

EARLY LINGUISTIC ENVIRONMENTS AND LANGUAGE DEVELOPMENT IN
CHILDREN WITH COCHLEAR IMPLANTS

By

Meisam Khalil Arjmandi

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Communicative Sciences and Disorders — Doctor of Philosophy

2019

ABSTRACT

EARLY LINGUISTIC ENVIRONMENTS AND LANGUAGE DEVELOPMENT IN CHILDREN WITH COCHLEAR IMPLANTS

By

Meisam Khalil Arjmandi

Prior research has documented tremendous variability in language outcomes of children with cochlear implants (CIs); despite more than a decade of research, a large portion of this variability remains unexplained. This study characterized the quantity and quality of early linguistic input in naturalistic environments of 14 early-implanted children with CIs to investigate variability across children as a possible source of variation that might help explain language outcome variability. In Chapter 2, daylong audio recordings from home environments of these children were analyzed to examine individual variability in language input they experienced in terms of lexical, morphosyntactic, and social-pragmatic properties. It was found that the quantity and quality of early language input varies substantially across children with CIs, where the degree of variability was comparable in magnitude to what has been reported between the most-advantaged and least-advantaged typically hearing children. In Chapter 3, estimates of the quantity and quality of language input were adjusted to consider environmental noise and reverberation to better represent the “useable” amount of input experienced by the children. It was found that children with CIs are differentially impacted by noise and reverberation in their naturalistic environments, such that some children are doubly disadvantaged in acquiring spoken language, both due to substantial variability in the amount and quality of linguistic input available to them, as well as due to their exposure and susceptibility to environmental noise and reverberation. In Chapter 4, an initial test was conducted to obtain preliminary results regarding how language input factors might shape development of language outcomes in children with CIs. The preliminary

results estimating the contribution of language input measures to language outcomes of the children with CIs suggested that measure of speech intelligibility tailored to children with CIs strongly predicted language outcomes. Overall, this study has provided evidence for substantial individual variability across children with cochlear implants in quantity and quality of their early language experience, which were mainly influenced by factors of child-directed speech and environmental noise and reverberation. This evidence-based knowledge can be used by parents and clinicians to effectively adjust early linguistic environments of children with CIs to maximize the advantage of using CIs.

ACKNOWLEDGEMENTS

I would like to thank everyone who has supported me in developing and execution of my dissertation work. The completion of this dissertation could not have been possible without the participation of families and their children and contribution of so many researchers and scholars who have been involved in this dissertation project. I would especially like to thank my family, my advisor, committee members, members of MSU Speech Perception and Production Lab, and the families and children who participated in this study.

I would like to express my sincere gratitude to my advisor, Dr. Laura C. Dilley, for the continuous support of my Ph.D. study and the related research projects. Her active and strategic guidance, encouragement, and immense knowledge have been exemplary and have tremendously helped me during my Ph.D. program and completion of my dissertation work.

I would also like to thank my committee members, Dr. Dimitar Deliyski, Dr. Mario Svirsky, Dr. Courtney Venker, and Dr. Joyce Chai for their great contribution in developing my dissertation proposal and their insightful comments and encouragement during execution of my dissertation project and writing the dissertation paper.

This work was supported by the National Institutes of Health - National Institute on Deafness and Other Communication Disorders - Grant No. R01DC008581. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This dissertation work was also partially supported with funding from the Graduate School of Michigan State University.

TABLE OF CONTENTS

LIST OF TABLES	viii
LIST OF FIGURES	xi
KEY TO ABBREVIATIONS.....	xvii
CHAPTER 1: Introduction	1
1.1. Individual variability in early linguistic environments may explain variability in language attainment	1
1.2. Quantity of early language experience	2
1.3. Quality of early language experience	3
1.4. The present study.....	5
CHAPTER 2: Variability in quantity and quality of early linguistic experience in children with cochlear implants	7
2.1. Introduction	7
2.2. Methods	22
2.2.1. Participants.....	22
2.2.2. Acquisition and selection of audio recordings.....	26
2.2.3. Sampling from day-long audio recordings	28
2.2.4. Analysis of audio by human analysts.....	30
2.2.5. Coding training and procedure.....	36
2.2.6. Inter-rater reliability analysis.....	37
2.2.7. Analysis of audio samples coded by human analysts	39
2.3. Results	45
2.3.1. Inter-rater reliability	45
2.3.2. Variability across children in the number of words experienced	46
2.3.3. Variability across children in lexical richness of their linguistic environment	52
2.3.4. Variability across children in morphosyntactic complexity of their linguistic environment.....	59
2.3.5. Variability across children in caregiver responsiveness	60
2.3.6. Variability across children in their experienced adults' speech rate.....	61
2.3.7. Variability across children based on a multidimensional space of language input.....	62
2.3.8. Variation in quality and quantity of linguistic input based on child-directed speech vs. adult-directed speech	63
2.4. Discussion	65

CHAPTER 3: Effects of environmental noise and reverberation on quality and quantity of early language and auditory environments of children with CIs	76
3.1. Introduction	76
3.2. Methods	86
3.2.1. Human coding for noise source and level of noise interference	87
3.2.2. Using human coding to provide updated estimates of number of “clean” words experienced by children with CIs	89
3.2.3. Using SRMR-CI to provide updated estimates of number of “clean” words experienced by children with CIs	91
3.2.4. Effect of noise on acoustic separability of child-directed and adult-directed speech	95
3.2.5. Variability across children in exposure to noise in their auditory environments	96
3.3. Results	99
3.3.1. Noise-adjusted estimates of numbers of words experienced by children with CIs	99
3.3.2. Effect of noise on acoustic separability between child-directed speech and adult-directed speech	120
3.3.3. Variability across children with cochlear implants in the profile of ambient noise in their linguistic environments	120
3.4. Discussion	125
CHAPTER 4: Toward understanding effects of quality and quantity of early language in linguistic environments in children with CIs on language outcomes	134
4.1. Introduction	134
4.2. Methods	140
4.2.1. Participants	140
4.2.2. Measures of language input (i.e., predictor variables)	140
4.2.3. Measures of language outcomes (i.e., predicted variables)	141
4.2.4. Statistical analysis	142
4.3. Results	143
4.3.1. To what extent are the total number of words and number of child-directed words (unadjusted for noise and reverberation) associated with language outcomes?	143
4.3.2. To what extent do the updated total number of words and number of child-directed words, adjusted for noise based on human coding, predict language outcomes?	145
4.3.3. To what extent does the updated total number of words and number of child-directed words (adjusted for noise and reverberation based on SRMR-CI metric) predict language outcomes?	146
4.3.4. Predictive value of SRMR-CI metric for language outcomes of children with CIs	148
4.4. Discussion	148

CHAPTER 5: Discussion and Conclusions	154
APPENDICES	164
Appendix A	165
Appendix B.....	173
Appendix C.....	175
Appendix D	176
Appendix E.....	178
Appendix F.....	179
Appendix G	180
BIBLIOGRAPHY	181

LIST OF TABLES

Table 2-1. Demographic information for 14 children with CIs who participated in this study. The information in this table includes children's ID, gender, age at CI activation, number of analyzed recordings, and children's chronological ages.....	23
Table 2-2. Demographic information for 14 children with CIs who participated in this study. The information in this table includes children's ID; hearing age at four time-intervals of 3m, 6m, 9m, and 12m; their type of hearing loss, and the laterality of their hearing loss.	24
Table 2-3. Demographic information for 14 children with cochlear implants participated in this study. The information in this table includes children's ID, their hearing age at four time-interval of 3m, 6m, 9m, and 12m, their type of hearing loss, and the laterality of their hearing loss.	25
Table 2-4. Demographic information for 14 children with CIs who participated in this study. The information in this table includes children's ID, the etiology of their hearing loss, their SES (given as maternal education, with HS = high school diploma and BA = bachelor's degree), and any additional diagnoses.....	25
Table 2-5. Hierarchical structure of information coding in the system across 8 <i>Praat</i> tiers. Columns labeled Level 1 to Level 4 present information analysts coded at each level and correspond to successive levels of detail used later to classify intelligible speech utterances. The details of each information on each tier and level are presented in "reference table of codes" in Table 1 in Appendix A. SNG = singing; NSV = non-speech vocalization. See Chapter 3 for more discussion of coding of degree of interference from noise and reverberation (cf. codes P = primary talker; L = low interference; M = medium interference; H = high interference).	35
Table 2-6. The results of inter-rater reliability analysis for various categories of interest in this study.....	46
Table 2-7. Measures of dispersion and central tendency for estimates of number of words per day for each child in his/her linguistic environment derived from total speech, child-directed speech, and overheard speech.	47
Table 2-8. Measures of dispersion and central tendency of type-token ratio derived across children with CIs. These measures are presented for total speech, child-directed speech, and overheard speech.....	53

Table 2-9. Measures of dispersion and central tendency for the AUTTC derived from language input of each child in his/her linguistic environments. These measures were calculated for total, speech, child-directed, and overheard speech. 59

Table 2-10. Measures of dispersion and central tendency for the MLU_m derived from language input of each child in his/her linguistic environments. These measures were calculated for total, speech, child-directed, and overheard speech. 60

Table 2-11. Measures of dispersion and central tendency for the average speech rate derived from language input of each child in his/her linguistic environments. These measures were calculated for total, speech, child-directed, and overheard speech in words per second. 62

Table 2-12. Z-score values of each child for each measure of quality and quantity of linguistic input. The measures from 1 to 11 include: (1) the total number of words per day, (2) the number of child-directed words per day, (3) TTR for total speech, (4) TTR for child-directed speech, (5) AUTTC for total speech, (6) AUTTC for child-directed speech, (7) MLU for total speech, (8) MLU for child-directed speech, (9) speech rate for total speech, (10) speech rate for child-directed speech, (11) caregivers' responsiveness. 63

Table 2-13. Results of statistical paired t-test to compare variation of measures of total number of words, TTR, AUTTC, MLU, and speech rate across child-directed speech and adult-directed speech conditions. 65

Table 3-1. Measures of central tendency and variability for the total number of word loss per day and its percent loss per day for two comparisons between "traditional and human-based method" and "traditional and SRMR-CI method" 103

Table 3-2. Table Results of statistical paired t-test to examine the significant effect of updating approaches of human-based and SRMR-CI-based on the amount of total words lost. 104

Table 3-3. Measures of dispersion and central tendency for the total number of words estimated through traditional, human-based, and SRMR-CI-based approaches..... 108

Table 3-4. Measures of central tendency and variability for the number of child-directed words lost and corresponding percentage word loss for two comparisons between traditional vs. human-based and traditional vs. SRMR-CI updating methods..... 113

Table 3-5. Results of statistical paired t-test to examine the significant effect of updating approaches of human-based and SRMR-CI-based on the amount of total words lost. 113

Table 3-6. Measures of dispersion and central tendency for number of child-directed words estimated through traditional, human-based, and SRMR-CI-based approaches. 115

Table 3-7. Measures of central tendency and variability for two measures of noise exposure: noise exposure index and ninety percent noise exposure index. 124

Table 4-1. Measures of quality and quantity of language input and language outcomes of the 13 children available for this study. Measures of language input were derived from analyzing the audio samples recorded from the natural linguistic environments of children. Measures of language outcomes were obtained by administering PLS5 tests..... 141

Table 4-2. The pairwise correlation coefficient, R-squared values, and *p*-values for the fitted simple linear regression models for examining the relationship between each language input variable derived from *total speech* and each of two variables of language outcomes..... 145

Table 4-3. The pairwise correlation coefficient, R-squared values, and *p*-values for the fitted simple linear regression models for examining the relationship between each language input variable derived from *child-directed speech* and each of two variables of language outcomes..145

Appendix A-Table 1. The reference table of codes and their descriptions at different levels on each tier.....165

Appendix G-Table 1. Pearson correlation coefficients for all possible pairs made by 16 measures of language input and 2 measures of PLS-AC and PLS-EC. *r* values that were statistically significant are bolded in the table .. 180

LIST OF FIGURES

Figure 2-1. (A) Total and (B) average amount of audio analyzed for each child with CIs participated in this study. 30

Figure 2-2. Block diagram of the general approach used by human analysts to code sampled audio files. Analyses presented in this chapter are based on coding steps shown by solid lines and solid boxes. The dashed lines and dashed-line boxes indicate steps associated with coding procedures for considering effects of noise and reverberation, which were developed for analyses of Chapter 3 and are discussed in that chapter in more detail. 33

Figure 2-3. Annotation scheme used by human analysts to code various sound events of interest in this study. The illustrated sample interval contains ~3.5 seconds of audio within a 30-second analysis interval drawn from a day-long audio recording of a child in the corpus, and was coded by a human analyst. For this interval, the CI child's father was communicating with his child. The two top rows of the display show the waveform and spectrogram, respectively. Coding consisted of *Praat Textgrid* tiers providing for annotation of the following information (top to bottom): (1) The Analyzed Interval tier indicated which 30-second sampled audio portions had been selected randomly for inclusion in the analysis (given with a 'y'); (2) The Adult Female tier contained speech from competent female talkers (or the primary female for the conversational situation, who was usually the target child's mother but necessarily always the mother, because our criteria specified that the primary talker was the one who held the floor); (3) The Adult Male tier was the same as Adult Female tier, except it was used to designate speech from male talkers; (4) The Target Child tier was used for the "target child" only and contained speech or speech-like or non-speech vocalization or cries from the target child; (5) The Other Talker tier contained vocal activities from other linguistically competent talkers (see text); (6) The Media tier contained any pre-recorded media-related content such as television, toys, tablets, etc.; (7) & (8) The Noise 1 and Noise 2 tiers contained noise sounds; (9) The Microphone Noise tier coded for microphone noise in the background which was detected. See Table 2-5 and Appendix A for further information on the structure and types of codes used in these nine tiers. Here, the father's conversation with the target child is coded within three intervals, two of which overlapped with a child's cry with medium interference determined by the "/M" code after "CRY". "T" label at Level 1 of these three coded intervals indicates that the speech was directed to the target child; see Table 2-0-5. The father's speech was transcribed for these three intervals while separated from Level 1 and Level 3 information using a ";" and "/", respectively. Further, an episode of the target child's crying is coded as partially overlapped with the father's speech with a level of interference judged to be medium (M) relative to his speech. See Ch. 3 for more details on the coding of noise. 33

Figure 2-4. Estimated (A) total words per day, (B) child-directed words per day, and (C) overhead words per day for each child in their home environments. In each panel a scatter plot shows the individual children, while the boxplot summarizes the distribution of language experienced in home environments. The data points are laid over a 1.96 standard error of the mean (95%

confidence interval) in red and 1 standard deviation shown by blue lines. The solid and dotted red line show the mean and median, respectively..... 47

Figure 2-5. The distribution of child-directed and overheard words per day in the language environments of each child with a CI. The ratio on the top of each bar is the proportion of child-directed words out of the total number of words experienced by each child per day..... 49

Figure 2-6. Estimated number of total words heard over years by the time of entry to kindergarten for each child. The "y axis" is presented in log scale. 50

Figure 2-7. Estimates of number of total words experienced over years by each child. Vertical numbers 1 through 5 on each bar represent the estimates for Year 1 through Year 5 post-implantation. 51

Figure 2-8. Estimated number of child-directed words heard over years by the time of entry to kindergarten for each child. The y-axis is shown on a log scale. 52

Figure 2-9. Variation in estimates of numbers of child-directed words experienced over years by each child. 52

Figure 2-10. Type-token ratio calculated from (A) total, (B) child-directed, and (C) overheard adult speech in the environment of each child with CIs. 53

Figure 2-11. Type-token curve and the area under them are shown for total words (blue line and blue shaded area), words directed to the target child (orange line and shaded area by orange), and adult-directed words (yellow line and shaded area by yellow)..... 56

Figure 2-12. Variation in the area under the type-token ratio calculated for 14 children participated in this study for (A) total speech, (B) child-directed speech, and (C) adult-directed speech. 58

Figure 2-13. Mean length of utterance of speech experienced by each child with CIs in her/his linguistic environments derived from (A) total speech, (B) child-directed speech, and (C) adult-directed speech..... 60

Figure 2-14. Variation across children in the percent of caregivers' responsiveness in the environment of each child..... 61

Figure 2-15. The relationship between number of child-directed words experienced by each child per day and caregiver responsiveness as a percentage.....	61
Figure 2-16. Variability across children in the rate of speech by which they experienced speech in their environments calculated from (A) total speech, (B) child-directed speech, and (C) and adult-directed speech.....	62
Figure 2-17. Distribution of measures of language input for child-directed vs. adult-directed speech in the environments of children with CIs.....	64
Figure 3-1. Structure of the approach for coding the audio samples for determining the presence of any source of overlapped sounds or speech and its level of interference.....	89
Figure 3-2. General approach for updating the number of words based on SRMR-CI value.	91
Figure 3-3. The relationship between level of interference identified by human and the corresponding SRMR-CI values for each speech interval. The central red line indicates the median of distribution and bottom and top edges of the box show the 25 th and 75 th percentiles, respectively.	94
Figure 3-4. Block diagram of the approach to construct the noise profile for environmental noise experienced by each child with a CI.	97
Figure 3-5. Frequency response of A-weighting filter in dB scale.....	99
Figure 3-6. The estimate total number of words per day for 14 children with CIs based on three approaches of traditional (unadjusted for noise), and for two methods adjusted for noise: human-based, and SRMR-CI-based.....	100
Figure 3-7. (A) The number of words lost per day and (B) percentage of words lost per day compared to the traditional method after considering the effect of noise on the number of total words based on human coding (Pair 1) and SRMR-CI metric (Pair 2). In each panel a scatter plot shows the individual children and boxplot summarizes the distribution of word loss for each child in his/her environment. The data points are laid over a 1.96 standard error of the mean (95% confidence interval) in red and a 1 standard deviation shown by blue lines. The solid and dotted red line show the mean and median, respectively.....	102
Figure 3-8. Distribution of clean and overlapped words in per-day estimates of the total number of words for each child.....	105

Figure 3-9. Distribution of various types of noise overlapped with speech in the environments of the 14 children with CIs. The code for various noise type are Speech = overlap speech, SNG = singing, NSV=non-speech vocalization, MUS=music, NOISE-GEN=general noise, NOISE-VOCAL=vocalization form other children, NOISE-BABB=multi-talker babbling; NOISE-ANIM=animal noise, NOISE-MECH=mechanized noise sources, NOISE-WATER=running water sound, NOISE-WIND=wind sound, CHILD-SP=speech from target child, CHILD-CRY=target child cry, CHILD-NSV=non-speech vocalization from the target child, MIX= multiple non-primary talkers overlapping or wo or more types of sound source above overlapping. See Appendix A for more information regarding these categories of noise or speech or sounds that overlapped with speech from the primary talker. 106

Figure 3-10. Estimated number of total words per day from traditional, human-based, and SRMR-CI-based approaches. 107

Figure 3-11. The total number of words extrapolated over 5 years after CI activation for each child estimated based on traditional approach. Numbers 1-5 next to data point circles reflect estimated child-directed words after 1, 2, 3, 4, or 5 years. 109

Figure 3-12. The total number of words extrapolated over 5 years after CI activation for each child estimated based on updated number of words by human auditory perceptual judgments. Numbers 1-5 next to data point circles reflect estimated child-directed words after 1, 2, 3, 4, or 5 years 110

Figure 3-13. The total number of words extrapolated over 5 years after CI activation for each child estimated based on updated number of words by SRMR-CI metric. Numbers 1-5 next to data point circles reflect estimated child-directed words after 1, 2, 3, 4, or 5 years. 110

Figure 3-14. The estimate total number of child-directed words per day for 14 children with CIs based on three approaches of traditional (unadjusted for noise; blue line), and for two methods adjusted for noise: human-based (orange), and SRMR-CI-based (yellow). 111

Figure 3-15. (A) The amount of words lost per day and (B) percent of words lost per day compared to the traditional method after considering the effect of noise on the number of total words based on human coding (Pair 1) and SRMR-CI metric (Pair 2). 112

Figure 3-16. Distribution of clean and overlapped words in per-day estimates of the number of child-directed words for each child..... 114

Figure 3-17. Estimated number of child-directed words per day from traditional, human-based, and SRMR-CI-based approaches. 115

Figure 3-18. The number of child-directed words extrapolated over 5 years after CI activation for each child based on the traditional approach. Numbers 1-5 next to data point circles reflect estimated child-directed words after 1, 2, 3, 4, or 5 years.	116
Figure 3-19. The number of child-directed words extrapolated over 5 years after CI activation estimated based on updated human auditory perceptual judgments of the level of noise interference. Numbers 1-5 next to data point circles reflect estimated child-directed words after 1, 2, 3, 4, or 5 years.	117
Figure 3-20. The number of child-directed words extrapolated over 5 years after CI activation estimated based on updated word counts based on SRMR-CI metric. Numbers 1-5 next to data point circles reflect estimated child-directed words after 1, 2, 3, 4, or 5 years.	117
Figure 3-21. Distribution of children for the total number of words per day (first column), clean/updated number of words per day (second column), and clean/updated number of child-directed words per day (third column) they experienced in their environments, estimated based on (A) traditional, (B) human-based, and (C) SRMR-CI-based approaches.	119
Figure 3-22. Mean and standard deviation of <i>Mahalanobis</i> distance between child-directed and adult-directed speech intervals at four levels of speech degradation.	120
Figure 3-23. The average spectral energy of noise under four sub-bands experienced by 14 children with CIs.	121
Figure 3-24. Distribution of children based on the experienced average spectral energy of noise under four sub-bands of 0-1, 1-3, 3-5, and 5-8 kHz.	122
Figure 3-25. Histogram of noise loudness for each child.	123
Figure 3-26. Variability across children in exposure to noise measured by (A) noise exposure index for all frames, (B) 90 percent noise exposure index, i.e., noise exposure index for noise frames with loudness of higher than 90 dBA.	123
Figure 3-27. Noise descriptors for 14 children with CIs.	125
Figure 3-28. Distribution of children based on L50-Leq index.	125

Figure 4-1. The relationship between total number of words and (A) PLS-AC and (B) PLS-EC standard scores, as well as the relationship between number of child-directed words and (C) PLS-AC and (D) PLS-EC standard scores. The r values are pairwise Pearson correlation coefficients between pairs of language input measure and language outcome. R-squared values are the coefficient of determination obtained from fitted ordinal linear regression models. The least-square regression lines are superimposed on each scatter plot. 144

Figure 4-2. The relationship between updated total number of words based on human coding of level of noise interference and (A) PLS-AC and (B) PLS-EC standard scores, and human-based updated number of child-directed words and (C) PLS-AC and (D) PLS-EC standard scores. The r values are pairwise Pearson correlation coefficients between pairs of language input measure and language outcome. *R-squared* values are the coefficient of determination obtained from fitted ordinal linear regression models. The least-square lines are superimposed on each scatter plot. 146

Figure 4-3. The relationship between updated total number of words based on SRMR-CI metrics and (A) PLS-AC and (B) PLS-EC standard scores, as well as SRMR-CI-based updated numbers of child-directed words and (C) PLS-AC and (D) PLS-EC standard scores. The r values are pairwise Pearson correlation coefficients between a given language input measure and the respective language outcome measure. *R-squared* values are the coefficient of determination obtained from fitted ordinal linear regression models. The least-square lines are superimposed on each scatter plot..... 147

Figure 4-4. The relationship between average SRMR-CI obtained over recordings and language outcome scores of PLS5-EC and PLS5-AC at 6 months post-implantation. The grey line is the least-square line calculated for each data. The R-squared values and p-values for the ordinary linear regression model are reported on each plot. 148

Appendix B-Figure 1. Type-token curves derived from total, child-directed and overheard speech in the environment of Child 1.....174

Appendix F-Figure 1. Correlation matrix scatterplot with least-square fitted lines for pairs of measures obtained from 16 measures of language input plus 2 measures of language outputs. Refer to Table 4-1 to match the numbers in this figure to their corresponding measures of language input.....179

KEY TO ABBREVIATIONS

CI	Cochlear Implants
CDS	Child-directed Speech
ADS	Adult-directed Speech
dB	Decibels
dBA	A-weighted Decibels
MD	Mahalanobis Distance
SRMR-CI	Speech-to-Reverberation Modulation Energy Ratio tailored for CI users
SES	Socioeconomic Status
LENA	Language ENvironment Analysis
SD	Standard Deviation
IQR	Interquartile Range
M	Male
F	Female
SE	Sensorineural
CO	Congenital
UN	Unknown
RL	Right and Left (bilateral implants)
HA-L	Hearing Aid and Left
R-HA	Right and Hearing Aid
Com	Communication Mode
HL	Hearing Loss

PTA	Pure-Tone Average
OC	Oral Communication
TC	Total Communication
HS	High School
BA	Bachelor
CMV	Cytomegalovirus Hearing Loss
EVA	Enlarged Vestibular Aqueducts
GDD	Generalized Developmental Delay
ANSD	Auditory Neuropathy Spectrum Disorder
ASD	Autism Spectrum Disorder
cCMV	Congenital Cytomegalovirus
AWC	Adult Word Count
L	Low
M	Medium
H	High
SNG	Signing
NSV	Non-Speech Vocalization
P	Primary
S	Speech
AF	Adult Female
AM	Adult Male
IRR	Inter-Rater Reliability
LOI	Level of Interference

TTR	Type-Token Ratio
AUTTC	Area Under the Type-Token Curve
MLU	Mean Length of Utterance
MLUm	Mean Length of Utterance in Morpheme
MIN	Minimum
MAX	Maximum
SNR	Signal-to-Noise Ratio
CVC	Consonant-Vowel-Consonant
SPL	Sound Pressure Level
SI	Speech Intelligibility
MFCC	Mel-Frequency Cepstral Coefficient
FFT	Fast Fourier Transform
NEI	Noise Exposure Index
SP	Speech
MECH	Mechanical
ANIM	Animal
BABB	Babbling
GEN	General
MUS	Music
PLS	Preschool Language Scale
PLS-AC	Preschool Language Scale-Auditory Comprehension
PLS-EC	Preschool Language Scale-Expressive Communication
C-ENT	Child-oriented Entertainment

CM	Child Male
CF	Child Female
C	Child
EDU	Educational
INFO	Informational
CLASS	Classical
MET	Metal
ENT	Entertainment
MOV	Movie
VOC	Vocalization

CHAPTER 1: Introduction

Approximately 188,000 individuals around the world and more than 41,000 adults and 26,000 children in the United States benefit from cochlear implants (CIs) to access sounds and speech in their environments (“National Institutes of Health, Cochlear Implants,” 2010). Despite a nearly successful development of spoken language in some pediatric recipients of CIs, the sources of enormous individual variability across children reported at all cochlear implant centers around the world is still unknown (Peterson, Pisoni, & Miyamoto, 2010). The unresolved question of sources of this enormous variability highlights the high risk of lack of developing age-appropriate language to the extent that some children with CIs may never develop functional spoken language skills (Peterson et al., 2010). Therefore, there is a substantial need to understand factors that contribute to this tremendous individual difference in language outcomes of children with CIs. The present dissertation aims to address this knowledge gap by studying an under-explored factor in language outcomes, variability in the quantity and quality of language input experienced by children with CIs in their naturalistic language environment.

1.1. Individual variability in early linguistic environments may explain variability in language attainment

There has been very little research on the impact of children’s language environment on their outcomes. Most of the research into variability in cochlear implant outcomes has been focused on device factors and implantation variables such as age at implantation and electrode configuration (Peterson et al., 2010). However, these factors together have explained a small fraction of the individual variability in language outcomes of this at-risk pediatric population, leaving a tremendous portion of sources of individual differences still unexplained (Niparko et al., 2010; Peterson et al., 2010). Few studies highlighted the key role of children’s language experience

on their language outcomes (Szagun & Stumper, 2012). However, these studies focused on specific properties of language input such as grammatical structure of maternal speech, which were derived from speech samples at cochlear implant centers rather than children's natural home environments. Therefore, while informative, such work skips basic questions about the home environment of children with CIs such as the quantity and quality of speech input experienced by children after their implantation.

1.2. Quantity of early language experience

There are many sources of evidence that nature of language input significantly affects language attainment in children with typical hearing, particularly in early childhood. Pioneering work by Hart and Risley (1995) established that there was a tremendous gap in the amount of language children experienced before kindergarten. They demonstrated a 30-million-word gap between children who lived in environments with highest amount of language input and those who experienced language environment with lowest quantity of language input by age 3. These findings have been refined and extended by recent studies where the original findings of 30-million-word gap has been associated to the characteristics of children's linguistic environments rather than solely to their socio-economic status (Sperry, Sperry, & Miller, 2018). More importantly, all these studies demonstrated that the extreme individual disparities in quantity of early language experience is tightly translatable to disparities in children's language outcomes. However, we have very limited knowledge about how quantity of early language experience varies across children with CIs in their home environment, knowing that these children are extremely variable in their language outcomes. The present project is inspired by the approach of Hart and Risley (1995) to understand to what extent individual children vary in quantity and quality of language input they experienced early after their implantation.

1.3. Quality of early language experience

Early language input experienced by children may also vary across individuals for its quality, which may significantly impact children's language development. Child-directed style is one of the major components of quality of early language input that best fosters language development in typically-hearing children (Weisleder & Fernald, 2013). This spoken communication style contains rich acoustic, linguistic and social-pragmatic information that together contribute significantly to improving the quality of speech experienced by children regardless of their hearing status (Dilley, Lehet, Wieland, Arjmandi, & Kondaurova, 2018; Weisleder & Fernald, 2013). Typically-hearing children who experience relatively greater amount of high-quality child-directed speech develop better language and cognitive skills (Golinkoff, Hoff, Rowe, Tamis-LeMonda, & Hirsh-Pasek, 2019; Weisleder & Fernald, 2013). This supportive effect of child-directed speech on children's vocabulary growth persists regardless of whether the talker is primary caregiver or other household members and siblings (Shneidman, Arroyo, Levine, & Goldin-Meadow, 2013). Therefore, characterizing variability across children in the amount of high-quality child-directed speech may help us to explain an important fraction of variability in their language outcomes, particularly in children with CIs who hear an impoverished and degraded version of speech input.

Furthermore, the amount of high-quality child-directed speech experienced in a day impacts other aspects of quality of language input. Children who live in environments with higher quantity of child-directed speech have greater chance of hearing more word types with higher repetition as well as experiencing higher examples of repeated and partially repeated utterances (Hoff, 2006). In addition, early language environments with higher amount of child-directed speech per day increases the chance of getting exposed to utterances that are semantically related,

leading to greater spoken communicative skills (Hart & Risley, 1995; Hoff-Ginsberg, 1991, 1994). Therefore, children who hear relatively lower amount of child-directed words in a day in their environments are at risk of developing smaller vocabulary size (Hart & Risley, 1995; Janellen Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991). The quality of the experienced child-directed speech may also differ across children for the length of utterances, leading to developing variable language skills across typically-hearing children (Hoff & Naigles, 2002). The emerging consensus from these studies is that, at least for children with typically hearing, individual variability in amount of child-directed speech experienced per day explains variability in their language attainment. However, there is no knowledge about how the amount of high-quality child-directed speech vary across early language environments of children with CIs.

More importantly, there is a highly degradative effect of noise, which disproportionately affects speech reception ability in individuals with cochlear implants (Neuman, Wroblewski, Hajicek, & Rubinstein, 2004). This effect has not yet been taken into consideration in a study of language development in children with cochlear implants. Listeners with CIs have great difficulties in understanding speech in background noise and/or reverberation compared to listeners with typical hearing, particularly when speech co-occurs with speech from multiple talkers (Caldwell & Nittrouer, 2013). This factor becomes more important knowing that children with CIs spend substantial portions of their daily lives in noisy environments (Busch, Vanpoucke, & van Wieringen, 2017). Therefore, noise in the environment is a crucial factor for language development of children with CIs that affects the quantity and quantity of language input experienced by children in their daily lives.

The findings from these studies together demonstrate that individual variability in language outcomes of children is tightly related to their differences in quantity and quality of the experienced

speech in their early language environments. However, we have very limited knowledge about how the quantity and quality of speech experienced by children with CIs vary across children early after implantation and how these variations may relate to variability in their language outcomes.

1.4. The present study

Variability across children with typically hearing in quantity and quality of early language experience is strongly associated with their differences in language outcomes. Chapter 2 aimed to investigate the extent of individual variability across 14 early-implanted children with CIs in the quantity and quality of their language experience soon after their implantation. We characterized language input experienced by these children in their home environments for the total amount of words, the amount of child-directed words, diversity of words, complexity of utterances, the degree of adults' engagement in child-adult interaction, and rate by which speech experienced by children in their environments. We examined the extent the individual children with CIs experienced language environments with different quantity and quality characterized by these traditional measures of language input. Considering that many children with CIs around the world are at high risk for lack of developing age-appropriate expressive and receptive language skills, this aim addresses a critical and timely question.

Quantity and quality of early linguistic environments have been mainly investigated based on traditional measures of language input. There is no knowledge about how noise and reverberation in the actual environments of children with CIs may impact the number of useful words they experience and how it differentially affects the properties of children's environments. Chapter 3 develops novel approaches to update the number of total words and high-quality child-directed words heard by children with CIs in their early linguistic environments to investigate the extent to which quantity and quality of linguistic environments of children vary after the noise-

based adjustment of language input. We analyzed speech input based on human coding of level of noise interference as well as a quantitative metric of speech intelligibility to adjust the quantity and quality of speech input for environmental noise and reverberation. This chapter addresses a crucial knowledge gap about the change in quantity and quality of language experienced by children with CIs after considering the effect of environmental noise and reverberation.

Chapter 4 uses the traditional and updated measures of language input to examine how well each of these measures predict language outcomes of the 14 early-implanted children with CIs participated in this study. The finding from this research project will provide new evidence about how quantity and quality of early linguistic input experienced by children with CIs vary across individuals and how these changes are associated with children's language outcomes in this at-risk pediatric population. We expect that outcomes of this study will provide evidence-based knowledge about quantity and quality of early language environments of children with CIs. This knowledge can be used to develop personalized early intervention strategies for clinicians and parents to maximize the likelihood of effective spoken communication skills in children with CIs.

CHAPTER 2: Variability in quantity and quality of early linguistic experience in children with cochlear implants

2.1. Introduction

Variability in language outcomes of children with CIs

Cochlear implants (CIs) are prosthetic biomedical devices which have dramatically impacted the lives of many children with severe-to-profound hearing loss by allowing them access to sounds and speech (Geers, Nicholas, & Sedey, 2003; Svirsky, 2017; Wilson & Dorman, 2008). Gaining access to sounds and speech permits children with CIs to acquire verbal language (Blamey et al., 2013; Cohen, Waltzman, Roland, Staller, & Hoffman, 1999; Stallings, Gao, & Svirsky, 2000; Summerfield & Marshall, 1995), to develop cognitive abilities (Pisoni, 2012), and to improve scholastic achievement (Spencer, Barker, & Tomblin, 2003). Despite all these advantages brought by cochlear implantation, there remains enormous unexplained variability among children with CIs in terms of their language outcomes after CI activation (Ertmer & Goffman, 2011; Geers et al., 2003; Holt, Beer, Kronenberger, Pisoni, & Lalonde, 2012; Niparko et al., 2010; Peterson et al., 2010; Pisoni, 2012; Svirsky, Teoh, & Neuburger, 2004a; Szagun, 2001; Tobey et al., 2013; Tyler et al., 2000). The magnitude of this variability is notably larger than that of children with typical hearing (Duchesne, Sutton, & Bergeron, 2009; Svirsky et al., 2004a; Szagun, 2001). After more than a decade of research on sources of this enormous variability in language outcomes of pediatric CI recipients, sources of such variability still remain poorly understood. Therefore, it is of the utmost importance to further focus on under-studied factors that potentially influence variability in the development of language in children who receive CI prosthetic devices.

Early linguistic environment has been shown to play a primary role in language development in children with typical hearing (Hart & Risley, 1995; Hoff, 2003; Newman, Rowe,

& Bernstein Ratner, 2016; Weisleder & Fernald, 2013). Importantly, the quantity and quality of language input has been shown to vary substantially across language environments of typically-hearing children. Hart and Risley (1995) demonstrated that there is substantial variability across children in properties of language they experience in their environments in a day. This large variability across children in amount of words they experience can be compounded over years to equate to a very large number. The scale of this growth was extremely large over years such that, by age 3, children living in linguistic environments with the highest relative quantity and quality experienced 30 million more words compared to those who lived in environments with the lowest quantity and quality of language input.

The basic finding of Hart and Risley (1995) – showing substantial differences across early linguistic environments of children with typical hearing – has been replicated in several recent studies (Golinkoff et al., 2019; Hirsh-Pasek et al., 2015; Romeo et al., 2018). For children with typical hearing, this extreme individual variability in exposure to speech early in childhood translates into individual differences in their language outcomes and cognitive skills (Golinkoff et al., 2019; Hart & Risley, 1995; Hoff, 2003). However, there is a major knowledge gap about the extent early linguistic environments vary across children with CIs early in childhood, recognizing that individual variability in this at-risk pediatric population in language outcomes is considerably higher than that of children with typical hearing.

In the current study, we aimed to address this knowledge gap by exploring how quantity and quality of early linguistic environments vary across children with CIs along multiple linguistic and indexical dimensions. Inspired by Hart and Risley (1995)’s work, we analyzed audio samples recorded from naturalistic auditory environments of children with CIs to estimate the quantity and

quality of language input experienced by children in a day. This permitted estimates of how this differential experience could impact language exposure over years.

To this end, in this chapter, we first discuss factors that have been shown to influence language development in children with CIs. Then, we will present evidence that the quantity and quality of early linguistic environments are key factors for language development in children both with typical hearing and with CIs, something that remains under-studied in explaining variability in language outcomes in this at-risk population. Next, we will present methods developed in this study to analyze early linguistic environment of children with CIs. We analyzed audio samples recorded from naturalistic environments of 14 early-implanted children with CIs to obtain estimates of the quantity and quality of language input they experienced. We particularly estimated the number of total words and child-directed words per day, as well as several other measures of lexical, morphosyntactic, and social-pragmatic information to characterize early language input. The results will be then presented in the *Results* section and will be further considered in the *Discussion*.

Known factors contributing to variability in language outcomes of children with CIs

Several factors have been studied in relation to pediatric CI users, which nevertheless fail to account for a large proportion of variance in these children's language outcomes. Known factors contributing to language acquisition in children with CIs relate to intervention characteristics (Blamey et al., 2013; Ertmer, Young, & Nathani, 2007a; Graham et al., 2009; Holden et al., 2013; Iler Kirk et al., 2002; Kirk, Miyamoto, Ying, Perdew, & Zuganelis, 2000; Lazard et al., 2010; Levine, Strother-garcia, Golinkoff, & Hirsh-pasek, 2016; Nicholas & Geers, 2006; Anu Sharma, Dorman, & Kral, 2005; Summerfield & Marshall, 1995; Szagun & Schramm, 2016; Szagun & Stumper, 2012; Tobey et al., 2013; Willstedt-Svensson, Löfqvist, Almqvist, & Sahlén, 2004),

characteristics of the CI devices (Blamey et al., 1992; Lazard et al., 2010), and factors related to device-brain interactions (Aschendorff, Kromeier, Klenzner, & Laszig, 2007; Bierer, Spindler, Bierer, & Wright, 2016; Blamey et al., 1992; Finley, C. C., & Skinner, 2008; Holden et al., 2013; Jahn & Arenberg, 2019; Yukawa et al., 2004). Factors related to intervention include age of hearing loss identification, age at CI implantation (Diaz, Labrell, Le Normand, Guinchat, & Dellatolas, 2019; Dunn et al., 2015; Ertmer, Young, & Nathani, 2007b; Iler Kirk et al., 2002; Kirk, Miyamoto, et al., 2000; Levine et al., 2016; Nicholas & Geers, 2006; Anu Sharma et al., 2005; Szagun & Schramm, 2016; Szagun & Stumper, 2012; Tobey et al., 2013; Tomblin, Barker, Spencer, Zhang, & Gantz, 2005; Willstedt-Svensson et al., 2004), duration of implant use, communication mode (Connor, Alexander, & Teresa, 2000; Stacey, Fortnum, Barton, & Summerfield, 2006; Tobey, Geers, Brenner, Altuna, & Gabbert, 2003), duration of deafness prior to implantation (Blamey et al., 2013; Graham et al., 2009; Summerfield & Marshall, 1995), the level of residual hearing before CI activation (Gomaa, Rubinstein, Lowder, Tyler, & Gantz, 2003), as well as the history of hearing aid use (Nicholas & Geers, 2006). Factors such as working memory (Pisoni & Cleary, 2012), gender, family SES, and age of hearing loss onset can be categorized as factors associated with children's personal characteristics. However, a large fraction of variability in language outcomes in children with CIs remains unexplained, even when these factors are taken into account.

Sources of variability in language outcomes in children with CIs: A critical gap in knowledge

Findings from language development studies and perceptual studies of speech recognition in children with CIs highlight that sources of variability in development of language in children with CIs have remained poorly understood. Children with CIs showed largely variable performance in perception and production of indexical and contextual information in speech

(Anderson, Lazard, & Hartley, 2017; S.-C. Peng, Tomblin, & Turner, 2008; Z. Peng, Hess, Saffran, Edwards, & Litovsky, 2019; Reed & Delhorne, 2005). For instance, although children with bilateral CIs showed above-chance performance in discrimination of consonant sounds, there were substantial differences among individual children in their performance (Peng et al., 2019). A similar pattern of large inter-subject variability was reported in pediatric CI patients in perceiving and producing speech intonation (Peng et al., 2008) and recognizing talkers' voices (Tamati, Janse, Pisoni, & Baskent, 2017). Adults with CIs performed quite differently in separation of ten categories of environmental sounds when tested on four closed sets (Reed & Delhorne, 2005). These studies raise a fundamental question: that is, which component involved in language learning of children with CIs has not yet been well-studied that might explain the tremendous variability in children's language outcomes? This fundamental question is addressed in the present dissertation.

As mentioned, a large fraction of variability in language outcomes of children with CIs has not yet been explained, even after controlling for all the above-named factors. Relevant to selecting a population for study, age at implantation have been identified as a strong predictor of language outcomes (Holt & Svirsky, 2008; Nicholas & Geers, 2007; Svirsky et al., 2004a; Tomblin et al., 2005). Results from these studies suggested that there is a sensitive period where children acquire language more effectively. These findings suggested that cochlear implantation is most effective when it occurs before 24 months of age (Geers, Nicholas, & Moog, 2007a; Holt & Svirsky, 2008; Svirsky, Teoh, & Neuburger, 2004b). However, recent findings by Szagun and Stumper (2012) demonstrated that the effect of age at implantation is highly dependent on the quality and quantity of language input and the duration of language experience. These findings demonstrate that although small portions of variability in language outcomes of children with CIs have been

explained by these factors, a large fraction of variance in language outcomes has not yet been explained, leading to high risk for language delay in some children.

Theoretical frameworks relevant to language development

A child's language system reflects interactions between the linguistic environment and the plastic pediatric brain. These two sources of variability together form a complex dynamic system for language development (Sharma, Dorman, & Spahr, 2002; Anu Sharma et al., 2005; Szagun & Schramm, 2016; Szagun & Stumper, 2012). Variability in the outcomes of this complex system can be understood within *dynamic systems theory* (Smith & Thelen, 2003; van Geert & van Dijk, 2002). Under this theoretical framework, complex cognitive systems, such as language, develop through interaction between physical and social components of environments (Smith & Thelen, 2003; Verspoor, Lowie, & Van Dijk, 2008). Therefore, the weighting system for perception and production of linguistic units is the outcome of active, long-term daily interactions with linguistic environments, constructing a complex, dynamic language-learning process. The present study is conceptually informed by this framework. As this theory states (Verspoor et al., 2008), examining patterns of variability in language input from children's environments is central to providing better explanations of sources of differences among individual children in language outcomes, particularly children with CIs (Geers, A. E., Strube, M. J., Tobey, E. A., & Moog, 2011).

Impacts of the language environment on language development can also be understood relative to widespread findings that children exploit statistical and distributional characteristics of speech input in forming linguistic representations (Johnson & Jusczyk, 2001; Saffran, 2002; Saffran & Kirkham, 2018). For instance, they learn within-word transitional probabilities as important cue for vocabulary acquisition (Saffran, Aslin, & Newport, 1996; Saffran, Johnson, Aslin, & Newport, 1999; Saffran, Newport, & Aslin, 1996; Saffran & Kirkham, 2018). Children

further follow transitional probabilities at longer time-scales relevant to both prosodic cues and hierarchical linguistic structures. Saffran et al. (1996) showed that children, by at least 8 months of age, track within-word transitional probabilities that are significantly higher than the between-word transitional probabilities. They also attend more to within-event probabilities between syllables and audiovisual events, leading to perception of a sequence as a unitary event. The same pattern of significantly higher within-event probabilities compared to between-events probabilities was observed between syllables and audiovisual events (Bulf, Johnson, & Valenza, 2011; Saffran, Newport, et al., 1996). From these findings it follows that higher exposures to language might render better estimates of within-word and across-word transitional probabilities. Therefore, these bodies of research on development as a dynamic system and statistical learning help inform understanding of how variability in exposure to language across children could lead to differential time-courses of development of linguistic systems.

Quantity of ambient language as a factor in language development in children with typical hearing abilities

Our core hypothesis was that sources of variability identified for children with typical hearing which help account for variability in their language development could also explain variability in language development in children with CIs. Prior studies have identified *the early linguistic environment* as a critical factor influencing later language outcomes for children with typical hearing (Hart & Risley, 1995; Hoff, 2003; Hurtado, Marchman, & Fernald, 2008; Janellen Huttenlocher et al., 1991; Neville & Bruer, 2001; Weisleder & Fernald, 2013). Seminal work by Hart and Risley (1995) compared the quantity of words experienced by children with either upper socioeconomic status (SES) (13 children), middle SES (10 children), or low SES (10 children). They analyzed children's language environments from the time children were 7-9 months old until

they were 3 years old based on a series of sequential monthly, hour-long recordings. Their estimates from analyzing these hours-long samples, based on linear extrapolation from the average number of words per hour over years, showed that after three years, the difference in number of words experienced by children in the highest SES bracket and the lowest SES bracket was approximately 30 million words, a word difference that is known as the so-called *30-million word gap*.

Sperry et al. (2018) examined the initial claim of Hart and Risley (1995) and showed that their findings from studying five American communities with widely variable SES levels did not support Hart and Risley (1995)'s claim that 30-million word gaps could be attributed to living in environments with lower SES versus those of higher SES families. Contrary to Hart and Risley's findings, Sperry et al.'s investigation of videotapes of children's spoken communication with their mother, grandmother and father showed that differences among children in exposure to amount of language was not associated with children's SES. Rather, there were also children living in low-income families experiencing a large amount of language input in a day. Sperry et al. (2018) also argued that overheard speech is as valuable as child-directed speech (CDS) for language learning, a claim that was rejected by Golinkoff et al. (2019). Golinkoff et al. (2019) further responded to Sperry et al. (2018)'s findings and argued that the difference across children is more about the amount of high-quality speech directed to children in their environments rather than solely the total amount of speech available in children's language environment. The finding that greater quantity of language experience leads to higher vocabulary growth and attainment has been supported in multiple studies (Ramírez-Esparza, García-Sierra, & Kuhl, 2014; Rowe, Raudenbush, & Goldin-Meadow, 2012; Weisleder & Fernald, 2013).

The main takeaway from these studies is that the quantity of speech experienced by children in their language environments matters for their language development. In other words, these findings highlighted the fact that the amount of spoken communicative experience provided to children in their linguistic environments may substantially vary across families, leading to large individual differences in rates of language growth, ability to successfully comprehend language, and produce speech. It was shown that the individual variability in amount of words experienced was not mediated through family's socioeconomic status, but rather the amount of high-quality language input. Therefore, variation in the quantity and quality of language input experienced by each child needs to be appropriately examined in order to better capture the connection between changes in specific components of language input and individual differences in language development (Hoff, 2006).

Findings from studies like Hart and Risley (1995) and Sperry et al. (2018) motivated the present study to focus on characterizing variability in amount (i.e., quantity) of language in children with CI's early linguistic environments before even targeting to seek their connection with language outcomes. We have almost no knowledge about how the quantity and quality of early linguistic environments of children with CIs vary across children. We examined the extent of variation across children with CIs in the quantity and quality of language input they experienced early after implantation to gain preliminary evidence of how children may be disproportionately benefited or disadvantaged due to individual differences in properties of early language input.

Quality of language input and its influences on language development

Studies with typical hearing children have also identified multiple factors which together are considered to relate to the *quality* of language input. This construct includes a notion of the person to whom individuals are speaking, how they pronounce their words and structure their

utterances, and the timing and conceptual relatedness of those utterances. All these factors were examined in the present study.

Children learn language in natural auditory environments, where often multiple adults and other children are present (Busch et al., 2017). Whether speech was directed to the child – namely, whether it was *child-directed speech* – or else whether the child overheard speech that was not addressed to him/her in any way, is a major component of quality of language input experienced by child (Fernald, 1989; Fernald, 2000; Frermann & Frank, 2016; Kuhl, 2000; Schachner & Hannon, 2011; Shneidman et al., 2013; Wang, Bergeson, & Houston, 2018). In fact, the positive effect of amount of speech input on children’s language outcomes depends largely on whether the speech was child-directed or overheard (Barnes, Gutfreund, Satterly, & Wells, 1983; Hoff-Ginsberg, 1986; Hoff, 2003, 2006; Janellen Huttenlocher et al., 1991; Shneidman et al., 2013; Weisleder & Fernald, 2013). The facilitative effect of child-directed speech on vocabulary growth is present regardless of whether the talker of speech is primary caregiver or other household members and siblings (Shneidman et al., 2013). Therefore, differences among children with CIs in the amount of child-directed speech translates into their difference in both *quantity* and *quality* of language input. It is still unknown how individual children with CIs differ in terms of amount of high-quality child-directed speech they experience in their early linguistic environments. In the following paragraph, we will clarify the notion of what is “*high quality*” in terms of how speech is spoken (pronunciation factors) and what is said – cf. the structure of utterances (word choice, MLU, etc.).

Child-directed speech provides acoustic, linguistic and extra-linguistic qualities that distinguish it from adult-directed speech or overheard speech. Children not only prefers listening to child-directed speech over adult-directed speech (Fernald & Kuhl, 1987; Wang, Shafto, &

Houston, 2018), but this preference also facilitates their language learning (Hoff, 2003; Weisleder & Fernald, 2013). Child-directed speech is often produced with acoustic qualities that distinguish it from adult-directed speech. Some of these distinguishing features are thought to facilitate speech processing and language learning both in children with typical hearing and children with CIs. For example, prosodic characteristics of CDS facilitates speech processing through enhancing their attention to speech (Bergeson, 2011; Fernald, 1992; Gout, Christophe, & Morgan, 2004; Kondaurova, Bergeson, & Dilley, 2012; Kondaurova, Bergeson, & Xu, 2013; Mehler, 2001; Nazzi, Dilley, Jusczyk, Shattuck-Hufnagel, & Jusczyk, 2005) and constrains children's lexical access (Levitin et al., 2003; Thiessen, Hill, & Saffran, 2005). Higher fluctuation of pitch over time in this speaking style was shown to provide valuable syllabic stress cues that fostered word segmentation in listening to speech for children with CIs (Spitzer, Liss, Spahr, Dorman, & Lansford, 2009). These findings together demonstrate that studies should consider variability across children in their early linguistic input based on the amount and quality of child-directed speech experienced by children with CIs.

The impact of child-directed speech on the quality of language input extends to other aspects of language such as lexical diversity and morphosyntactic complexity. For example, exposing children to a higher amount of high-quality child-directed speech in a day increases children's chances of experiencing both more word types and higher repetitions of words (Hoff, 2006). Mothers who speak more to their children provide more examples of semantically relevant utterances, which also facilitates children's language learning (Hart & Risley, 1995; Hoff-Ginsberg, 1994). These considerations show that child-directed speech contains high-quality features in terms of linguistic information. Taken together, these findings demonstrate that using a child-directed speaking style enriches linguistic and extra-linguistic information experienced by

children to the extent that differential exposure to this speech style is potentially a strong source of variability in language outcomes of children with CIs. We have minimal knowledge about individual variability across children with CIs in experiencing high-quality child-directed speech.

