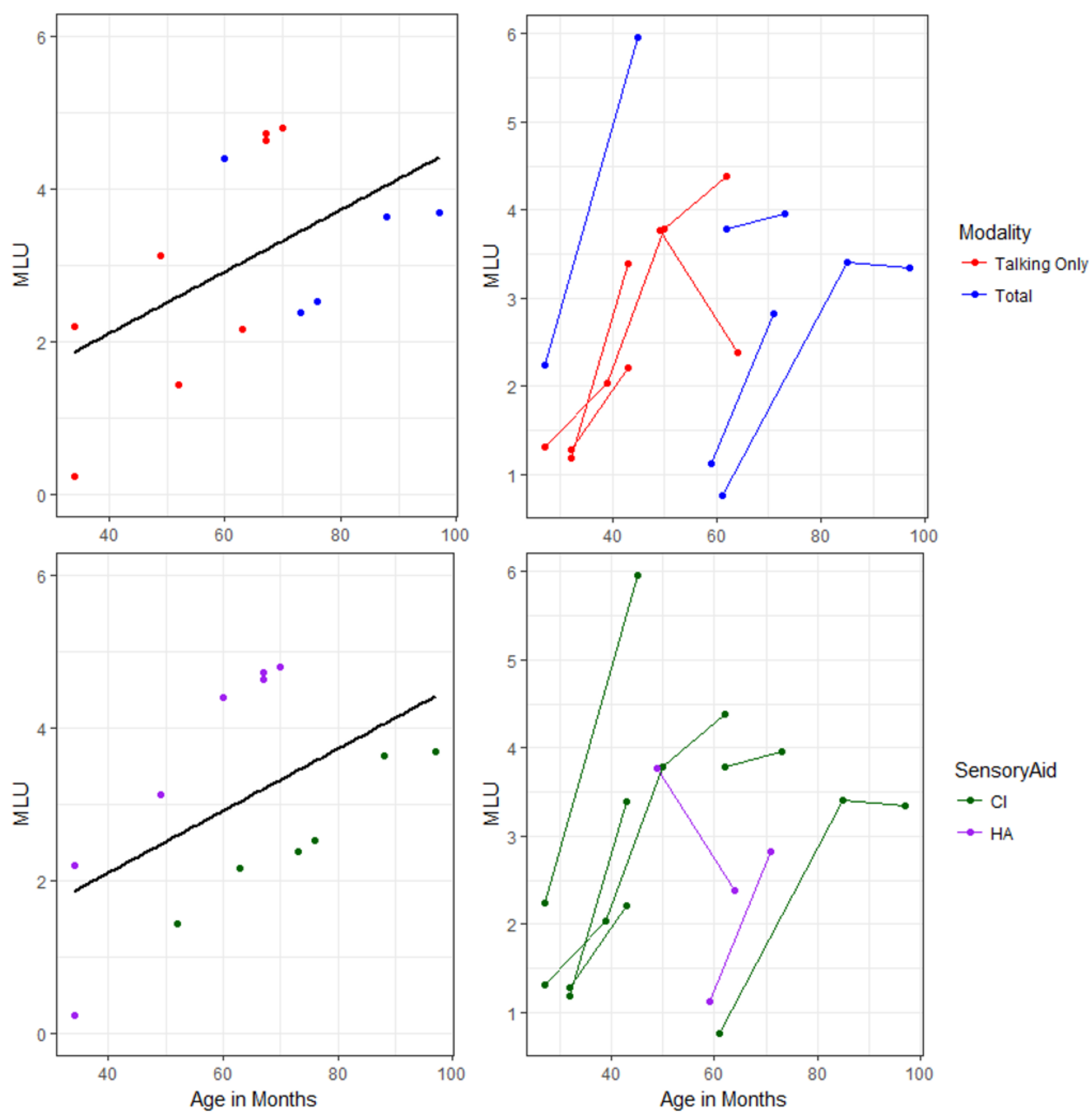
Language characteristics of children with children with hearing loss were examined using MLU, TTR, and the number of different word types. Like the speech measures, these language outcomes from the three language measures were described by type of sensory aids and communication modality using linear mixed models via the lme4 package (v1.1-14) in R (v3.4.2). In each model, a value of MLU, TTR or the number of different words for the 32 speech samples was the dependent variable. Fixed effects included chronological age, hearing device and communication modality and a random effect was participant. In this model, the intercept represents a reference condition indicating a value of dependent variable (e.g. MLU) for a child at the mean of chronological age in months, with cochlear implants, and in a speaking only group. The relationship of MLU was also examined using R-squared from linear mixed models.

4.2.1 Mean Length of Utterances (MLU)

The linear mixed model analysis revealed that age is a significant predictor of MLU for these children with hearing loss. Controlling for hearing device and communication modality, the older their chronological age, the higher score on MLU they exhibited. This is consistent with the speech measures described earlier. Specifically, for two children who have the same types of hearing device and communication modality, a child who is 1 month older than the other child is predicted to have 0.05 longer MLU than the other child (B = 0.050, t (10) = 3.487, p = .006). Figure 10 shows relationships between chronological age and MLU for children in cross-sectional data set (left) and longitudinal data set (right).

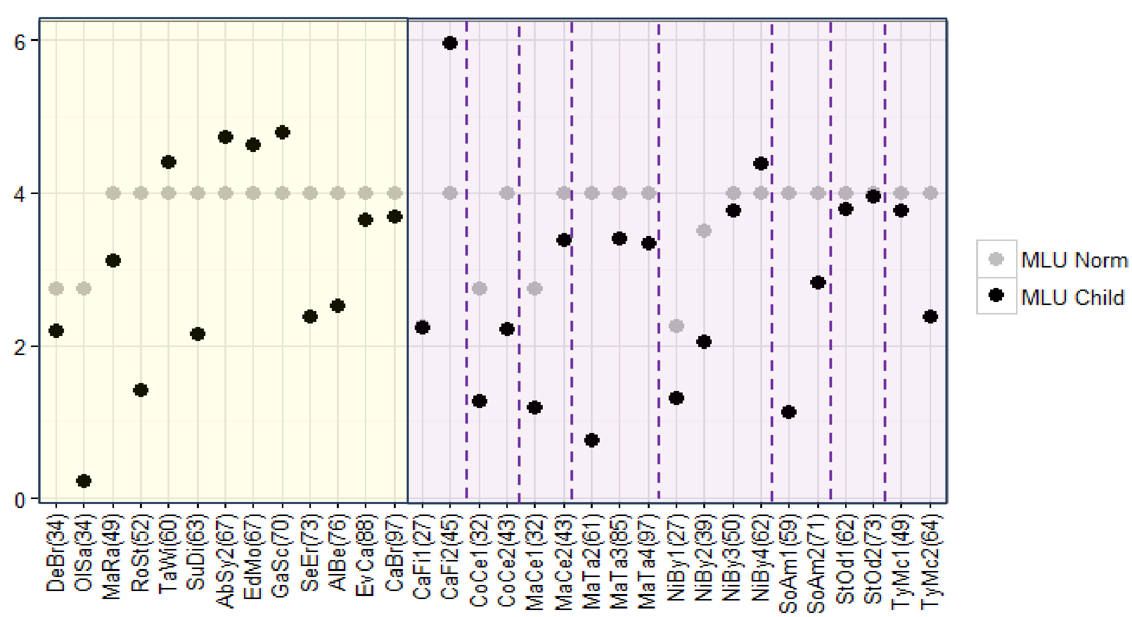
When the other predictors are controlled in this model, neither of the two categorical fixed effects, hearing device or communication modality, are significant predictors in the regression with MLU (hearing device: B = 0. 506, t (18) = 0.872, p = 0.395, communication modality: B = -0.395, t (18) = - 0.612, p = 0.549). The median value of MLU for the talking only group was 2.29 and the value of MLU for the total communication group was 3.37. The median value of MLU for the CI group was 2.52, and for the HA group was 3.13.

Figure 10. MLU across chronological age in months. The left graph includes cross-sectional child group data and the right graph includes longitudinal child group data. Red colored dots are children who use Talking Only and blue dots denote children who use Total Communication for the upper graphs. Purple colored dots are children with HA and green dots denote children with CI for the lower graphs. The linear black line in the left graph is a "smoothed conditional mean" to show trends in the raw data for MLU across chronological ages. The eight lines in the right graph connect data within a child in the longitudinal data sets for the 8 children.



MLU for the children with hearing loss was compared to normative data for hearing children (Brown, 1973). Figure 11 shows MLU for cross-sectional (yellow) and longitudinal data (purple) arranged by these children's chronological age within a data set. MLU values in children with CIs demonstrate delay compared to normative values for MLU based on typical developmental milestones (Brown, 1973). Six children caught up with values for age matched hearing peers. Within the longitudinal data set, children demonstrated better performance on MLU when they get older except for TyMc.

Figure 11. MLU ordered by chronological age in months. The left box (yellow) indicates children who had only one data sample (13 children, 13 samples). The right box (purple) shows the children who had more than 1 data sample (8 children, 19 samples). Dashed lines distinguish individuals within the longitudinal data group. Black dots are MLU children's scores. Gray dots are MLU norm for age in months of chronological age matched peers.



R-squared values for the linear mixed models examined the relationships between MLU and every speech measure, including PCC-R, PVC-R, PWP, WCM-Ratio, and one parameter of WCM-Ratio for final consonants were descriptively compared across the speech measures. Every speech measure is a significant predictor for MLU. This means that children who had greater scores on speech measures appeared to produce longer MLUs. The results of linear mixed models with MLU and each speech measure and the final consonant parameter of WCM-Ratios are reported in Table 6 below. The greatest value of R-squared from the models is 0.69 for WCM-Ratio. R-squared values for PWP is 0.66 and PCC-R is 0.63. The values of R-squared for all three speech measures (PCC-R, PWP, WCM-Ratio) are similar across measures. Only PVC-R showed lower values of R-squared for the model to predict MLU. WCM-Ratio for the final consonant is also a significant predictor for MLU and R-squared is 0.58 from the model.

Table 6. Regression results for each speech measure to examine relationship with MLU.

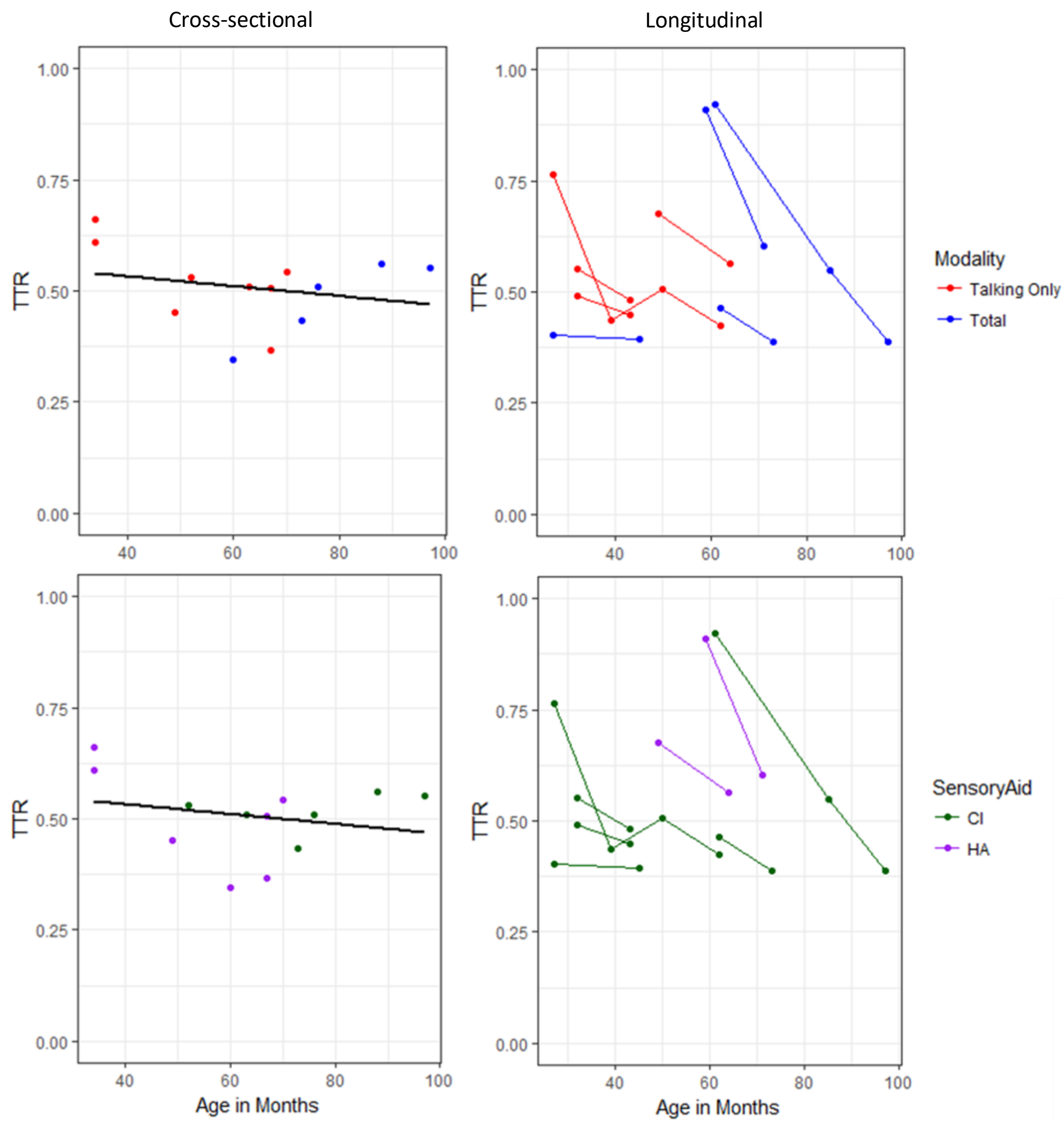
	Each speech measure as a fixed effect for MLU outcome				
	PCC-R	PVC-R	PWP	WCM-Total Ratio	WCM-FC-Ratio *
Estimate	5.70	6.11	9.14	5.50	4.47
R-squared	0.63	0.19	0.66	0.69	0.58
p-value	<0.001	0.021	< 0.001	< 0.001	< 0.001

*WCM-FC-Ratio is WCM-Ratio for the parameter of final consonant.

