

NORTHWESTERN UNIVERSITY

Language Processing Differences in U.S. Monolingual English-Speaking Late Talkers and
Typical Talkers: Specificity of Phonological Representations
and Sensitivity to Lexical-Semantic Relationships

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Communication Sciences and Disorders

By
Philip R Curtis

EVANSTON, ILLINOIS

June 2021

ABSTRACT

Late talkers (LTs) are a heterogeneous group of children who experience delayed language development in the absence of other known causes (Fisher, 2017). Many LTs “catch up” to their typically talking peers (TTs), but some go on to experience continued difficulties with speech and language development (Dale & Hayiou-Thomas, 2013), with little variance in long-term outcomes predicted by current language measures (Fisher, 2017). While typical language measures used in clinical practice and research with LTs involve broad-based language assessments, fine-grained measures of LTs’ language processing may elucidate mechanisms of early language delay and offer incremental utility in predicting later outcomes. In this dissertation, a group of TTs and LTs completed two eye-tracking experiments to assess two fine-grained aspects of language processing. Children were from U.S. monolingual English-speaking speaking families and the sample was mostly white and of higher socioeconomic status.

In Study One, children completed a mispronunciation Looking-While-Listening (LWL) paradigm (Swingley & Aslin, 2000; White & Morgan, 2008). Children heard words that were pronounced correctly or contained a word-initial mispronunciation. Results revealed that both LTs and TTs were sensitive to these mispronunciations, but that LTs’ processing may have been more disrupted than that of TTs. These results suggest that, at least on word-initial phonemes, LTs have created accurate auditory-phonological representations.

In Study Two, children completed a target-absent LWL paradigm (Chow et al., 2017; Huettig & McQueen, 2007) to assess children’s sensitivity to taxonomic semantic relationships. Results showed that both LTs and TTs were sensitive to the taxonomic relationships used in the experiment and looked to semantic competitors in the target-absent trials. Both groups also

looked more to the target in the target-present condition than in the target-absent condition, suggesting both groups had formed specific, differentiated representations of these lexical items.

Taken together, these results reveal both similarities and differences in how LTs and TTs process language. Future research should incorporate additional aspects of language processing, such as assessments of children's motor-phonological representations, and include more diverse participants that are representative of the target population.

ACKNOWLEDGEMENTS

I would like to dedicate this dissertation to the memory of my late sister, Frieda, whose creativity, passion, and love continue to inspire me every day.

This work would not have been possible without tremendous support and mentorship from many people throughout my graduate training. First and foremost, my wonderful committee. Thank you to my first co-advisor, Megan Roberts, who taught me to be passionate and relentless in using research to improve the lives of children and families. Your energy and drive are always a source of inspiration to me. Thank you to my second co-advisor, Adriana Weisleder, who taught me to think deeply but also to take a step back (and take a deep breath) when needed. Thank you for keeping your door (and Zoom room) always open, and for helping me think through the many, many big questions, as well as the little details that came up along the way. Your thoughtfulness and incredible depth of knowledge are so inspiring. Additionally, your passion for racial and ethnic equity in research has deeply shaped how I think and what I want to accomplish in my career. And thank you to Ryne Estabrook, who patiently mentored me and encouraged me to think critically about the statistics we use to draw inferences from data. You have inspired me to be cautious and thorough in my design and analysis choices.

I had the pleasure and privilege of being in two amazing labs: the Early Intervention Research Group (EIRG) and the Child Language Lab (CLL). Both of these labs are filled with the most incredible, passionate, thoughtful, creative, and dedicated students, staff, and clinicians. You have all been *constant* sources of inspiration to me. Also, I owe everyone in the EIRG a massive debt of gratitude for looking at dozens of pictures of noodles, coats, jackets, boots, pizza, horses, etc. I think all your effort improved this study immensely. I would like to thank

Hannah Fipp-Rosenfield who recorded the audio stimuli for both experiments in this project. Thank you so much for your patience and time spent in many recording sessions. This study literally couldn't have happened without your vocal talents. To McKenna Lanter: I cannot imagine completing this study without you. You were such an integral part at each step of this journey. You were always thorough, attentive, and helped me to laugh along the way. To Talia Apter: you have contributed to this project in so many ways and I am so grateful for your careful attention to detail. I would also like to thank Rahma Almajid for your time and effort in reviewing the many surveys that parents completed and ensuring the quality of the data we collected. To Laura Sudec: thank you for always answering my questions about the many logistical aspects of running a research study – your passion and thoroughness are truly remarkable. Thank you to Tara Salomon, Katie Zanzinger, and Bailey Sone for your clinical mentorship and collaboration throughout this project. Thank you also to Anele Villanueva, Murielle Standley, and Gaby Espana for your thoughtfulness in reviewing these documents with an eye towards equity and inclusion.

Thank you to the many amazing clinical supervisors I have worked with throughout my time at Northwestern. Your passion and amazing depth of knowledge helped me to keep myself grounded in the purpose of this research project: to build knowledge that will hopefully help children and their families. You have been such a source of inspiration to me during my time at Northwestern.

And to Yael, my constant support and companion throughout our entire graduate school journey. I will never be able to express my gratitude and appreciation for you. Thank you for pushing me to get out of the weeds and look at the big picture and for helping me to think deeply

and make connections. Thank you for your constant emotional support – through frustrations and setbacks, you found ways to make me laugh and persevere. I cannot imagine the last seven years without you!

Thank you to my family for your *endless* support and encouragement throughout this process. You helped me to put things in perspective and to see the big picture. You always expressed excitement and curiosity about this project, which in turn re-ignited my own passion for this work. And thank you to Mark, who not only provided amazing amounts of emotional support, but lent me your creativity in finding new ways to recruit families.

Finally, thank you to the families who participated in this project. It was an honor and privilege to work with each family, and I cannot express my gratitude for their time and energy. Even during a global pandemic, they continued to engage in this work.

Funding sources

This study was supported by a National Institute of Deafness and Communication Disorders fellowship (F31DC017631) as well as a scholarship from Council of Academic Programs in Communication Sciences and Disorders and a Northwestern Graduate Research Grant to P.C., and a NIDCD grant to A.W. (R21DC018357).

TABLE OF CONTENTS

ABSTRACT.....	2
ACKNOWLEDGEMENTS.....	4
TABLE OF CONTENTS.....	7
LIST OF TABLES.....	9
LIST OF FIGURES.....	10
Chapter 1: Introduction.....	11
Chapter 2: Sensitivity to Mispronunciations in U.S. Monolingual Later Talkers and Typical Talkers.....	15
Phonological Development in Early Childhood.....	16
Phonological Development in Late Talkers.....	19
The Current Study.....	21
Methods.....	22
Participants.....	22
Measures.....	23
Procedure.....	27
Eye Tracking Data Preparation.....	28
Analysis Plan.....	29
Results.....	30
Window Analysis.....	32
Growth Curve Analyses.....	35
Exploratory Analyses.....	37
Discussion.....	38
Design Considerations.....	43
Limitations and Constraints to Generality.....	45
Conclusion.....	47
Chapter 3: Sensitivity to Semantic Relationships in U.S. Monolingual English-Speaking Typical Talkers and Late Talkers.....	49
Introduction.....	49
Sensitivity to Semantic Relationships Among TT Toddlers.....	50
Semantic Deficits in Children with Language Disorders.....	51

The Current Study	52
Methods	52
Participants	52
Measures	53
Procedure	57
Eye Tracking Data Preparation	57
Analysis Plan	58
Results	60
Window Analysis	62
Growth Curve Modeling	65
Reaction Time Analyses	66
Alternative Explanation and Exploratory Analysis: Syllable Structure Matching	67
Discussion	70
Type of Semantic Knowledge Assessed	71
Receptive vs. Expressive Tasks	72
Heterogeneity Among Typical and Late Talkers	73
Sensitivity to Syllable Structure Similarities	73
Individual Differences and Future Language Development	74
Limitations and Constraints on Generality	75
Conclusion	76
Chapter 4: General Discussion	78
Auditory, Phonological, and Motor Speech Processes	79
Heterogeneity in causal mechanisms within LTs	81
Additional Considerations	82
Future Directions	85
Limitations	86
Conclusion	87
REFERENCES	88

LIST OF TABLES

Table 2.1. <i>Lexical Stimuli Properties</i>	25
Table 2.2. <i>Participant Characteristics</i>	30
Table 2.3. <i>Window Analysis, 300 - 1800 ms</i>	32
Table 2.4. <i>Estimated Marginal Means by Talker Type and Phonological Condition in Log-Odds Units</i>	32
Table 2.5. <i>Window Analysis, 300 - 1800 ms, Continuous Vocabulary Scores</i>	34
Table 2.6. <i>Growth Curve Analysis, 300 - 1800 ms</i>	37
Table 3.1. <i>Semantic Similarity Ratings within Stimuli Quartets</i>	55
Table 3.2. <i>Participant Characteristics</i>	60
Table 3.3. <i>Window Analysis, 300-1800 ms, LT and Target-Absent Conditions as Reference</i>	62
Table 3.4. <i>Estimated Marginal Means in Log-Odds Units from Main Window Analysis Model, 300-1800 ms</i>	62
Table 3.5. <i>Window Analysis with Vocabulary as Continuous Predictor, 300-1800 ms</i>	64
Table 3.6. <i>Reaction Time Analysis on Distractor-Initial Trials</i>	66
Table 3.7. <i>Reaction Time Analysis on Target-Initial Trials</i>	67

LIST OF FIGURES

Figure 2.1. <i>Stimulus Presentation</i>	26
Figure 2.2. <i>Time Course from 300 – 1800 ms</i>	31
Figure 2.3. <i>Proportion Looks to Target by Talker Type and Phonological Condition</i>	33
Figure 2.4. <i>Relation Between Productive Vocabulary Size and Proportion Looks to Target by Phonological Condition</i>	35
Figure 2.5. <i>Item-Level Plots of Proportion Looks to the Target</i>	38
Figure 3.1. <i>Example Target-Present and Target-Absent Trial Sequences</i>	56
Figure 3.2. <i>Proportion Looks to the Target in Main Window Analysis, 300-1800 ms</i>	63
Figure 3.3. <i>Growth Curve Model by Talker Type and Semantic Condition</i>	65
Figure 3.4. <i>Growth Curve of Target-Absent Trials by Talker Status and Syllable Cue Status</i>	69

Chapter 1: Introduction

There is a great deal of variability in children's early language development (Fenson et al., 2000; Fenson et al., 1994; Fenson et al., 2006). While general guidelines exist for major linguistic milestones, such as children saying their first words around the age of 12 months and producing their first word combinations around the age of 18 months, there is variation in both the onset and trajectory of language acquisition. This variability can be difficult to interpret for both caregivers (Yimgang et al., 2017) and clinicians (Singleton, 2018), as delays in vocabulary development or word combinations may simply reflect normal variation in development or be early indications of persistent language learning difficulties (Dale & Hayiou-Thomas, 2013; Singleton, 2018).

Children who experience delayed expressive language development are typically referred to as "Late Talkers" (LTs); other terms used in the literature are early language delay (Dale et al., 2003), delayed expressive language development (Hawa & Spanoudis, 2014), expressive language delay (Fischel et al., 1989), children with late language emergence (Rice et al., 2008), and specific expressive language delay (Rescorla & Schwartz, 1990). Definitions of LTs vary widely but typically utilize a measure of caregiver-reported vocabulary size, such as the Language Development Survey (Rescorla, 1989) or MacArthur-Bates Communicative Development Inventory (Fenson et al., 2006) and percentile-score thresholds of the 10th, 15th, or 20th percentiles using age-based normative scores from these measures. Other methods of categorizing children as LTs or typical talkers (TTs) include using specific vocabulary sizes (50 words or 70 words) and sometimes the absence of 2-word-combinations by 24 months of age.

The prevalence of LTs is estimated to be anywhere between 10-20% of children (Desmarais et al., 2008; Klee et al., 1998; Rescorla, 2002).

While both caregivers and clinicians may display concern when children's language development seems to lag behind their same-aged peers, the majority of LTs go on to "recover" from their initial delay. Their performance on language measures often fall in the typical range, although their performance may continue to fall below that of children who had no history of LT (Rescorla, 2005, 2009). In addition to the variability in defining LTs mentioned earlier, there is also variability in terms and criteria used to refer to children with persistent language learning difficulties. This variability in definitions makes it difficult to give relative risk of continued language deficits for LTs and TTs. Defining LTs as children who had fewer than 70 words or no two-word combinations, Rice et al. (2008) found that 20% of former LTs achieved general language scores one standard deviation or more below normative levels at age 7, compared to only 11% of children who had not been LTs. Using cutpoints of the 10th percentile on the MCDI at age 2 and the 10th percentile on the Peabody Picture Vocabulary Test at age 4, Hammer et al. (2017) found that 25% of LTs continued to have vocabulary deficits at age 4 (an odds ratio of about 4 in a logistic regression model controlling for covariates). These studies suggest that while the majority of LTs will not show clinically significant language impairments later in development, LTs are at increased risk compared to their TT peers.

While numerous studies have attempted to identify predictors that might help clinicians to identify children most at risk for continued delays, even statistically significant predictors account for only a small percentage of variance in long-term outcomes (Fisher, 2017). Previous studies have found that the degree of delay in expressive language skills, receptive language

abilities, and socioeconomic status are significant predictors of later outcomes (Fisher, 2017). These studies have mostly relied on broad behavioral measures of children's language abilities, such as parent checklists or standardized language assessments. While these tools are clinically efficient and useful, recent work in communication sciences and disorders and development psychology has explored more fine-grained aspects of children's language development that may add incremental utility in predicting later development.

Several decades of research has revealed that children's developing language systems contain many of the complex psycholinguistic processes observed in adults that suggest the presence of interactive architecture. For instance, during eye tracking paradigms, toddlers have been shown to process speech incrementally. Toddlers can identify a target object given only partial phonetic information for that word (Fernald et al., 2001) and are slowed in identifying a target image if the initial phonemes of that target overlap with those of a competitor image (Swingley et al., 1999). Toddlers activate both phonological and semantic competitors during language processing (Chow et al., 2017) and are sensitive to even minor changes in pronunciation (Mani et al., 2008; Mani & Plunkett, 2007; Swingley, 2009; White & Morgan, 2008). Priming studies have demonstrated that children are sensitive to phonological (Mani & Plunkett, 2011b), semantic (Arias-Trejo & Plunkett, 2009), phono-semantic (Angulo-Chavira & Arias-Trejo, 2018; Mani et al., 2012), and semantic-phono (Angulo-Chavira & Arias-Trejo, 2018) primes. These results suggest that children's early lexicons display both lexical-semantic and lexical-phonological organization and elucidate the early connections that children form between lexical items in their lexicons. Other work has shown that toddlers can use verbs (Mani & Huettig, 2012) and thematic relationships between objects (Bobb et al., 2016) to predict future

referents during online language processing. Taken together, these findings suggest that toddlers' lexicons already display many of the complex, interactive processes observed in adults.

Assessing some of these processes in LTs may both elucidate specific disrupted systems or processes that underly these children's difficulty with language acquisition, and potentially serve as predictors of later language development. Work by (Fernald & Marchman, 2012) has demonstrated that measuring the *speed* of children's familiar-word processing shows promise as a predictor of later development among LTs. It could be that assessing other aspects of children's online language processing may offer further predictive and diagnostic capabilities. The current project sought to explore two such fine-grained aspects of children's language processing: the specificity of children's early phonological representations, and children's sensitivity to taxonomic relationships among familiar lexical-semantic concepts.

This dissertation is organized as follows. In Chapter 2, I describe a study investigating typical and late talkers' sensitivity to mispronunciations in order to assess the specificity of their internal phonological representations for early-acquired words. In Chapter 3, I describe a study investigating typical and late talkers' sensitivity to taxonomic relationships, in order to assess children's lexical-semantic organization. In Chapter 4, I provide an overall discussion of the studies presented in chapters 2 and 3, as well as a discussion of implications and direction for future research.

Chapter 2: Sensitivity to Mispronunciations in U.S. Monolingual Later Talkers and Typical Talkers

One of the most common concerns that parents raise during well-child visits is delayed language development (Coonrod & Stone, 2004; Yimgang et al., 2017). This high rate of concerns about communication development is consistent with epidemiological data that show that approximately 10 - 20% of 24-month-olds experience delayed communication development (Collisson et al., 2016; Horwitz et al., 2003; Reilly et al., 2007; Zubrick et al., 2007). These children are frequently referred to as “late talkers” (LTs) or children with late language emergence. Children are classified as LTs when they present with a reduced expressive vocabulary in comparison with their age-matched peers (variably defined as either at or below the 10th, 15th, or 20th percentile on a parent-reported expressive vocabulary checklist), or when they are not combining words by the age of 24 months. The long-term outcomes for these LTs are highly variable. Approximately 20% of LTs exhibit language below expected levels at 7 years of age, compared to 11% of TTs (Rice et al., 2008). Others go on to develop language in the low-average range and may struggle with reading and other academic skills later in life (Rescorla, 2009). A core barrier in clinical intervention for this population is the lack of quality predictors that exist to determine which children will recover spontaneously, and which children would benefit from early language intervention.

Many clinical assessments used to assess the language abilities of toddlers rely on parent-reported vocabulary checklists or standardized behavioral assessments of children’s general language abilities. However, lower scores on these measures may be caused by a number of underlying differences in speech-language processing (Stackhouse & Wells, 1997). Difficulties

in multiple psycholinguistic processes, such as poor phonological encoding (Nash & Donaldson, 2005), weak semantic representations (Nash & Donaldson, 2005), and difficulties with phonological production planning (Stackhouse & Wells, 1997) may all result in the common presentation of a reduced expressive vocabulary. However, different processes may have different long-term outcomes and may require different forms of intervention.

Phonological Development in Early Childhood

In order to become competent and efficient language users, children need to acquire the phonological system(s) of the language(s) in their environment, a process which is closely tied to children's ability to form lexical representations. The development of a child's phonological system can be observed both in children's language *production* as well as in children's language *perception*. Children begin to produce consonant-vowel combinations, frequently referred to as canonical babbling, at about 6 months of age (Fagan, 2009) and the frequency of canonical babbling increases as children age (Cychosz et al., 2019). Children's early phonemic repertoires consist of only a few early-acquired speech sounds (Robb & Bleile, 1994) and syllable structures (Levelt et al., 1999). Both phonetic inventory and syllable structure diversity increase as children get older. At around 12 months, children say their first words. These first words typically contain a limited number of speech sounds and a limited set of syllable shapes, and are more likely to contain sounds that are already in children's productive phonemic repertoire (Davis et al., 2018; Schwartz & Leonard, 1982). Assessment of children's early phonological development typically involve assessing children's phonemic inventories, including phonemes produced in multiple positions within word productions, and the diversity of syllable structures children

produce (Paul & Jennings, 1992). These measures are thought to reflect children's acquisition of the rule-based phonological system of the language or languages to which they are exposed.

While these linguistic behaviors are evidence of the productive manifestations of children's developing phonological systems, other work has used perception and comprehension tasks to investigate the nature of children's internal representations of lexical forms. Early theories posited that infants formed indistinct, holistic representations of the phonological structure of words they understand (Charles-Luce & Luce, 1995; Metsala & Walley, 1998), and only create representations with fine-grained phonetic detail as their lexicons expand and acquire words with similar phonological structures that require precise encoding. However, a preponderance of evidence has accumulated demonstrating that early in development, children are sensitive to small phonemic differences in word comprehension tasks. For example, Fennell and Werker (2003) found that 14-month-olds were sensitive to contrasts in known words that differed on only one phonological feature (e.g., 'ball' and 'doll').

One method for assessing children's internal phonological representations is the Looking While Listening (LWL) Mispronunciation paradigm (Swingley & Aslin, 2000). In this paradigm, children are shown two images on the screen representing familiar objects (e.g., dog, ball) and hear a word naming one of the images (e.g., *Look! Dog!*). There are two conditions in this paradigm: the correct pronunciation condition, in which the target item is labeled correctly, and the mispronunciation condition, in which children hear the target label with a mispronunciation on one phoneme. Children's looking behavior (most often the proportion of time spent looking at the target image) is then compared between the two conditions to derive a measure of 'mispronunciation sensitivity'.

Using this paradigm, Swingley and Aslin (2002) found that 14- and 15-month-olds were sensitive to mispronunciations on the onsets of familiar words. These findings have been replicated in multiple conditions across the second year of life. Toddlers can detect mispronunciations across multiple phonological features (White & Morgan, 2008) and are sensitive to mispronunciations on word onsets (Swingley & Aslin, 2000; White & Morgan, 2008), vowels (Mani & Plunkett, 2011a), codas (Swingley, 2009), and lexical tone (in languages which possess this feature) (Wewalaarachchi et al., 2017). Taken together, these findings suggest that young children without known language delays or disorders have formed fine-grained phonological representations of words in their receptive vocabularies prior to age 2.