Another aspect of verbal communication thought to influence child language development is *parental responsiveness*. Parental responsiveness relates to both timing and “contingency” – a construct that unites the notions of child-directedness and semantic cohesion. It is, therefore, an important aspect of assessing quality of early linguistic environments that significantly influences children’s vocabulary growth early in childhood. Parents’ prompt and contingent responses to children’s spoken communication enhances children’s understanding of pragmatic aspects of their native language and facilitates construction of word-referent mappings (Goldstein & Schwade, 2008; Tamis-LeMonda, Kuchirko, & Song, 2014). Mothers’ responses to their prelinguistic infants’ emotional actions at 9 months was a strong predictor of infants’ language skills at 21 months (Nicely, Tamis-Lemonda, & Bornstein, 1999). The same positive effect of parental responsiveness was observed on infants vocabulary growth (Tamis-LeMonda, Bornstein, Kahana-Kalman, Baumwell, & Cyphers, 1998), communication diversity (Beckwith & Cohen, 1989), and faster rate of language development (Tamis-LeMonda, Bornstein, & Baumwell, 2001; Tamis-LeMonda et al., 1998). Variability in the level of parents’ responsiveness to their children’s verbal communicative activities explained variability in infants’ success in achieving language milestones (Tamis-LeMonda et al., 2001, 1998). The rate of language development in infants who experienced linguistic environments with high-responsivity mothers (90th percentile) at 9 and 13 months was remarkably higher than those from families with low-responsivity mothers (10th percentile), with a time lag of 4 to 6 months in milestone achievements (Tamis-LeMonda et al., 2001, 1998). These findings show that parental responsiveness is an important component of early

linguistic environments for understanding the connection between variability in the quality of language input and children's language outcomes. The role of this factor becomes more important in children with cochlear implants who have limited access to fine-grained acoustic cues in speech and whose speech recognition is highly susceptible to environmental noise.

Taken together, studies on early linguistic environments of children who are typically developing have demonstrated that variability in quantity and quality of language input strongly translates into variability in children's language outcomes. Therefore, systematic investigation of variability across children with CIs in their early language experience may help us to better explain the sources of variability in these children's language outcomes.

Current Study

There has been little research into variability in the quantity and quality of language input experienced by children with CIs in their natural language environments. Investigating variability across children with CIs in their early language experience may help us to better explain variability across children in their language outcomes. It further allows us to examine how properties of language input contribute to variability in outcomes above and beyond factors associated with child's characteristics and intervention strategies (Geers et al., 2003; Holt et al., 2012). The findings from the literature discussed above support the notion that, at least for children with typical hearing, language outcomes are predicted by factors having to do with quantity and quality of language input. The present dissertation provides the first comprehensive study analyzing auditory environments of children with CIs in their home environments. The goal was to characterize language input along multiple linguistic dimensions that have been widely shown to play important roles in language outcomes of children with normal hearing as well as, to a much lesser extent, children with CIs.

In the present study, children had their CIs activated between approximately 8 months and 22 months of age. Prior research has shown that this early age range is a time period of optimal neuroplasticity for learning language in children with CIs (Sharma et al., 2005, 2004; Svirsky, Robbins, Kirk, Pisoni, & Miyamoto, 2000). We characterized the *quantity* of language input experienced by each child based on the number of words they heard in a day. To characterize the quality of language input, the amount of child-directed speech, lexical diversity of speech, morphosyntactic complexity of speech input, rate of speech input, and caregiver responsiveness were measured to understand the range of individual differences among children with CIs in the quantity and quality of language in their home environments. We studied the magnitude of variability across children with CIs in their linguistic environments as a potential factor contributing to the observed enormous variability in language outcomes of this at-risk population.

We analyzed audio recordings from 14 children with CIs to examine the quantity and quality of their early language experiences. Inspired by the pioneering work of Hart and Risley (1995), we randomly sampled from naturalistic audio recordings obtained from children's home environments to estimate quantity and quality of language input. Quantity and quality of language input were characterized for lexical richness of input (i.e., the amount of words (Hart & Risley, 1995) and lexical diversity (Carroll, 1938) as well as morphosyntactic complexity of speech input (Hoff, 2003). We also measured one aspect of caregiver responsivity, which was taken here as the frequency with which children's speech or speech-like vocalization were responded to promptly with child-directed speech utterances.

This chapter aimed to answer the following questions:

Question 1: To what extent do children with CIs vary in the amount of total words and high-quality child-directed words they experienced per day in their early linguistic environments?

It was expected that children with CIs will substantially vary in the amounts of total and child-directed words they experienced in a day in their naturalistic environments.

Question 2: To what extent do quantity and quality differences across children with CIs “scale up” when extrapolated over multiple years to a degree of variability comparable to the 30-million word gap between most- and least-advantaged children in terms of quantity and quality of their language input (Hart & Risley, 1995)?

We expected to observe a similar pattern involving a gap of approximately 30 million words between most-advantaged and least-advantaged children in this study, consistent with prior findings of Hart and Risley (1995). We anticipated that this word gap would appear for the amount of high-quality child-directed speech experienced by children and will be a function of the quantity and quality of children’s linguistic environments rather than their socioeconomic status.

Question 3: How does the quality of the early linguistic environment vary across children with CIs in terms of lexical, morphosyntactic, and social-pragmatic information, as well as speech rate?

It was expected that substantial variability across children would exist in the quality of their language input, which would extend to other domains of language, including lexical diversity, morphosyntactic complexity, parental responsiveness, and rate of speech, whether overheard or directed to the child.

Question 4: How does speech directed to children with CIs differ from adult-directed speech across multiple language attribute dimensions, including lexical and morphosyntactic information as well as speech rate?

We hypothesized that child-directed speech provides simpler models of language for children with CIs than adult-directed speech. We expected to observe that child-directed speech would have lower lexical diversity, lower morphosyntactic complexity, and lower speech rate.

2.2. Methods

2.2.1. Participants

This research project was part of a larger ongoing research project involving 14 families with a child with a CI who enrolled in a longitudinal study conducted by the Ohio State University Wexner Medical Center Department of Otolaryngology. The audio data was shared with Michigan State University for studying the early linguistic environments of these children. As part of their participation, families agreed to record their home environments using LENA devices for at least one day within the first year following activation of their child's CI(s). This work analyzes these audio recordings to understand how individual children are different in terms of the quality and quantity of the experienced environments, and how these variations are eventually related to differences in their language outcomes. Recordings were collected at approximately 3 months, 6 months, 9 months, or 12 months after CIs activation. The recordings for each child were not fully available at all four time-intervals due to attrition in participants, and also some families agreed to record their home environments for only one day (see Table 2-1). Therefore, the descriptive information in this section is derived from the available recordings.

Children who participated in this study were 14 prelingually deaf children [4 boys (~29%) and 10 girls (~71%)] with an average of 3.57 months ($SD = 0.26$), 6.33 months ($SD = 0.3$), 8.74 months ($SD = 0.19$), and 12 months ($SD = 0.58$) post-implantation hearing experience at the 3 month, 6 month, 9 month, and 12 month post-CI activation recording intervals, respectively. Table 2-1 presents children's mean chronological ages and associated standard deviations at each of these four time-intervals of recording. Table 2-2 shows mean hearing ages and associated standard deviations. Children's age at CI activation ranged from 8.12 months to 22.57 months ($M_{age} = 14.97$ months, $SD_{age} = 4.1$ months).

Of these 14 children, 11 children were bilaterally implanted (~76%) and three had a hearing aid in one ear and a CI in the other (~24%). Due to the small, heterogeneous clinical sample, some children had additional comorbid diagnoses (N = 4). Measures of children's preoperative hearing status (preimplantation residual hearing) and their CI device information are presented in Table 2-3. All infants were living in monolingual English language environments. Children's socioeconomic status (SES) was indexed by their parents' education level. Additional demographic and audiological information such as children's SES and their CI device characteristics can be found in Tables 2-1 to 2-4.

Table 2-1. Demographic information for 14 children with CIs who participated in this study. The information in this table includes children's ID, gender, age at CI activation, number of analyzed recordings, and children's chronological ages.

Subject ID Code	Subject ID	Gender	Age at CI Activation	Number of Analyzed Recordings	Chronological Age			
					3m	6m	9m	12m
BT0001	1	F	22.57	1			31.98	
BT0004	2	M	17.16	4	20.29	23.15	26.04	29.61
BT0005	3	F	19.77	2	22.55	25.87		
BT0010	4	F	14.89	3	18.45	21.00	22.68	
BT0032	5	F	19.27	4	22.83	25.12	28.09	31.07
BT0071	6	F	18.30	2	22.20	24.75		
BT0141	7	M	13.58	2	17.23	20.29		
BT0155	8	M	14.17	3*		20.94*		26.59
BT0179	9	F	11.02	2*		17.68*		
BT0189	10	F	15.25	2	18.44	21.47		
BT0191	11	F	8.125	2	11.64	14.20		
BT0195	12	M	10.10	2		16.53		21.34
BT0210	13	F	12.40	2*		18.53*		
BT0227	14	F	13.03	1			21.65	
<i>M</i>			14.97	2.29				
(<i>SD</i>)			(4.1)	(0.91)	19.20 (3.7)	20.79 (3.6)	26.08 (4.18)	27.15 (4.3)

Note. All the reported ages are in months. Gender: F= Female, M= Male. * There was more than one recording at the specified time intervals for these participants, and the chronological age is the average chronological age derived from the multiple recordings in that time-interval. No recording was available for the time intervals marked with grey color.

Table 2-2. Demographic information for 14 children with CIs who participated in this study. The information in this table includes children's ID; hearing age at four time-intervals of 3m, 6m, 9m, and 12m; their type of hearing loss, and the laterality of their hearing loss.

Subject ID Code	Subject ID	Hearing Experience Age				Type of Hearing Loss		Bilateral/ Unilateral
		3m	6m	9m	12m	R	L	
BT0001	1			8.54		SE	SE	RL
BT0004	2	3.16	6.02	8.92	12.48	CO+SE	CO+SE	RL
BT0005	3	3.81	6.13			SE	SE	RL
BT0010	4	3.59	6.14	8.82		SE	SE	RL
BT0032	5	3.60	5.89	8.86	11.83	SE	SE	R-HA
BT0071	6	3.93	6.48			SE	SE	RL
BT0141	7	3.68	6.74			SE	SE	RL
BT0155	8		6.81*		12.45	CO+SE	CO+SE	HA-L
BT0179	9		6.69*			CO+SE	CO+SE	R-HA
BT0189	10	3.23	6.25			CO+SE	CO+SE	RL
BT0191	11	3.55	6.11			UN	UN	RL
BT0195	12		6.47		11.27	UN	UN	RL
BT0210	13		6.16*			CO+SE	CO+SE	RL
BT0227	14			8.65		UN	UN	RL
<i>M</i>		3.57	6.32	8.76	12.01			
<i>(SD)</i>		(0.26)	(0.3)	(0.16)	(0.58)			

Note. All the reported ages are in months. Hearing Experience Age is the number of months since CI activation for which the child had been receiving auditory input.

* More than one recording at the specified time interval was available for these participants which were a few days apart; in these cases, the listed chronological age is the mean chronological age derived from the two recordings. No recording was available for the time intervals marked with grey color. Type of Hearing Loss: SE = Sensorineural, CO = Congenital, UN = Unknown. Bilateral/Unilateral: RL: bilateral implants, HA-L: CI in left ear and hearing aid in right ear, R-HA: CI in the right ear and hearing aid in the left ear.

Table 2-3. Demographic information for 14 children with cochlear implants participated in this study. The information in this table includes children's ID, their hearing age at four time-interval of 3m, 6m, 9m, and 12m, their type of hearing loss, and the laterality of their hearing loss.

Subject ID	Subject ID	Mean PTA unaided (dB)	Com mode	Degree of HL
BT0001	1	86.25	OC	Profound
BT0004	2	120.00	TC	Profound
BT0005	3	120.00	TC	Profound
BT0010	4	120.00	TC	Profound
BT0032	5	92.50	TC	Profound
BT0071	6	119.50	TC	Profound
BT0141	7	113.75	OC	Profound
BT0155	8	110.00	TC	Profound
BT0179	9	109.37	OC	Profound
BT0189	10	70.56	TC	Severe
BT0191	11	Unknown	TC	Profound
BT0195	12	120.00	TC	Profound
BT0210	13	120.00	TC	Profound
BT0227	14	120.00	TC	Profound
<i>M</i>				
<i>(SD)</i>				
<i>Note.</i> PTA is pure-tone average before implantation (across the frequencies of 250, 500, 1000, 2000, and 4000 Hz); Com Mode is the type of communication program that the child was following in speech-language therapy; HL: hearing loss; OC: oral communication (exclusively spoken); TC: total communication (a combination of spoken language and Signed Exact English)				

Table 2-4. Demographic information for 14 children with CIs who participated in this study. The information in this table includes children's ID, the etiology of their hearing loss, their SES (given as maternal education, with HS = high school diploma and BA = bachelor's degree), and any additional diagnoses.

Subject ID	Subject ID	Etiology of hearing loss	SES	Additional Diagnosis
BT0001	1	UN	Graduate	None
BT0004	2	Connexin	HSD	GDD
BT0005	3	Auditory Neuropathy	Graduate	GDD
BT0010	4	Auditory Neuropathy	Less than HS	ANSD; ASD
BT0032	5	EVA	HSD	None
BT0071	6	UN	Elementary	None
BT0141	7	CMV	Associates	cCMV; Global Delay
BT0155	8	Connexin	BA	None
BT0179	9	Connexin	HSD	None
BT0189	10	EVA	BA	None
BT0191	11	UN	Graduate	None
BT0195	12	UN	Associates	None
BT0210	13	Connexin	BA	None
BT0227	14	UN	BA	None
<i>M</i>				
<i>(SD)</i>				
<i>Note.</i> Etiology of hearing loss. EVA: Enlarged vestibular aqueducts, CMV: Cytomegalovirus hearing loss. Additional diagnosis. GDD: Generalized developmental delay, ANSD: Auditory neuropathy spectrum disorder, ASD: Autism spectrum disorder, cCMV: congenital cytomegalovirus.				

2.2.2. Acquisition and selection of audio recordings

This project involved analysis of daylong audio recordings of natural environments for the 14 early-implanted children with CIs described above. Recordings were made using a digital audio recorder, the Language ENvironment Analysis (LENA) device. LENA is a wearable audio recorder that collects children’s daily spoken interactions and automatically classifies stretches of audio as conversational blocks or pauses. LENA’s algorithms classify stretches of audio as *female adult speech*, *male adult speech*, *key child*, *another child*, *overlapping vocalization*, *TV/electronic media*, *noise*, *silence*, or *uncertain*. Seven categories except *silence* are then divided by LENA into “near-field” or “far-field” speech or speech-like vocalizations. The stretches identified as “near-field” are grouped together into units, which are called “conversational blocks”. The remaining speech is labeled as “Pause”. LENA also provides automated measures of adult word counts (AWC) and other vocalization measures such as conversational turns (Gilkerson & Richards, 2009; Oetting, Hartfield, & Pruitt, 2009; Ye Wang et al., 2017).

Overall, 32 day-long LENA audio recordings were collected and analyzed for the families participating in this study across four time points of 3-, 6-, 9-, and 12-month post-implantation (hereafter 3m-, 6m-, 9m-, and 12m post-implantation). At each interval, up to four daylong LENA recordings were made (see Table 2-1). Between one and four recordings were analyzed for each child (*Mean number of recordings* = 2.28, *SD* = 0.91). For 11 children, there were at least two recordings at two separate time-intervals. Two children had LENA recordings from all four time-intervals. We analyzed a maximum of four LENA recordings per child in the first-year post-implantation. For children with fewer than four recordings, we included the maximum number of recordings available for each child (see Table 2-1).

In the parent NIH-funded project, it was planned to use metrics derived automatically from LENA device as indices of the amount of language children with CIs were exposed to. However, LENA was not developed for performing the full kinds of analyses that were targeted in this project. Notably, child-directed speech versus overheard speech was one of the core classification categories in this project, which LENA is not developed to perform. Classification of child-directed speech from overheard speech was very important for the purpose of this project, because its strong positive effect has been widely attested in prior studies (Hoff, 2003, 2006; Weisleder & Fernald, 2013). Therefore, to assess the amount of high-quality child-directed speech input, it was necessary to collect auditory perceptual judgments from human listeners to characterize audio samples. Some studies proposed using a binary decision of child-directed speech versus overheard speech on entire stretches of audio within a conversational block, the reliability of which has not yet been tested and is questionable (Hurtado et al., 2008; Soderstrom & Wittebolle, 2013; Weisleder & Fernald, 2013).

To further evaluate LENA's reliability for this project, a complementary study (which was not part of the present dissertation project) was conducted in our lab leading up to the present dissertation that assessed LENA's reliability. This study revealed that LENA's automatic classification and word count measures were insufficiently accurate to assess individual differences in a language input across children in a small sample (Lehet, Arjmandi, Dilley, & Houston, under review). Classification accuracy for speech of any type and adult speech varied widely, from 53% to 86% across families for any speech (corresponding to 14% to 47% false negative/error rates) and from 45% to 82% for adult speech (corresponding to a range between 18% to 55% false negative/error rates) (Lehet, Arjmandi, Houston, & Dilley, under review). It was also shown that accuracy of LENA's count of adult words was significantly affected by the talker's

gender (male or female) and speaking style (child-directed speech or overheard speech). LENA's inaccurate and widely variable classification of adult speech from other types of speech and child-directed speech from overheard speech solidified the decision not to rely on LENA for studies of this research project. Overall, these findings showed that LENA's classification was not reliable enough to be used for purposes of characterizing children's linguistic environments, particularly when focusing on individual differences, as in this small sample study (Lehet et al., under review). As a result, an alternative method of assessing the amount and quality of language in children's natural auditory environments was followed, as described in the next section.

2.2.3. Sampling from day-long audio recordings

In response to LENA's limitations, we therefore developed a comprehensive coding system tailored toward characterizing the quality and quantity of early linguistic input in children with CIs. In this study, the basic metric for assessing quantity of language in children's home environments was to determine *per-day estimates* of language input. Such per-day estimates were derived through analyzing chunks of audio sampled randomly from the day-long audio recordings, a common approach that is necessary to deal with the infeasibility of analyzing the entire day-long audio (Hart & Risley, 1995; Shneidman et al., 2013; Weisleder & Fernald, 2013). Accordingly, the amount of speech derived from randomly selected portions of audio recordings can be extended to the entire day to estimate the total amount of speech experienced by children per day (Hart & Risley, 1995; Shneidman et al., 2013; Weisleder & Fernald, 2013). Such per-day estimates of amounts of words have provided valuable knowledge in modeling individual variability in linguistic environments experienced by children in naturalistic settings (Hart & Risley, 1995; Weisleder & Fernald, 2013).

Recordings in this study were up to 16 hours long, which made it prohibitive for humans to listen and analyze the entire day-long audio. We therefore used random sampling as the most representative sampling method (Hart & Risley, 1995; Weisleder & Fernald, 2013). From within this set of day-long audio recordings, each with a length of around 16 hours, we randomly sampled intervals constituting 5% of the waking time of the child. To accomplish this sampling in a computationally tractable way, the entire day-long recording was first split into 30-second intervals as a preliminary to random selection. Next, those 30-second intervals that included any time during which the child was judged to be asleep were removed. The time intervals where the child was not awake were excluded by hand from the analysis based on contextual cues such as heavy breathing, parents discussing the child sleeping, etc. From the remaining set, 5% of the 30-second intervals were randomly selected for the analysis.

Figure 2-1 shows the total (Figure 2-1A) and average (Figure 2-1B) amount of audio analyzed per recording for each CI child. Overall, a total of 17.7 hours of audio was analyzed by hand by human listeners for this project in 2,118 30-second samples. On average, 1.3 hours of audio were analyzed per child with a standard deviation of 0.58 hours. The variability in total amount of analyzed speech reflects the variability across children in the number of analyzed recordings and the amount of total speech available for analysis after excluding the audio segments where the child was asleep. As Table 2-1 presents, from 14 children, two had only one recording (~14%), eight had two recordings (~57%), two had three recordings (14%), and two had four recordings (~14%). The average total amount of day-long audio for each child after excluding sleeping time also varied between children, with the minimum amount being 6.5 hours and the maximum being 13.8 hours ($SD = 2.1$ hours). The average original length of day-long audio recordings for each child was also variable across children ($min = 8.4$ hours, $max = 16$ hours, SD

= 2.2 hours). This was because 4 of 14 families turned off the LENA recorder before it reached its full-day 16-hour recording buffer. In summary, there were three sources of variability affecting amount of audio analyzed: the number of day-long recordings available per child, variability across children in the amount of sleep time, and variability across children in the length of the original day-long recordings. The chosen method for calculating per-day estimates of language input took account of all this variability.

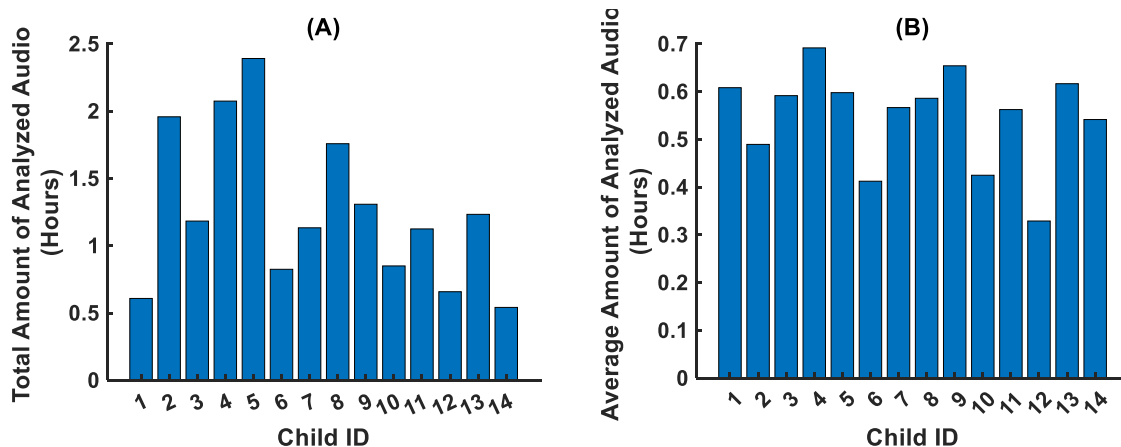


Figure 2-1. (A) Total and (B) average amount of audio analyzed for each child with CIs participated in this study.

2.2.4. Analysis of audio by human analysts

Comprehensive coding system for characterizing early linguistic environments

A comprehensive coding system was developed for this study to perform an accurate and informed analysis of auditory environments of the 14 children with CIs in this study. Our coding system was developed to prioritize coding portions of audio where there was an informative and useful speech-language model generated by a competent talker. The goal was to use this coding system to characterize early linguistic environments of each child in terms of total number of words, number of child-directed words (Hart & Risley, 1995; Montag, Jones, & Smith, 2018; Weisleder & Fernald, 2013), lexical diversity (Carroll, 1938; Montag et al., 2018), and morphosyntactic information (Hoff, 2003). Additionally, our coding system was designed to

permit human coders to analyze speech and background noise experienced by each child with CIs in his/her auditory environment.

Below, we describe the basic elements of the coding system for talkers and the words they spoke in each child's home environment. Using this comprehensive coding system, we were able to characterize the auditory information experienced by the child wearing the LENA device, which we frequently call the *target child* hereafter to distinguish him/her from other children in the environment. Additional details about the coding system focused on identifying background noise and masking and reverberation effects are discussed in Ch. 3, and readers are referred to that chapter for more details.

Coding linguistic input experienced by the target child by human analysts

The general sequence of actions for coding is shown in Figure 2-2. The audio samples were labeled by trained human analysts using *Praat* (Boersma & Weenink, 2001) textgrids using annotation conventions described in these sections while these textgrids were time-aligned with their corresponding audio files. The day-long recordings and textgrids were split into 15-minute intervals for coding so that context around each 30-second interval was available to human analysts. Therefore, when human coders analyzed 30-second audio samples they had clearly labeled samples embedded within longer sections of audio that could provide contextual information to inform their coding decisions.

For each 30-second audio interval, human analysts first identified whether there was any human-generated sound, either live or recorded. If so, analysts marked the temporal starts and ends of the human-generated sound on the relevant tier in *Praat* (Boersma & Weenink, 2001), breaking contiguous speech into separate utterances as needed. Coding conventions further considered the gender of talker and whether the sound at the time of LENA recording appeared to have been live

or pre-recorded (such as television) (see Figure 2-3). For stretches judged to contain a human-generated sound, analysts determined whether it was a speech by a “competent talker”, defined as a talker over the age of five who articulated words in an intelligible, audible fashion. For portions of audio judged to be speech by a competent talker, analysts determined the following information, all of which was captured in coding conventions in Praat textgrid annotations: (a) whether the speech was understandable, in which case they transcribed the words within a contiguous stretch of speech into the relevant *Praat* textgrid interval (including utterance with conventionalized sound-meaning mappings such as *whoosh*, *moo*, *choo-choo*, *woo-hoo*, *yeah*, etc.), (b) who spoke the speech (whether adult male, adult female, or child); (c) who the speech was directed to (whether the target child, and/or other children or adults); and (d) whether the speech was “in the clear” or whether it co-occurred with some type of noise or interfering signal. For simplicity, details on coding noise interference will be covered only briefly below. These noise coding conventions will be considered in more detail in Chapter 3, at which point we will consider analyses that utilized noise coding to develop updated measures of numbers of words experienced by each child over the day (See Figure 2-3 in *Praat*). Since our focus was on the auditory input experienced by the target child, any sounds judged to be from the target child were noted separately and coded in the appropriate tier (See Figure 2-3 and Table 2-5 and Table in Appendix I).

Human analysts followed these structures closely for inserting the labels after they finalized their decisions. They used these structures when coding audio intervals on eight tiers to provide multiple types of information with the designated codes (see Table 2-5 and Appendix I). To increase the accuracy of the coding, human analysts were instructed to actively incorporate information from three sources. These sources were their perceptual auditory impression, the

waveform, and the spectrogram. Human analysts actively cross-checked information across these three sources to assist making better coding decisions.

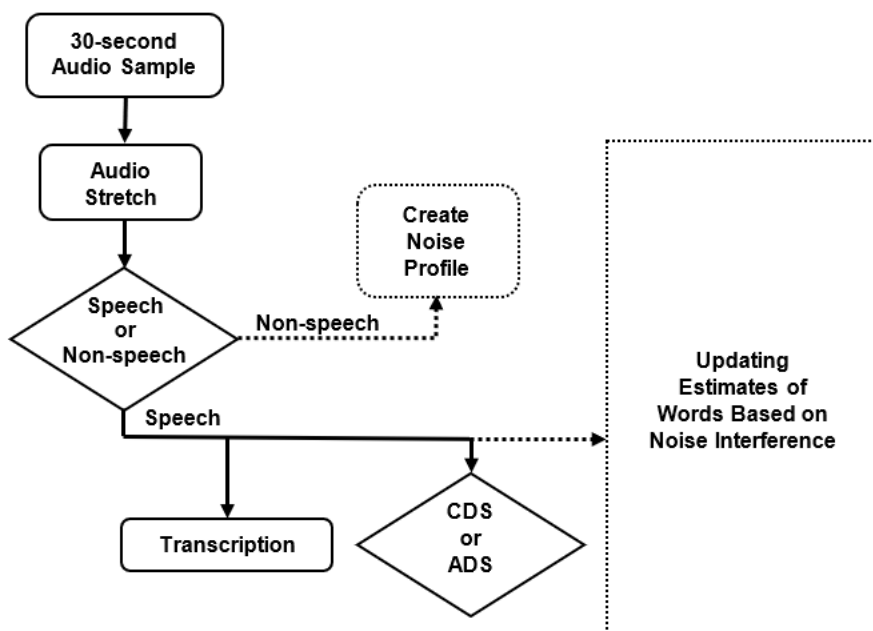


Figure 2-2. Block diagram of the general approach used by human analysts to code sampled audio files. Analyses presented in this chapter are based on coding steps shown by solid lines and solid boxes. The dashed lines and dashed-line boxes indicate steps associated with coding procedures for considering effects of noise and reverberation, which were developed for analyses of Chapter 3 and are discussed in that chapter in more detail.

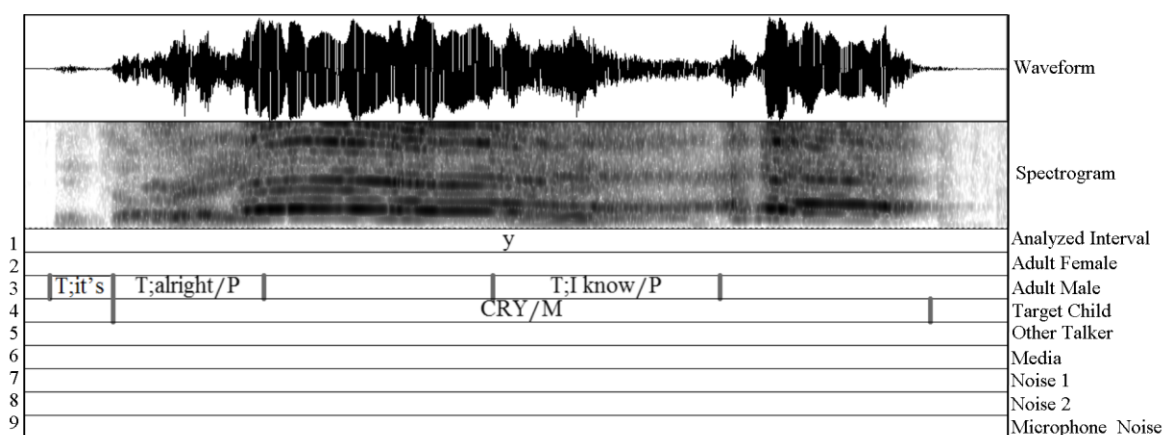


Figure 2-3. Annotation scheme used by human analysts to code various sound events of interest in this study. The illustrated sample interval contains ~3.5 seconds of audio within a 30-second analysis interval drawn from a day-long audio recording of a child in the corpus, and was coded by a human analyst. For this interval, the CI child's father was communicating with his child. The two top rows of the display show the waveform and spectrogram, respectively. Coding consisted of *Praat Textgrid* tiers providing for annotation of the following information (top to bottom): (1)

Figure 2-3. (cont'd) the Analyzed Interval tier indicated which 30-second sampled audio portions had been selected randomly for inclusion in the analysis (given with a 'y'); (2) The Adult Female tier contained speech from competent female talkers (or the primary female for the conversational situation, who was usually the target child's mother but necessarily always the mother, because our criteria specified that the primary talker was the one who held the floor); (3) The Adult Male tier was the same as Adult Female tier, except it was used to designate speech from male talkers; (4) The Target Child tier was used for the "target child" only and contained speech or speech-like or non-speech vocalization or cries from the target child; (5) The Other Talker tier contained vocal activities from other linguistically competent talkers (see text); (6) The Media tier contained any pre-recorded media-related content such as television, toys, tablets, etc.; (7) & (8) The Noise 1 and Noise 2 tiers contained noise sounds; (9) The Microphone Noise tier coded for microphone noise in the background which was detected. See Table 2-5 and Appendix A for further information on the structure and types of codes used in these nine tiers. Here, the father's conversation with the target child is coded within three intervals, two of which overlapped with a child's cry with medium interference determined by the "/M" code after "CRY". "T" label at Level 1 of these three coded intervals indicates that the speech was directed to the target child; see Table 2-0-5. The father's speech was transcribed for these three intervals while separated from Level 1 and Level 3 information using a ";" and "/", respectively. Further, an episode of the target child's crying is coded as partially overlapped with the father's speech with a level of interference judged to be medium (M) relative to his speech. See Ch. 3 for more details on the coding of noise.

Table 2-5 depicts the hierarchical approach to coding information in the current system. In Table 2-5, Levels 1, 2, 3 and 4 depict successive levels of detail specified for coded audio. For instance, an audio stretch determined to contain *Adult Female* or *Adult Male* speech consisting of intelligible words (including lexicalized utterances with conventionalized sound-meaning mappings like *choo-choo*) utilized further annotation codes on Tier 1 or 2, respectively, as specified in Table 2-5 and the Table 1 in Appendix A and the coding description in Appendix C.

If a Level 1 code was designated, then further detail about the vocalization was specified in higher levels of coding (e.g., Level 1, 2, 3, etc.) (See Table 2-5 and Appendix A). For example, when human analysts found an audio interval with speech, they demarcated the start and end of the speech and identified whether the talker was male or female. Then, depending on the gender of the talker and overall number of talkers, they used either Tier 1, 2, or 4 to insert various types of information, including who the speech was directed (see Appendix A for full list of code for addressee) corresponding to increasing levels of hierarchical detail (e.g., Level 1 of Tier 1 and 2

or Level 2 of Tier 4). If the speech was intelligible human analysts used Level 2 on tier 1 or 2 or Level 3 on tier 4 to insert the transcription. A similar procedure was followed for other types of information such as vocalization from target child and sound or speech from media. The codes for a given coded time interval that represented different hierarchical levels of information on a given tier were separated with a semi-colon (“;”); further, coding the level of *noise and reverberation interference* (i.e., **Low**, **Medium**, **High**) was distinguished by adding a slash separator (“/”). (See Chapter 3 for more discussion of coding and analyses related to noise and reverberation.) These separators later were used to extract the information at different coding levels on each tier.

Table 2-5. Hierarchical structure of information coding in the system across 8 *Praat* tiers. Columns labeled Level 1 to Level 4 present information analysts coded at each level and correspond to successive levels of detail used later to classify intelligible speech utterances. The details of each information on each tier and level are presented in “reference table of codes” in Table 1 in Appendix A. SNG = singing; NSV = non-speech vocalization. See Chapter 3 for more discussion of coding of degree of interference from noise and reverberation (cf. codes P = primary talker; L = low interference; M = medium interference; H = high interference).

<i>Tier ID</i>	<i>Tier</i>	<i>Level 1</i>	<i>Level 2</i>	<i>Level 3</i>	<i>Level 4</i>
1&2	Adult Female & Adult Male	Addressee code	Lexical transcription (for intelligible content only) or designation as SNG or NSV	“P” and/or Noise interference {L, M, H}	----
3	Target Child	Type of vocalization (speech/speech-like vocalization, cry, or NSV, etc.)	Lexical transcription (for S only)	Noise interference {L, M, H}	----
4	Other Talker	Codes indicating talker gender and age (adult or child)	Addressee	Lexical transcription (for intelligible content only) or designation as SNG or NSV	“P” and/or Noise interference {L, M, H}
5	Media	Type	Purpose	Transcription	Noise interference {L, M, H}
6&7	Noise 1 & 2	Type	Noise interference {L, M, H}	----	----
8	Microphone Noise	Type	Noise interference {L, M, H}	----	----

2.2.5. Coding training and procedure

Thirteen human analysts who were all undergraduate research assistants from related majors (e.g., communicative sciences and disorders or linguistics) received extensive training to learn and implement the designed coding for this study. The overall process of training and coding took around 8 months to complete coding the selected audio for this study. We conducted an extensive training regimen to reduce variability in human analysts' auditory perceptual judgments for various categories of interest in this study (Hallgren, 2012; McHugh, 2012). The instruction procedure involved lecture-style instruction in a group where coders guided to implement the coding on sample files with multiple levels of coding difficulties. Follow-up individualized meetings were regularly scheduled to keep track of analysts' coding proficiency and accuracy. The training procedure included (a) instructing human coders to effectively use *Praat* software for analyzing and coding the sampled audios; (b) providing brief background of the purpose of the project and how the developed coding system is related to the aims of this project; (c) instructing human analysts to read carefully through the coding document and the relevant addendum; (d) quizzing them on the coding system to make sure that they had a good understanding of the rules of coding; (e) assigning them to code a series of practice files with various levels of coding difficulties to make them familiar with the complexity of auditory environments from easy to very hard scenario; (f) evaluating coding of human analysts on practice files and providing them with individualized feedback based on their coding, with instructions to use the provided feedback to refine their coding for the practice files; and (e) performing a final evaluation to decide whether a human coder was subjectively qualified to start coding the actual audio files. After assuring that a coder had acquired the requisite level of proficiency to start coding the actual files, the trained coders were assigned to analyze the randomly selected 30-second audio chunks for each LENA

recording in *Praat* while following closely the instructions given in the developed language environment coding document.

Error checking procedure. The coded textgrids by human analysts were checked for various types of minor errors during the course of coding procedure. From the pool of the files coded by each human analyst at the middle of coding procedure, ten files were randomly selected for each of thirteen analysts and assigned to a new analyst for evaluation. The results of this coding evaluation were that the new analyst generated instructive feedback for the original analysts to recheck their coding for prior files, to correct them, and to enhance their quality for the remaining files. Several *Matlab* scripts (Bořil & Skarnitzl, 2016) were developed to automatically check the accuracy of labels on all textgrids on nine tiers and at different levels as illustrated in Table 2-5. For example, a label of “AF;T;yeah” on the Adult Female tier would be detected by this error-checking process for having incorrect label at Level 1 of this tier (AF is not in the set of defined codes for tier *Adult Female* and the nature of error was probably in this case due to choosing an incorrect tier to insert the label.). These minor errors were in turn passed to analysts for correction. This process also included checking for the separators of “;” and “/” to make sure that they represented the correct coding format. We reran the scripts for these types of errors multiple times to assure their accuracy, since these labels were later used to capture various measures of linguistic input.

2.2.6. Inter-rater reliability analysis

An inter-rater reliability analysis was conducted to assure acceptable agreement between human analysts in analyzing and coding the LENA audio samples. Multiple inter-rater reliability analyses were performed from the same common set of data coded by all analysts for these

analyses to examine the degree of agreement between coders in coding audio samples for different variables of interest.

To conduct the inter-rater reliability analysis, the 13 human analysts were asked to code 2% of the total 5% of randomly selected audio samples. Overall, about 0.36 hours (21.5 minutes) of audio from the total of 17.7 hours was coded by analysts for the inter-rater reliability analysis. The audio selected for the inter-rater reliability analysis had not been previously seen by coders in this study, and analysts coded these files independently of one another. The design of the coding study was a fully-crossed design where all randomly selected audio was coded by all thirteen coders (Hallgren, 2012).

A frame-based inter-rater reliability analysis was conducted to assess the agreement between pairs of coders based on *Cohens' Kappa* analysis as well as among coders for *Fleiss Kappa* analysis at the level of frames with 100 ms length. From these frames, inter-rater reliability was calculated based on two measures of reliability of *Cohen's kappa index* and *Fleiss' Kappa*. These metrics have been shown to be the most reliable ways to assess agreement among human coders for different categories of interest (Hallgren, 2012; McHugh, 2012). Since more than two human coders were involved in this project, the *Cohen's kappa* statistics were estimated by calculating the average of kappa for all coder-pairs (Light, 1971). The final estimates of *Cohen's kappa* were derived by averaging across coder pairs for each of the categories of interest. The primary categories for this study were “*speech*” vs. “*non-speech*”, “*child-directed speech*” vs. “*adult-directed speech*”, “*clean vs. overlapped speech*”, and “*level of interference*”. We calculated the *Cohen's kappa* for other categories as well for reference. However, the number of levels were large for other categories and high agreement was not expected due to variability and higher resolutions for those categories.

Finally, a word-transcription reliability analysis was also conducted to identify how much coders agreed with each other in transcribing the speech of competent talkers at the level of the word. The analysis involved first identifying regions that coders coded as speech, followed by determining the degree of word similarity between coders' transcriptions. The inter-rater analysis for word-transcription was computed by averaging over the percent of agreement between coders over the intersection of intervals identified as speech.

2.2.7. Analysis of audio samples coded by human analysts

Following application of the coding system to the 5% of randomly sampled 30-second audio stretches by the 13 analysts, we analyzed speech intervals and their transcriptions in each 30-second audio sample to capture the following linguistic measures: (1) *total number of words per-day*, (2) *type-token ratio (TTR)*, (3) *area under the type-token curve (AUTTC)*, (4) *mean length of utterance (MLU)*, (5) *caregiver responsiveness*, and (6) *speech rate*. A discussion of how these six measures were calculated is presented below. All these language measures (except *caregiver responsiveness*) were also calculated separately for subsets of intelligible speech consisting of either (a) child-directed speech, (b) overheard speech, and/or (c) adult-directed speech.

Total word count per day. Total word count per day is a measure of the quantity of words that each child was exposed to in a day. To calculate this measure, the stretches containing live speech were first identified using their labels depending on their tier and level of information (see 2-5 and Appendix A for the structure and details of codes on each tier and at each level of information.). In cases where speech from two or more talkers overlapped, the speech utterance from the talker designated *primary* in the conversation was identified and analyzed as the basis of counting the number of words. The speech stretches that overlapped with other sources of sound or/and speech and were split into smaller stretches due to variable level of noise over that speech

were merged and treated as one utterance as the focus of this study was on all available speech utterances without considering the effect of noise to update number of words. Utterances produced by media were excluded, since analysis of media was beyond the focus of this study. Further, prior studies showed that speech from media do not significantly contribute to language learning in children, particularly in early childhood (Ambrose, VanDam, & Moeller, 2014; Kuhl, 2007).

Following the above approach, the number of words for all 5% randomly sampled 30-second audio chunks were summed. The space delimiters between words in utterances were used in our automatic analysis of textgrids to separate words in transcriptions and count them. Contracted words were counted as one single word. This estimate was then extrapolated over a full 16-hour day based on the duration of 5% audio sampled from each day-long recording. For example, for a daylong recording with N words within its 5% analyzed audio with length of L minutes, the estimated number of words was $(N/L)*60 \text{ minutes} * 16 \text{ hours}$. These analyses were repeated for data subsets (e.g., child-directed speech, overheard speech, and adult-directed speech).

To estimate cumulative exposures experienced by each child over multiple years after CI activation, estimates derived as above were used to calculate estimates of N -year post-implantation exposures (where N is the number of years after implantation). These were calculated by using the per-day estimate (i.e., the average estimate for each child based on a 16-hour day derived as above), and multiplying by 365 days, times N years. For example, if the estimated total number of words per day for a child was 2,000 words, then the number of words experienced by the child after three years of implantation was estimated as $2,000 \times 365 \times 3 = 2,190,000$ words.

Type-token ratio (TTR). TTR is a measure of lexical richness in language. This traditional measure of language input has been commonly used in studies of language development to characterize the redundancy of speech input (Broen, 1972; Cheskin, 1981; Montag et al., 2018).

To calculate this ratio measure, word types determined across all 5% selected audio stretches, and then divided by total word tokens determined across all 5% selected audio stretches to calculate the TTR. A child who experienced language input with relatively higher TTR compared to other children would hear more word types, assuming that the number of word tokens in input was equal across children.

For each recording, a measure of TTR was calculated based on the word transcriptions. Prior studies showed that TTR is not able to properly reflect the lexical richness of speech input across joint dimensions of variability in word types and word tokens (MacWhinney, 2017; Montag et al., 2018) because its value is sensitive to both diversity of word types and the total number of word tokens. Therefore, differences among children in the amount of TTR in their experienced speech is not necessarily translatable to lexical diversity, due to the possibility of exposure to variable numbers of words in their environments in a day. For example, if children differ from each other in the number of words they experienced per day while the word diversity in speech input is the same, those with a smaller number of words per day would be judged to experience a higher TTR. This would imply exposure to more diverse speech, which does not necessarily reflect individual differences in exposure to diverse words. In fact, these children experienced fewer unique word types with fewer repetitions of words compared to those who experienced an environment with a higher number of word tokens. To deal with this caveat in using TTR for characterizing lexical richness of language input, we used *type-token curves* to model the variation of input along both dimension of number of word tokens, and the number of words types (Montag et al., 2018), discussed in the next section.

Type-token curve. We calculated the *type-token curve* as another measure of lexical richness (Montag et al., 2018). Type-token curves are an innovation of Montag et al. (2018), who

identified problems with the traditional TTR measure; Montag et al. suggested that *type-token curves* provide a better way of evaluating the richness of lexical input experienced by each child by quantifying how the number of *word types* and *tokens* grow together for an individual child. Following their approach, we analyzed the transcriptions of adults' speech utterances within the 5% sampled audio for each day-long recording to construct a type-token curve for each day/recording, following the approach discussed in (Montag et al., 2018).

To compute type-token curves, the complete set of morphemes was first extracted from transcriptions of available competent talkers' speech utterances. Separator characters (including the space between words and the "+" sign for separating morphemes in words coded in the transcriptions) were used as markers to identify words in utterances and separate them into morphemes as the smallest meaningful morphological units. Note that any audio intervals labeled with "X", indicating words in transcriptions that were unintelligible to analysts, were excluded from this analysis. The pairs of word types and word tokens for constructing the type-token curve were then derived through a sequence of random selections of words from the entire set of words of a recording, followed by calculating word types for that set of randomly selected word tokens. We used morphemes instead of words to give us more resolution in exploring distributions in these smaller meaningful units of language. See further details of type-token curve determination in Appendix B.

Area under the type-token curve (AUTTC): *Area under the type-token curve (AUTTC)* is a summary measure that we propose in this study to appropriately and simultaneously evaluate language input for the degree of lexical diversity experienced by each child. We suggest that AUTTC is able to reflect both the amount of speech experienced by each child per day, as well as the diversity of these words as an appropriate, single summary measure. In fact, a child who

experiences speech input with a relatively higher AUTTC value per day is exposed to a greater number of words, as well as higher word diversity compared to other children.

To calculate this measure, we first constructed the type-token curve for each recording based on the approach discussed above. Then, we calculated the AUTTC for each day-long recording as a summary measure of lexical richness of speech input. Finally, the lexical richness of linguistic environments of each child was summarized by averaging AUTTC values over the recordings available for each child.

Mean length of utterance (MLU). Mean length of utterance (MLU) in morphemes of adults' speech (Brown, 1973) has been calculated as a traditional index of morphosyntactic complexity of speech input (DeThorne, Johnson, & Loeb, 2005; Parker & Duluth, 2005; Szagun & Stumper, 2012). This measure reflects the average length of utterances in speech experienced by children in a day in their linguistic environments. Experiencing speech input with a relatively larger MLU indicates children's exposure to longer utterances (Bornstein, Haynes, & Painter, 1998; Furrow, Nelson, & Benedict, 1979; Hoff & Naigles, 2002). Since increases in MLU can be due to either more words or more bound morphemes, MLUs more than 3.0 are advised to be interpreted cautiously (DesJardin & Eisenberg, 2007; Eisenberg, Fersko, & Lundgren, 2001).

We analyzed all speech utterances available for each recording to evaluate the structure of language input in terms of MLU with respect to the number of words spoken in each utterance. We calculated this measure for total speech, child-directed speech, and overheard speech from the day-long recordings. To calculate this measure for each recording, the number of morphemes for a recording was counted from all utterances available in that recording. This count was then divided to the total number of utterances to calculate the MLU of speech experienced by child in his/her environments (Parker & Duluth, 2005).

Caregiver responsiveness. Caregiver responsiveness was calculated by counting the instances of speech or speech-like *target child* vocalization (i.e., the child with CIs) which were responded to by an adult producing child-directed speech within 3 seconds after initiation of child's speech or speech-like vocalization. Note that this operational definition implies an assumption that the child-directed speech provided a relevant response to a child's vocalization. We assumed that these two criteria – that the response occurred within this 3-second interval and identifying child-directed speech as a response to child's vocalization – satisfies two main features of *contiguity* and *contingency* of responsiveness (Tamis-LeMonda et al., 2014). Since this exploratory approach has not yet been validated, the results need to be interpreted cautiously.

To calculate this metric of caregiver responsiveness, we first identified the intervals containing speech or speech-like vocalization from the target child and searched for child-directed speech produced by an adult within 3 seconds after initiation of the target child's speech or speech like vocalization. This count was then divided by the total number of child vocalizations to give the estimated metric of caregiver responsiveness to the child's verbal activities. This measure was then averaged over recordings for children with more than one recording.

Speech rate. Speech rate was calculated to evaluate the rate by which these children with CIs heard speech in their environments. Speech with a lower rate has been shown to be a strong predictor of speech clarity and intelligibility (Ferguson & Poore, 2010; Hargus Ferguson & Quené, 2014). Speech rate for each recording was calculated as words per second by summing up all the words available in speech utterances and dividing the value to the total duration of the speech utterances for a recording. For children with more than one recording the final speech rate was calculated by averaging speech rate value across available recordings. We acknowledge that this method of rate calculation has a limitation that words per second might not be as accurate as

estimating rate differences as syllables per second. This is because word complexity might differ between adult and child-directed speech; therefore, the words/second measure might underestimate rate differences between adult-directed and child-directed speech as compared with a syllables per second measure.

Software and coding routines in implementing analyses. To implement these analyses, the coded *Praat* textgrids were called and analyzed in *Matlab* software to characterize the quality of linguistic environments based on above linguistic measures. The coded Textgrids, together with the 30-second audio intervals sampled from each daylong recording were analyzed using *Matlab*, the *mPraat toolbox* (Bořil & Skarnitzl, 2016), and *Praat* (Boersma & Weenink, 2001) to derive per-day estimates of each of the above linguistic measures for each child. For children with more than one recording, estimate for each day were averaged together across multiple days (i.e., recordings) to give the final estimate for each child.

2.3. Results

2.3.1. Inter-rater reliability

Good agreement was found for most coding category distinctions; see Table 2-6 for results of the inter-rater reliability for different categories of interest in this study. These results show that human coders demonstrated substantial agreement (0.61–0.80) for most categories that were a focus of this project (Landis & Koch, 1977). The categories with low agreement were not the focus of this study, although the human analysts were instructed to double-check those categories for the coded files and adjust or revise the interval and labels to have coding of acceptable quality.

Additionally, pairwise correlations of word count for speech intervals were summed within each sample and calculated for the pairwise correlations between each rater. This analysis revealed high agreement, with $r = 0.95$ ($SD = 0.01$, range: 0.94-0.97). These results show that the

implemented training procedure was effective in permitting human coders to implement the complex coding system with high reliability for the purpose of characterizing children's early linguistic environments.

Table 2-6. The results of inter-rater reliability analysis for various categories of interest in this study.

<i>Category</i>	<i>Measure of Inter-rater Reliability Analysis</i>	
	Cohen's Kappa	Fleiss' Kappa
<i>Speech vs. Non-speech</i>	0.68	0.69
<i>Addressee (child-directed & adult-directed)</i>	0.88	0.60
<i>Clean vs. Overlapped Speech</i>	0.67	0.61
<i>Level of Interference</i>	0.87	0.2
<i>Type of Adult Vocalization</i>	0.92	0.67
<i>Gender of Talker</i>	0.61	0.61
<i>Presence of Target Child Vocalization</i>	0.65	0.66
<i>Type of Target Child Vocalization</i>	0.75	0.58
<i>Type of Media</i>	0.24	0.47
<i>Purpose of Media</i>	0.21	0.34
<i>Type of Noise</i>	0.80	0.28

2.3.2. Variability across children in the number of words experienced

Total number of words experienced

We first investigated how the number of words in environments of children with CIs differed across children. Based on extrapolations from our audio samples, children heard an average of 25,134 words per day (range: 37,745 words per day), as shown in Figure 2-4A. The first row of Table 2-7 summarizes the dispersion of the total number of words per day experienced across children.

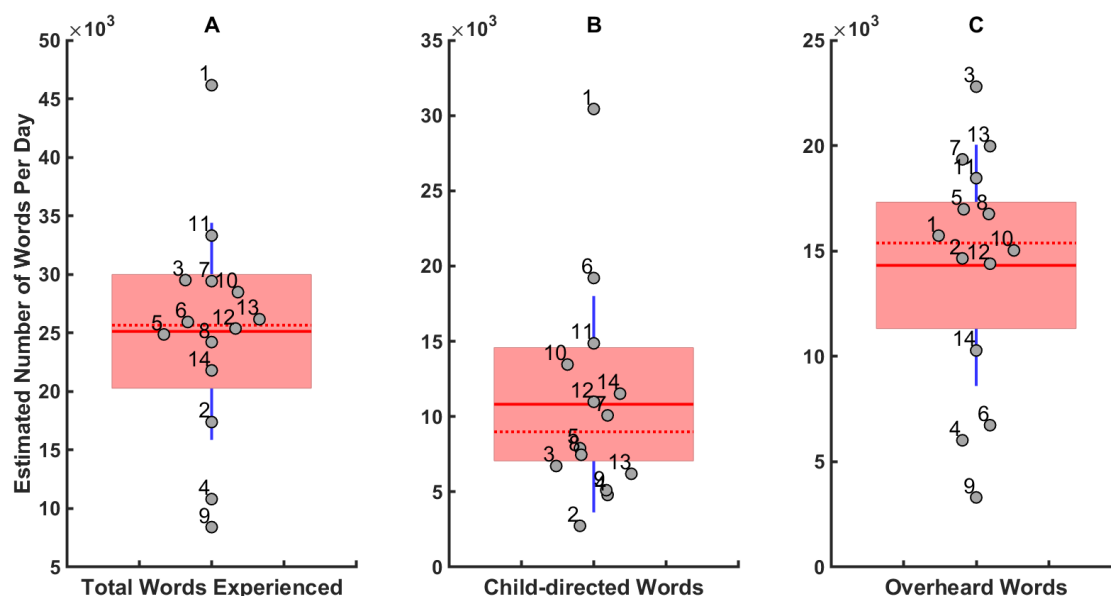


Figure 2-4. Estimated (A) total words per day, (B) child-directed words per day, and (C) overhead words per day for each child in their home environments. In each panel a scatter plot shows the individual children, while the boxplot summarizes the distribution of language experienced in home environments. The data points are laid over a 1.96 standard error of the mean (95% confidence interval) in red and 1 standard deviation shown by blue lines. The solid and dotted red line show the mean and median, respectively.

