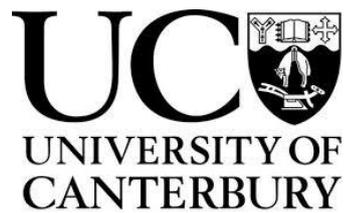


Sublexical Correlates of the Accuracy of Nonword Production in Preschool Children

A thesis submitted in partial fulfilment of the requirements for the Degree of

Master of Science in Speech and Language Sciences

Department of Communication Disorders



by Desma Mitchel D'Souza

July, 2015

Table of Contents

Preface	vi
Acknowledgements	vii
Abstract	viii
List of tables	i
List of figures	iii
List of Abbreviations	v
1.0 Literature review	1
1.1 Introduction	1
1.2 Two accounts of the nature of children's lexical representations	2
1.3 Phonotactic probability	4
1.3.1 Phonotactic probability and nonword repetition	7
1.3.2 Phonotactic probability and the onset of lexical Development	9
1.3.3 Phonotactic probability and short term memory	11
1.3.4 Phonotactic probability, vocabulary size and change over time	12
1.4 Variability in development of the spoken lexicon	15
1.5 Late Talkers, Language Impairment, and sensitivity to phonotactic probability	19
1.6 Rationale for the current study	20
1.7 Aims of the study	23
1.8 Research questions	23

2.0 Method	25
2.1 Ethical considerations	25
2.2 Participants	25
2.3 Materials/Tests	27
2.4 Procedure	28
2.5 Transcription	29
2.5.1 Transcription guidelines	30
2.6 Inter-transcriber agreement	32
2.7 Scoring	33
2.7.1 Word match	33
2.7.2 PCC/PVC (Standard)	34
2.7.3 APCC/APVC (Aligned)	34
2.8 Data analysis	35
2.8.1 Positional Segment Frequency	36
2.8.2 Biphone Frequency	36
3.0 Results	38
3.1 Descriptive statistics	38
3.2 Correlation analysis	41
3.3 Regression analysis	51
3.4 Partial correlations	52
3.5 Summary	61
4.0 Discussion	63
4.1 Nonword repetition accuracy	65
4.2 Sensitivity to sublexical correlates	66

4.2.1	Sensitivity to Positional Segment Frequency	67
4.2.2	Sensitivity to Biphone Frequency	68
4.3	Developmental trends in phonotactic probability- Continuity versus discontinuity	69
4.4	Performance variability in Late Talkers	71
4.4.1	Nonword repetition performance in Late Talkers	71
4.4.2	Late Talkers' sensitivity to phonotactic probability	72
4.5	Phonotactic probability and short term memory	73
4.6	Vocabulary, age and their association with phonotactic Probability	74
4.7	Reliability between the two forms of nonword repetition scoring using TENR-R	75
5.0	Study limitations	76
6.0	Future directions and conclusion	76
7.0	References	78
8.0	Appendices	93
	Appendix A: Test of Early Nonword Repetition- Revised (TENR-R)	93
	Appendix B: R script used to calculate phonotactic probability	94
	Appendix C: Phonotactic probability averages for the nonwords on the TENR-R	104

List of Tables

Table 1	<i>Summary of Languages Exposed to in Addition to English at T1 (N=22) and T3 (N=17)</i>	26
Table 2	<i>Additional Notes Added to Phon during Transcription</i>	32
Table 3	<i>Language Profile of Participants based on PLS4 (N=123) Raw Scores</i>	38
Table 4	<i>Descriptive Statistics for Typically Developing Children at T1 and T3</i>	39
Table 5	<i>Descriptive Statistics for Late Talkers at T1 and T3</i>	40
Table 6	<i>Regression Analysis Summary for Phonotactic Probability Variables Predicting Percentage of Nonwords Repeated Accurately in Typically Developing Children</i>	51
Table 7	<i>Regression Analysis Summary for Phonotactic Probability Variables Predicting Percentage of Nonwords Repeated Accurately in Late Talkers at T3</i>	52
Table 8	<i>Partial Correlation among Proportion of Nonwords Correctly Imitated and Phonotactic Probability Measures when one of the two measures is controlled across both groups at T3</i>	53
Table 9	<i>Descriptive Statistics for all Participants in the Study at T3</i>	54