School-aged children are similarly able to detect mispronunciations of varying kinds in both auditory lexical discrimination tasks (McNeill & Hesketh, 2010) and word identification tasks (Krueger & Storkel, 2020). While typically developing school-aged children readily detect mispronunciations in the auditory lexical discrimination task, children diagnosed with Developmental Language Disorder are less sensitive to these mispronunciations (Claessen & Leitão, 2012; Van Alphen et al., 2004). When asked whether a word produced by the examiner is correct or incorrect in these tasks, children with DLD are more likely to accept mispronunciations as correct, while their typically-developing peers typically identify them as incorrect. Claessen and Leitão (2012) suggest that children with DLD have poorly specified phonological forms for these words, leading to difficulty in detecting errors in others' productions.

These findings suggest that children's sensitivity to mispronunciations, or more generally the specificity of their phonological representations, may be an indicator of their language

abilities. There is some empirical evidence to support this idea. For instance, Werker et al. (2002) found that when 14-month-olds were taught new words that differed by only one phoneme ('bih' and 'dih'), only infants with larger vocabularies could detect changes between the two words. Using a mispronunciation paradigm, Law and Edwards (2015) found that 30- to 46-month-olds' sensitivity to mispronunciations was associated with their vocabulary size: children with larger vocabularies were more sensitive to mispronunciations. However, other studies have failed to find any correlation between vocabulary size and sensitivity to mispronunciations (e.g., Swingley & Aslin, 2002). While these studies looked at children with typical language development, fewer studies have investigated the specificity of phonological representations in children who are LTs.

Phonological Development in Late Talkers

As a group, LTs experience delays in multiple areas of expressive phonological development, including the onset of canonical babbling, the size of their early consonant inventories, the complexity of early learned words, the inclusion of atypical phonological patterns in their word productions, and increased variability in sound productions (Williams & Elbert, 2003). These delays persist into the preschool period for some LTs (Paul, 1993; Whitehurst et al., 1991). In a sample of 4- to 5-year-olds with a history of late talking, Neam et al. (2020) found that these children had significant difficulty producing polysyllabic words, compared to typically developing peers. Additionally, the degree to which LTs' phonological systems are delayed – assessed productively via consonant and syllable shape inventories - has been found in some studies to be predictive of their later language development (Carson et al., 2003). Delays in LTs' expressive phonology may be the result of difficulty with the motoric

aspect of speech production, or may result from difficulty retrieving phonological representations (Oller et al., 1999). Observations of children's language production are not sufficient to differentiate between difficulties with motor-speech production and difficulties with the creation or retrieval of phonological representations.

A few studies have used the LWL mispronunciation paradigm to investigate links between children's lexical phonological representations and language ability. For instance, Höhle et al. (2009) administered a LWL mispronunciation task to a group of children at 18 months; then, at 30 months, children were classified as having either typical language development or low language performance. The children classified as having typical language at 30 months had shown sensitivity to mispronunciations at 18 months, looking to the target image only when it was labeled correctly. The children classified as having low language performance at 30 months, on the other hand, only looked to the target image when the label was mispronounced, and not when it was pronounced correctly. The authors suggest that this difference may be due to the low language performance group having unstable phonological representations.

Another study used a variant of the LWL mispronunciation task to compare children concurrently classified as LTs and TTs (MacRoy-Higgins et al., 2013). The authors taught groups of LTs and TTs 12 novel words over 10 sessions. The authors then used multiple measures to assess children's learning of these nonwords, including traditional comprehension measures that require the child to point to physical objects, production tasks that required children to label the novel objects, and a LWL mispronunciation task. When data from the LWL task was analyzed separately for LTs, they showed no evidence of sensitivity to mispronunciations for the novel words. One interpretation of these findings is that LTs have less

precise phonological representations and, like older children with DLD, ‘accept’ mispronunciations as correct. However, LTs also performed worse than TTs on traditional comprehension measures of the novel words. Thus, it may be that LTs did not show sensitivity to the mispronunciations in this task because they had not had enough experience with the novel words. Evidence suggests that children with language difficulties require many more exposures to novel words before they are added to the child’s lexicon. It may be that with additional exposures to the novel words, LTs may have shown sensitivity in the mispronunciation condition.

In order to better understand the extent to which LTs form detailed phonological representations of words after repeated exposure, the current study investigates LTs’ sensitivity to mispronunciations for familiar real words. The results from this study will elucidate whether LTs possess fine-grained phonological representations for familiar words they have been exposed to many times, in contrast to the novel words used by MacRoy-Higgins et al. (2013). The results will also help us understand whether difficulties in creating detailed phonological representations may be a causal mechanism in LTs’ delayed expressive vocabulary development.

The Current Study

In the current study, we administered a LWL mispronunciation task to a sample of LTs and TTs. The study was designed to test whether LTs show a mispronunciation effect as TTs do when the stimuli involve familiar items. Our hypothesis is that LTs as a group, like older children with DLD, will show reduced sensitivity to mispronunciations compared to the TTs.

Methods

Participants

Participants were recruited through online advertisements, flyers distributed in the community, and through existing research databases. Participants were eligible if they had a child between the ages of 22 and 30 months of age, reported no concerns about their child's hearing, reported that their child only heard English at home, the child was born at 37 weeks gestation or later, and if the caregiver had no concerns that their child may have an autism spectrum disorder. Additional inclusion criteria were assessed during research visits. These included a T-Score of 35 or higher on the Mullen Scales of Early Learning – Visual Reception scale (MSEL-VR), and a score of less than 2 on the Screening Tool for Autism in Toddlers (STAT). Two participants did not complete a second study visit due to a university-mandated suspension of research activities due to the Covid-19 pandemic. For this reason, these children did not complete the STAT. Instead, we utilized the Childhood Autism Rating Scales, 2nd Edition (CARS-2). Both children scored in the “minimal to no risk of autism spectrum disorder” category. Because recruitment and data collection were curtailed due to the COVID-19 pandemic, we may have reduced power to detect interaction effects of interest.

In total, 57 participants were enrolled in the study. Twelve participants were excluded from the analysis: six children received a score of 2 or higher on the STAT; two children received a T-Score of less than 35 on the Mullen; one child's caregiver did not complete the MCDI; and three children did not contribute at least 4 trials in each condition during the Looking-While-Listening task.

A total of 45 children were retained for analyses. Twenty-three children qualified as TTs, and 22 as LTs. Participant characteristics are given in Table 2.2. Children did not differ on STAT or MSEL-VR scores. Children in the LT group were significantly younger than children in the TT group ($t(42.24) = 2.60, p = 0.01$). Consistent with our analysis plan, we include child age as a covariate in all models, together with race (White vs. not White), biological sex, and caregiver education (college or less vs. post-secondary degree). A detailed discussion of how the age difference between groups might impact the results of the current study are presented in the discussion.

Measures

Screening Tool for Autism in Toddlers and Young Children (STAT; Stone & Ousley, 2008). The STAT is a Level 2 autism screener for children between the ages of 24 and 36 months, with studies demonstrating its validity in younger children as well (Stone et al., 2008). The STAT assesses three broad domains: Play, Social Communication, and Imitation Skills. Assessors use a series of presses throughout the assessment, including rolling a ball back and forth with the assessor, giving the child a sealed jar of snack that the child cannot open themselves, and giving the child a bag of toys. The STAT classifies children receiving a score of 2 or higher as having a higher likelihood of having autism. These children were referred for further assessment through Illinois' Early Intervention System.

Childhood Autism Rating Scale, Second Edition (CARS-2; Schopler et al., 2010). Two children in the study did not complete the STAT because their second study visit was cancelled due to the 2020 Covid-19 pandemic. To screen for the presence of autism in these children, we utilized the CARS-2. The CARS-2 is an observational coding system. Observers rate children on

15 items, including items that assess how the child relates to others, how the child uses objects, and how the child adapts to change. The CARS-2 can be scored based on observations of the child in a variety of settings. For the current study, a trained rater scored the CARS-2 based on recordings of a 15-minute caregiver-child interaction.

Mullen Scales of Early Learning, Visual Reception Scale (MSEL-VR; Mullen, 1995).

This subscale assesses children's non-verbal cognitive skills. Children complete a range of activities, such as completing a foam-board puzzle, making object associations, and matching shapes. The Visual Reception scale produces a T-Score based on the child's age. Children needed to achieve a T-Score of 35 (1.5 standard deviations below the mean) or higher to be included in the study.

MacArthur-Bates Communicative Development Inventory – Words and Sentences

(MCDI; Fenson et al., 2006). The MCDI is a caregiver report measure that assesses children's expressive communication development, including expressive vocabulary and syntax. The MCDI was used to classify children into the Later Talker (LT) and Typical Talker (TT) groups. Children who received a score at or below the 15th percentile (Beckage et al., 2011; Hodges et al., 2017; Lu et al., 2020), and/or children who had no reported two-word combinations were classified as LTs. All other children were classified as TTs.

Looking While Listening Mispronunciation Task

Design. Twelve monosyllabic words that are acquired early in development (Fenson et al., 2006) were selected. Six yoked pairs were created by matching word pairs on age of acquisition, lexical frequency in child-directed speech, and phonological neighborhood size. For each word, a mispronunciation was created by altering the initial phoneme's voicing (6 words) or

place of articulation (6 words). Previous work has shown that toddlers are sensitive to changes in both voicing and place of articulation in the mispronunciation paradigm (White & Morgan, 2008). See Table 2.1 for stimuli descriptions.

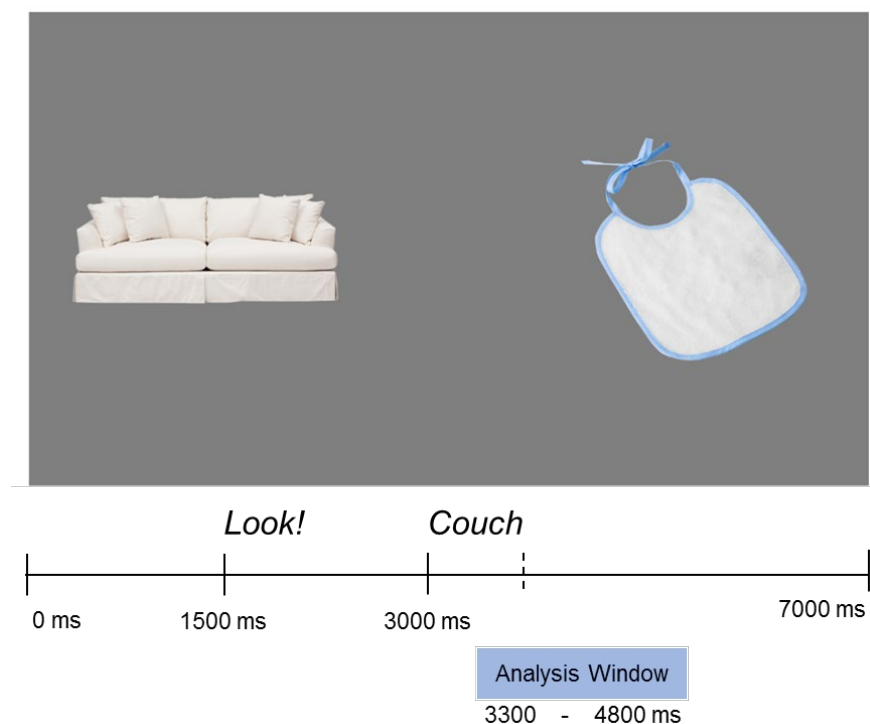
Table 2.1. *Lexical Stimuli Properties*

Yoked Pair	Word	Initial Phoneme (IPA)		Type of change
		Correct	Mispronunciation	
1	fork	/f/	/s/	Place
	purse	/p/	/b/	Voicing
2	foot	/f/	/v/	Voicing
	book	/b/	/p/	Voicing
3	duck	/d/	/g/	Place
	bed	/b/	/p/	Voicing
4	ball	/b/	/g/	Place
	car	/k/	/g/	Voicing
5	bib	/b/	/g/	Place
	couch	/k/	/t/	Place
6	bath	/b/	/g/	Place
	keys	/k/	/g/	Voicing

Each yoked pair was presented 4 times during the experiment. Each item was presented once as a target with the correct pronunciation, once as a target with the mispronunciation, and twice as a distractor for its yoked word. Four pseudo-random counterbalanced orders were created. In each order, side of target presentation was counterbalanced. Orders were adjusted so that the first item was a correct pronunciation and no more than two mispronunciations were presented consecutively. Additional filler trials with colorful images and cartoon characters were interspersed every 3 to 4 trials to maintain children's attention.

In each trial, two images were presented on the screen for a total of 7000 ms. After 1500 ms, children heard the speaker say "Look!" At 3000 ms, the target stimulus word was played. The images remained on the screen for another 4000 ms. The trials thus consisted of a pre-naming phase (0-3000 ms) and a post-naming phase (3000 – 7000 ms). See Figure 2.1.

Figure 2.1. *Stimulus Presentation*



Auditory Stimuli. Audio stimuli were recorded by a female General American English speaker who used infant-directed prosody. Multiple examples of each word were recorded. A group of adults then listened to both correct and mispronunciation word tokens in isolation and were asked to rate their clarity and to phonetically transcribe the word. Only tokens on which all listeners correctly identified the initial phoneme and had an average rating of 6 or higher on a 7-point Likert scale measuring the words' clarity were included as stimuli in the experiments. Words were on average 0.7315 (sd = 0.0856) seconds in duration. The correct pronunciation and mispronunciation word sets did not differ in their length ($t(21.784) = -0.378, p = 0.709$), mean pitch ($t(21.719) = 1.028, p = 0.315$), or listeners' average clarity ratings ($t(21.997) = 1.481, p = 0.153$).

Visual Stimuli. Two sets of images were selected for each yoked pair of stimuli (e.g., there were two image pairs for the duck-bed pair, two for the foot-book pair, etc.). For each word, multiple photographs were selected and rated by a group of adults. Adults were asked to rate the degree to which each image matched their internal representation of each word on a scale of 1-100. Images with a value of at least 80% were selected (average rating of 95.437). Image pairings for each yoked pair were created by analyzing the distribution of luminance values for the pixels in each picture. The imager package in R (Barthelme, 2018) was used to extract an array of luminance values for all pixels in each picture, and a Kolmogorov-Smirnov test was used to assess the similarity of each image's luminosity distribution. Image pairs with similar distributions were selected in order to reduce the likelihood that children would prefer one image because of its visual properties.

Images were resized and centered in a 600 x 600 pixel area. The narrower of the two images was adjusted so its maximum length would be 600 pixels. The other image was then resized so that it had approximately the same number of pixels. Images were placed on a 1920 x 1080 gray background (hex color code #7f7f7f).

Procedure

The LWL experiment was presented using Tobii Studio Pro. Children either sat in a highchair or in their caregiver's lap in a sound-attenuating area of the testing room. Caregivers were given a pair of occluded sunglasses to wear and were instructed not to repeat any of the target words and to only speak in order to redirect their child's attention to the stimuli. Visual stimuli were presented on a 24" ASUS VS248H-P monitor. Auditory stimuli were presented at 65 dB through Bose Companion 2 Series III Multimedia Speakers on either side of the monitor.

Children's gaze was captured using a Tobii X3-120 using a 120 Hz sampling rate and via video recording. Children completed the MSEL-VR and the STAT behavioral testing after a brief warm-up period while caregivers completed the MCDI and demographic forms.

Eye Tracking Data Preparation

Raw gaze data was exported from the Tobii Pro Studio software. The average gaze coordinates provided by Tobii Pro Studio were used in analyses. Gaps in data less than 75 ms, most likely representing blinks or brief periods of track loss, were interpolated using scripts from Unger, Vales, et al. (2020). An area of interest (AOI) was drawn around each image. Previous studies utilizing other Tobii products report lower precision and accuracy in toddler-aged participants (Dalrymple et al., 2018). Likewise, other eye tracking studies have demonstrated that specifying larger AOI's reduces the influence of noise in the data (Hessels et al., 2016). For this reason, we added an additional 4 degrees to all sides of the area of interest, resulting in AOI's that are 764 pixels x 764 pixels, with the image centered in the middle of the AOI.

Data from six participants were hand coded using peyecoder (v 1.1.13; Olson et al., 2020). Two participants had extreme track loss, despite optimal viewing position and angle. The first and last author then reviewed plots that aggregated fixations from each child in a cartesian plane to identify calibration drift. Plots were reviewed in a random order, and identifying information was removed from each plot. The authors were blind to LT status, and plots contained data from both correct pronunciation and mispronunciation conditions. From this inspection, four children were identified as having significant drift. Upon reviewing the videos for these children, it was observed that these children frequently shifted their head position during the experiment, despite remaining attentive. These children's data were hand-coded in

peycoder as well. A separate participant with optimal data from the Tobii eye tracker was also hand coded, and the Tobii gaze data was compared with hand-coded data. The correlation between proportion of looks to the target image in 100 ms bins was 0.93, demonstrating that the two methods give very similar results.

Data were processed in R (v 4.0.3) (R Core Team, 2014) using the eyetrackingR package (Dink & Ferguson, 2015) and the lme4 package (Bates et al., 2015). Looks to neither AOI (shifts, looking at parts of the monitor outside the AOI, trackloss, looks away from the monitor, etc. were removed). Data were then collapsed into 100 ms bins for growth curve analyses, or into one bin for window analyses.

Analysis Plan

Statistical models.

The analysis plan was pre-registered on OSF prior to analyzing any main effects (<https://osf.io/hda87>). In all analyses, the caregiver's education, the child's race (White vs. non-White), the child's age, and the child's biological sex were included as covariates. We utilized an analysis window of 300-1800 ms post-word onset (Fernald et al., 2008).

Window Analysis. Consistent with prior studies using the Looking While Listening (LWL) paradigm, we first conducted a window analysis (Swingley & Aslin, 2000; White & Morgan, 2008). In this type of analysis, the total amount of time spent looking at the target versus the distractor during the analysis window were entered into a model. We conducted the window analysis using a logistic mixed effects model, with talker status (LT vs TT), experimental condition (correct pronunciation vs. mispronunciation), and their interaction included in the model, along with covariates.

Growth Curve Modeling. The window analysis approach collapses all data within the analysis window, potentially obscuring fine-grained temporal differences in children’s processing of stimuli. In order to capitalize on the temporal precision offered by the LWL methodology, we analyzed data using growth curve modeling (Barr, 2008; Mirman et al., 2008). While previous studies have utilized an empirical logit transformation of gaze data, there are several statistical shortcomings with this approach (see Donnelly and Verkuilen (2017) for a detailed review and simulation). For this reason, we utilized a logistic mixed effects model, which uses a vector of successes (a sample in which the child looks to the target) and failures (a sample in which the child looks to the distractor) within each timebin or window (Donnelly & Verkuilen, 2017; Mirman, 2014).