Table 2-7. Measures of dispersion and central tendency for estimates of number of words per day for each child in his/her linguistic environment derived from total speech, child-directed speech, and overheard speech.

Measure of Language Input (per day)	Measures of Variability						
	Min	Max	Range	Mean	SD	Median	IQR
Number of words	8,414	46,159	37,745	25,134	9,267	25,656	7,623
Number of child-directed words	2,741	30,431	27,690	10,817	7,187	8,986	7,258
Number of overheard words	3,306	22,797	19,491	14,317	5,724	15,377	8,176

Number of words experienced in child-directed speech

Figure 2-4B shows the distribution of estimated exposures to child-directed speech per day across the 14 children with CIs. Children heard an average of 10,817 child-directed words per day, with a difference of 27,690 words per day between the child hearing the least (Child 2) and the most (Child 1) child-directed words. The summary measures of variability in the second row of

Table 2-7 provide further information about how dispersed the children with CIs are with respect to the amount of high-quality child-directed words they experience in a day in their environments.

An interesting observation relates to the change in the rank order of children relative to one another in how much child-directed speech they experience vs. total speech. For example, Child 11 heard more total words than Child 6 (Figure 2-4A). However, Child 11 heard more child-directed words than Child 6 (Figure 2-4B). Finally, Figure 2-4C shows the distribution of overheard words per day across children, which is just the difference between total words and child-directed words; overheard words include any words not judged to be spoken directly to the target child, including words directed to adults or other children. The third row of Table 2-7 summarizes variability and central tendency in overheard words across children.

Figure 2-5 shows the distribution of child-directed vs. overheard words per day for each child. The ratio on the top of each bar is the proportion of child-directed words out of the total number of words experienced by each child per day. This plot suggests that children who experienced a linguistic environment with a greater total number of words per day do not necessarily hear the most child-directed words per day as well. For example, the vast portion of speech in Child 3's environment was overheard speech (Figure 2-5), which changes his/her distance from other children in terms of relative qualities of linguistic environments.

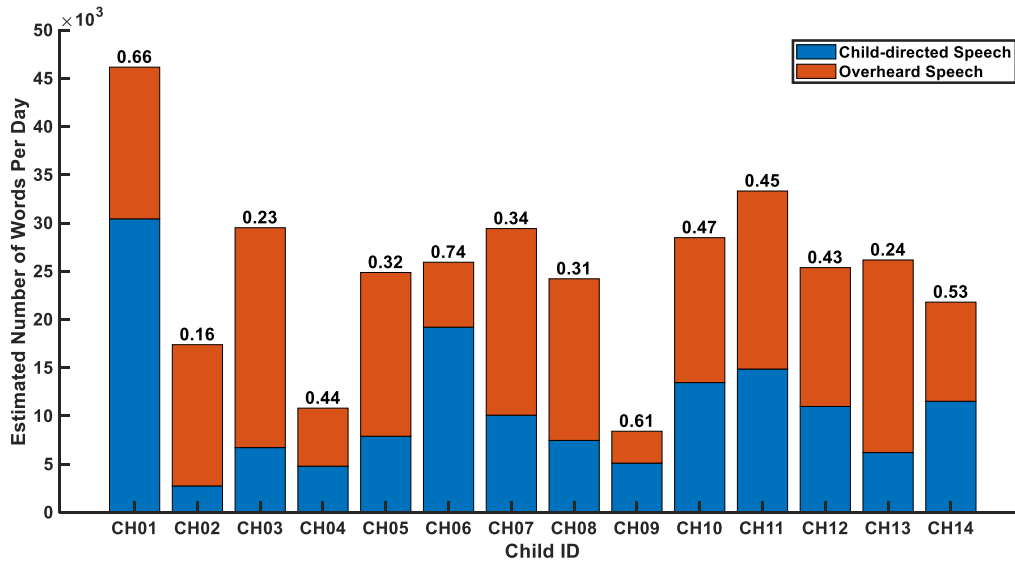


Figure 2-5. The distribution of child-directed and overheard words per day in the language environments of each child with a CI. The ratio on the top of each bar is the proportion of child-directed words out of the total number of words experienced by each child per day.

Number of words experienced over years after CI activation

To estimate cumulative effects of differential word exposures over years, we extrapolated per-day word exposure estimates out to 5 years post-implantation.

Projected word exposures: Total words

Figure 2-6 shows estimated total words experienced by each child over 1, 2, 3, 4, and 5 years after CI activation. The y-axis is shown on a log scale to better present the variation across children and over years after CI implantation. These rough estimates suggest that, over a three-year span, participants are expected to have experienced on average 27.5 million total words. However, estimated total words varied substantially across children (*min* = 9.2 million words, *max* = 50.5 million words; *range* = 41.3 million words; *SD* = 10.1 million words; *IQR* = 8.3 million words). These finding replicates and extends the basic finding of Hart and Risley (1995), demonstrating that children who are the most advantaged in language exposure will have heard in excess of 30 million more words than the children who are least advantaged.

Figure 2-7 presents the same data from another direction and shows how estimates of the total numbers of words vary over years for each child with CIs. Vertical numbers 1-5 on each bar represent estimates of total numbers of words at 1 through 5 years post-implantation. As this figure suggests, the absolute magnitude of individual differences between children will largely expand over years, signaling substantial variability between children in the quantity of speech they are expected to have experienced before kindergarten. This has been shown to be a key factor leading to variability in children's language and scholastic achievement (Golinkoff et al., 2019; Rowe, 2012; Weisleder & Fernald, 2013).

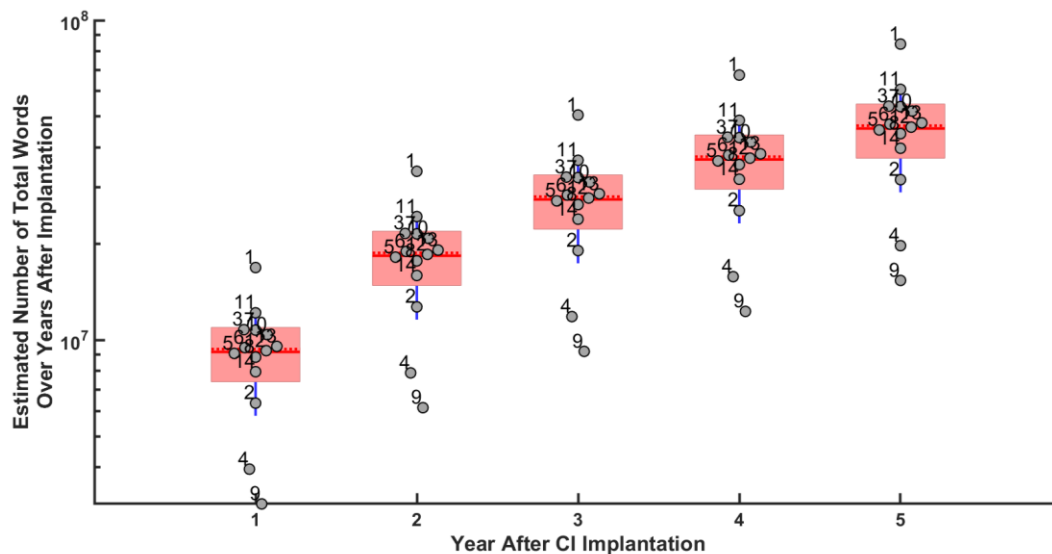


Figure 2-6. Estimated number of total words heard over years by the time of entry to kindergarten for each child. The "y axis" is presented in log scale.

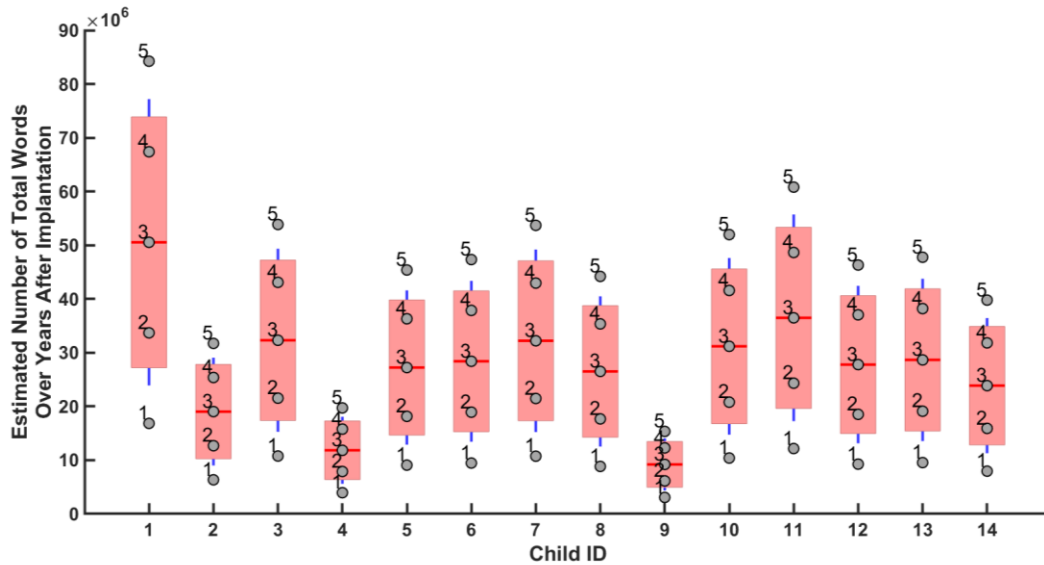


Figure 2-7. Estimates of number of total words experienced over years by each child. Vertical numbers 1 through 5 on each bar represent the estimates for Year 1 through Year 5 post-implantation.

Projected word exposures: Child-directed words

Figure 2-8 and 2-9 show estimated child-directed words experienced by each child over several years after CI activation in two different ways. Over a three-year span, participants are expected to have experienced on average 11.8 million child-directed words (*min* = 3.0 million words, *max* = 33.3 million words; *range* = 30.3 million words; *SD* = 7.9 million words; *IQR* = 7.9 million words). This substantial variability across children suggests that by 3 years after CI activation, a gap of 30.3 million words will separate the child most advantaged in terms of child-directed speech exposure from the child least advantaged.

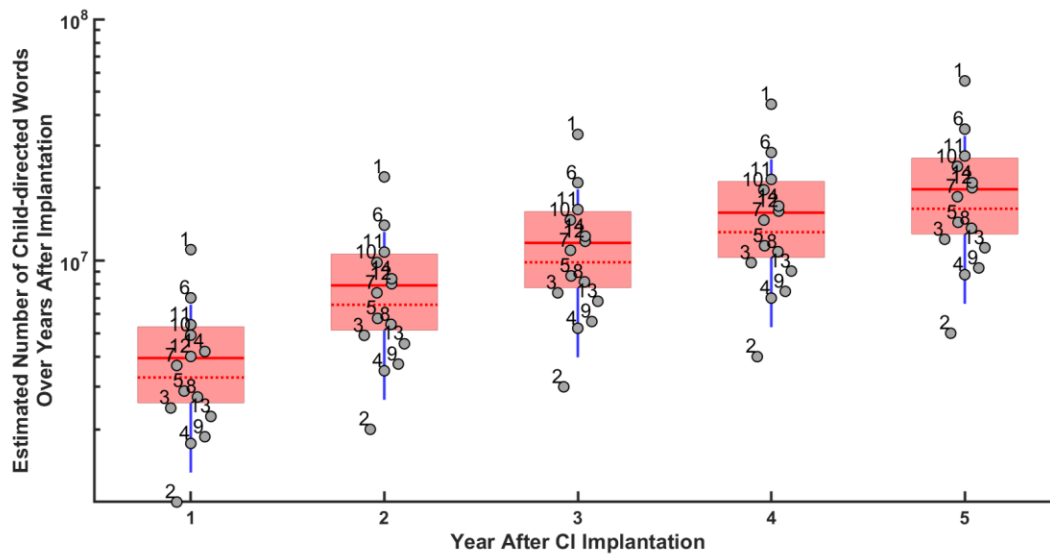


Figure 2-8. Estimated number of child-directed words heard over years by the time of entry to kindergarten for each child. The y-axis is shown on a log scale.

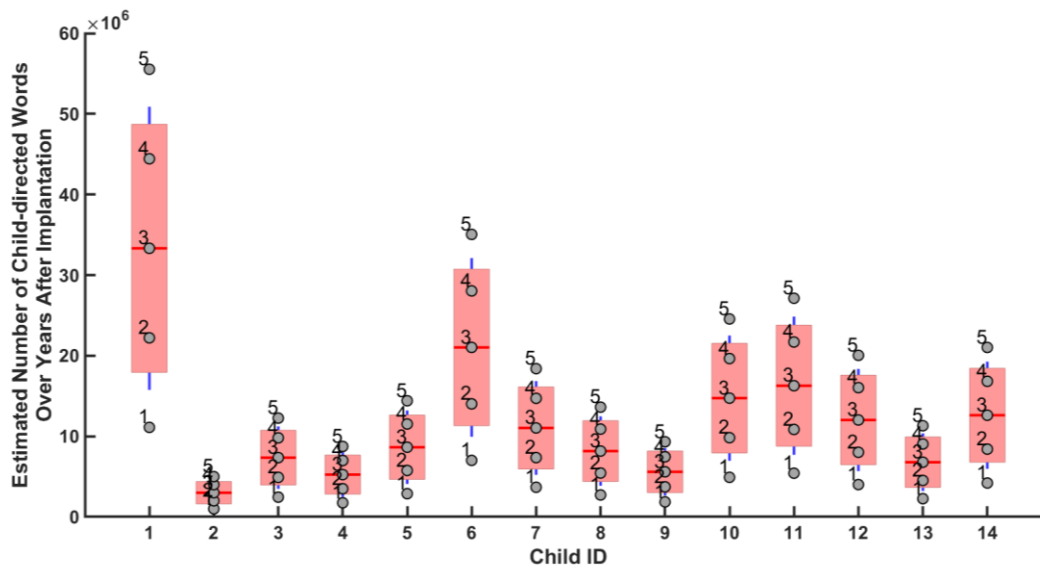


Figure 2-9. Variation in estimates of numbers of child-directed words experienced over years by each child.

2.3.3. Variability across children in lexical richness of their linguistic environment

Lexical richness of experienced speech based on type-token ratio (TTR)

Lexical diversity of speech input, measured as type-token ratio (TTR), is a strong predictor of vocabulary growth in children with normal hearing (Hart & Risley, 1995; Hoff & Naigles, 2002;

Huttenlocher et al., 1991; Pan, Rowe, Singer, & Snow, 2005; Rowe, 2012; Shneidman et al., 2013; Weizman & Snow, 2001). A child who experiences language input with a relatively high type-token ratio will have heard a greater diversity of words, compared with a child experiencing a relatively low type-token ratio, contingent on exposure to the same number of words. Figure 2-10 shows scatterplots of this measure for 14 children with CIs for total, child-directed, and overheard speech.

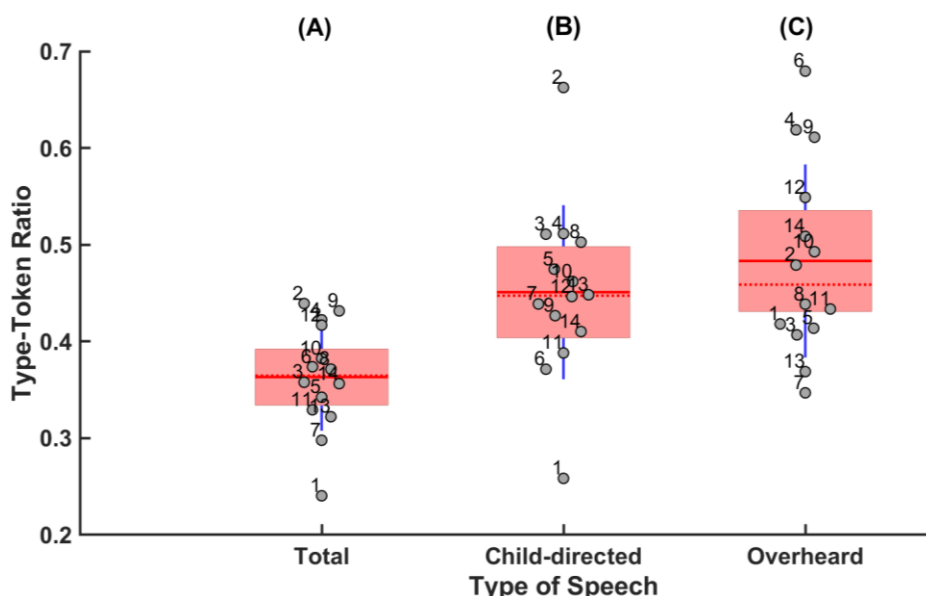


Figure 2-10. Type-token ratio calculated from (A) total, (B) child-directed, and (C) overheard adult speech in the environment of each child with CIs.

Table 2-8. Measures of dispersion and central tendency of type-token ratio derived across children with CIs. These measures are presented for total speech, child-directed speech, and overheard speech.

Measure of Language Input (per day)	Measures of Variability						
	Min	Max	Range	Mean	SD	Median	IQR
Type-token ratio of total speech	0.24	0.44	0.20	0.36	0.05	0.36	0.09
Type-token ratio of child-directed speech	0.26	0.66	0.40	0.45	0.10	0.45	0.09
Type-token ratio of overheard speech	0.35	0.68	0.33	0.48	0.10	0.46	0.13

The plot shows a substantial range of variability of TTR from speech experienced by children in their environments (See also Table 2-8). A linguistic environment with relatively higher TTR value compared to other environments is considered to be a high-quality environment

due to the higher ratio between word types and tokens in speech input, reflecting a greater diversity of words. According to this measure, it appears that Child 1 experienced the lowest quality child-directed input compared to other children, whereas it was previously revealed that this child experienced the best linguistic environment based on both total number of words and number of child-directed words. The TTR measure suggests this child experienced a great deal of repetition in the types of words heard.

However, one weakness of TTR as a measure of children's early linguistic experience is that TTR does not reflect merely the diversity of vocabulary in speech input, but is also sensitive to number of tokens in the analyzed sample, as observed in our data (e.g., TTR for Child 1) (Herdan, 1960; Montag et al., 2018). This means children who heard fewer words per day in their environments have a higher chance of having a higher TTR suggestive of hearing more diverse speech. However, the total number of unique word types for these children was not higher; due to exposure to smaller number of words per day, the repetition of words was lower compared to other children. Comparing Figure 2-10 with Figure 2-4 highlights the point that the TTR measure for child-directed speech shows more or less the reverse ordering of children compared with the number of child-directed words, indicating the weakness of TTR for appropriate representation of lexical richness of input.

Lexical richness of experienced speech based on type-token curves

To overcome the limitation of the TTR measure's dependence on total numbers of word tokens, we also constructed *type-token curves* to characterize different degrees of lexical diversity across children's linguistic environments. Type-token curves relate total word tokens to word types and were constructed through the random sampling process described in Appendix B (Malvern, Richards, Chipere, & Durán, 2004; Montag et al., 2018). As such, they plot how word type

exposures change as a function of number of words randomly selected from the entire words in a day. Therefore, a *type-token curve* that stands higher for the same amount of words indicates a speech input per day with higher word diversity.

Figure 2-11 shows this curve for the recordings of each child at four time-intervals of 3m, 6m, 9m, and 12m after CI implantation whenever a recording was available. The curves derived from total words, child-directed words, and adult-directed words are shown by *blue*, *red*, and *yellow* lines, respectively. As these curves suggest, individual variability across children in terms of diversity of vocabulary they are exposed to is partly a function of the numbers of words the experience in a day. Since these curves are constructed based on two values of number of words and word types, it was not practical and logical to average them over recordings. So, we presented them individually. Later, we will use a summary measure based on the area under this curve to make it possible to represent these two-dimensional data in one summary measure.

Comparing the curves across children highlights the large individual variability in the quality and quantity of children's linguistic environments. A striking observation from these figures is the extent to which the areas under the type-token curve change across children for high-quality child-directed speech. Child 1 is distinguished from other children, such as Child 2, primarily due to the higher number of child-directed words, which also increased the chance of exposure to greater numbers of word types. It is also interesting how language environments of some children vary from day to day, particularly with respect to the lexical richness of high-quality child-directed speech.

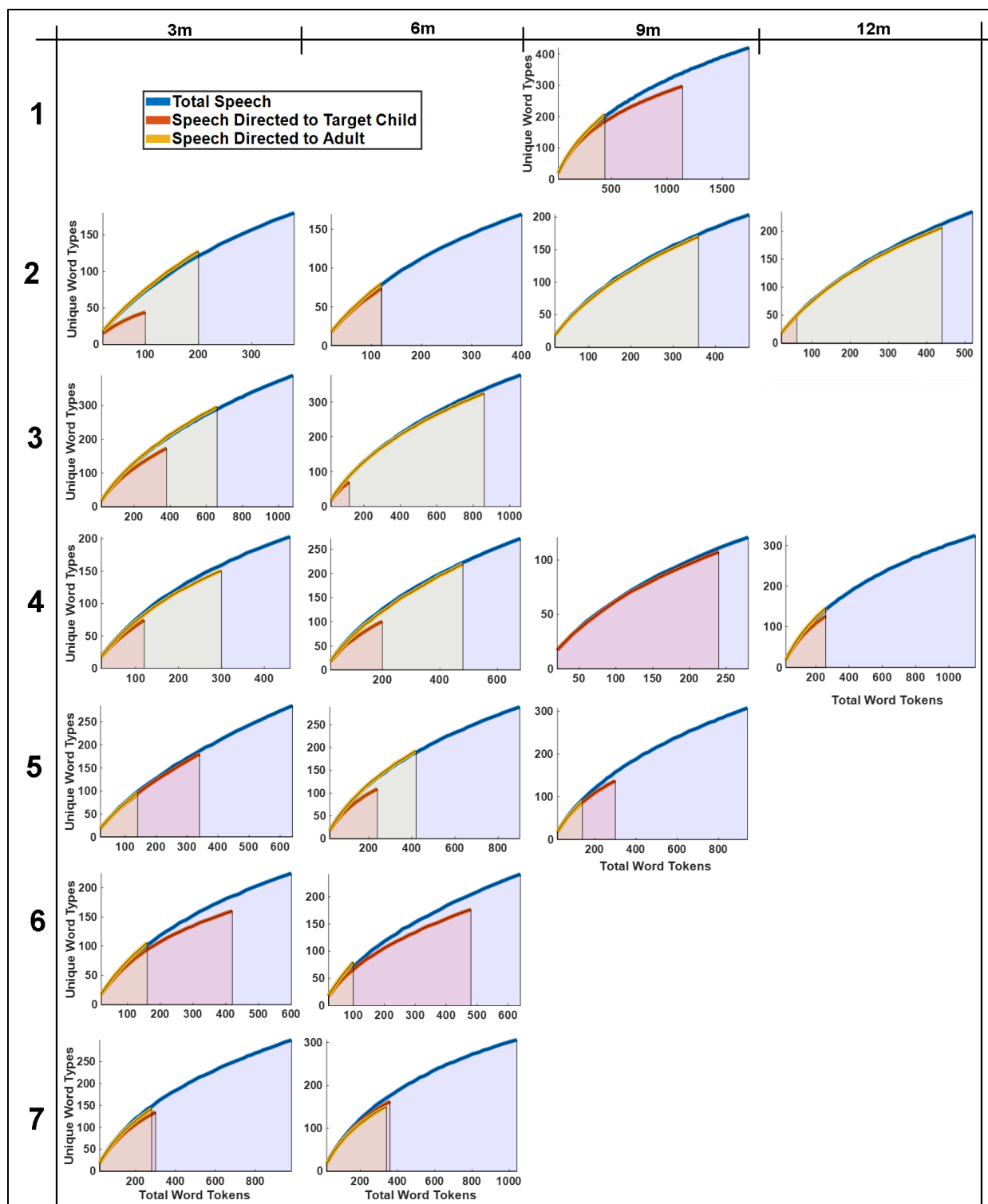
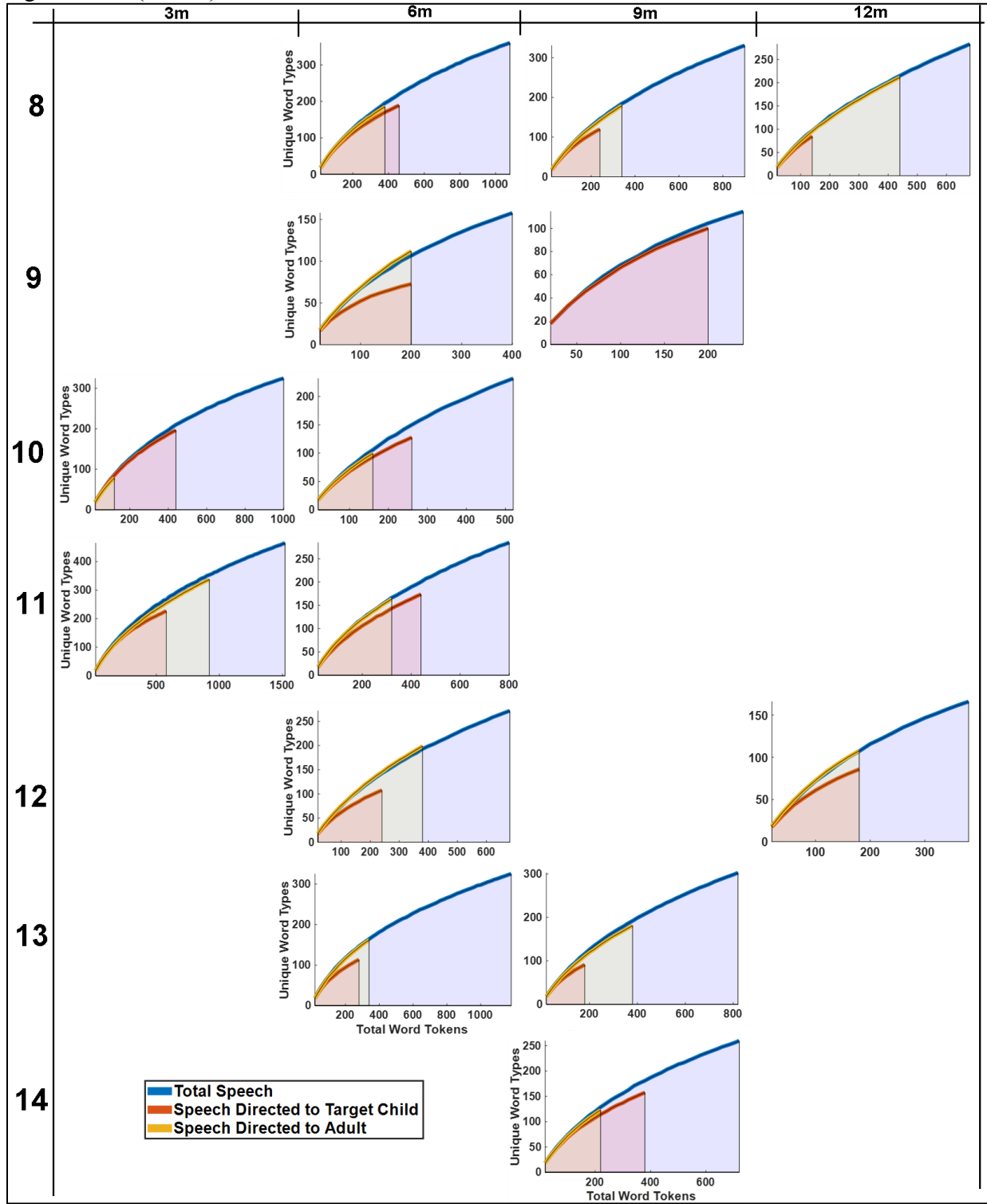


Figure 2-11. Type-token curve and the area under them are shown for total words (blue line and blue shaded area), words directed to the target child (orange line and shaded area by orange), and adult-directed words (yellow line and shaded area by yellow).

Figure 2-11. (cont'd)



To better summarize the quality of linguistic environments of each child across lexical dimensions, we proposed calculating the *area under the type-token curve (AUTTC)* as a summary measure that jointly reflects the amount of daily speech in children’s environments and lexical diversity in speech input. Figure 2-12 presents the distribution of children based on this measure of lexical richness derived from three types of speech: total, child-directed, and overheard. The values are the average values of *AUTTC* calculated over the recordings available for each child. *AUTTC* is able to better represent richness of linguistic input experienced by each child along lexical dimension. Table 2-9 represents the central tendency and dispersion values for *AUTTC* measure for three types of total, child-directed, and overheard speech. The large differences between children and with respect to the average *AUTTC* value suggest that children with CIs do not benefit equally in experiencing high-quality early language input. Further investigation is required to understand the magnitude of contribution of this variability in lexical information to the large differences between individual children with CIs in their language outcomes.

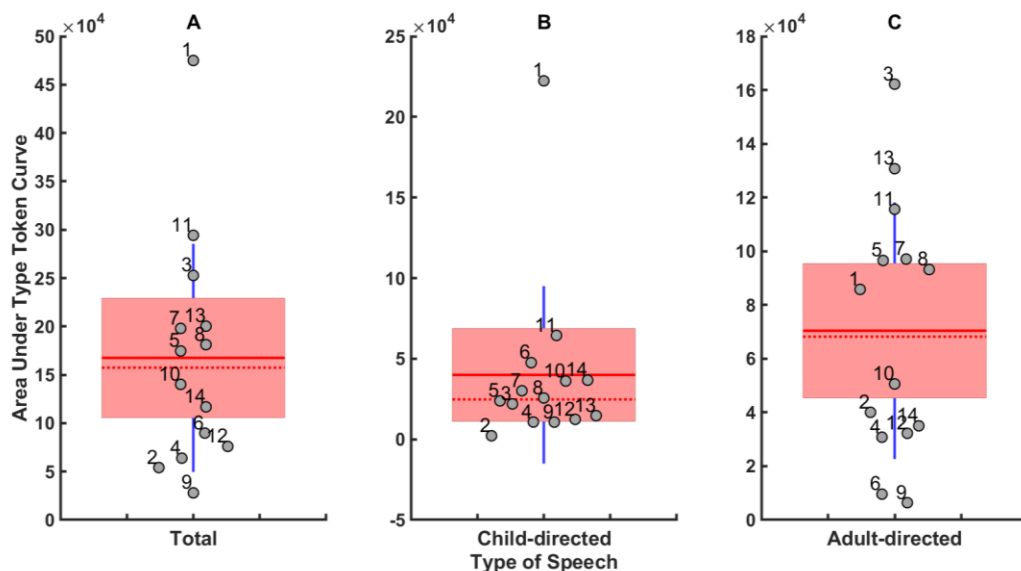


Figure 2-12. Variation in the area under the type-token ratio calculated for 14 children participated in this study for (A) total speech, (B) child-directed speech, and (C) adult-directed speech.

Table 2-9. Measures of dispersion and central tendency for the AUTTC derived from language input of each child in his/her linguistic environments. These measures were calculated for total, speech, child-directed, and overheard speech.

Measure of Language Input (per day)	Measures of Variability						
	Min	Max	Range	Mean	SD	Median	IQR
AUTTC from total speech ($\times 10^4$)	2.8	47.5	44.7	16.7	11.8	15.7	12.4
AUTTC from child-directed speech ($\times 10^4$)	0.2	22.2	22.0	4.0	5.5	2.5	2.4
AUTTC from overheard speech ($\times 10^4$)	0.6	16.2	15.6	7.0	4.8	6.8	6.5

2.3.4. Variability across children in morphosyntactic complexity of their linguistic environment

Figure 2-13 shows the scatterplot of MLUs of adults' speech in the environments of children in this study for total words, child-directed speech, and overheard speech. Table 2-10 shows the dispersion and central tendency of this measure across children for total words, child-directed words, and overheard words. The variability in MLU of speech across children is particularly important because of the impact of high-quality child-directed speech on children's language outcomes. Comparing Figure 2-4 and Figure 2-13 reveals an interesting pattern in this data. Language environments that attest a higher total number of words (e.g., Child 1 and 6) do not necessarily provide children with better morphosyntactic information (MLU_m), suggesting that these two measures of language input (number of words and MLU_m) probably characterize speech for two different qualities along largely orthogonal dimensions.

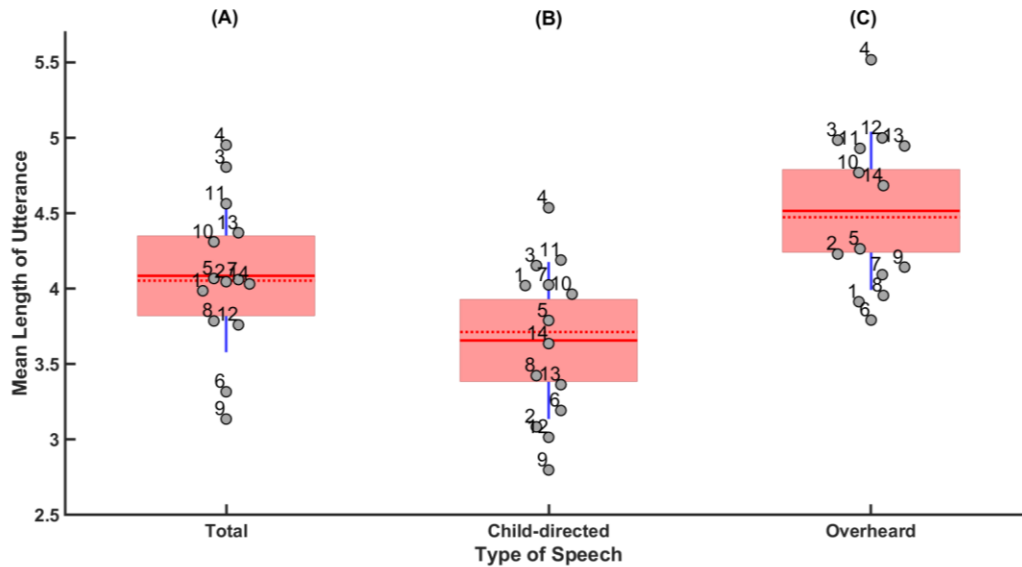


Figure 2-13. Mean length of utterance of speech experienced by each child with CIs in her/his linguistic environments derived from (A) total speech, (B) child-directed speech, and (C) adult-directed speech.

Table 2-10. Measures of dispersion and central tendency for the MLU_m derived from language input of each child in his/her linguistic environments. These measures were calculated for total, speech, child-directed, and overheard speech.

Measure of Language Input (per day)	Measures of Variability						
	Min	Max	Range	Mean	SD	Median	IQR
MLU_m of total speech	3.1	4.9	1.8	4.1	0.51	4.0	0.6
MLU_m of child-directed speech	2.8	4.5	1.7	3.7	0.52	3.7	0.8
MLU_m of overheard speech	3.8	5.5	1.7	4.5	0.52	4.5	0.8

2.3.5. Variability across children in caregiver responsiveness

Figure 2-14 shows differences across children in terms of caregiver responsiveness, defined as the percentage of children's speech (or speech-like) vocalizations which were responded to promptly with child-directed speech by adult caregivers in the child's home environment. The average caregiver responsiveness was 21% across children with a large range of 40.1% across children ($SD = 12.9\%$, $IQR = 14.1\%$). This large amount of variability across children in caregivers' responsiveness during the first year after CI activation may help us to better explain individual differences in language outcomes.

Further, we found a high positive correlation between *caregivers responsiveness* and number of child-directed words ($r = 0.86$, $p < 0.01$). Figure 2-15 shows the relationship between these two measures across children in this study. Later, we will return to considering what this significant positive correlation could suggest about language development.

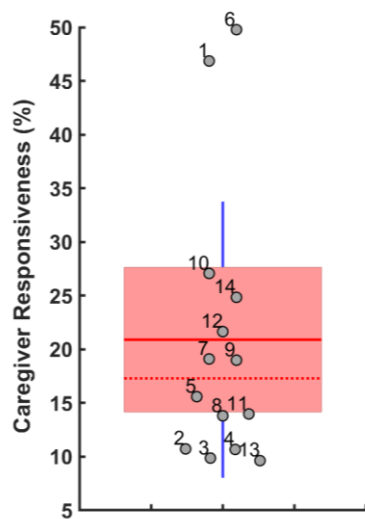


Figure 2-14. Variation across children in the percent of caregivers' responsiveness in the environment of each child.

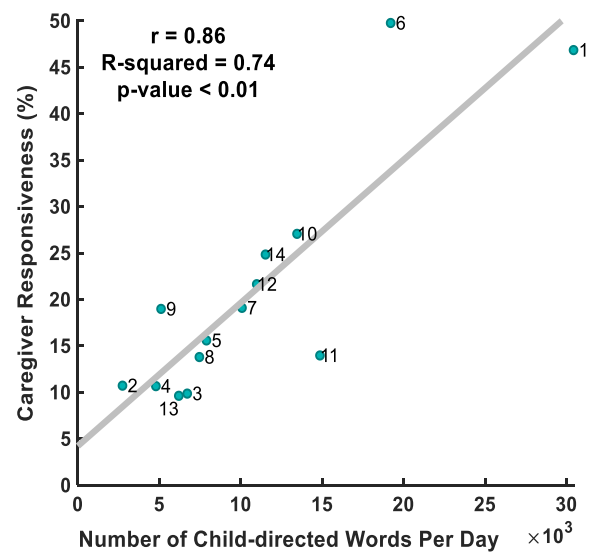


Figure 2-15. The relationship between number of child-directed words experienced by each child per day and caregiver responsiveness as a percentage.

2.3.6. Variability across children in their experienced adults' speech rate

The distribution of speech rates in speech heard by 14 children with CIs in our study is shown in Figure 2-16 for total speech, child-directed speech and adult-directed speech. Adult-directed speech is presented instead of overheard speech to exclude the confounding effect of different types of addressees on adult talkers' speech rate in overhead speech (i.e., when addressees are other children or unknown). Within the 5% samples, the average rate of speech that children experienced in their environments was 3.5 words/second ($range = 1.1$, $SD = 0.35$, $IQR = 0.7$). Figures 2-16B and C suggest that this pattern is not consistent across children; some children

experienced faster child-directed speech compared to the rate of adult-directed speech in their environments, although the overall distribution supports the prior findings in (Bergeson, Miller, & McCune, 2006).

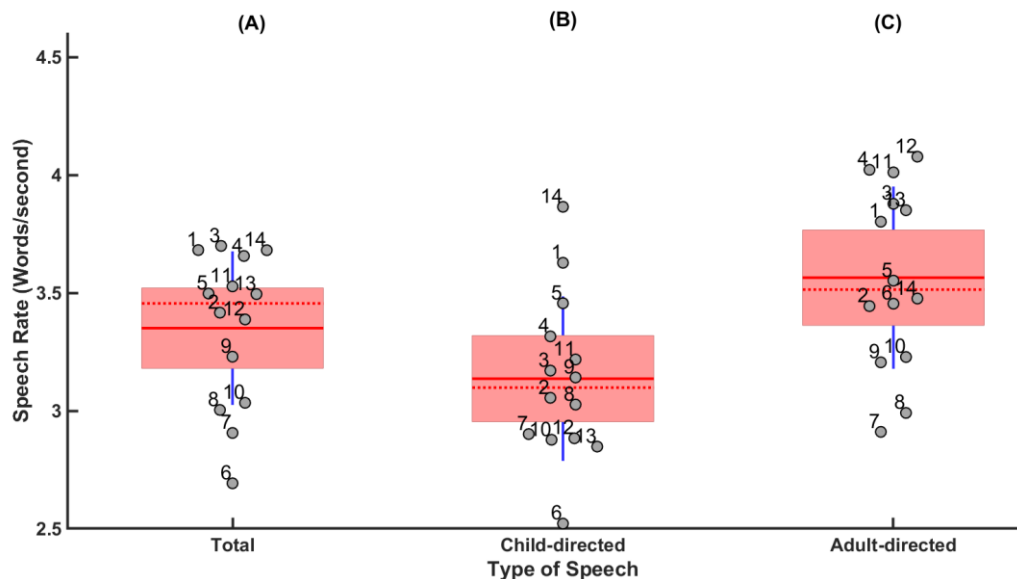


Figure 2-16. Variability across children in the rate of speech by which they experienced speech in their environments calculated from (A) total speech, (B) child-directed speech, and (C) and adult-directed speech.

Table 2-11. Measures of dispersion and central tendency for the average speech rate derived from language input of each child in his/her linguistic environments. These measures were calculated for total, speech, child-directed, and overheard speech in words per second.

Measure of Language Input	Measures of Variability						
	Min	Max	Range	Mean	SD	Median	IQR
Rate of total speech	2.7	3.7	1.0	3.4	0.30	3.5	0.6
Rate of child-directed speech	2.5	3.9	1.3	3.1	0.35	3.1	0.4
Rate of overheard speech	2.9	4.1	1.2	3.6	0.39	3.5	0.7

2.3.7. Variability across children based on a multidimensional space of language input

The measures presented thus far together provide multiple distinct, complementary estimates of the quantity and quality of language input in children's early linguistic environments across lexical, morphosyntactic, and pragmatic language subsystems. Table 2-12 presents z-scores for each child for each of the measures of quantity and quality early linguistic environments. We

would like to point out that some of the linguistic measures used in this study are not yet validated, and therefore should be interpreted cautiously, such as caregivers' responsiveness.

Table 2-12. Z-score values of each child for each measure of quality and quantity of linguistic input. The measures from 1 to 11 include: (1) the total number of words per day, (2) the number of child-directed words per day, (3) TTR for total speech, (4) TTR for child-directed speech, (5) AUTTC for total speech, (6) AUTTC for child-directed speech, (7) MLU for total speech, (8) MLU for child-directed speech, (9) speech rate for total speech, (10) speech rate for child-directed speech, (11) caregivers' responsiveness.

CHILD	MEASURE OF LANGUAGE INPUT										
ID	1	2	3	4	5	6	7	8	9	10	11
1	2.27	2.73	-2.21	-2.14	2.61	3.31	-0.2	0.7	-1.02	-1.41	2.02
2	-0.84	-1.12	1.37	2.35	-0.96	-0.69	-0.08	-1.1	-0.2	0.23	-0.79
3	0.47	-0.57	-0.1	0.67	0.72	-0.33	1.42	0.96	-1.07	-0.1	-0.86
4	-1.55	-0.84	1.07	0.67	-0.88	-0.53	1.71	1.69	-0.94	-0.51	-0.79
5	-0.03	-0.41	-0.37	0.26	0.06	-0.29	-0.03	0.26	-0.45	-0.92	-0.41
6	0.09	1.17	0.19	-0.88	-0.66	0.14	-1.52	-0.89	2.02	1.76	2.24
7	0.46	-0.1	-1.18	-0.13	0.26	-0.18	-0.05	0.71	1.36	0.67	-0.14
8	-0.1	-0.47	0.15	0.58	0.12	-0.26	-0.59	-0.45	1.07	0.31	-0.55
9	-1.8	-0.79	1.23	-0.27	-1.18	-0.53	-1.87	-1.65	0.37	-0.01	-0.15
10	0.36	0.37	0.35	0.12	-0.23	-0.07	0.45	0.59	0.97	0.74	0.48
11	0.88	0.56	-0.61	-0.7	1.07	0.45	0.94	1.03	-0.54	-0.23	-0.54
12	0.03	0.02	0.97	-0.05	-0.78	-0.5	-0.64	-1.23	-0.11	0.72	0.06
13	0.11	-0.64	-0.74	-0.03	0.28	-0.46	0.56	-0.56	-0.44	0.83	-0.88
14	-0.36	0.1	-0.12	-0.45	-0.43	-0.06	-0.11	-0.04	-1.01	-2.09	0.31

2.3.8. Variation in quality and quantity of linguistic input based on child-directed speech vs. adult-directed speech

The characteristics of speech spoken to children with CIs reflects similar modifications made by adults (mothers or caregivers) to their normal-hearing children when addressing their child (Bergeson, 2011). These differences across features of child-directed and adult-directed speech have not yet been studied for speech experienced by children with CIs in natural settings. Understanding how, and to what extent, changes in the addressee (target child vs. adults) modifies linguistic features of adults' speech in naturalistic environments will allow us to better formulate the relationship between exposure to child-directed speech and improved language outcomes.

Figure 2-17 shows the distributions of five language measures of number of total words per day, TTR, AUTTC, MLU_m, and speech rate for child-directed and adult-directed speech

derived from analyzing the 5% audio samples in this study. These measures were the average measures for each child across his/her available recordings. The distributions are kernel distributions which are non-parametric representations of the probability density functions of each measure (Hill, 1985). Using kernel distributions allowed us to avoid making any assumption about distribution of the data. These plots corroborate prior findings that showed speech directed to children with CIs is lexically and morphosyntactically simpler than adult-directed speech (Bergeson, 2011).

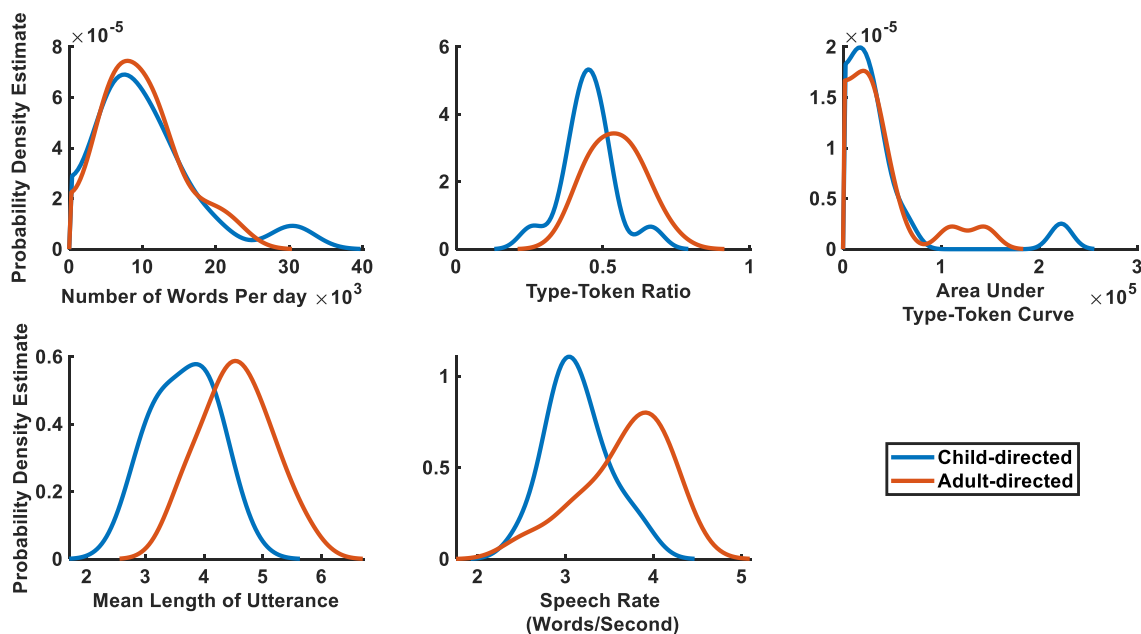


Figure 2-17. Distribution of measures of language input for child-directed vs. adult-directed speech in the environments of children with CIs.

A set of paired *t*-tests, shown in Table 2-13, were conducted to identify whether the observed differences between child-directed speech and adult-directed speech across each of these dimensions of linguistic input were statistically significant. The results showed that child-directed speech was significantly different from adult-directed speech for TTR ($p < 0.05$), MLU_m ($p < 0.01$), and speech rate ($p < 0.01$). child-directed speech and adult-directed speech were not significantly different for the total number of words or AUTTC.

Table 2-13. Results of statistical paired t-test to compare variation of measures of total number of words, TTR, AUTTC, MLU, and speech rate across child-directed speech and adult-directed speech conditions.

Measure of Linguistic Input	Paired Differences				t	df	p-value
			95% Confidence Interval of the Difference				
	Mean	Std. Deviation	Lower	Upper			
Number of Words	3.5×10 ³	8.6×10 ³	-4.2×10 ³	-5.7×10 ³	0.35	13	0.73
TTR	0.09	0.12	-0.16	-0.02	-2.8	13	0.01*
AUTTC	126.46	6.22×10 ⁴	-3.6×10 ⁴	3.6×10 ⁴	0.0076	13	0.99
MLU	0.89	0.70	-1.29	-0.48	-4.77	13	0.00*
Speech Rate	0.53	0.45	-0.79	-0.26	-4.34	13	0.00*

2.4. Discussion

The current study investigated early linguistic environments of 14 prelingually deaf children who received cochlear implants (CIs) early in childhood (age at activation < 24 months) to characterize variability across these children in terms of the quantity and quality of language experienced during the first year after implantation. Day-long audio recordings from the natural auditory environments of these children were extensively and systematically analyzed to evaluate inter-subject variability in the quality and quantity of language input experienced across a set of lexical, morphosyntactic, and pragmatic dimensions. The analysis of the sampled audio was based on a system of hand coding which was developed due to LENA being found to be unreliable and insufficient for the intended purpose of this study.

Variability across children in the total amount of words experienced

Our results showed significant individual variability across children with CIs in total words they experienced per day in their linguistic environments. This large variability is observable from the standard deviation of around 9,000 words per day with an interquartile range of around 7,000 words per day (25th percentile: ~22,000 words per day; 75th percentile: ~29,000 words per day). Exposure to relatively larger amount of total words per day in some children (e.g., Child 1 at

46,000 total words per day) compared to other children (e.g., Child 9 at 8,000 words per day) suggests that children vary substantially in their chances to experience extra “nonregulatory” speech that is argued by Hart and Risley (1995) to be a source of individual differences in language outcomes. These differences between children with CIs in their exposure to very different numbers of total words per day can potentially impact their learning and developing words and a linguistic system. The striking range of disparity of ~38,000 words per day between the child who experienced the highest number of words and the one who experienced the lowest number of words highlights a potential contributor to the observed enormous variability across children in their language outcomes.

Variability across children in the amount of high-quality child-directed speech

It is not merely the quantity of words that a child hears in his/her linguistic environment that matters, but also the quality of speech experienced by children which is important. One of the major elements of quality of early linguistic environments is child-directed speech. Children with CIs in our study differed considerably in the amount of high-quality child-directed speech they experienced. The standard deviation of around 7,000 child-directed words across an average of ~10,000 child-directed words per day suggests that children are substantially different from each other in terms of the amount of high-quality child-directed words they experience per day. The distance of around 27,500 words per day between the child living in a linguistic environment with lowest amount of high-quality child-directed speech (Child 2) and the child living in an environment with the highest amount of high-quality child-directed input (Child 1) may help us to explain variability between these children in their language outcomes later on. Experiencing a higher number of child-directed words per day not only exposes children to a greater number of

high-quality words but also increases the chance of hearing more words types with higher repetition (Huttenlocher et al., 1991; Naigles & Hoff-Ginsberg, 1998).

These results extended – for the first time – the prior well-established findings in children with typical hearing to a pediatric CI population, namely, there is tremendous variability across children in amount of child-directed speech they experience per day (Greenwood, Thiemann-Bourque, Walker, Buzhardt, & Gilkerson, 2011; Hart & Risley, 1995; Hoff, 2006; Sperry et al., 2018). To our knowledge, this is the first study that demonstrates the patterns of large individual differences among children with CIs in exposure to total number of words early after their cochlear implantation, which may conceivably account for large differences between individual children with CIs in their language outcomes, a possibility explored in Chapter 4.