4.2.2. Type Token Ratio (TTR)

The linear mixed model showed that none of the three fixed effects are significant for TTR in these children with hearing loss [chronological age: B = -0.002, t (10) = -1.420, p = 0.186, hearing device: B = 0. 063, t (18) = 1.080, p = 0.294, communication modality: B = 0059, t (18) = 0.900, p = 0.380]. TTR appeared to decrease when children's chronological ages increase even though there was no significant effect of chronological age for TTR (see Figure 12). The median value of TTR for the talking only group was 0.51 and the value of TTR for the total communication group was 0.56. The median value of TTR for the CI group was 0.49, and for the HA group was 0.56.

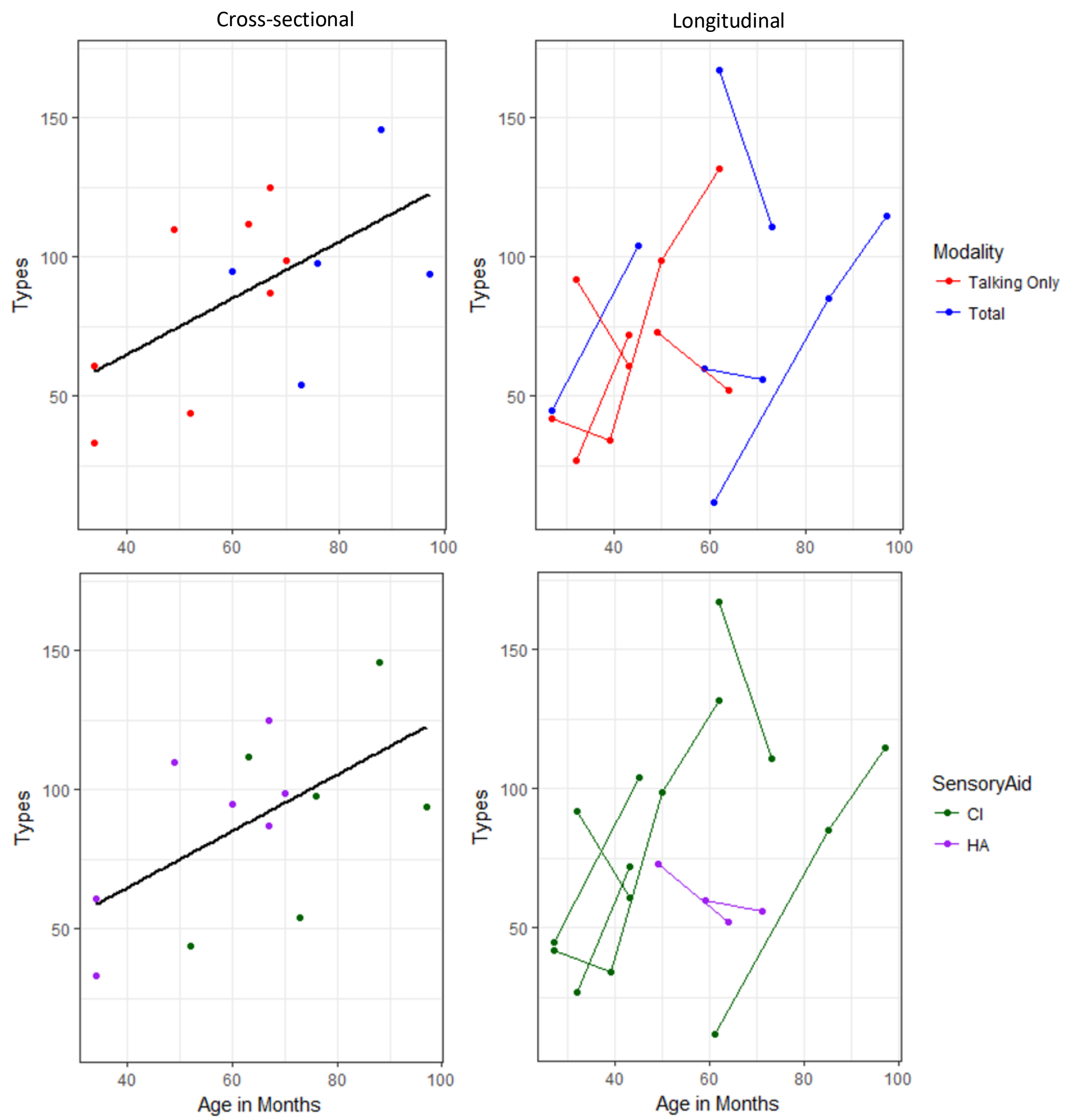
Figure 12. TTR across chronological age in months. The left graphs include cross-sectional child group data and the right graphs include longitudinal child group data. Red colored dots are children who use Talking Only. Blue dots denote children who use Total Communication. Purple colored dots are children with HA and green dots denote children with CI. The linear black line in the left graph is a "smoothed conditional mean" to show trends in the raw data for TTR across chronological ages. The eight lines in the right graph connect data within a child in the longitudinal data sets for the 8 children.



4.2.3 Number of different word types

The number of different words in a session was examined with the three fixed factors. Controlling for hearing device and communication modality, the older their chronological age, the more different words the children produced. Specifically, for two children who have the same type of hearing device and communication modality, a child who is 1 month older than the other child is predicted to produce one more different word than the other child (B = 1.235, t (10) = 3.299, p = .008). Neither fixed effects of hearing device or communication modality are significant [hearing device: B = -5.126, t (18) = -0.365, p = 0.719, communication modality: B = -12.605, t (18) = - 0.797, p = 0.436]. Figure 13 shows relationships between chronological age and the number of different words in a session for the children in cross-sectional (left) and longitudinal (right) data sets. The median value of the number of different word types for the talking only group was 72.5 and for the total communication group was 94.6. The median value of the number of different words for the CI group was 92, and for the HA group was 93.

Figure 13. The number of different word types across chronological age in months. The left graphs include cross-sectional child group data and the right graphs include longitudinal child group data. Red colored dots are children who use Talking Only. Blue dots denote children who use Total Communication. Purple colored dots are children with HA and green dots denote children with CI. The linear black line in the left graph is a "smoothed conditional mean" to show trends in the raw data for the number of different word types across chronological ages. The eight lines in the right graph connect data within a child in the longitudinal data sets for the 8 children.



R-squared values for the linear mixed models for the relationships between the number of different words and every speech measure including PCC-R, PVC-R, PWP and WCM-Ratio were descriptively compared across the speech measures. Except for PVC-R, each speech measure is a significant predictor for the number of different word types. This indicates that children who had greater scores on speech measures produced more different words in a session. The results of linear mixed models with the number of different word types and each speech are reported in Table 7 below. The greatest value of R-squared from the models is 0.47 for PWP. The R-squared value for PCC is 0.40 and for WCM-Ratio is 0.45. The values of R-squared for all three speech measures (PCC-R, PWP, WCM-Ratio) are similar for predicting vocabulary size. In addition, the three speech measures including PCC-R, PWP, and WCM-Ratio are significant predictors for the number of different word types respectively. Based on these findings, it could be expected that as a child's speech skills develop, the child increases his or her vocabulary size for these children with hearing loss.

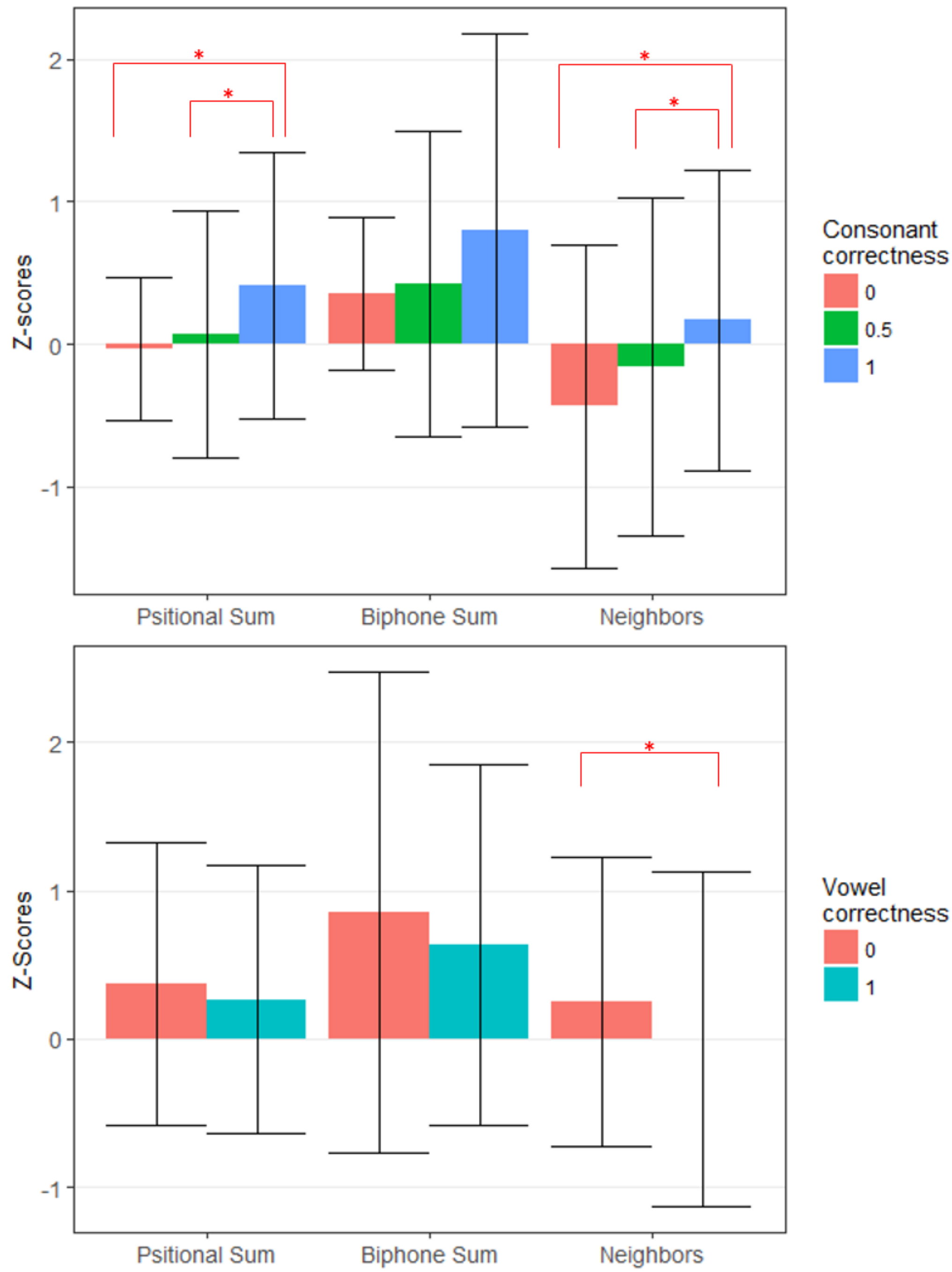
Table 7. Regression results for each speech measure to examine relationship with number of different words.

	Each speech measure as a fixed effect			
	PCC-R	PVC-R	PWP	WCM-Ratio
Estimate	123.63	93.57	201.30	124.02
R-squared	0.40	0.06	0.47	0.45
p-value	0.003	0.182	0.002	0.002

4.3. VOCABULARY INPUT FACTORS ON SPEECH OUTCOMES IN CHILDREN WITH HEARING LOSS (AIM 3).

One goal of the study was to examine the influence of neighborhood density and phonotactic probability on speech characteristics such as consonant correctness for children with hearing loss. The values of ND and PP for CVC words in the child corpora were analyzed from supplementary material from Storkel (2013) due to lack of availability of the comprehensive corpus. One speech sample (MaTa2) among 32 speech samples was excluded because it only included only one CVC word. The total number of CVC word tokens was 1,748. The number of CVC word tokens for each speech sample was between 13 and 144 across participants. Mean z-values of positional segment sum, biphone sum, and neighborhood density for each consonant correctness and vowel correctness score are reported in Figure 14 below.

Figure 14. Average values of positional segment sum, biphone sum, and neighborhood density for consonant correctness (upper) and vowel correctness (lower).



Consonant and vowel correctness were analyzed using mixed effects logistic regression models with the lme4 package (v1.1-14) in R (v3.4.2). The z scores for positional segment sum, biphone sum, and neighborhood density were tested as fixed effects and child was included as a random intercept. A CVC word form included three types of scores for consonant correctness (i.e. 0, 0.5, & 1) and two types of scores for vowel correctness (i.e. 0, 1). The dependent variable for consonant correctness is a multinomial variable (more than 2 levels in a categorical variable). Three linear mixed logistic models were performed to test the three contrasts for consonant correctness for phonotactic probability and neighborhood density. The number of CVC word tokens for the comparison between "1" and "0" was 1,225 in the regression model. The second regression model included 1623 CVC word tokens for the contrast between "0.5" and "1". The last regression model had 648 CVC word tokens for the contrast between "0" and "0.5" for the dependent variable.