Results

Table 2.2. *Participant Characteristics*

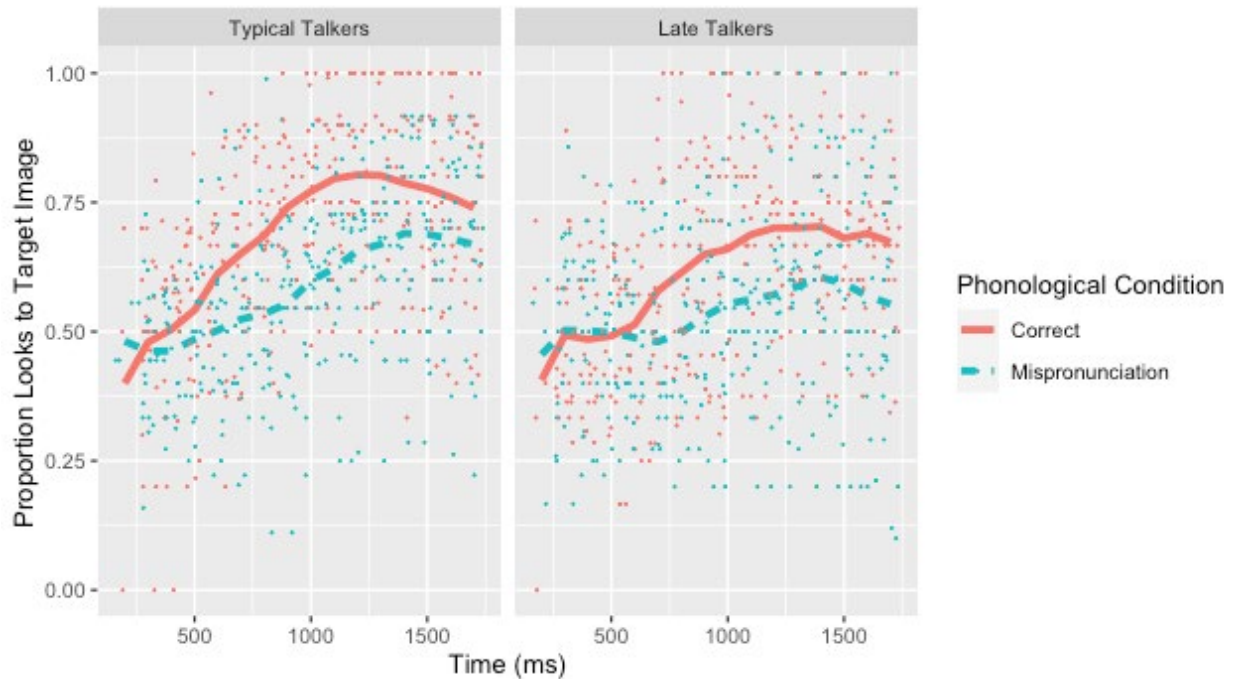
Variable	Typical Talkers	Late Talkers	d	p
	Mean (SD) or <i>n</i>			
N	23	22		
Age (Months)	26.27 (2.38)	24.57 (1.99)	0.77	0.01
STAT Score ^a	0.42 (0.34)	0.61 (0.50)	-0.45	0.15
MSEL-VR ^b	51.61 (19.51)	50.00 (8.24)	0.11	0.72
Caregiver education				
Some college	1	0		
College graduate	9	10		
Graduate/professional training or above	13	12		
Caregiver report of child race				
White	18	16		
Black	1	1		
Asian	0	1		
More than one race	4	4		

Note. Abbreviations: MSEL VR, Mullen Scales of Early Learning Visual Reception; STAT, Screening Tool for Autism in Toddlers

^aScore range of 0 – 4

^bT-Score (Mean = 50; SD = 10)

Figure 2.2. Time Course from 300 – 1800 ms



Out of a total of 1,080 trials, 867 (80.28%) were retained for analysis. Trials were excluded for the following reasons: equipment failure or aborted trial ($n=6$), child not looking before or after target word ($n=41$), noise during the critical word ($n=78$), child was inattentive during critical word ($n=13$), trackloss or inattentiveness ($n=75$). On average, TTs contributed 20.087, and LTs contributed 18.409 trials ($t=1.444$ (37.722), $p = 0.157$).

An analysis of looks to the target image during the pre-naming phase (-1500 to 0 ms before target word onset) revealed no systematic preference for the target image before naming (TTs = 50.633%, LTs = 48.737%, $t(40.173) = 0.811$, $p = 0.423$). Accordingly, we did not control for pre-naming gaze behavior in analyses.

Window Analysis

Looks to the target and looks to the distractor were entered into a logistic multilevel model which included random effects for participant, participant by condition, lexical item, and lexical item by condition. Both LTs and TTs were included in this model. Results for this analysis are given in Table 2.3. Estimated marginal means were obtained using the emmeans package (Lenth, 2020) in order to quantify TTs' and LTs' performance in each condition.

Table 2.3. *Window Analysis, 300 - 1800 ms*

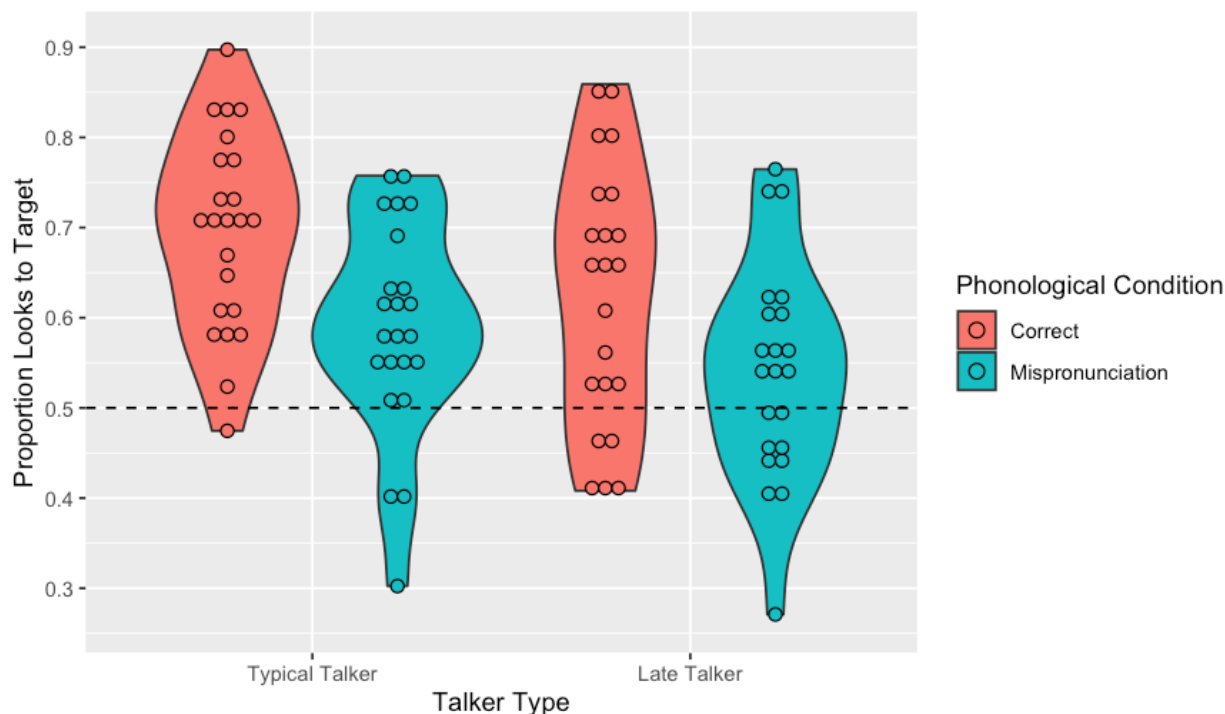
Predictors	Log-Odds	Std. Error	Wald Z	<i>p</i>
Intercept	0.966	0.176	5.479	< 0.001
Late Talker	-0.340	0.180	-1.894	0.058
Condition (Mispronunciation)	-0.569	0.150	-3.793	< 0.001
Late Talker : Condition	0.191	0.168	1.136	0.256

Table 2.4. *Estimated Marginal Means by Talker Type and Phonological Condition in Log-Odds Units*

	Estimated Marginal Mean (SE)	Asymptotic 95% CI
Correct Pronunciation Condition		
TTs	0.968 (0.176)	[0.623, 1.312]
LTs	0.627 (0.171)	[0.291, 0.963]
Mispronunciation Condition		
TTs	0.399 (0.176)	[0.054, 0.743]
LTs	0.250 (0.171)	[-0.086, 0.586]

Note. Abbreviations: LT, Late Talker; TT, Typical Talker

Figure 2.3. *Proportion Looks to Target by Talker Type and Phonological Condition*



The window analysis revealed that typical talkers looked more to the target image during the post-naming window in the correct pronunciation condition (Wald = 5.479, $p < .001$). Late Talkers looked less at the target during the correct pronunciation trials, although this was only marginally significant ($p = 0.058$). There was a highly significant effect of mispronunciation condition, showing that TTs looked less to the target image when it was mispronounced (Wald = -3.793, $p < .001$). The LT by Condition interaction did not reach significance ($p = 0.256$), although this interaction effect was in the hypothesized direction. This suggests that LTs, like TTs, were sensitive to mispronunciations.

To further investigate the performance of LTs, the same model was run for LTs only. Like TTs, LTs looked more to the target image than the distractor in the correct pronunciation

condition (Wald = 2.641, $p = 0.008$), and looked less to the target image in the mispronunciation condition than in the correct pronunciation condition (Wald = -1.926, $p = 0.054$). The effect of mispronunciation just failed to reach significance when analyzing LTs on their own. When analyzing TTs on their own, there was a significant effect of condition (Wald = -2.719, $p = 0.007$). The estimated marginal means in Table 2.4 revealed that TTs looked at above-chance levels to the target image in both conditions, while LTs looked at above-chance levels in only the correct pronunciation condition. In the mispronunciation condition, LTs did not look at above-chance levels to the target.

Productive Vocabulary as a Continuous Predictor.

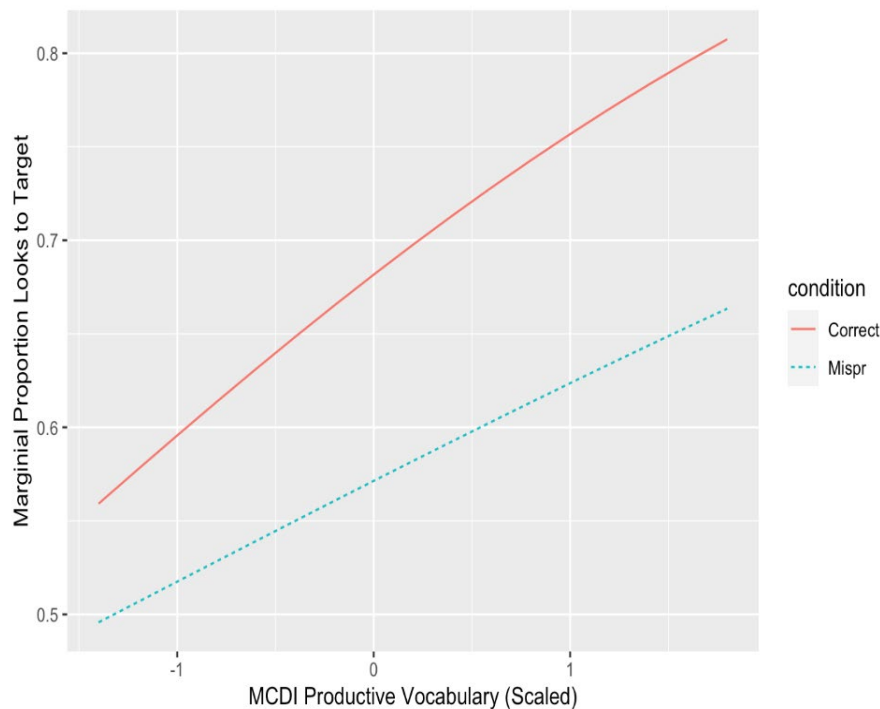
Because various definitions of Late Talkers exist in the literature, we also performed a window-analysis with total vocabulary scores from the MCDI entered as scaled continuous predictor in place of the talker status variable.

Table 2.5. *Window Analysis, 300 - 1800 ms, Continuous Vocabulary Scores*

Predictors	Log-Odds	Std. Error	Wald Z	p
Intercept	0.827	0.145	5.711	<0.001
MCDI Total Vocabulary	0.374	0.099	3.776	<0.001
Condition (Mispronunciation)	-0.474	0.124	-3.816	<0.001
MCDI : Condition	-0.156	0.085	-1.830	0.067

Note. Abbreviations: MCDI, MacArthur Bates Communicative Development Inventories

Figure 2.4. *Relation Between Productive Vocabulary Size and Proportion Looks to Target by Phonological Condition*



Once again, Mispronunciation Condition was highly significant (Wald = -3.816, $p < 0.001$), as was MCDI scores (Wald = 3.776, $p < 0.001$). As a group, children were sensitive to mispronunciations, and children with larger productive vocabularies looked to the target image more than children with smaller productive vocabularies in the correct pronunciation condition. The coefficient of the interaction term between MCDI and Condition was negative and marginally significant (Wald = -1.830, $p = 0.067$). Consistent with our hypothesis, there was a trend for children with larger vocabularies to show a greater sensitivity to mispronunciations.

Growth Curve Analyses

Growth curve analyses were conducted using logistic fixed effects models. The main model contained orthogonal polynomials for linear and quadratic time (beginning 300 ms after

stimulus onset to 1800 ms), as well as their interaction with talker status and mispronunciation condition and previously stated covariates. The model was first run with maximal random effects of intercept, linear time, and quadratic time nested in both participants and participants by condition. This model did not converge, however, and so the random effect of linear time nested within participant was removed, and the intercept and quadratic time random effects nested in participant were set to be uncorrelated (Barr et al., 2013).

There was a main effect of condition (Wald $Z = 7.437, p < 0.001$), demonstrating that on average, TTs (the reference group) looked to the target image less in the mispronunciation condition. There was also a main effect of talker type (Wald $Z = -2.025, p = 0.043$). Late talkers looked to the target image less in the correct pronunciation condition than do typical talkers. The main coefficients of interest investigating the effect of talker type were the two-way interaction between condition and talker type, and the three-way interactions between condition, talker type, and linear and quadratic time. None of these interactions were significant (talker by condition: Wald $Z = 1.177, p = 0.239$; talker by condition by linear time: Wald $Z = 0.055, p = 0.956$; talker by condition by quadratic time: Wald $Z = -0.587, p = 0.557$).

Table 2.6. *Growth Curve Analysis, 300 - 1800 ms*

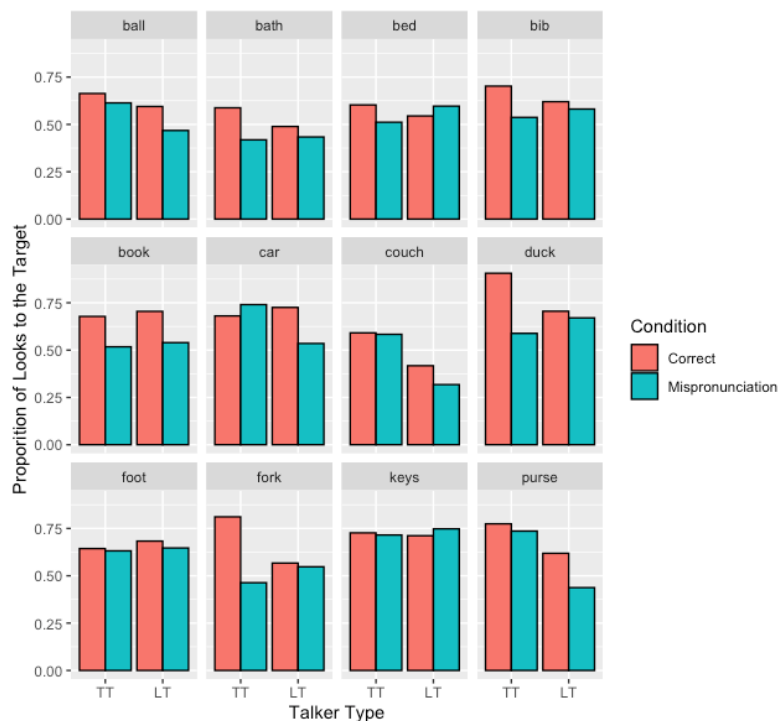
Predictors	Log-Odds	Std. Error	Wald Z	<i>p</i>
(Intercept)	1.080	0.145	7.437	<0.001
Phonological Condition	-0.649	0.140	-4.641	<0.001
Late Talker	-0.402	0.198	-2.025	0.043
Late Talker : Phonological Condition	0.236	0.200	1.177	0.239
Late Talker : Linear Time : Phonological Condition	0.048	0.866	0.055	0.956
Late Talker : Quadratic Time : Phonological Condition	-0.303	0.516	-0.587	0.557

Note. Abbreviations: LT, Late Talker; ot1, orthogonal time – linear; ot2, orthogonal time – quadratic

Exploratory Analyses

In order to explore item-by-item variability, effect plots were created for each item (see Figure 2.5). Considerable item-by-item variability existed in the effect of the mispronunciation condition between LTs and TTs. In order to analyze item characteristics that may have led to this variability, we constructed an additional model that included item-level variables. Previous literature has suggested that words' phonotactic probability may affect children's sensitivity to mispronunciations (MacRoy-Higgins et al., 2014). Because the design was not balanced to assess the impact of phonotactic probability, we included other item-level covariates, including age of acquisition, frequency in child-directed speech, number of phonological neighbors, and mispronunciation type (change in voicing or change in place). Random effects of participant, participant by condition, target word, and target word by condition were entered as well. In order to assess whether phonotactic probability had a differential impact on moderating the effect of condition for LTs compared to TTs, the main term of interest was entered as a 3-way-interaction between talker type, phonotactic probability, and mispronunciation condition.

Figure 2.5. *Item-Level Plots of Proportion Looks to the Target*



Results revealed that this 3-way-interaction was significant; phonotactic probability altered the mispronunciation effect, and the effect of phonotactic probability varied based on talker status.

Discussion

This study used a Looking-While-Listening paradigm to investigate typical talkers' and late talkers' sensitivity to mispronunciations. The primary research question was whether LTs, who as a group show differences in their *expressive* phonological development (Carson et al., 2003; Williams & Elbert, 2003), showed similar differences in their receptive phonological development. The present findings do not support our hypothesis that LTs show difficulty in forming specific phonological representations of frequently encountered words in their

environment, as both LTs and TTs showed sensitivity to mispronunciations for the early-acquired words used in the current study. When including vocabulary as a continuous predictor in statistical models, there is some evidence that children with smaller vocabularies may be less sensitive to mispronunciations than children with larger vocabularies. However, there was no significant effect of late talker status on sensitivity to mispronunciations, at least for the word-initial mispronunciations on monosyllabic words used in this study.

Although we did not observe differences in sensitivity to mispronunciations between LTs and TTs, we did observe differences in their processing of correctly pronounced and mispronounced familiar words. Compared to TTs, LTs looked less to the target image on correct pronunciation trials, suggesting they were less efficient in processing the familiar word and identifying its referent, consistent with previous studies (Fernald & Marchman, 2012). In addition, LTs looked less to the target image on mispronunciation trials relative to TTs, and were barely different from chance on this condition. This suggests that LTs struggled with identifying the referent on mispronunciation trials. Overall, this pattern of results suggests that both TTs and LTs were sensitive to the differences between correctly pronounced and mispronounced words, yet TTs were better able to recover from the mispronunciation to identify its more likely referent while LTs did not always associate the mispronounced word with the intended target. While mispronunciation paradigms have largely focused on analyzing the reduction in looking time in the mispronunciation condition compared to the correct pronunciation condition (Von Holzen & Bergmann, 2021), the current findings also draw attention to the fact that TTs were able to identify the target image even in the presence of a mispronunciation. LTs, on the other hand, did

not recognize these mispronunciations as well – even relatively minor changes in pronunciation almost completely disrupted LTs’ ability to understand the intended referent of the speaker.

What led TTs to identify the intended target in the mispronunciation condition, while LTs were less successful? Previous work has demonstrated that typically developing children, like adults, process speech incrementally (Chow et al., 2017; Swingley et al., 1999). For example, when looking at a picture of a baby, hearing the first sound in the word “dog” is sufficient for a two-year-old to shift their eyes away from the baby and towards the dog. When a listener encounters an initial phoneme that does not match that of any referent in the environment (here the two pictures on the screen), they must use the remaining phonological information in the auditory stimulus to select the most likely match (McMurray et al., 2019). Adult speakers are able to accommodate disturbances to voice onset times in word-initial position (McMurray et al., 2009). The current results suggest that TTs were also able to use phonological information in the auditory stimulus other than the word onset to infer the intended referent.

LTs’ diminished ability to recover after a mispronunciation may be caused by a number of factors. It may be that while LTs have well-specified representations of word-onsets, they have poorly specified representations of vowels or offsets, leading to difficulty in matching the remainder of the stimulus to the target referent. Another explanation may involve differences in TTs’ and LTs’ activation of the referents’ phonological forms prior to word onset. Some studies show that typically-developing infants activate the phonological form of a pictured object before hearing its name (Mani et al., 2012; Mani & Plunkett, 2010a). It may be that LTs in the current sample did not do this, or did so to a lesser degree than TTs. This in turn may have led to greater difficulty in matching the limited correct phonological information in the input to the less-

activated phonological forms of the referents. It could also be that LTs' overall decreased efficiency in language processing slowed their ability to recover from the mispronunciation. However, even when extending the time window to 3000 ms post word-onset, LTs' proportion of looks to the target in the mispronunciation condition remained close to chance.

As noted by Marchman and Fernald (2013), performance on the LWL may be impacted by multiple cognitive and linguistic processes. Given that differences between TTs and LTs were observed in both the correct pronunciation and mispronunciation conditions, it is possible that factors outside of phonological processing, such as working memory, attentional control, or general processing speed, may have impacted LTs' performance. Indeed, research suggests that older children with a diagnosis of developmental language disorder exhibit impairments on nonlinguistic processing tasks, in addition to language processing tasks (e.g., Kohnert et al., 2009; Leonard et al., 2007). Regardless of the cause of LTs' performance, it may be that in everyday settings, LTs are not only less efficient in processing the language in their environment (Fernald & Marchman, 2012), but that even minor inaccuracies in language input disrupt LTs' language comprehension to a greater extent than their TT peers.