Individual variability in the quantity and quality of language input over years after implantation

Estimates of the amount of speech input were extrapolated to estimate cumulative effects of numbers of words over years on variability across children in the quantity and quality of language in their environments. Extrapolating to three years, we found evidence of a projected gap of near 30 million words separating the children experiencing the most and the least child-directed words, which is shown to be an important contributor to individual differences in children's language development (Hart & Risley, 1995; Hoff, 2003; Weisleder & Fernald, 2013). This gap of 30 million words in exposure to child-directed corroborates prior findings by Hart & Risley (1995). However, this substantial difference across children with CIs in exposure to high-quality child-directed words was not associated with children's socioeconomic status (*Pearson $r=0.35$, p -value* $=0.22$), consistent with the findings from Sperry et al. (2018).

Comparing the results of Hart and Risley (1995) and these results, we observe that a similar magnitude of difference across children exists, which could potentially have a major impact on their

school readiness (Rowe, 2012). The impact of these substantial differences across children in the amount of exposure to child-directed speech per day on vocabulary growth and early scholastic achievements in school is expected to be higher for children with CIs compared to children with typical hearing. This is because the reliance by children with CIs on facilitative properties of high-quality child-directed speech is conceivably much higher than for typical hearing children.

Individual variability in lexical diversity of language input

The degree of word diversity has been recognized as a strong predictor of children's vocabulary growth in children with typical hearing (Hart & Risley, 1995; Hoff & Naigles, 2002; Huttenlocher et al., 1991; Pan et al., 2005; Rowe, 2012; Shneidman et al., 2013; Weizman & Snow, 2001). The results suggest that children were different in terms of exposure to more types of words based on TTR. However, these differences were not purely due to lexical diversity and influenced by the number of total words per day as well. To solve this caveat in using TTR to assess lexical diversity of speech input, we constructed type-token curves to appropriately represents the joint variation in number of words and word diversity (Montag et al., 2018).

From type-token curves obtained from each recording, we demonstrated that children with CIs are quite different from each other in experiencing high-quality speech input characterized as exposure to more words with higher diversity in a day. Using type-token curve, we observed that children with CIs are largely different in experiencing child-directed speech with rich lexical information. The total number of words and word types are two important factors in explaining the individual differences in language outcomes of children with normal hearing (Hart & Risley, 1995; Hoff & Naigles, 2002; Montag et al., 2018).

We summarized the variation of language input across two dimensions of number of word types and word tokens by calculating the area under the type-token curve for each child. Our results

demonstrated a sizeable difference in lexical richness of language input experienced by children with CIs. An extreme individual difference in AUTTC for child-directed speech was observed between the child with the greatest and the lowest quality of linguistic environment. The variability in lexical richness among children is particularly large for child-directed speech, which may help us to explain a good portion of poor vocabulary growth in some children compared to others (Szagun & Stumper, 2012).

Individual variability in morphosyntactic complexity in speech input

Additionally, we examined how children with CIs varied in experienced linguistic environments based on morphosyntactic information in speech input, which is shown to influence language outcomes both in children with typical hearing (Furrow et al., 1979; Harkness, 1977; Hoff-Ginsberg, 1998; Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002) and those with CIs (Szagun & Schramm, 2016; Szagun & Stumper, 2012). The results from this analysis showed that MLU in speech experienced by the 14 children with CIs in a day varies from an MLU_m as short as 2.8 to a long MLU_m of 4.5. Although a few studies showed that shorter MLUs in speech input have a positive impact on syntactic development in children with typical hearing (e.g., Furrow et al., 1979), several others showed the opposite pattern both in typical hearing children (Harkness, 1977; Hoff-Ginsberg, 1998; Huttenlocher et al., 2002) and in children with CIs (Szagun & Schramm, 2016; Szagun & Stumper, 2012). This discrepancy across prior findings about the effect of MLU of speech input on language outcomes makes the interpretation of these observed variabilities across children with CIs in MLU_m of their experienced language input difficult. Future studies with a higher-constraint design will be required to elaborate on the casual effect of variation in MLU on children's language outcomes.

Individual variability in caregiver responsiveness

One of the major aspects of quality of language environments is the quality of communicative interaction between children and adults, which triggers social-pragmatic components of language learning (Kuhl, 2004, 2007). We therefore derived a measure of caregiver responsiveness, taken as the percentage of the time that adults in children's language environments responded promptly to speech or speech-like activities of the children with CIs with a child-directed speech utterance. The range of variability was about 40% between the child who experienced the highest and lowest caregiver responsiveness. This suggests extreme individual variability across children in experiencing linguistic environments with rich social-pragmatic features.

More importantly, the degree of responsiveness of adults was strongly associated with number of child-directed words per day. This could be interpreted to suggest that exposing children to a higher number of child-directed words increases their engagement in parent-adult interaction, and encourages them to initiate a conversation, leading to higher recruitment of language subsystems involved in receptive and expressive skills and better language outcomes (Evans, Maxwell, & Hart, 1999; Tamis-LeMonda et al., 2014). It should be noted that a higher number of child-directed words per day does not automatically satisfy the conditions for caregivers' responsiveness. It is possible that both represent independent, converging measures of the quality of language environments provided by caregivers for children with CIs.

Individual variability in the rate by which speech experienced

Moreover, we examined the variability of rate of speech experienced by each child in his/her environments as one factor that might relate to clarity and intelligibility of perceived speech (Bradlow, Kraus, & Hayes, 2003). Different choices of speaking style often impact speech rate,

where slower rate is thought to contribute to enhanced intelligibility of clear speech compared to conversational speech both in typical listeners (Ferguson & Poore, 2010) and recipients of CIs (Li et al., 2011; Zanto, Hennigan, Östberg, Clapp, & Gazzaley, 2013). Our results showed that the range of difference between children in average of rate of speech input corresponded to approximately one more word per second both in child-directed and overheard speech, which may negatively impact children's ability to segment words and comprehend speech in real-life settings.

Prior studies showed that performance of listeners with CIs is considerably poorer than listeners with normal hearing (~90% decline) in understanding sentences spoken with high rate (Zanto et al., 2013). Four of 14 children with CIs in our study (Child 1, 4, 5, and 14) experienced child-directed speech with an average rate of greater than the mean of "normal speech rate" for adult (i.e., ~5 words/second) as reported in Zanto et al. (2013). This suggests that these children might be at risk for not taking advantage of high-quality child-directed speech input because adults used a fast speaking style when addressing the child with CIs. These rate differences across speech addressed to children with CIs may result in individual differences in effective isolation of speech sounds (vowels and consonants) and syllable boundaries during listening to continuous speech, processes critical for word segmentation and vocabulary learning during language development (Jusczyk, Hohne, & Bauman, 1999; Jusczyk, Houston, & Newsome, 1999; Mattys & Jusczyk, 2001). Such possibilities remain to be tested.

Individual variability in early linguistic environments in a multidimensional language space

We computed the z-scores for each child for measures of number of words per day, type-token ratio, area under the type-token curve, mean length of utterance, speech rate, and caregivers' responsiveness to assess to what extent children are spread around the mean values. Based on these z-scores, we expect that *Child 1* experienced language input with highest quality along multiple

dimensions, while *Child 9* is at high risk for experiencing language input with lowest quality. These results in multidimensional language space corroborate the general trends of large variability across children in richness of language input experienced early after CI implantation. These results highlight the fact that children's linguistic environments can be better modeled by measuring language input along multiple dimensions of vocabulary, syntax, and morphology for appropriate characterization of children's language experience (Montag et al., 2018). This approach might be more effective particularly for studying children with CIs who are at high risk for developing poor language outcomes. Further research will establish the extent to which each of these measures of language input influences language outcomes, which can be further used to better model the connection between language input and outcomes.

Indices of separation between child-directed speech and adult-directed speech

We also examined how child-directed speech and adult-directed speech in early linguistic environments of children with CIs are different for number of words, TTR, AUTTC, MLU, and speech rate. The results from this analysis corroborate prior findings that child-directed speech provides simpler models of language input (see Hoff, 2006 for an extensive review). Our results suggest that child-directed speech has shorter MLU, lower TTR, and lower rate compared to adult-directed speech. Our study is the first one that extends these results to characterizations of ecologically valid, naturalistic environments. However, the number of words and AUTTC in child-directed speech and adult-directed speech obtained across children were not significantly different. These two measures mostly reflect the quantity of speech that are not expected to be a function of speaking style and there is no prior evidence for such a difference. Comparing the results of number of words for overheard speech (see Figure 2-5) with number of words for adult-directed speech suggests that this non-significant pattern is due to excluding the amount of words that were

identified as overheard words but not adult-directed words. In fact, a large portion of overheard words was not necessarily adult-directed words. These results provide preliminary evidence for language features that distinguish child-directed speech from adult-directed speech, which have been shown both in children with typical hearing and children with CIs (Bergeson, 2011; Bergeson et al., 2006). These patterns can be further studied in the future to understand how they relate to supportive effects of child-directed speech on language outcomes.

Overall, we have provided evidence that the quality and quantity of early linguistic environments vary substantially across children with CIs. Children with CIs are at high risk for lack of developing age-appropriate language skills and the language outcomes are shown to be substantially different across children. Based on analyzing the audio samples recorded from natural linguistic environments of 14 early-implanted children with CIs, this study showed that individual children with CIs are at risk for experiencing extremely variable language input- measured through multiple features of quantity and quality of language input to represent lexical, morphosyntactic, and social-pragmatic aspects of input. Results from this study suggest that some children with CIs may be doubly disadvantaged in acquiring spoken language: both due to degradation associated with electronic hearing, as well as due to substantial variability in the amount and quality of linguistic input available to them.

More importantly, this study has provided – for the first time – evidence of a cumulative 30-million-word gap between children with CIs who experienced the highest quality language environment and the one who experienced the lowest quality language environment over three years’ time, mirroring the patterns in Hart and Risley (1995). This new evidence of substantial variability across children in their early language experience may conceivably account for a large portion of variability in children’s communicative outcomes in real-world language development

in this at-risk clinical population. Prior studies have shown that such large differences among children in exposure to high-quality child-directed speech strongly predict individual differences in children's vocabulary skills and scholastic achievement (Dickinson, Golinkoff, & Hirsh-Pasek, 2010; Hart & Risley, 1995; Hoff, 2003; J. Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010; Rowe et al., 2012; Walker, Greenwood, Hart, & Carta, 1994). Future research will establish the extent to which variability in the amount and quality of language input in the clinical population of children with cochlear implants predicts their language outcomes, and whether any such relationships reflect a causal mechanism. We will provide a preliminary test of this relationship in Chapter 4 of the current dissertation.

Limitations of this study

The relatively small sample in this study suggests that these results should be cautiously interpreted. Further studies with higher numbers of participants is necessary to assure that these results are generalizable. Also, our small sample of children was relatively heterogeneous. For instance, some children had mixed combinations of hearing aids and CIs, and others had multiple diagnoses. Each of these points of population variation deserves a well-powered investigation. The distribution of daylong audio recordings for each child was also not equal for all children, such that for some children four recordings were available, whereas others had only one recording. Conducting studies with a higher number of recordings during the first years after implantation will reduce the possible effect of noisy data in our small number of recordings on the estimated per-day measures. In addition, the results of this study were based on analyzing 5% of audio that was randomly sampled from each daylong audio recording. Although prior studies demonstrated that estimates based on analyzing several hours long of audio fairly represent quality and quantity of linguistic environments of children with typical hearing, further studies are required to examine

this for children with CIs. Another limitation of this study is related to the calculation of caregivers' responsiveness, where we assumed that a child-directed speech within 3 seconds after initiation of speech or speech-like vocalization from the target child was semantically and contextually relevant to child's vocalization. Considering these limitations, the results should not be taken as the final determination of how the quality and quantity of linguistic environments of children with CIs vary across children, but rather as preliminary findings to guide further exploration of these questions.

Despite these limitations, the current study provides new evidence on how early language environments of children with CIs vary across children during early childhood based on analyzing their natural home environments. This is the first study to conduct an extensive and fine-grained investigation of the early language experience of children with CIs to assess individual differences in exposure to high-quality language input across lexical, morphosyntactic, and social dimensions, while considering the speaking style of adult talkers in children's environments.

CHAPTER 3: Effects of environmental noise and reverberation on quality and quantity of early language and auditory environments of children with CIs

3.1. Introduction

The prior chapter articulated the central problem in the present dissertation: tremendous unexplained variability in language outcomes of children with CIs. A large literature with typically-hearing children motivated this dissertation project's focus on variability in the quantity and quality of language input as a focus of likely unexplained variation in language outcomes of children with CIs. Specifically, building on seminal work of Hart and Risley (1995), we sought to determine if the magnitude of difference between children with CIs experiencing the most and the least language input approached the 30-million word gap reported by Hart and Risley for their typically-hearing children who heard the most and the least language, respectively.

Very little research has focused on the effects of language input factors on language outcomes in children with CIs (see Szagun & Stumper, 2012 and Szagun & Schramm, 2016). However, crucially, no studies so far have examined other sources of variability that is likely essential to consider for this population: noise and reverberation in children's linguistic environments. Populations of children and adults with CIs are extremely sensitive to noise, which can highly degrade speech perception in this population (Caldwel & Nitttrouer, 2008; Dorman & Gifford, 2017; Hazrati & Loizou, 2012; Neuman et al., 2004; Neuman, Wroblewski, Hajicek, & Rubinstein, 2010a).

The focus of this chapter is to address a fundamental knowledge gap in studies of language development in children with CIs, which is the effect of environmental noise and reverberation on

words experienced by children with CIs early after implantation. To this end, the present chapter develops analyses which characterize noise and reverberation affecting language input. We then derive estimates of the amount of “useable” language expected to survive noise and reverberation, updating Ch. 2 metrics to provide noise-adjusted estimates of numbers of words experienced by children in their language environment. We specifically update per-day estimates of number of total words and child-directed words experienced by each child through two separate novel approaches: (1) human perceptual judgments of the level of noise interference, and (2) using a quantitative metric of speech intelligibility tailored to children with CIs. Below we review the literature related to the effect of noise and reverberation on speech recognition and language development in listeners with CIs.

Environmental noise and language development in children with cochlear implants

A variety of studies support that environmental noise has severely degrading effects on speech perception and spoken word recognition for both adult and child CI users. Living in environments with high ambient noise levels may have serious negative consequences on children with CIs’ abilities to comprehend speech and learn language effectively (Caldwell & Nittrouer, 2013; Davidson, Geers, Blamey, Tobey, & Brenner, 2011). Recipients of CIs have great difficulties in understanding speech in background noise and reverberation, particularly when speech occurs in a multi-talker setting (Caldwell & Nittrouer, 2008; Dorman & Gifford, 2017; Hazrati & Loizou, 2012; Neuman et al., 2004, 2010a). Talking and listening over a cellphone is also extremely challenging for users of cochlear implants (Adams, Hasenstab, Pippin, & Sismanis, 2004; Cohen, Waltzman, & Shapiro, 1989). In addition, cochlear implants are mainly designed to transfer the frequency range that is responsible for speech perception without the ability to resolve fine musical patterns (Kong, Stickney, & Zeng, 2005; Svirsky, 2017). Therefore, recipients of

cochlear implants often do not derive pleasure enjoy listening to music compared to individuals with biologically intact hearing. Among these listening difficulties, speech recognition in noisy and/or reverberant environments makes for the greatest impact on language learning for children with CIs, as they are expected to spend substantial portions of their daily lives in noisy environments (Busch et al., 2017).

Vulnerability to environmental noise is a critical factor for speech perception and language learning in pediatric users of CIs because of their significantly impaired ability to separate and process sound sources, due to their having only partial access to fine-grained spectro-temporal cues (Baskent & Gaudrain, 2016; Fu & Nogaki, 2005). Typically developing children with normal hearing need an average of 7 dB higher signal-to-noise ratio (SNR) at age 5 to show the same performance in word recognition as 12 years old children (Bradley & Sato, 2004), demonstrating the significant effect of age on word learning in noisy environments. The size of this age-noise interaction effect on lexical growth is expected to be considerably higher in children with CIs who have already shown deficits in word learning compared to children with typical hearing in quiet environments (Walker & McGregor, 2013).

The effect of environmental noise and loudness on speech recognition of children with CIs

As outlined in the prior sections, research with typically-hearing children has identified a number of language-input related factors which shape language outcomes. Further, a very small amount of research has investigated how variability in language input might shape language outcomes in children with CIs (Szagun & Schramm, 2016; Szagun & Stumper, 2012). However, no studies previously have considered in detail the effects of noise and reverberation on degradation of linguistic signals in examining such outcomes. Quantifying the effects of noise and

reverberation on language input to children with CIs was therefore a novel innovation, which was a focus of this dissertation.

Language development in children with cochlear implants may be negatively influenced by living in complex listening environments where noise frequently exists. Speech recognition in noisy conditions is extremely challenging for recipients of CIs compared to their normal hearing peers (Busch et al., 2017; Dawson, Decker, & Psarros, 2004; Hazrati & Loizou, 2012; Neuman et al., 2004; Schafer & Thibodeau, 2006; Zeng et al., 2005). Naturalistic auditory environments of children with CIs is typically highly complex and full of these challenging conditions (Busch et al., 2017). Children with CIs showed significantly poorer performance in understanding speech in noise (Caldwell & Nittrouer, 2013; Hochberg, Boothroyd, Weiss, & Hellman, 1992; Munson & Nelson, 2005; Neuman et al., 2010a). Results from testing the performance of 27 kindergarteners with CIs on recognition of 10 phonetically balanced consonant-vowel-consonant (CVC) words in variable noise levels (SNR: -3 dB, 0 dB, and +3 dB) and quiet condition showed an average of about 25% and 16% difference between these children and their normal hearing peers in phoneme recognition and words recognition, respectively. Speech recognition in children with CIs were significantly lower than their typically-hearing peers in all three noisy conditions and surprisingly even in quiet (Caldwell & Nittrouer, 2013).

In addition, Firszt et al. (2004) showed that susceptibility of adult CI recipients to background noise for sentence comprehension (e.g., 60 dB SPL, SNR +8) was much higher than their difficulties in understanding soft speech analogous to quiet conversational scenes (e.g., 50 dB SPL). A significant drop of about 22% in scores of open-set word recognition was reported in adults with CIs in response to 20 dB decrease (70 to 50 dB) in the sound pressure level (SPL) of the presented stimuli (Davidson, 2006), highlighting the extreme vulnerability of children with CIs

to background noise and soft speech (Firszt et al., 2004). Background noise and loudness of speech can vary dramatically in a wide range in natural linguistic environments, which may impact children's processing of words in their language environment. Therefore, individual variability across children in experiencing speech in the presence of noise may translate into variability in their vocabulary growth and, by extension, language development.

The effect of reverberation and noise on speech recognition of children with CIs

Another common source of speech degradation in naturalistic environments that has substantial, negative impacts on speech comprehension in children with CIs is reverberation. Understanding speech in reverberant conditions is very challenging for recipients of CIs (Hazrati & Loizou, 2012; Neuman et al., 2004, 2010a); this may, in turn, impact the degree to which they benefit from speech in their environments toward developing language skills. Neuman et al. (2004) compared speech perception of 7 children with CIs in reverberant and quiet conditions with that of 18 normal hearing children in quiet condition in a virtual classroom. The results from their study showed a significant and substantial decline of approximately 20% in scores of speech recognition of children with cochlear implants in reverberant condition (0.8 s) compared to non-reverberant condition. Children were also largely different in their degree of susceptibility to reverberation for speech recognition. Performance of children with CIs further decreased when noise was added to the reverberation. The amount of SNR that children with CIs needed to obtain 50% speech recognition was substantially higher than the SNR level required by children with typical hearing (around 1.6 times of typical hearing listeners).

Hazrati & Loizou (2012) showed that detrimental effect of reverberation on speech intelligibility is higher than noise with an exacerbated effect in combined conditions (i.e., noise-plus-reverberation). Overall, these studies have demonstrated that children with CIs are at risk for

experiencing extremely challenging situations for understanding speech in their daily lives in natural linguistic environments, where speech frequently coexists with noise and/or reverberation. Therefore, studies of sources of variability in language outcomes of children with CIs require considering the impact of noise and/or reverberation on the quality and quantity of speech experienced by children with CIs in their linguistic environments.

Given the goal of this study to quantify how language input was impacted by environmental noise and reverberation, the question is how this goal can be methodologically achieved. One possible solution is to use auditory perceptual judgments of humans as a “gold standard” to identify the level by which environmental noise interferes with speech that is most useful for children in their environments. Identifying the level of masking and disruption imposed by background noise on speech based on human auditory perceptual judgments has several advantages, including robustness to indexical variability in speech input and using contextual information to identify the extent speech in the environments of children with CIs are degraded due to background noise and other sources of speech degradation. In addition, human judgments of the level of intelligibility of speech is robust to variation of speech loudness and can readily and accurately separate the reference speech signal from sources of noise and speech degradation to identify the level of noise interference (Bregman, 1994).

Additionally, a central innovation of this dissertation was to use a state-of-the-art, non-intrusive quantitative acoustic metric, which has recently been developed to validly estimate intelligibility of speech input for CI users under various noisy and clean conditions. In particular, this study used a metric known as the *speech-to-reverberation modulation energy ratio* (SRMR) with its specific implementation tailored to listeners with CIs (SRMR-CI) in order to estimate the intelligibility of speech experienced by children with CIs in their home environments (Falk et al.,

2015; Santos & Falk, 2014). The SRMR-CI quantitative, computational metric for estimating speech intelligibility for listeners with CIs emulates the frequency selectivity of CI speech processing to approximate speech intelligibility in this population. This measure is non-intrusive because it does not require a reference clean signal to measure the relative distortion between the original and degraded speech signal (Santos & Falk, 2014), which is a critical factor for studying the level of speech degradation in natural environments. This quantitative measure of speech intelligibility in listeners with CIs has been shown to be reliably correlated to speech intelligibility for CI users from clinical perceptual tests under different conditions of clean, noisy, reverberant, and noise-plus-reverberation (Falk et al., 2015). We calculated this metric for each speech utterance identified in the environment of children; we further mapped the obtained SRMR-CI value to a corresponding estimate of speech intelligibility as a percent. This speech intelligibility percentage value was used to update estimates of the number of words experienced by each child.

Effects of environmental noise on the separability of child-directed speech from adult-directed speech

In Chapter 2, we reviewed findings suggesting that children are uniquely attuned to child-directed speech. As discussed, child-directed speech is characterized by a diverse set of acoustic properties that affect segmental and suprasegmental attributes. Environmental noise may negatively influence the extent to which children with CIs benefit from high-quality child-directed speech for learning language by distorting the acoustic information responsible for distinguishing child-directed from adult-directed speech. This potential distortion impacts the quality of linguistic environments of children with cochlear implants by making one environment less favorable for a child compared to another environment. Similar to their typically-hearing peers, children with impaired hearing are able to recognize speech directed to them from adult-directed speech, and

they also prefer to attend and listen to child-directed speech over adult-directed speech (Cooper, 1990; Fernald, 1985; Grieser & Kuhl, 1988; Wang, Bergeson, & Houston, 2017; Werker, Pegg, & McLeod, 1994). Child-directed speech also facilitates language learning in children with cochlear implants (DesJardin & Eisenberg, 2007; Dilley et al., 2018). However, we have no knowledge about how various levels of environmental noise may deteriorate acoustic separability of child-directed speech from adult-directed speech, changing the quality of linguistic environments to lower or higher quality. Therefore, this dissertation pursued analyses to answer these questions.

Current Study

Very little prior work has explored characteristics of early linguistic environments of children with CIs; crucially, this work has done so without considering the effect of noise and reverberation on the quantity and quality of speech input (Szagun & Stumper, 2012). However, language learning occurs in spectrally complex natural environments where children frequently face with challenging noisy or/and reverberant conditions (Dorman & Gifford, 2017; Fetterman & Domico, 2002; Fu & Nogaki, 2005; Neuman et al., 2004; Whitmal, Poissant, Freyman, & Helfer, 2007). It is therefore critical to understand how background noise and/or reverberation may degrade lexical and acoustic information in speech experienced by or spoken to infants with CIs in natural environments.

This study aimed to quantify the effects of noise and reverberation on quality and intelligibility of speech experienced by young children with CIs and to evaluate how these sources may influence the lexical experiences of these children. It was hypothesized that environmental noise and reverberation would significantly reduce the total amount of words and amount of high-quality child-directed words experienced by children with CIs in their language environments. Characterizing the effect of noise on the number of useful words in children's linguistic

environments is not only novel but will enhance our understanding both of how noise and reverberation in children's linguistic environments may impact the quantity of useful speech and how this effect vary across children.

Our approach involved a focus on deriving two updated measures - the total number of words and the number of child-directed words - adjusting for noise and reverberation to consider effects on early linguistic experience in these two dimensions for children with CIs. The word count adjustment process was implemented based on both human perceptual judgments of level of noise interference and the SRMR-CI quantitative measure of speech intelligibility that is tailored to listeners with CIs. These adjustments reflect how the estimates of useable numbers of words heard by children with CIs in their natural environments should be updated to account for the probable effects of noise degradation. This study also attempted to elaborate on the notion of "quality" to consider both (a) individual differences across caregivers (e.g., in child-directed speech vs. adult-directed speech separability), in light of (b) a factor especially important for this population, namely noise. This work addresses how separability of child-directed speech from adult-directed speech varies across caregivers, and how this separability is affected by background noise.

This study will provide crucial new knowledge on how environmental noise degrades linguistic input at the word level. Our approach will involve using SRMR-CI quantitative acoustic measures of speech intelligibility to predict the quality and intelligibility of speech experienced by children with CIs in real-life settings, which will provide quantitative metrics for assessment of the amount of useable linguistic input in their environments. We also analyzed the spectro-temporal characteristics of auditory environments of children with CIs to evaluate how the characteristics of noise (e.g., exposure levels, loudness and spectral distributional characteristics)

compared across children with CIs as an additional measure of quality of their linguistic environments.

Below are the research questions that are investigated in the present study:

Question 1: To what extent do environmental noise and/or reverberation influence the number of total words and child-directed words experienced by children with CI in their linguistic environments?

We hypothesized that noise degrades input to a differential extent across individual home environments, such that the picture of amount of language input experienced by children will look substantially different before versus after considering the effects of noise degradation on language input.

Question 2: How does consideration of noise and reverberation change the picture of the word gap across children in the sample, for total number of words and child-directed words?

It was hypothesized that although noise and reverberation will significantly reduce the amount of words experienced by children, children still are substantially variable in quality and quantity of their linguistic environments.

Question 3: Do adults change the acoustic properties of how they speak to children to compensate for interference due to ambient, overlapping noise in the environment?

We predicted that the acoustic separability between child-directed and adult-directed speech (as gauged by the *Mahalanobis distance* metric) would decrease as the level of interfering noise increases. In particular, we hypothesized that higher degree of noise interference will negatively influence the acoustic separability between child-directed and adult-directed speech such that less discriminative acoustic information is available as the level of noise interference

becomes higher, leading to a lower *Mahalanobis distance* measure in higher noise condition compared with lower noise condition.

Question 4: To what extent do the characteristics of environmental noise in terms of spectral energy and loudness vary across children?

We hypothesized that spectral characteristics and loudness of background noise vary considerably from child to child, such that some children experience spectral noise profiles which are more detrimental than others.

To investigate these questions, we utilized the database of audio which had been coded by humans according to procedures described in Chapter 2. We elaborated on this coding by developing a human coding system in which judgments were made about (which words were affected by noise, the source of the noises(s) and the extent to which the noise interfered with understanding the word). These human judgments were then used to update estimates of how many clean intelligible words were heard per day by each child. We also quantified the intelligibility of the signal using a computational method based on SRMR-CI metric. We then updated counts of words using each of these two measures to obtain a picture of variation to derive estimates of “clean, intelligible” words per day and estimate the growth in language input over years. We further characterized the spectral and loudness characteristics of ambient noise in each child’s environment and determined how acoustics of adult speech changed under different levels of ambient, interfering noise.

3.2. Methods

Participants. The participants in this study were the same 14 children with CIs studied in the chapter 2. Refer to Section 2.2.1. of Chapter 2 for more information about these children and their demographic characteristics.

Audio samples and database. In Chapter 2, we described a process for robust, accurate assessment of the intelligible word content in home environments of our sample of children with CI. Recall that audio samples were 30-second audio intervals randomly selected from the daylong audio recording after excluding the intervals where the child was not awake. Five percent of the audio was randomly selected and coded for intelligible words, including who was speaking (whether an adult or child, including the child with CIs) and who the speech was addressed to (whether the target child or someone else). In the present analyses, we expanded on coding derived from the same 5% audio samples as were studied in Chapter 2; these samples had been randomly selected from 32 LENA audio recordings recorded from the auditory environments of 14 early-implemented CIs participated in this study. More information regarding the recordings and sampling process were presented in the *Method* section of Chapter 2.

To recap, a main goal of analyses described in the present chapter was to leverage human annotation of degree of noise interference on word perception described briefly in Chapter 2 to derive estimates of clean or “useable” words which were predicted to more accurately depict input to these children. In addition, we explored applicability of a new measure, SRMR-CI, for a similar purpose in characterizing language input in naturalistic environments of children with CIs. We further investigated profiles of noise across children’s environments and how this noise might affect acoustic separability of child-directed speech vs. adult-directed speech.

3.2.1. Human coding for noise source and level of noise interference

Recall that Chapter 2, Section 2.2.4.2. under *Method* described the method used by human analysts for coding words in selected 30-second audio samples using *Praat* software in linguistic environments of children with CIs. As discussed in that chapter, coding involved indication of where the speech from an adult talker occurred and whether the talker was a man, woman, or child.

The human analysts identified whether the talker’s speech was directed to the target child with a CI, another child, or an adult. A final step was that human analysts identified whether the speech involved perceptual interference from any overlapping noise, and they coded the type and extent of the noise that overlapped and perceptually interfered with the words. Below, we describe the details of the noise coding, elaborating on points briefly discussed in Ch. 2. We then discuss how the noise coding was used in two ways to update estimates of word count.

Figure 3-1 presents a schematic diagram of the human coding process, updated from Figure 2-2, which shows the additional steps involved in analysis of perceptual interference from overlapping noise. In particular, after human analysts coded the selected audio for the above information, they coded whether there was one or more sound sources (e.g., background noise or competing talker(s) or some combination, etc.) that overlapped with words identified as “primary” (e.g., those from the talker holding the floor). When words were identified as having overlapping noise from one or more competing sound sources, human analysts next determined the perceived level of interference with understanding the words – whether *low (L)*, *medium (M)*, or *high (H)* – based on the judged relative loudness and masking by the competing sound source(s). Additionally, human analysts were instructed to identify (as best they could) the type of noise that overlapped with the primary speech. Information about the different types of noise categories and the relevant coding description are indicated in Appendix A (See Chapter 2 for more details on training and evaluation of the coding system.) In cases where there were multiple simultaneous talkers, analysts followed the steps explained in Appendix D.

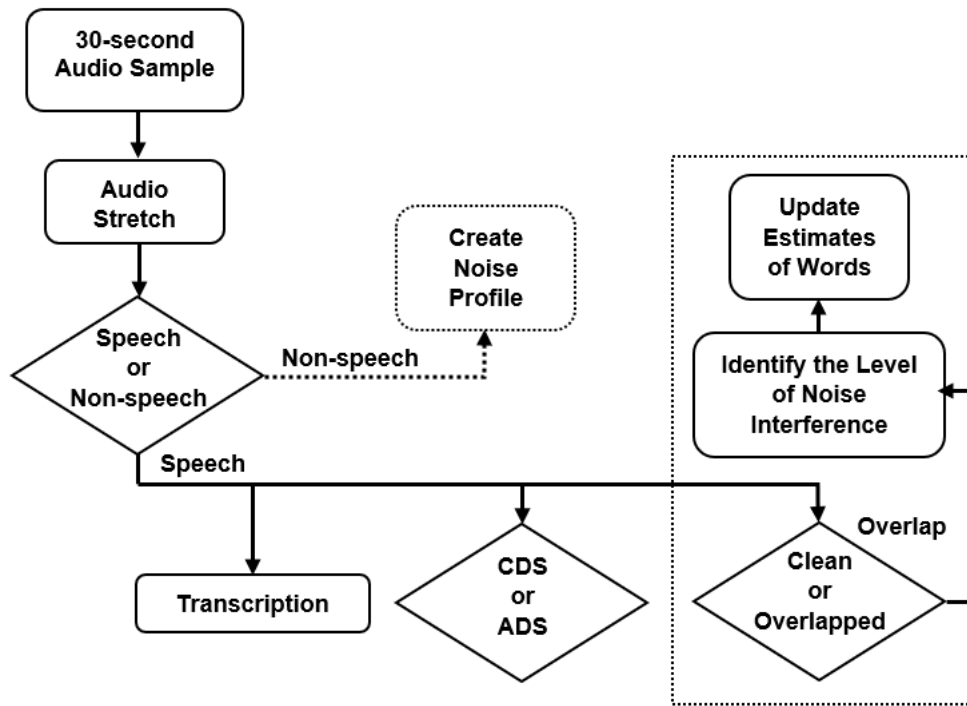


Figure 3-1. Structure of the approach for coding the audio samples for determining the presence of any source of overlapped sounds or speech and its level of interference.

3.2.2. Using human coding to provide updated estimates of number of “clean” words experienced by children with CIs

From the noise interference coding described above, the number of total words and child-directed words experienced by each child were updated based on two methods. The first of these, which is described in this section, uses coding provided by human for the level of noise interference when it overlapped with a speech interval to update the number of useful words experienced by the target child in his/her environments for that speech utterance. The goal was to use the auditory perceptual judgments of word masking made by human analysts to update the useful number of words experienced by each child for developing their vocabulary and language learning.

The computational approach to using human coding of noise interference to update total word counts and child-directed word counts is as follows. The coded textgrids were analyzed in *Matlab* and the *mPraat* toolbox (Bořil & Skarnitzl, 2016) to capture the level of interference of each speech interval and use that information to update the number of useable words experienced by each child in his/her environments. For each interval containing adults' speech, we checked whether there was(were) source(s) of sound or speech that overlapped with the *primary speech* (see Chapter 2, Method section 2.2.4.2. for the definition of primary speech). If so, we searched for the level of interference code (L, M or H) on other tiers at its defined level as illustrated in Table 2-5 in chapter 2. For utterances with different levels of noise interference that were split into shorter utterances with its own level of noise interference, the number of words were updated separately for each shorter interval. This information was extracted in addition to the previous information described in Chapter 2, which involved whether speech was child-directed, adult-directed, or overheard (which included adult-directed speech, but also speech directed to children other than the target child).

To compute updated estimates of numbers of clean and “useable” words, the number of words for each speech interval was first extracted from transcriptions. Then, the following formula was used to update number of words based on low, medium or high level of interference (LOI):

$$f(x) = \begin{cases} 1.00 * x, & LOI = 0 \\ 0.75 * x, & LOI = Low \\ 0.50 * x, & LOI = Medium \\ 0.25 * x, & LOI = High \end{cases} \quad (3-1)$$

In the above, x is the number of words for each speech interval identified by the human coder, and $f(x)$ is the updated number of words based on the level by which a speech interval overlapped with a source of interference. This formula is based on a major assumption that the auditory perceptual

distance in terms of the size of the effect of noise and masking on speech has equal perceptual distance between these levels of interference with linear behavior.

3.2.3. Using SRMR-CI to provide updated estimates of number of “clean” words experienced by children with CIs

The second method of updating the number of total words and child-directed words experienced by each child relied on a quantitative measure of speech intelligibility tailored for listeners with CIs, SRMR-CI. SRMR-CI was used to predict the degree of intelligibility of speech utterances and used the predicted intelligibility to update the number of useable words experienced by each child. The general approach for updating the number of words based on this quantitative metrics of speech quality and intelligibility is shown in Figure 3-2.

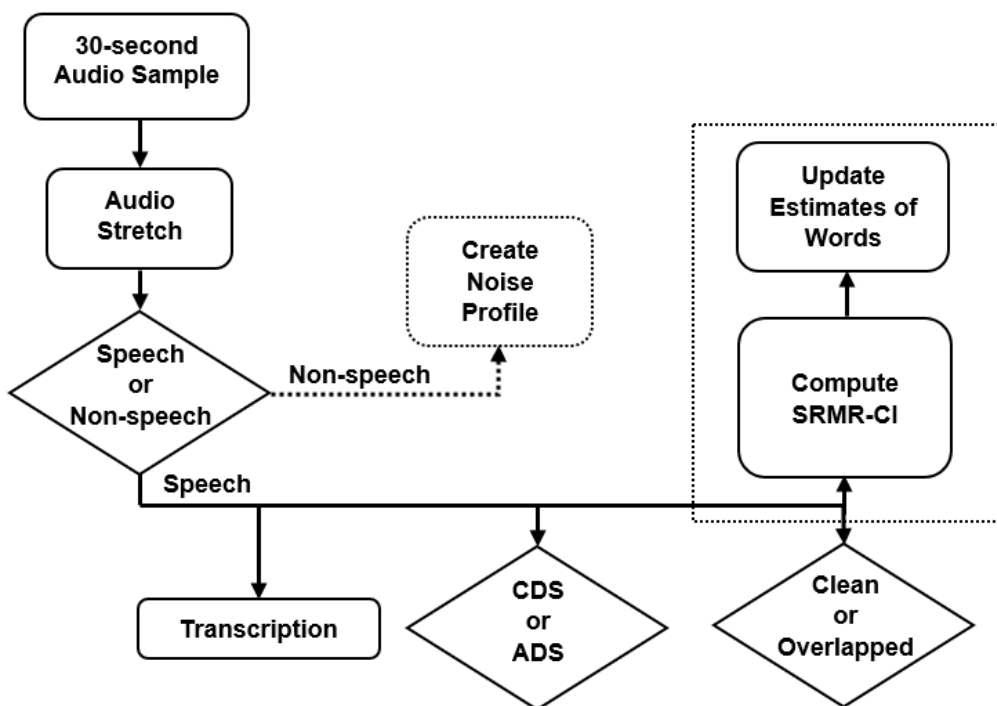


Figure 3-2. General approach for updating the number of words based on SRMR-CI value.

Our hypothesis was that the SRMR-CI could be mapped to a percent intelligibility (SI%), which further could be validly used as a multiplier for the estimated useable number of words in the speech interval. In fact, the updated number of words for each speech interval is the number of words for that interval multiplied by the speech intelligibility percentage value obtained through SRMR-CI. To this end, we first constructed a mapping function from the SRMR-CI ratio to its corresponding speech intelligibility percentage value as proposed in Santos and Falk (2014). We calculated the SRMR-CI values for all speech intervals from all 5% audio samples derived from each of the 32 recordings. (See Ch. 2 for more details.) The intervals that were completely unintelligible (i.e., no intelligible words were transcribed) or partially intelligible (i.e., some words were intelligible and some were not) or overlapped with microphone noise were excluded from this analysis to avoid the effect of artifacts that could interfere with process of SRMR-CI calculation. We particularly used the updated version of this measure (Santos & Falk, 2014), which is expected to be more suitable in studies of naturalistic environments because it reduces the variability of the ratio to indexical factors such as talker gender (which affects e.g., F0) and contextual factors (Santos & Falk, 2014). To make this measure robust to these sources of variability, two methods of modulation energy thresholding and modulation frequency range limitation were applied (Santos & Falk, 2014).

Following the above pre-processing steps, SRMR-CI values were mapped to their corresponding speech intelligibility percentage values based on a sigmoidal mapping function, as proposed in Falk et al. (2015). As shown in Figure 7-d in Falk et al. (2015), mapping the calculated SRMR-CI to a corresponding speech intelligibility percentage value using this sigmoidal function requires first scaling the SRMR-CI ratio to values between 0 and 1. In Figure 7-d in Falk et al. (2015), an SRMR-CI of zero corresponds to the metric's minimum value, while an SRMR-CI of

1 corresponds to the average value attributed to clean speech samples. To carry out the required scaling, therefore, we computed the average value of the SRMR-CI of speech intervals identified by human analysts as clean; the raw SRMR-CI values were then scaled between 0 and 1 based on the following procedure. To compute an SRMR-CI scaled between 0 and 1, we first assumed that $SRMR_CI_{max}$ was the average SRMR-CI value obtained from SRMR-CIs of clean speech intervals. Further, we assumed that $SRMR_CI_{min}$ was the minimum SRMR-CI value from SRMR-CIs of all speech intervals from all 32 recordings. The SRMR-CI values were, then, scaled between 0 and 1 based on the following *minmax* normalization formula:

$$SRMR_CI_{scaled} = \frac{(SRMR_CI - SRMR_CI_{min})}{(SRMR_CI_{max} - SRMR_CI_{min})} \quad (3-2)$$

In the above, $SRMR_CI_{scaled}$ is the scaled SRMR-CI value between 0 and 1. For the speech samples in this study, the obtained $SRMR_CI_{max}$ was 1.65 and the $SRMR_CI_{min}$ was 0.286. Note that $SRMR_CI_{max}$ was the average of SRMR-CI values from speech segments in the clean category (i.e., coded as having no noise interference).

Figure 3-3 shows the distribution of SRMR-CI values for the four levels of interference identified by human: clean (i.e., no noise interference), L, M, or H); higher SRMR-CI values reflects estimates of higher intelligibility. Across increasing levels of noise interference (left to right), this figure suggests that there is a fair negative correlation between a SRMR-CI value and increasing levels of noise interference for speech identified by human coders. This suggests that human coders and SRMR-CI were in rough agreement regarding trends for estimating levels of noise interference. However, it is expected that the results from human auditory perceptual judgments of the effect of noise/overlap on speech will be far from what listeners with CIs perceive, as shown in prior studies (Neuman et al., 2004). In fact, such judgments from adult listeners are a partial reflection of what children with CIs perceive.

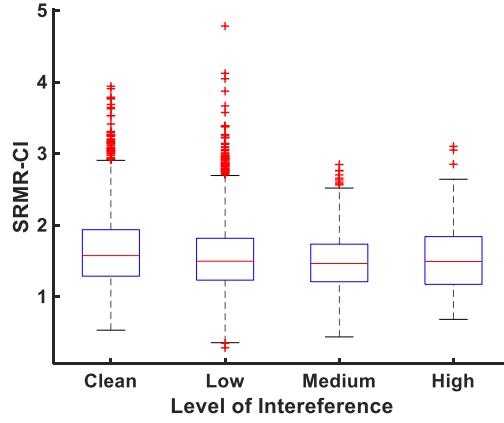


Figure 3-3. The relationship between level of interference identified by human and the corresponding SRMR-CI values for each speech interval. The central red line indicates the median of distribution and bottom and top edges of the box show the 25th and 75th percentiles, respectively.

Having determined scaled SRMR-CI values as described above, we then reconstructed the sigmoidal function given in Falk et al. (2015; Santos & Falk, 2014) that relates SRMR-CI values to estimated speech intelligibility (%). The pairs of scaled SRMR-CI values and their corresponding percent speech intelligibility (SI%) were obtained from (Falk et al., 2015). We used a *non-linear least-square curve fitting* approach to reconstruct this sigmoidal function that maps SRMR-CI values to their estimated speech intelligibility (%). The obtained sigmoidal function from this curve-fitting process was:

$$SI(\%) = \frac{K}{1 + e^{-(\alpha_1 X - \alpha_2)}} \times 100\% \quad (3-3)$$

The model parameters α_1 , α_2 , k are given in Appendix E. The model fitted the observations very well ($R\text{-Squared} = 0.99$), reflecting the fact that the predicted SI values from the model used here were very close to the observed data values derived from Figure 3-3.

The number of words for each speech utterance was then updated according to the following overall procedure. We first calculated the SRMR-CI value for that speech signal and then scaled it between 0 and 1 based on the scaling formula (3-1). If the SRMR-CI value was equal or bigger than $SRMR\text{-}CI_{max}$, we considered it to be a speech sample from the clean category with

SRMR-CI value of 1. We then estimated the speech intelligibility as a percent by using the sigmoidal mapping function constructed above. This speech intelligibility percentage value was used to update the number of words for each speech segment by multiplying the SI ratio with the number of words obtained from the transcription: *updated number of words* = *number of words* x *SI*. For example, a speech interval which contained 6 words with a raw SRMR-CI value for this speech of 1.5 (from a scaled value of 0.89 with $SRMR-CI_{max}= 1.65$ and $SRMR-CI_{min}= 0.28$) would result in a $SI = \frac{0.88}{1+e^{-(13.43*0.89-7.7)}} = 0.87$ (87%). The updated number of words for this interval would then be $6 \times 0.87 = 5.22$. This implies a loss of approximately one word ($0.78 \approx 1$) for CI listeners in listening to this speech, as adjusted by quantitative measure of SRMR-CI.

Finally, the updated number of words across speech intervals was then determined by summing the updated numbers of words across all 5% samples for that recording. Then, these word counts within the sampled audio were extrapolated over a full 16-hour day, following the methodology explained in Chapter 2. For children with more than one recording, the final updated per-day word estimate for each child was the average of values across multiple days (i.e., multiple recordings).

3.2.4. Effect of noise on acoustic separability of child-directed and adult-directed speech

The coding applied by human coders was used to further analyze the characteristics of early linguistic environments of children with CIs to better understand the effect of noise on separability of child-directed speech from adult-directed speech. The goal was to understand how separation of child-directed speech from adult-directed speech may be affected by four levels of noise masking ranging from no degradation (clear speech) to the speech that is overlapped with other sound source signals, causing *low*, *medium*, or *high* levels of noise interference. To this end, the labeled utterances within each 30-second chunk of audio were analyzed to describe the audio

characteristics in feature space. Speech utterances were partitioned into frames of 100 ms, on the assumption that they were quasi-stationary over these short segments. Twelve Mel-frequency cepstral coefficients (MFCCs) were then extracted for each frame. The MFCC model is inspired by the human auditory system and approximates its filtering structure and frequency resolution (Shaneh & Taheri, 2009).

Each speech interval was binned into 30-ms frames and then windowed using a Hamming window of the same length. Twelve MFCCs were captured for each frame. Next, the acoustic-phonetic distance between adult-directed speech and CDS speech segments at each of four levels of degradation was computed using *Mahalanobis* distance (MD) (Masnan et al., 2015). The change in the acoustic distance between child-directed speech and adult-directed speech across different levels of noise interference was then evaluated based on variation in MD.

3.2.5. Variability across children in exposure to noise in their auditory environments

We analyzed auditory environments of each child when no speech was present to create a noise profile. Within each 5% randomly selected 30-second interval of audio from daylong recordings, portions of audio that only contained noise without any other sorts of communicative activities (e.g., live speech, media) were identified. These portions were then analyzed to capture two general forms of information: (1) the spectral energy of the noise under four sub-bands of 0-1, 1-3, 3-5, and 5-8 KHz; and (2) loudness of the noise measured base on a designed A-weighted sound level meter (Genuit, 1999; Schomer, Suzuki, & Saito, 2002; Su, Rubin, & Terman, 2004). Figure 3-4 shows the block diagrams of the steps implemented to calculate these two metrics of noise characteristics in environments of children with cochlear implants.

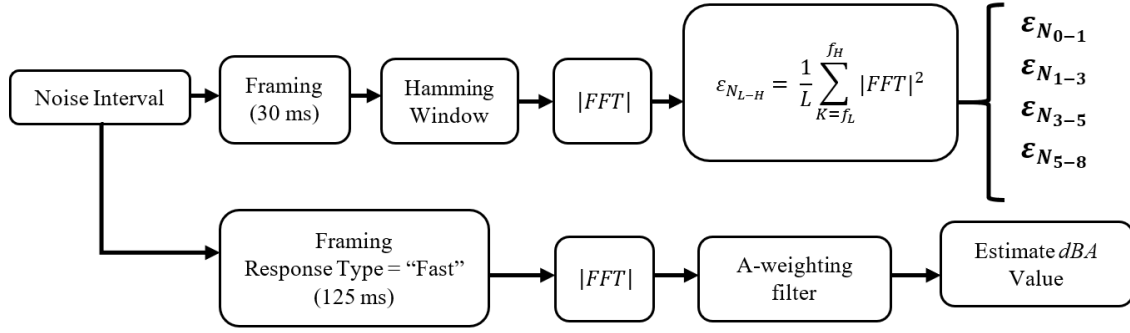


Figure 3-4. Block diagram of the approach to construct the noise profile for environmental noise experienced by each child with a CI.

Recall that human analysts labeled noise only when it overlapped with speech. So, only portions of audio that were (a) within randomly selected 5% of 30-second audio intervals but (b) not coded as speech were analyzed to characterize noise in the environments in these steps. First, to estimate the spectral energy of the noise under each sub-band, each noisy audio interval (i.e., portion of selected 30-second intervals without any live or pre-recorded speech activities) was first framed and windowed using a Hamming window (using a frame and window length of 30 msec). The absolute Fourier Transform of noise signal in each frame was then calculated by performing a Fast Fourier Transform (FFT) analysis. The spectrum was divided into four sub-bands spanning 0-1 kHz, 1-3 kHz, 3-5 kHz, and 5-8 kHz. The energy of the FFT spectrum under each of four sub-bands was calculated to characterize the spectral behavior of noise experienced by each child in his/her environment. The spectral energy under each frequency sub-band was calculated as following:

$$\epsilon_N = \frac{1}{L} \sum_{k=f_L}^{f_H} |X_N[k]|^2 \quad (3-4)$$

where $X_N[k]$ is the spectrum of the noise signal obtained from applying FFT on the noise signal $x_N[n]$. f_L and f_H are the lower and higher bounds of the frequency range. We defined four frequency bands to compute spectral energy of noise spectrum under these sub-bands. The spectral energy of noise under each sub-band was calculated to summarize the spectral characteristics of

each noise frame by four measures of spectral energy ($\epsilon_{N_{0-1}}$, $\epsilon_{N_{1-3}}$, $\epsilon_{N_{3-5}}$, $\epsilon_{N_{5-8}}$). $\epsilon_{N_{f_L-f_H}}$ is the energy of the noise spectrum between the lower frequency f_L and the higher frequency f_H . The overall spectral energy of the noise for a recording was summarized by averaging the spectral energy under each sub-band over all frames within the selected regions of a recording.

In addition, we derived an analysis of environmental noise based on two major principal dimensions of “*loud-soft*” and “*high-low*”. This involved first measuring the loudness of noises in children’s environments in regions of audio discussed above using the A-weighting sound level, which is a common measure for assessment of environmental sounds, particularly the level of noise exposure and annoyance (Parmanen, 2007; Schultz, 1978). The A-weighting filter is a filter structure designed to mimic the perceptual response of the human hearing system to loudness, particularly loudness due to noise; see Figure 3-5. The frequency attenuation of the A-weighting filter corresponds to an empirical average obtained across a broad sample of perceptual experiments. This filtering scheme is frequently used in sound level meters to appropriately measure loudness levels in dB. As Figure 3-5 shows, the A-weighting filter has a peak response near 3 kHz and its frequency response rapidly decreases for lower frequencies (i.e., below 100 Hz). The attenuation is nonmonotonic across the frequency range; and shows a gradual decrease from 3000 Hz to 8 kHz. As previously discussed, this is a crude approximation of the human hearing system, allowing us to quantify the level of sound sources with respect to a standard “human” observer.

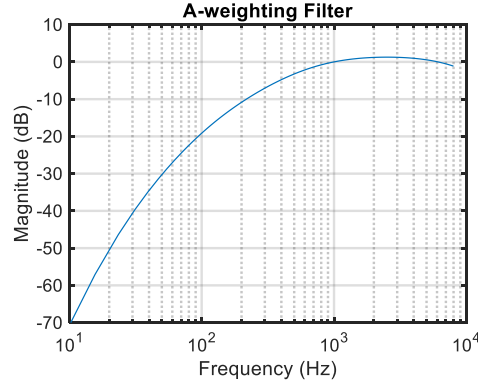


Figure 3-5. Frequency response of A-weighting filter in dB scale.

We used an A-weighting filter to modify the frequency spectrum of noise sound to better mimic the human hearing response to various levels of noise loudness and annoyance in children's auditory environments. The signal level is measured in the frequency domain by applying Parseval's relation (Schomer et al., 2002). We evaluated the amount and degree of noise exposure for each child based on this A-weighted sound pressure level, and also the energy of spectrum under four frequency sub-bands. We then defined a noise exposure index which was taken to be the summation of the duration of each noisy frame multiplied by its loudness level.

$$NEI_i = \sum_{l=1}^L dBA_l \times T_l \quad (3-5)$$

where NEI_i is the *noise exposure index* for the recording i . dBA_l is the noise loudness level of frame l in dBA (A-weighted decibels) and T_l is the duration of the frame. This index thus reflects both the duration of noisy auditory input and its loudness and annoyance experienced by the target child in her/his auditory environments.