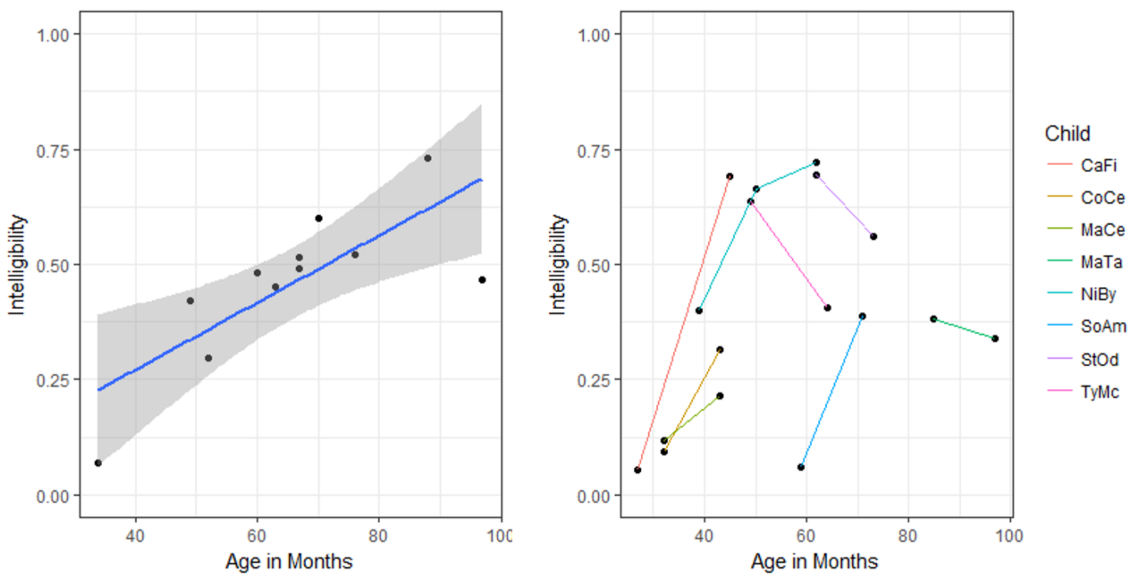
For every one-unit change in the positional segment sum, the odds of a CVC word with "1" of consonant correctness (versus "0": B = 1.718, z = 2.651, $p = 0.008$, and versus "0.5": B = 1.351, z = 2.735, $p = 0.006$) significantly increased when the biphone sum and neighborhood density are held at a fixed value. The positional segment sum was not a significant factor for the difference between "0.5" and "0" of consonant correctness (B = 1.213, z = 0.908, $p = 0.364$). For every unit change in the neighborhood density, the odds of CVC word with "1" of consonant correctness (versus "0": B = 1.614, z = 4.086, $p < 0.001$, and versus "0.5": B = 1.151, z = 2.209, $p < 0.027$) significantly increased when the positional segment sum and the biphone sum are held at a fixed value. Neighborhood density was not significant for the difference between “0” and “0.5” (B = 1.193 z = 1.771, p = 0.077). Biphone sum was not significant in consonant correctness across the three scores.

On the contrary, neighborhood density was negatively associated with vowel correctness. For every unit change in neighborhood density, the odds of CVC word with “1” of vowel correctness (versus “0”) significantly decreased, holding the positional segment sum and biphone sum (B = 0.797, z = -2.411, p = 0.016). The segment positional sum and the biphone sum were not significant in the difference between “1” and “0” of vowel correctness.

4.4. RELATIONSHIP OF LISTENER INTELLIGIBILITY FOR SPEECH MEASURES IN CHILDREN WITH HEARING LOSS (AIM 4).

Prior to a statistical analysis, individual variance in intelligibility judged by adult listeners was described. In the cross-sectional data set, chronologically older children with hearing loss tended to be perceived to have higher intelligibility scores than the chronologically younger children in the group. In the longitudinal set, 5 children with hearing loss improved their intelligibility judged by adult listeners over time whereas three children did not increase intelligibility values at their older ages (see figure 15). Three children (StOd, TyMc, & MaTa) demonstrated a decrease in intelligibility scores at older chronological ages. StOd who had CI demonstrated a decrease in his intelligibility at an older chronological age (from 70% at 5 years 2 months to 56% at 6 years 1 month). TyMc who had HA showed 64% of intelligibility at his early age (4 years 1 month) whereas he had 41% of intelligibility at older age (5 years 4 months). MaTa's intelligibility at her early age (7 years 1 month) was 38% whereas her intelligibility at her older age (8 years 1 month) was 34%.

Figure 15. Intelligibility score by chronological age for children with hearing loss in the cross-sectional data set (left) and longitudinal data set (right).

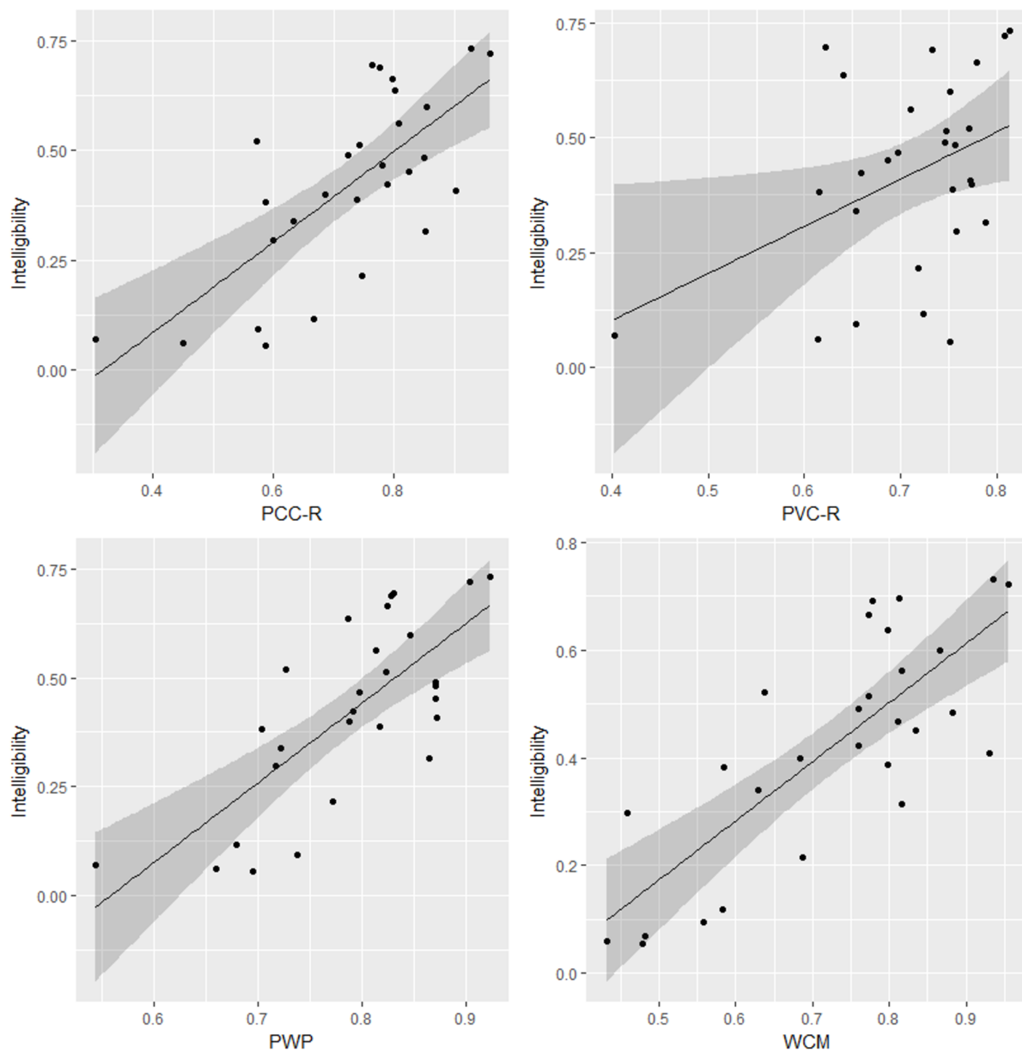


In the first analysis for the relationships between speech measures and intelligibility, a mean intelligibility value for each child was computed to evaluate relationships between intelligibility and speech values (i.e., PCC-R, PVC-R, PWP, & WCM). Twenty-eight speech samples for 19 children with hearing loss were included to analyze relationships between the four speech measures and intelligibility values. Fifteen utterances from each speech sample were selected. Preliminary statistical analysis was performed to compare the relationships with intelligibility across the speech measures. The range of the average scores for intelligibility was between 5% and 73% for these children with hearing loss. The average intelligibility score per child was preliminarily analyzed by using a linear model for each speech measure with the lme4 package (v1.1-14) in R (v3.4.2). The speech measure was a fixed effect within the model. A random intercept of child was excluded because there was no significant difference between the null model without a random intercept of child and the model with a random intercept (child) even though eight children had two speech samples. Results indicated that PCC, PWP, and WCM were highly correlated with each other (all three $r > .90$). This high correlation amongst independent variables causes multicollinearity (Miles & Shevlin, 2001). Multicollinearity makes it difficult to interpret the role of individual predictors in a regression model. Thus, four linear models were applied for each speech measure. The models were descriptively compared by using R-squared and AIC (Akaike Information Criterion). This analysis revealed that all speech measures were significant positive predictors for the intelligibility score. Among the four measures, WCM had the lowest AIC and the highest R-squared (see Table 8). Figure 16 demonstrates fitted models for the intelligibility score across the values of each speech measure. All speech measures showed positive relationships with intelligibility scores.

Table 8. Summary of models for speech measures on intelligibility

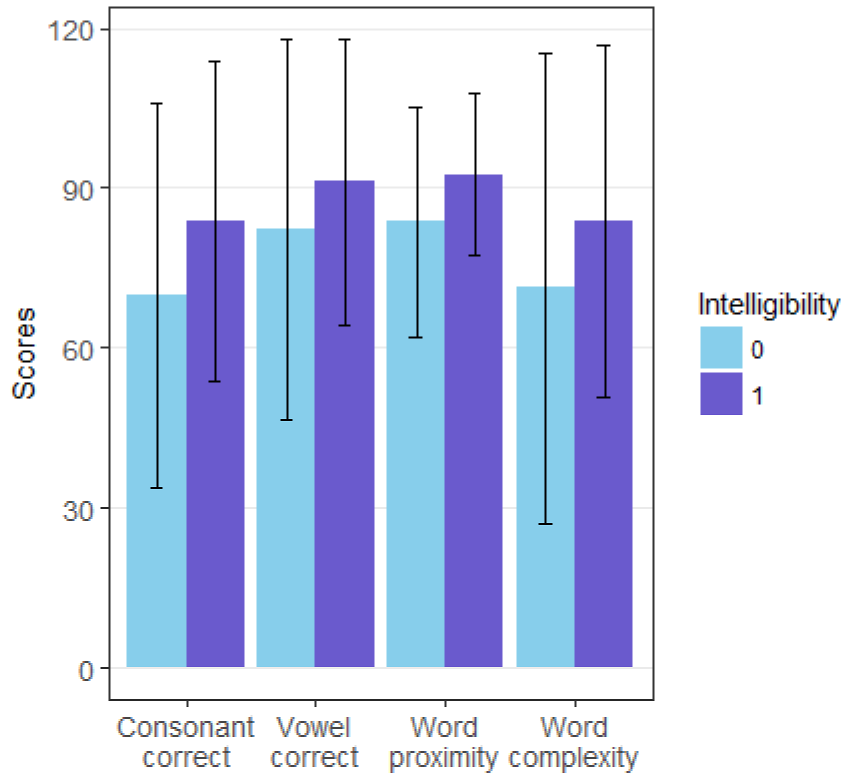
	Estimate	t value	P-value	R-squared	AIC
PCC-R	1.032	5.313	<0.001	0.52	-23.622
PVC-R	1.0259	2.303	0.030	0.17	-8.236
PWP	1.8314	5.68	<0.001	0.55	-25.63
WCM-Ratio	1.0975	6.53	<0.001	0.62	-30.221

Figure 16. Fitted linear models of intelligibility with each speech measure. Each linear line in the gray area denotes the confidence interval for predicted values.



Intelligibility was scored by word identification from adult listeners. To consider variance across listeners, four mixed effects logistic regression models were used with the lme4 package (v1.1-14) in R (v3.4.2) for each speech measure. Each mixed effects logistic regression included consonant correctness, vowel correctness, word proximity or word complexity as a fixed effect. Two random intercepts of child and listener were also added because the model fit was improved by adding these two random intercepts. The dependent variable for each model was intelligibility score coded by two values. Zero, "0" denotes an incorrect response and one, "1" denotes a correct response. Figure 17 shows scores of the four speech measures related to intelligibility scores. All the four speech measures for incorrect words have lower values than correct words.

Figure 17. Intelligibility scores across the speech outcomes. The Y-axis indicates the average score on each speech measure from zero to a hundred. The light blue bar (zero) denotes incorrect responses and the bright blue bar is correct responses on intelligibility.



All the four regressions indicated that the four speech measures are significant predictors for intelligibility. For one unit increase in consonant correctness (0.1), the odds of being correct in an intelligibility task increased by 13.9% (B = 1.139, z = 24.251, $p <$ 0.001). For one unit increase in vowel correctness (0.1), the odds of being correct in an intelligibility task increase by 10.9% (B = 1.109, z = 19.460, $p < 0.001$). When 0.1 of word proximity increased, the odds of being correct in an intelligibility task increased by 28.1 (B = 1.281, z = 27.310, $p < 0.001$). For one unit increase in word complexity (0.1), the odds of being correct in an intelligibility task increased by 6.3% (B = 1.063, z = 13.648, p

< 0.001). The model for word complexity has the smallest AIC (Akaike Information Criterion). The fixed effects across the four models are summarized in Table 9 below.