These findings help to refine our understanding of what underlies LTs' delays in expressive vocabulary development. There is extensive research documenting delays in LTs' expressive phonological development, including reduced phonemic inventories, reduced diversity of syllable shapes, and lower accuracy in speech sound production. We had hypothesized these delays may have been driven by weaknesses in LTs' forming specific phonological representations. However, the current results do not support this idea. What can

explain these differences in LTs' expressive phonological development in light of the current findings?

As theorized in psycholinguistic models of speech processing (Stackhouse & Wells, 1997; Terband et al., 2019), many different processes might explain difficulties in producing a target word, even when a child has a well-formed internal phonological representation of that word. Difficulties in motor planning, motor programming, or motor execution could all lead to delays or deviations in spoken language. Recent studies have linked toddlers' speech-motor control to both concurrent and future expressive language development (Alcock & Krawczyk, 2010; Nip et al., 2011), highlighting links between motor and language development. It may be that differences in aspects of speech-motor control prevent LTs from using their well-formed internal phonological representations to produce speech.

As demonstrated in older children with speech delays (Geronikou & Rees, 2016), it may also be that individual children show delays or impairments in different aspects of speech processing. Within both TDs and LTs in the current sample, there was heterogeneity in children's sensitivity to mispronunciations, supporting the notion that children – TD and LT alike – may show differences in specific aspects of speech processing. If this is the case, then individual differences in sensitivity to mispronunciations may still predict later language development. Given that many LTs go on to develop language abilities in the average range, it may be that only the subset of LTs who will continue to show protracted language development will display reduced sensitivity to mispronunciations. We are currently collecting longitudinal follow-up data on this sample of children to address this possibility. In addition, recent work shows that preschoolers with speech sound disorders are less sensitive to mispronunciations than

children without speech sound disorders (Brosseau-Lapre & Schumaker, 2020). It may be that only children (TTs and LTs alike) who are susceptible to later speech sound disorders, and not to language delays more generally, show a reduced sensitivity to mispronunciations.

Children may also differ in their sensitivity to mispronunciations on speech sounds that they can and cannot produce accurately. Future investigations may use the LWL mispronunciation task to probe sensitivity to mispronunciations on speech sounds that are both in and out of toddlers' phonemic inventories. We are currently conducting analyses of expressive phonological development for children in the current sample, but future work may include designs that use individualized stimuli based on each child's phonemic inventory and error patterns.

Finally, results from the exploratory analyses support previous findings suggesting that lexical properties, such as phonotactic probability, impact children's sensitivity to mispronunciations (MacRoy-Higgins et al., 2014). The current analyses expand on these findings and suggest that lexical properties may have a differential impact on processing of mispronunciations in LTs and TTs. Future studies should investigate lexical properties that may impact LTs lexical processing of mispronunciations in familiar words.

Design Considerations

In addition to the explanations discussed above, there are several design choices that may have impacted our ability to detect a difference in TTs' and LTs' sensitivity to mispronunciations.

Mispronunciation Type

This study used 1-feature mispronunciations that occurred in word-initial position on monosyllabic words. Given that recent work has found that children with a history of being a LT have difficulty producing polysyllabic words (Neam et al., 2020), sensitivity to mispronunciations on longer words may reveal differences we did not see in the current study. In future work, we plan to explore LTs' sensitivity to mispronunciations of vowels and codas, as well as mispronunciations of polysyllabic words.

In the current study, we created mispronunciations by manipulating either place of articulation or voicing of target words. While White and Morgan (2008) found that 18-month-old children were sensitive to both of these types of mispronunciations, other studies of younger children (Mani & Plunkett, 2010b) using familiar-distractor paradigms showed that children did not show a mispronunciation effect in reaction to changes in voicing. In the current study, when data from only TTs is analyzed, TTs showed a smaller mispronunciation effect for changes in voicing compared to changes in place of articulation (interaction between phonological condition and mispronunciation type: Wald $Z = 2.115$, $p = 0.034$). When compared to White and Morgan (2008)'s finding in a similar sample of 18-month-old English-speaking North American children, this suggests that the TTs in our sample "accommodated" the voicing mispronunciations when presented with a familiar distractor. Future studies that utilize the familiar-distractor mispronunciation paradigm may want to avoid mispronunciations formed by changes in only voicing, as any mispronunciation effect may be masked by accommodation of the mispronunciation.

Novel vs. Familiar Distractor Images

In the current study, we chose to use familiar images for both the distractor and the target image. This is in contrast to some other studies, in which a novel distractor object is paired with a familiar target (White & Morgan, 2008). In the presence of a novel image, children may assume that a mispronunciation is in fact a novel word and refers to this unfamiliar object; however, in the presence of two familiar objects, children may “accommodate” the mispronunciation and look to the target image because it is the closest match available to the auditory stimulus (White & Morgan, 2008). Thus, the difference in looking time to the target between correct and mispronunciation conditions may be smaller in familiar-distractor paradigms compared to novel-distractor paradigms. In a study using a novel-distractor mispronunciation paradigm in 30-46-month-olds, Law and Edwards (2015) found that children with larger vocabularies were more sensitive to mispronunciations than were children with smaller vocabularies.

Limitations and Constraints to Generality

A limitation of the study is that the group of late talkers was significantly younger than the group of typical talkers. Though we controlled for age in all statistical models to account for this, one might nevertheless worry that differences between the groups could be explained by age rather than late talker status. We next discuss this limitation in the context of the findings. The main goal of this study was to determine whether late talkers and typical talkers differed in sensitivity to mispronunciations. We found that both groups were sensitive to mispronunciations and did not find evidence that late talkers differed from typical talkers in the magnitude of their sensitivity to mispronunciations. Given this result, it is unlikely that the age difference between groups influenced the main findings.

On the other hand, the finding that late talkers did not look at the target in the MP condition above chance levels may have resulted from their younger age. However, previous mispronunciation studies with typically developing children have found above-chance looking in the mispronunciation condition in children as young as 14 months (Swingley & Aslin, 2002). Late talkers in this study had a mean age of 24.57 months and the youngest was 22 months, suggesting that age is unlikely to explain their performance in the MP condition. This suggests that the pattern of results in the current study reflects differences based on children's late talking status rather than age. Finally, the main effect of talker status (indicating that late talkers looked less to the images in the correct pronunciation condition) may reflect an age difference between groups and so should be interpreted with caution. However, the results of our analyses do not hinge on this effect.

The children in this sample were monolingual English-learning children from the U.S. Children were primarily white (80%), and children's caregivers had completed high levels of education (only one caregiver had not completed college). These children in the current sample passed a screening test for ASD, scored within 1.5 SD of the mean on a test of nonverbal cognitive abilities, and were not reported to have been born prematurely. We expect these results to generalize to similar groups of monolingual U.S. toddlers learning English with similar characteristics. Children learning languages other than English, children learning more than one language, children at risk for ASD or general cognitive delay, and children who were born prematurely may perform differently on similar measures. We also used a fairly inclusive threshold for classification as a LT (age- and gender-matched scores at or below the 15th percentile, and/or no caregiver-reported two-word combinations). Studies using different

thresholds to categorize children as LTs or TTs may find different results. Future studies should strive to include more diverse participants with varying characteristics in order to increase the generalizability of findings.

One of the largest limitations of this study is a reduced sample size. The study was interrupted by the COVID-19 2020 pandemic, which reduced our recruitment and testing efforts. While interaction terms assessing the difference in sensitivity to mispronunciations between LTs and TTs failed to reach statistical significance, they were in the expected direction of LTs showing smaller mispronunciation effects. With a larger sample, these effects may have been statistically significant. The fact that TTs showed a smaller mispronunciation effect to changes in voicing also means that the overall mispronunciation effect for TTs was reduced, potentially further obscuring differences that did exist between LTs and TTs. Although the current study leaves open the possibility that there may be differences in the magnitude of the mispronunciation effect between TTs and LTs, it is quite clear from the present findings that LTs – as a group – are sensitive to mispronunciations.

Conclusion

In the current study, two-year-old late talkers and typical talkers showed sensitivity to mispronunciations of early acquired familiar words. This suggests that children in both groups possess lexical representations with sufficient phonological detail to allow detection of minor changes in pronunciation, at least in word-initial position. In contrast to previously documented delays in late talkers' expressive phonological development, these findings do not find evidence of comparable delays in receptive phonological development. On the other hand, findings suggest that late talkers may be more impacted by mispronunciations than are typical talkers,

such that relatively minor inaccuracies in language input may disrupt late talkers' ability to understand a speaker's intended message.

In the current study, both LTs and TTs showed sensitivity to mispronunciations. This suggests in both groups, stored phonological representations contain sufficient detail to detect minor changes in pronunciations, at least in word-initial position. This stands in contrast to previous findings demonstrating delayed expressive phonological development in LTs. LTs may be more impacted by mispronunciations than are TTs, such that even minor inaccuracies in language input disrupt LTs ability to understand intended messages.

Chapter 3: Sensitivity to Semantic Relationships in U.S. Monolingual English-Speaking

Typical Talkers and Late Talkers

Introduction

Late talkers (LTs) are a heterogeneous group of toddlers who experience delayed language development in the absence of known genetic, cognitive, or other developmental conditions. LTs experience varied language outcomes later in development, with some children “catching up” to their typically developing peers (a group referred to as “Late Bloomers”) while others continue to experience delays in various aspects of language processing (Dale et al., 2003; Fernald & Marchman, 2012; Rescorla, 2005, 2009, 2011).

Numerous studies have attempted to find child characteristics that clinicians might use to predict which LTs will catch up and which will not, in the hopes of targeting early intervention services to children most at risk for continued language difficulties. One meta-analysis revealed that initial expressive language, receptive language, and socioeconomic status predicted continued difficulties with expressive language, while gender, family history, and the presence of phrase speech did not (Fisher, 2017). However, these predictors explained only a small percentage of the variance in later outcomes (Fisher, 2017). While studies so far have largely used broad-based measures of children’s receptive and expressive language development, such as standardized assessments and caregiver-reported vocabulary checklists, there is the possibility that more fine-grained analysis of LTs’ language processing may reveal more nuanced differences that may better differentiate those LTs who will continue to show difficulties with language and those who will develop language skills in the typical range.

Sensitivity to Semantic Relationships Among TT Toddlers

There is an extensive literature suggesting that lexical-semantic concepts (words) in both children's and adults' lexicons are organized around taxonomic categories and thematic relations, such that items sharing taxonomic and/or thematic associations co-activate one another (Mirman et al., 2017). For instance, using a visual world paradigm, Mirman and Graziano (2012) found that adults look to distractor images that are either taxonomically or thematically related to the target image more than they do completely unrelated distractors. Similarly, Unger, Savic, et al. (2020) found that children looked more to an image that was either taxonomically related to a named word, or frequently co-occurred with that word, compared to an unrelated object. This sensitivity to semantic relationships is also evident in the toddler-period. For instance, using a priming task, Arias-Trejo and Plunkett (2009) found that 21-month-olds' lexical processing was disrupted by an unrelated prime word, but was not affected by a semantically-related prime, suggesting these children were sensitive to the semantic relationship between prime and target. Other work using a diverse set of methodologies has revealed that young toddlers are sensitive to taxonomic relations between lexical items (Chow et al., 2017; Delle Luche et al., 2014; Johnson et al., 2011; Mani et al., 2013; Ross, 1980; Styles & Plunkett, 2009; von Koss Torkildsen et al., 2007; Willits et al., 2013). Furthermore, some work has shown that individual differences in the sensitivity to these relationships is related to children's language abilities (Chow et al., 2017). These findings suggest that children's sensitivity to semantic relationships among lexical-semantic items in their developing lexicon may be an important indicator of underlying language development.

Semantic Deficits in Children with Language Disorders

The storage of a word's meaning, or its semantic representation, has been implicated as an area of weakness in children with language learning deficits. For instance, older children with Developmental Language Disorder (DLD) demonstrate difficulties with giving detailed semantic information about the words they know (McGregor et al., 2013), have difficulty producing semantic associates (Sheng & McGregor, 2010), and show differences in a number of other tasks assessing semantic knowledge (Alt et al., 2004; Botting & Adams, 2005; Sabisch et al., 2006; Sheng & McGregor, 2010). This work has demonstrated that many preschool and kindergarten children with language difficulties show differences in tasks that require knowledge of semantic relationships. It may be that LTs also show similar differences in early lexical-semantic organization.

While the defining characteristic of LTs has been a delay in expressive vocabulary *size*, recent work has shown that late talkers also differ in their vocabulary *structure*. For instance, Beckage and colleagues (Beckage et al., 2011) found that the vocabularies of late talkers contained fewer semantic connections than did those of typically-developing children, even after controlling for vocabulary size. Additionally, work from our lab has shown that early differences in the semantic structure of late-talkers' vocabularies predicted their language development one year later (Curtis et al., 2017). While these studies have relied on children's *expressive* vocabularies, other studies of typically developing infants and toddlers have explored children's knowledge of semantic relationships using *receptive* language measures. In studies using both eye tracking and semantic priming paradigms, children as young as 21 months of age show knowledge of various semantic relationships (Chow et al., 2017; Johnson et al., 2011; Styles &

Plunkett, 2009). While many expressive language tasks assessing lexical-semantic organization are not feasible for toddlers, these receptive language tasks may be an ideal way to assess lexical-semantic organization in LTs.

The Current Study

In order to assess children's sensitivity to semantic relationships between lexical items, we employed a target-absent Looking-While-Listening Paradigm. In target-absent paradigms, children see images on a screen and hear a target word that is not pictured, but that is semantically related to one of the images. Target-absent paradigms have been used previously to study semantic processing in both children (Chow et al., 2017) and adults (Huettig & Altmann, 2005; Huettig & McQueen, 2007). We hypothesized that LTs, as a group, would show reduced sensitivity to semantic relationships among familiar words compared to TTs. This study will improve our knowledge of the lexical-semantic structure of LTs' early lexicons without relying on children's expressive vocabulary.

Methods

Participants

Participants were recruited via flyers distributed in the community, online advertisements, and existing research registries. Children were eligible for the study if they were between 22 and 30 months of age, heard only English in the home, were born at 37 weeks gestation or later, and had no caregiver-reported concerns regarding hearing loss or the presence of autism spectrum disorder. Other eligibility criteria were assessed during laboratory visits and included obtaining a T-score of 35 or higher on the Visual Reception scale of the Mullen Scales

of Early Learning, and a passing score of less than 2 on the Screen Tool for Autism in Toddlers (STAT).

In total, 57 participants were eligible for the study; 46 participants were retained for the current analyses. Participants were excluded for the following reasons: six participants received a score of 2 or higher on the STAT; two participants did not meet eligibility criteria on the Mullen; one child did not have a completed MCDI; and two children did not complete the eye tracking experiment because of visit cancellation due to the COVID-19 pandemic.

Measures

Screening Tool for Autism in Toddlers and Young Children (STAT; Stone & Ousley, 2008). The STAT is a level 2 screening measure for children between the ages of 24-36 months, with studies showing its efficacy in children younger than 24 months (Stone et al., 2008). The STAT is a 20-minute structured assessment, in which the examiner completes a number of presses to assess aspects of the child's social communication, imitation skills, and play. A score of 2 or higher indicates a higher likelihood of having autism. Children who received a score of 2 or higher were referred to Illinois' Early Intervention System and were excluded from further analyses.

Mullen Scales of Early Learning – Visual Reception Scale (MSEL-VR; Mullen, 1995). The Mullen Visual Reception (VR) scale is a measure of children's nonverbal cognitive abilities. Children are asked to complete a number of activities, including sorting objects by shape, matching visual images, and nesting cups together. In the current study, children with a T-score below 35 (more than 1.5 standard deviations from the mean of age-matched peers) were excluded.

MacArthur-Bates Communicative Development Inventory – Words and Sentences

(MCDI; Fenson et al., 2006). The MCDI is a checklist of early-acquired vocabulary completed by the child's caregiver. In the current study, children were classified as LTs if their vocabulary score was at or below the 15th percentile (Beckage et al., 2011; Hodges et al., 2017; Lu et al., 2020) compared to age- and gender-matched peers, and/or if the child's caregiver reported that their child was not combining words. Children with vocabularies at or above the 20th percentile and who were reported to combine words were classified as TTs.

Target-absent Looking-While-Listening Task.

Design. Four groups of four words each (quartets) were chosen from words that are acquired early in development (Fenson et al., 2006). Each quartet was made up of words from two of the following taxonomic categories: clothing, animals, food, or vehicles. Each quartet contained two words from one category, and two words from the other. From each quartet, two yoked pairs were created, combining one word from each taxonomic category (e.g., pizza-shirt; hat-crackers).

Quartets were chosen by matching words on their frequency in child-directed speech, number of phonological neighbors, and age of acquisition. The semantic relationships between words in each quartet was determined by using the Buchanan feature norms (Buchanan et al., 2019). The root value was extracted for each pair of words in the quartet. The root value is formed by comparing lists of semantic features generated by participants when presented with individual lexical items. A cosine value is calculated from the feature vectors of each word pair. Words were considered semantically related when they came from the same taxonomic category and their cosine value was greater than 0.2 (average = 0.3649). Words were considered

semantically unrelated when they came from separate taxonomic categories and their cosine value was less than 0.1 (average = 0.013). A full table of stimuli and their root values is given in Table 3.1.

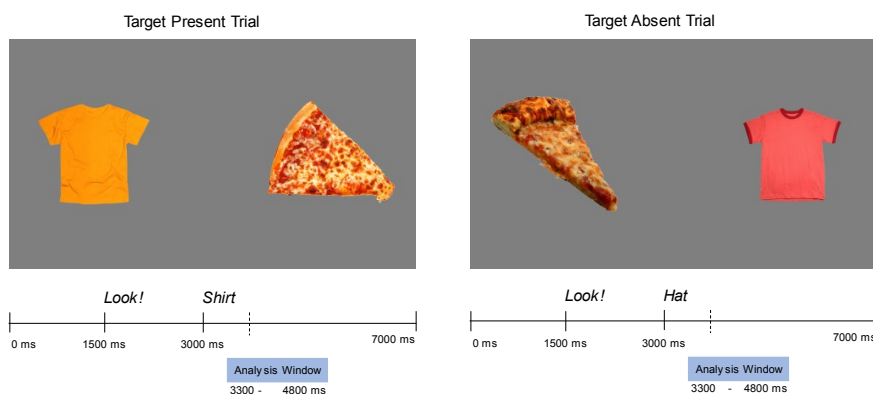
Table 3.1. *Semantic Similarity Ratings within Stimuli Quartets*

Quartet 1	Pants	Boots	Cereal	Toast
	Pants	0.323	0	0
	Boots		0	0
	Cereal			0.374
	Toast			
Quartet 2	Shirt	Hat	Crackers	Pizza
	Shirt	0.571	0	0
	Hat		0	0
	Crackers			0.2382
	Pizza			
Quartet 3	Airplane	Truck	Bread	Cheese
	Airplane	0.226	0	0
	Truck		0	0
	Bread			0.444
	Cheese			
Quartet 4	Monkey	Lion	Bus	Train
	Monkey	0.359	0.096	0.037
	Lion		0.061	0.017
	Bus			0.384
	Train			

Each word was presented twice: once in the target present condition, in which an image of that word appeared on the screen, and once in the target-absent condition, in which an image of the taxonomically related word from the same quartet appeared. Four pseudo-randomized counter-balanced orders were created. Targets appeared equally often on the left and right side of the screen. Each order began with a target present trial, and no more than two target-absent trials were presented consecutively, as children this age may become upset when presented with multiple target-absent trials consecutively (Chow et al., 2017).

Images were presented for a total of 7,000 ms. At 1500 ms, the speaker said “Look!” At 3000 ms, the speaker said the target word for that trial. The images remained on the screen for an additional 4,000 ms. See Figure 3.1 for an example of a target-present and target-absent trial.

Figure 3.1. *Example Target-Present and Target-Absent Trial Sequences*



Auditory Stimuli. Auditory stimuli were recorded by a female North American Midwest General American English speaker. Target words were on average 0.824 seconds in duration with a mean pitch of 346.54 Hz.

Visual Stimuli. In order to increase child engagement during the task and reduce repetition of the same images, two images were chosen for each word. Images were selected by the first author, and a group of adult raters were asked to label each picture and rate how well it matched their internal representation of that word. Adult raters correctly labeled every picture. Average ratings of the representativeness of each image were 97.597% (min 94.6, max 99.74). Images were matched to create pairs by comparing the distribution of luminosity of each pixel in that image. Images with similar distributions, as measured with Kruskal Wallis H Test, were combined into pairs. Each image pair was presented twice during the experiment.