3.3. Results

3.3.1. Noise-adjusted estimates of numbers of words experienced by children with CIs

Total number of words. We first focused on the change in the total number of words due to their co-occurrence with overlapped noise. Updated estimates of the total number of words

experienced by each child are shown in Figure 3-6. Blue bars show the estimates of words per day for the 14 children for the *Traditional* method that does not consider noise interference; these values are slightly different from those in Chapter 2 because the speech intervals that overlapped with microphone noise were excluded here in order to have fair comparison with results from the two updated methods based on human coding and SRMR-CI values. Orange bars reflect estimated numbers of useable words, updated to consider noise interference based on *human auditory perceptual judgments* of level of interference (human-based). Yellow bars reflect estimated numbers of useable words, updated to consider noise interference based on *SRMR-CI*. Comparing these bars for each child and across children highlights the detrimental effect of noise and/or reverberation on the number of useable words experienced by children in their language environments.

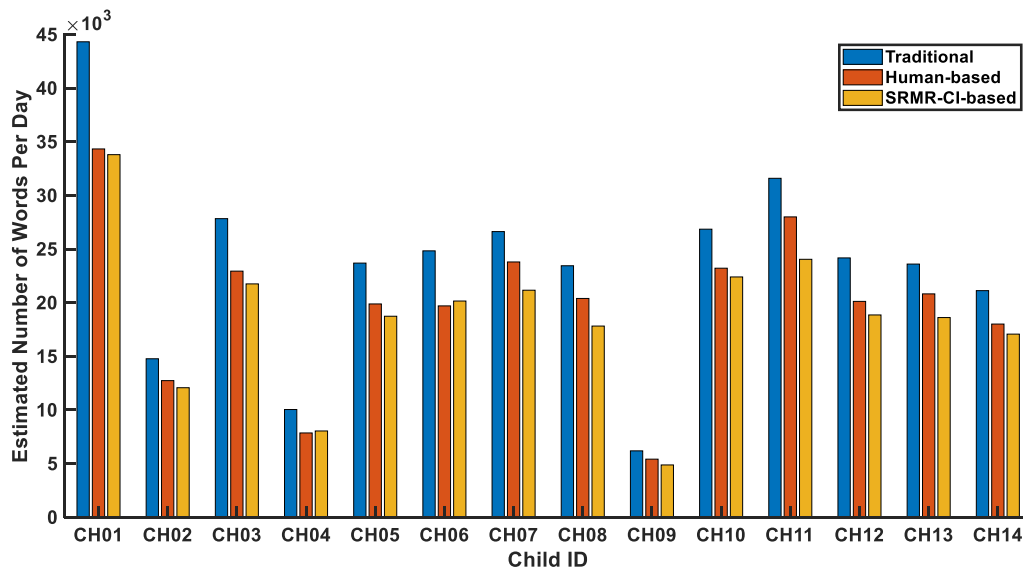


Figure 3-6. The estimate total number of words per day for 14 children with CIs based on three approaches of traditional (unadjusted for noise), and for two methods adjusted for noise: human-based, and SRMR-CI-based.

Figure 3-7 shows the amount and distribution of word loss in number of words (panel A) and in percent (panel B) obtained from two word updating methods of human-based and SRMR-CI-based compared to the number of words estimated from traditional method for the 14 children with CIs participated in this study. *Pair 1* is the word loss between traditional method and human-based method and *Pair 2* is the word loss between traditional and SRMR-CI-based method.

On average, children heard around 37,000 less words per day after updating the number of words for the effect of noise based on human coding of level of noise interference (see *Pair 1* in Figure 3-7A). This means the total number of words experienced by children decreased ~15.5%, on average, across children after updating the number of useful words based on human auditory perceptual judgments of the level of speech degradation (see *Pair 1* in Figure 3-7B). Table 1- presents the measures of central tendency and variability for the amount and percent of the word lost because of noise. This table suggests a gap of 9,216 words loss per-day between the child living in environments with lowest quality of language input and the one experienced the linguistic environments with highest quality in terms of word loss. The degree of individual variability across children can be characterized by standard deviation of 2,134 words per day and *IQR* of 1,273 words per day based on human coding of the level of noise interference.

Updating the number of words based on SRMR-CI metric, children lost 4,978 words on average (Table 3-1 and pair 2 in Figure 3-7A). These figures imply that the amount of decline in exposure to useful words was approximately 21% on average across children after considering noise and reverberation as quantified by SRMR-CI (Table 3-1 and pair 2 in Figure 3-7B). Such a large difference in the effective number of words available to children with typical hearing and those with CIs indicates the high degree of risk faced by children with CIs for gaining less-developed language skills compared to those with normal hearing. The high degree of variability

across children in the amount of words they experienced per day based on updated values from SRMR-CI-based method can be observed by $SD = 2,288$ words and $IQR = 1,569$ words (Table 3-1). Although the loss in exposures to useful words increases after taking into account the effect of noise and reverberation together based on SRMR-CI, the variability and word gap is almost the same regardless of the updating approach (i.e., standard deviation of around 2,000 words and word gap of around 9,000 words per day).

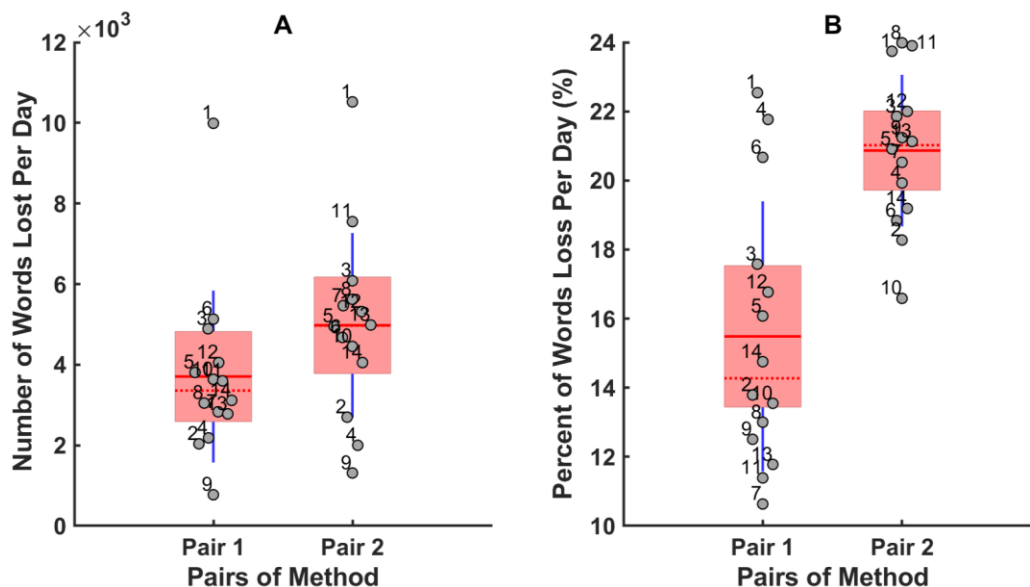


Figure 3-7. (A) The number of words lost per day and (B) percentage of words lost per day compared to the traditional method after considering the effect of noise on the number of total words based on human coding (Pair 1) and SRMR-CI metric (Pair 2). In each panel a scatter plot shows the individual children and boxplot summarizes the distribution of word loss for each child in his/her environment. The data points are laid over a 1.96 standard error of the mean (95% confidence interval) in red and a 1 standard deviation shown by blue lines. The solid and dotted red line show the mean and median, respectively.

Table 3-1. Measures of central tendency and variability for the total number of word loss per day and its percent loss per day for two comparisons between “traditional and human-based method” and “traditional and SRMR-CI method”.

Pair of Method		Measures of Central Tendency and Variability						
		Min	Max	Range	Mean	SD	Median	IQR
Pair 1: Traditional vs. Human	Words	773	9,989	9,216	3,706	2,134	3,357	1,273
	Percent	10.6	22.5	11.9	15.5	3.9	14.3	5.1
Pair 2: Traditional vs. SRMR-CI	Words	1,313	10,521	9,208	4,978	2,288	4,97	1,569
	Percent	16.6	24	7.4	20.9	2.2	21	2.8

Two paired *t*-tests were conducted to determine whether there was a statistically significant differences between estimates derived from traditional measures and those derived from either the human-based or SRMR-CI-based approaches. The first test compared the difference between traditional and human-based methods for the total number of words (Table 3-2). Results of these paired *t*-tests showed that the difference in the estimates of total number of words were statistically significant when comparing traditional ($M = 23,503$, $SD = 9,214$) and human-based ($M = 19,797$, $SD = 7,443$) approaches ($t(13) = 6.49$, $p < 0.05$). The second test compared the difference between traditional and SRMR-CI-based methods for the total number of words (Table 3-2). The difference in the total number of words estimated by traditional and SRMR-CI-based ($M = 18,525$, $SD = 7,014$) approaches was also statistically significant ($t(13) = 8.1$, $p < 0.001$). In terms of the difference between human-based and SRMR-CI-based approach in estimating the total number of words, there was an average difference of 1,272 words per day ($t(13) = 3.99$, $p < 0.05$). However, the Pearson's linear correlation coefficient of 0.99 ($p < 0.05$) suggests that these two approaches probably behave equally in terms of modeling the effects of noise and reverberation on speech intelligibility and the number of useful words.

Table 3-2. Table Results of statistical paired t-test to examine the significant effect of updating approaches of human-based and SRMR-CI-based on the amount of total words lost.

Pairs of Method	Paired Differences				T	df	p-value
			95% Confidence Interval of the Difference				
	Mean	Std. Deviation	Lower	Upper			
Traditional vs. Human	3,795	2,133	2,473	4,937	6.49	13	<0.001*
Traditional vs. SRMR-CI	4,978	2,288	3,657	6,299	8.14	13	<0.001*

We were also interested in examining the distribution of the words that overlapped with other sources of sound or speech and clean words for each child. Figure 3-8 shows the distribution of clean and overlapped words for each child per day. The ratio on top of each bar graph is the *overlap ratio*, which is obtained by counting the number of words per day that overlapped with a source of noise and dividing that count to the original number of words per day. This figure suggests that the traditional approach of word counts is not able to appropriately represent the actual total amount of words experienced by children with CIs early after CI activation. A vast portion of words in the environments of children are contaminated by other sources of sound or speech that overlap with useful speech input, which is expected to impact the number of useable words heard by children with cochlear implants.

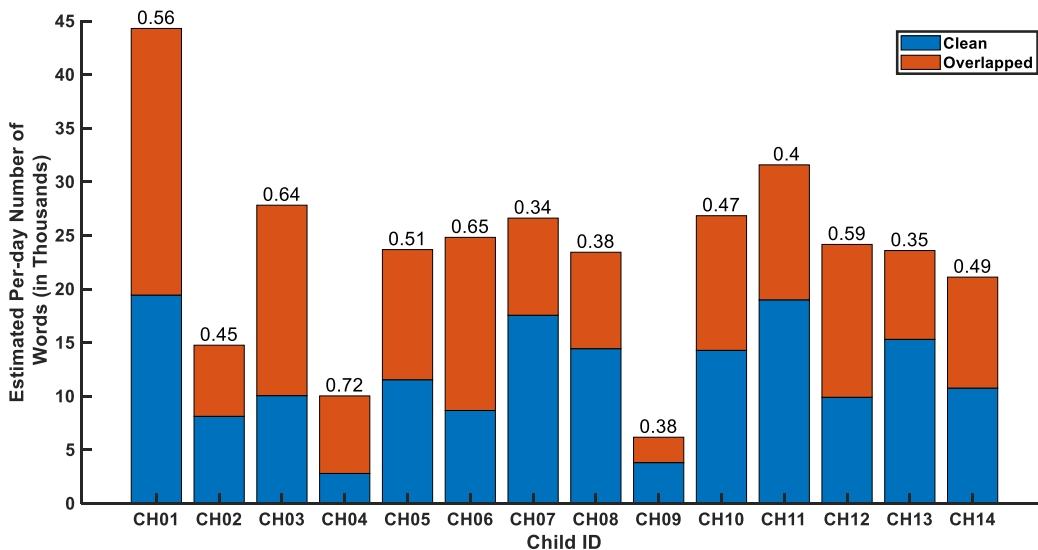


Figure 3-8. Distribution of clean and overlapped words in per-day estimates of the total number of words for each child.

Figure 3-9 shows how the distribution of noise categories overlapping with speech varied across children's auditory environments. As this figure suggests, two categories of overlapped speech (i.e., a competing talker) and mix noise had the greatest prevalence across children. The category of mix noise was for either multiple non-primary talkers overlapping or for two or more types of sound source above overlapping or alternating rapidly (see Table 2-5 and Appendix A in Chapter 2). The distribution of categories of noise types were, for the most part, highly variable across children. The overall pattern of distributions of noise overlapped with speech in the environments of these 14 children highlights the complexity of the auditory environments of children with CIs with respect to understanding speech.

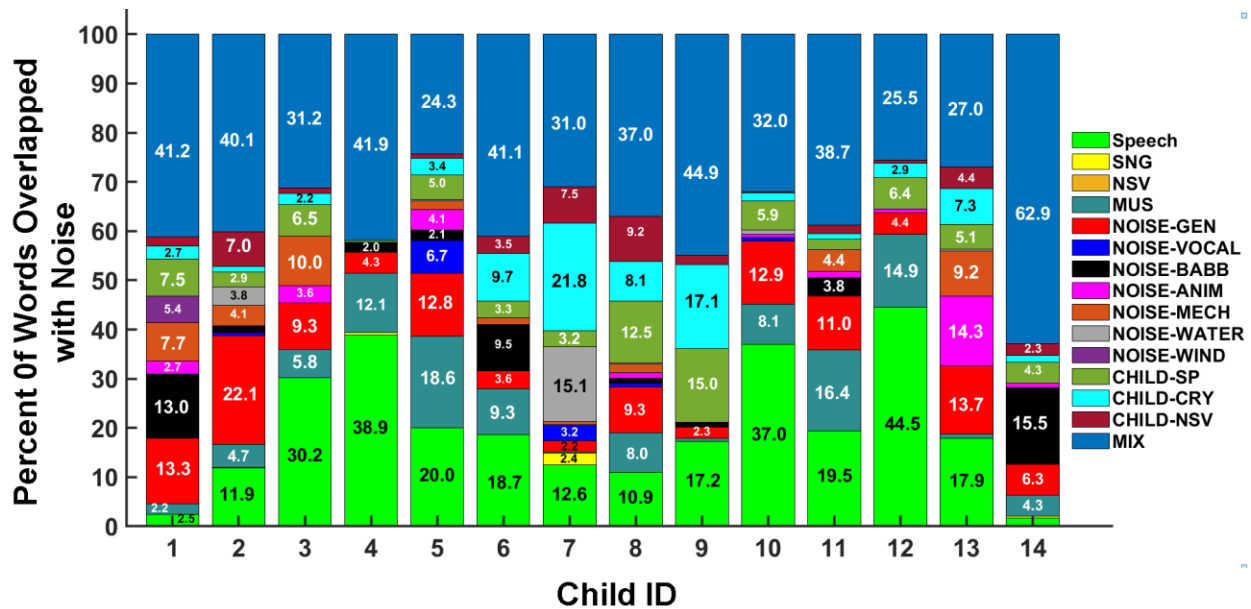


Figure 3-9. Distribution of various types of noise overlapped with speech in the environments of the 14 children with CIs. The code for various noise type are Speech = overlap speech, SNG = singing, NSV=non-speech vocalization, MUS=music, NOISE-GEN=general noise, NOISE-VOCAL=vocalization from other children, NOISE-BABB=multi-talker babbling; NOISE-ANIM=animal noise, NOISE-MECH=mechanized noise sources, NOISE-WATER=running water sound, NOISE-WIND=wind sound, CHILD-SP=speech from target child, CHILD-CRY=target child cry, CHILD-NSV=non-speech vocalization from the target child, MIX=multiple non-primary talkers overlapping or two or more types of sound source above overlapping. See Appendix A for more information regarding these categories of noise or speech or sounds that overlapped with speech from the primary talker.

The variability across children in the total number of words they experienced changes when we update the number of words based on effects of noise and reverberation co-occurring with experienced words. Figure 3-10 shows the distribution of children based on the total number of words estimated by three approaches of *traditional*, *human-based* (i.e., updated based on human perceptual judgments of the level of noise interference), and *SRMR-CI-based* (i.e., updated based on *SRMR-CI* values). Table 3-3 further presents the measures of central tendency and variability for the distribution of total number of words obtained from the three methods of traditional, human-based, and SRMR-CI-based approaches. Within the sampled audio, children experienced an average of 23,503 total words per day based on the traditional method, where effect of noise and reverberation were not modeled in the usefulness of heard speech ($SD = 9,215$ and $IQR = 5,729$).

This average total number of words was updated to 19,797 words after adjusting for the effect of noise through human perceptual judgments of word masking and utterance intelligibility ($SD=7,443$ and $IQR =5,207$). Updating the number of words based on quantitative metric of SRMR-CI resulted in an average total number of words of 18,525 words per day with standard deviation of 7,014 words and inter-quantile range of 4,681 words. These results suggest that patterns of variability across children persist after updating the estimates of the number of words per day. However, children with CIs are expected to be disadvantaged more compared to children with normal hearing, due to the amount of word loss estimated by these updating approaches that consider the effect of noise and/or reverberation in children’s environments on the amount of useable words.

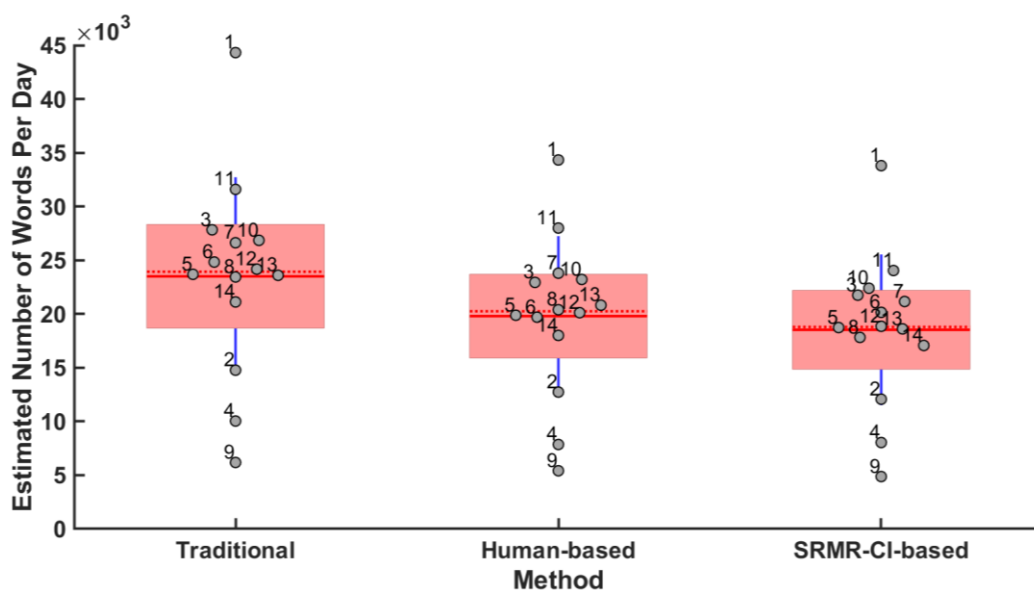


Figure 3-10. Estimated number of total words per day from traditional, human-based, and SRMR-CI-based approaches.

Table 3-3. Measures of dispersion and central tendency for the total number of words estimated through traditional, human-based, and SRMR-CI-based approaches.

Method of Word Estimate	Measures of Variability						
	Min	Max	Range	Mean	SD	Median	IQR
Traditional	6,183	44,317	38,134	23,503	9,215	23,928	5,729
Human-based	5,409	34,328	28,918	19,797	7,443	20,254	5,2075
SRMR-CI	4,869	33,797	28927	18,525	7,014	18,793	4,681

Hart & Risley (1995) showed how the individual differences between children with normal hearing in the number of heard words, could by age 3, extend to a 30-million-word gap between children hearing the most and the least spoken input in their linguistic environments. They also showed that these substantial individual differences in the number of heard words was strongly correlated to children's language outcomes. We next examined how the cumulative values of the total number of words changed after updating the number of words for the damaging effect of noise and/or reverberation to better estimate the useable words heard by each child. Figures 3-11, 3-12, and 3-13 show cumulative values of per-day estimates of total words over five years after CI activation, calculated based on the same extrapolation method over years described in Chapter 2.

These figures highlight the extent to which the size of variability across children in the amount of exposure to words can grow over years after CI implantation, assuming children experience linguistic environments with the same characteristics over these years. A word gap of approximately 40 million words (41.76 million words) in the total number of heard words is expected by three years post-implantation between the children with the most and the least spoken language experienced in their environments (*Mean* = 30.5 million words, *SD* = 16 million words) (Figure 3-11). This word gap turns to around 31.2 million words by this age after updating the number of useful words based on human perceptual judgments of word degradation (*Mean* = 21.7 million words, *SD* = 8.1 million words) (Figure 3-12). Updating the number of words based on the SRMR-CI method leads to an estimated 31.7 million word gap between the children experiencing

the most and the fewest useable exemplars by the time they have 3 years hearing experience (*Mean* = 20.8 million words, *SD*= 7.7 million words) (Figure 3-13). The striking pattern in these estimates is the amount of variability across children that remains even after updating the number of words.

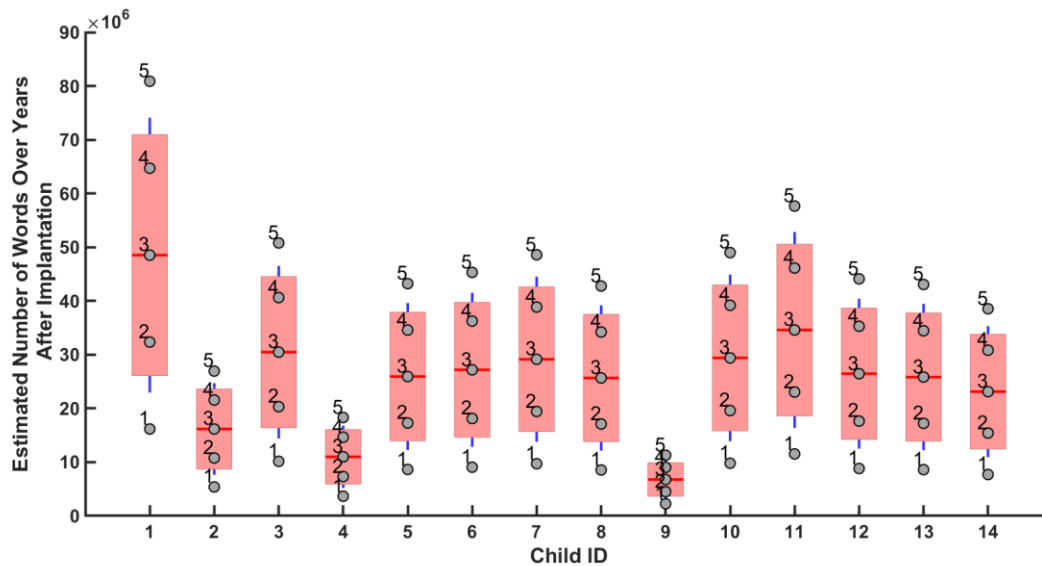


Figure 3-11. The total number of words extrapolated over 5 years after CI activation for each child estimated based on traditional approach. Numbers 1-5 next to data point circles reflect estimated child-directed words after 1, 2, 3, 4, or 5 years.

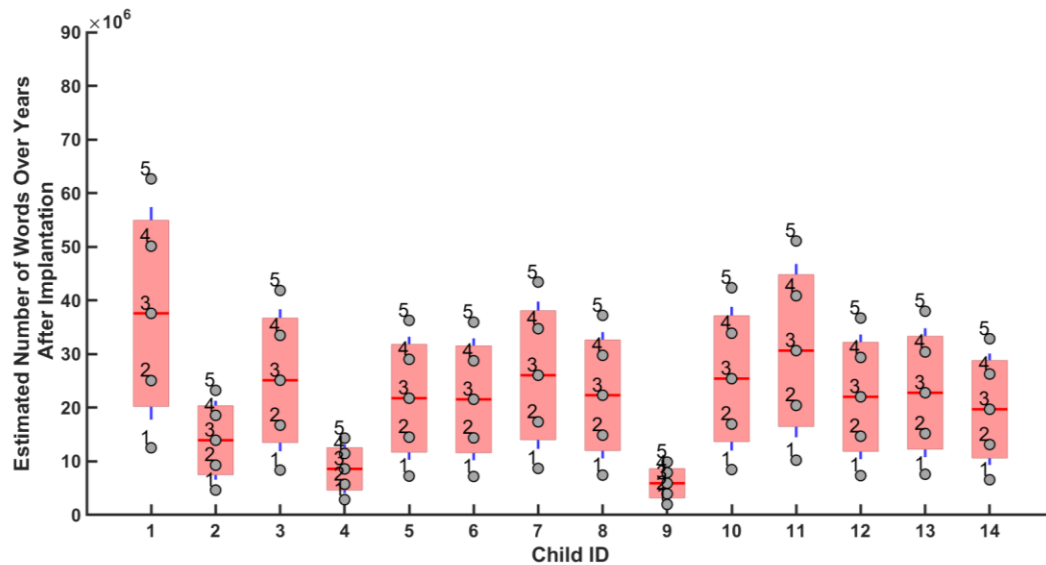


Figure 3-12. The total number of words extrapolated over 5 years after CI activation for each child estimated based on updated number of words by human auditory perceptual judgments. Numbers 1-5 next to data point circles reflect estimated child-directed words after 1, 2, 3, 4, or 5 years

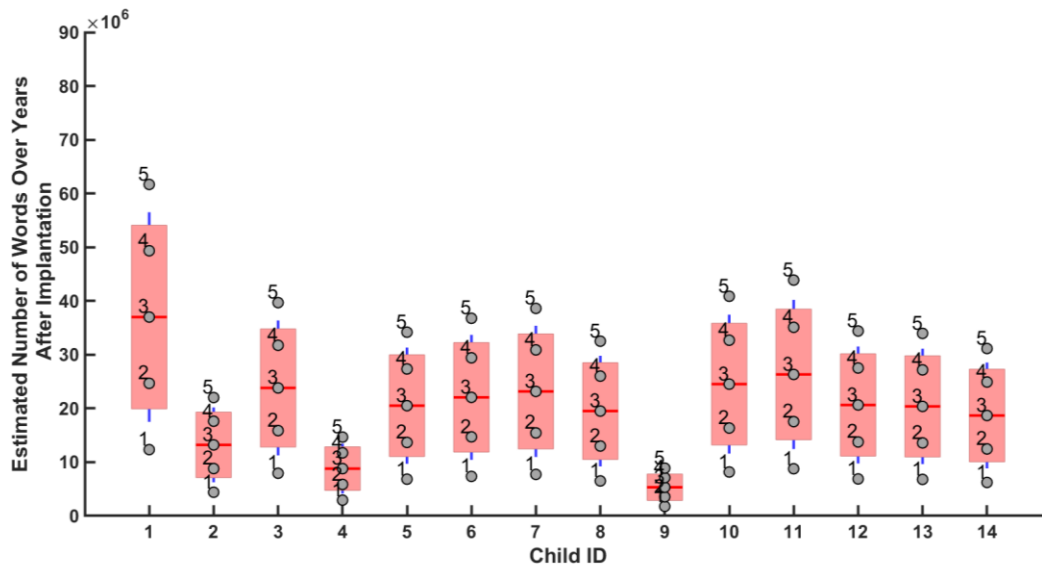


Figure 3-13. The total number of words extrapolated over 5 years after CI activation for each child estimated based on updated number of words by SRMR-CI metric. Numbers 1-5 next to data point circles reflect estimated child-directed words after 1, 2, 3, 4, or 5 years.

Change in numbers of child-directed words. The positive impact of words experienced by children significantly increases if the words are directed to them, i.e., if speech is child-directed.

After investigating the effects of noise and/or reverberation on the total number of words, we turned here to examining how updating the number of words for co-occurrence with noise and reverberation affected the amount of high-quality child-directed speech experienced by each child. Figure 3-14 shows the amount of child-directed words per day estimated through traditional, human-based, and SRMR-CI-based methods as shown by blue bars, orange bars, and yellow bars, respectively. Comparing the number of child-directed words per day estimated from these three measures within and across children highlights the large, and variable, degree of disadvantage brought by exposure to noise and/or reverberation for children with CIs in their auditory environments.

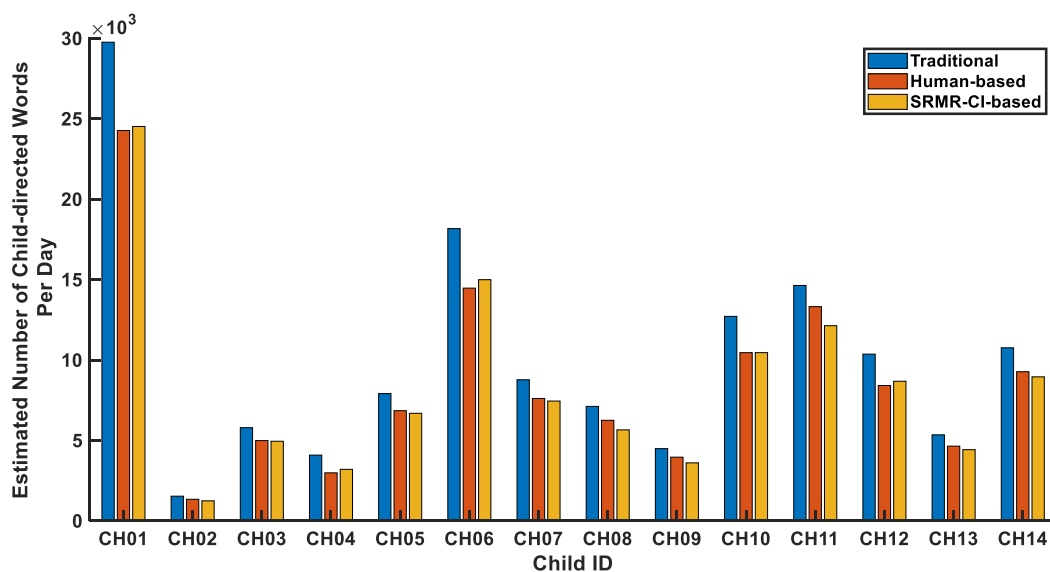


Figure 3-14. The estimate total number of child-directed words per day for 14 children with CIs based on three approaches of traditional (unadjusted for noise; blue line), and for two methods adjusted for noise: human-based (orange), and SRMR-CI-based (yellow).

Figure 3-15 shows the amount of word loss per day (panel A) and the percent of word loss per day (panel B) for child-directed speech based on our two updating approaches, human-based and SRMR-CI-based. The measures of central tendency and variability for the amount and percent of child-directed word loss are further presented in Table 3-4. On average, children lost 1,616

words per day due to the effect of noise that co-occurred with words. As this figure shows, there was a loss of 15.4% of child-directed words after considering the effect of noise based on human perceptual judgments of levels of inference (Table 3-4). The standard deviation of 1,414 in word loss and the range of 5,297 word loss (Table 3-4) shows the degree of variability and the range of difference across children experiencing high-quality child-directed words in the presence of environmental noise.

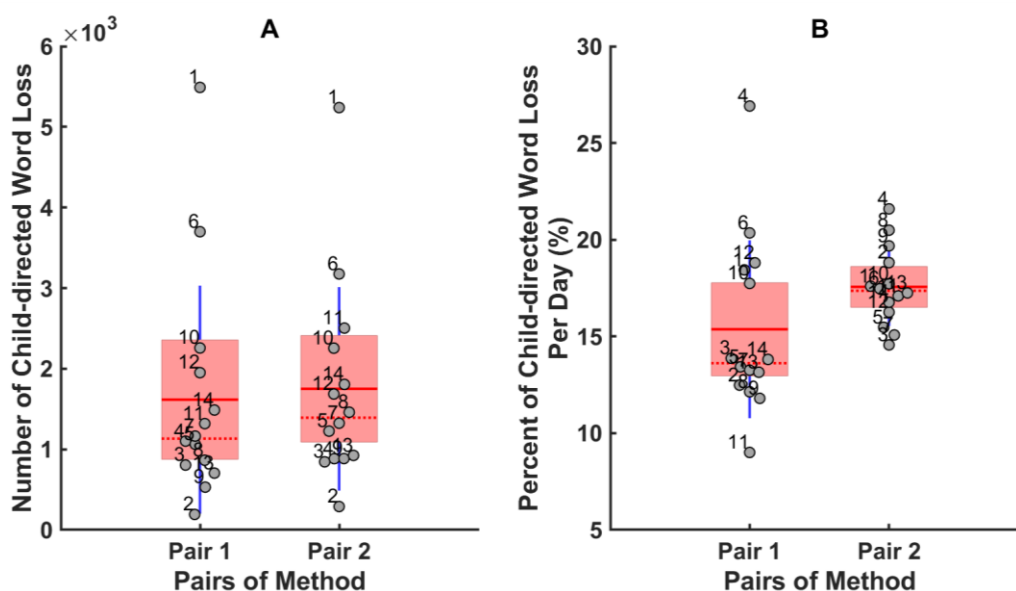


Figure 3-15. (A) The amount of words lost per day and (B) percent of words lost per day compared to the traditional method after considering the effect of noise on the number of total words based on human coding (Pair 1) and SRMR-CI metric (Pair 2).

Results of Figure 3-15 (pair 2) represents the amount and percent of child-directed word loss for each child after updating the number of child-directed words based on the SRMR-CI metric. Updating the number of child-directed words based on the SRMR-CI approach suggests that, on average, children lost the benefit of approximately 1,750 words per day due to background noise and/or reverberation, which corresponds to a loss of 17.6% of child-directed words (Table 3-4). The variability across children in the amount of word loss per day based on the SRMR-CI

metric amounted to a standard deviation of 1,262 words and an interquartile range (IQR) of 1,370 words.

Table 3-4. Measures of central tendency and variability for the number of child-directed words lost and corresponding percentage word loss for two comparisons between traditional vs. human-based and traditional vs. SRMR-CI updating methods.

Pair of Methods		Measures of Central Tendency and Variability						
		Min	Max	Range	Mean	SD	Median	IQR
Traditional-Human	Words	192	5,490	5,297	1,616	1,414	1,133	1,146
	Percent	9	26.9	17.9	15.4	4.6	13.6	6
Traditional-SRMR-CI	Words	290	5,239	4,949	1,750	1,262	1392	1,370
	Percent	14.6	21.6	7	17.6	2	17.4	2.6

Two paired *t*-tests were conducted to examine whether the difference between traditional and human-based methods and also between traditional and SRMR-CI-based methods had statistically significant effects on the estimated numbers of child-directed words. Table 3-5 presents the results of these two paired *t*-tests. This effect of noise and reverberation in children's linguistic environments on the amount of word loss was statistically significant both based on human-based word updating approach ($t(13)=4.27$, $p < 0.05$) and the SRMR-CI-based one ($t(13)=5.18$, $p < 0.05$).

Table 3-5. Results of statistical paired *t*-test to examine the significant effect of updating approaches of human-based and SRMR-CI-based on the amount of total words lost.

Pairs of Method	Paired Differences						
			95% Confidence Interval of the Difference				
	Mean	Std. Deviation	Lower	Upper	t	df	p-value
Traditional vs. Human	1,616	1,414	799	2,433	4.27	13	0.00*
Traditional vs. SRMR-CI	1,750	1,262	1,021	2,478	5.18	13	0.00*

We examined what portion of high-quality child-directed words in the environments of each child overlapped with one or more competing sources of speech degradation. Figure 3-16 shows the distribution of clean and overlapped words per day in child-directed speech experienced

by each child in his/her environment. The ratio on top of each bar in the graph is the *overlap ratio*, which is obtained by counting the number of words per day that overlapped with a source of noise and dividing that count to the original number of words per day. This plot suggests that some children who appeared to be relatively advantaged when considering traditional, raw word counts might not be so advantaged after considering the noise that co-occurred with words in their environments; see Child 6 and Child 7 as examples.

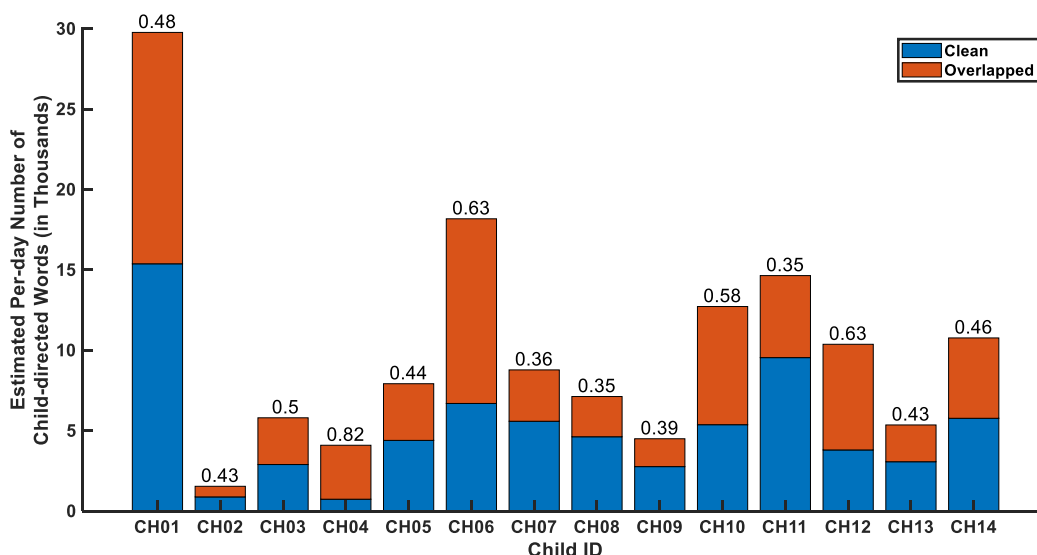


Figure 3-16. Distribution of clean and overlapped words in per-day estimates of the number of child-directed words for each child.

Figure 3-17 summarizes comparisons of estimated numbers of child-directed words derived from a traditional approach and the two novel approaches that consider noise. Table 3-6 further shows the measures of central tendency and variability for child-directed words obtained from these three approaches. The variability across children in the amount of child-directed words based on traditional approach are characterized by $SD = 7,217$ words per day and $IQR = 7,365$ words per day (Table 3-6). The individual variability across children suggested by the human-based approach is characterized by $SD = 5,883$ words and $IQR = 5,812$ words. The variability

across children in the amount of child-directed words based on SRMR-CI approach are characterized by $SD = 5,962$ words and $IQR = 6,034$ words (Table 3-6).

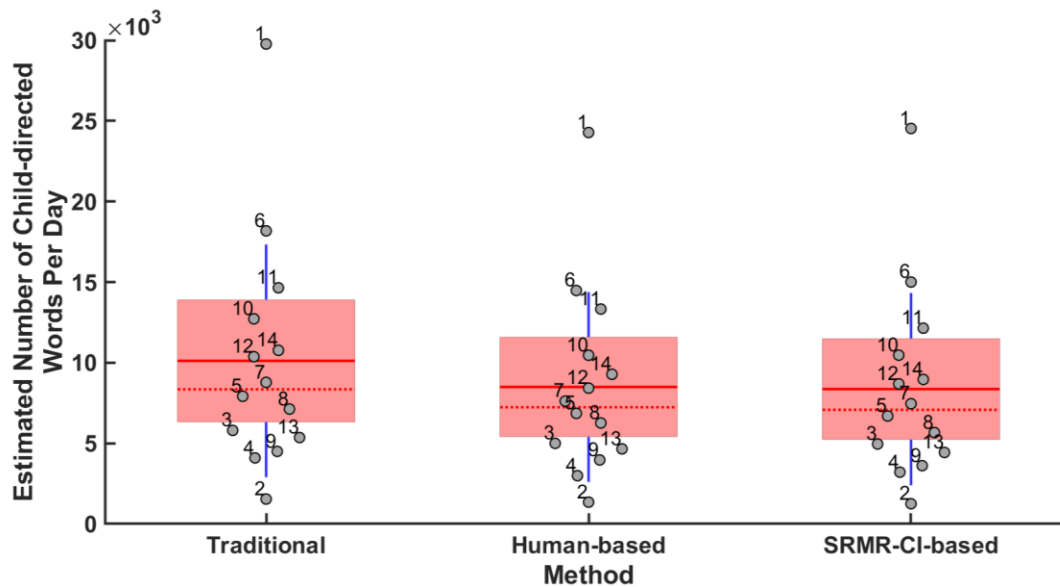


Figure 3-17. Estimated number of child-directed words per day from traditional, human-based, and SRMR-CI-based approaches.

Table 3-6. Measures of dispersion and central tendency for number of child-directed words estimated through traditional, human-based, and SRMR-CI-based approaches.

Method of Word Estimate	Measures of Variability						
	Min	Max	Range	Mean	SD	Median	IQR
Traditional	1,541	29,767	28,226	10,111	7,217	8,350	7,365
Human-based	1,348	24,278	22,929	8,495	5,883	7,236	5,812
SRMR-CI	1,251	24,528	23,277	8,361	5,962	7,075	6,034

Previous studies showed that the cumulative number of child-directed words experienced by children during early childhood predicted their language and cognitive outcomes at 3 years old and later in school (Hart & Risley, 1995; Hoff, 2006; Weisleder & Fernald, 2013). Therefore, we estimated the cumulative number of child-directed words experienced by each child over years after CI implantation through performing the same extrapolation on the number of child-directed words per day as calculated from traditional, human-based, and SRMR-CI approaches. Figure 3-18, 3-19, and 3-20 present the extrapolated data for the 14 early-implanted children in this study

derived from the three approaches of number of words estimation. The word gap between children with the highest-quality linguistic environments and lowest-quality environments for the number of child-directed words experienced by hearing age 3 was around 30 million words (30.9 million words, *mean* = 11.1 million words, *SD* = 7.9 million words), which is close to what was reported by Hart and Risley (1995). However, this gap in the number of child-directed speech slightly reduced to 25.1 million words (*mean* = 9.3 million words, *SD* = 6.4 million words) after updating word estimates based on human perceptual judgments. The word gap between children with lowest- and highest-quality language experience was estimated at a similar value of around 25.5 million words by three years after implantation based on SRMR-CI method (*Mean* = 9.2 million words, *SD* = 6.5 million words).

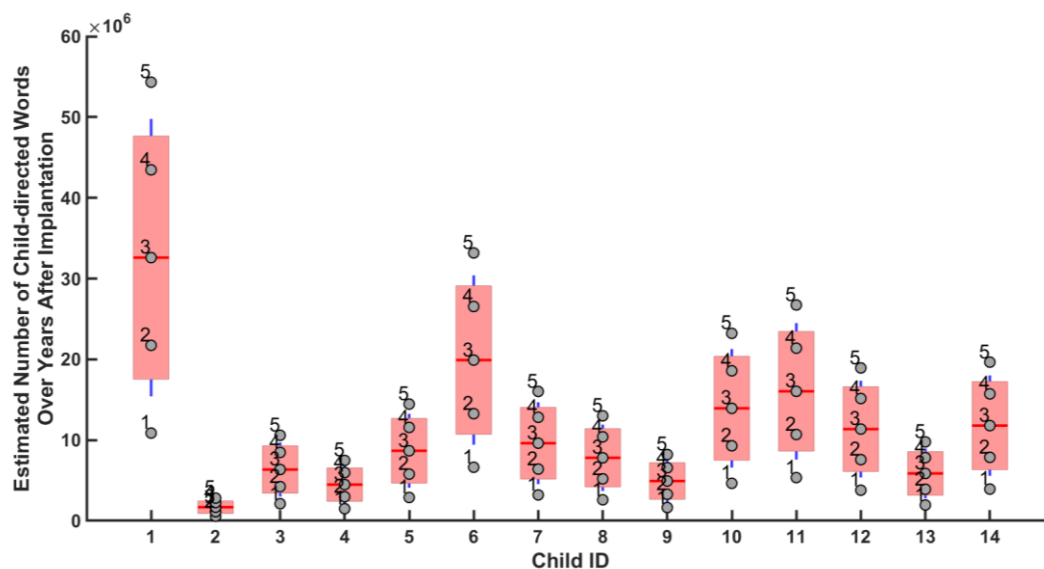


Figure 3-18. The number of child-directed words extrapolated over 5 years after CI activation for each child based on the traditional approach. Numbers 1-5 next to data point circles reflect estimated child-directed words after 1, 2, 3, 4, or 5 years.

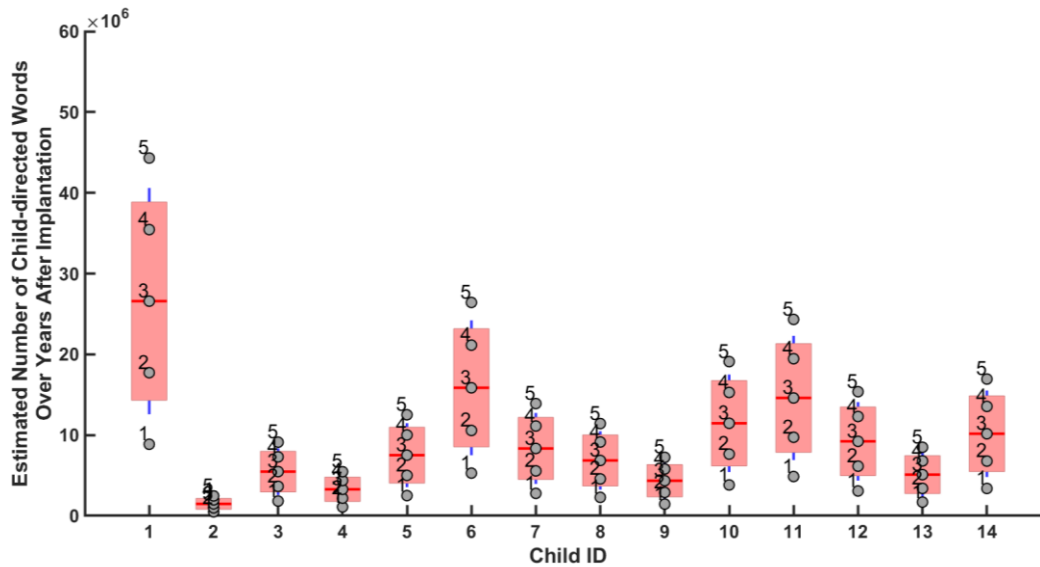


Figure 3-19. The number of child-directed words extrapolated over 5 years after CI activation estimated based on updated human auditory perceptual judgments of the level of noise interference. Numbers 1-5 next to data point circles reflect estimated child-directed words after 1, 2, 3, 4, or 5 years.

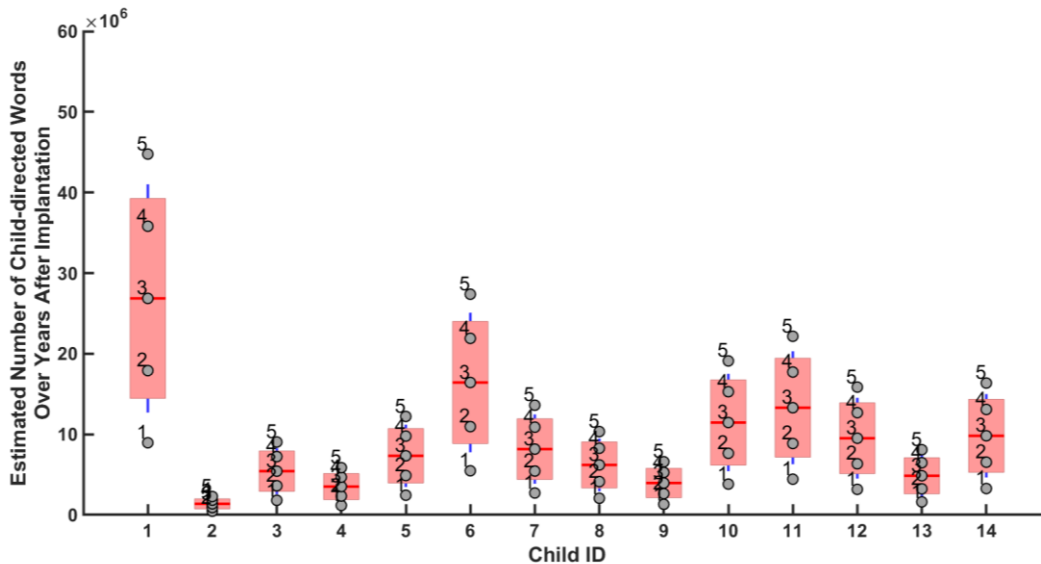


Figure 3-20. The number of child-directed words extrapolated over 5 years after CI activation estimated based on updated word counts based on SRMR-CI metric. Numbers 1-5 next to data point circles reflect estimated child-directed words after 1, 2, 3, 4, or 5 years.

The effect of noise degradation and speaking style on variability across children. Speech with the highest quality in the environments of children with cochlear implants is child-directed speech with no degradation (i.e., clean). Figure 3-21 summarizes the patterns of variability across individual children with CIs in the number of experienced words based on two factors of speaking style (child-directed vs. total speech) and environmental noise and reverberation. Human-based and SRMR-CI based approaches provide higher resolution for updating the number of useful words experienced by each child and the numbers compared to simply making a binary decision of clean vs. overlapped in considering whether a word exemplar is useful for a child for building vocabulary. These plots point out two important messages in characterizing the quality and quantity of early linguistic environments in children with CIs: (1) environmental noise and reverberation may greatly reduce the quality and quantity of early linguistic input experienced by children with CIs compared to children with normal hearing, as highlighted by word loss of around 9,000 words per day estimated for children with CIs compared to unadjusted word estimates (analogous to what a normal ear is expected to hear), and (2) the variability across children in the amount of high-quality words experienced per day remains substantial, even after updating the number of words based on the effects of noise and reverberation in children's auditory environments.

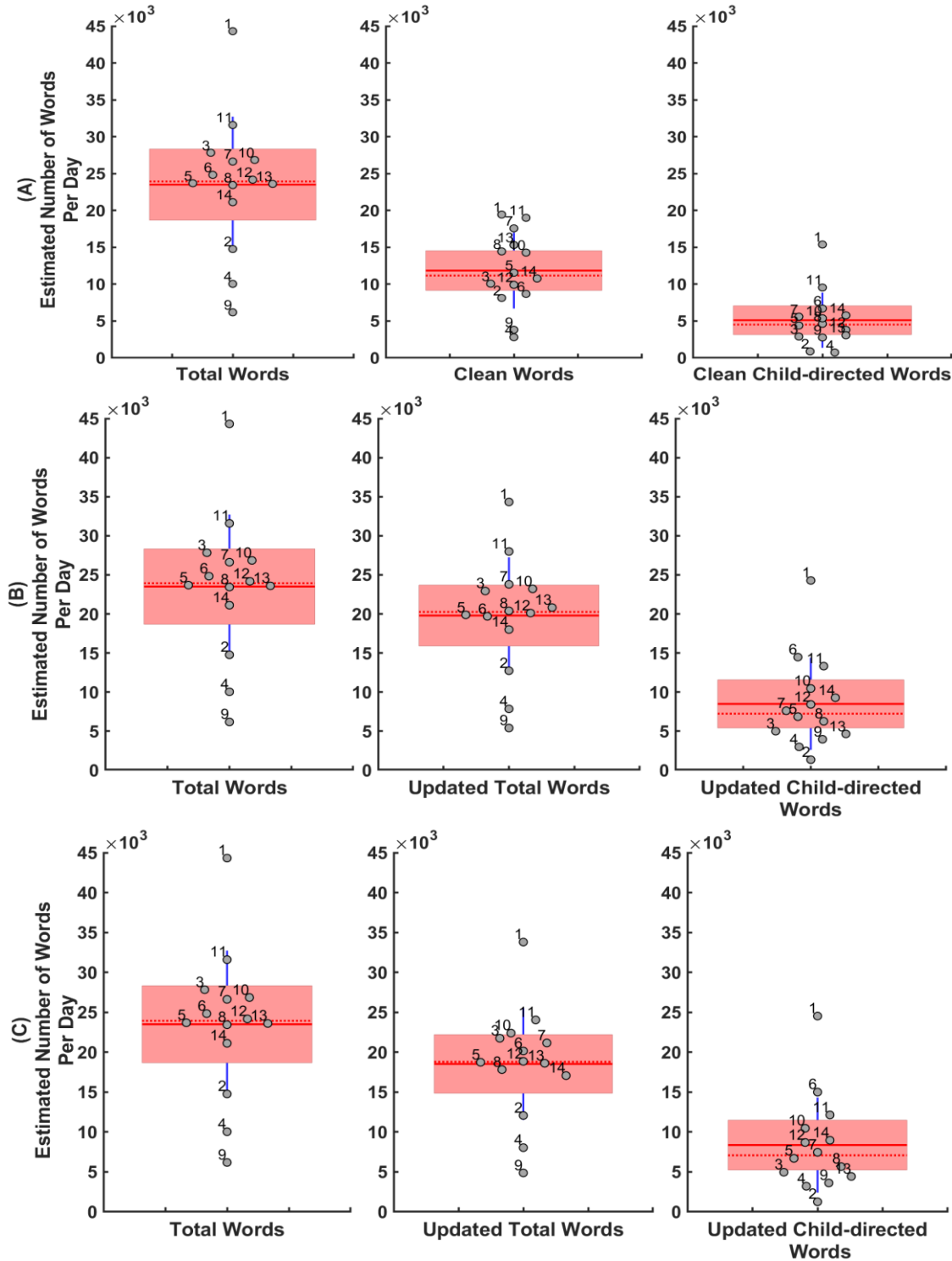


Figure 3-21. Distribution of children for the total number of words per day (first column), clean/updated number of words per day (second column), and clean/updated number of child-directed words per day (third column) they experienced in their environments, estimated based on (A) traditional, (B) human-based, and (C) SRMR-CI-based approaches.