Table 9. Fixed effects for the four mixed logistic regression models

Predictor	Exp_estimate	Std.Error	Z	P-value	AIC
Consonant Correctness	1.139	0.005	24.251	<0.001	28275.9
Vowel Correctness	1.109	0.005	19.460	<0.001	30237.8
Word Proximity	1.281	0.009	27.310	<0.001	29826.2
Word Complexity	1.063	0.004	13.648	<0.001	24644.3

5. Discussion

Early identification and instrumentation has been found to positively impact achievement of age-appropriate speech and language skills, speech intelligibility, and general oral communication skills in children with significant hearing impairment identified early (e.g. Yoshinaga-Itano, 2013). However, despite early implantation, considerable individual differences in speech development occur across these children identified early with hearing loss and fit with hearing instrumentation. In particular, children with hearing loss have shown diverse profiles relative to acquisition of speech production patterns compared with normative values for children with typical hearing acuity. As a result, comprehensive approaches to evaluating and assessing spoken language abilities in children with hearing loss have important implications for understanding the basis of individual differences in their acquisition of spoken language abilities. While there are many studies of speech output in young children with hearing loss, most studies focus only on one particular area rather than on potential relationships between multiple capacities and measures within the same child, including speech and language output patterns, language input variables, and intelligibility judgements by listeners.

To evaluate comprehensively different outcomes in achievement of age appropriate speech and language skills in children with hearing loss, a view of both input and output capacities with various types of measures is extremely important for both researchers and clinicians. Accordingly, this study investigated whether speech and language acquisition patterns relate to each other and which speech measure is the most informative for assessing functional communication intelligibility for young children identified early with hearing loss. Speech patterns derived from contemporary speech measures were compared to language skills, including morpho-syntactic abilities.

Word characteristics related to phonotactic probability and neighborhood density based on frequency of input values were explored relative to speech accuracy as well. Previous work on these input characteristics in children developing typically has employed carefully controlled experimental tasks (e.g., Storkel & Lee, 2011). However, experimental paradigms may have limits relative to validly reflecting the natural process of word production and phonological acquisition in functional communication settings. Observational data from conversational speech samples can be uniquely informative for characterizing children's functional communication skills relative to consideration of how they reflect input values in word forms produced.

Listener ratings of speech intelligibility in spontaneous conversational language samples were analyzed. The goal was to reveal which speech measures most strongly correlate with intelligibility judgements by listeners. The broader goal of this dissertation was to more fully understand input-output relationships in speech development profiles of young children identified early with hearing loss. The impact of comprehensive characterization of the capacities of children with early sensory deficits can contribute to more refined assessment and intervention protocols.

Two child groups were followed; a longitudinal group for which two samples were gathered over time and a cross-sectional group where only one sample was available. Four research questions were posed related to the goal of increasing the comprehensiveness of characterization of speech acquisition patterns in young children with hearing loss: 1) What are speech characteristics for children with hearing loss compared to speech patterns observed in typically developing hearing peers? 2) What language features do children with hearing loss show, and how do their language capacities relate to speech patterns observed? 3) Are input characteristics present in a functional spontaneous child output corpus (Storkel, 2013) related to the children's speech accuracy? 4) Which speech measure is the most relevant to adult listener intelligibility judgement outcomes?

5.1. SPEECH AND LANGUAGE CHARACTERISTICS OF CHILDREN WITH HEARING LOSS (AIM 1 & 2).

To answer the first research question related to speech characteristics in children with hearing loss, four speech measures used in contemporary clinical and research were applied to analyze the two group's spontaneous speech samples: PCC-R, PVC-R, PWP, and WCM. In the longitudinal data group (i.e., two samples or more were available across development), children with hearing loss tended to improve their performance in speech skills over time based on results on the four speech measures. In the cross-sectional data group (i.e. only one sample was available), chronologically older children all showed speech scores higher than the chronologically younger children. This outcome indicates that early adaptation of contemporary hearing instruments can help children with hearing loss to improve their speech production capacities without significant deviations in the process of acquiring speech production abilities. Despite perturbations in early auditory processing, these children with hearing loss followed the developmental trajectory of speech acquisition. However, most of the children with hearing loss in the two groups did not achieve the normative score expected for children with typical hearing profiles for either PCC-R or PVC-R. These findings are consistent with a large body of previous work (e.g., Ertmer & Goffman, 2011; Moeller et al., 2007; Warner-Czyz & Davis, 2008). This delayed development in speech acquisition for the children with hearing loss reflects their limited auditory input due to hearing loss although most of the children have used hearing aids or cochlear implants before 3 years of age.

In the longitudinal data sets, children showed much greater increases in PCC-R and WCM-Ratios in comparison to their PVC-R and PWP values when they were chronologically older. Vowel productions account for PVC-R and PWP scores. PCC-R only considers consonant accuracy. Most of parameters of the WCM-Ratio also included consonant production. This outcome indicates that consonant productions were highly dependent on chronological age whereas growth of vowel productions was less sensitive to chronological age-related development in these children with hearing impairment. Results reflect the difference in the rate of acquisition of sound inventories between consonants and vowels. Accurate vowel production is achieved by 3 years of age in children with typical hearing profiles (Stoel-Gannon & Pollock, 2008, Pollock, 2002, Templin, 1957) whereas accuracy for some consonants (i.e. /s/, /r/) is not fully achieved by 7 years of age, even in typically developing children (Sander,1972).

The current study examined the PCC-R - PWP intersect analysis whether children with hearing loss exhibit speech delay (i.e. speech acquisition that is slower but looks like hearing children's profiles) or disorder (i.e., looks different from sound acquisition patterns reported for hearing children). To distinguish speech disorder from speech delay, children were divided into Linear and Nonlinear groups following Ingram (2015). The linear group included children who showed smaller differences between PCC-R and PWP for simpler words (i.e., words without consonant clusters) than for complex words (i.e., words with consonant clusters). In contrast, children in the Nonlinear group did not show any relationship between PCC-R - PWP intersect across the levels of word complexity. According to previous PCC-R - PWP intersect findings (Cummings, A., & Larson, 2015; Ingram, 2015), children with Speech Sound Disorder (SSD) who demonstrated Nonlinear PCC-R - PWP intersect patterns could be identified as having a speech sound "disorder" while children with SSD who had a Linear intersect pattern could be identified as having a speech sound "delay". However, the current study did not show any patterns to differentiate between "disorder" versus "delay" in this group of children with hearing loss. Instead, the Nonlinear group in this cohort showed higher PCC-R and PWP, and bigger MLU values in comparison to the Linear group. Children who produced consonants less accurately, demonstrated a relationship between consonant accuracy and word complexity. Children in the Linear group tended to produce words containing consonant clusters less accurately than words that did not contain consonant clusters whereas children in the Nonlinear group tended show values for word accuracy regardless of whether the word contained consonant clusters. The children with hearing loss who showed lower PCC-R scores in the Linear group might simply have speech delay instead of disorder. In the Nonlinear group, children who showed higher PCC-R scores but did not achieve the norm expected for PCC-R at their chronological age, may have difficulties in producing specific sounds regardless of word complexity related to their precise level of perceptual access.

Within the longitudinal data set, 6 out of 8 children demonstrated non-linearity in their later occurring speech samples. Thus, as those 6 children increased PCC-R values, they did not show any relation between word complexity (i.e. singleton versus clusters) and consonant correctness. However, 2 out of 8 of these children demonstrated linearity when they achieved higher scores on PCC-R. The two children did not produce any multi-syllable words with clusters in their speech sample at an older chronological age.

Ingram (2015) analyzed the PCC-R - PWP intersect for children who were diagnosed with speech sound disorders. He did not report their PCC scores, but the young children he studied might likely have much lower PCC values relative to their chronological age based on their diagnosis of SSD. Ingram’s earlier pilot study (2012) reported PCC-PWP relationships for children diagnosed with sound speech disorders. The highest PCC score in his Nonlinear group was less than 40%. PCC for his Linear group ranged from 20 % to 80%. A mean value of PCC-R for the children with hearing loss in the Linear group was 64% and the value of PCC-R for the Nonlinear group was 76%. Their PCC scores were much higher than the scores for the children in Ingram’s pilot study.

In addition, previous studies have tested words equally from monosyllabic words with singletons (i.e., no-consonant clusters) to multisyllabic words with consonant clusters based on formal single word assessments. Ingram (2012) used speech data from the *Assessment of Phonological Processes* (taken from Hodson, 1983, 1991). Cummings and Larson (2015) used speech outcomes from the *GFTA-2* (Goldman & Fristoe, 2000). Speech outcomes in the current study were collected via spontaneous conversational speech sampling only. Multisyllabic words with consonant clusters might not occur often enough to evaluate complex words related to the children's language and vocabulary use in their spontaneous functional communication. This lack of distribution of word types may be a drawback to spontaneous speech sampling because the child may not have the opportunity or ability to express many multisyllable words in a play sample. Both spontaneous speech sampling and formal word-based speech sound testing would be optimal in assessment protocols with these children to make a more decisive diagnosis of speech delay versus speech disorder. This approach would require that the children were developmentally ready to comply with formalized testing that includes naming pictures.

The results of the PCC-R - PWP intersect analysis in these children with hearing loss revealed that when a child with hearing loss showed low accuracy of consonant production, consonants errors observed can be dependent on word complexity. When a child can produce consonants relatively more accurately, the child may have difficulty producing a certain sound regardless of word complexity. According to these results, clinicians should consider relationships between word complexity and sound accuracy to evaluate and plan intervention for children with hearing loss. This analysis of relationship between word complexity and sound accuracy could be another resource to assess the child's speech production patterns more precisely.

The WCM-Ratio was one of the speech measures employed to examine whole word complexity (Stoel-Gammon, 2010). The parameters of word patterns were less meaningful and weakly reliable for the ratio analysis because only a few occurrences were observed in the WCM-Actuals (i.e. a child's actual form) and Targets (i.e. the adult form for the word a child produced) across this cohort of children with hearing loss. Among the syllable and sound class parameters, the velar consonant parameter was the highest value for these children with hearing loss. Velar consonants are produced at a 90% level for all typically developing children by 4 years of age. In contrast, other phonemes present within these WCM parameters related to sound patterns (i.e., fricatives, affricates, liquids & rhotic vowels) are mastered after 4 years of age (Templin, 1957; Wellman et al., 1931). In the cross-sectional data set, both chronologically younger and older children produced velar consonants almost equally relative to how frequently they attempted those sounds. In the longitudinal data set, most of children demonstrated large improvement in production of velar consonants relative to their WCM-Ratio score. Overall, the 32 speech samples analyzed for these children with hearing loss showed the lowest WCM-Ratios for voiced fricative/affricates and consonant clusters. These phonemes are mastered at 90% level for all typically developing children between 4 and 8 years of age (Templin, 1957; Wellman, Case, Mengert, & Bradbury, 1931; Waring, Fisher, Atkin, 2001). The lower scores for later developing consonant sounds and sound combinations are consistent with typically developing children's speech production patterns. The results indicate that children with hearing loss can follow the developmental trajectory of speech production even though they may show delayed development in speech production due to limited early ambient language input.

According to the results of the linear mixed model analysis to examine the relationships between PCC-R and parameters of the WCM-Ratio, most consonant errors can be explained by consonant clusters, final consonants, liquid/rhotics, and fricative/affricates among the 8 WCM-Ratio parameters for these children with hearing loss. Consonant error patterns for these children with hearing loss are related to errors on final fricative/affricate and liquid/rhotic sounds. This outcome indicates that characteristics of speech patterns for these children with hearing loss could be validly described using analysis results of the relationships between PCC-R and parameters of the WCM-Ratio.

The PCC or PCC-R measures based on conversational speech sampling reflect a pervasively used segment-oriented measure for characterizing phonological development. PCC-R simply evaluates a child's phonological development for consonants based on normative values for accuracy by chronological age for typically developing hearing children (Campbell et al., 2007). However, the PCC-R measurement is unable to capture details of the types of errors such as sound classes or syllable shapes because the measure is simply computed by identifying correct/incorrect consonants in a word. By employing the WCM-Ratio, clinicians can compare the more detailed phonetic and phonological aspects of production patterns necessary to plan each child's intervention protocol. Analysis of relationships between the PCC-R and WCM-Ratio values consequently has important and positive implications for planning assessment and intervention protocols for these children with early identified hearing loss. The combined analyses describe target sound patterns that a child needs to work on for improving his or her speech patterns toward mastery of ambient language sound patterns.

According to the results of all the four speech measures for these children with hearing loss, clinicians should understand that children with hearing loss demonstrated delayed consonant and vowel acquisitions despite early implantation. Based on the relationship between PCC-R and whole word measures, consonant errors should be evaluated in relationship to word complexity. A whole word approach is necessary to understand speech characteristics and provide appropriate interventions for children with hearing loss.

Speech outcomes on these speech measures were analyzed relative to each child's chronological age, hearing device and communication modality to consider the variations observed in their speech skills. For these environmental factors, the Talking Only group showed significantly higher PCC-R and PWP values than children using Total Communication. This finding is consistent with a recently published study (Geers et al., 2017). Geers and colleagues examined effects of long-term use of sign. Their results revealed that children who used Total Communication exhibited lower speech skills compared with speech performance in a Talking Only group of children. In the current data corpus, chronologically older children (mean age: 70 months) tended to use both sign language and verbal communication whereas younger children (mean age: 49 months) used only spoken English without sign language. Even though their chronological age was controlled to examine for a communication modality effect on their speech outcomes, the discrepancy in chronological age between Talking Only and Total Communication groups should not be ignored. Further data collection is necessary to establish communication modality effects on speech skills for children with hearing loss.