Images were centered in 600 x 600 pixel areas of interest (AOI). Image size was manipulated so that the total number of pixels of each image was approximately equal. This was done to minimize the chance that broader stimulus items took up a larger proportion of the screen than narrower stimulus items.

Procedure

Children either sat in a highchair or on their caregiver's lap in a sound-attenuated area of the laboratory. Caregivers were instructed not to repeat any of the words used during the experiment and only to provide general verbal encouragement or redirection. Caregivers were given a pair of darkened sunglasses to wear to obscure their vision of the stimuli. The experiment was presented on a 24" ASUS VS248H-P monitor using Tobii Studio Pro. Audio was presented at 65 dB using Bose Companion 2 Series III Multimedia Speakers situated on the left and right sides of the monitor. A Tobii X3-120 eye tracker was used to collect children's gaze data at a sampling rate of 120 Hz. Children completed the Mullen VR and STAT while parents completed survey measures.

Eye Tracking Data Preparation

Gaps in eye tracking data less than 75 ms were interpolated using analysis scripts from Unger, Vales, et al. (2020). A 600 x 600 pixel area of interest (AOI) was defined around each target and distractor image. An additional 4 degrees of visual angle were added to all sides of the AOI, as previous studies have suggested that using larger AOI's may reduce the influence of noisy gaze data (Hessels et al., 2016).

Several children were hand-coded because the Tobii system was not able to track them. Children's eye tracking data were hand coded if Tobii calibration failed (4 children), if an

excessive number of trials were lost due to trackloss (one child), or if significant calibration drift was detected after inspecting visual plots (4 children). These data were coded using peyecoder (Olson et al., 2020).

All data processing and analyses were conducted in R (v 4.0.3) (R Core Team, 2014), using the eyetrackingR (Dink & Ferguson, 2015), lme4 (Bates et al., 2015), lmerTest (Kuznetsova et al., 2017), and emmeans (Lenth, 2020) packages. Gaze samples in which children looked outside of an AOI were removed from analyses.

Analysis Plan

Prior to data analysis, the analysis plan was registered on OSF (<https://osf.io/x3vsj/>). In all models, the child's race as reported by their caregiver, the caregiver's education (coded as graduate or professional degree vs. college degree or less), the child's biological sex, and the child's age (scaled and centered) were included as covariates. Consistent with previous analyses using the LWL paradigm (Fernald et al., 2008), we utilized an analysis window of 300 – 1800 ms after word onset.

Window Analyses

The first analysis we performed was a “window analysis.” In this type of analysis, the proportion of time spent looking at the target image throughout the entire analysis window is analyzed. We utilized a logistic mixed effects regression, with talker status and semantic condition (target-absent condition serving as the reference group), and their interaction as variables in the model, in addition to the above-mentioned covariates. The estimated marginal means for each talker group in each condition was also estimated using the emmeans package (Lenth, 2020).

Growth Curve Modeling

Because window analyses obscure fine-grained temporal dynamics of children's language processing, we also utilized growth curve modeling to capture any potentially fine-grained differences in children's processing (Barr, 2008; Mirman et al., 2008). We utilized logistic multilevel modeling, which analyzes a vector of successes (looks to the target) and failures (looks to the distractor) in each time bin (Donnelly & Verkuilen, 2017; Mirman et al., 2008). Our pre-registered analysis plan stated that linear and quadratic orthogonal polynomials would be used to model the data, but visual inspection of growth curves suggested more than one inflection point. For this reason, orthogonal polynomials for linear, quadratic, and cubic time, as well as their interactions with talker status and condition were entered into the models, as well as random intercepts and slopes for linear, quadratic, and cubic time nested in participants and participants by condition. Data were collapsed into 50 ms bins to reduce autocorrelation in the data. Sensitivity analyses were conducted with 100 ms and 200 ms windows, and results remained consistent across bin sizes.

Reaction Time Analyses

In addition to children's accuracy in the window and growth curve analyses, children's speed of lexical processing (or reaction time) was also analyzed. Consistent with previous studies using the LWL paradigm, the latency to shift from the distractor to the target image was analyzed on trials in which children were looking at the distractor image at word onset (distractor-initial trials) (Fernald & Marchman, 2012; Fernald et al., 2008). We additionally looked at children's latency to shift from the target image on target-initial trials. Differences in latency to shift on target-initial trials between the target-present and target-absent conditions may

index the specificity of children’s semantic representations of each target word. If children have specific definitions of each word, then they may shift off of the target image more quickly in the target-absent condition compared to the target-present condition.

Results

Visual inspection of the data revealed one outlier. This child, a late talker, looked to the target image 22.5% of the time in the target-absent condition (i.e., they looked reliably to the distractor). Given the small sample size of the current study, the outlier exerted an outsized influence on model coefficients. This participant was therefore excluded from all analyses presented here. Note that excluding this participant shifts results to be *less* consistent with our hypothesis (increasing the proportion of looks to the target among LTs in the target-absent trials), not more.

Table 3.2. *Participant Characteristics*

Variable	Typical Talker	Late Talker	Cohen’s d	p
	Mean (SD) or n			
<i>n</i>	24	21		
Child Age (months)	26.28 (2.29)	24.28 (1.98)	0.892	P = 0.004
MSEL VR ^a	52.38 (19.45)	50.48 (8.05)	0.125	P = 0.665
STAT ^b	0.406 (0.344)	0.631 (0.430)	-0.581	P = 0.063
Female	11	9		
Caregiver-report of child race				
White	19	17		
Black	1	0		
More than one race	4	4		
Caregiver education				
Some college	1	1		
College graduate	9	9		
Graduate/Professional training	14	11		

Note. Abbreviations: MSEL VR, Mullen Scales of Early Learning Visual Reception; STAT, Screening Tool for Autism in Toddlers

^aT-Score (Mean = 50; SD = 10)

^bScore range of 0 - 4

A total of 45 children were included in the following analyses (24 TTs, 21 LTs).

Participant characteristics are provided in Table 3.2. Children in the LT group were significantly younger than children in the TT group ($t(42.998) = 3.013, p = 0.004$). Because of this, age was entered as a covariate in all models. This limitation is discussed in the discussion section of this chapter, as well as in the general discussion in Chapter 4. Children did not differ on their Mullen-VR scores or STAT scores.

Out of 1,440 total trials from the LWL task, 1,083 were retained for analyses (75.21%). Trials were excluded for the following reasons: testing was aborted early or there was equipment failure (15); the child did not look at least one frame before and after word onset (97); the child made a noise during the critical word (88); there was an external noise during the critical word (28); the child was inattentive during the critical word (21); or there was greater than 50% trackloss during the critical window (108).

On average, children contributed 24.1 trials to analyses. This number did not differ significantly by talker status (LTs = 22.714 trials, TTs = 25.25 trials, $t(39.1) = 1.522, p = 0.136$). Gaze behavior prior to word onset was analyzed to determine whether children systematically preferred either distractor or target images. Proportion of looks to the target in the pre-naming phase did not differ from chance and did not differ significantly by talker status (TTs = 49.94%, LTs = 49.185%, $t(39.801) = 0.366, p = 0.717$). Accordingly, pre-gaze behavior was not included in analyses.

Window Analysis

A logistic multilevel model was used to analyze looking behavior across the 300 ms – 1800 ms window. The model included random effects for participant, participant by condition, item, and item by condition, as well as previously stated covariates.

Table 3.3. *Window Analysis, 300-1800 ms, LT and Target-Absent Conditions as Reference*

Predictors	Log-Odds	Std. Error	Statistic	<i>p</i>
(Intercept)	0.440	0.158	2.788	0.005
TT	0.007	0.137	0.053	0.957
Target-present condition	0.369	0.156	2.372	0.018
TT : Target-present condition	0.128	0.157	0.818	0.413

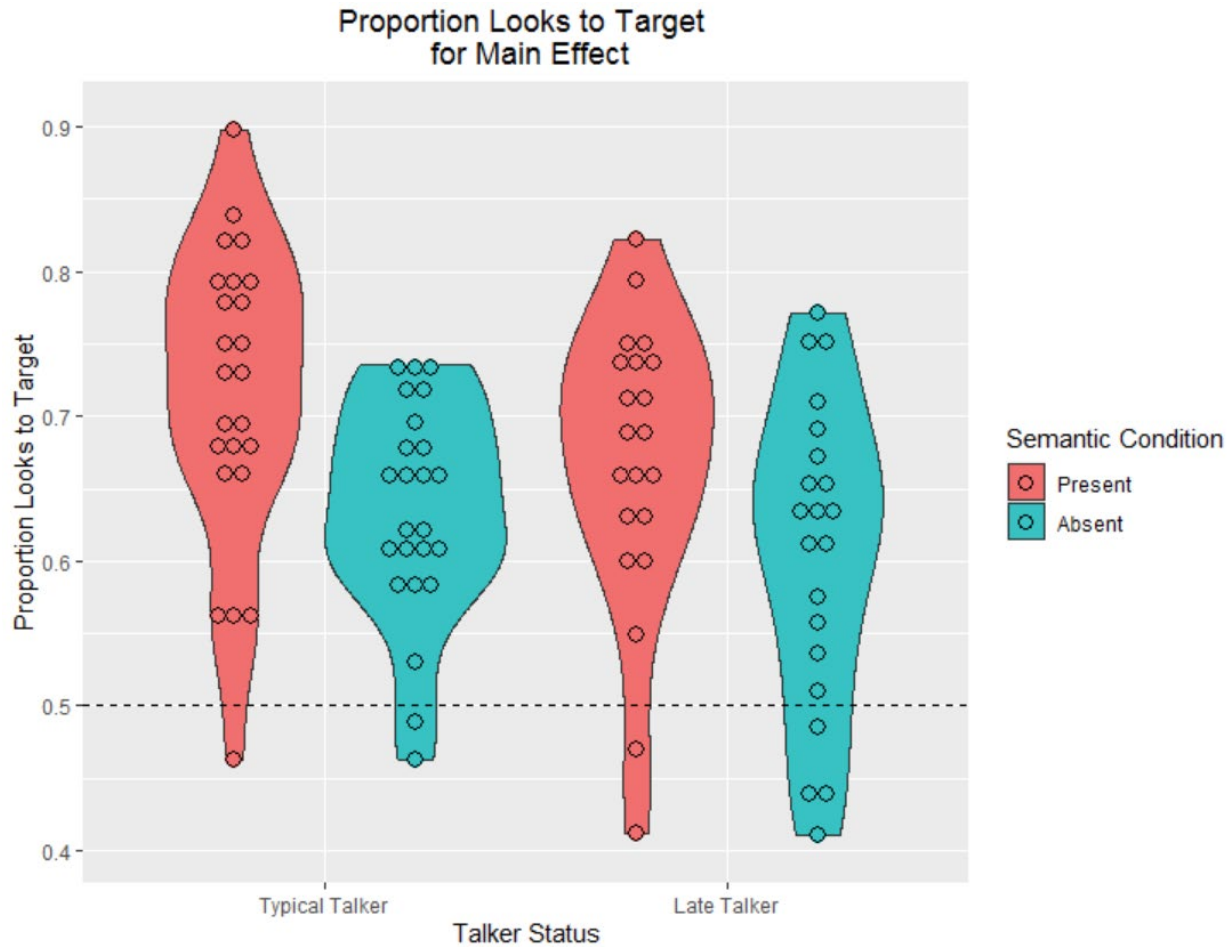
Abbreviations: TT, Typical Talker

Table 3.4. *Estimated Marginal Means in Log-Odds Units from Main Window Analysis Model, 300-1800 ms*

	EMM (SE)	95% Asympt. CI
Target-absent		
LT	0.433 (0.158)	[0.133, 0.753]
TT	0.451 (0.155)	[0.147, 0.754]
Target-present		
LT	0.813 (0.158)	[0.502, 1.123]
TT	0.948 (0.155)	[0.644, 1.252]

Note. Abbreviations: LT, Late Talker; TT, Typical Talker

Figure 3.2. *Proportion Looks to the Target in Main Window Analysis, 300-1800 ms*



The window analysis revealed that LTs looked significantly above chance at the target image in the target-absent condition (Wald $Z = 2.788$, $p = 0.005$).

Additionally, LTs looked significantly more at the target during the target-present condition than in the target-absent condition (Wald $Z = 2.372$, $p = 0.018$).

Reviewing the estimated marginal means in logit units from Table 3.4 reveals that TTs looked more to the target than the distractor in both target-absent and target-present conditions. There was no significant difference in looking to the

target between TTs and LTs in the target-absent condition (Wald $Z = 0.053$, $p = 0.957$) and no interaction between talker status and condition (Wald $Z = 0.818$, $p = 0.413$).

MCDI as a continuous variable

As various definitions of LT exist in the literature, we also used expressive vocabulary as a continuous predictor. A logistic multilevel model was used to analyze looking behavior across the 300 ms – 1800 ms window. Instead of talker status, children’s scaled and centered total expressive vocabulary scores from the MCDI were entered into the model. The model included random effects for participant, participant by condition, item, and item by condition, as well as previously stated covariates.

Table 3.5. *Window Analysis with Vocabulary as Continuous Predictor, 300-1800 ms*

Predictors	Log-Odds	Std. Error	Statistic	p
(Intercept)	0.445	0.140	3.181	0.001
MCDI Score (scaled)	0.086	0.079	1.097	0.272
Target-present condition	0.432	0.131	3.291	0.001
MCDI Score (scaled) * Target-present condition	0.073	0.080	0.908	0.364

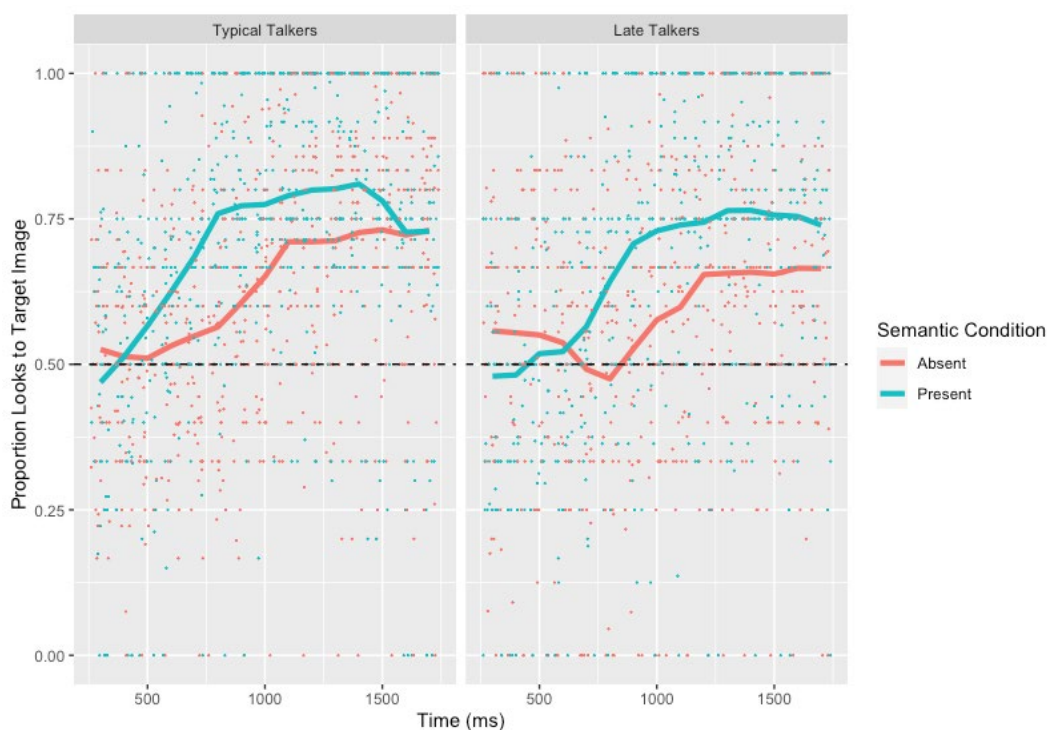
Note. Abbreviations: MCDI, MacArthur Bates Communicative Development Inventories

Results revealed that, after controlling for expressive vocabulary, children looked to the target in the target-absent condition significantly above chance. They also looked more in the target-present condition. There was no effect of expressive vocabulary on looks to the target in the target-absent condition (Wald $Z = 1.097$, $p = 0.272$) and no significant interaction between expressive vocabulary and condition (Wald $Z = 0.908$, $p = 0.364$).

Growth Curve Modeling

In order to capitalize on the temporal precision offered by eye tracking, a growth curve model was created to analyze data across the 300 – 1800 ms window. The model contained linear, quadratic, and cubic orthogonal polynomials for time, as well as their interaction with talker status and semantic condition.

Figure 3.3. *Growth Curve Model by Talker Type and Semantic Condition*



A model converged with random intercepts and slopes for linear and quadratic time nested within participants, and random intercepts and slopes for linear, quadratic, and cubic time nested within a participant by condition interaction. Because orthogonal polynomials were used in these models, the intercept represents overall looks to the target throughout the analysis window. As in the window analysis above, LTs looked to the target significantly above chance

in the target-absent condition (Wald $Z = 3.423$, $p = 0.001$). They also looked significantly more in the target-present condition than in the target-absent condition (Wald $Z = 2.165$, $p = 0.030$). There was no effect of talker status (Wald $Z = 0.252$, $p = 0.801$), and there was no significant interaction either between talker type and condition (Wald $Z = 1.103$, $p = 0.270$), talker type, condition, and linear time (Wald $Z = -0.961$, $p = 0.337$), talker type, condition, and quadratic time (Wald $Z = -0.763$, $p = 0.446$), talker type, condition, and cubic time (Wald $Z = 1.140$, $p = 0.254$).

Reaction Time Analyses

Children who did not provide at least 2 trials in each condition were excluded from analyses. This resulted in 3 participants being excluded from the distractor-initial analysis and 7 participants being excluded from the target-initial analysis. Reaction times were log-transformed prior to analyses.

Distractor-Initial Trials.

A multilevel model was run and p-values derived using the lmerTest package (Kuznetsova et al., 2017). Random effects of participant, participant by condition, and lexical item were included in the model, as well as previously stated covariates.

Table 3.6. *Reaction Time Analysis on Distractor-Initial Trials*

Predictors	Estimates	Std. Error	Statistic	p
(Intercept)	0.563	0.042	13.265	<0.001
TT	0.031	0.054	0.579	0.565
Target-present condition	0.042	0.042	1.011	0.318
TT : Target-present condition	-0.091	0.056	-1.633	0.111

Note. Abbreviations: TT, Typical Talker

Results revealed no differences between LTs and TTs in the target-absent condition (Wald $Z = 0.579$, $p = 0.565$), no significant interaction between talker type and condition (Wald $Z = -1.633$, $p = 0.111$) trials, nor an effect of semantic condition on LTs (Wald $Z = 1.011$, $p = 0.318$).

Target-Initial Trials.

A multilevel model was fit using lmerTest, and included random effects of participant, participant by condition, and lexical item, as well as previously stated covariates.

Table 3.7. *Reaction Time Analysis on Target-Initial Trials*

<i>Predictors</i>	<i>Estimates std. Error Statistic</i>			<i>p</i>
(Intercept)	0.653	0.059	11.069	<0.001
TT	-0.126	0.072	-1.759	0.086
Target-present condition	0.116	0.065	1.795	0.082
TT : Target-present condition	0.025	0.085	0.300	0.766

Abbreviations: TT, Typical Talker

Results revealed that TTs shifted away from the target image more quickly in the target-absent condition than did LTs, but this was only marginally significant (Wald $Z = -1.759$, $p = 0.086$). There was no interaction between talker status and condition (Wald $Z = 0.300$, $p = 0.766$).