3.3.2. Effect of noise on acoustic separability between child-directed speech and adult-directed speech

In Figure 3-22, the average value of *Mahalanobis distance* between child-directed speech and adult-directed speech intervals at each level of interference are shown, averaged across 14 children. This figure suggests an opposite pattern compared to our expectation, which had been for a higher level of noise interference to result in lesser acoustic distance between child-directed and adult-directed speech. Possible reasons for such an unexpected pattern will be discussed in the Discussion section.

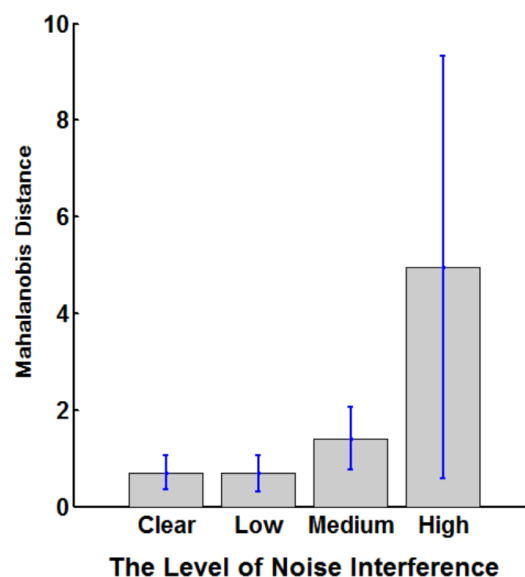


Figure 3-22. Mean and standard deviation of *Mahalanobis distance* between child-directed and adult-directed speech intervals at four levels of speech degradation.

3.3.3. Variability across children with cochlear implants in the profile of ambient noise in their linguistic environments

Variation in noise energy spectrum. Figure 2-23 shows the average spectral energy for 14 children calculated over snippets of noise identified within 30-second audio samples. The patterns of variation in spectral energy in these plots show that the average spectral distribution of noise is fairly similar across children's auditory environments. Further, average noise profiles reveal higher

amounts of average spectral energy in lower frequency regions (0-1 kHz) compared to higher frequency regions of the spectrum.

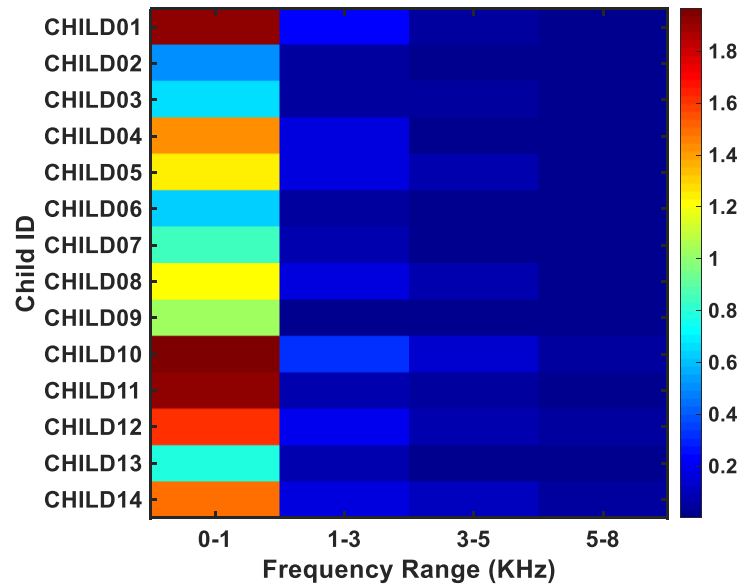


Figure 3-23. The average spectral energy of noise under four sub-bands experienced by 14 children with CIs.

Figure 3-24 shows the distribution of noise energy in environments of the 14 children as a scatterplot. The plot converges with the results shown in Figure 3-23 in suggesting that children differ from each other mainly in energy of low-frequency noise between 0-1 kHz.

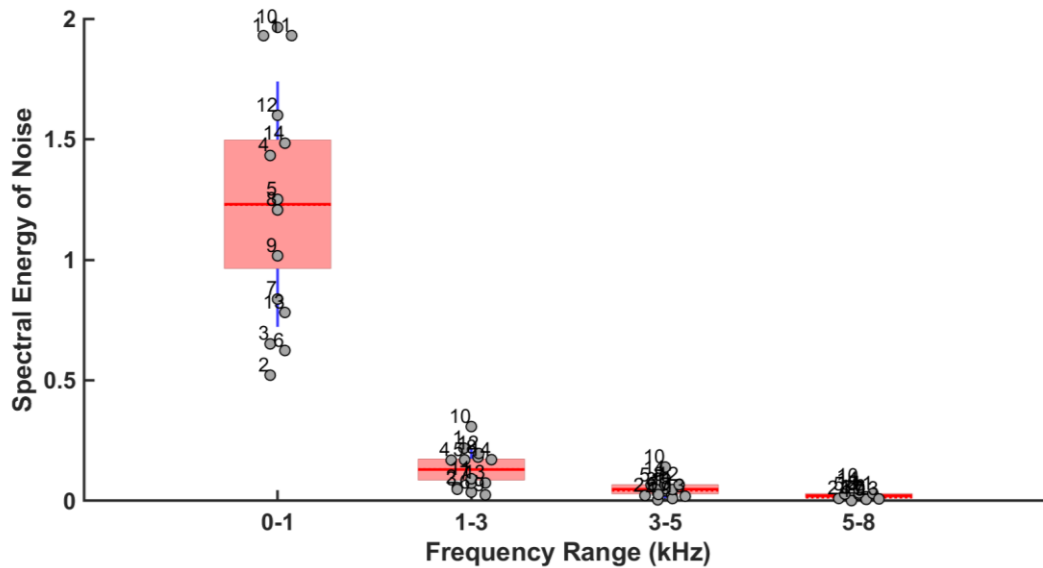


Figure 3-24. Distribution of children based on the experienced average spectral energy of noise under four sub-bands of 0-1, 1-3, 3-5, and 5-8 kHz.

Variation in noise loudness. Figure 3-25 presents a histogram of dBA values for frames of noise ranging from soft to loud in the environments of the children. As is clear, children experience noise with various loudness levels and durations. Some children, such as Child 13, were not exposed to loud noises very often, and most of the noise in his/her environment had less than 40 dBA loudness. Figure 3-26 shows the distribution of children in terms of exposure to noise measured by the noise exposure index and 90 percent noise exposure index for each child. The ninety percent noise exposure index was calculated by including only the frames for which the dBA loudness exceeded 90 dBA. Table 3-7 shows the measures of central tendency and variability for both noise exposure index and 90 percent noise exposure index. The values of standard deviation and IQR suggest that children vary largely in terms of exposure to noise in their auditory environments.

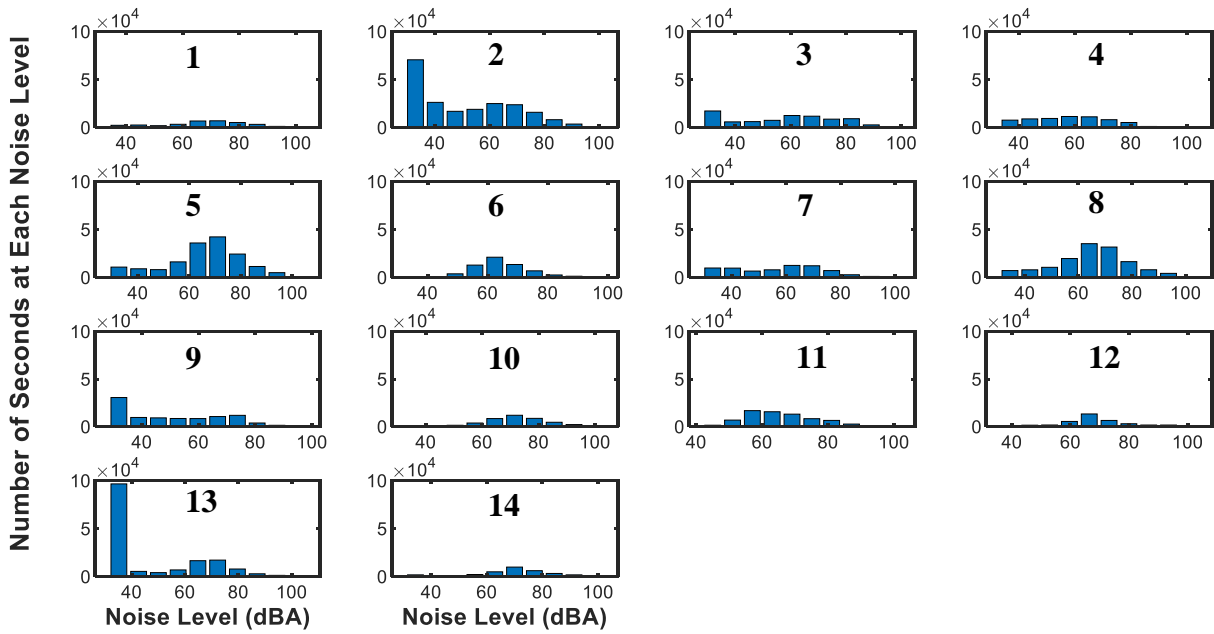


Figure 3-25. Histogram of noise loudness for each child.

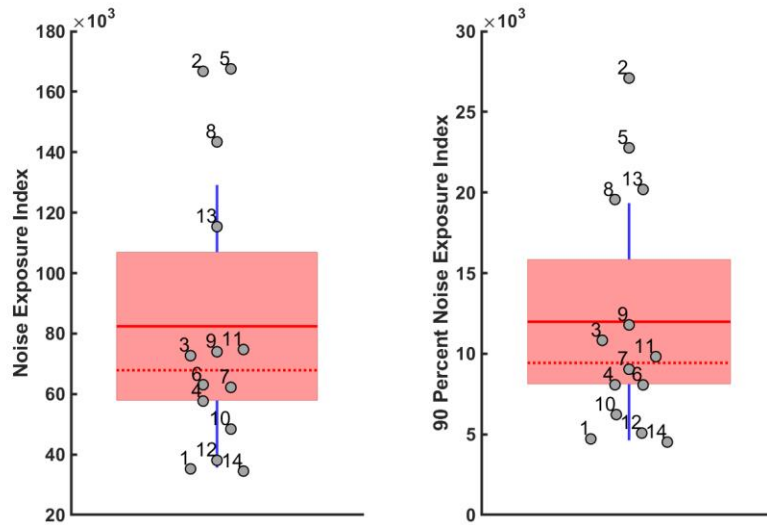


Figure 3-26. Variability across children in exposure to noise measured by (A) noise exposure index for all frames, (B) 90 percent noise exposure index, i.e., noise exposure index for noise frames with loudness of higher than 90 dBA.

Table 3-7. Measures of central tendency and variability for two measures of noise exposure: noise exposure index and ninety percent noise exposure index.

Measure of Noise Exposure	Measures of Variability						
	Min	Max	Range	Mean	SD	Median	IQR
Noise Exposure Index ($\times 10^3$)	34.5	167.5	133	82.4	46.8	67.8	66.978
90% Noise Exposure Index ($\times 10^3$)	4.5	27.1	22.6	11.9	7.4	9.4	13.3

Finally, the quality of exposure to noise for each child is quantified based on four descriptors that are commonly used to monitor the loudness and annoyance of environmental noises (Bockstael et al., 2011; Parmanen, 2007; Schomer et al., 2002). These descriptors are obtained from the histogram of dBA for each child. L_n is the noise level exceeded n% of the time. For instance, L_{50} is the median level that reflects the noise level that is exceeded 50% of the time during measuring the noise interval. Leq is the mean level of the noise. The difference between the mean level and the median level ($L_{50}-Leq$) is shown as an index of the direction and amount of skewness of noise loudness with respect to the softest and loudest noise level in the noise profile of each child (Figure 3-27). For example, for Child 1, $L_{50}-Leq = 1.6$ is positive, which means the distribution of noise loudness is negatively skewed, suggesting that child 1 experienced relatively high amounts of loud noise. On the other hand, $L_{50}-Leq$ for child 13 is negative with a value around -12.5, which means that the distribution of noise level is positively skewed for this child. Therefore, this child experienced a relatively healthy auditory environment in terms of low exposures to loud noise. Figure 3-28 shows variability across children in amount of loud noise they experienced in 5% audio samples. This figure suggests that children vary considerably in the amount of loud noise they experienced in their environments.

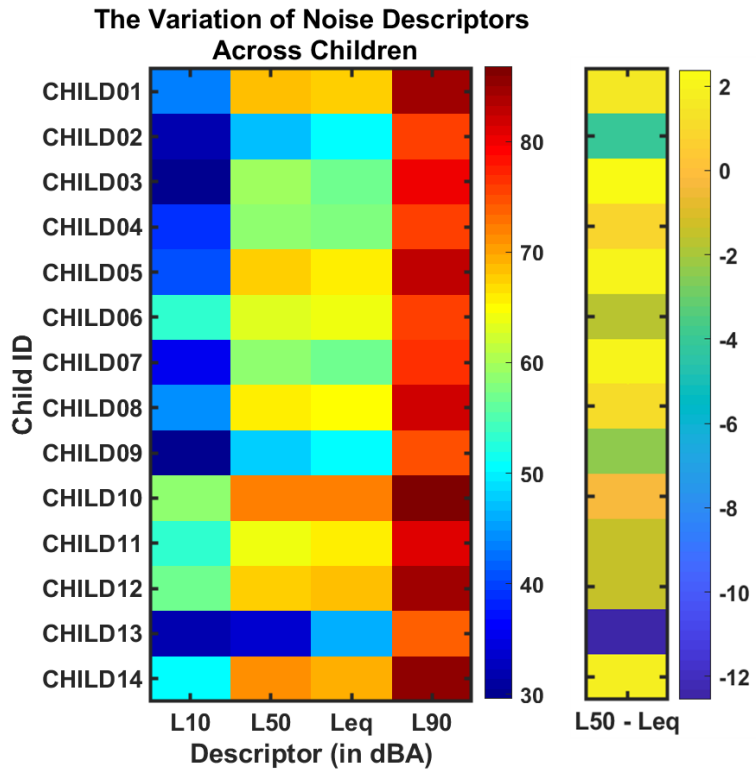


Figure 3-27. Noise descriptors for 14 children with CIs.

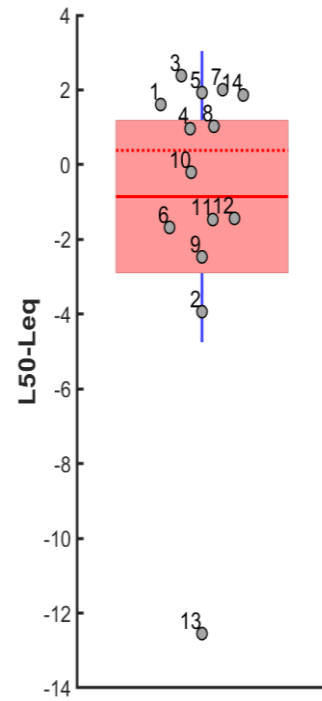


Figure 3-28. Distribution of children based on *L50-Leq* index.

3.4. Discussion

In the present study, audio recordings from linguistic environments of 14 prelingually deaf children with sensorineural hearing loss were analyzed to investigate how environmental noise and/or reverberation impact the estimates of number of useable words experienced by children in their environments. We further used the updated estimates of number of total words and child-directed words for each child to examine individual differences in early linguistic environments for these 14 early-implemented children.

The influence of environmental noise and reverberation on numbers of total words and child-directed words

The results from this study demonstrated that a significant number of both total words and child-directed words are likely lost in children's language environments due to background noise and/or reverberation. Updating the number of words based on human perceptual judgments of the level of noise interference showed a statistically significant decline of ~15.5%, on average, in number of useful words experienced by children with CIs per day in their environments. A similar degree of word loss per day (~21%) was obtained by updating the number of words experienced by children with CIs after adjusting numbers of useable words based on the SRMR-CI computational metric of speech intelligibility tailored to listeners with CIs. These results from analyzing speech input experienced by children with CIs in naturalistic environments corroborate prior findings on the large detrimental effect of noise and reverberation on word and phoneme recognition from lab and virtual classroom (Caldwell & Nittrouer, 2008; Caldwell & Nittrouer, 2013; Neuman et al., 2004), within a more ecologically valid investigation.

Differential effect of noise across children's language environment

The results from this study demonstrated large individual differences across children in the amount of total words lost in their environment due to noise and/or reverberation. The degree of variability across children in the amount of word loss per day was comparably large across updating approaches derived from human-based noise level interference ($SD \cong 2,000$ words loss per day, $mean \cong 3,700$ words loss per day) and the quantitative SRMR-CI metric ($SD \cong 2,200$ words loss per day, $mean \cong 4,900$ words loss per day). This pattern of individual differences in the amount of total word loss may be interpreted to suggest that some children are at higher risk of developing delayed receptive and expressive vocabulary skills compared to others. However,

there was a significant correlation between number of words per day obtained from traditional method and number of words per day obtained from each of the updated approaches (i.e., human-based and SRMR-CI-based) ($r = 0.99$, $p\text{-value} < 0.001$), suggesting that although individual children were differentially impacted by noise and reverberation, the final impact on the absolute number of words experienced were almost the same. The updated number of child-directed words experienced by each child per day based on SRMR-CI method is expected to provide the most accurate picture of the quality by which these children perceived the child-directed words in their environments compared to the traditional and human-based methods.

Although the differences among children in amount of word loss is large, variability across children in total words experienced remained sizable ($SD = 7,443$ words for human-based word adjustments and $SD = 7,014$ words for SRMR-CI-based word adjustments). Updating the number of words based on the effect of environmental noise and reverberation is expected to provide a better representation of the quality and quantity by which children experienced words in their environment. These results also showed that individual children were differentially impacted by the effect of noise and reverberation on the number words in their environment, which may conceivably explain an important portion of variability in language outcomes of these children.

The effect of noise on separability of child-directed speech from adult-directed speech

The current research used a signal-based approach to examine how the separability of child-directed speech from adult-directed speech is influenced by the level of noise interference with speech. Contrary to our initial expectation that the acoustic distance between child-directed and adult-directed speech would decrease as the level of noise interference increases, a higher level of background noise *increased* the acoustic distance between child-directed speech and adult-directed speech. Our expectation was that noise that overlaps with speech would disrupt acoustic

cues that are responsible to distinguish child-directed speech from adult-directed speech. This unexpected result might be due to multiple sources of variability that were not controlled in this study due to the nature of the study design. One of these uncontrolled factors was the speech context that was not necessarily the same across two categories of child-directed and adult-directed speech utterances. Another uncontrolled factor related to the variable spectral characteristics of background noise interfering with speech, which was not controlled due to the exploratory nature of this study. Although the level of noise interference was controlled at four levels of null/clean, low, medium, and high, the variability in spectral characteristics of noise could lead to higher distances between MFCCs of the two classes of child-directed speech and adult-directed speech samples when the level of noise was higher, as the results suggest. Future perceptual and computational modeling studies with careful design will elaborate our understanding of how various levels of noise may impact separability of child-directed speech from adult-directed speech. Being able to acoustically separate child-directed speech from adult-directed speech is the preliminary step for children to prefer child-directed speech over adult-directed speech (Yuanyuan Wang, Bergeson, et al., 2018).

Variability across children in the exposure to noise

We also analyzed the non-speech regions in auditory environments of these 14 early-implemented children to characterize only environmental noise and identify the extent to which children with CIs are differentially exposed to noise. Children clearly differed descriptively in the extent to which they were exposed to different sources of noise in their environments. However, our results suggest that children are not largely different in terms of the spectral characteristics of noise in their environment when focusing on the variation of spectral energy of noise in the

proposed four frequency bands. Specifically, our results suggested that children principally differed in the extent to which they were exposed to low-frequency (< 1 kHz) energy.

Further, we examined how children varied in the amount of time they were exposed to loud noises in their environments. Two measures of noise exposure index and *L50-Leq* index were calculated to assess the extent of difference between children in their auditory diets. Our evaluation of the amount of loud noise in the environment of children showed that children varied widely in exposure to noise, ranging from as small as noise exposure index of 34.5×10^4 for some children to as large as noise exposure index of 133×10^4 for others with a large standard deviation of 46.8×10^4 in a day. The same large amount of individual difference in noise exposure index was shown when only noise frames with loudness of higher than 90th percentile were included for noise exposure index calculation ($range=22.6 \times 10^4$ and $SD= 7.4 \times 10^4$). We further assessed the annoyance of environmental noise for each child based on four commonly used descriptors of *L10*, *L50*, *Leq*, and *L90* (Bockstael et al., 2011; Parmanen, 2007; Schomer et al., 2002). The variability of *L50-Leq* as an index of the direction and amount of skewness of noise loudness with respect to the softest and loudest noise level showed that children vary largely in exposure to noise in their auditory environments. The variability of this index was quantified by standard deviation of 3.9 around mean of -0.85. This large variation among children in their auditory diets may contribute to the observed variability in language outcomes of these children with CIs. Therefore, exposure to sudden high-density noise or chronic noise with high energy may add to other challenges that children with CIs may face in their auditory environments, potentially contributing to differences among children in language and cognitive development.

Overall, the results from this study have provided new evidence of how the quality and quantity of language experienced by children with CIs during the first year after implantation

significantly decreases due to background noise and/or reverberation in their environments. Based on rigorous analysis of samples of audio recorded from the auditory environments of 14 children with CIs, we showed that the number of total words and child-directed words experienced by children with CIs in a day significantly decreased because of the disrupting effects of background noise and/or reverberation on speech recognition – as modeled by two word count updating approaches: (1) human perceptual judgments of levels of noise interference, and (2) a quantitative measure of speech intelligibility tailored to listeners with CIs which is called the speech-to-reverberation modulation energy ratio for CI users (SRMR-CI) (Falk et al., 2015; Santos & Falk, 2014). These findings corroborate prior findings showing children with CIs are extremely susceptible to noise and reverberation in recognition of speech at multiple levels of granularity from phonemes, to words, to utterances (Caldwell & Nittrouer, 2013; Iglehart, 2009; Neuman et al., 2004).

This study marks one step further in comparing to prior studies in findings of sources of variability in language outcomes of children with CIs. However, we have shown the extent to which the number of total words and high-quality child-directed words are influenced by noise and reverberation in naturalistic environments. Furthermore, the results from this study suggested that children with CIs are differentially impacted by the amount of noise and reverberation in their environments that co-occurred with speech. These results suggest that some children are relatively at higher disadvantage compared to others due to degradation imposed by noise and reverberation in their environments.

We also demonstrated that early language environments of children with CIs are largely variable across children in amounts of loud environmental noise in a day, which suggests that distinct profiles of noise in the auditory environment may be a contributing factor for individual

variability in language outcomes of children with CIs. The findings from this study highlight the need for monitoring daily auditory environments of children with CIs for the amount of high-quality child-directed speech and exposure to environmental noise and reverberation. These results provide novel evidence for parents, caregivers, and clinicians to pay careful attention to environments of their children, particularly during early childhood after cochlear implantation when children's brain is highly plastic and sensitive to learn speech units (i.e., phonemes, syllables, words, and sentences) and rules (i.e., semantic, phonotactic, and syntactic).

Limitations of this study

This study is restricted by the same limitations discussed in the discussion of Chapter 2. In addition to those limitations, there are some limitations with the computational approaches that are proposed in this study to update the number of words based on their co-occurrence with noise and reverberation in the environment. In particular, to translate the human coding for level of noise interference into updates of the number of words, a simple mapping function was proposed that relied on an assumption of linear and equal distance between low, medium, and high levels of noise interference. Further auditory perceptual studies should be planned to investigate the validity of this assumption.

Another limitation of these analyses related to lack of robustness of the SRMR-CI metric to sources of artifacts in the recorded audio. This metric of speech intelligibility has been only tested on controlled speech stimuli from standard datasets. We used this metric to analyze audio environments of children with CIs recorded by LENA device under the major assumption that this metric would be robust to various artifacts related to recording device and environment. An important line of research for the future is to enhance this metric to be reliable for analyzing audio signals recorded from naturalistic environments. In addition, our noise-adjustment approach

included two other steps following calculation of a SRMR-CI metric. In the first step, raw metrics of SRMR-CI were scaled to 0 and 1. Several scaling function were tested to address inconsistent reports of the type of scaling used in the relevant papers (Falk et al., 2015; Santos, Cosentino, Hazrati, Loizou, & Falk, 2013; Santos & Falk, 2014). The *minmax* formula was similar to the approach in these papers and provided sensical results. The second step involved constructing a sigmoidal function, which was used to map an SRMR-CI value to its corresponding speech intelligibility percentage value, which depended on the distribution of SRMR-CI values for speech utterances categorized as clean speech by human analysts. Therefore, because SRMR-CI is a signal-based approach, its value for utterances in clean category might be affected by uncontrolled factors of the natural environments such as distance of microphone from talkers. The second step involved updating the number of words by simply multiplying numbers of words in a speech utterance by the percent speech intelligibility. However, this mapping has not yet been validated and requires further investigation in the future.

Furthermore, although we used the results of inter-rater reliability on the amount of agreement between human analysts to check all the noise coding, coding the level of noise might still have been negatively impacted by the performance of some coders. This was due to this study's limitation in having access only to audio, not video, recordings, leading to increased ambiguity for coders about sources of noise.

Despite these limitations and the exploratory nature of this study, the study is nevertheless conceptually and methodologically novel in terms of focusing on an unexplored and crucial aspects of early linguistic input. In particular, this is the first study to examine to estimate effects of noise and reverberation on estimates of the amount of words experienced by children with CIs in natural environments. We have demonstrated how children with CIs may be doubly disadvantaged in

experiencing high-quality language input: both due to degradation associated with electronic hearing and environmental noise and reverberation as well as due to substantial variability in the amount and quality of linguistic input available to them. The results from this study have highlighted the great difficulties that children with CIs face in understanding speech in naturalistic environments. As an at-risk population, children with CIs face the task of building a mental lexicon from heard exemplars of words which are especially susceptible to signal degradation due to environmental noise and reverberation. The findings from this study have provided evidence for clinicians, parents, and caregivers that early linguistic environments of children with CIs are highly complex. Our results highlight that effective intervention strategies will likely require developing means of reducing the detrimental impact of noise and reverberation on children's language learning and cognitive development.

CHAPTER 4: Toward understanding effects of quality and quantity of early language in linguistic environments in children with CIs on language outcomes

4.1. Introduction

For children with typical hearing, prior studies have established that the quality and quantity of language input early in childhood has an important impact on their later language outcomes. However, the extent to which quality and quantity of language input to children with CIs plays a role in shaping their language outcomes is unclear. The purpose of the study presented in this chapter was to begin to establish links between differences in early language input for a small sample of children with CIs and the language outcomes in this population. In so doing, the analyses presented in this chapter aim to provide an initial test of the hypothesis that language input factors shape development of language outcomes in children with CIs. An exploratory approach was therefore taken in this chapter which was intended to investigate the extent to which variability in language input to a sample of children with CIs, as indexed by variability measures developed in Chapters 2 and 3, could conceivably account for variability in a set of outcome (or predicted) variables, as measured by standard clinical speech-language measures. These statistical relationships were examined to determine the strength of correlations between predictor variables and predicted variables for the small set of children with CIs examined in this dissertation project.

Background: Review of language input factors which affect language outcomes in children with typical hearing

Prior chapters discussed several factors that have been identified to contribute to large individual differences in speech and language outcomes in infants with normal hearing (Hart &

Risley, 1995; Hoff-ginsberg & Hoff-ginsberg, 1985; Hoff, 2003, 2006) as well as those with CIs (Connor et al., 2000; Geers, Hayes, & Louis, 2009; Geers, Nicholas, & Moog, 2007b; Niparko & Tobey, 1989; Peterson et al., 2010; Svirsky et al., 2004b; Szagun, 2001; Szagun & Stumper, 2012; Tomblin et al., 2005). Among these factors, very few studies have provided preliminary evidence that quality and quantity of early linguistic environment plays a crucial role in the development of speech and language in children with CIs. To the extent that this has been tested, studies have primarily focused on samples of continuous speech recorded in constrained situational settings, rather than actual linguistic environments of children with CIs (Szagun & Schramm, 2016; Szagun & Stumper, 2012). Therefore, there is limited knowledge about the relationship between aspects of early linguistic input and language outcomes of this at-risk population based on naturalistic data. To begin to understand these relationships, in this chapter's analyses we used the language measures derived in Chapters 2 and 3 to investigate the relationship between lexical, morphosyntactic and social-pragmatic aspects of early language input and language outcomes of the sample of early-implanted children with CIs under study.

Multiple factors associated with variability in language input have been identified previously as likely influencing language outcomes in typically-hearing children. As discussed in Chapter 2 and 3, child-directed speech is one of the key components of early linguistic environments that enhances development of lexical, semantic, and syntactic knowledge (Barnes et al., 1983; Huttenlocher et al., 1991; Szagun & Stumper, 2012) and improves children's cognitive processing skills (Bradley & Caldwell, 1976a, 1976b; Clarke-Stewart, 1973; McCartney, 1984). Experiencing linguistic environments with a relatively higher amount of child-directed speech facilitates language learning through exposing children to more diverse vocabulary (Hart & Risley, 1995; Hoff & Naigles, 2002), higher instances of word repetition (Hoff-Ginsberg & Shatz, 1982),

and more instances of semantically coherent utterances (Golinkoff et al., 2019; Hoff-Ginsberg, 1991, 1994). Collectively, these prior studies suggest that children benefit from child-directed speech due to its high-quality lexical, morphosyntactic and social-pragmatic information. Child-directed speech is therefore expected to be particularly crucial for children with CIs who may face greater challenges in listening and understanding overheard speech that is not spoken directly to them.

Consistent with the likely advantages of child-directed speech for children with CIs, a few studies have begun to show that children with CIs who experience language input with higher quality and quantity early in childhood showed better language development (DesJardin & Eisenberg, 2007; Fagan, M. K., Bergeson, T. R., & Morris, 2014; Szagun & Schramm, 2016; Szagun & Stumper, 2012). In particular, Szagun and Stumper (2012)'s study on 25 early-implanted children with CIs showed that children's MLU or expansions at 24 and 30 months post-implantation were predicted by maternal MLU or expansions at 12 months and 18 months after implantation. A large portion of the variance in children's MLU (between 21% and 44%) was explained by maternal MLU. The same pattern was reported for maternal expansion, which accounted for 19% to 25% of variance in children's MLU. Notably, the effect of the quality of maternal input on children's language outcomes was greater than the effect of age at implantation. A similarly strong association between maternal input and children's language outcomes was reported by Szagun and Schramm (2016). However, these two studies focused primarily on differences in grammatical structure of language input (Szagun & Schramm, 2016; Szagun & Stumper, 2012) and parental style (DesJardin & Eisenberg, 2007) and were not based on analyzing the actual home environments of children with CIs. These studies also did not address how

disassociating the effect of language input based on child-directed versus overheard speech might influence contribution of various properties of language input on children's language outcomes.

Background: Factors of language input which may uniquely affect language development in children with CIs

An innovation of the present study involved considering in detail the potential impacts of ambient noise and reverberation on language development in children with CIs. For children with CIs, the amount of noise and reverberation experienced while listening to speech is expected to be a factor that might dramatically impact language development, by affecting the quantity and quality of speech-language input signals. In particular, noise and reverberation might reduce the quantity of words accessible as language input to children with CIs, because words which occur in the presence of substantial noise and reverberation may be too degraded to be understood as an instance of the spoken word. A focus of the present dissertation was therefore to consider the possibility that *noise-adjusted word counts* – estimates of the numbers of total words or child-directed words that were minimally impacted by noise – might better predict language outcomes than raw, traditional word counts which did not consider signal degradation due to noise.

Complementary research supports the premise that noise and reverberation significantly affect perception of audio signals – including speech – experienced by children with CIs. Children with CIs face several challenges including appreciating music, listening and talking over cellphones, understanding speech in noisy and reverberant conditions. Among these challenges speech recognition in noisy and/or reverberant conditions is the most dangerous factor that may impact their language outcomes and contribute to individual differences (Caldwell & Nittrouer, 2008; Neuman et al., 2004; Neuman, Wroblewski, Hajicek, & Rubinstein, 2010b). This issue is of substantial importance, considering that users of CIs spend significant portions of their daily lives

in noisy environments (Busch et al., 2017). In one study (Caldwell & Nitttrouer, 2013), normal hearing children outperformed children with CIs by about 25% and 16% on recognizing phonemes and words, respectively, in noisy conditions. Listeners with CIs have higher difficulties in understanding speech in noise compared to understanding soft speech (Firszt et al., 2004). Natural auditory environments present highly complex conditions involving background noise and/or reverberation which can add to challenges of recognizing soft speech. Therefore, characterization of quantity and quality of language input without considering effects of noise and reverberation will present an incomplete representation of what children with CIs experience in processing and learning speech and language.

Current study

Building on prior findings with typical hearing children and children with CIs, this study aimed to provide preliminary results by exploring the potential impacts of multiple variables associated with linguistic environments on standard scores of language outcomes in a small group of children with CIs. The relation between characteristics of early linguistic input and language outcomes were explored using standard statistical analysis to gain an initial understating of the contribution of these factors on children's language outcomes. We will also present preliminary results of how noise-adjusted total word counts and amount of noise-adjusted child-directed words may affect prediction of language outcomes.

The following research questions were addressed in the present study:

Question 1: Do variability in the *total words per day* or *child-directed words per day* experienced by children with CIs predict language outcomes (from clinical speech-language assessments), without regard to noise and reverberation?

We hypothesized that variability in *child-directed words* will predict variability in language outcomes. It was also hypothesized that *child-directed speech* provides better language input to foster language development in children with CIs than *total speech*. We hypothesized that child-directed speech provides better language input to foster language development in children with CIs than overheard speech. This led to the prediction that variability in *child-directed words* would better predict language outcomes than variability in *total words*.

Question 2: Do *noise-adjusted estimates* of words per day (total or child-directed) predict outcomes better than traditional estimates that do not consider degradation due to noise and reverberation?

We hypothesized that *noise-adjusted estimates of child-directed words* provide the best prediction of language outcomes.

Question 3: What is the relative contribution of lexical, morphosyntactic, and social-pragmatic measures of language input on language outcomes of children with CIs at 6 months post-implantation?

We hypothesized that children who experienced richer lexical, morphosyntactic, and social-pragmatic language environments early in childhood would show better language outcomes.

Question 4: What is the best way of taking into account noise and degradation to language input – human-based judgments or SRMR-CI estimates?

We hypothesized that SRMR-CI will better represent the effect of noise and reverberation on speech intelligibility compared to human-based judgments.

4.2. Methods

4.2.1. Participants

Thirteen children from the total of fourteen children with CIs studied in prior chapters were further studied in this chapter. Language outcomes were not available for *Child 14*; therefore, this child was excluded from the analyses of this chapter.

4.2.2. Measures of language input (i.e., predictor variables)

The general approach taken in this chapter was to compute statistically the strength of predictive relationships between various predictor measures described in Chapters 2 and 3 and predicted variables, namely, language outcome measures determined from the *Preschool Language Scale-Fifth Edition (PLS-5; Zimmerman, Steiner, & Pond, 2011)*. Table 4.1. lists the measures of quality and quantity of linguistic input and language outcomes studied in this chapter to investigate how characteristics of early linguistic input of children with CIs were related to their language skills early after implantation.

Measures of language input for traditional and noise-updated measures were calculated in Chapters 2 and 3 to characterize early linguistic environments of children with CIs participating in this study. In Chapter 2, we measured the number words, type-token ratio and curves, MLU, caregiver responsiveness, and speech rate to characterize early linguistic environments of children with CIs. Chapter 3 addressed a major gap in understanding language learning in children with CIs by updating the number of words (both total and child-directed) by considering effects of noise and reverberation in the environment.

Table 4-1. Measures of quality and quantity of language input and language outcomes of the 13 children available for this study. Measures of language input were derived from analyzing the audio samples recorded from the natural linguistic environments of children. Measures of language outcomes were obtained by administering PLS5 tests.

Measures of Language Input	Measures of Language Outcomes
(1) Total words per day ¹ (2) Child-directed words per day ¹ (3) Noise-adjusted total words per day (<i>estimated from human coding</i>) ² (4) Noise-adjusted child-directed words per day (<i>estimated from human coding</i>) ² (5) Noise-adjusted total words per day (estimated from transfer function from SRMR-CI to speech intelligibility) ² (6) Noise-adjusted child-directed words per day (estimated from transfer function from SRMR-CI to speech intelligibility) ² (7) Type-token ratio from total words ¹ (8) Type-token ration from child-directed words ¹ (9) Area under the type-token curve from total words ¹ (10) Area under the type-token curve from child-directed words ¹ (11) Mean length of utterance from total words ¹ (12) Mean length of utterance from child-directed words ¹ (13) Speech rate from total words ¹ (14) Speech rate from child-directed words ¹ (15) Caregiver responsiveness ¹ (16) SRMR-CI ²	(1) Pre-school Language Scale–Fifth Edition-Auditory Comprehension (PLS5-AC) (2) Pre-school Language Scale–Fifth Edition-Expressive Communication (PLS5-EC)

¹ These measures of language input were obtained in Chapter 2.

² These measures of language input were obtained in Chapter 3.

4.2.3. Measures of language outcomes (i.e., predicted variables)

Language outcomes of the 13 early-implanted children were measured using the Preschool Language Scale-Fifth Edition (PLS5; Zimmerman et al., 2011), which is a standardized test to assess children’s receptive and expressive language skills. This test was administered at 6 months post-implantation.

The Preschool Language Scale-Fifth Edition (PLS-5): PLS-5 was administered to assess the expressive and receptive language skills of the children at 6 months after cochlear implantation. It is a standardized test that is individually administered to evaluate children’s expressive and receptive language skills. Items address children’s pre-verbal skills, interaction-based skills, as well as their language milestones until early literacy. The test is designed for administration to ages ranging from birth to 7 years and 11 months (Zimmerman et al., 2011). The chronological

age of children in the present study at 6 months post-implantation are given in table 2-1. This test is commonly used for testing language skills in children with hearing loss (Fitzpatrick, Durieux-Smith, Eriks-Brophy, Olds, & Gaines, 2007; Geers et al., 2009). It has two components, Auditory Comprehension (PLS-AC) and Expressive Communication (PLS-EC), with standard scores to assess children's language perception and production. The test was administered at Ohio State University by certified speech-language pathologists with extensive experience in assessing children with hearing loss.

The PLS-5 standard scores were available for 13 children at 6 months post-implantation. We assumed that the per-day estimates of language measures calculated in Chapter 2 and 3 can be generalized to the entire first year after implantation as proposed in prior studies (Hart & Risley, 1995; Shneidman et al., 2013; Weisleder & Fernald, 2013). This assumption was based on the notion that there is no evidence that the characteristics of language spoken to children with CIs vary systematically and significantly over the first year after they receive their implantation. Therefore, a per-day estimate derived at 6 months or later can be a reasonable basis for estimating how much speech was spoken in the household prior to six months. Further studies in the future with more participants and recordings can provide more evidence on whether such an assumption is valid.

4.2.4. Statistical analysis

A series of simple linear regression analyses was conducted to examine the relative contribution of each measure of language input to language outcomes. The results from conducting the *Anderson-Darling* test showed that measures of language input were from a population with a normal distribution. We also used Pearson correlation coefficients to create correlation matrices for all pairs of measures of language input and language outcomes to simply examine how well

each of 16 measures of language input predict the two standard scores derived from PLS-AC and PLS-EC.

4.3. Results

4.3.1. To what extent are the total number of words and number of child-directed words (unadjusted for noise and reverberation) associated with language outcomes?

Figure 4-1 shows four plots depicting the relationship between total number of words and outcome measures (top plots) and number of child-directed words per day and outcome measures (bottom plots). The analyses showed that none of these correlations were statistically significant.

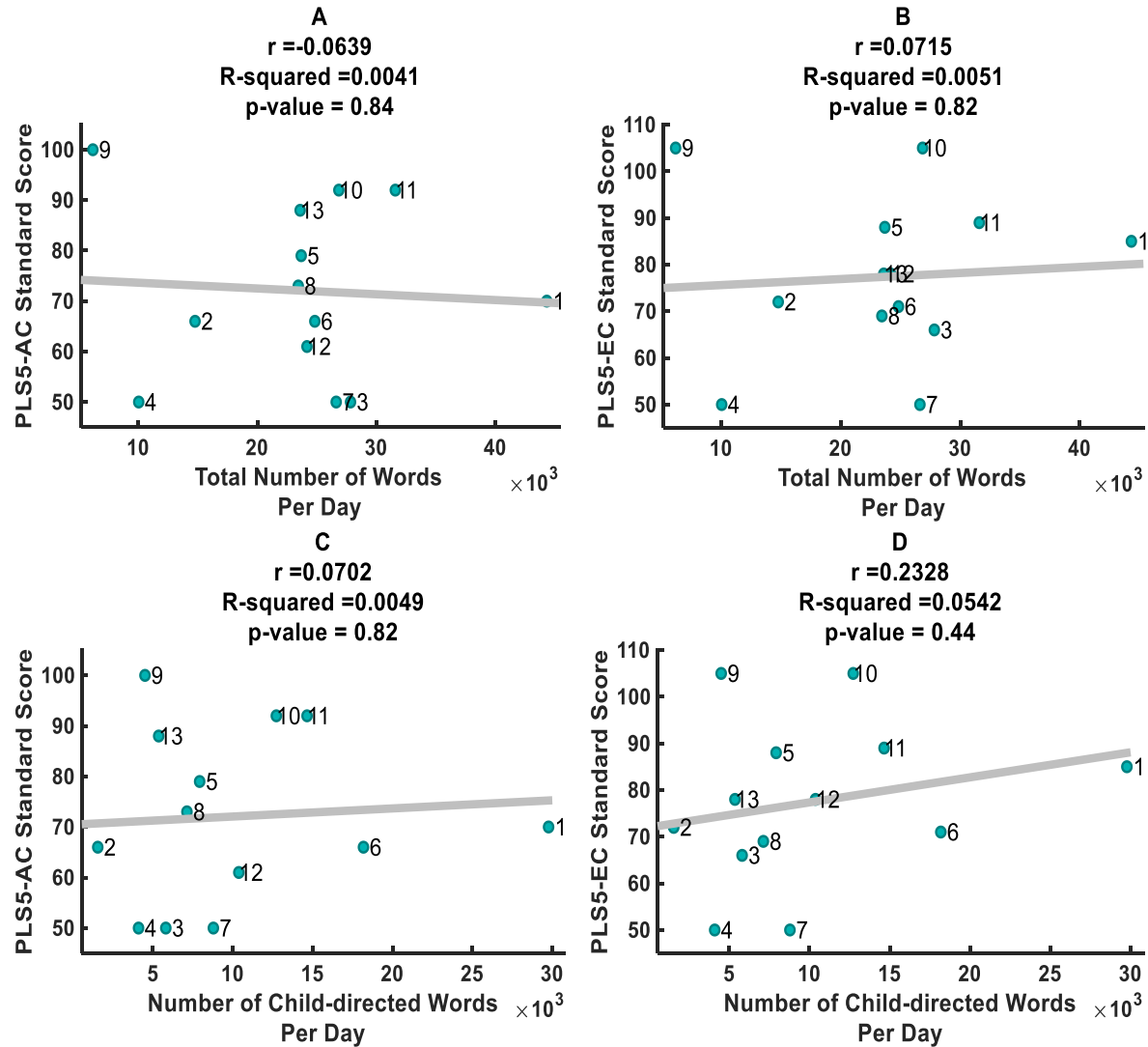


Figure 4-1. The relationship between total number of words and (A) PLS-AC and (B) PLS-EC standard scores, as well as the relationship between number of child-directed words and (C) PLS-AC and (D) PLS-EC standard scores. The r values are pairwise Pearson correlation coefficients between pairs of language input measure and language outcome. R-squared values are the coefficient of determination obtained from fitted ordinal linear regression models. The least-square regression lines are superimposed on each scatter plot.

Several simple linear regression analyses were also performed to investigate how well traditional measures of language predicted PLS-AC and PLS-EC scores at 6 months post-implantation. These measures of language input were type-token ratio, area under the type-token curve, mean length of utterance, caregiver responsiveness, and speech rate. Tables 4-2 and 4-3 show that there were no statistically significant correlations among these variables.

Table 3-2. The pairwise correlation coefficient, R-squared values, and *p*-values for the fitted simple linear regression models for examining the relationship between each language input variable derived from *total speech* and each of two variables of language outcomes.

Measure of Language Input	PLS5-AC			PLS5-EC		
	r	R-squared	p-value	r	R-squared	p-value
TTR	0.0184	0.0003	0.95	0.0471	0.0022	0.88
AUTTC	-0.0025	0.0	0.99	0.0501	0.0025	0.87
MLU	-0.3143	0.0988	0.3	-0.3546	0.1271	0.23
Caregiver Responsiveness	0.02	0.0002	0.96	0.2	0.0403	0.51
Speech Rate	-0.0794	0.0063	0.8	0.011	0.0001	0.97

Table 3-3. The pairwise correlation coefficient, R-squared values, and *p*-values for the fitted simple linear regression models for examining the relationship between each language input variable derived from *child-directed speech* and each of two variables of language outcomes.

Measure of Language Input	PLS5-AC			PLS5-EC		
	<i>r</i>	<i>R-squared</i>	<i>p-value</i>	<i>r</i>	<i>R-squared</i>	<i>p-value</i>
TTR	0.224	0.502	0.46	-0.279	0.0779	0.36
AUTTC	0.0371	0.0014	0.9	0.177	0.0313	0.56
MLU	0.3323	0.1104	0.27	-0.3312	0.1097	0.27
Caregiver Responsiveness	-	-	-	-	-	-
Speech Rate	0.0103	0.0001	0.97	0.1187	0.0141	0.7

4.3.2. To what extent do the updated total number of words and number of child-directed words, adjusted for noise based on human coding, predict language outcomes?

Figure 4-2 shows measures of total number of words and number of child-directed words per day adjusted based on noise interference from human judgments plotted against language outcomes as measured by PLS-AC and PLS-EC. There was no significant relationship between any of these language input measures and either of the language outcome measures.

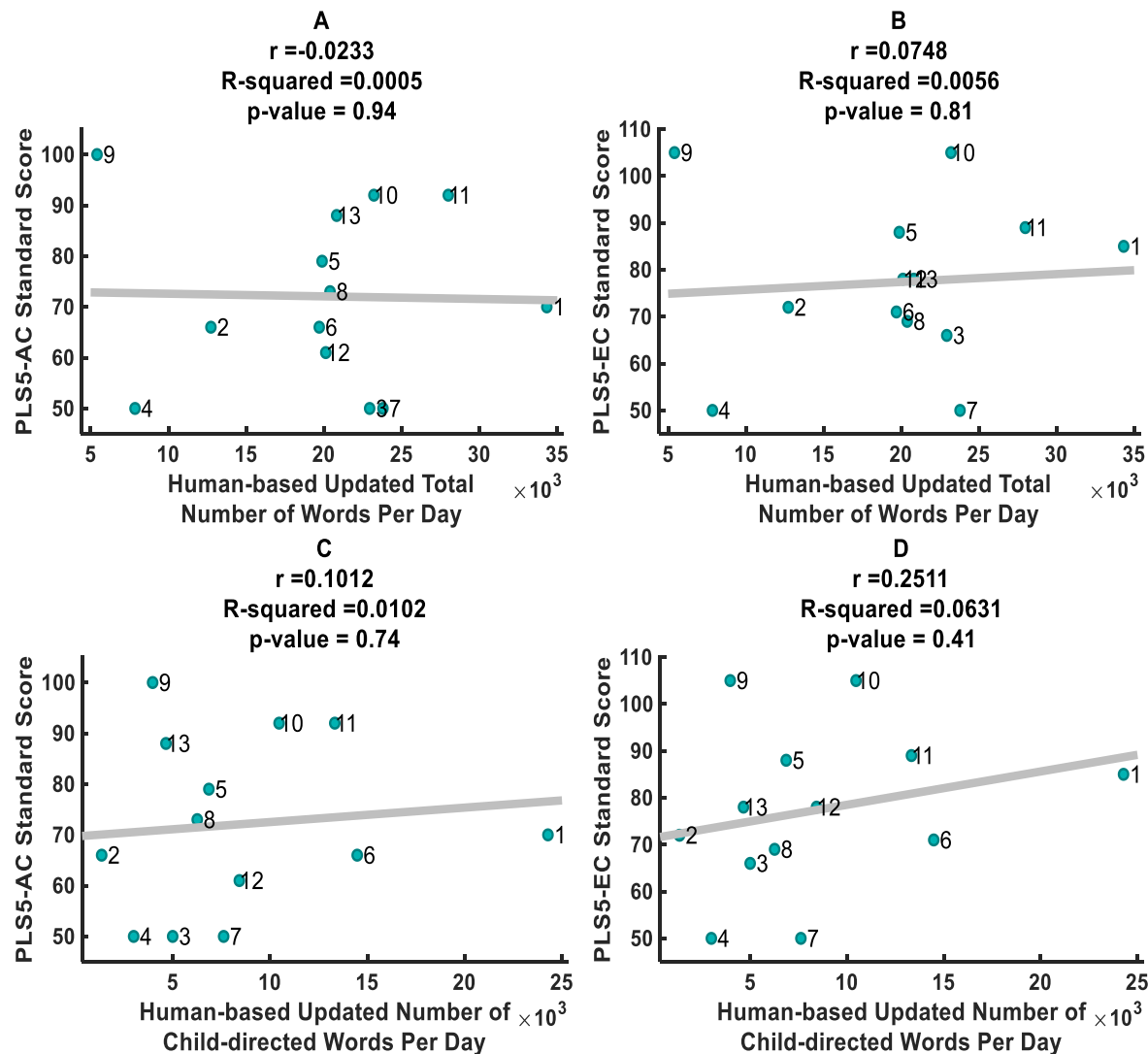


Figure 4-2. The relationship between updated total number of words based on human coding of level of noise interference and (A) PLS-AC and (B) PLS-EC standard scores, and human-based updated number of child-directed words and (C) PLS-AC and (D) PLS-EC standard scores. The r values are pairwise Pearson correlation coefficients between pairs of language input measure and language outcome. $R\text{-squared}$ values are the coefficient of determination obtained from fitted ordinal linear regression models. The least-square lines are superimposed on each scatter plot.

4.3.3. To what extent does the updated total number of words and number of child-directed words (adjusted for noise and reverberation based on SRMR-CI metric) predict language outcomes?

Environmental noise and/or reverberation can greatly interfere with speech recognition in listeners with CIs, and therefore likely interfere with the language learning process. Figure 4-3

shows the relationship between number of words and number of child-directed words updated based on SRMR-CI metrics and PLS-AC and PLS-EC standard scores. This figure suggests that there was no significant correlation between updated number of words based on the SRMR-CI metric and either PLS-AC or PLS-EC standard scores. The updated number of child-directed words also did not predict expressive and receptive language outcomes.