Analysis from all four speech measures and three language measures showed there are no significant differences in speech scores between children with HA and with CI instruments. Blamey et al. (2001) revealed that both CI and HA groups showed approximate performance in speech perception, production, and language scores while an average hearing loss for the children with CI (106dB HL) was higher than an average hearing loss for children with HA (78dB HL). The current study also reported consistent results with Blamey and colleagues (2001) by showing no differences between children with HA and CI. In the current study, all of the children with CI had profound deafness whereas the children with HA had from mild to severe levels. The current FDA guidelines permit cochlear implantation in children who are 2 years and older with severe-to-profound deafness (i.e., pure tone average thresholds of 70 dB HL or greater), and in children 12 to 23 months of age with profound deafness (i.e., pure tone average thresholds of 90 dB HL or greater) (ASHA, 2004). Hearing level for recipients of HA is between mild and severe. Thus, it is hard to establish an equivalent hearing level to compare speech outcomes between children with HA and CI. In addition, Fitzpatrick, Olds, Gaboury, McCrae and colleagues (2012) reported inconsistent speech and language outcomes between children with CI and HA. Fitzpatrick et al. (2012) reported that children with HA (severe hearing loss) performed better than children with CI (profound hearing loss) on measures of receptive vocabulary, language, and phonological memory. Further investigation is required to compare effectiveness of hearing devices between CI and HA related to the metrics measured in this study.

For these children, the growth of speech performance across the speech measures was individually varied. One child in the longitudinal data set achieved a normative PCC-R and MLU values relative to typically developing children when he was 5 years old. This child with CI who used Oral Communication only (NiBy) had age appropriate language ability based on his *Preschool Language Scale*, fourth edition (PLS-4) at the ages of both 4 and 5 years. Based on the measures gathered, he achieved age appropriate language and speech skills by the age of 5 years. He received a CI when he was one year old. His early identification and intervention may have positively affected his speech and language development. On the other hand, another child with CI (MaTa) increased her PCC-R score over time, but still showed a moderate delay in speech development. MaTa received CIs when she was 4 years 2 months of age. Her age at implantation is relatively older compared to the other children in this cohort. The relatively late implantation may be associated with her moderate speech development delay. The finding of early identification benefit on speech development from these individual children with hearing loss is consistent with findings in numerous previous studies (e.g. Connor et al., 2006; Kirk et al., 2002).

Two children in the longitudinal data group did not show growth in PVC-R scores over time. One child (CaFi) did not improve his PVC-R (i.e., it was 75% to 73% over time). In contrast, he achieved a steep increase in MLU score from 2.23 to 5.96 in that same period. He produced the longest MLU among the children. His longer sentence use potentially reduced vowel accuracy once he began to produce longer utterances rapidly in functional communication. However, his PCC-R improved over time even though he did not achieve an age-appropriate level of PCC-R. One other child (MaCe) also did not improve vowel accuracy at an older chronological age (both PVC-R scores were 72%). However, he also demonstrated a relatively abrupt increase in MLU score from 1.19 to 3.39. These two children illustrate profiles that might support further intervention for vowel accuracy.

The second research question evaluated language characteristics and relationships between language capacities and speech patterns. Mean Length of Utterance (MLU, Brown, 1973), and Type Token Ratio (TTR, Templin, 1957) were applied to examine language characteristics. Older children with hearing loss in this cohort produced longer sentences shown in higher MLU values. This result is consistent with the chronological age effect on speech measures described earlier. In the longitudinal data group, 7 of 8 children with hearing loss increased their length of utterance over time. The length of utterance for TyMc abruptly decreased from 3.78 to 2.38 as his chronological age increased from 49 months to 64 months. However, his first speech sample was collected with his mother at his home in an exceptional way. His second speech sample was collected by graduate clinicians in the UTSHC as usual. This result implies that a child might produce longer utterances when he/she communicates with a familiar communication partner in a familiar context. This environmental factor indicates the potential importance of familiar versus unfamiliar communication partners for maximizing spontaneous speech outcomes of these children within functional communication settings.

In the cross-sectional data group, chronologically older children achieved longer utterances. However, only six of 21 children caught up with values for age matched hearing peers. These results are consistent with previous studies where children with CI increased MLU values over time even though the children with CI continued to demonstrate delayed patterns relative to their chronological age (Flipsen & Kangas, 2014) and children with HA showed delayed development in their morpho-syntactic abilities as measured by MLU (Koehlinger et al., 2013).

According to the results of the linear mixed model analysis to examine the relationships between MLU and speech measures, PCC-R, PVC-R, PWP, and WCM-Ratio, all four speech measures are significant predictors for MLU respectively. Children who showed higher values on these speech measures, produced longer utterance lengths. However, it is hard to simply indicate which speech measure is the most highly correlated to MLU based on the results of the models because the values of the R-squared for all three measures including WCM-Ratio, PWP, and PCC-R are approximate.

In this regard, children with hearing loss are likely to be less accurate in producing phonological forms, especially final fricative consonants (e.g. /s/ and /z/), implicated in English morphology (Moeller, 2013). The WCM-Ratio for the final consonants in children with hearing loss might be associated with their MLU outcomes because the parameter of WCM-Ratio is related to final consonant deletion. The lower score on WCM-Ratio for final consonants means that the adult form of the word a child tries to produce has a final consonant, but the final consonant is not observed the child's actual production of the word. The result of the linear mixed model with WCM-Ratio of the final consonants for MLU demonstrated that the parameter of final consonant is a significant predictor for MLU. This finding indicates that production of final consonants should be considered as a potential factor in examining children's grammatical abilities in clinical settings. In addition, the parameter of WCM-Ratio can provide a valid tool for clinicians to catch latent phonological factors relative to child's language abilities.

TTR is the only measure among all speech and language measures, which did not show any significant chronological age effect. Even though there is no age effect reflected by TTR, children with hearing loss appeared to decrease TTR as they got older especially in the longitudinal data set. The linear mixed model with the number of different word types revealed chronological age is a significant predictor for TTR. This result indicates that children produced different word types more frequently as they got older. Also, the children increased more steeply in producing word tokens compared to word types. This outcome results in a declining TTR pattern for these children with hearing loss. The results for TTR and the number of different word types indicate that the number of different word types can be a more valid measure to describe vocabulary growth for children with hearing loss than word tokens.

5.2 Vocabulary input factors on speech outcome in children with hearing loss (Aim 3).

The third research question focused on investigating effects of input factors including phonotactic probability (i.e. positional segment sum, biphone sum) and neighborhood density on speech production. Phonotactic probability is the relative frequency with which individual sounds and sound sequences occur in syllables and words (Jusczyk & Luce, 2002). Words that contain highly likely occurring sounds and sound sequences in a given language are referred to as having high phonotactic probability whereas words that contain rarely occur in the lexicon are considered to having low phonotactic probability. The current study preliminarily examined these input factors only for consonant and vowel correctness of CVC word forms due to the status of the database used for analysis (Storkel & Hoover, 2010). These children with hearing loss increased in their probability of producing CVC words with two correct consonants compared to words with no correct consonants for a one-unit increase in phonotactic probability (i.e. positional segment sum). This finding indicates that the children with hearing loss produced CVC words that contain frequently occurring sounds more accurately than words with infrequent sounds, consistent with previous work that has shown a facilitative effect of phonotactic probability in word learning and speech production of children with typical hearing profiles (Storkel, 2001 & 2003, Zamuner, 2009; Edwards et al., 2004).

Neighborhood density is defined by the number of words that differ by one phoneme from a given word (Vitevitch & Luce, 1999). Words with high neighborhood density have many neighbors whereas words with low neighborhood density contain few neighbors in the lexicon. The result of this input analysis revealed that children are likely to produce accurate consonants in CVC words within dense neighborhoods. This indicates facilitative effects of high neighborhood density on consonant accuracy in functionally employed CVC words within conversational samples. Previous research mainly demonstrated increasing neighborhood density is associated with less accurate and slow processing in word recognition and word learning for typically developing children (German & Newman, 2002; Munson, Edwards, & Beckman, 2005; Heisler & Goffman, 2016). Based on inhibitory effects of neighborhood density in the previous research, researchers argued that when lexical word forms are strongly activated by the input of a real word, competition occurs among the phonological word forms. The discrepancy between the current study of children early identified with hearing loss and previous studies with typically developing hearing children indicates that children with hearing loss may have less dense neighbors for the same word compared to hearing peers. This difference may cause lack of competition due to less neighbors for a word when they want to activate the word. However, this potential lack of competition in children with hearing loss is not sufficient to explain more accurate consonant production for words with high neighborhood density compared to words with low neighborhood density. Children with hearing loss may have different input experiences from their daily communication environment compared to hearing peers. For example, parents of children with hearing loss may potentially over-articulate every word for their children to perceive their spoken words regardless of word characteristics. Further research including child-directed speech directed to children with hearing loss is required to examine their input characteristics more precisely.

The paradigm in this study could be another aspect to explain different results between the current study and previous studies. Previous studies of spoken language processing controlling lexical and phonological values featured designs which dichotomized words into high and low phonotactic probability and neighborhood density categories. Experimental designs that include dividing high and low phonotactic probability and neighborhood density may disregard the possibility of a continuous set of values across the whole spectrum of phonological and lexical representation of speech and language. This issue should be also examined further using experimental designs. Despite the inhibitory effects of neighborhood density on word recognition and learning, some studies have shown facilitative effects of neighborhood density on speech production in more naturalistic settings. For example, Sosa and Stoel-Gammon (2012) demonstrated positive relationships between neighborhood density and word proximity (Whole Word Proximity; PWP, Ingram, 2002). They elicited five productions of target words and found that words with dense neighborhoods had higher proximity, but no phonotactic probability advantages on PWP. Storkel (2004) showed that early acquired words come from dense neighborhoods whereas later acquired words have sparse neighborhoods according to parent report using the *MacArthur Communicative Development Inventories* (M-CDI, Fenson et al., 2006) in children with intact hearing capacities. Children with CIs were also found to be more likely to learn words with dense neighborhoods based on M-CDI results and the *Expressive One-Word Picture Vocabulary Test-3* (EOWPVT-3, Brownell, 2000). The neighborhood advantage for the children with hearing loss in this preliminary study is consistent with the previous research within a context of more natural processes of word production studies.

On the contrary, the results of vowel correctness relative to input factors demonstrated an inhibitory effect of neighborhood density. Gahl, Yao, and Johnson (2012) demonstrated that words with many neighbors were phonetically reduced in production in a study of adult participants with typical hearing profiles. They analyzed phonetic variables including word duration and vowel dispersion in connected speech. Vowels were shortened in duration and produced as more centralized in connected speech production from hearing adults. The current study provides evidence of the reduction of vowel accuracy for words with high neighborhood density by showing decreased neighborhood density for CVC words with a correct vowel in children with hearing loss.

However, Munson and Solomon (2004) reported an inconsistent result of neighborhood density on speech outcomes in ten hearing adults for selected CVC words. Thirty CVC words were categorized into lexically difficult words (low frequency with dense neighborhoods) and lexically easy words (high frequency with low neighborhood density). These words were selected based on lexical competition on speech production. Vowel spaces for high neighborhood density words were more expanded than those for low density words. They explained that due to difficulty in perception of words with dense neighborhoods, speakers tended to produce clearer speech by showing more expanded vowel spaces in order to facilitate speech perception for listeners. Both previous studies only reported relationships between vowel production and neighborhood density with hearing adults and focused on acoustic properties of vowel production. Further study is needed of child populations, particularly of children with hearing impairment to establish whether inhibitory or facilitative effects of neighborhood density occur in child populations. While typically hearing adults reflect listeners' difficulty in perception of words with dense neighborhoods by showing more expanded vowel spaces in order to facilitate speech perception for listeners, children with hearing loss may have less sensitive to responding to needed expansions for supporting perception in their listeners.

Relative to clinical assessment and intervention planning, it is important for clinicians that children with hearing loss can more accurately produce consonants of words with frequently occurring phoneme combinations compared to words with less frequently occurring phoneme combinations. That patterning is consistent with typically developing hearing children. At the same time, clinicians should explore the possibility of different lexical acquisition in children with hearing loss based on facilitative effects of neighborhood density on consonant and vowel production. A more practical guideline could be suggested based on the results of the current study. If the facilitative neighborhood effect on consonant accuracy is noted in children with hearing loss, clinicians can select words with more neighbors first and then words with lower numbers of neighbors when their goal is to start intervention with 'easier' words.

5.3 Relationships between Intelligibility and Speech measures of children with hearing loss (Aim 4).

Intelligible speech production is an important functional component for understanding oral communication competence in speaker-listener interactions for children with hearing loss. Intelligibility scores can be a tool to examine hearing devices, to build goals in speech intervention, and to assess progress on successful oral communication for children with hearing loss who use hearing devices (Ertmer, 2010 & 2011). An aim of this study was to consider the best speech measures to predict intelligibility. Intelligibility scores were measured using word identification by typical adult listeners for children with hearing loss. Two analysis procedures were performed for the intelligibility data.