Alternative Explanation and Exploratory Analysis: Syllable Structure Matching

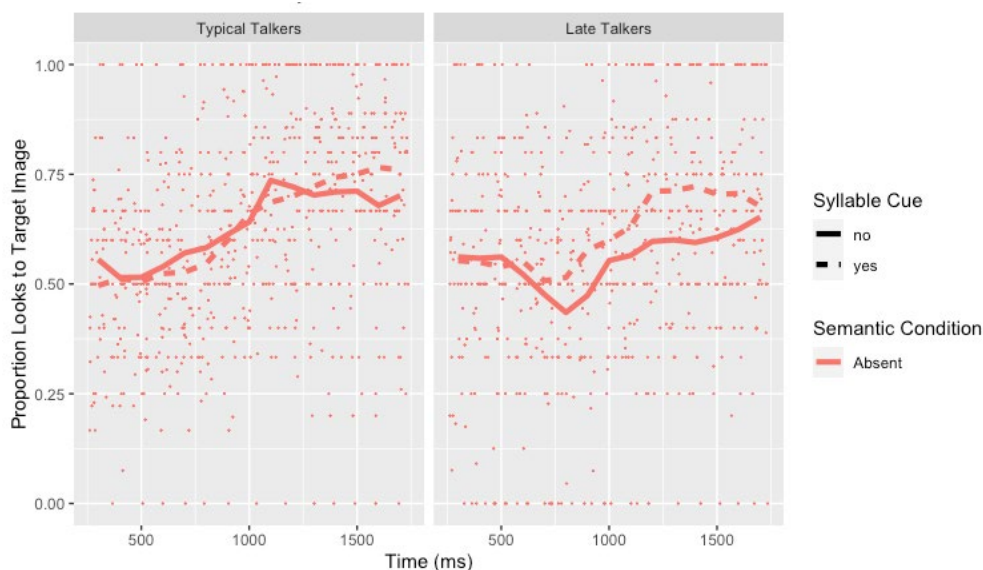
The findings above show that both TTs and LTs looked longer at an image that was taxonomically related to the target word than to an unrelated distractor. However, we wondered whether there were other differences between the semantically related and unrelated items that

could have influenced children's performance. Although the semantically related and unrelated items were matched on parameters known to influence lexical processing in young children (frequency in child-directed speech, age of acquisition, and number of phonological neighbors), we realized that for some word pairs, the semantically related item and the target word shared the same number of syllables (e.g., monkey and lion) while the unrelated item did not (e.g., bus). Thus, it is possible that on some target-absent trials, children may have looked to the target because of a match in the number of syllables (referred to as "syllable structure" hereon) between the auditory stimulus and the taxonomically-related image, and not necessarily due to their semantic relationship.

In order to explore whether syllable structure congruency impacted children's processing, we created a variable for each trial indicating whether syllable structure was an informative cue for identifying the target word (i.e., when the target word and target item matched in syllable structure while the target word and the distractor did not, as in monkey – lion [target] vs monkey – bus [distractor]).

This syllable structure congruency variable ("syllable cue") was added to a growth curve model. This analysis was done only for target-absent trials, as the main concern over syllable structure was whether it could fully explain the above results in the target-absent condition. The syllable cue variable was entered in 3-way interactions with time and talker status. As in previous models, random intercepts and slopes for linear, quadratic, and cubic time nested in both participant and participant by condition were included in the model. Models did not converge with the full random effects structure, and so random slope for cubic time nested in participant was removed from the model.

Figure 3.4. *Growth Curve of Target-Absent Trials by Talker Status and Syllable Cue Status*



Results revealed that LTs (as well as TTs) looked significantly above chance at the target even without syllable cues (Wald $Z = 2.727$, $p = 0.006$), suggesting that children looked at the target image on these target-absent trials on the basis of semantic relatedness alone. However, LTs looked significantly more to the target in target-absent trials in the presence of a syllable structure cue (Wald $Z = 5.325$, $p < .001$). While TTs did not look significantly more to the target than did LTs (Wald $Z = 0.857$, $p = 0.392$) in the target-absent condition, there was a significant interaction between talker group and syllable cue, indicating that TTs were not as impacted by syllable cues compared to LTs (Wald $Z = -3.590$, $p < 0.001$).

As in previous analyses, a separate model was run with scaled and centered continuous expressive vocabulary from the MCDI entered instead of talker status. In this model, there was a

significant effect of vocabulary on looks to the target without a syllable cue (Wald $Z = 2.282$, $p = 0.022$). Children with larger vocabularies looked to the target more during target-absent trials in the absence of a syllable cue. The effect of vocabulary size on looks to the target was weaker in the presence of a syllable cue (Wald $Z = -8.080$, $p < .001$). When using vocabulary as a continuous predictor, children with larger vocabularies looked more to the target in target-absent trials without a syllable cue, but this effect was weaker in trials in which syllable structure could be used as a cue.

Discussion

The current study utilized a target-absent Looking-While-Listening paradigm to assess typical and late talkers' sensitivity to semantic relationships among early-acquired lexical items. Contrary to our hypothesis, late talkers looked towards a visual referent (e.g., truck) after hearing a semantically related word (e.g., bus) to the same extent than typical talkers, suggesting similar levels of co-activation between taxonomically related items. Additionally, late talkers looked significantly more to the target in the target-present condition compared to the target-absent condition, suggesting that, as a group, late talkers in the current sample had specific definitions of these lexical items. This suggests that their gaze behavior in target-absent trials was not due to a vague or overly general representation of the words used in the study, but instead that they were sensitive to the taxonomic relationship between the items used. There were no statistically significant differences in sensitivity to semantic relationships between typical and late talkers, although there was some evidence that children with larger expressive vocabularies looked more to targets in the target-absent condition in trials in which syllable structure did not provide an additional cue.

These findings extend previous literature demonstrating that typically developing children are sensitive to semantic relationships among early-acquired words (Arias-Trejo & Plunkett, 2009; Chow et al., 2017). These findings also stand in contrast to previous literature that has found differences between late talkers' and typical talkers' lexical-semantic structures using network analysis (Beckage et al., 2011), as well as to studies of older children with DLD demonstrating difficulties in tasks assessing lexical-semantic organization. The discrepancy between the current findings and these previous investigations may be due to a number of factors, which are discussed below.

Type of Semantic Knowledge Assessed

There is emerging consensus that in adults, conceptual semantic knowledge for objects can be divided into two broad systems: taxonomic semantic relationships, and thematic semantic relationships (Mirman et al., 2017). Broadly, taxonomic relationships exist among objects with shared perceptual features (lions and monkeys both have eyes and tails) (McRae et al., 2005), while thematic relationships exist among objects that co-occur in similar contexts or are present within similar events (pizza and plates are often used together) (Estes et al., 2011). In adults, evidence from a number of independent investigations has demonstrated activation of distinct neural circuitry for taxonomic and thematic relationships, suggesting this is a theoretical distinction that influences linguistic and cognitive processing (see Mirman et al. (2017) for a comprehensive review of this literature). Developmental research has demonstrated that children are sensitive to both thematic and taxonomic relationships (Chow et al., 2017; Nguyen, 2007), and that early lexical networks show structural properties when analyzed using perceptual features (Peters & Borovsky, 2019) and co-occurrence metrics (Beckage et al., 2011).

While late talkers in the current sample showed sensitivity to the taxonomic relationships represented in our items, it may be that late talkers are less sensitive to thematic relationships. Beckage et al. (2011)'s study that found that the semantic networks of late talkers show less clustering than do the networks of typical talkers utilized co-occurrence relationships to construct the networks. These networks may have been indexing children's sensitivity to thematic relationships instead of taxonomic relationships. Chow et al. (2017) also found that children's language abilities were related to their sensitivity to thematic relationships, also lending support to the idea that late talkers may be less sensitive to thematic relationships. This possibility warrants further research.

Receptive vs. Expressive Tasks

Another possible explanation for the current findings is the fact that the current study used a receptive task, whereas studies of older children that found differences between typically developing children and children with DLD used production tasks (McGregor et al., 2002; McGregor et al., 2013). Utilizing the notion that taxonomic relationships exist between objects with shared semantic features, it may be that in the current task, in which late talkers are presented with images and lexical labels, shared features are activated to a sufficient degree to allow children to identify the target image. In production tasks, children are asked to access semantic features without visual cues in order to name taxonomic associates or give detailed word descriptions. It may be that late talkers and children with DLD are slower to access shared features in the absence of visual referents. There is evidence from semantic fluency tasks, which assess children's ability to name as many items from a specific category as possible (e.g., foods), that while children with DLD are slower than their age-matched peers and produce fewer

category exemplars, their response patterns do not show any evidence of differences in lexical-semantic *organization* (Mengisidou et al., 2020). Difficulties in language tasks assessing children's semantic knowledge may be unduly influenced by slower lexical access, and not differences in lexical-semantic organization per se (Mengisidou et al., 2020).

Heterogeneity Among Typical and Late Talkers

Previous studies have demonstrated that a majority of late talkers go on to develop language in the typical range (Dale & Hayiou-Thomas, 2013), although there is evidence that these children continue to perform below the level of same-age peers (Rescorla, 2005, 2009). Thus, it may be that while the majority of late talkers have developed a lexical-semantic organization similar to that of typical talkers, a subset of these late talkers have differently-organized lexico-semantic networks, and that it is these late talkers who will later persist in having language difficulties with diminished sensitivity to semantic relationships. Longitudinal follow-up of this cohort will help adjudicate between these possibilities.

Sensitivity to Syllable Structure Similarities

Upon further exploration of our stimuli, we discovered that on some trials, children could use a match in syllable count between the auditory stimuli and pictured items in the target-absent trials to identify the target, and thus could have used syllable matching rather than semantic relatedness to succeed in the task. Further analyses indicated that children looked significantly longer at the target image even on trials in which syllable count was not informative, confirming that their performance was driven by their sensitivity to the semantic relationship between the auditory and visual stimuli.

These additional analyses also revealed an incidental but interesting finding. Although typical talkers were unaffected by the presence or absence of syllable count congruency (in both trials with and without syllable cues, typical talkers looked to the target image), late talkers showed a clear and statistically significant effect of syllable structure congruency: on target-absent trials, late talkers looked more to the target in trials containing a syllable cue compared to non-congruent trials. This suggests that late talkers were sensitive to the taxonomic relationship between auditory and visual stimuli, but also influenced by the syllable count congruency of the stimuli.

These findings are consistent with previous work demonstrating that children are sensitive to the prosodic structure of lexical items (Jusczyk et al., 1999; Vihman et al., 2004). Late talkers seemed to utilize syllable structure as an additional cue in target-absent trials. This could imply that typical talkers activated taxonomically related concepts to a greater degree across both congruent and non-congruent trials, and so did not show a processing difference between these trials. Future research using this target-absent paradigm should carefully control syllable count among stimuli.

Individual Differences and Future Language Development

As can be seen in Figure 3.2, there was a large degree of variability in both typical talkers and late talkers in looking behavior in target-absent trials, with some children in both groups displaying greater or lesser activation of taxonomically related lexical-semantic concepts. While there was not strong evidence for differences between typical and late talkers' sensitivity to taxonomic relationships across all stimuli, it may be that a subset of children shows reduced activation of semantically related items. Individual differences in this sensitivity may be related

to either: a) children's vocabulary composition (Borovsky et al., 2016) and/or (b) later vocabulary development. We plan to utilize network science techniques (Beckage et al., 2011; Borovsky et al., 2016; Peters & Borovsky, 2019) to investigate the role of individual differences in the structure of children's early lexicons on online language processing in target-absent trials, and we are currently collecting follow-up data to investigate the impact of individual differences in sensitivity to taxonomic relationships on language outcomes in this sample.

Limitations and Constraints on Generality

It is important to note that the late talker group in this study was significantly younger than the typical talker group. The main aim of this study was to assess whether late talkers differed from typical talkers in their sensitivity to taxonomic relationships. Both groups looked at above-chance levels to the target image in the target-absent condition, and there was no effect of talker status. Therefore, the age difference between groups did not influence the main finding of this paper.

The current sample was composed of monolingual English-speaking children from the U.S. Toddlers in the sample were predominately white (80%) and caregivers had high levels of formal education (all but 2 caregivers had completed college). These children also passed a screener for ASD and a nonverbal cognitive skills screener. We expect the results from this study to generalize to other U.S. monolingual English-speaking children with similar characteristics. Children at risk for ASD, children born prematurely, or children with delays in non-verbal cognitive development may behave differently than the current sample. It is important to note that we used a relatively inclusive criteria for late talker status (scores at or below the 15th percentile using age- and gender-matched norms on the MCDI, and/or no two-word

combinations). Studies that utilize different criteria for late talker status may find different results. It may also be that toddlers exposed to more than one language may show different sensitivity to taxonomic relationships on target-absent trials. There is evidence that bilingual toddlers are sensitive to both within- and between-language semantic primes (Jardak & Byers-Heinlein, 2019; Singh, 2014), although in some contexts the priming effect may vary based on language dominance (Singh, 2014).

We utilized words that are acquired early in development based on normative data from the MCDI, which itself is biased towards white North American monolingual children. Utilizing words that are learned later in development may result in different behavior. The stimuli chosen represented a relatively restricted set of taxonomic categories (vehicles, clothing, food, and animals), and so work that investigates other categories is needed. In the literature, there is evidence that both culture and experience may influence the processing of semantic relationships (Imai et al., 2010; López-Pérez et al., 2016; Medin et al., 2006; Mirman et al., 2017). For this reason, sensitivity to taxonomic relationships between specific stimuli may vary based on children's culture and experiences across samples.

Conclusion

The current results suggests that U.S. English-speaking monolingual late talkers and typical talkers are both sensitive to taxonomic relationships among early-acquired lexical-semantic items. Variability existed within both groups, and the role of individual differences in sensitivity to these relationships will be explored in future research. Late talkers specifically benefited from similarities in syllable count as an additional cue in language processing. In

future research, it may be promising to examine the interplay in toddlers' use of prosodic and semantic cues during spoken word recognition.

Chapter 4: General Discussion

This dissertation investigated both phonological and semantic language processes in U.S. monolingual English-speaking typical and late talkers. Two experiments were conducted to investigate children's sensitivity to word-initial mispronunciations and children's sensitivity to taxonomic relationships between early-acquired words. The results from these two experiments revealed both similarities and differences in typical and late talkers' language processing. With regards to phonological processing, both LTs and TTs showed sensitivity to word-initial mispronunciations. However, these mispronunciations seemed to disrupt LTs' language processing to a greater degree than they did TTs' processing. When it comes to semantic processing, both LTs and TTs were sensitive to taxonomic relationships between early-acquired lexical-semantic items and looked more to target images than distractors in the target-absent condition. There was some evidence that children with larger vocabularies were more sensitive to these relationships when analyzing vocabulary as a continuous predictor. Taken together, these findings suggest that as a group, LTs' developing lexicons share some of the lexical-semantic and lexical-phonological organization that is seen in TTs, older children, and adults.

While the results from these two experiments did not find statistically significant group differences between TTs and LTs, this fact may not be surprising given the heterogeneity amongst LTs and their developmental trajectories (Rescorla, 2002, 2005, 2009; Rescorla et al., 2000). Additionally, the reduced sample size may have masked significant differences that do exist between groups. Because many LTs "catch up" with their typically developing peers (Dale & Hayiou-Thomas, 2013), there may not be group-level differences in lexical organization between TTs and LTs. Only a subset of LTs will go on to experience continued language

difficulties, and it may be that only this subset will show differences on the language-processing tasks used in these experiments.

Indeed, there may even be a subset of TTs who show differences in these tasks, further weakening group-level differences at baseline. While there is evidence demonstrating that LTs are at increased risk for DLD compared to their TT peers (Dale & Hayiou-Thomas, 2013; Rice et al., 2008), many children with DLD were not LTs (Rice et al. (2008) estimates the risk of later language difficulties for TTs as 11%). This may suggest that there are TTs in the current sample who show reduced specificity of phonological representations and/or reduced sensitivity to semantic relationships. This possibility is reflected in variability among both TTs and LTs across both tasks. In order to assess whether performance on these two assessments can predict future outcomes for TTs and LTs alike, we are currently collecting longitudinal caregiver-reported speech and language development, as well as video observations of children's language in naturalistic contexts.

The findings from these two studies raise a number of questions about the various processes that may be contributing to the expressive language delays observed in late talkers and the intersections between these processes. Below I discuss these larger issues, limitations from the current work, and outline directions for future research.

Auditory, Phonological, and Motor Speech Processes

The two experiments conducted in this project probed children's lexical-phonological and lexical-semantic organization. However, there are several other possible aspects of language processing that may differentiate TTs and LTs. Observations of dissociations in speech recognition and speech production abilities in adults with acquired language disorders have led

some theorists to posit separate auditory and motor lexical representations (Caramazza, 1991; Jacquemot et al., 2007). This has led several psycholinguistic models to include stored oral-motor representations of either phonemes or words (phonological output lexicon in the Ellis and Young model (Ellis & Young, 1993), stored motor program in the Stackhouse and Wells model (Stackhouse & Wells, 1997), or motor-phonological representation in the SLAM model (Walker & Hickok, 2016)). Using the framework from the SLAM model, the studies in this project investigated aspects of children's lexical-semantic and auditory-phonological representations. However, motor-phonological representations (or auditory-phonological to motor-phonological paths) may also be an area of deficit for LTs. This aspect of children's lexical representations is referenced in multiple models of speech sound disorders (McLeod & Baker, 2017), and offers the possibility that even in the presence of intact semantic representations and accurate auditory phonological representations, children may still have difficulty producing words they understand.

Indeed, there is some evidence that LTs have impairments in motor development as well as language development (Viholainen et al., 2002). There are also documented longitudinal relationships between speech-motor control (specifically lip and jaw movements) and vocabulary development (Nip et al., 2011); children with better speech-motor control have larger vocabularies, suggesting that the development of motor control is an important domain to consider in early communication development.

The deficits long documented in LTs' expressive phonological systems (Whitehurst et al., 1991; Williams & Elbert, 2003) may be due to either deficits in auditory-phonological representations or motor-phonological representations. Accurate auditory-phonological representation of phonemes must be present in order for children to learn to execute articulatory

gestures that reproduce those sounds (Hickok et al., 2011). The fact that LTs in the current sample were sensitive to word-initial mispronunciation seems to imply that, at least for initial phonemes, LTs' auditory-phonological representations are well-specified. This might suggest that for these children, it is only the motor-phonological representations that are impaired in some way. The fact that these children were not able to use the remaining phonological information in the word to infer the intended target (e.g., utilizing [ʌk] in “guk” to infer that the target was “duck” and not “bed”) could suggest that non-initial phonological information in the auditory-phonological representations are not well specified. Or, it could suggest that LTs' slower lexical processing impeded successful word recognition in this online task (Fernald & Marchman, 2012). Further studies targeting vowel and coda mispronunciations will be needed to elucidate these possibilities. In addition, to further disentangle the relationships between auditory-phonological and motor-phonological representations, analyses of phoneme-level sensitivity to mispronunciations will need to be compared to measures of these same phonemes in speech production.

Other aspects of psycholinguistic speech processing models include sensorimotor and auditory feedback monitoring (Hickok, 2012). There is some evidence that toddlers monitor their own speech and will alter their vocalizations in response to perturbations in auditory feedback (Scheerer et al., 2020). Utilizing this methodology with LTs may elucidate whether there are disturbances in error monitoring that may contribute to differences in LTs speech output.

Heterogeneity in causal mechanisms within LTs

While the two experiments here were analyzed separately, it may be that heterogeneity in LTs extends to the underlying processes that lead to their language delays. The evidence from

these studies suggests that, as a group, LTs co-activate taxonomically related lexical-semantic concepts, and that, as a group, LTs have specific phonological representations and are sensitive to word-initial mispronunciations. However, it may be that individual children have deficits in only one of these areas and not the other, and that comparing all LTs as a group against TTs obscures processing differences that do exist among subgroups of LTs. Further analyses that examine individual profiles of speech processing abilities may be especially useful in predicting later outcomes and developing treatment goals (Stackhouse & Wells, 1997). While motor-phonological output was not analyzed in the current experiments, both real word and non-word repetition tasks were administered to the current sample of children. Future analyses will simultaneously evaluate children's sensitivity to lexical-semantic relationships, sensitivity to mispronunciation, and speech motor output to assess the possibility of clusters of performance among LTs and TTs and their impact on later language development.

Additional Considerations

The current sample of children is composed of typical talkers and late talkers between the ages of 22 and 30 months. We utilized the age- and gender-matched normative data provided by the MCDI in month-by-month intervals to classify children as late talkers and typical talkers. These norms account for the non-linear changes in children's vocabulary development over time, and so all children, regardless of age, are classified based on data from same-age peers. As discussed in Chapters 2 and 3, the group of late talkers was significantly younger than the group of typical talkers. For the main findings in this study, namely that late talkers *were* sensitive to mispronunciations and *were* sensitive to taxonomic relationships between lexical-semantic items, this age difference does not unduly affect interpretation of these findings. Even the younger late

talker group were sensitive to the manipulations in each study. The finding that late talkers may not identify the target at above-chance levels in the mispronunciation condition in Chapter 2 may be influenced by difference in children's age. While our models did statistically adjust for the effects of age, it may be that a group of late talkers who were matched in age with the typical talkers would identify the target at above-chance levels in the mispronunciation condition.