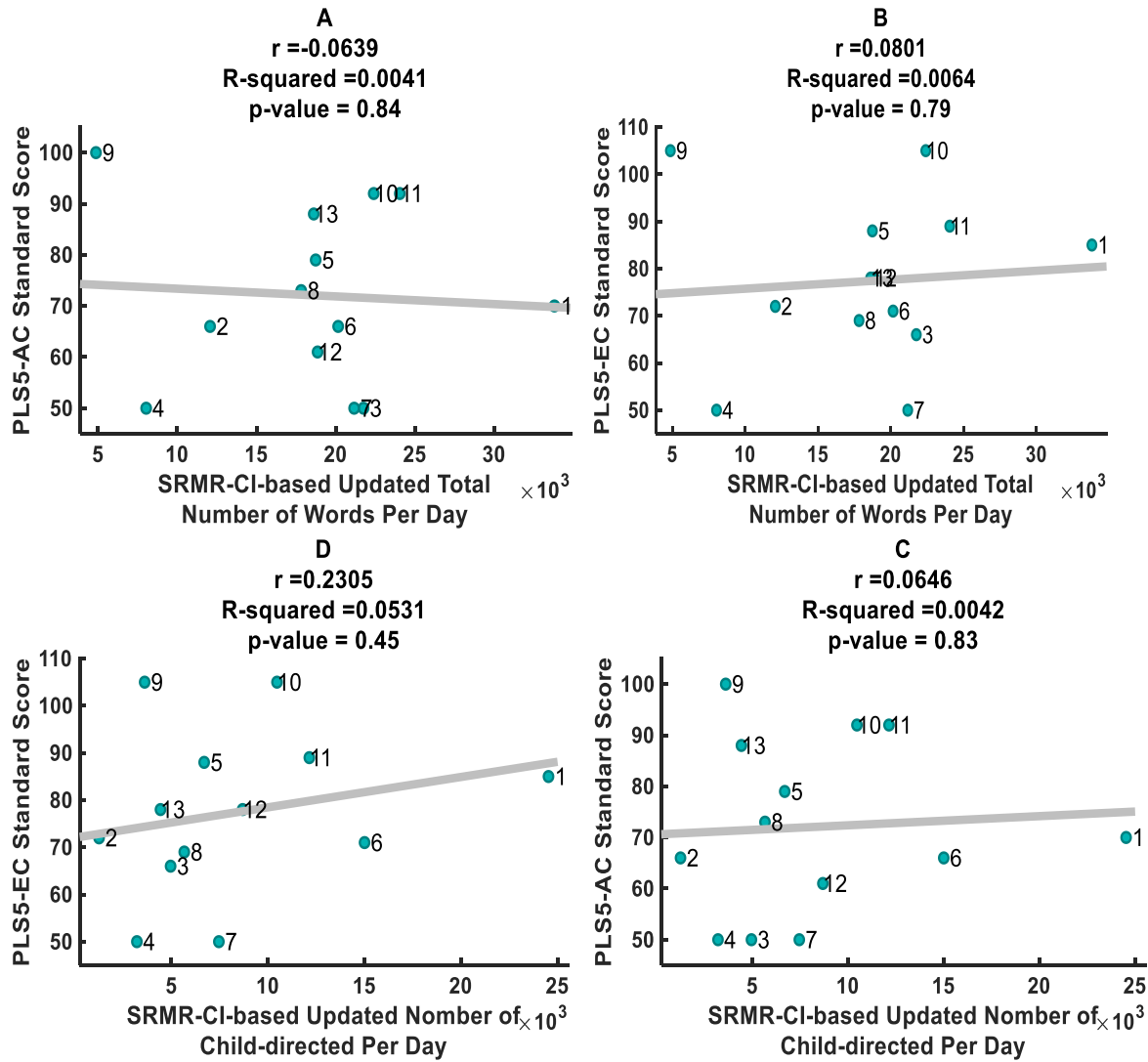


Figure 4-3. The relationship between updated total number of words based on SRMR-CI metrics and (A) PLS-AC and (B) PLS-EC standard scores, as well as SRMR-CI-based updated numbers of child-directed words and (C) PLS-AC and (D) PLS-EC standard scores. The r values are pairwise Pearson correlation coefficients between a given language input measure and the respective language outcome measure. $R\text{-squared}$ values are the coefficient of determination obtained from fitted ordinal linear regression models. The least-square lines are superimposed on each scatter plot.

4.3.4. Predictive value of SRMR-CI metric for language outcomes of children with CIs

Figure 4-4 plots average SRMR-CI from speech experienced by children in their home environments against children's expressive and receptive language outcomes. Notably, SRMR-CI strongly predicted both PLS-AC and PLS-EC. Higher values of SRMR-CI were significantly associated with higher PLS-AC and PLS-EC standard scores, consistent with our expectation (Santos & Falk, 2014). Appendix F and G present correlation matrix scatterplots and their corresponding *Pearson coefficients* for all pairs made by 16 measures of language input and 2 measures of PLS-AC and PLS-EC.

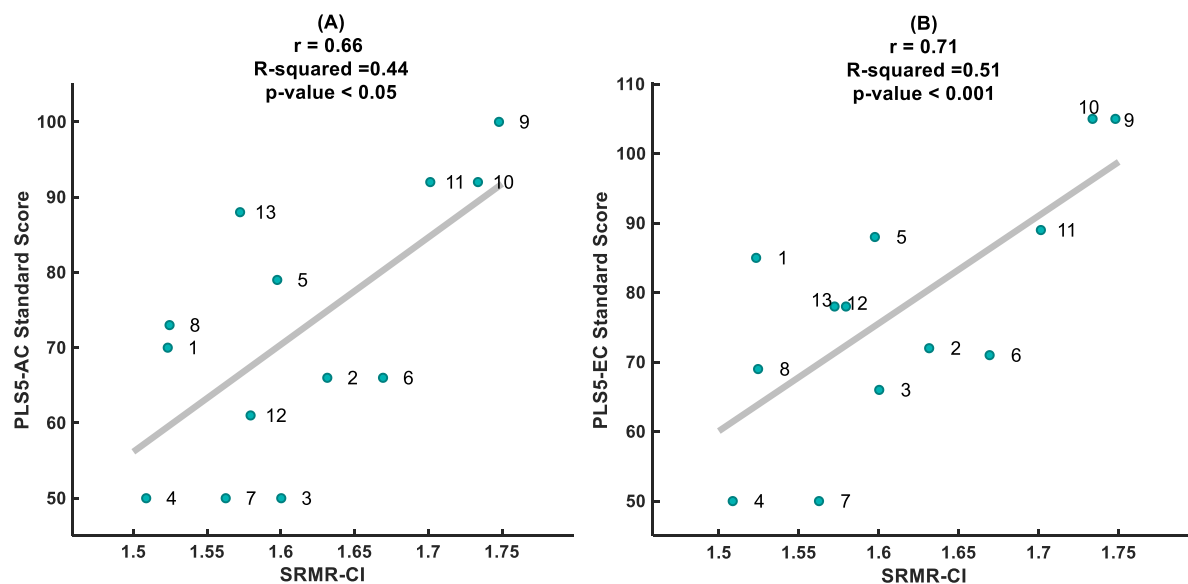


Figure 4-4. The relationship between average SRMR-CI obtained over recordings and language outcome scores of PLS5-EC and PLS5-AC at 6 months post-implantation. The grey line is the least-square line calculated for each data. The R-squared values and p-values for the ordinary linear regression model are reported on each plot.

4.4. Discussion

In this study, we used traditional and noise-adjusted measures of language input derived from Chapters 2 and 3 to explore how well each measure of language input predicted language outcomes at 6 months after cochlear implantation for early-implanted children with CIs participating in this study. Our first question concerned whether individual variability in the total

number of words per day would predict language outcomes, as measured by PLS5-AC and PLS5-EC standard scores. From these preliminary results based on small sample of children with CIs, no predictive value was found for total number of words per day. Our second question aimed to explore whether adjusting the number of words per day (total and child-directed) by considering the effect of noise and reverberation would better explain variability in language outcomes. We examined the predictive power of each of two methods for adjusting the number of words, i.e., human judgments of level of noise interference and the SRMR-CI metric, separately for PLS-AC and PLS-EC. Our preliminary results did not support our expectation, showing no predictive value of noise-adjustment estimates of words per day for explaining variability in language outcomes of this small group of children. However, these null results were anticipated to some extent because prior studies showed that the effect of child-directed speech appears after about 9 months (Rüter, 2011; Szagun & Rüter, 2009). Another reason for these null results could be the limited statistical power in the present study due to the small sample size. In contrast, the language outcome scores used for the present study were measured at 6-months post-implantation and thus were available 3 months before this minimum time lag for observing effects of language input.

We considered whether the average SRMR-CI derived over speech utterances within children's audio samples would predict language outcomes of these 13 children with CIs. The results showed that, indeed, the SRMR-CI quantitative measure of speech intelligibility was able to explain substantial variability in language outcomes of the 13 children. In particular, the average SRMR-CI values derived from analyzed sample audio explained 43% of the variance in receptive language skills (PLS5-AC) and 51% of the variance in expressive language skills (PLS5-EC) in these children. These results provide encouraging evidence, in that they suggest that SRMR-CI

value can be effectively used to assess the degree of intelligibility of speech experienced by children with CIs in their naturalistic home environments.

Our exploration of the relative contributions of lexical, morphosyntactic, and social-pragmatic measures of language input for children with CIs at 6 months post-implantation also did not provide any preliminary evidence of the connection between these measures of language input and language outcomes. We particularly explored the predictive value of type-token ratio (TTR), area under the type-token curve (AUTTC), mean length of utterance (MLU), caregiver responsiveness, and speech rate for PLS5-AC and PLS5-EC standard scores. None of the results from our simple regression analysis indicated a strong and significant correlation between each of these measures of language input and language outcomes.

Caution should be taken in making strong claims about the predictive value of SRMR-CI, due to several limitations and possible confounding factors. First, the robustness of this metric to various sources of variability in real-life scenarios of audio recording has not yet been verified. In addition, analyzing a greater number of daylong recording should be planned for future studies to assure that the average values of SRMR-CI over 5% audio samples obtained from anywhere from one to four daylong recordings for each child in this study provide a fair representation of children's auditory environments.

A surprising result was that updating the number of words (total and child-directed) to consider effects of noise and reverberation did not increase the predictive strength of these variables for language outcomes. These unexpected results can be explained in several ways. Children with CIs have access to amplitude envelopes of speech signals but have very limited access to fine-grained spectro-temporal cues (Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995; Shannon, 2002; Svirsky, 2017). This suggests that the word level might not be the right level of

granularity for understanding processes by which children make sense of degraded speech input. Prior studies also showed that listeners with CIs are able to organize and access to words in the broader utterance context (Kirk, Hay-McCutcheon, Sehgal, & Miyamoto, 2000; Kirk, Pisoni, & Osberger, 1995; McMurray, Ellis, & Apfelbaum, 2019). This implies that updating word counts might not reflect the actual pathway through which noise impacts children's understanding of speech.

Overall, the results from this exploratory study have not provided strong evidence to support a role of the quality and quantity of language input on children's language outcomes at 6 months after implantation. One of the primary reason behind this null result may be that according to prior studies, there needs to be at least 9 months lag between exposure to high-quality child-directed speech and language outcome assessments to observe effects of language input (Rüter, 2011; Szagun & Rüter, 2009) as well as the methodological issue related to the small sample size. Given these major limitations, we observed that adjusting the number of words experienced by each child to consider effects of background noise and reverberation did not lead to better prediction of language outcomes in children with CIs. Nevertheless, the high predictive value of the SRMR-CI metric suggests that this quantitative measure of speech intelligibility, which is specifically tailored to listeners with CIs, can be used in conjunction with other clinical approaches to provide more evidence for parents and clinicians in developing effective strategies to modify linguistic environments of children. Training parents to become familiar with various noise and reverberation patterns in real-life scenarios can be a strategy for controlling the quality and quantity of speech input that children experience every day. Results from the correlations between SRMR-CI and outcomes can be considered as preliminary evidence for researchers, clinicians, and parents

to pay more attention to the complexity of early linguistic environments experienced by children with CIs, particularly with respects to noise and reverberation.

Limitations of this study

The current study has multiple imitations which were already discussed in Chapters 2 and 3. In general, the exploratory nature of this study, including the small number of children with CIs, encourage us to be cautious about interpretation and generalization of these findings. In addition, the language outcomes of PLS-AC and PLS-EC were only available at 6 months post-implantation, which essentially limited this study for the minimum time-lag required to observe the expected effect of several measures of quantity and quality of language input that their strong effects on language development have been largely shown in children with typical hearing. Another limitation of this study, which was due to attrition, was the number of recording available for each child and their sparsity. Although our approach for coding the auditory environments were meticulous and comprehensive, the representativeness of the calculated language measures was limited by the number of recording available for each child.

Despite these limitations, this work has provided preliminary evidence that controlling the amount of noise and reverberation in linguistic environments of children with CIs may be an important consideration for intervention to improve language learning in children with CIs and decrease individual differences in their language skills. These preliminary results also provided evidence for SRMR-CI as a valid quantitative metric of speech intelligibility in listeners with CIs for use in monitoring quality and quantity of speech input experienced by children with CIs (Santos & Falk, 2014; Falk et al., 2015). Future studies along several lines can be targeted to provide more evidence for whether, and how, child-directed speech facilitates language development in children with CIs, as well as how environmental noise and reverberation may possibly impact language

outcomes in this at-risk population. Some possible future directions to improve on current research include: (1) studying more children who received their CIs early after implantation (i.e., less than 24 months of age), (2) including more daylong audio recordings across the first year of implantation, (3) videotaping language environments of children with CIs to enhance the coding process, (4) improving the reliability and accuracy of the LENA device, while extending it to enable classification of major categories such as child-directed speech versus adult-directed speech, (5) improving the transfer function responsible for mapping the level of noise to word adjustment, (6) testing and validating the robustness of SRMR-CI to sources of variability in speech and across environments, (7) improving characterizations of how SRMR-CI maps to percent speech intelligibility and updated words.

CHAPTER 5: Discussion and Conclusions

Despite more than a decade of research on factors influencing language outcomes in children with CIs, the enormous variability in these outcomes is still unexplained. Findings from prior studies largely demonstrated that the amount and quality of language input during early childhood significantly influences children language development as well as cognitive abilities (Hart & Risley, 1995; Hoff, 2003). Drawing inspiration from Hart and Risley (1995), we focused on characterizing variability in language input across children with CIs to determine the language gap between children with CIs receiving the most and the least language input. This study is the first one to use data from naturalistic environments of 14 prelingually deaf children who received CIs before the age of 24 months to understand the degree of individual variability in experienced quantity and quality of language input. Given the limitations with the LENA device for studying individual variability (Lehet, Arjmandi, Dilley, & Houston, under review), we developed an extensive coding system in order for human listeners to assess the amount of speech, who the speech was directed to, whether the speech overlapped with noise, the level of noise interference, and the type of overlapped noise. We used this coding system to extensively analyze the audio samples from natural linguistic environments of these children to explore the extent of variability across children with CIs in their language input along lexical, morphosyntactic, and social-pragmatic dimensions, focusing particularly on child-directed speech. To this end, we measured the number of words, type-token ratio, type-token curve, area under the type-token curve, mean length of utterance, caregiver responsiveness, and rate of speech experienced by children in their home environments. Treating the first year after implantation as a unit of time, the quality and quantity of language experienced by each child over the first year after implantation were summarized by per-day estimates of these measures. Using these measures, we demonstrated that

children's early language environments vary substantially in terms of quantity and quality of language input.

Prior studies showed how variability in the quantity and quality of language input across children with typical hearing translates into large variability in their language outcomes (Hart & Risley, 1995). Consistent with prior findings in typically-hearing children, our estimates demonstrated that children with CIs varied largely in the number of child-directed words they experienced per day. We showed how these per-day child-directed word differences across children may scale up remarkably over years to an extent comparable to that between most-advantaged and least-advantaged typically hearing children (Hart & Risley, 1995). We found a similar pattern of an approximately 30-million-word gap between the most-advantaged and least-advantaged children with CIs, consistent with prior findings of Hart and Risley (1995). This substantial word gap across children in experiencing high-quality, child-directed speech is of special importance, knowing that child-directed speech carries rich acoustic, linguistic, and extra-linguistic cues that significantly facilitate speech and language learning in children (Hart & Risley, 1995; Hoff, 2003; Weisleder & Fernald, 2013).

The supportive effect of child-directed speech on children's language outcomes becomes much more important for children with CIs. This is because these children have partial access to spectro-temporal cues in speech and are extremely vulnerable to sources of disturbances and interferences when listening to speech (Caldwell & Nittrouer, 2013; Hazrati & Loizou, 2012; Neuman et al., 2004). Children who experienced a relatively lower amount of child-directed speech had reduced chances to benefit from multimodal cues available in this type of speech to learn and construct language subsystems. It is therefore conceivable that the large differences between children with CIs observed in this study will contribute substantially to the large disparity

in their language outcomes. We further showed that how these per-day differences in exposure to varying amounts of child-directed speech can expand to around 30 million words after 3 years of hearing through CIs, which largely distinguishes some children from others in terms of the amount of child-directed speech experienced. Therefore, more attention to parents and caregivers' spoken communication style during their interaction with children with CIs may be a simple and effective remedy to the observed large differences among children in their language skills.

The quality and quantity of early linguistic environments of these 14 children with CIs were not dispersed only in the amount of speech they experienced. In addition, the results from our investigation demonstrated that children experienced language input with largely different lexical, morphosyntactic and pragmatic richness. Our novel measure of area under the type-token curve (AUTTC), which we proposed to model the quality and quantity of language input along two joint dimensions of number of words and word types, showed that children were largely dispersed in terms of their exposures to lexically rich, high-quality child-directed speech. Exposure to child-directed speech with richer lexical input both in terms of total number of words and number of unique word types has been shown to be an influential contributing factor to language outcomes in children with typical hearing (Golinkoff et al., 2019; Hoff-Ginsberg, 1991, 1994; Hoff & Naigles, 2002; Weisleder & Fernald, 2013). Therefore, this large variability across children in exposure to lexically rich speech input requires careful attention toward developing effective and evidence-based family-centered intervention strategies, particularly during the early period after implantation, in order to achieve an optimal match between plasticity of the brain and the early auditory diet.

We also showed how speech directed to children with CIs provides simpler models of language compared to adult-directed speech, which is consistent with prior findings in typically

developing children (Furrow et al., 1979) and children with CIs (Bergeson, 2011). Child-directed speech in the environments of children with CIs was less lexically complex, as shown by smaller TTRs compared to adult-directed speech. In addition, the sentences spoken to the children with CIs had shorter MLUs and were spoken with slower rate compared to adult-directed speech, which suggested that children with CIs have access to language models with less lexical and morphosyntactic complexity when speech is directed to them. It is not yet clear whether and how these simpler models of language in child-directed speech may contribute to language learning of children with CIs. However, in this study, using the data from natural auditory environments of children with CIs, we have showed that child-directed speech is distinguished from adult-directed speech based on lower TTR, shorter MLUs, and lower speech rate. Such findings from analyzing data in naturalistic environments may suggest that adults modify their speaking style when talking to children with CIs to assure that children benefit most from the language input (Bergeson, 2011; Wang, Bergeson, et al., 2018). Future studies can elaborate whether these discriminant properties of child-directed speech significantly contribute to the development of language in children with CIs.

Examinations of effects of environmental noise and reverberation on the quantity and quality of language input experienced by children with CIs has been a missing piece in studies of language development in this at-risk population. This is a critical knowledge gap, because listening to speech in noisy and reverberant conditions is extremely difficult for children with CIs, compared to children with typical hearing (Caldwell & Nittrouer, 2013; Neuman et al., 2004). We addressed this gap by developing two novel approaches to update the number of actual words heard by children with CIs in their home environments. In the first approach, results from auditory perceptual judgments of level of noise interference made by adult human analysts at four levels of

null clean, low, medium, and high were used to update the number of clean words experienced by each child. In the second approach, hearing the words through CIs was emulated by passing speech utterances through a computational model of speech hearing developed for listeners with CIs (i.e., SRMR-CI) to consider the effect of noise and reverberation on intelligibility of speech experienced by these children. The results of noise-adjusted estimates of number of child-directed words demonstrated – for the first time – that environmental noise and reverberation significantly reduce the amount of high-quality child-directed speech experienced by children with cochlear implants.

Our results suggested that variability across children with CIs in terms of amount of word loss due to noise and reverberation was large. Some children lost a relatively small amount of child-directed words, on the order of two hundred words per day, whereas other children missed a large amount, on the order of *five thousand* words per day. However, this difference was essentially due to the original amount of unadjusted child-directed words, rather than an effect of noise or reverberation on the updated words *per se*. Our data show evidence that children were differentially impacted by noise and reverberation, although the overall weight of variability across children in amount of high-quality child-directed speech experienced by children was almost the same before and after noise-adjustment. Results of this part basically suggest that the predictive value of the amount of child-directed speech is potentially improved after updating the number of words for noise and reverberation. Future studies with more children and recordings should be conducted to shed more light on this finding.

In addition, we found a similar pattern of a 30-million-word gap in our estimated child-directed words over three years after implantation based on the adjusted word counts, which could conceivably account for variability in language outcomes of children as demonstrated by Hart & Risley (Hart & Risley, 1995). The results from our study show how children with CIs are extremely

disadvantaged in experiencing high-quality child-directed words in their environments due to three factors. First, they must overcome a baseline disadvantage in ability to acquire language due to the extreme susceptibility of their impaired hearing to the amount and level of background noise and reverberation in their home environments. Second, the results from this study showed dramatic differences in the amount of high-quality child-directed speech experienced across different home environments of children with CIs, which raises the possibility of an additional disadvantage (e.g., Hart & Risley, 1995), namely that some children with CIs may receive very minimal language input, further challenging language acquisition. Unlike prior studies that merely focused on structural complexity of child-directed speech (Szagun & Schramm, 2016; Szagun & Stumper, 2012), we were inspired by the seminal work of Hart and Risley (1995). Here, we showed that children with CIs not only are substantially different in terms of experiencing high-quality child-directed words, but they are also at risk for missing a large amount of high-quality child-directed words available in their environments.

Our analyses also showed that children with CIs experience various environments in terms of distribution of types of noise overlapping with speech in their environments. For example, the susceptibility of recipients of CIs to mixed noise and multi-talker babbling noise in understanding speech is much higher than background noise with less spectral variation, such as white noise (Soleymani, Selesnick, & Landsberger, 2018; Tobey, Shin, Sundarajan, & Geers, 2011). The results of distributions of various types of noise that overlapped with words highlighted the complexity of auditory environments of children with CIs, while also showing that variability between children in exposure to various types of overlapping noise was substantial. This corroborates prior findings regarding large differences among users of CIs in the quality of speech presented in background noise (Busch et al., 2017). Our results further demonstrated that

characteristics of environmental noise itself varied substantially across children with CIs, such that some children experienced environments with considerably higher-intensity noise compared to others. Exposure to high-intensity noise during early development may have detrimental consequences on the learning rates of these children with CIs, thereby contributing to differences in their language outcomes (Bureš, Grécová, Popelář, & Syka, 2010; Busch et al., 2017).

We conducted an initial test on this small corpus of children with CIs to gain some preliminary data concerning the relationship between traditional (i.e., number of words, TTR, AUTTC, MLU, caregivers responsiveness, and speech rate) and updated measures of language input (i.e., noise-adjusted number of words), along with receptive and expressive language outcomes at 6 months post-implantation, as gauged through PLS5-AC and PLS5-EC standard scores. Our preliminary results did not show any significant correlation between measures of language input and language outcomes. These null results were anticipated, due to several limitations of this exploratory study, including, but not limited to, small sample size and access to measures of language outcome at a time interval before the 9-month time-lag required to observe the effect of noise (Rüter, 2011; Szagun & Rüter, 2009). However, we observed that the SRMR-CI metric nevertheless strongly predicted language outcomes of children at 6 months post-implantation, suggesting that interventions based on this quantitative metric may benefit children with CIs. The predictive value of this metric can be associated with its nature of emulating hearing speech through CIs (Falk et al., 2015), which is expected to closely approximate the degree of speech intelligibility from the ears of users of CIs. However, more studies in the future with a higher number of children and more recordings will help us to further examine the predictive ability of this metric.

The current research has several limitations. Perhaps the most significant limitation of this study was the small sample size drawn from a heterogeneous population, which is the nature of most longitudinal studies on the pediatric CIs population, particularly in recording home environments of children, which generates major challenges for participant recruitment. The number of audio recordings available per child was small, which raises concerns about the representativeness of the number of recordings and their generalizability to the entire first year after implantation. Considering these limitations, the results should be viewed as preliminary findings for a very new line of research which has its own constraints and difficulties. Still, this line of research has potential to significantly benefit the field of language development in children with CIs, as well as typical-hearing children, by addressing contributions of one of the major unexplored components to variability in such development: early linguistic environments. Another limitation of this study concerned the uncontrolled quality of recordings made in naturalistic environments using LENA devices. Although we tested the reliability and accuracy of LENA devices for word counts and major categories of interest in this study (e.g., speech vs. non-speech, child-directed speech or adult-directed speech), further study is required to assure that the signal-based metric of SRMR-CI behaves robustly in the presence of other sources of variability, such as variable microphone distance and device position with respect to talkers. Further, the two transfer functions used for mapping the codes for level of noise interference and SRMR-CI values to updated numbers of words should be further studied through conducting perceptual studies to provide improved emulation of performance of children with CIs in speech recognition in the presence of noise and reverberation.

Despite these limitations, the present work has contributed to the field both conceptually and methodologically. For the first time, this study focused on several elements of

early linguistic environment which have been shown to best support language learning in typically-hearing children, but which are understudied in children with CIs to examine how these elements vary across children. We showed that the pattern of a 30-million-word gap between most-advantaged and least-advantaged typically hearing children (Hart & Risley, 1995) exists in children with CIs, which could conceivably help us to explain an important portion of variability in children's language outcomes. In addition, this study has provided the first evidence showing that child-directed speech provides simpler models of language for children with CIs based on naturalistic data. More importantly, this is the first study that investigated variability across children in the quality and quantity of language input available to them after considering the effect of noise and reverberation on the amount of high-quality child-directed speech experienced by children. We demonstrated that environmental noise reduces the amount of child-directed words, but the 30-million-word gap still exists. This study has also provided evidence of substantial variability across children in exposure to various types of noise either in the presence of speech or noise itself. Our results suggest that SRMR-CI is a valuable quantitative metric for monitoring the quality of early linguistic environments of children with CIs for the intelligibility of speech in those environments. Furthermore, the methods are novel in terms of developing a Praat-Matlab interface tailored to analyze codes from a rich descriptive system implemented by human analysts, as well as the actual audio signal, in order to answer questions of this study by providing data at linguistic and acoustic levels. Other aspects of novelty include the proposed approaches for updating the number of words based on human coding of noise level and SRMR-CI metrics.

Future work

This project opens up several lines of research for the future, all centered around a primary question of how properties of early linguistic environments contribute to tremendous variability in language outcomes of children with CIs. The primary focus of this project was early linguistic environments of children with CIs. Further studies with more enhanced and reliable tools should be conducted to better characterize properties of language environments of children with CIs that best support their language learning. The goal of these planned studies for the future is to provide evidence-based feedback for parents and clinicians to improve the quality of children's language environments, particularly for those children who are at higher risk for underdeveloped language skills. Studies with higher participants and more controlled populations should be planned to reduce the possible confounding effects of uncontrolled variables in this study, such as additional comorbid diagnoses. Future studies equipped with video monitoring tools will significantly enhance the accuracy and reliability of several categories that were coded and analyzed in this study. Enhancing signal processing algorithms related to the SRMR-CI metric can further be examined and improved to achieve greater robustness for this metric against undesirable variables of naturalistic input, while conducting auditory perceptual studies on children with CIs to further improve the metric. The results from these envisioned perceptual studies can be specifically incorporated to improve the two transfer functions used in this study to map human codes for levels of noise interference and the SRMR-CI metric to corresponding levels of speech intelligibility.

APPENDICES

APPENDIX A

The Reference Table of Codes

Appendix A-Table 1. The reference table of codes and their descriptions at different levels on each tier.

<i>Tier 2 & 3: Adult Female and Adult Male Tiers</i>	
<i>Tier 2 & 3: Level 1 - Addressee Codes</i>	
Code/Contents	Description
T	Addressing Target Child. Analysts used this code if they were reasonably confident that the adult was talking to the target child, from context or otherwise.
O	Addressing Other Child. Analysts used this label when the addressee was <i>clearly another child</i> besides the target child (for example, if the parent said the sibling's name directly)
A	Addressing Adult. Analysts used this code when an adult speaker was talking to another adult.
X	Other Addressee. Analysts used this code when it was clear who the addressee is, but it was another category than above (<i>e.g.</i> , self-directed, pet-directed, or mobile/tablet app-directed), or the addressee was not clear.
<i>Tier 2 & 3: Level 2 – Type of Vocalization</i>	
Code/Contents	Description
Transcription	Analysts transcribed the words and utterances with conventionalized meanings. Partial, incorrect, and unintelligible words were indicated by an "X".
SNG	Singing. Analysts used this code when words were produced as singing.
NSV	Non-Speech Vocalization. Any noise produced by talker that was not a speech utterance such as <i>burps</i> , <i>yawns</i> , etc. were coded as NSV, but only if they overlapped with speech.
<i>Tier 2 & 3: Level 3 - Level of Interference or Primary Talker</i>	

Appendix A-Table 1. (cont'd)

Code/Contents	Description
L, M, H	The level of interference (Low, Medium, or High). Analysts coded the level of interference of speech only when it overlapped with the speech of the primary talker. The primary speech could be from any tiers of Adult Male or Other Talker.
P	Primary talker. Analysts used this code when there was overlap among multiple talkers to identify whether the speech was from the primary talker, the one whose speech was most informative to the target child.
<i>Tier 4: Target Child Tier</i>	
<i>Tier 4: Level 1 - Type of Vocalization</i>	
Code/Contents	Description
S	Speech or Speech-like Vocalization. Analysts used this label to code any vocalization from the target child that was remotely speech or speech-like.
CRY	Crying or Whining. Analysts used this code when a child was loudly and clearly crying or was making a whining noise.
NSV	Non-Speech Vocalization. Analysts used this code to label any non-speech vocalization events such as <i>yawns</i> , <i>sighs</i> , and vegetative noises (burps, hiccups) only if they were overlapping with adult speech.
<i>Tier 4: Level 2 - Transcription</i>	
Code/Contents	Description
Transcription	Analysts transcribed child's speech if it was clear enough to understand specific syllables.
<i>Tier 4: Level 3 - Level of Interference</i>	
Code/Contents	Description

Appendix A-Table 1. (cont'd)

L, M, H	The level of interference (Low, Medium, or High). Analysts used these labels to code the level of interference when child's vocalization overlapped with a useful speech.
<i>Tier 5: Other Talker Tier</i>	
<i>Tier 5: Level 1 – Talker Codes</i>	
Code/Contents	Description
AM	Adult male. Analysts used this code on this tier when a situation involved a second adult male talker.
AF	Adult female. Analysts used this code on this tier when a situation involved a second adult female talker.
CM	Child Male. Analysts used this code to label speech from other male children who (1) were not the target child, and (2) were a competent speaker (> 5 years old).
CF	Child Female. Analysts used this code to label speech from other female children who (1) were not the target child, and (2) a were competent speaker (> 5 years old).
C	Child of uncertain gender. Analysts used this code to label speech from children for whose gender was not clear from the context.
MIX	Multiple non-primary talkers overlapping. Analysts used this code in cases where multiple speakers were talking simultaneously and it was possible to code different speakers on different tiers of Adult Male, Adult Female and this tier. Analysts made sure that this multiple talker situation did not reach the state of babbling (see the description for Noise 1 and 2 tiers).
<i>Tier 5: Level 2 – Addressee Codes</i>	
Code	Description
T	Addressing Target Child. Analysts used this code if they were reasonably confident that the adult was talking to the target child, from context or otherwise.

Appendix A-Table 1. (cont'd)

O	Addressing Other Child. Analysts used this label when the addressee was clearly another child besides the target child (for example, if the parent said the sibling's name directly)
A	Addressing Adult. Analysts used this code when an adult speaker was talking to another adult.
X	Other Addressee. Analysts used this code when it was clear who the addressee is, but it was another category than above (<i>e.g.</i> , self-directed, pet-directed, or mobile/tablet app-directed) or the addressee was not clear.
Tier 5: Level 3 – Words and Utterances with Conventionalized Meanings	
Code	Description
Transcription	Analysts transcribed the words and utterances with conventionalized meanings. Partial, incorrect, and unintelligible words were indicated by an "X".
SNG	Singing. Analysts used this code when words were produced as singing.
NSV	Non-Speech Vocalization. Any noise produced by a talker that was not a speech utterance, such as <i>burps</i> , <i>yawns</i> , etc., were coded as NSV, but only if they overlapped with speech.
Tier 5: Level 4 - Level of Interference or Primary Talker	
Code	Description
L, M, H	The level of interference (Low, Medium, or High). Analysts coded the level of interference of speech only when it overlapped with the speech of the primary talker. The primary speech could be from either the Adult Male or Adult Female tiers.
P	Primary talker. Analysts used this code when there was overlap among multiple talkers to identify whether the speech was from the primary talker, the one whose speech was most informative to the target child.

Appendix A-Table 1. (cont'd)

Tier 6: Media Tier Codes	
Tier 6: Level 1 - Type of Media	
Code	Description
SP	Speech. Analysts used this code when they heard speech but were not able to determine more about the talkers' age or gender.
MUS	Music (with instruments/notes). Analysts used this label to code any sound with pitched instruments including electronic instruments and voice plus instruments.
SNG	Singing. Analysts used this label to code any pre-recorded media containing <i>unaccompanied</i> singing (voice with no instruments).
AM	Adult male. Analysts used this label to code events of adult male speaker that was unaccompanied by any other concurrent sound (e.g., a newscast delivered by a male talker where there is no music or other sound track).
AF	Adult female. Analysts used this label to code events of adult female speaker that was unaccompanied by any other concurrent sound.
C	Child. Analysts used this label to code a child speaker in the foreground and it was non-infant entertainment or content.
MIX	Analysts used this label to code two or more types of sound source above overlapping or alternating rapidly (for example, a person talking with a musical soundtrack in the background) and/or two or more media sources simultaneously.
Tier 6: Level 2 - Purpose of Pre-recorded Content	
Code	Description
G	General. This was the default label used by analysts when they did not have any sense of content.
EDU	Educational. This label was used to code any educational content that was judged to be <i>child-oriented</i> .

Appendix A-Table 1. (cont'd)

C-ENT	Child-oriented entertainment. This label was used by analysts to code events of child-oriented entertainment, specifically content that would be entertaining for a very young child (i.e., infant or toddler). Examples were Dora, Wiggles, Sesame Street.
TOY	Toy. This label used by analysts to code any pre-recorded content that were judged to be coming from a <i>child's toy</i> of some kind. Examples were toy car sounds, baby doll sound, etc.
INFO	Informational. Analysts used this label to code informational, adult-oriented content. Examples were news broadcasts, “serious” informational podcasts (unless content is specifically oriented to infants, in which case, analysts used EDU or C-ENT), announcer’s voice on a classical music station, etc.
POP, CLASS, JAZZ, MET	Specific genres for music. analysts used these labels to indicate the specific genres such as pop, classical, jazz, metal, etc.
ENT	Entertainment. Analysts used this label to code any entertainment for middle school or older. Examples were Youtube video content oriented to teens or older, science shows geared toward middle schoolers, traditional TV shows like sitcoms, etc.
MOV	Movie. This label was used to code movies that were geared toward older-age children or older. If the movie was clearly oriented toward an infant or toddler, analysts labeled it as C-ENT.
Tier 6: Level 3 - Words and Conventionalized Utterances	
Code	Description
Transcription	Analysts transcribed words for rare cases where the media was audible enough to be understood.
Tier 6: Level 4 - Level of Interference	
Code	Description
L, M, H	Analysts used these labels to identify the level of interference when media overlapped with an informative speech.
Tiers 7 & 8: Noise Tier Codes	
Tiers 7 & 8: Level 1 - Type of Noise	

Appendix A-Table 1. (cont'd)

Code	Description
G	General. Analysts used this label to code events that did not fit neatly into one of the below categories.
VOC	Vocalization. Analysts used this label to code vocalizations from other children that were not clear speech or coming from children under the age of 5.
BAB	Multi-Talker Babble. Analysts used this label to code audio intervals where there were so many people talking in the background such that the conversation could not be made out.
ANIM	Animal. Analysts used this label to code a noise judged to be generated by any type of animal.
MIX	Mixed noise source. This label was used to code events with more than two simultaneous noises.
MECH	Mechanical noise. Analysts used this label to code sounds from mechanized sources, including appliances or motorized machines of any type.
WATER	Water noise. Analysts used this label to code the running water sounds such as running bath water, doing dishes, shower, rain, etc.
WIND	Wind noise. Analysts used this label to code wind sound that was in the child's environments.
<i>Tier 7 & 8: Level 2 - Level of Interference</i>	
Code	Description
L, M, H	The level of interference (Low, Medium, or High). Analysts used these labels to identify the level of interference when media overlapped with an informative speech.
<i>Tiers 9: Microphone Noise Tier Codes</i>	
<i>Tier 9: Level 1 - Type of Microphone Noise</i>	
Code	Description

Appendix A-Table 1. (cont'd)

M	Microphone noise. Analysts used this code to indicate whether microphone noise occurred with intelligible speech or no speech at all.
O	Other. Analysts used this label when there were other oddities artefactual and associated with the LENA device itself.
<i>Tier 9: Level 2- Level of Interference</i>	
Code	Description
L, M, H	The level of interference (Low, Medium, or High). Analysts used these labels to indicate how much the microphone noise seemed to obscure other important sounds and speech that were happening.

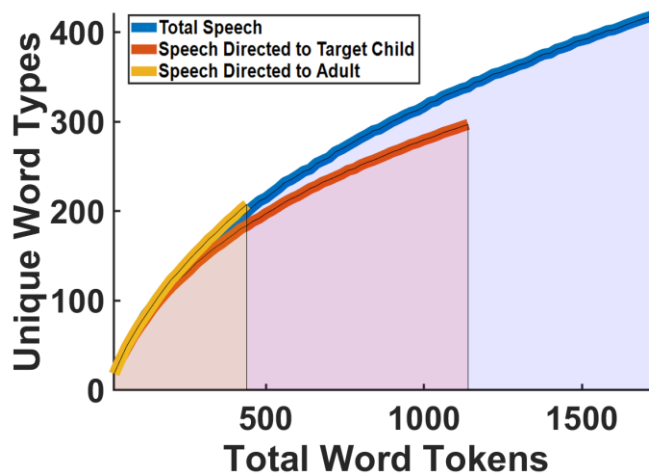
APPENDIX B

The Procedure for Constructing Type-token Curves

The procedure for constructing type-token curves started with randomly selecting word samples from all the words derived from the transcription of speech utterances belonging to a day-long recording (from the 5% of randomly sampled 30-second audio portions). The number of words that were randomly selected in each trial increased in increment of 20 morphemes until the size of the random selection was equal to the total number of words available for a recording (i.e., the size of random selection changed from 20 to 40 to 60 until the maximum number of words was achieved). For each random selection, we counted the number of unique word *types*, leading to an ordered pair consisting of a number of word *tokens* and word *types*. At each size of random selection (e.g., 40), the sampling procedure and calculation of the number of word types was repeated 100 times and the final number of unique word types was the average of the values of unique word types over those repetition at each sample size of random selection. This was done to reduce the possible effect of random selection on the calculated number of words types (Montag et al., 2018). The final values for constructing the type-token curve were counts of the number of unique word tokens at each of these sample sizes, which spanned the range from 20 words to the total number of words within the 5% of selected audio from a daylong recording.

An example of type-token curve derived from three types of total, child-directed, and overheard speech was shown in Figure 2-4. As explained before, the values on the graphs are the pairs of (number of word tokens, number of word types) derived for each sample size of random selection (i.e., 20, 40, 60, and so forth). The graph shows how the relationship between the number of word tokens and word types varies as the number of randomly chosen words in each selection

increases by a constant increment of 20. More information about the implementation process can be found in (Montag et al., 2018).



Appendix B-Figure 1. Type-token curves derived from total, child-directed and overheard speech in the environment of Child 1.

APPENDIX C

Further information about the coding procedure

In transcribing the words, analysts decided whether there was a conventionalized meaning for the utterances. Examples of utterances with conventionalized meanings were onomatopoeia (e.g., whoosh), sounds that animals or other entities made (e.g., moo, neigh, choo-choo, etc.) or other conventionalized speech-like sound sequences with conventionalized meanings (e.g., woo-hoo). If the meaning of an utterance was judged to not have been conventionalized for the linguistic community, then analysts did not code it as a word on the corresponding tier. A list of approved conventionalized meanings and spellings was created that analysts actively referred to during the coding process. This list was updated regularly during the coding project. In cases where human analysts were unsure whether an utterance had a conventionalized sound-mapping meaning, a project manager who was a native English speaker made the final decision. The morphemes in words in the transcriptions were separated by a “+” sign, which was later used to count the number of morphemes in transcriptions. Connected words (e.g., “uh-oh”) and contracted words (e.g., “we’re”) were treated as one word when the textgrids were analyzed to capture measures of linguistic input. The spelling of all lexical transcriptions and the spellings of terms with conventionalized sound-meaning mappings (e.g., choo-choo) were double-checked by a senior research assistant with advanced coding skills and high familiarity with the coding system.

APPENDIX D

The coding procedure followed by human analysts to code instances where there were multiple simultaneous talkers

In cases where there were multiple simultaneous talkers, analysts first identified the Primary talker among multiple talkers, and then used the three tiers of *AdultFemale* or *AdultMale*, and/or *OtherTalker* to code speech of each talker on the appropriate tier (see Table 2-5 and Appendix A in Chapter 2). The primary talker was the one whose speech was judged to be most informationally relevant to the target child while providing the best word examples for her/him to learn from. As shown in Table 2-5 and Appendix A, as well as Figure 2-3 in Chapter 2, human analysts used the code of “P” at Level 3 of the *AdultFemale* and *AdultMale* tiers and Level 4 of the *OtherTalker* tier to identify the primary speech when there were multiple talkers. Human analysts used the following information and cues to identify the primary talker (“*primary speech*” hereafter): (a) whether speech was directed to the target child, (b) whether it was produced by a caregiver, (c) whether it was more intelligible to the ear of human analysts, (d) whether it was louder than other speech, indicating the closeness to the target child, and/or (e) whether it began before the interference (i.e., temporal precedence). After analysts identified the speech from the primary talker and marked that specific portion of audio on the textgrid, they identified the start time and end time within the utterance that other competing sound/speech sources (either noise or other speech) overlapped with the *primary speech*. Then they coded the stretches of overlapped sound/speech on the appropriate tier. (See Table 2-5 and Appendix A and Figure 2-3 in Chapter 2.) Analysts then judged the level of interference of competing sources in terms of their masking and interfering effects on intelligibility of the *primary speech*, and coded the corresponding audio

stretches on the appropriate tiers. (See Table 4 in Chapter 2.) This was a decision relative to the loudness and intelligibility of the primary speech to understand to what extent one or multiple sources of interference may mask (or obscure) words that the target child heard.

Overlapping sound or speech from any sources of media, human, and other noise types were labeled in their appropriate tiers. Their level of interference with the primary speech was coded as either low (“L”) or medium (“M”) or high (“H”). (See Table 2-5, Chapter 2.) If an interval with *primary speech* (i.e., an utterance) was partially contaminated by another overlapping source of sound/speech, analysts broke the interval up into separate intervals each with its own level of interference (i.e., *clean, low, medium, high*). (See Figure 2-3 in Chapter 2 for an example of this condition.) Words within an interval that was partially disrupted by noise were categorized into one of the two intervals with different levels of interference from overlapping sounds based on the portion of the word that was degraded. (The cut-off was fifty percent word degradation.) For example, in the case of the mother speaking to the target child and saying “*Can you come back and play?*”, if the child bangs blocks together during the word “*back*” the entire utterance would be split into three intervals coded as: {Boundary} T;can you come {Boundary} T;back {Boundary} T;and play {Boundary}. The type of overlapping noise and its level of interference would be coded on the *Target Child* tier for this case. (See Figure 2-3 in Chapter 2 for a similar scenario.)

APPENDIX E

Parameters of the model for fitting SRMR-CI values to SI(%)

The parameters of the model that fits the scaled SRMR-CI values to their corresponding SI(%) are presented below. The model parameters calculated from the non-linear curve-fitting procedure with their 95% confidence bounds are:

$$\alpha_1 = 13.43 (13.13, 13.73), \alpha_2 = 7.701 (7.533, 7.868), k = 0.89 (0.88, 0.893)$$

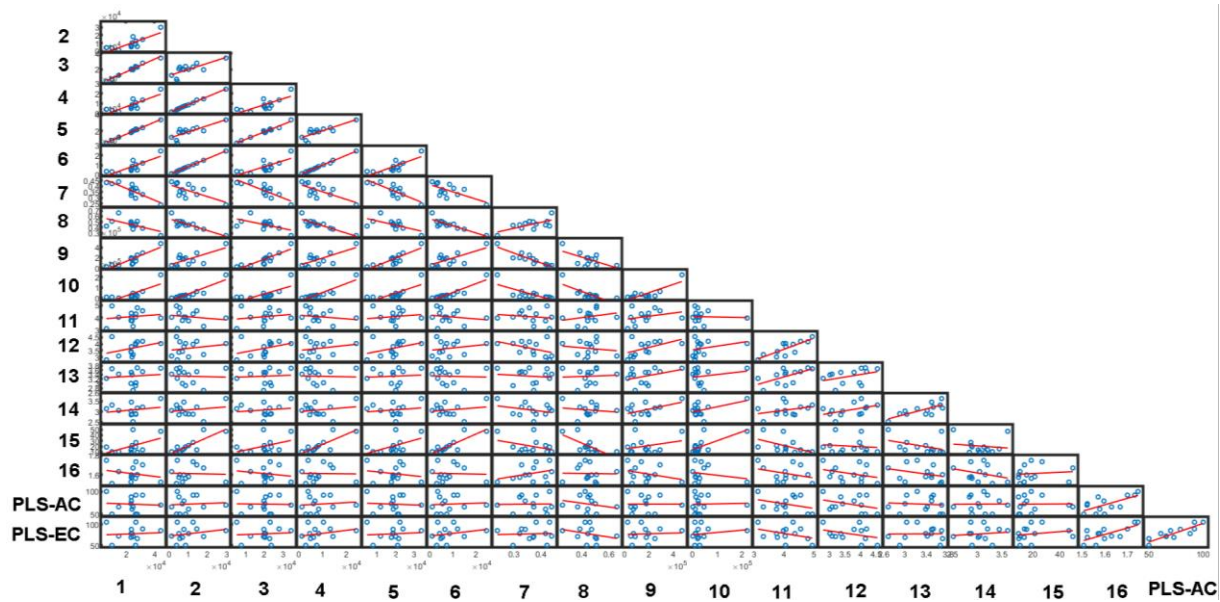
$$R\text{-Squared} = 0.99$$

These parameters belong to the following sigmoidal mapping function:

$$SI(\%) = \frac{K}{1 + e^{-(\alpha_1 X - \alpha_2)}} \times 100\%$$

APPENDIX F

Correlation matrix scatterplot with least-square fitted lines for pairs of measures of language input and measures of language outputs.



Appendix F-Figure 1. Correlation matrix scatterplot with least-square fitted lines for pairs of measures obtained from 16 measures of language input plus 2 measures of language outputs. Refer to Table 4-1 to match the numbers in this figure to their corresponding measures of language input.

APPENDIX G

Pearson correlation coefficients for all possible pairs made by measures of language input and measures of language outcomes.

Appendix G-Table 1. Pearson correlation coefficients for all possible pairs made by 16 measures of language input and 2 measures of PLS-AC and PLS-EC. *r* values that were statistically significant are bolded in the table.