The first analysis of the intelligibility data considered the relationship of mean values for intelligibility relative to each of the four speech measures per child. All speech measures, including PCC-R, PVC-R, PWP, and WCM-Ratio, were significant predictors for intelligibility scores for this cohort of children with hearing loss. Speech outcomes with higher values of PCC-R, PVC-R, PWP or WCM-Ratio tended to be judged as more intelligible for these children. In addition, the r-squared values for PCC-R, PWP, and WCM-Ratio were higher than .5. This outcome means that more than 50% of the intelligibility scores can be explained by each speech measure respectively. Ertmer (2010) reported a weak association between the single word articulation test results and connected speech intelligibility. These speech measures for connected speech production could provide an indirect but function-based way to assess intelligibility in children with hearing loss.

The linear regression model with WCM-Ratio showed the highest R-squared and the lowest AIC values among the four models with the speech measures. This result indicates that the model fit with WCM-Ratio is better than the other models with PCC-R, PVC-R, or PWP. Based on this result, WCM-Ratio appears the most strongly related to intelligibility outcomes for the current cohort of children with hearing loss. WCM quantifies the presence of phonological characteristics at the word and syllable levels in addition to sound class parameters. WCM results also emphasize analysis results for later developing phonological characteristics (e.g. fricatives, affricates, consonant clusters, etc.). The WCM-Ratio is computed by ratios comparing overall occurrences of each parameter in the two samples of ‘target’ and ‘actual’ words. For example, when a child produces more fricative sounds, he or she has a higher WCM-Actual. In this case, the WCM-Ratio for the child would be also higher. Thus, when children with hearing loss produce more complex syllable structures and later developing sounds, their speech production is identified as more intelligible by listeners.

The second analysis of the intelligibility data was performed using correct/incorrect responses for each word instead of using an average score for intelligibility. All of the four speech measures were also calculated using word by word comparisons. The mixed logistic regression models showed all speech measures are significant predictors for intelligibility scores. Words with more correct consonants and vowels are likely to be judged more intelligible by listeners. More intelligible words appear to have a higher ratio of actual word forms (i.e., what a child actually produced) with target word forms (i.e., adult forms that a child attempted to produce) in terms of word and syllable structure and sound complexity. Odds ratios across the four speech measures were descriptively compared to find the best predictor for listener intelligibility judgements for these children with hearing loss. The highest odds of being correct in an intelligibility task was word proximity compared to consonant correctness, vowel correctness, and word complexity. Word proximity considers word length and syllable complexity as well as consonant accuracy. Thus, if a word has more complex syllable structure and more correct consonants, the probability of being correctly identified by listeners increases. Ingram (2002) proposes that PWP is a potential indicator for intelligibility. The consistency of syllable structures and consonant correctness between target word and child's actual production could reflect intelligibility based on the result. However, based on comparing AIC values across the four regression models, the model for the WCM-Ratio had the smallest AIC. AIC is an estimator of the relative quality of statistical models for a given set of data. The smaller AIC is, the better model fit is. The model with WCM-Ratio has the best model fit among the four models even though for one-unit increase in WCM, the probability of being less correctly identified by listeners increases compared to the other measures.

The first analysis leads us to pick the WCM-Ratio as the best predictor for intelligibility whereas the second analysis introduces word proximity as an additional important variable. The discrepancy of these two results could be attributed to adaptation in calculating WCM-Ratio for the second analysis. Stoel-Gammon (2010) introduced how to calculate a WCM-Ratio procedure. She clearly stated that the ratios are not based on word-by-word comparisons. We should compute the ratio based on the total numbers of occurrences of each parameter in the word target and actual production samples respectively. However, in the second analysis, word by word comparison is inevitable if a regression model needs to assess the reliable effects taking into account listener variation. The result of the second analysis for the mixed logistic regression with word complexity showed the smallest estimate of word complexity compared to the other three measures. This result should be carefully interpreted to generalize the result to the original speech measure relative to intelligibility.

However, even though the two analyses did not show identical results, both results highlight the use of whole word approaches within available speech measures to provide reliable estimates of intelligibility for children with hearing loss. Both PWP and WCM-Ratio score complexity of word and syllable structures. These two measures give a higher score to a child who produces more complex words than simple words when the child's actual production matches target words. The capacity for producing longer words with complex structures can support more intelligible speech production judgements. Segment accuracy (i.e. PCC-R) is most commonly used to assess speech production in the clinical field. However, this relational measure is limited for assessing functional speech intelligibility that can assist with intervention planning and evaluating speech intervention outcomes for children with hearing loss.

The average intelligibility score for children with hearing loss in this cohort was 42%. This intelligibility score is much lower in comparison to hearing children who achieve over 90% intelligibility by age 4 years (e.g. Flipsen, 2008). Two issues can be associated with the low intelligibility scores in this study. First, conversational speech production can be difficult to understand due to lack of context information for listeners. The listeners had no information about conversation topics, utterances of conversational partners, and visual cues. When a child plays with toys without any attention to his or her speech, the child may not attend to production efforts on his or her words. Another reason for the extremely reduced intelligibility score can be low sound quality of recordings for stimuli. Even though speech outcomes overlapped with other speaker's production or noise caused by toys excluded for the experimental stimuli, there was still a low sound quality of utterances with background noise.

The improvement in speech intelligibility of children measured longitudinally provides indirect evidence of the efficacy of sensory aids (i.e., whether the children receive full benefit from their sensory aids). Five children demonstrated improvement in speech intelligibility judged by listeners at older chronological ages. However, three children did not improve their speech intelligibility over time in the longitudinal data set. StOd demonstrated a decrease in measures of intelligibility at an older chronological age. His

PWP score also declined over time and he showed very little increase in WCM-Ratio. His lack of improvement on whole word productions was associated with a decrease his intelligibility scores at the second measurement period. MaTa's intelligibility at her first sample is 38% whereas her intelligibility at the second sample is 34%. She also demonstrated little improvement in PWP (increasing from 70% to 72%). Her small increase in PWP scores may also be related to lack of improvement in listener intelligibility outcomes over time. These two children may have produced more complex word forms at older chronological age based on somewhat higher WCM-Ratio scores at their older ages (StOd: 81% → 82%, MaTa: 58% → 63%). A third child (TyMc) showed a decrease in his speech intelligibility scores over time also. Again, his first speech sample was collected with his mother at his home in an exceptional way. His second speech sample was collected by graduate clinicians in the UTSHC as usual. This type of result reinforces the implication that a child demonstrates more intelligible speech production when he/she communicates with a familiar communication partner in a familiar context.

Intelligible speech production is a predominant goal of speech intervention for children with hearing loss. The current study supports the proposal that whole word measures could be indirect intelligibility tests for children with hearing loss. When clinicians have difficulty recruiting unfamiliar listeners to evaluate children's speech intelligibility, a whole word measure could be a tool to evaluate and plan treatment for speech production patterns for children with hearing loss.

5.4 Overview and Implications for Future Research

This study investigated speech and language characteristics of spontaneous speech production for children with hearing loss by using multiple speech and language measures. The findings from all four speech measures and MLU & NDW reveal children with hearing loss demonstrate chronological age-related development in speech and language production. However, most of these children did not fully achieve age-appropriate scores on speech measures or MLU. Based on additional analyses for each speech measure, these children with hearing loss demonstrate a delay in acquiring later developing consonants and more complex word forms. The relationship between MLU and the WCM-Ratio for final consonants are significant within this cohort of children with hearing loss.

These results have important clinical implications for providing proper assessment and intervention for children with hearing loss. The findings may serve as a resource for clinicians in developing preliminary expectations of performance in speech and language for the children with hearing loss. Overall, clinicians may expect a speech delay for a child with CI or HA. In addition, it is necessary for clinicians to know that particular phonological skills (e.g. final consonant production) can be a potential factor to examine within children's grammatical abilities. The data presented here may also provide a resource for clinicians to select and use speech measures in an appropriate way. In order to use a measure of PCC-R- PWP intersect to make a more decisive diagnosis of speech disorder versus speech delay, both spontaneous speech sampling and formal word-based speech testing would be optimal in assessment protocols with these children. This approach provides two types of environments for children with hearing loss to produce more complex words. The well-organized formal testing allows the clinician to elicit target words without missing form of words and spontaneous speech sample can facilitate producing more natural functional speech outcomes. This approach supports a full analysis of interactions between consonant correctness and word complexity.

Environmental factors are also important to clinical planning. Communication modality is a significant factor related to speech measures, especially PCC-R and PWP measures. The Talking Only group showed higher values on PCC-R and PWP than the Total Communication group. Using both ASL and spoken English may be related to less effective speech production development. There was no significant difference between the HA and CI groups using any of these speech measures. Despite lack of intact hearing levels for the children with hearing loss in this cohort, children with CI show comparable speech skills to children with HA despite much more severe hearing loss levels for children with CI. This finding may extend the evidence of CI benefits (Blamey et al., 2001).

Vocabulary input factors including phonotactic probability and neighborhood density related to speech production only contained CVC word forms due to lack of availability of the online comprehensive corpus (Storkel & Hoover, 2010). Despite the limitation on data availability for analysis, these preliminary results revealed that the children with hearing loss produced CVC words containing frequently occurring sounds or more neighbors more accurately than words with infrequent sounds or fewer neighbors. These results indicate facilitative effects of high phonotactic probability and neighborhood density on consonant accuracy in CVC words within conversational samples. These findings allow clinicians to make precise intervention plans to select target words for articulation therapy for children with hearing loss. The children would benefit by beginning with working on words with more neighbors or frequently occurring sounds within target sounds.

Intelligibility outcomes identified from adult listeners were highly associated with speech measures incorporating a whole word approach to analysis of patterns; PWP and WCM-Ratio. When these children with hearing loss produce the more complex words and later developing consonants that match with target words and sounds more fully, their speech production is judged as more intelligible. That relationship of intelligibility with word complexity is not revealed by PCC-R and PVC-R measures of speech patterns. Measures of intelligibility have been used to help determine the need for intervention and provide an index of the severity of a speech disorder. However, assessing intelligibility via word identification from multiple listeners is difficult to conduct in clinical practice due to time limitations. PWP and WCM-Ratio would be indirect measures to assess intelligibility to assist with intervention planning and to evaluate speech intervention outcomes for children with hearing loss.

Speech outcomes of spontaneous speech production for children with hearing loss were comprehensively assessed using multiple speech and language measures. For all speech measures, chronologically older children achieved higher scores than younger children even though most children did not achieve age-appropriate expectations for speech production skills. This result indicates the importance of early identification, amplifications, and educational interventions. Early amplification or implantation can assist children with hearing loss to achieve speech and language milestones without significant disturbance in speech production. They can follow the developmental trajectory of speech production via early adaption of hearing aids or cochlear implants.

On the contrary, despite early adaption of hearing aids or cochlear implantation, the majority of the children with hearing loss in this cohort demonstrated delayed speech and language development based on multiple speech measures and MLU. This delayed speech and language development highlights the importance of early input from the ambient language (even earlier than 1 year of age) to achieve age appropriate speech and language capacities.

Speech production and speech perception are revealed in a dynamic system across the early period of development. Growth and maturity of the biological systems underlying auditory perception, speech production, and neural aspects of development impact children's cognitive capacities. Growth in these capacities is also founded in socially mediated input from children's environment based on functional daily communication interactions (Davis & Bedore, 2013). The capacity of the auditory system to receive and transmit information to the central nervous system is a critical dimension of development affecting maturing speech production capacities in young children (Sininger, Grimes, & Christensen, 2010). Even though the early use of hearing devices may be instantiated in children with hearing loss, they still have lack of intact auditory capacities for hearing vocalizations from the environment. They are likely to have an insufficiently diverse set of experiences with both perception and production aspects of development, leading to difficulties with fully developing speech and language capacities relative to chronological age expectations.

One major goal of this study was to explore potential reasons for diverse outcomes of speech development in children identified early with hearing loss and fit with instrumentation. The results of comprehensive analyses would help understand individual diversity in speech outcomes for children with hearing loss. The current study revealed that most of this cohort of children with hearing loss demonstrate common patterns in their speech production. They followed the typical developmental trajectory by showing lower accuracy for later developing consonants (e.g. fricatives) compared to early developing consonants (e.g. labial sounds). In addition, higher scores on speech measures were observed in chronologically older children than the younger children in the group. Most of children with hearing loss increased their speech scores when they were chronologically older. No individual differences were observed in their PCC-R scores. However, in the longitudinal data set, a few children with hearing loss demonstrated different characteristics in their speech developing patterns based on PVC-R, PWP, WCM-Ratio, MLU & intelligibility scores.

The potential factors contributing to individual differences in their speech production can be chronological age at adaption or implantation, mean length of utterances, or familiarity of speech sampling context. One child, (MaTa), who showed a moderate delay in PCC-R decreased her intelligibility at the second sample (older). Another child (StOd) also showed decreased intelligibility at an older chronological age. Both children demonstrated lack of improvement on whole word productions. They are the only two children who did not make meaningful progress in PWP. StOd, instead, showed a decrease in PWP scores and very little improvement in WCM-Ratio in his second speech sample. Their lack of increase in whole word properties may be related to lack of improvement in listener intelligibility outcomes over time. In addition, they adapted CIs after 3 years of age despite their profound hearing loss. The other two children who did not show improvement in their PVC scores produced much longer utterances in their later (older) data samples compared to the other children in the longitudinal data set. Vowel production could be reduced or slurred when they produce longer sentences in a conversational environment. One of the two children (TyMc) showed a decrease in his speech intelligibility over time also. His two spontaneous speech samples were collected in different contexts. The more familiar environment with his mother at the first speech sample likely affected more intelligible speech production. However, these potential factors should be descriptively interpreted within this cohort due to lack of statistical analysis. Further studies are required to fully understand individual diversity in speech development with a larger number of children with hearing loss in the various environments including a broader range of ages at adaption of hearing devices. To monitor individual differences, multiple data collections in a long-term period is also essential for assessment of potential factors.