There are additional limitations caused by the relatively wide age range used in the current study (although some studies use even broader age ranges, e.g. Beckage et al. (2011)). One of these limitations is that we do not have data on the older participants' language development at the age of the youngest participants in this study (i.e. the vocabulary size of a 27-month-old participant when they were 22 months of age). This means that some older typical talkers may have previously met criteria for classification as a late talker. Any underlying processing differences that differentiate late talkers and typical talkers may change across development, such that the impact of talker status on looking behavior in the two tasks used in this dissertation may not be constant across development. These are limitations imposed by the structure of the current data set. Future researchers may wish to restrict the age of participants to a narrower developmental window, or to recruit multiple age groups with sample sizes that support investigating developmental differences in the influence of talker status.

There are also broader methodological considerations that impact the current study, as well as other research conducted on late talkers.

Classifications Rely on Normative Samples

Studies of late talkers typically rely on caregiver-report measures of children's vocabulary to classify children as either typical or late talkers. Researchers typically use the

normative data provided by those instruments (often the MacArthur-Bates Communicative Development Inventories (MCDI; Fenson et al., 2006) and a selected threshold (10th, 15th, 20th, or 25th percentiles) to categorize participants. The validity of this classification is thus influenced by the extent to which the study sample matches the characteristics of the normative sample. The current study utilized the MCDI to classify children as late talkers and typical talkers. The study sample matches the MCDI normative sample well on children's race – since both the current sample and the MCDI normative sample are primarily white. The samples are also relatively well matched with regards to SES, although the caregivers in the current study completed more years of education than the average in the MCDI normative sample. This suggests that for the current study, the MCDI norms may provide accurate classification of children as late or typical talkers in relation to the normative sample. In other samples that do not match normative samples – for instance with predominately non-white samples – the MCDI norms may mis-classify children as late talkers or typical talkers.

Classifications Utilize a Single Timepoint

As previously discussed, most studies classify children as typical talkers or late talkers at a single timepoint. Often, longitudinal data are collected on children's later development. However, data on children's language development prior to classification as an LT or TT is often not available, as was the case in the current study. Because many LTs go on to develop language in the typical range (Dale & Hayiou-Thomas, 2013; Rescorla et al., 2000; Rescorla & Schwartz, 1990), even after only a few months, a child classified as a typical talker at age 24 or 27 months may have in fact met criteria for classification as a late talker just a few months prior. This fact points to the need to examine vocabulary size not just at a single timepoint, but instead at

multiple timepoints in order to characterize the (often nonlinear) trajectory of a child's vocabulary over time. This method was used by Rowe et al. (2012), who found that both the linear and quadratic coefficients describing children's vocabulary development trajectories helped to explain later vocabulary. It may be that underlying language processing differences may have stronger associations with children's vocabulary *trajectories* than with their vocabulary at a single time point. While we did collect longitudinal data on children in the current sample after initial classification as a typical or late talker, we did not collect data prior to children's first study visits. Future studies may wish to collect this data from early in development in order to classify children based on the properties of their vocabulary trajectories.

Future Directions

One future direction we plan to pursue is to utilize network analyses, or graph theoretical analytical methods, to model the semantic and/or phonological properties of children's emerging lexica. This sort of analysis has been conducted to analyze both the semantic (Beckage et al., 2011; Hills et al., 2009; Jimenez & Hills, 2017; Peters & Borovsky, 2019) and phonological structure of children's vocabularies. Multiplex, or multi-layer networks, have also been used to model multiple properties simultaneously (Stella et al., 2017). In all of these techniques, words are entered as nodes in the network. Semantic or phonological relationships between words are then used to create "links" or "ties" between nodes, and mathematical formulae are used to analyze various structural properties of the derived network. One limitation of this approach, however, is the very non-trivial task of choosing ties. In phonological networks, for instance, researchers have used Levenshtein distances of 1 (Stella et al., 2017), Levenshtein distance of 2 (Fourtassi et al., 2020), inverse of the Levenshtein distance (Beckage, 2016), degree of

phonological overlap (Beckage, 2016), and more complicated metrics based on the euclidean distance of shared phonological features (Laing, 2020). Similarly, a variety of measures of semantic relatedness have been used in semantic network analyses (Beckage et al., 2011; Beckage, 2016; Jimenez & Hills, 2017; Peters & Borovsky, 2019). Many of these studies also then choose an edge “threshold” value in order to create a binary network, where edges are either present or absent, introducing further researcher degrees of freedom. Before these measures can be used to compare groups or to predict later development, further work is needed to investigate the psychological validity of these edge metrics within both typical talkers and in late talkers. I plan to conduct further research using experimental data from eye tracking paradigms and real- and nonword-repetition tasks in order to investigate the psychological validity of these various edge metrics in future work.

Limitations

One important limitation of the current study is a reduced sample size. Due to the COVID-19 Pandemic, recruitment was cut short prior to achieving our planned sample size. This reduction in power may have resulted in failing to detect differences between groups. In addition to the reduction in power, the current sample largely self-identified as white and caregivers were highly educated. These characteristics constrain the generalizability to other groups, as discussed in the discussion section of each experiment. Additionally, the current sample consisted of only monolingual English-speaking children. This choice was made for several reasons. Current definitions of LT are typically based on monolingual vocabulary norms, and valid methods for classification of multi-language learners as LTs or TTs are still being developed (Pearson, 2013). The two eye tracking paradigms used here have not been used extensively with bilingual

children, and so their validity in bilingual populations is not yet known. Future research should recruit children that are more representative of the target research population and extend this work to dual language learners in order to enhance the generalizability of these findings.

Conclusion

Results from this dissertation revealed important aspects of language processing in U.S. monolingual English-speaking late talkers. Like typical talkers, late talkers were sensitive to mispronunciations on word-initial phonemes. This suggests that, unlike older children with DLD or SSD, late talkers have formed accurate auditory-phonological representations of words in their lexicons (at least on the initial phoneme). Likewise, late talkers were also sensitive to taxonomic relationships between early-acquired words. This suggests that, like typical talkers, late talkers' emerging lexicons contain the complicated semantic organizational structure found in older children and adults. Heterogeneity in both typical and late talkers' performance on these two experimental paradigms reinforces previous research demonstrating varied outcomes and language development trajectories among children this age. Longitudinal follow-up is needed to assess whether individual differences on these tasks may be useful predictors of later development. Additional analyses will incorporate assessments of children's motor-phonological representations to assess whether this area of language processing the language delays in the current sample. Additionally, profile analyses that simultaneously incorporate children's lexical-semantic, auditory-phonological, and motor-phonological representations will be conducted to assess the presence of distinct language processing profiles among late talkers.

REFERENCES

- Alcock, K. J., & Krawczyk, K. (2010). Individual differences in language development: relationship with motor skill at 21 months. *Dev Sci*, *13*(5), 677-691. <https://doi.org/10.1111/j.1467-7687.2009.00924.x>
- Alt, M., Plante, E., & Creusere, M. (2004). Semantic features in fast-mapping: performance of preschoolers with specific language impairment versus preschoolers with normal language. *Journal of Speech, Language, and Hearing Research*, *47*(2), 407-420. [https://doi.org/10.1044/1092-4388\(2004\)033](https://doi.org/10.1044/1092-4388(2004)033)
- Angulo-Chavira, A., & Arias-Trejo, N. (2018). Development of bidirectional phono-semantic activation in toddlers. *Quarterly Journal of Experimental Psychology (2006)*, *71*(9), 1968-1979. <https://doi.org/10.1177/1747021817737214>
- Arias-Trejo, N., & Plunkett, K. (2009). Lexical-semantic priming effects during infancy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1536), 3633-3647. <https://doi.org/10.1098/rstb.2009.0146>
- Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, *59*(4), 457-474. <https://doi.org/10.1016/j.jml.2007.09.002>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255-278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Barthelme, S. (2018). *imager: Image Processing Library Based on 'CImg'. R package version 0.41.1*. <https://CRAN.R-project.org/imager>
- Bates, D., Machler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1-48. <https://doi.org/10.18637/jss.v067.i01>
- Beckage, N., Smith, L., & Hills, T. (2011). Small worlds and semantic network growth in typical and late talkers. *PloS One*, *6*(5), e19348. <https://doi.org/10.1371/journal.pone.0019348>
- Beckage, N. M. (2016). *Predictive modeling to capture the words a toddler will learn next* [Dissertation, University of Colorado at Boulder]. Boulder, Colorado.
- Bobb, S. C., Huettig, F., & Mani, N. (2016). Predicting visual information during sentence processing: Toddlers activate an object's shape before it is mentioned. *Journal of Experimental Child Psychology*, *151*, 51-64. <https://doi.org/10.1016/j.jecp.2015.11.002>
- Borovsky, A., Ellis, E. M., Evans, J. L., & Elman, J. L. (2016). Lexical leverage: category knowledge boosts real-time novel word recognition in 2-year-olds. *Dev Sci*, *19*(6), 918-932. <https://doi.org/10.1111/desc.12343>
- Botting, N., & Adams, C. (2005). Semantic and inferencing abilities in children with communication disorders. *International Journal of Language and Communication Disorders*, *40*(1), 49-66. <https://doi.org/10.1080/13682820410001723390>
- Brosseau-Lapre, F., & Schumaker, J. (2020). Perception of Correctly and Incorrectly Produced Words in Children With and Without Phonological Speech Sound Disorders. *Journal of Speech Language and Hearing Research*, *63*(12), 3961-3973. https://doi.org/10.1044/2020_Jslhr-20-00119

- Buchanan, E. M., Valentine, K. D., & Maxwell, N. P. (2019). English semantic feature production norms: An extended database of 4436 concepts. *Behavior Research Methods*, *51*(4), 1849-1863. <https://doi.org/10.3758/s13428-019-01243-z>
- Caramazza, A. (1991). Some aspects of language processing revealed through the analysis of acquired aphasia: The lexical system. In *Issues in reading, writing and speaking: a neuropsychological perspective* (Vol. 3, pp. 15-44). Springer Science + Business Media. <https://doi.org/10.1007/978-94-011-3740-9>
- Carson, C. P., Klee, T., Carson, D. K., & Hime, L. K. (2003). Phonological profiles of 2-year-olds with delayed language development: Predicting clinical outcomes at age 3. *American Journal of Speech-Language Pathology*, *12*(1), 28-39. [https://doi.org/10.1044/1058-0360\(2003/050\)](https://doi.org/10.1044/1058-0360(2003/050))
- Charles-Luce, J., & Luce, P. A. (1995). An examination of similarity neighbourhoods in young children's receptive vocabularies. *Journal of Child Language*, *22*(3), 727-735. <https://doi.org/10.1017/s0305000900010023>
- Chow, J., Davies, A. A., & Plunkett, K. (2017). Spoken-word recognition in 2-year-olds: The tug of war between phonological and semantic activation. *Journal of Memory and Language*, *93*, 104-134. <https://doi.org/10.1016/j.jml.2016.08.004>
- Claessen, M., & Leitão, S. (2012). Phonological representations in children with SLI. *Child Language Teaching and Therapy*, *28*(2), 211-223. <https://doi.org/10.1177/0265659012436851>
- Collisson, B. A., Graham, S. A., Preston, J. L., Rose, M. S., McDonald, S., & Tough, S. (2016). Risk and protective factors for late talking: an epidemiologic investigation. *The Journal of pediatrics*, *172*, 168-174. e161. <https://doi.org/10.1016/j.jpeds.2016.02.020>
- Coonrod, E. E., & Stone, W. L. (2004). Early concerns of parents of children with autistic and nonautistic disorders. *Infants and Young Children*, *17*(3), 258-268. <https://doi.org/10.1097/00001163-200407000-00007>
- Curtis, P. R., Beckage, N. M., McWeeny, S., & Roberts, M. Y. (2017). *Predicting language growth in toddlers with language delay using semantic network analysis* [Conference session]. Cognitive Network Science Symposium, Indianapolis, IN.
- Cychoz, M., Cristia, A., Bergelson, E., Casillas, M., Baudet, G., Warlaumont, A. S., Scaff, C., Yankowitz, L., & Seidl, A. (2019). Canonical babble development in a large-scale crosslinguistic corpus. *PsyArXiv*.
- Dale, P. S., & Hayiou-Thomas, M. E. (2013). Outcomes for late talkers. In L. A. Rescorla & P. S. Dale (Eds.), *Late talkers: Language development, interventions, and outcomes* (pp. 241-257). Paul H. Brookes Publishing.
- Dale, P. S., Price, T. S., Bishop, D. V. M., & Plomin, R. (2003). Outcomes of early language delay: I. Predicting persistent and transient language difficulties at 3 and 4 years. *Journal of Speech Language and Hearing Research*, *46*(3), 544-560. [https://doi.org/10.1044/1092-4388\(2003/044\)](https://doi.org/10.1044/1092-4388(2003/044))
- Dalrymple, K. A., Manner, M. D., Harmelink, K. A., Teska, E. P., & Elison, J. T. (2018). An examination of recording accuracy and precision from eye tracking data from toddlerhood to adulthood. *Frontiers in Psychology*, *9*, 803. <https://doi.org/10.3389/fpsyg.2018.00803>

- Davis, B., Van der Feest, S., & Hoyoung, Y. (2018). Speech sound characteristics of early words: influence of phonological factors across vocabulary development. *Journal of Child Language*, 45(3), 673. <https://doi.org/10.1017/S0305000917000484>
- Delle Luche, C., Durrant, S., Floccia, C., & Plunkett, K. (2014). Implicit meaning in 18-month-old toddlers. *Developmental Science*, 17(6), 948-955. <https://doi.org/10.1111/desc.12164>
- Desmarais, C., Sylvestre, A., Meyer, F., Bairati, I., & Rouleau, N. (2008). Systematic review of the literature on characteristics of late-talking toddlers. *International Journal of Language and Communication Disorders*, 43(4), 361-389. <https://doi.org/10.1080/13682820701546854>
- Dink, J. W., & Ferguson, B. (2015). *eyetrackingR: An R Library for Eye-tracking Data Analysis*. In
- Donnelly, S., & Verkuilen, J. (2017). Empirical logit analysis is not logistic regression. *Journal of Memory and Language*, 94, 28-42. <https://doi.org/10.1016/j.jml.2016.10.005>
- Ellis, A. W., & Young, A. W. (1993). *Human cognitive neuropsychology: A textbook with readings*. Psychology Press.
- Estes, Z., Golonka, S., & Jones, L. L. (2011). Thematic thinking: The apprehension and consequences of thematic relations. In *Psychology of learning and motivation* (Vol. 54, pp. 249-294). Elsevier.
- Fagan, M. K. (2009). Mean length of utterance before words and grammar: Longitudinal trends and developmental implications of infant vocalizations. *Journal of Child Language*, 36(3), 495. <https://doi.org/10.1017/S0305000908009070>
- Fennell, C. T., & Werker, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech*, 46(2-3), 245-264. <https://doi.org/10.1177/00238309030460020901>
- Fenson, L., Bates, E., Dale, P., Goodman, J., Reznick, J. S., & Thal, D. (2000). Reply: Measuring variability in early child language: Don't shoot the messenger. *Child Development*, 71(2), 323-328. <https://doi.org/10.1111/1467-8624.00147>
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 59(5), i-185. <http://www.ncbi.nlm.nih.gov/pubmed/7845413>
- Fenson, L., Marchman, V. A., Thal, D., Dale, P. S., Reznick, J. S., & Bates, E. (2006). *MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual, 2nd Edition*. Brookes Publishing.
- Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development*, 83(1), 203-222. <https://doi.org/10.1111/j.1467-8624.2011.01692.x>
- Fernald, A., Swingle, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development*, 72(4), 1003-1015. <https://doi.org/10.1111/1467-8624.00331>
- Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. M. Fernandez, & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (Vol. 44, pp. 97-136). John Benjamins Publishing Company.

- Fischel, J. E., Whitehurst, G. J., Caulfield, M. B., & DeBaryshe, B. (1989). Language growth in children with expressive language delay. *Pediatrics*, *83*(2), 218-227.
<http://www.ncbi.nlm.nih.gov/pubmed/2913552>
- Fisher, E. L. (2017). A Systematic Review and Meta-Analysis of Predictors of Expressive-Language Outcomes Among Late Talkers. *Journal of Speech, Language, and Hearing Research*, *60*(10), 2935-2948. https://doi.org/10.1044/2017_Jslhr-L-16-0310
- Fourtassi, A., Bian, Y., & Frank, M. C. (2020). The growth of children's semantic and phonological networks: Insight from 10 languages. *Cognitive Science*, *44*(7), e12847.
<https://doi.org/10.1111/cogs.12847>
- Geronikou, E., & Rees, R. (2016). Psycholinguistic profiling reveals underlying impairments for Greek children with speech disorders. *Child Language Teaching and Therapy*, *32*(1), 95-110. <https://doi.org/10.1177/0265659015583915>
- Hammer, C. S., Morgan, P., Farkas, G., Hillemeier, M., Bitetti, D., & Maczuga, S. (2017). Late talkers: A population-based study of risk factors and school readiness consequences. *Journal of Speech, Language, and Hearing Research*, *60*(3), 607-626.
https://doi.org/10.1044/2016_JSLHR-L-15-0417
- Hawa, V. V., & Spanoudis, G. (2014). Toddlers with delayed expressive language: An overview of the characteristics, risk factors and language outcomes. *Research in Developmental Disabilities*, *35*(2), 400-407. <https://doi.org/10.1016/j.ridd.2013.10.027>
- Hessels, R. S., Kemner, C., van den Boomen, C., & Hooge, I. T. (2016). The area-of-interest problem in eyetracking research: A noise-robust solution for face and sparse stimuli. *Behavior Research Methods*, *48*(4), 1694-1712. <https://doi.org/10.3758/s13428-015-0676-y>
- Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature reviews neuroscience*, *13*(2), 135-145. <https://doi.org/10.1038/nrn3158>
- Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron*, *69*(3), 407-422.
<https://doi.org/10.1016/j.neuron.2011.01.019>
- Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009). Longitudinal analysis of early semantic networks preferential attachment or preferential acquisition? *Psychological Science*, *20*(6), 729-739. <https://doi.org/10.1111/j.1467-9280.2009.02365.x>
- Hodges, R., Baker, E., Munro, N., & McGregor, K. K. (2017). Responses made by late talkers and typically developing toddlers during speech assessments. *International Journal of Speech-Language Pathology*, *19*(6), 587-600.
<https://doi.org/10.1080/17549507.2016.1221452>
- Höhle, B., van de Vijver, R., & Weissenborn, J. (2009). Word processing at 19 months and its relation to language performance at 30 months: A retrospective analysis of data from German learning children. *Advances in Speech Language Pathology*, *8*(4), 356-363.
<https://doi.org/10.1080/14417040600970614>
- Horwitz, S. M., Irwin, J. R., Briggs-Gowan, M. J., Bosson Heenan, J. M., Mendoza, J., & Carter, A. S. (2003). Language delay in a community cohort of young children. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*(8), 932-940.
<https://doi.org/10.1097/01.CHI.0000046889.27264.5E>

- Huettig, F., & Altmann, G. T. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition*, *96*(1), B23-B32. <https://doi.org/10.1016/j.cognition.2004.10.003>
- Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language*, *57*(4), 460-482. <https://doi.org/10.1016/j.jml.2007.02.001>
- Imai, M., Saalbach, H., & Stern, E. (2010). Are Chinese and German children taxonomic, thematic, or shape biased? Influence of classifiers and cultural contexts. *Frontiers in Psychology*, *1*, 194. <https://doi.org/10.3389/fpsyg.2010.00194>
- Jacquemot, C., Dupoux, E., & Bachoud-Lévi, A.-C. (2007). Breaking the mirror: Asymmetrical disconnection between the phonological input and output codes. *Cognitive Neuropsychology*, *24*(1), 3-22. <https://doi.org/10.1080/02643290600683342>
- Jardak, A., & Byers-Heinlein, K. (2019). Labels or concepts? The development of semantic networks in bilingual two-year-olds. *Child Development*, *90*(2), e212-e229. <https://doi.org/10.1111/cdev.13050>
- Jimenez, E., & Hills, T. T. (2017). Network Analysis of a Large Sample of Typical and Late Talkers. CogSci, London, England.
- Johnson, E. K., McQueen, J. M., & Huettig, F. (2011). Toddlers' language-mediated visual search: They need not have the words for it. *Quarterly Journal of Experimental Psychology*, *64*(9), 1672-1682. <https://doi.org/10.1080/17470218.2011.594165>
- 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>
- Klee, T., Carson, D. K., Gavin, W. J., Hall, L., Kent, A., & Reece, S. (1998). Concurrent and predictive validity of an early language screening program. *Journal of Speech, Language, and Hearing Research*, *41*(3), 627-641. <https://doi.org/10.1044/jslhr.4103.627>
- Kohnert, K., Windsor, J., & Ebert, K. D. (2009). Primary or “specific” language impairment and children learning a second language. *Brain and Language*, *109*(2-3), 101-111. <https://doi.org/10.1016/j.bandl.2008.01.009>
- Krueger, B. I., & Storkel, H. L. (2020). Children's Response Bias and Identification of Misarticulated Words. *Journal of Speech, Language, and Hearing Research*, *63*(1), 259-273. https://doi.org/10.1044/2019_JSLHR-19-00140
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1-26. <https://doi.org/10.18637/jss.v082.i13>
- Laing, C. (2020, October 22-23). *Measuring phonological systematicity in infants' early words* [Conference session]. Many Paths to Language, Online.
- Law, F., & Edwards, J. R. (2015). Effects of vocabulary size on online lexical processing by preschoolers. *Language Learning and Development*, *11*(4), 331-355. <https://doi.org/10.1080/15475441.2014.961066>
- Lenth, R. V. (2020). *emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.5.3*. <https://CRAN.R-project.org/emmeans>
- Leonard, L. B., Weismer, S. E., Miller, C. A., Francis, D. J., Tomblin, J. B., & Kail, R. V. (2007). Speed of processing, working memory, and language impairment in children.