Measure of Language Input	2	0.80															
	3	0.98	0.74														
	4	0.82	0.99	0.76													
	5	0.99	0.8	0.98	0.81												
	6	0.81	0.99	0.75	0.99	0.81											
	7	-0.84	-0.66	-0.84	-0.69	-0.83	-0.67										
	8	-0.62	-0.85	-0.57	-0.86	-0.6	-0.85	0.67									
	9	0.88	0.69	0.87	0.72	0.86	0.70	-0.9	-0.6								
	10	0.77	0.9	0.7	0.91	0.75	0.90	-0.73	-0.75	0.83							
	11	0.15	-0.17	0.18	-0.16	0.15	-0.17	-0.15	0.25	0.29	-0.04						
	12	0.37	0.2	0.37	0.21	0.37	0.2	-0.44	-0.14	0.5	0.29	0.81					
	13	0.10	-0.04	0.07	-0.04	0.07	-0.04	-0.12	0.06	0.37	0.22	0.61	0.38				
	14	0.20	0.19	0.15	0.2	0.16	0.19	-0.31	-0.15	0.51	0.49	0.33	0.46	0.74			
	15	0.48	0.85	0.4	0.82	0.5	0.84	-0.35	-0.71	0.28	0.67	-0.5	-0.11	-0.38	-0.13		
	16	-0.21	-0.03	-0.18	-0.01	-0.18	-0.03	0.32	-0.02	-0.31	-0.21	-0.32	-0.31	-0.31	-0.31	0.11	
Language Outcome	1	-0.06	0.07	-0.02	0.1	-0.06	0.06	0.01	-0.22	0	0.03	-0.31	-0.33	-0.07	0.01	0.01	0.66
	2	0.07	0.23	0.07	0.25	0.08	0.23	0.04	-0.27	0.05	0.17	-0.35	-0.33	0.01	0.11	0.2	0.71
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
		Measures of Language Input															
																	Language Outcome

BIBLIOGRAPHY

BIBLIOGRAPHY

- Adams, J. S., Hasenstab, M. S., Pippin, G. W., & Sismanis, A. (2004). Telephone use and understanding in patients with cochlear implants. *Ear, Nose and Throat Journal*, 83(2), 96–103.
- Ambrose, S. E., VanDam, M., & Moeller, M. P. (2014). Linguistic input, electronic media, and communication outcomes of toddlers with hearing loss. *Ear and Hearing*, 35(2), 139–147. <https://doi.org/10.1097/AUD.0b013e3182a76768>.Linguistic
- Anderson, C. A., Lazard, D. S., & Hartley, D. E. H. (2017). Plasticity in bilateral superior temporal cortex: Effects of deafness and cochlear implantation on auditory and visual speech processing. *Hearing Research*, 343, 138–149. <https://doi.org/10.1016/j.heares.2016.07.013>
- Aschendorff, A., Kromeier, J., Klenzner, T., & Laszig, R. (2007). Quality control after insertion of the nucleus contour and contour advance electrode in adults. *Ear and Hearing*, 28(SUPPL.2), 75–79. <https://doi.org/10.1097/AUD.0b013e318031542e>
- Barnes, S., Gutfreund, M., Satterly, D., & Wells, G. (1983). Characteristics of Adult Speech which Predict Children's Language Development. *Journal of Child Language*, 10(1), 65–84. <https://doi.org/10.1017/S0305000900005146>
- Baskent, D., & Gaudrain, E. (2016). Perception and Psychoacoustics of Speech in Cochlear Implant Users. *Scientific Foundations of Audiology. Perspectives from Physics, Biology, Modelling, and Medicine*, 185–320. Retrieved from https://books.google.de/books?hl=de&lr=&id=EtAyDAAQBAJ&oi=fnd&pg=PA285&dq=Scientific+Foundations+of+Audiology&ots=cfEfTicv7h&sig=1cTQmXsc_FR7oNQiwWYgpkIOkN0
- Beckwith, L., & Cohen, S. E. (1989). Maternal responsiveness with preterm infants and later competency. *New Directions for Child and Adolescent Development*, 1989(43), 75–87. <https://doi.org/10.1002/cd.23219894308>
- Bergeson, T. R. (2011). Maternal speech to hearing-impaired infants in the first year of hearing aid or cochlear implant use: a preliminary report. *Cochlear Implants International*, 12(sup1), S101–S104. <https://doi.org/10.1179/146701011X13001035752741>
- Bergeson, T. R., Miller, R. J., & McCune, K. (2006). Mother's speech to hearing-impaired infants and children with cochlear implants. *Infancy*, 10(3), 221–240. https://doi.org/10.1207/s15327078in1003_2

- Bierer, J. A., Spindler, E., Bierer, S. M., & Wright, R. (2016). An Examination of Sources of Variability Across the Consonant-Nucleus-Consonant Test in Cochlear Implant Listeners. *Trends in Hearing*, 20, 1–8. <https://doi.org/10.1177/2331216516646556>
- Blamey, P., Artieres, F., Başkent, D., Bergeron, F., Beynon, A., Burke, E., ... Lazard, D. S. (2013). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: An update with 2251 patients. *Audiology and Neurotology*, 18(1), 36–47. <https://doi.org/10.1159/000343189>
- Blamey, P., Pyman, B., Clark, G., Dowell, R., Gordon, M., Brown, A. M., & Hollow, R. D. (1992). Factors predicting postoperative sentence scores in postlinguistically deaf adult cochlear implant patients. *Annals of Otolaryngology, Rhinology & Laryngology*, 101(4), 342–348. <https://doi.org/10.1177/000348949210100410>
- Bockstael, A., Dekoninck, L., De Coensel, B., Oldoni, D., Can, A., & Botteldooren, D. (2011). Wind turbine noise: Annoyance and alternative exposure indicators. *Proceedings of Forum Acusticum*, 345–350.
- Boersma, P., & Weenink, D. (2001). Praat, a system for doing phonetics by computer. *Glott International*, 5:9/10, 341–345. Retrieved from <http://www.praat.org>
- Bořil, T., & Skarnitzl, R. (2016). Tools rPraat and mPraat. *Text, Speech, and Dialogue. TSD 2016. Lecture Notes in Computer Science*, 9924, 367–374. https://doi.org/10.1007/978-3-319-45510-5_42
- Bornstein, M. H., Haynes, M. O., & Painter, K. M. (1998). Sources of child vocabulary competence: A multivariate model. *Journal of Child Language*, 25(2), 367–393. <https://doi.org/10.1017/S0305000998003456>
- Bradley, J. S., & Sato, H. (2004). Speech intelligibility test results for grades 1, 3 and 6 children in real classrooms. *ICA*, 1191–1194.
- Bradley, R. H., & Caldwell, B. M. (1976a). Early home environment and changes in mental test performance in children from 6 to 36 months. *Developmental Psychology*, 12(2), 93. *Developmental Psychology* 12, 93(2), 1689–1699. <https://doi.org/10.1017/CBO9781107415324.004>
- Bradley, R. H., & Caldwell, B. M. (1976b). The relation of infants' home environments to mental test performance at fifty-four months: A follow-up study. *Child Development*, 1172–1174.
- Bradlow, A. R., Kraus, N., & Hayes, E. (2003). Speaking clearly for children with learning disabilities: Sentence perception in noise. *Journal of Speech, Language, and Hearing*

Research, 46(1), 80–97. [https://doi.org/10.1044/1092-4388\(2003/007\)](https://doi.org/10.1044/1092-4388(2003/007))

- Bregman, A. S. (1994). Auditory scene analysis : The perceptual organization of sound Auditory scene analysis : The perceptual organization of sound Computational auditory scene analysis : Principles , algorithms , and applications Auditory Scene Analysis : The Perceptual Organ. *MIT Press*, 3–4.
- Broen, P. (1972). The verbal environment of the language learning child. In *ASHA Monographs* (Vol. 17).
- Brown, R. (1973). A First Language: The Early Stages. In *Language*.
<https://doi.org/10.2307/412904>
- Bulf, H., Johnson, S. P., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition*, 121(1), 127–132. <https://doi.org/10.1016/j.cognition.2011.06.010>
- Bureš, Z., Grécová, J., Popelář, J., & Syka, J. (2010). Noise exposure during early development impairs the processing of sound intensity in adult rats. *European Journal of Neuroscience*, 32(1), 155–164. <https://doi.org/10.1111/j.1460-9568.2010.07280.x>
- Busch, T., Vanpoucke, F., & van Wieringen, A. (2017). Auditory Environment Across the Life Span of Cochlear Implant Users: Insights From Data Logging. *Journal of Speech Language and Hearing Research*, 60(5), 1362. https://doi.org/10.1044/2016_JSLHR-H-16-0162
- Caldwel, A., & Nittrouer, S. (2008). *Speech Perception in Noise by Children With Cochlear Implants*. 64(12), 2391–2404. <https://doi.org/10.1038/jid.2014.371>
- Caldwell, A., & Nittrouer, S. (2013). Speech Perception in Noise by Children With Cochlear Implants. *JSLHR*, 56(February), 13–31. [https://doi.org/10.1044/1092-4388\(2012/11-0338\)a](https://doi.org/10.1044/1092-4388(2012/11-0338)a)
- Carroll, J. B. (1938). Diversity of vocabulary and the harmonic series law of word-frequency distribution. *The Psychological Record*, 2(16), 379–386.
<https://doi.org/10.1007/bf03393224>
- Cheskin, A. (1981). The verbal environment provided by hearing mothers for their young deaf children. *Journal of Communication Disorders*, 14(6), 485–496.
[https://doi.org/10.1016/0021-9924\(81\)90029-0](https://doi.org/10.1016/0021-9924(81)90029-0)
- Clarke-Stewart, K. A. (1973). Interactions between mothers and their young children: Characteristics and consequences. *Monographs of the Society for Research in Child Development*, 1–109.
- Cohen, N. L., Waltzman, S. B., Roland, J. T., Staller, S. J., & Hoffman, R. A. (1999). Early

- results using the nucleus CI24M in children. *American Journal of Otology*, Vol. 20, pp. 198–204. Retrieved from <http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L29476647%5Cnhttp://sfx.library.uu.nl/utrecht?sid=EMBASE&issn=01929763&id=doi:&atitle=Early+results+using+the+nucleus+CI24M+in+children&stitle=Am.+J.+Otol.&title=American+Journal+of+>
- Cohen, N. L., Waltzman, S. B., & Shapiro, W. H. (1989). Telephone speech comprehension with use of the nucleus cochlear implant. *Annals of Otology, Rhinology & Laryngology*, 98(8_suppl), 8–11. <https://doi.org/10.1177/00034894890980S802>
- Connor, C. M., Alexander, H., & Teresa, A. (2000). *Speech , vocabulary , and the education of children using cochlear implants : Oral or total communic ...*
- Cooper, R. P. (1990). *Preference for Infant-Directed Speech in the First Month after Birth*. 61(5), 1584–1595.
- Davidson, L. (2006). Effects of stimulus level on the speech perception abilities of children using cochlear implants or digital hearing aids. *Ear and Hearing*, 27(5), 493–507. <https://doi.org/10.1097/01.aud.0000234635.48564.ce>
- Davidson, L. S., Geers, A. E., Blamey, P. J., Tobey, E., & Brenner, C. (2011). Factors contributing to speech perception scores in long-term pediatric CI users. *Ear and Hearing*, 32(1 Suppl), 19S–26S. <https://doi.org/10.1038/jid.2014.371>
- Dawson, P. W., Decker, J. A., & Psarros, C. E. (2004). Optimizing dynamic range in children using the nucleus cochlear implant. *Ear and Hearing*, 25(3), 230–241. <https://doi.org/10.1097/01.AUD.0000130795.66185.28>
- DesJardin, J. L., & Eisenberg, L. S. (2007). Maternal contributions: supporting language development in young children with cochlear implants. *Ear & Hearing*, 28(4), 456–469. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=106012820&site=ehost-live>
- DeThorne, L. S., Johnson, B. W., & Loeb, J. W. (2005). A closer look at MLU: What does it really measure? *Clinical Linguistics and Phonetics*, 19(8), 635–648. <https://doi.org/10.1080/02699200410001716165>
- Diaz, L., Labrell, F., Le Normand, M. T., Guinchat, V., & Dellatolas, G. (2019). School achievement of deaf children ten years after cochlear implantation. *Neuropsychiatrie de l'Enfance et de l'Adolescence*, 67(1), 50–57. <https://doi.org/10.1016/j.neurenf.2018.07.006>

- Dickinson, D. K., Golinkoff, R. M., & Hirsh-Pasek, K. (2010). Speaking out for language: Why language is central to reading development. *Educational Researcher*, 39(4), 305–310. <https://doi.org/10.3102/0013189X10370204>
- Dilley, Lehet, M., Wieland, E. A., Arjmandi, M. K., & Kondaurova, M. (2018). Quality and quantity of maternal speech predicts speech-language clinical outcomes in deaf infants with cochlear implants. Manuscript submitted for publication. *Journal of Speech, Language, and Hearing Research*.
- Dorman, M. F., & Gifford, R. H. (2017). Speech Understanding in Complex Listening Environments by Listeners Fit With Cochlear Implants. *Journal of Speech, Language and Hearing Research*, 60(October), 3019–3026. https://doi.org/10.1044/2017_JSLHR-H-17-0035
- Duchesne, L., Sutton, A., & Bergeron, F. (2009). Language achievement in children who received cochlear implants between 1 and 2 years of Age: Group trends and individual patterns. *Journal of Deaf Studies and Deaf Education*, 14(4), 465–485. <https://doi.org/10.1093/deafed/enp010>
- Dunn, C. C., Walker, E. A., Oleson, J., Kenworthy, M., Voorst, T. Van, Tomblin, J. B., ... Gantz, B. J. (2015). Longitudinal speech perception and language performance in pediatric cochlear implant users. The effect of age at implantation. *Ear and Hearing*, 35(2), 148–160. <https://doi.org/10.1097/AUD.0b013e3182a4a8f0.Longitudinal>
- Eisenberg, S. L., Fersko, T. M. G., & Lundgren, C. (2001). The Use of MLU for Identifying Language Impairment in Preschool Children: A Review. *American Journal of Speech-Language Pathology*, 10(4), 323–342. [https://doi.org/10.1044/1058-0360\(2001/028\)](https://doi.org/10.1044/1058-0360(2001/028))
- 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. [https://doi.org/10.1044/1092-4388\(2010/09-0165\).Speech](https://doi.org/10.1044/1092-4388(2010/09-0165).Speech)
- Ertmer, D. J., Young, N. M., & Nathani, S. (2007a). Profiles of vocal development in young cochlear implant recipients. *Journal of Speech, Language, and Hearing Research*, 50(2), 393–407. [https://doi.org/10.1044/1092-4388\(2007/028\)](https://doi.org/10.1044/1092-4388(2007/028))
- Ertmer, D. J., Young, N. M., & Nathani, S. (2007b). Profiles of Vocal Development in Young Cochlear Implant Recipients. *Journal of Speech Language and Hearing Research*, 50(2), 393. [https://doi.org/10.1044/1092-4388\(2007/028\)](https://doi.org/10.1044/1092-4388(2007/028))
- Evans, G. W., Maxwell, L. E., & Hart, B. (1999). Parental language and verbal responsiveness to children in crowded homes. *Developmental Psychology*, 35(4), 1020–1023. <https://doi.org/10.1037/0012-1649.35.4.1020>

- Fagan, M. K., Bergeson, T. R., & Morris, K. J. (2014). *Synchrony, Complexity and Directiveness in Mothers' Interactions with Infants Pre- and Post-Cochlear Implantation*. 37(3), 249–257. <https://doi.org/10.1088/1367-2630/15/1/015008.Fluid>
- Falk, T. H., Parsa, V., Santos, J. F., Arehart, K., Hazrati, O., Falk, T. H., ... Scollie, S. (2015). *Objective Quality Prediction for Users of and Intelligibility Assistive Listening Devices*. (march), 114–124.
- Ferguson, S. H., & Poore, M. A. (2010). Acoustic Correlates of Reported Clear Speech Strategies. *Jara*, XLIII, 45–64.
- Fernald. (1989). *Intonation and communicative intent in mothers' speech to infants: Is the melody the message?*
- Fernald, A. (1992). Meaningful melodies in mothers' speech to infants. In *Nonverbal vocal communication: Comparative and developmental approaches* (p. 262).
- Fernald, Anne. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, 8(2), 181–195. [https://doi.org/10.1016/S0163-6383\(85\)80005-9](https://doi.org/10.1016/S0163-6383(85)80005-9)
- Fernald, Anne. (2000). Speech to infants as hyperspeech: Knowledge-driven processes in early word recognition. *Phonetica*, Vol. 57, pp. 242–254. <https://doi.org/10.1159/000028477>
- Fernald, & Kuhl. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behaviour and Development*, 10, 279–293.
- Fetterman, B. L., & Domico, E. H. (2002). Speech recognition in background noise of cochlear implant patients. *Otolaryngology - Head and Neck Surgery*, 126(3), 257–263. <https://doi.org/10.1067/mhn.2002.123044>
- Finley, C. C., & Skinner, M. W. (2008). Role of electrode placement as a contributor to variability in cochlear implant outcomes. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 29(7), 920–928. <https://doi.org/10.1097/MAO.0b013e318184f492.Role>
- Firszt, J. B., Holden, L. K., Skinner, M. W., Tobey, E. A., Peterson, A., Gaggl, W., ... Wackym, P. A. (2004). Recognition of speech presented at soft to loud levels by adult cochlear implant recipients of three cochlear implant systems. *Ear and Hearing*, 25(4), 375–387. <https://doi.org/10.1097/01.AUD.0000134552.22205.EE>
- Fitzpatrick, E., Durieux-Smith, A., Eriks-Brophy, A., Olds, J., & Gaines, R. (2007). The impact of newborn hearing screening on communication development. *Journal of Medical*

Screening, 14(3), 123–131.

- Frermann, L., & Frank, M. C. (2016). *Prosodic Features from Large Corpora of Child-Directed Speech as Predictors of the Age of Acquisition of Words*.
- Fu, Q. J., & Nogaki, G. (2005). Noise susceptibility of cochlear implant users: The role of spectral resolution and smearing. *JARO - Journal of the Association for Research in Otolaryngology*, 6(1), 19–27. <https://doi.org/10.1007/s10162-004-5024-3>
- Furrow, D., Nelson, K., & Benedict, H. (1979). Mothers' speech to children and syntactic development: Some simple relationships. *Journal of Child Language*, 6(3), 423–442. <https://doi.org/10.1017/S0305000900002464>
- Geers, A. E., Strube, M. J., Tobey, E. A., & Moog, J. S. (2011). (2011). Epilogue: factors contributing to long-term outcomes of cochlear implantation in early childhood. *Ear and Hearing*, 32(1 Suppl), 84S. <https://doi.org/10.1097/AUD.0b013e3181ffd5b5>.Epilogue
- Geers, A. E., Hayes, H., & Louis, S. (2009). *Spoken Language Scores of Children Using Cochlear Implants Compared to Hearing Age-Mates at School Entry*. <https://doi.org/10.1093/deafed/enn046>
- Geers, A. E., Nicholas, J. G., & Moog, J. S. (2007a). *Estimating the influence of cochlear implantation on language development in children*. 262–273. <https://doi.org/10.1080/16513860701659404>
- Geers, A. E., Nicholas, J. G., & Moog, J. S. (2007b). Estimating the influence of cochlear implantation on language development in children. *Audiological Medicine*, 5(4), 262–273. <https://doi.org/10.1080/16513860701659404>
- Geers, A. E., Nicholas, J. G., & Sedey, A. L. (2003). Language Skills of Children with Early Cochlear Implantation. *Ear and Hearing*, 24(Supplement), 46S-58S. <https://doi.org/10.1097/01.AUD.0000051689.57380.1B>
- Genuit, K. (1999). The use of psychoacoustic parameters combined with A-weighted SPL in noise description. *Inter-Noise 1999*, (January 1999), 1887–1892.
- Gilkerson, J., & Richards, J. A. (2009). The LENA Natural Language Study. *LENA Foundation Technical Report*, (September 2008), 1–26.
- Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, 19(5), 515–523. <https://doi.org/10.1111/j.1467-9280.2008.02117.x>

- Golinkoff, R. M., Hoff, E., Rowe, M. L., Tamis-LeMonda, C. S., & Hirsh-Pasek, K. (2019). Language Matters: Denying the Existence of the 30-Million-Word Gap Has Serious Consequences. *Child Development*, 90(3), 985–992. <https://doi.org/10.1111/cdev.13128>
- Gomaa, N. A., Rubinstein, J. T., Lowder, M. W., Tyler, R. S., & Gantz, B. J. (2003). Residual Speech Perception and Cochlear Implant Performance in Postlingually Deafened Adults. *Ear and Hearing*, 24(6), 539–544. <https://doi.org/10.1097/01.AUD.0000100208.26628.2D>
- Gout, A., Christophe, A., & Morgan, J. L. (2004). Phonological phrase boundaries constrain lexical access II. Infant data. *Journal of Memory and Language*, 51(4), 548–567. <https://doi.org/10.1016/j.jml.2004.07.002>
- Graham, J., Throat, R. N., Cochlear, E. H., Robinson, P., Midgley, E., & Implant, B. C. (2009). Predictors of audiological outcome following cochlear implantation in adults. *Cochlear Implants International*, 10(July), 119–141. <https://doi.org/10.1002/cii>
- Greenwood, C. R., Thiemann-Bourque, K., Walker, D., Buzhardt, J., & Gilkerson, J. (2011). Assessing children’s home language environments using automatic speech recognition technology. *Communication Disorders Quarterly*, 32(2), 83–92. <https://doi.org/10.1177/1525740110367826>
- Grieser, D. A. L., & Kuhl, P. K. (1988). Maternal Speech to Infants in a Tonal Language: Support for Universal Prosodic Features in Motherese. *Developmental Psychology*, 24(1), 14–20. <https://doi.org/10.1037/0012-1649.24.1.14>
- Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutor Quant Methods Psychol*, 8(1), 23–34. <https://doi.org/10.1093/carcin/9.8.1355>
- Hargus Ferguson, S., & Quené, H. (2014). Acoustic correlates of vowel intelligibility in clear and conversational speech for young normal-hearing and elderly hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 135(6), 3570–3584. <https://doi.org/10.1121/1.4874596>
- Harkness, S. (1977). Aspects of the social environment and first language acquisition in rural Africa. *Talking to Children: Language Input and Acquisition*, 309–318.
- Hart, B., & Risley, T. R. (1995). Meaningful differences in the everyday experience of young American children. In *Paul H Brookes Publishing*.
- 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. <https://doi.org/10.3109/14992027.2012.658972>

- Herdan, G. (1960). *Type-token mathematics (Vol. 4)*. Mouton.
- Hill, P. D. (1985). Kernel Estimation of a Distribution Function. *Communications in Statistics - Theory and Methods*, 14(3), 605–620. <https://doi.org/10.1080/03610928508828937>
- Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., ... Suma, K. (2015). The Contribution of Early Communication Quality to Low-Income Children's Language Success. *Psychological Science*, 26(7), 1071–1083. <https://doi.org/10.1177/0956797615581493>
- Hochberg, I., Boothroyd, A., Weiss, M., & Hellman, S. (1992). Effects of noise and noise suppression on speech perception by cochlear implant users. *Ear and Hearing*, 13(4), 263–271. <https://doi.org/10.1097/00003446-199208000-00008>
- Hoff-Ginsberg, E. (1986). Function and Structure in Maternal Speech. Their Relation to the Child's Development of Syntax. *Developmental Psychology*, 22(2), 155–163. <https://doi.org/10.1037/0012-1649.22.2.155>
- Hoff-Ginsberg, E. (1991). Mother-Child Conversation in Different Social Classes and Communicative Settings. *Child Development*, 62(4), 782–796.
- Hoff-Ginsberg, E. (1994). Influences of Mother and Child on Maternal Talkativeness. *Discourse Processes*, 18(1), 105–117. <https://doi.org/10.1080/01638539409544886>
- Hoff-Ginsberg, E. (1998). The relation of birth order and socioeconomic status to children's language experience and language development. *Applied Psycholinguistics*, 19(4), 603–629. <https://doi.org/10.1017/s0142716400010389>
- Hoff-ginsberg, E., & Hoff-ginsberg, E. (1985). Some contributions of mothers' speech to their children's syntactic growth Some contributions of mothers' speech to their children's syntactic growth *. *Journal of Child Language*, 12(2), 367–385. <https://doi.org/10.1017/S0305000900006486>
- Hoff-Ginsberg, E., & Shatz, M. (1982). Linguistic input and the child's acquisition of language. *Psychological Bulletin*, 92(1), 3–26. <https://doi.org/10.1037/0033-2909.92.1.3>
- Hoff, E. (2003). *The Specificity of Environmental Influence: Socioeconomic Status Affects Early Vocabulary Development Via Maternal Speech*. 74(5), 1368–1378. <https://doi.org/10.1111/1467-8624.00612>
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, 26(1), 55–88. <https://doi.org/10.1016/j.dr.2005.11.002>

- Hoff, E., & Naigles, L. (2002). How children use input in acquiring a lexicon. *Child Development*, 73(2), 418–433.
- Holden, L. K., Finley, C. C., Firszt, J. B., Holden, T. A., Brenner, C., Potts, L. G., & Skinner, M. W. (2013). Factors Affecting Open-Set Word Recognition in Adults with Cochlear Implants. *Ear and Hearing*, 34(3), 342. <https://doi.org/10.1038/jid.2014.371>
- Holt, R. F., Beer, J., Kronenberger, W. G., Pisoni, D. B., & Lalonde, K. (2012). Contribution of family environment to pediatric cochlear implant users' speech and language outcomes: some preliminary findings. *Journal of Speech, Language, and Hearing Research : JSLHR*, 55(3), 848–864. [https://doi.org/10.1044/1092-4388\(2011/11-0143\)](https://doi.org/10.1044/1092-4388(2011/11-0143))
- Holt, R. F., & Svirsky, M. A. (2008). An Exploratory Look at Pediatric Cochlear Implantation: Is Earliest Always Best? *Ear and Hearing*, 29(4), 492–511. <https://doi.org/10.1097/AUD.0b013e31816c409f>
- Hurtado, N., Marchman, V. A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science*, 11(6), 31–39. <https://doi.org/10.1111/j.1467-7687.2008.00768.x>
- Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive psychology*, 61(4), 343–365. *Cognitive Psychology*, 61(4), 343–365. <https://doi.org/10.1016/j.cogpsych.2010.08.002>
- Huttenlocher, Janellen, Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early Vocabulary Growth: Relation to Language Input and Gender. *Developmental Psychology*, 27(2), 236–248. <https://doi.org/10.1037/0012-1649.27.2.236>
- Huttenlocher, Janellen, Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology*, 45(3), 337–374. [https://doi.org/10.1016/S0010-0285\(02\)00500-5](https://doi.org/10.1016/S0010-0285(02)00500-5)
- Iglehart, F. (2009). Combined effects of classroom reverberation and noise on speech perception by students with typical and impaired hearing. Paper presented at Inter-Noise 2009, Ottawa, Ontario, Canada. *Inter-Noise*. Ottawa, Ontario, Canada.
- Iler Kirk, K., Elizabeth Ying, I., Indianapolis, M., Richard Miyamoto, I. T., Tara, I. O., Cara Lento, I. L., ... Chesterfield, M. (2002). Effects of age At Implantation in Young Children. *Ann Oral Rhinal Laryngol III*, 69–73. Retrieved from <http://journals.sagepub.com/doi/pdf/10.1177/00034894021110S515>
- Jahn, K. N., & Arenberg, J. G. (2019). Evaluating Psychophysical Polarity Sensitivity as an Indirect Estimate of Neural Status in Cochlear Implant Listeners. *JARO - Journal of the*

- Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, 44(4), 548–567. <https://doi.org/10.1006/jmla.2000.2755>
- Jusczyk, P. W., Hohne, E. A., & Bauman, A. (1999). Infants' sensitivity to allophonic cues for word segmentation. *Perception and Psychophysics*, 61(8), 1465–1476. <https://doi.org/10.3758/BF03213111>
- Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The Beginnings of Word Segmentation in English-Learning Infants. *Cognitive Psychology*, 39(3–4), 159–207. <https://doi.org/10.1006/cogp.1999.0716>
- Kirk, K. I., Hay-McCutcheon, M., Sehgal, S. T., & Miyamoto, R. T. (2000). Speech perception in children with cochlear implants: Effects of lexical difficulty, talker variability, and word length. *Annals of Otology, Rhinology and Laryngology*, 109(12 II SUPPL.), 79–81. <https://doi.org/10.1177/0003489400109s1234>
- Kirk, K. I., Miyamoto, R. T., Ying, E. a, Perdew, a E., & Zuganelis, H. (2000). Cochlear implantation in young children: Effects of age at implantation and communication mode. *Volta Review*, 102(4), 127–144.
- Kirk, K. I., Pisoni, D. B., & Osberger, M. J. (1995). Lexical Effects on Spoken Word Recognition by Pediatric Cochlear Implant Users. *Ear and Hearing*, 16(5), 470. <https://doi.org/10.1097/MCA.0000000000000178>.Endothelial
- Kondaurova, M. V., Bergeson, T. R., & Dilley, L. C. (2012). Effects of deafness on acoustic characteristics of American English tense/lax vowels in maternal speech to infants. *The Journal of the Acoustical Society of America*, 132(2), 1039–1049. <https://doi.org/10.1121/1.4728169>
- Kondaurova, M. V., Bergeson, T. R., & Xu, H. (2013). Age-related changes in prosodic features of maternal speech to prelingually deaf infants with cochlear implants. *Infancy*, 18(5), 825–848. <https://doi.org/10.1111/infa.12010>
- Kong, Y.-Y., Stickney, G. S., & Zeng, F.-G. (2005). Speech and melody recognition in binaurally combined acoustic and electric hearing. *The Journal of the Acoustical Society of America*, 117(3), 1351–1361. <https://doi.org/10.1121/1.1857526>
- Kuhl, P K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences of the United States of America*, 97(22), 11850–11857. <https://doi.org/10.1073/pnas.97.22.11850>

- Kuhl, Patricia K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience*, 5(11), 831–843. <https://doi.org/10.1038/nrn1533>
- Kuhl, Patricia K. (2007). Is speech learning “gated” by the social brain? *Developmental Science*, 10(1), 110–120. <https://doi.org/10.1111/j.1467-7687.2007.00572.x>
- Lazard, D. S., Bordure, P., Lina-Granade, G., Magnan, J., Meller, R., Meyer, B., ... Truy, E. (2010). Speech perception performance for 100 post-lingually deaf adults fitted with Neurelec cochlear implants: Comparison between Digisonic® Convex and Digisonic® SP devices after a 1-year follow-up. *Acta Oto-Laryngologica*, 130(11), 1267–1273. <https://doi.org/10.3109/00016481003769972>
- Lehet, M., Arjmandi, M., Dilley, L., & Houston, D. (2018). Accuracy of the Language ENvironment Analysis (LENA) system for quantifying adult speech experienced by infants in naturalistic settings. *Journal of Speech Language and Hearing Research, Under revi.*
- Lehet, M., Arjmandi, M. K., Houston, D., & Dilley, L. C. (n.d.). Circumspection in using automated measures: Talker gender and addressee affect error rates for adult speech detection in the Language ENvironment Analysis (LENA) system. *Behavior Research Methods*.
- Levine, Æ. D., Strother-garcia, K., Golinkoff, M., & Hirsh-pasek, Æ. K. (2016). *Language Development in the First Year of Life : What Deaf Children Might Be Missing Before Cochlear Implantation*. (19), 56–62.
- Levitin, D. J., Menon, V., Schmitt, J. E., Eliez, S., White, C. D., Glover, G. H., ... Reiss, A. L. (2003). Neural correlates of auditory perception in Williams syndrome: An fMRI study. *NeuroImage*, 18(1), 74–82. <https://doi.org/10.1006/nimg.2002.1297>
- Li, Y., Zhang, G., Kang, H., Liu, S., Han, D., & Fu, Q.-J. (2011). Effects of speaking style on speech intelligibility for Mandarin-speaking cochlear implant users. *The Journal of the Acoustical Society of America*, 129(6), EL242–EL247. <https://doi.org/10.1121/1.3582148>
- Light, R. J. (1971). Measures of response agreement for qualitative data : some generalizations and alternatives. *Psychological Bulletin*, 76(5), 365.
- MacWhinney, B. (2017). *Tools for Analyzing Talk Part 2 : The CLAN Program*. (2000). <https://doi.org/10.21415/T5G10R>
- Malvern, D., Richards, B., Chipere, N., & Durán, P. (2004). *Lexical diversity and language development*. New York: Palgrave Macmillan.
- Masnan, M. J., Mahat, N. I., Shakaff, A. Y. M., Abdullah, A. H., Zakaria, N. Z. I., Yusuf, N., ...

- Aziz, A. H. A. (2015). Understanding Mahalanobis distance criterion for feature selection. *AIP Conference Proceedings*, 1660(February 2015). <https://doi.org/10.1063/1.4915708>
- Mattys, S. L., & Jusczyk, P. W. (2001). Phonotactic cues for segmentation of fluent speech by infants. In *Cognition* (Vol. 78). [https://doi.org/10.1016/S0010-0277\(00\)00109-8](https://doi.org/10.1016/S0010-0277(00)00109-8)
- McCartney, K. (1984). Effect of quality of day care environment on children's language development. *Developmental Psychology*, 20(2), 244–260. <https://doi.org/10.1037/0012-1649.20.2.244>
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, 22(3), 276–282. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/23092060> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3900052>
- McMurray, B., Ellis, T. P., & Apfelbaum, K. S. (2019). How Do You Deal With Uncertainty? Cochlear Implant Users Differ in the Dynamics of Lexical Processing of Noncanonical Inputs. *Ear and Hearing*, 40(4), 961–980. <https://doi.org/10.1097/AUD.0000000000000681>
- Mehler, A. C. and J. (2001). Perception of Prosodic Boundary Correlates by Newborn Infants. *Infancy*, 2(3), 385–394.
- Montag, J. L., Jones, M. N., & Smith, L. B. (2018). Quantity and Diversity: Simulating Early Word Learning Environments. *Cognitive Science*, 42, 375–412. <https://doi.org/10.1111/cogs.12592>
- Munson, B., & Nelson, P. B. (2005). Phonetic identification in quiet and in noise by listeners with cochlear implants. *The Journal of the Acoustical Society of America*, 118(4), 2607–2617. <https://doi.org/10.1121/1.2005887>
- Naigles, L. R., & Hoff-Ginsberg, E. (1998). Why are some verbs learned before other verbs? Effects of input frequency and structure on children's early verb use. *Journal of Child Language*, 25(1), 95–120. <https://doi.org/10.1017/S0305000997003358>
- National Institutes of Health, Cochlear Implants. (2010). *NIH Cochlear Implant Fact Sheet, Updated Oc.* Retrieved from [https://report.nih.gov/NIHfactsheets/Pdfs/CochlearImplants\(NIDCD\).pdf](https://report.nih.gov/NIHfactsheets/Pdfs/CochlearImplants(NIDCD).pdf) <https://www.nidcd.nih.gov/staticresources/health/hearing/FactSheetCochlearImplant.pdf>
- Nazzi, T., Dilley, L. C., Jusczyk, A. M., Shattuck-Hufnagel, S., & Jusczyk, P. W. (2005). English-learning infants' segmentation of verbs from fluent speech. *Language and Speech*, 48(3), 279–298. <https://doi.org/10.1177/00238309050480030201>

- Neuman, A. C., Wroblewski, M., Hajicek, J., & Rubinstein, A. (2004). Measuring Speech Recognition in Children With Cochlear Implants in a Virtual Classroom. *Journal of Speech Language and Hearing Research, 13*(6), 1684–1692. [https://doi.org/10.1044/1092-4388\(2011/11-0058\)b](https://doi.org/10.1044/1092-4388(2011/11-0058)b)
- Neuman, A. C., Wroblewski, M., Hajicek, J., & Rubinstein, A. (2010a). Combined effects of noise and reverberation on speech recognition performance of normal-hearing children and adults. *Ear and Hearing, 31*(3), 336–344. <https://doi.org/10.1097/AUD.0b013e3181d3d514>
- Neuman, A. C., Wroblewski, M., Hajicek, J., & Rubinstein, A. (2010b). Combined effects of noise and reverberation on speech recognition performance of normal-hearing children and adults. *Ear and Hearing, 31*(3), 336–344. <https://doi.org/10.1097/AUD.0b013e3181d3d514>
- Neville, H. J., & Bruer, J. T. (2001). Language processing: How experience affects brain organization. *Critical Thinking about Critical Periods*, 151–172.
- Newman, R. S., Rowe, M. L., & Bernstein Ratner, N. (2016). Input and uptake at 7 months predicts toddler vocabulary: The role of child-directed speech and infant processing skills in language development. *Journal of Child Language, 43*(5), 1158–1173. <https://doi.org/10.1017/S0305000915000446>
- Nicely, P., Tamis-LeMonda, C. S., & Bornstein, M. H. (1999). Mothers' attuned responses to infant affect expressivity promote earlier achievement of language milestones. *Infant Behavior and Development, 22*(4), 557–568. [https://doi.org/10.1016/S0163-6383\(00\)00023-0](https://doi.org/10.1016/S0163-6383(00)00023-0)
- 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. [https://doi.org/10.1044/1092-4388\(2007/073\).Will](https://doi.org/10.1044/1092-4388(2007/073).Will)
- Nicholas, J. G., & Geers, A. E. (2006). Effects of early auditory experience on the spoken language of deaf children at 3 years of age. *Ear and Hearing, 27*(3), 286–298. <https://doi.org/10.1097/01.aud.0000215973.76912.c6>
- Niparko, J. K., & Tobey, E. A. (1989). *Spoken Language Development in Children Following Cochlear Implantation. 15*(15), 1498–1506. <https://doi.org/10.1001/jama.2010.451.Spoken>
- Niparko, J. K., Tobey, E. A., Thal, D. J., Eisenberg, L. S., Wang, N.-Y., Quittner, A. L., & Fink, N. E. (2010). Spoken language development in children following cochlear implantation. *Jama, 303*(15), 1498–1506. <https://doi.org/10.1001/jama.2010.451>
- Oetting, J. B., Hartfield, L. R., & Pruitt, S. L. (2009). Exploring LENA as a Tool for Researchers

- and Clinicians. *ASHA Leader*, 14(6), 20–22.
<https://doi.org/10.1044/leader.FTR3.14062009.20>
- Pan, B. A., Rowe, M. L., Singer, J. D., & Snow, C. E. (2005). Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Development*, 76(4), 763–782.
<https://doi.org/10.1111/j.1467-8624.2005.00876.x>
- Parker, M. D., & Duluth, M. (2005). *A comparative study between mean length of utterance in morphemes (MLUm) and mean length of utterance in words (MLUw)*. 25(3), 365–376.
<https://doi.org/10.1177/0142723705059114>
- Parmanen, J. (2007). A-weighted sound pressure level as a loudness/annoyance indicator for environmental sounds - Could it be improved? *Applied Acoustics*, 68(1), 58–70.
<https://doi.org/10.1016/j.apacoust.2006.02.004>
- Peng, S.-C., Tomblin, J. B., & Turner, C. (2008). Production and Perception of Speech Intonation in Pediatric Cochlear Implant Recipients and Individuals with Normal Hearing. *Ear and Hearing*, 29(3), 336–351. <https://doi.org/10.1097/AUD.0b013e318168d94d>
- Peng, Z., Hess, C., Saffran, J. R., Edwards, J. R., & Litovsky, R. Y. (2019). *Assessing Fine-Grained Speech Discrimination in Young Children With Bilateral Cochlear Implants*.
<https://doi.org/10.1097/MAO.0000000000002115>
- Peterson, N. R., Pisoni, D. B., & Miyamoto, R. T. (2010). Cochlear implants and spoken language processing abilities: Review and assessment of the literature. *NIH Public Access*, 28(2), 237–250.
- Pisoni, D. B. (2012). *Cognitive Factors and Cochlear Implants - Some Thoughts on Perception, Learning, and Memory in Speech Perception*. 21(1), 70–78.
- Pisoni, D. B., & Cleary, M. (2012). *Measures of Working Memory Span and Verbal Rehearsal Speed in Deaf Children after Cochlear Implantation*. 24, 1–23.
<https://doi.org/10.1097/01.AUD.0000051692.05140.8E.Measures>
- Ramírez-Esparza, N., García-Sierra, A., & Kuhl, P. K. (2014). Look who's talking: Speech style and social context in language input to infants are linked to concurrent and future speech development. *Developmental Science*, 17(6), 880–891. <https://doi.org/10.1111/desc.12172>
- Reed, C. M., & Delhorne, L. A. (2005). Reception of environmental sounds through cochlear implants. *Ear and Hearing*, 26(1), 48–61. <https://doi.org/10.1097/00003446-200502000-00005>
- Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., &

- Gabrieli, J. D. E. (2018). Beyond the 30-Million-Word Gap: Children's Conversational Exposure Is Associated With Language-Related Brain Function. *Psychological Science*, 29(5), 700–710. <https://doi.org/10.1177/0956797617742725>
- Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech vocabulary development. *Child Development*, 83(5), 1762–1774. <https://doi.org/10.1111/j.1467-8624.2012.01805.x>
- Rowe, M. L., Raudenbush, S. W., & Goldin-Meadow, S. (2012). The pace of vocabulary growth helps predict later vocabulary skill. *Child Development*, 83(2), 508–525. <https://doi.org/10.1111/j.1467-8624.2011.01710.x>
- Rüter, M. (2011). Einfluss von Expansionen auf den Grammatikerwerb von Kindern mit Cochleaimplantat [The influence of expansions on the acquisition of grammar in children with cochlear implants]. *Hno*, 59(4), 360–365. <https://doi.org/10.1007/s00106-010-2234-3>
- Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language*, 47(1), 172–196. <https://doi.org/10.1006/jmla.2001.2839>
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-months-old infants. *Science*, 274(5294), 1926–1928. <https://doi.org/10.1126/science.274.5294.1926>
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70(1), 27–52. [https://doi.org/10.1016/S0010-0277\(98\)00075-4](https://doi.org/10.1016/S0010-0277(98)00075-4)
- Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word Segmentation: The Role of Distributional Cues. *Journal of Memory and Language*, 35(4), 606–621. <https://doi.org/10.1006/jmla.1996.0032>
- Saffran, & Kirkham. (2018). Infant Statistical Learning. *Annu Rev Psychol.*, 165(2), 255–269. <https://doi.org/10.1016/j.trsl.2014.08.005>.The
- Santos, J. F., Cosentino, S., Hazrati, O., Loizou, P. C., & Falk, T. H. (2013). Objective speech intelligibility measurement for cochlear implant users in complex listening environments. *Speech Communication*, 55(7–8), 815–824. <https://doi.org/10.1016/j.specom.2013.04.001>
- Santos, J. F., & Falk, T. H. (2014). Updating the SRMR-CI Metric for Improved Intelligibility Prediction for Cochlear Implant Users. *IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING*, 22(12), 2197–2206. <https://doi.org/10.1109/TASLP.2014.2363788>
- Schachner, A., & Hannon, E. E. (2011). Infant-Directed Speech Drives Social Preferences in 5-

- Month-Old Infants. *Developmental Psychology*, 47(1), 19–25.
<https://doi.org/10.1037/a0020740>
- Schafer, E. C., & Thibodeau, L. M. (2006). Speech recognition in noise in children with cochlear implants while listening in bilateral, bimodal, and FM-system arrangements. *American Journal of Audiology*, 15(2), 114–126. [https://doi.org/10.1044/1059-0889\(2006/015\)](https://doi.org/10.1044/1059-0889(2006/015))
- Schomer, P. D., Suzuki, Y., & Saito, F. (2002). Evaluation of loudness-level weightings for assessing the annoyance of environmental noise. *The Journal of the Acoustical Society of America*, 110(5), 2390–2397. <https://doi.org/10.1121/1.1402116>
- Schultz, T. J. (1978). Synthesis of social surveys on noise annoyance. *The Journal of the Acoustical Society of America*, 64(2), 377–405. <https://doi.org/10.1121/1.382013>
- Shaneh, M., & Taheri, A. (2009). *Voice Command Recognition System Based on MFCC and VQ Algorithms*. 3(9), 2231–2235.
- Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech Recognition with Primarily Temporal Cues. *Science*, 270(5234), 303–304.
<https://doi.org/10.1126/science.270.5234.303>
- Shannon, Robert V. (2002). The Relative Importance of Amplitude Temporal Spectral Cues for CIs. *The American Journal of Audiology*, Vol. 11, pp. 124–127.
- Sharma, A, Dorman, M., & Spahr, A. (2002). A sensitive period for the development of the central auditory system in children with cochlear implants: implications for age of implantation. *Ear and Hearing*, 23(6), 532–539.
<https://doi.org/10.1097/01.AUD.0000042223.62381.01>
- Sharma, Anu, Dorman, M. F., & Kral, A. (2005). The influence of a sensitive period on central auditory development in children with unilateral and bilateral cochlear implants. *Hearing Research*, 203(1–2), 134–143. <https://doi.org/10.1016/j.heares.2004.12.010>
- Sharma, Anu, Tobey, E., Dorman, M., Bharadwaj, S., Martin, K., Gilley, P., & Kunkel, F. (2004). Central Auditory Maturation and Babbling Development in Infants with Cochlear Implants. *Archives of Otolaryngology - Head and Neck Surgery*, 130(5), 511–516.
<https://doi.org/10.1001/archotol.130.5.511>
- Shneidman, L., Arroyo, M. E., Levine, S. C., & Goldin-Meadow, S. (2013). What counts as effective input for word learning? *Journal of Child Language*, 40(3), 672–698.
<https://doi.org/10.3174/ajnr.A1256.Functional>
- Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive*

- Sciences*, 7(8), 343–348. [https://doi.org/10.1016/S1364-6613\(03\)00156-6](https://doi.org/10.1016/S1364-6613(03)00156-6)
- Soderstrom, M., & Wittebolle, K. (2013). *When Do Caregivers Talk ? The Influences of Activity and Time of Day on Caregiver Speech and Child Vocalizations in Two Childcare Environments*. 8(11). <https://doi.org/10.1371/journal.pone.0080646>
- Soleymani, R., Selesnick, I. W., & Landsberger, D. M. (2018). SEDA: A tunable Q-factor wavelet-based noise reduction algorithm for multi-talker babble. *Speech Communication*, 96(October 2017), 102–115. <https://doi.org/10.1016/j.specom.2017.11.004>
- Spencer, L., Barker, B., & Tomblin, J. (2003). Exploring the language and literacy outcomes of pediatric cochlear implant users. *Ear and Hearing*, 24(3), 236–247. <https://doi.org/10.1097/01.AUD.0000069231.72244.94.Exploring>
- Sperry, D. E., Sperry, L. L., & Miller, P. J. (2018). *Reexamining the Verbal Environments of Children From Different Socioeconomic Backgrounds*. 00(0), 1–16. <https://doi.org/10.1111/cdev.13072>
- Spitzer, S., Liss, J., Spahr, T., Dorman, M., & Lansford, K. (2009). The use of fundamental frequency for lexical segmentation in listeners with cochlear implants. *The Journal of the Acoustical Society of America*, 125(6), EL236–EL241. <https://doi.org/10.1121/1.3129304>
- Stacey, P. C., Fortnum, H. M., Barton, G. R., & Summerfield, A. Q. (2006). Hearing-impaired children in the United Kingdom, I: Auditory performance, communication skills, educational achievements, quality of life, and cochlear implantation. *Ear and Hearing*, 27(2), 161–186. <https://doi.org/10.1097/01.aud.0000202353.37567.b4>
- Stallings, L. M., Gao, S., & Svirsky, M. A. (2000). Assessing the language abilities of pediatric cochlear implant users across a broad range of ages and performance abilities. *Volta Review*, 102(4), 215–235.
- Su, J., Rubin, J., & Terman, D. (2004). Effects of noise on people. *Nonlinearity*, 17(1), 133–157. <https://doi.org/10.1088/0951-7715/17/1/009>
- Summerfield, A. Q., & Marshall, D. H. (1995). Preoperative predictors of outcomes from cochlear implantation in adults: Performance and quality of life. *Annals of Otology, Rhinology and Laryngology*, 106, 105.
- Svirsky, M. A. (2017). Cochlear implants and electronic hearing. *Physics Today*, 70(8), 53–58. <https://doi.org/10.1063/PT.3.3661>
- Svirsky, M. A., Robbins, A. M., Kirk, K. I., Pisoni, D. B., & Miyamoto, R. T. (2000). Language Development in Profoundly Deaf Children with Cochlear Implants. *Psychological Science*,

11(2), 153–158. <https://doi.org/10.1111/1467-9280.00231>

- Svirsky, M. A., Teoh, S., & Neuburger, H. (2004a). Development of Language and Speech Perception in Congenitally Profoundly Deaf Children as a function of age at cochlear implantation. *Audiology and Neurotology*, 9(4), 224–233.
- Svirsky, M. A., Teoh, S. W., & Neuburger, H. (2004b). Development of language and speech perception in congenitally, profoundly deaf children as a function of age at cochlear implantation. *Audiology and Neurotology*, 9(4), 224–233.
- 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., & Rüter, M. (2009). The influence of parents’ speech on the development of spoken language in German-speaking children with cochlear implants. *Revista de Logopedia, Foniatria y Audiologia*, 29(3), 165–173. [https://doi.org/10.1016/S0214-4603\(09\)70025-7](https://doi.org/10.1016/S0214-4603(09)70025-7)
- Szagun, G., & Schramm, S. A. (2016). Sources of variability in language development of children with cochlear implants: Age at implantation, parental language, and early features of children’s language construction. *Journal of Child Language*, 43(3), 505–536. <https://doi.org/10.1017/S0305000915000641>
- Szagun, G., & Stumper, B. (2012). Age or Experience? The Influence of Age at Implantation and Social and Linguistic Environment on Language Development in Children With Cochlear Implants. *Journal of Speech, Language, and Hearing Research*, 55(6), 1640–1654.
- Tamati, T. N., Janse, E., Pisoni, D. B., & Baskent, D. (2017). Talker variability in real-life speech recognition by cochlear implant users. *The Journal of the Acoustical Society of America*, 141(5), 2017–2020. <https://doi.org/10.1121/1.4989293>
- Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children’s achievement of language milestones. *Child Development*, 72(3), 748–767. <https://doi.org/10.1111/1467-8624.00313>
- Tamis-LeMonda, C. S., Bornstein, M. H., Kahana-Kalman, R., Baumwell, L., & Cyphers, L. (1998). Predicting variation in the timing of language milestones in the second year: An events history approach. *Journal of Child Language*, 25(3), 675–700. <https://doi.org/10.1017/S0305000998003572>
- Tamis-LeMonda, C. S., Kuchirko, Y., & Song, L. (2014). Why Is Infant Language Learning Facilitated by Parental Responsiveness? *Current Directions in Psychological Science*, 23(2), 121–126. <https://doi.org/10.1177/0963721414522813>

- Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant directed speech facilitates word segmentation. *Infancy*, 7(1), 53–71.
- 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(Supplement), 36S–45S.
<https://doi.org/10.1097/01.AUD.0000051688.48224.A6>
- Tobey, E. A., Shin, S., Sundarajan, M., & Geers, A. E. (2011). Spoken word recognition in adolescent cochlear implant users during quiet and multispeaker babble conditions. *Otology and Neurotology*, 32(3), 413–418. <https://doi.org/10.1097/MAO.0b013e31820d9613>
- Tobey, E. A., Thal, D., Niparko, J. K., Eisenberg, L. S., Quittner, A. L., & Wang, N. Y. (2013). Influence of implantation age on school-age language performance in pediatric cochlear implant users. *International Journal of Audiology*, 52(4), 219–229.
<https://doi.org/10.3109/14992027.2012.759666>
- Tomblin, J. B., Barker, B. A., Spencer, L. J., Zhang, X., & Gantz, B. J. (2005). *The Effect of Age at Cochlear Implant Initial Stimulation on Expressive Language Growth in Infants and Toddlers*. 48(4), 853–867. [https://doi.org/10.1044/1092-4388\(2005/059\).The](https://doi.org/10.1044/1092-4388(2005/059).The)
- Tyler, R. S., Teagle, H. F., Kelsay, D. M., Gantz, B. J., Woodworth, G. G., & Parkinson, A. J. (2000). Speech perception by prelingually deaf children after six years of Cochlear implant use: effects of age at implantation. *The Annals of Otology, Rhinology {&} Laryngology. Supplement*, 185, 82–84.
- van Geert, P., & van Dijk, M. (2002). Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behavior and Development*, 25(4), 340–374.
[https://doi.org/10.1016/S0163-6383\(02\)00140-6](https://doi.org/10.1016/S0163-6383(02)00140-6)
- Verspoor, M., Lowie, W., & Van Dijk, M. (2008). Variability in second language development from a dynamic systems perspective. *Modern Language Journal*, 92(2), 214–231.
<https://doi.org/10.1111/j.1540-4781.2008.00715.x>
- Walker, D., Greenwood, C., Hart, B., & Carta, J. (1994). Prediction of School Outcomes Based on Early Language Production and Socioeconomic Factors. *Child Development*, 65(2), 606–621. <https://doi.org/10.1111/j.1467-8624.1994.tb00771.x>
- Walker, E. a, & McGregor, K. K. (2013). *Word learning processes in children with cochlear implants*. 56(2), 375–387. [https://doi.org/10.1044/1092-4388\(2012/11-0343\).Word](https://doi.org/10.1044/1092-4388(2012/11-0343).Word)
- Wang, Ye, Hartman, M., Aziz, N. A. A., Arora, S., Shi, L., & Tunison, E. (2017). A Systematic Review of the Use of LENA Technology. *American Annals of the Deaf*, 162(3), 295–311.

<https://doi.org/10.1353/aad.2017.0028>

- Wang, Yuanyuan, Bergeson, T. R., & Houston, D. M. (2017). Infant-Directed Speech Enhances Attention to Speech in Deaf Infants With Cochlear Implants. *Journal of Speech Language and Hearing Research*, 60(11), 3321. https://doi.org/10.1044/2017_JSLHR-H-17-0149
- Wang, Yuanyuan, Bergeson, T. R., & Houston, D. M. (2018). Preference for Infant-Directed Speech in Infants With Hearing Aids: Effects of Early Auditory Experience. *Journal of Speech, Language & Hearing Research*, 61(9), 2431–2439. https://doi.org/10.1044/2018_JSLHR-H-18-0086
- Wang, Yuanyuan, Shafto, C. L., & Houston, D. M. (2018). Attention to speech and spoken language development in deaf children with cochlear implants: a 10-year longitudinal study. *Developmental Science*, (July 2017), e12677. <https://doi.org/10.1111/desc.12677>
- Weisleder, A., & Fernald, A. (2013). Talking to Children Matters: Early Language Experience Strengthens Processing and Builds Vocabulary. *Psychological Science*, 24(11), 2143–2152.
- Weizman, Z. O., & Snow, C. E. (2001). Lexical input as related to children's vocabulary acquisition: effects of sophisticated exposure and support for meaning. *Developmental Psychology*, 37(2), 265–279. <https://doi.org/10.1037/0012-1649.37.2.265>
- Werker, J. F., Pegg, J. E., & McLeod, P. J. (1994). A cross-language investigation of infant preference for infant-directed communication. *Infant Behavior and Development*, 17(3), 323–333. [https://doi.org/10.1016/0163-6383\(94\)90012-4](https://doi.org/10.1016/0163-6383(94)90012-4)
- Whitmal, N. A., Poissant, S. F., Freyman, R. L., & Helfer, K. S. (2007). Speech intelligibility in cochlear implant simulations: Effects of carrier type, interfering noise, and subject experience. *The Journal of the Acoustical Society of America*, 122(4), 2376–2388. <https://doi.org/10.1121/1.2773993>
- Willstedt-Svensson, U., Löfqvist, A., Almqvist, B., & Sahlén, B. (2004). Is age at implant the only factor that counts? The influence of working memory on lexical and grammatical development in children with cochlear implants. *International Journal of Audiology*, 43(9), 506–515. <https://doi.org/10.1080/14992020400050065>
- Wilson, B. S., & Dorman, M. F. (2008). Cochlear implants: A remarkable past and a brilliant future. *Hearing Research*, 242(1–2), 3–21. <https://doi.org/10.1016/j.heares.2008.06.005>
- Yukawa, K., Cohen, L., Blamey, P., Pyman, B., Tungvachirakul, V., & O'Leary, S. (2004). Effects of insertion depth of cochlear implant electrodes upon speech perception. *Audiology and Neuro-Otology*, 9(3), 163–172. <https://doi.org/10.1159/000077267>

- Zanto, T. P., Hennigan, K., Östberg, M., Clapp, W. C., & Gazzaley, A. (2013). Effect of speaking rate on recognition of synthetic and natural speech by normal-hearing and cochlear implant listeners. *Ear and Hearing, 34*(3), 313.
<https://doi.org/10.1016/j.cortex.2009.08.003>. Predictive
- Zeng, F.-G., Nie, K., Stickney, G. S., Kong, Y.-Y., Vongphoe, M., Bhargave, A., ... Cao, K. (2005). Speech recognition with amplitude and frequency modulations. *Proceedings of the National Academy of Sciences, 102*(7), 2293–2298.
<https://doi.org/10.1073/pnas.0406460102>
- Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). *Preschool Language Scales—Fifth Edition (PLS-5)*. Bloomington, MN: Pearson.