The current data was collected via spontaneous speech sampling for these children with hearing loss. Analysis of spontaneous speech is the best procedure for evaluating functional speech capacities in daily communication. However, PCC or PCC-R is the only speech measure for spontaneous speech sample analysis, that has been mostly used to evaluate child's speech characteristics in addition to standardized single word tests. A lack of normative data and difficulties in relevant interpretation strategies for the other measures would limit general use in clinical fields. The current research findings with various speech measures would provide a guideline on how to use these measures and interpret the resulting scores. In the future, it would be more valuable to collect speech samples for age-matched hearing peers to directly compare children with hearing loss to hearing children.

A lack of standardized procedures for eliciting speech samples makes it difficult to collect representative speech samples. One child's speech data was collected by child' mother at home whereas the other speech sample for the same child was collected via clinicians in the UTSHC as usual. All of children except that child demonstrated positive associations between chronological age growth and speech outcomes. This type of result indicates the potential importance of familiar versus unfamiliar communication partners for maximizing spontaneous speech outcomes. Future work should require a standardized procedure for consistently eliciting speech samples in a valid and reliable way.

Chronological age, communication modality and hearing device were environmental factors that the current study reported to investigate influence of these factors on speech and language development in children with hearing loss. However, the current results on environmental factors relative to speech outcomes need to be very carefully interpreted within this data set. The current data set only includes thirty-two speech samples of 21 children with hearing loss. There is no complete information about their hearing levels and details about hearing devices. Future work with more participants including full hearing information must continue to clarify effects of communication modality and hearing device for children with hearing loss to receive optimal care.

Results of vocabulary input factors indicate facilitative effects of higher phonotactic probability and greater neighborhood density on consonant accuracy. A majority of experimental studies have reported the inhibitory effects of neighborhood density on word recognition and word learning. The current study of spontaneous CVC word forms inconsistently demonstrated facilitative effects of neighborhood density on speech production. This inconsistent result can be associated to the different study approach. However, the current data included only CVC word forms. Word frequency which contributes to categorizing words into lexically difficult and easy categories (Munson & Solomon, 2004) was not considered to analyze the current data set. Future research with broad range of word forms and controlling other input factors (e.g. word frequency) are required to examine whether words with dense neighborhoods facilitate acquiring speech production accuracy.

The two statistical analyses suggest that the listener intelligibility score is associated with higher scores on PWP and WCM-Ratio. However, in the first analysis, small sample size (only 32 observations) and repeated observations in a longitudinal data set reduced the power of statistical results. In the second analysis, using word-by-word comparisons of speech measures and intelligibility should be carefully interpreted because adaptation of the original procedures of the speech measures include both speaker and listeners variations. Both analyses did not include all four speech measures in a model due to multicollinearity. Thus, there is a limit to statistically comparison of these four measures to predict intelligibility.

Conclusions

Comprehensive analysis of speech outcomes with contemporary speech and language measures is critical to understand speech characteristics of children with hearing loss and the relationship of auditory input system and production capacities. These results reinforce early input from children's environment based on functional daily communication interactions by demonstrating early identification benefits and delayed speech development in these children fit with HA or CI instrumentation. The analyses reported in the current study can provide a resource for clinicians to comprehensively describe speech and language patterns and their relationships to input values and listener intelligibility judgements for children with hearing loss. In addition, these findings can support clinicians in establishing valid assessment protocols and intervention plans. However, these results should be carefully interpreted due to the relatively small number of participants and lack of detail about their hearing levels and other environmental factors (e.g., hearing age). Future work should include more children with hearing loss from broad environmental backgrounds to evaluate environmental factors on speech outcomes more fully. It would be valuable to gather speech data for age matched hearing peers for all these measures to directly compare between these two groups. Documenting comprehensive outcomes via multiple measures can support understanding of developmental profiles related to the interface of perception and production capacities in young children with hearing loss. In addition, they provide some precision regarding measures that can reveal the sources of reduced intelligibility for planning of clinical assessment and intervention in young children who use CI or HA.

Appendix

Appendix. Excluded mono-syllable function words, be-verb, and interjection.

Type	Word	Type	Word
article	a	preposition	at
article	an	preposition	in
article	the	preposition	with
auxVerb	be	preposition	to
auxVerb	been	conjunction	and
auxVerb	is	conjunction	but
auxVerb	am	adverb	so
auxVerb	are	interjection	ah
auxVerb	being	interjection	um
auxVerb	can	interjection	uh
auxVerb	can't	interjection	oh
auxVerb	could	interjection	uh oh
auxVerb	couldn't	preposition	as
auxVerb	did	preposition	for
auxVerb	didn't	preposition	of
auxVerb	do	preposition	on
auxVerb	does	interjection	ooh
auxVerb	doesn't	conjunction	or
auxVerb	doing	interjection	ouch
auxVerb	don't	interjection	ow
auxVerb	get	interjection	whoa
auxVerb	getting	interjection	wow
auxVerb	got	interjection	yea
auxVerb	will	interjection	yeah
auxVerb	would	--	no
		--	yes

References

Allen, C., Nikolopoulos, T. P., Dyar, D., & O'donoghue, G. M. (2001). Reliability of a rating scale for measuring speech intelligibility after pediatric cochlear implantation. *Otology & Neurotology*, *22*(5), 631-633.

American Speech-Language-Hearing Association (ASHA). (2004). *Cochlear implants* [Technical Report]. Available from www.asha.org/policy.

Anderson, C., & Cohen, W. (2012). Measuring word complexity in speech screening: Single-word sampling to identify phonological delay/disorder in preschool children. *International Journal of Language & Communication Disorders / Royal College of Speech & Language Therapists, 47(5),* 534.

Bergeson, T. R., Pisoni, D. B., & Davis, R. A. (2003). A longitudinal study of audiovisual speech perception by children with hearing loss who have cochlear implants. *The Volta Review*, *103*(4), 347.

Bewick, V., Cheek, L., & Ball, J. (2003). Statistics review 7: Correlation and regression. *Critical care*, *7*(6), 451.

Blamey, P. J., Sarant, J. Z., Paatsch, L. E., Barry, J. G., Bow, C. P., Wales, R. J., ... & Tooher, R. (2001). Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. *Journal of Speech, Language, and Hearing Research*, *44*(2), 264-285.

Bow, C. P., Blamey, P. J., Paatsch, L. E., & Sarant, J. Z. (2004). The effects of phonological and morphological training on speech perception scores and grammatical judgments in deaf and hard-of-hearing children. *Journal of Deaf Studies and Deaf Education*, *9*(3), 305-314.

Bozdogan, H. (2000). Akaike's information criterion and recent developments in information complexity. *Journal of mathematical psychology*, *44*(1), 62-91.

Brown, R. (1973). *A First language: the early stages.* Cambridge, MA: Harvard University Press.

Brownell, R. (2000). *Expressive one-word picture vocabulary test-3*. Navato, CA: Academic Therapy Publications.

Burrows, L., & Goldstein, B. A. (2010). Whole word measures in bilingual children with speech sound disorders. *Clinical linguistics & phonetics, 24*, 357-368.

Campbell, T. F., Dollaghan, C., Janosky, J. E., & Adelson, P. D. (2007). A performance curve for assessing change in Percentage of Consonants Correct–Revised (PCC-R). *Journal of Speech, Language, and Hearing Research*, *50*(4), 1110-1119.

Castellanos, I., Kronenberger, W. G., Beer, J., Henning, S. C., Colson, B. G., & Pisoni, D. B. (2014). Preschool speech intelligibility and vocabulary skills predict long-term speech and language outcomes following cochlear implantation in early childhood. *Cochlear implants international*, *15*(4), 200-210.

Chin, S. B., Tsai, P. L., & Gao, S. (2003). Connected speech intelligibility of children with cochlear implants and children with normal hearing. *American Journal of Speech-Language Pathology*, *12*(4), 440-451.

Ching, T. Y. (2015). Is early intervention effective in improving spoken language outcomes of children with congenital hearing loss? *American journal of audiology*, *24*(3), 345-348.

Ching, T. Y., Dillon, H., Button, L., Seeto, M., Van Buynder, P., Marnane, V., ... & Leigh, G. (2017). Age at intervention for permanent hearing loss and 5-year language outcomes. *Pediatrics*, *140*(3), e20164274.

Coalson, G. A., Byrd, C. T., & Davis, B. L. (2012). The influence of phonetic complexity on stuttered speech. *Clinical linguistics & phonetics*, *26*(7), 646-659.

Cummings, A., & Larson, A. (2015, June). Identification of subclasses of children with speech sound disorders using the PCC, PWP Intersect. In *Conference Program* (p. 21).

Connor, C. M., Craig, H. K., Raudenbush, S. W., Heavner, K., & Zwolan, T. A. (2006). The age at which young deaf children receive cochlear implants and their vocabulary and speech-production growth: is there an added value for early implantation? *Ear and hearing*, *27*(6), 628-644.

Davis, B. L., MacNeilage, P. F., & Matyear, C. L. (2002). Acquisition of serial complexity in speech production: A comparison of phonetic and phonological approaches to first word production. *Phonetica*, *59*(2-3), 75-107.

Davis, B. L., & Bedore, L. M. (2013). *An emergence approach to speech acquisition: Doing and knowing*. Psychology Press.

Davis, B., van der Feest, S., & Yi, H. (2017). Speech sound characteristics of early words: influence of phonological factors across vocabulary development. *Journal of child language*, 1-30 Nusbaum.

DeCasper, A.J., & Fifer, W.P. (1980). Of human bonding: Newborns prefer their mother's voices. *Science,* 208, 1174-6.

Dunn, L. M., & Dunn, L. M. (1997). *PPVT-III: Peabody picture vocabulary test*. American Guidance Service.

Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4: Peabody picture vocabulary test*. Pearson Assessments.

Dunn, C. C., Walker, E. A., Oleson, J., Kenworthy, M., Van Voorst, T., Tomblin, J. B., ... & Gantz, B. J. (2014). Longitudinal speech perception and language performance in pediatric cochlear implant users: the effect of age at implantation. *Ear and hearing*, *35*(2), 148.

Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of speech, language, and hearing research*, *47*(2), 421-436.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science, 171(3968),* 303-306.

Eriks-Brophy, A., Gibson, S., & Tucker, S. (2013). Articulatory error patterns and phonological process use of preschool children with and without hearing loss. *The Volta Review, 113*(2), 87-125.

Ertmer, D. J. (1999). *The First Words Speech Test*. Unpublished manuscript, Department of Speech, Language, and Hearing Sciences, Purdue University, West Lafayette, IN.

Ertmer, D. J. (2007). Speech Intelligibility in Young Cochlear Implant Recipients: Gains During Year Three. *Volta Review*, *107*(2).

Ertmer, D. J. (2010). Relationships between speech intelligibility and word articulation scores in children with hearing loss. *Journal of Speech, Language, and Hearing Research*, *53*(5), 1075-1086.

Ertmer, D. J. (2011). Assessing speech intelligibility in children with hearing loss: Toward revitalizing a valuable clinical tool. *Language, Speech, and Hearing Services in Schools*, *42*(1), 52-58.

Ertmer, D. J., & Goffman, L. A. (2011). Speech production accuracy and variability in young cochlear implant recipients: Comparisons with typically developing age-peers. *Journal of Speech, Language, and Hearing Research*, *54*(1), 177-189.

Ertmer, D. J., Strong, L. M., & Sadagopan, N. (2003). Beginning to Communicate After Cochlear ImplantationOral Language Development in a Young Child. *Journal of Speech, Language, and Hearing Research*, *46*(2), 328-340.

Faes, J., Gillis, J., & Gillis, S. (2015). Syntagmatic and paradigmatic development of cochlear implanted children in comparison with normally hearing peers up to age 7. *International journal of pediatric otorhinolaryngology*, *79*(9), 1533-1540.

Fenson, L., Marchman, V. A., Thal, D., Dale, P. S., Bates, E. & Reznick, J. S. (2006). The MacArthur-Bates Communicative Development Inventories user's guide and technical manual, 2nd ed. Baltimore, MD: Brookes.

Ferguson, C. A., & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language*, 419-439.

Fitzpatrick, E. M., Olds, J., Gaboury, I., McCrae, R., Schramm, D., & Durieux-Smith, A. (2012). Comparison of outcomes in children with hearing aids and cochlear implants. *Cochlear Implants International*, *13*(1), 5-15.

Flipsen, P. (2008). Intelligibility of spontaneous conversational speech produced by children with cochlear implants: a review. *International Journal of Pediatric Otorhinolaryngology*, *72*(5), 559-564.

Flipsen, P., & Parker, R. G. (2008). Phonological patterns in the conversational speech of children with cochlear implants. *Journal of Communication Disorders*, *41*(4), 337-357.

Flipsen, P., & Kangas, K. (2014). Mean length of utterance (MLU) in children with cochlear implants. *The Volta Review*, *114*(2), 135-155.

Gahl, S., Yao, Y., & Johnson, K. (2012). Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of memory and language*, *66*(4), 789-806.

Geers, A. E. (2004). Speech, language, and reading skills after early cochlear implantation. *Archives of Otolaryngology–Head & Neck Surgery*, *130*(5), 634-638.

Geers, A. E., Mitchell, C. M., Warner-Czyz, A., Wang, N. Y., Eisenberg, L. S., & CDaCI Investigative Team. (2017). Early Sign Language Exposure and Cochlear Implantation Benefits. *Pediatrics*, e20163489.

Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation–Second Edition*. Circle Pines, MN: American Guidance Service.

Gordon-Brannan, M., & Hodson, B. W. (2000). Intelligibility/severity measurements of prekindergarten children's speech. *American Journal of Speech-Language Pathology*, *9*(2), 141-150.