- Journal of Speech, Language, and Hearing Research*, 50(2), 408-428.
[https://doi.org/10.1044/1092-4388\(2007\)029](https://doi.org/10.1044/1092-4388(2007)029)
- Levelt, C. C., NSchiller, N. O., & Levelt, W. J. (1999). A developmental grammar for syllable structure in the production of child language. *Brain and Language*, 68(1-2), 291-299.
<https://doi.org/10.1006/brln.1999.2092>
- López-Pérez, B., Gummerum, M., Wilson, E., & Dellaria, G. (2016). Studying Children's Intrapersonal Emotion Regulation Strategies from the Process Model of Emotion Regulation. *The Journal of Genetic Psychology*, 1-16.
<https://doi.org/10.1080/00221325.2016.1230085>
- Lu, H.-H., Tsao, F.-M., & Tsai, J.-D. (2020). Behavioral problems of Mandarin-speaking late-talking toddlers and preschool aged children: A prospective case-control study in Taiwan. *Medicine*, 99(48). <https://doi.org/10.1097/MD.0000000000023341>
- MacRoy-Higgins, M., Shafer, V. L., Schwartz, R. G., & Marton, K. (2014). The influence of phonotactic probability on word recognition in toddlers. *Child Language Teaching and Therapy*, 30(1), 117-130. <https://doi.org/10.1177/0265659013487534>
- MacRoy-Higgins, M., Schwartz, R. G., Shafer, V. L., & Marton, K. (2013). Influence of phonotactic probability/neighbourhood density on lexical learning in late talkers. *International Journal of Language and Communication Disorders*, 48(2), 188-199.
<https://doi.org/10.1111/j.1460-6984.2012.00198.x>
- Mani, N., Coleman, J., & Plunkett, K. (2008). Phonological specificity of vowel contrasts at 18-months. *Language and Speech*, 51(1-2), 3-21.
<https://doi.org/10.1177/00238309080510010201>
- Mani, N., Durrant, S., & Floccia, C. (2012). Activation of phonological and semantic codes in toddlers. *Journal of Memory and Language*, 66(4), 612-622.
<https://doi.org/10.1016/j.jml.2012.03.003>
- Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—But only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance*, 38(4), 843.
- Mani, N., Johnson, E., McQueen, J. M., & Huettig, F. (2013). How yellow is your banana? Toddlers' language-mediated visual search in referent-present tasks. *Developmental Psychology*, 49(6), 1036. <https://doi.org/10.1037/a0029382>
- Mani, N., & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language*, 57(2), 252-272.
<https://doi.org/10.1016/j.jml.2007.03.005>
- Mani, N., & Plunkett, K. (2010a). In the infant's mind's ear: Evidence for implicit naming in 18-month-olds. *Psychological Science*, 21(7), 908-913.
<https://doi.org/10.1177/0956797610373371>
- Mani, N., & Plunkett, K. (2010b). Twelve-month-olds know their *cups* from their *keps* and *tups*. *Infancy*, 15(5), 445-470. <https://doi.org/10.1111/j.1532-7078.2009.00027.x>
- Mani, N., & Plunkett, K. (2011a). Does size matter? Subsegmental cues to vowel mispronunciation detection. *Journal of Child Language*, 38(3), 606.
<https://doi.org/10.1017/S0305000910000243>
- Mani, N., & Plunkett, K. (2011b). Phonological priming and cohort effects in toddlers. *Cognition*, 121(2), 196-206. <https://doi.org/10.1016/j.cognition.2011.06.013>

- Marchman, V. A., & Fernald, A. (2013). Variability in real-time spoken language processing in typically developing and late-talking toddlers. In L. A. Rescorla & P. S. Dale (Eds.), *Late Talkers: Language Development, Interventions, and Outcomes* (pp. 145-168). Brookes Publishing.
- McGregor, K. K., Newman, R. M., Reilly, R. M., & Capone, N. C. (2002). Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 45*(5), 998-1014. [https://doi.org/10.1044/1092-4388\(2002/081\)](https://doi.org/10.1044/1092-4388(2002/081))
- McGregor, K. K., Oleson, J., Bahnsen, A., & Duff, D. (2013). Children with developmental language impairment have vocabulary deficits characterized by limited breadth and depth. *International Journal of Language and Communication Disorders, 48*(3), 307-319. <https://doi.org/10.1111/1460-6984.12008>
- McLeod, S., & Baker, E. (2017). *Children's speech: An evidence-based approach to assessment and intervention*. Pearson Boston, MA.
- 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 non-canonical inputs. *Ear and Hearing, 40*(4), 961. <https://doi.org/10.1097/AUD.0000000000000681>
- McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2009). Within-category VOT affects recovery from “lexical” garden-paths: Evidence against phoneme-level inhibition. *Journal of Memory and Language, 60*(1), 65-91. <https://doi.org/10.1016/j.jml.2008.07.002>
- McNeill, B. C., & Hesketh, A. (2010). Developmental complexity of the stimuli included in mispronunciation detection tasks. *International Journal of Language and Communication Disorders, 45*(1), 72-82. <https://doi.org/10.3109/13682820902745479>
- McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods, 37*(4), 547-559. <https://doi.org/10.3758/bf03192726>
- Medin, D. L., Ross, N. O., Atran, S., Cox, D., Coley, J., Proffitt, J. B., & Blok, S. (2006). Folkbiology of freshwater fish. *Cognition, 99*(3), 237-273. <https://doi.org/10.1016/j.cognition.2003.12.005>
- Mengisidou, M., Marshall, C. R., & Stavrakaki, S. (2020). Semantic fluency difficulties in developmental dyslexia and developmental language disorder (DLD): poor semantic structure of the lexicon or slower retrieval processes? *International Journal of Language and Communication Disorders, 55*(2), 200-215. <https://doi.org/10.1111/1460-6984.12512>
- Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89-120). Lawrence Erlbaum Associates, Inc.
- Mirman, D. (2014). *Growth curve analysis and visualization using R*. CRC Press.
- Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language, 59*(4), 475-494. <https://doi.org/10.1016/j.jml.2007.11.006>
- Mirman, D., & Graziano, K. M. (2012). Individual differences in the strength of taxonomic versus thematic relations. *Journal of Experimental Psychology: General, 141*(4), 601. <https://doi.org/10.1037/a0026451>

- Mirman, D., Landrigan, J.-F., & Britt, A. E. (2017). Taxonomic and thematic semantic systems. *Psychological Bulletin*, 143(5), 499. <https://doi.org/10.1037/bul0000092>
- Mullen, E. M. (1995). *Mullen scales of early learning*. American Guidance Service.
- Nash, M., & Donaldson, M. L. (2005). Word learning in children with vocabulary deficits. *Journal of Speech, Language, and Hearing Research*, 48(2), 439-458. [https://doi.org/10.1044/1092-4388\(2005/030\)](https://doi.org/10.1044/1092-4388(2005/030))
- Neam, S. Y., Baker, E., Hodges, R., & Munro, N. (2020). Speech production abilities of 4-to 5-year-old children with and without a history of late talking: The tricky tyrannosaurus. *International Journal of Speech-Language Pathology*, 22(2), 184-195. <https://doi.org/10.1080/17549507.2019.1638968>
- Nguyen, S. P. (2007). Cross-classification and category representation in children's concepts. *Developmental Psychology*, 43(3), 719. <https://doi.org/10.1037/0012-1649.43.3.719>
- Nip, I. S., Green, J. R., & Marx, D. B. (2011). The co-emergence of cognition, language, and speech motor control in early development: A longitudinal correlation study. *Journal of Communication Disorders*, 44(2), 149-160. <https://doi.org/10.1016/j.jcomdis.2010.08.002>
- Oller, D. K., Eilers, R. E., Neal, A. R., & Schwartz, H. K. (1999). Precursors to speech in infancy: The prediction of speech and language disorders. *Journal of Communication Disorders*, 32(4), 223-245. [https://doi.org/10.1016/s0021-9924\(99\)00013-1](https://doi.org/10.1016/s0021-9924(99)00013-1)
- Olson, R. H., Pomper, R., Potter, C. E., Hay, J. F., Saffran, J. R., Ellis Weismer, S., & Lew-Williams, C. (2020). Peyecoder: An open-source program for coding eye movements. *Zenodo*. <https://doi.org/10.5281/zenodo.4313832>
- Paul, R. (1993). Patterns of development in late talkers: Preschool years. *Journal of Childhood Communication Disorders*, 15(1), 7-14. <https://doi.org/10.1177/152574019301500103>
- Paul, R., & Jennings, P. (1992). Phonological behavior in toddlers with slow expressive language development. *Journal of Speech, Language, and Hearing Research*, 35(1), 99-107. <https://doi.org/10.1044/jshr.3501.99>
- Pearson, B. (2013). Distinguishing the bilingual as a late talker from the late talker who is bilingual. In L. A. Rescorla & P. S. Dale (Eds.), *Late talkers: Language development, interventions, and outcomes* (pp. 67-87). Paul Brookes.
- Peters, R., & Borovsky, A. (2019). Modeling early lexico-semantic network development: Perceptual features matter most. *Journal of Experimental Psychology: General*, 148(4), 763. <https://doi.org/10.1037/xge0000596>
- R Core Team. (2014). *R: A language and environment for statistical computing*. In R Foundation for Statistical Computing.
- Reilly, S., Wake, M., Bavin, E. L., Prior, M., Williams, J., Bretherton, L., Eadie, P., Barrett, Y., & Ukoumunne, O. C. (2007). Predicting language at 2 years of age: a prospective community study. *Pediatrics*, 120(6), e1441-e1449. <https://doi.org/10.1542/peds.2007-0045>
- Rescorla, L. (1989). The Language Development Survey: A screening tool for delayed language in toddlers. *Journal of Speech and Hearing Disorders*, 54(4), 587-599. <https://doi.org/10.1044/jshd.5404.587>

- Rescorla, L. (2002). Language and reading outcomes to age 9 in late-talking toddlers. *Journal of Speech, Language, and Hearing Research*, 45(2), 360-371. [https://doi.org/10.1044/1092-4388\(2002/028\)](https://doi.org/10.1044/1092-4388(2002/028))
- Rescorla, L. (2005). Age 13 language and reading outcomes in late-talking toddlers. *Journal of Speech, Language, and Hearing Research*, 48(2), 459-472.
- Rescorla, L. (2009). Age 17 language and reading outcomes in late-talking toddlers: support for a dimensional perspective on language delay. *Journal of Speech, Language, and Hearing Research*, 52(1), 16-30. [https://doi.org/10.1044/1092-4388\(2008/07-0171\)](https://doi.org/10.1044/1092-4388(2008/07-0171))
- Rescorla, L. (2011). Late talkers: do good predictors of outcome exist? *Developmental Disabilities Research Reviews*, 17(2), 141-150. <https://doi.org/10.1002/ddrr.1108>
- Rescorla, L., Mirak, J., & Singh, L. (2000). Vocabulary growth in late talkers: Lexical development from 2; 0 to 3; 0. *Journal of Child Language*, 27(02), 293-311. <https://doi.org/10.1017/s030500090000413x>
- Rescorla, L., & Schwartz, E. (1990). Outcome of toddlers with specific expressive language delay. *Applied Psycholinguistics*, 11(4), 393-407. <https://doi.org/10.1017/S0142716400009644>
- Rice, M. L., Taylor, C. L., & Zubrick, S. R. (2008). Language outcomes of 7-year-old children with or without a history of late language emergence at 24 months. *Journal of Speech, Language, and Hearing Research*, 51(2), 394-407. [https://doi.org/10.1044/1092-4388\(2008/029\)](https://doi.org/10.1044/1092-4388(2008/029))
- Robb, M. P., & Bleile, K. M. (1994). Consonant inventories of young children from 8 to 25 months. *Clinical Linguistics & Phonetics*, 8(4), 295 - 320. <https://doi.org/10.3109/02699209408985314>
- Ross, G. S. (1980). Categorization in 1-to 2-yr-olds. *Developmental Psychology*, 16(5), 391. <https://doi.org/10.1037/0012-1649.16.5.391>
- 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.
- Sabisch, B., Hahne, A., Glass, E., von Suchodoletz, W., & Friederici, A. D. (2006). Lexical–semantic processes in children with specific language impairment. *Neuroreport*, 17(14), 1511-1514. <https://doi.org/10.1097/01.wnr.0000236850.61306.91>
- Scheerer, N. E., Jacobson, D. S., & Jones, J. A. (2020). Sensorimotor control of vocal production in early childhood. *Journal of Experimental Psychology: General*, 149(6), 1071. <https://doi.org/10.1037/xge0000706>
- Schopler, E., Van Bourgondien, M., & Wellman, G. (2010). *Childhood autism rating scale*. Western Psychological Services.
- Schwartz, R. G., & Leonard, L. B. (1982). Do children pick and choose? An examination of phonological selection and avoidance in early lexical acquisition. *Journal of Child Language*, 9(2), 319-336. <https://doi.org/10.1017/S0305000900004748>
- Sheng, L., & McGregor, K. K. (2010). Lexical–semantic organization in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 53(1), 146-159. [https://doi.org/10.1044/1092-4388\(2009/08-0160\)](https://doi.org/10.1044/1092-4388(2009/08-0160))
- Singh, L. (2014). One world, two languages: Cross-language semantic priming in bilingual toddlers. *Child Development*, 85(2), 755-766. <https://doi.org/10.1111/cdev.12133>

- Singleton, N. C. (2018). Late Talkers: Why the Wait-and-See Approach Is Outdated. *Pediatric Clinics*, 65(1), 13-29. <https://doi.org/10.1016/j.pcl.2017.08.018>
- Stackhouse, J., & Wells, B. (1997). *Children's Speech and Literacy Difficulties: A Psycholinguistic Framework*. Whurr Publishers.
- Stella, M., Beckage, N. M., & Brede, M. (2017). Multiplex lexical networks reveal patterns in early word acquisition in children. *Scientific Reports*, 7(1), 1-10. <https://doi.org/10.1038/srep46730>
- Stone, W. L., McMahon, C. R., & Henderson, L. M. (2008). Use of the Screening Tool for Autism in Two-Year-Olds (STAT) for children under 24 months: an exploratory study. *Autism*, 12(5), 557-573. <https://doi.org/10.1177/1362361308096403>
- Stone, W. L., & Ousley, O. Y. (2008). *Screening tool for autism in toddlers and young children (STAT)*. Vanderbilt University.
- Styles, S. J., & Plunkett, K. (2009). How do infants build a semantic system? *Language and Cognition*, 1(1), 1-24. <https://doi.org/10.1515/LANGCOG.2009.001>
- Swingle, D. (2009). Onsets and codas in 1.5-year-olds' word recognition. *Journal of Memory and Language*, 60(2), 252-269. <https://doi.org/10.1016/j.jml.2008.11.003>
- Swingle, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, 76(2), 147-166. [https://doi.org/10.1016/S0010-0277\(00\)00081-0](https://doi.org/10.1016/S0010-0277(00)00081-0)
- Swingle, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, 13(5), 480-484. <https://doi.org/10.1111/1467-9280.00485>
- Swingle, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition*, 71(2), 73-108. [https://doi.org/10.1016/S0010-0277\(99\)00021-9](https://doi.org/10.1016/S0010-0277(99)00021-9)
- Terband, H., Maassen, B., & Maas, E. (2019). A Psycholinguistic Framework for Diagnosis and Treatment Planning of Developmental Speech Disorders. *Folia Phoniatrica et Logopaedica*, 71(5-6), 216-227. <https://doi.org/10.1159/000499426>
- Unger, L., Savic, O., & Sloutsky, V. M. (2020). Statistical regularities shape semantic organization throughout development. *Cognition*, 198, 104190. <https://doi.org/10.1016/j.cognition.2020.104190>
- Unger, L., Vales, C., & Fisher, A. V. (2020). The Role of Co-Occurrence Statistics in Developing Semantic Knowledge. *Cognitive Science*, 44(9), e12894. <https://doi.org/10.1111/cogs.12894>
- Van Alphen, P., De Bree, E., Gerrits, E., De Jong, J., Wilsenach, C., & Wijnen, F. (2004). Early language development in children with a genetic risk of dyslexia. *Dyslexia*, 10(4), 265-288. <https://doi.org/10.1002/dys.272>
- Vihman, M. M., Nakai, S., DePaolis, R. A., & Hallé, P. (2004). The role of accentual pattern in early lexical representation. *Journal of Memory and Language*, 50(3), 336-353. <https://doi.org/10.1016/j.jml.2003.11.004>
- Viholainen, H., PhD, T. A., PhD, M. C., Lyytinen, P., & Lyytinen, H. (2002). Development of early motor skills and language in children at risk for familial dyslexia. *Developmental Medicine and Child Neurology*, 44(11), 761-769. <https://doi.org/10.1111/j.1469-8749.2002.tb00283.x>

- Von Holzen, K., & Bergmann, C. (2021). The Development of Infants' Responses to Mispronunciations: A Meta-Analysis. *Developmental Psychology*, 57(1), 1-18. <https://doi.org/10.1037/dev0001141>
- von Koss Torkildsen, J., Syversen, G., Simonsen, H. G., Moen, I., & Lindgren, M. (2007). Electrophysiological correlates of auditory semantic priming in 24-month-olds. *Journal of Neurolinguistics*, 20(4), 332-351. <https://doi.org/10.1016/j.jneuroling.2007.02.003>
- Walker, G. M., & Hickok, G. (2016). Bridging computational approaches to speech production: The semantic–lexical–auditory–motor model (SLAM). *Psychonomic bulletin & review*, 23(2), 339-352. <https://doi.org/10.3758/s13423-015-0903-7>
- Werker, J. F., Fennell, C. T., Corcoran, K. M., & Stager, C. L. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy*, 3(1), 1-30. https://doi.org/10.1207/S15327078in0301_1
- Wewalaarachchi, T. D., Wong, L. H., & Singh, L. (2017). Vowels, consonants, and lexical tones: Sensitivity to phonological variation in monolingual Mandarin and bilingual English–Mandarin toddlers. *Journal of Experimental Child Psychology*, 159, 16-33. <https://doi.org/10.1016/j.jecp.2017.01.009>
- White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, 59(1), 114-132. <https://doi.org/10.1016/j.jml.2008.03.001>
- Whitehurst, G. J., Smith, M., Fischel, J. E., Arnold, D. S., & Lonigan, C. J. (1991). The continuity of babble and speech in children with specific expressive language delay. *Journal of Speech, Language, and Hearing Research*, 34(5), 1121-1129.
- Williams, A. L., & Elbert, M. (2003). A prospective longitudinal study of phonological development in late talkers. *Language, Speech, and Hearing Services in Schools*, 34(2), 138-153. [https://doi.org/10.1044/0161-1461\(2003/012\)](https://doi.org/10.1044/0161-1461(2003/012))
- Willits, J. A., Wojcik, E. H., Seidenberg, M. S., & Saffran, J. R. (2013). Toddlers activate lexical semantic knowledge in the absence of visual referents: Evidence from auditory priming. *Infancy*, 18(6), 1053-1075. <https://doi.org/10.1111/infa.12026>
- Yimgang, D. P., Albury, R. A., & Leppert, M. L. (2017). Do Parental Concerns Predict Developmental and Behavioral Diagnoses in a Developmental Clinic? *Clinical Pediatrics*, 56(3), 263-267. <https://doi.org/10.1177/0009922816648289>
- Zubrick, S. R., Taylor, C. L., Rice, M. L., & Slegers, D. W. (2007). Late language emergence at 24 months: An epidemiological study of prevalence, predictors, and covariates. *Journal of Speech Language and Hearing Research*, 50(6), 1562-1592. [https://doi.org/10.1044/1092-4388\(2007/106\)](https://doi.org/10.1044/1092-4388(2007/106))