Haley, K. L., & Jacks, A. (2014). Single-word intelligibility testing in aphasia: Alternate forms reliability, phonetic complexity and word frequency. *Aphasiology*, *28*(3), 320-337.

Han, M. K., Storkel, H. L., Lee, J., & Yoshinaga-Itano, C. (2015). The influence of word characteristics on the vocabulary of children with cochlear implants. *Journal of deaf studies and deaf education*, env006.

Hazrati, O., & Loizou, P. C. (2012). The combined effects of reverberation and noise on speech intelligibility by cochlear implant listeners. *International journal of audiology*, *51*(6), 437-443.

Heisler, L., & Goffman, L. (2016). The Influence of Phonotactic Probability and Neighborhood Density on Children's Production of Newly Learned Words. *Language Learning and Development*, *12*(3), 338-356.

Hodge, M. and Whitehill, T. (2010) *Intelligibility Impairments*, in The Handbook of Language and Speech Disorders (eds J. S. Damico, N. Müller and M. J. Ball), Wiley-Blackwell, Oxford, UK.

Ingram, D. (1976). Phonological Analysis of a Child. *Glossa*.

Ingram, D. (2002). The measurement of whole-word productions. *Journal of Child Language, 29*, 713–733.

Ingram, D. (2012). A comparison of two measures of correctness: PCC & PWP. Presented at the Annual meeting of the American Speech-Language-Hearing Association, San Diego, CA.

Ingram, D., & Ingram, K. D. (2001). A whole-word approach to phonological analysis and intervention. *Language, speech, and hearing services in schools*, *32*(4), 271-283.

Ireton, H., & Thwing, E. (1974). *Minnesota child development inventory*. Minneapolis: Behavior Science Systems.

Jusczyk, P. W., & Luce, P. A. (2002). Speech perception and spoken word recognition: Past and present. *Ear and hearing*, *23*(1), 2-40.

Khan, L., & Lewis, N. (2002). *Khan-Lewis Phonological Analysis 2 (KLPA-2*). Circle Pines, MN; American Guidance Service.

Kirk, K., Miyamoto, R., Ying, E., Perdew, A., & Zuganelis, H. (2002). Cochlear implantation in young children: effects of age at implantation and communication mode. *Volta Review*, *102*(4), 127-144.

Kolson (1960). *The vocabulary of kindergarten children*. Unpublished doctoral dissertation, University of Pittsburgh, Pittsburgh.

Koehlinger, K. M., Van Horne, A. J. O., & Moeller, M. P. (2013). Grammatical outcomes of 3-and 6-year-old children who are hard of hearing. *Journal of Speech, Language, and Hearing Research*, *56*(5), 1701-1714.

Kral, A., Kronenberger, W. G., Pisoni, D. B., & O'Donoghue, G. M. (2016). Neurocognitive factors in sensory restoration of early deafness: a connectome model. *The Lancet Neurology*, *15*(6), 610-621.

Kuhl, P. K., & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *The Journal of the Acoustical Society of America 100 (4),* 2425-2438.

Lecanuet, J. P., & Schaal, B. (1996). Fetal sensory competencies. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, *68*, 1-23.

Longman (1993). *Longman dictionary of American English.* White Plains, NY: Longman.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. Journal of Speech, Language, and Hearing Research, 50(4), 940-967.

Miles, J., & Shevlin, M. (2001). *Applying regression and correlation: A guide for students and researchers*. Sage.

Moe, A.J., Hopkins, C.J. & Rush, R.T. (1982). *The vocabulary of first grade children*.

Montag, J. L., AuBuchon, A. M., Pisoni, D. B., & Kronenberger, W. G. (2014). Speech Intelligibility in Deaf Children After Long-Term Cochlear Implant Use. *J Speech Lang Hear Res*, 57(6), 2332-2343.

Moeller, M. P., Hoover, B., Putman, C., Arbataitis, K., Bohnenkamp, G., Peterson, B., ... & Stelmachowicz, P. (2007). Vocalizations of infants with hearing loss compared with infants with normal hearing: Part I–Phonetic development. *Ear and hearing*, *28*(5), 605-627.

Moeller, M. P. (2000). Early intervention and language development in children who are deaf and hard of hearing. *Pediatrics*, *106*(3), e43-e43.

Moeller, M. P., McCleary, E., Putman, C., Tyler-Krings, A., Hoover, B., & Stelmachowicz, P. (2010). Longitudinal development of phonology and morphology in children with late-identified mild-moderate sensorineural hearing loss. *Ear and hearing*, *31*(5), 625.

Moeller, M. P., & Tomblin, J. B. (2015). An introduction to the Outcomes of Children with Hearing Loss study. *Ear and hearing*, *36*, 4S-13S.

Munson, B., & Solomon, N. P. (2004). The effect of phonological neighborhood density on vowel articulation. *Journal of speech, language, and hearing research*, *47*(5), 1048-1058.

Moon, C. M., & Fifer, W. P. (2000). Evidence of transnatal auditory learning. *Journal of Perinatology*, *20*(8).

Newman, R. S., & German, D. J. (2002). Effects of lexical factors on lexical access among typical language-learning children and children with word-finding difficulties. *Language and Speech*, *45*(3), 285-317.

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*(4), 1048-1062.

Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words. *Research on speech perception progress report*, *10*(10), 357-376.

Osberger, M. J., Robbins, A. M., Todd, S. L., & Riley, A. I. (1994). Speech intelligibility of children with cochlear implants. Volta Review, 96, 169–180.

Peng, S. C., Spencer, L. J., & Tomblin, J. B. (2004). Speech intelligibility of pediatric cochlear implant recipients with 7 years of device experience. *Journal of Speech, Language, and Hearing Research*, *47*(6), 1227-1236.

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of neuroscience methods*, *162*(1), 8-13.

Pisoni, D. B. (2000). Cognitive factors and cochlear implants: Some thoughts on perception, learning, and memory in speech perception. *Ear and Hearing, 21(1),* 70-78.

Pollock, K. E. (2002). Identification of vowel errors: Methodological issues and preliminary data from the Memphis Vowel Project. *Vowel disorders*, 83-113.

Rice, M. L., Redmond, S. M., & Hoffman, L. (2006). Mean Length of Utterance in Children with Specific Language Impairment and in Younger Control Children Shows Concurrent Validity and Stable and Parallel Growth Trajectories Rice et al.: MLU Holds Up Over Time. *Journal of Speech, Language, and Hearing Research*, *49*(4), 793-808.

Roberts, J. E., Burchinal, M., & Footo, M. M. (1990). Phonological process decline from 212 to 8 years. *Journal of Communication Disorders*, *23*(3), 205-217.

Rose, Y., MacWhinney, B., Byrne, R., Hedlund, G., Maddocks, K., O'Brien, P., & Wareham, T. (2006, April). Introducing Phon: A software solution for the study of phonological acquisition. In *Proceedings of the... Annual Boston University Conference on Language Development. Boston University Conference on Language Development* (Vol. 2006, p. 489). NIH Public Access.

Sander, E. K. (1972). When are speech sounds learned?. *Journal of speech and hearing disorders*, *37*(1), 55-63.

Semel, E., Wiig, E. H., & Secord, W. A. Clinical Evaluation of Language Fundamentals . 2003. *The Psychological Corporation*.

Sininger, Y. S., Grimes, A., & Christensen, E. (2010). Auditory development in early amplified children: Factors influencing auditory-based communication outcomes in children with hearing loss. *Ear and hearing*, *31*(2), 166.

Shriberg, L. D., Austin, D., Lewis, B. A., McSweeny, J. L.,& Wilson, D. L. (1997). The Percentage of Consonants Correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language, and Hearing Research, 40,* 708–722.

Schauwers, K., Gillis, S., & Govaerts, P. J. (2008). The characteristics of prelexical babbling after cochlear implantation between 5 and 20 months of age. *Ear and hearing*, *29*(4), 627-637.

Schwartz, R. G., Steinman, S., Ying, E., Mystal, E. Y., & Houston, D. M. (2013). Language processing in children with cochlear implants: A preliminary report on lexical access for production and comprehension. *Clinical linguistics & phonetics*, *27*(4), 264-277.

Sosa, A. V., & Stoel-Gammon, C. (2012). Lexical and phonological effects in early word production. *Journal of Speech, Language, and Hearing Research*, *55*(2), 596-608.

Spencer, P. E. (2004). Individual differences in language performance after cochlear implantation at one to three years of age: Child, family, and linguistic factors. *Journal of Deaf Studies and Deaf Education*, *9*(4), 395-412.

Stoel-Gammon, C., & Pollock, K. (2008). Vowel development and disorders. *The handbook of clinical linguistics*, 525-548.

Stoel-Gammon, C. (2010). The Word Complexity Measure: Description and application to developmental phonology and disorders. *Clinical linguistics & phonetics*, *24*(4-5), 271-282.

Storkel, H. L. (2001). Learning New WordsPhonotactic Probability in Language Development. *Journal of Speech, Language, and Hearing Research*, *44*(6), 1321-1337.

Storkel, H. L. (2003). Learning new words II: Phonotactic probability in verb learning. *Journal of Speech, Language, and Hearing Research*, *46*(6), 1312-1323.

Storkel, H. L. (2004). Do children acquire dense neighborhoods? An investigation of similarity neighborhoods in lexical acquisition. *Applied Psycholinguistics*, *25*(2), 201-221.

Storkel, H. L., & Hoover, J. R. (2010). An online calculator to compute phonotactic probability and neighborhood density on the basis of child corpora of spoken American English. *Behavior Research Methods*, *42*(2), 497-506.

Storkel, H. L., & Lee, S. Y. (2011). The independent effects of phonotactic probability and neighbourhood density on lexical acquisition by preschool children. *Language and Cognitive Processes*, *26*(2), 191-211.

Storkel, H. L. (2013). A corpus of consonant–vowel–consonant real words and nonwords: Comparison of phonotactic probability, neighborhood density, and consonant age of acquisition. *Behavior Research Methods*, *45*(4), 1159-1167.

Szagun, G. (2000). The acquisition of grammatical and lexical structures in children with cochlear implants: A developmental psycholinguistic approach. *Audiology and Neurotology*, *5*(1), 39-47.

Szagun, G. (2001). Language acquisition in young German-speaking children with cochlear implants: Individual differences and implications for conceptions of a ‘sensitive phase’. *Audiology and Neurotology*, *6*(5), 288-297.

Szagun, G. (1998). Language acquisition in children with cochlear implants; First results of a development psycholinguistic study. *Sprache Stimme Gehör*, *22*, 133-8.

Templin, M. C. (1957). Certain language skills in children; their development and interrelationships.

Texas Speech Language Hearing Association Articulation Guidelines (2010): CLD Companion

Tobey, E. A., Geers, A. E., Brenner, C., Altuna, D., & Gabbert, G. (2003). Factors associated with development of speech production skills in children implanted by age five. *Ear and hearing*, *24*(1), 36S-45S.

Tomblin, J. B., Oleson, J. J., Ambrose, S. E., Walker, E., & Moeller, M. P. (2014). The influence of hearing aids on the speech and language development of children with hearing loss. *JAMA Otolaryngology–Head & Neck Surgery*, *140*(5), 403-409.

Vihman, M., & Croft, W. (2007). Phonological development: Toward a “radical” templatic phonology.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, *40*(3), 374-408.

Vitevitch, M. S., & Luce, P. A. (2016). Phonological neighborhood effects in spoken word perception and production. *Annual Review of Linguistics*, *2*, 75-94.

Von Hapsburg, D., & Davis, B. L. (2006). Auditory sensitivity and the prelinguistic vocalizations of early-amplified infants. *Journal of Speech, Language, and Hearing Research*, *49*(4), 809-822.

Waring, R., Fisher, J., Atkin, N., Wilson, L., & Hewat, S. (2001). The articulation survey: Putting numbers to it. In *Proceedings of the 2001 Speech Pathology Australia national conference: Evidence and innovation. Melbourne: Speech Pathology Australia* (pp. 145-151).

Warner-Czyz, A. D., & Davis, B. L. (2008). The emergence of segmental accuracy in young cochlear implant recipients. *Cochlear Implants International*, *9*(3), 143-166.

Watson, M. M., & Terrell, P. (2012). Longitudinal changes in phonological whole-word measures in 2-year-olds. *International journal of speech-language pathology*, *14*(4), 351-362.

Weiss, C. E. (1982). *Weiss intelligibility test*. CC Publications.

Wellman, B. L., Case, I. M., Mengert, I. G., & Bradbury, D. E. (1931). Speech sounds of young children. *University of Iowa Studies: Child Welfare*.

Werker, J. F., & Tees, R. C. (1983). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant behavior & Development, 7,* 49-63.

Yoshinaga-Itano, C., Baca, R. L., & Sedey, A. L. (2010). Describing the trajectory of language development in the presence of severe to profound hearing loss: A closer look at children with cochlear implants versus hearing aids. *Otology & neurotology: official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, *31*(8), 1268.

Yoshinaga-Itano, C., Sedey, A. L., Coulter, D. K., & Mehl, A. L. (1998). Language of early-and later-identified children with hearing loss. *Pediatrics, 102(5)*, 1161-1171.

Yoshinaga-Itano, C. (2013). Principles and guidelines for early intervention after confirmation that a child is deaf or hard of hearing. *Journal of Deaf Studies and Deaf Education*, ent043.

Zamuner, T. S. (2009). Phonotactic probabilities at the onset of language development: Speech production and word position. *Journal of Speech, Language, and Hearing Research*, *52*(1), 49-60.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (1992). *PLS-3: Preschool language scale-3*. Psychological Corporation.