	<i>Bivariate Correlations among Proportion of Nonwords Correctly</i>	
Table 10	<i>Imitated, Positional Segment Frequency Average and Biphone Frequency Average for the Entire Sample at T3</i>	55
	<i>Regression Analysis Summary for Biphone Frequency and Positional Segment Frequency Predicting the Proportion of Nonwords Imitated Correctly for all Participants in the Study at T3</i>	
Table 11		58
	<i>Partial Correlation among Proportion of Nonwords Correctly Imitated and Phonotactic Probability Measures When One of the Two Measures is Controlled for all Participants at T3</i>	
Table 12		59

List of Figures

<i>Figure 1</i>	Scatterplot of the Relationship between Proportion of Nonwords Correctly Imitated at T1 by Positional Segment Frequency Average for Typically Developing Children	43
<i>Figure 2</i>	Scatterplot of the Relationship between Proportion of Nonwords Correctly Imitated at T1 by Positional Segment Frequency Average for Late Talkers	44
<i>Figure 3</i>	Scatterplot of the Relationship between Proportion of Nonwords Correctly Imitated at T1 by Biphone Frequency Average for Typically Developing Children	45
<i>Figure 4</i>	Scatterplot of the Relationship between Proportion of Nonwords Correctly Imitated at T1 by Biphone Frequency Average for Late Talkers	46
<i>Figure 5</i>	Scatterplot of the Relationship between Proportion of Nonwords Correctly Imitated at T3 by Positional Segment Frequency Average for Typically Developing Children	47
<i>Figure 6</i>	Scatterplot of the Relationship between Proportion of Nonwords Correctly Imitated at T3 by Positional Segment Frequency Average for Late Talkers	48

	Scatterplot of the Relationship between Proportion of Nonwords	
<i>Figure 7</i>	Correctly Imitated at T3 by Biphone Frequency Average for Typically Developing Children	49
	Scatterplot of the Relationship between Proportion of Nonwords	
<i>Figure 8</i>	Correctly Imitated at T3 by Biphone Frequency Average for Late Talkers	50
	Scatterplot of the Relationship between Proportion of Nonwords	
<i>Figure 9</i>	Correctly Imitated at T3 by Positional Segment Frequency Average for all Participants	56
	Scatterplot of the Relationship between Proportion of Nonwords	
<i>Figure 10</i>	Correctly Imitated at T3 by Biphone Frequency Average for all Participants	57

List of Abbreviations

APCC	Aligned Percentage of Consonants Correct
APPC	Aligned Percentage of Phonemes Correct
APVC	Aligned Percentage of Vowels Correct
BF	Biphone Frequency average
CDI	Communicative Development Inventory
CLC	Child Language Centre
IBM	International Business Machines
LT	Late Talkers
LTT	Learning to talk
NZ CDI: WS	New Zealand adaptation of the MacArthur-Bates Communicative Development Inventory: Word and Sentences
PCC	Percentage of Consonants Correct
PLS4	The Preschool Language Scale, Fourth Edition
PPC	Percentage of Phonemes Correct
PSF	Positional Segment Frequency average
PVC	Percentage of Vowels Correct
T1	Time 1
T2	Time 2
T3	Time 3
TD	Typically Developing children
TENR-R	Test of Early Nonword Repetition - Revised

Preface

This dissertation is submitted for the degree of Master of Science, in Speech and Language Sciences, at the University of Canterbury. The research described herein was carried out between July 2014 and July 2015 in the Department of Communication Disorders at the University of Canterbury. The research was based at Child Language Centre, University of Canterbury and was conducted under the supervision and guidance of Professor Stephanie Stokes and Professor Thomas Klee. This work is to the best of my knowledge original, except as acknowledged in the text and has not been previously submitted either in part or in whole, for a degree at this or any other University.

Acknowledgements

My journey through Masters would not have been possible without the contributions from so many people in so many ways. My first words of gratitude are to my dear parents. I would never have made it this far without the love and support from you, Mum and Dad. Your vision for me is what has brought me here today. I will always be indebted to you. I am extremely grateful to my sisters Crystal & Dale, brother –in-law Magton and lil Cayle for always being there when I needed them. I feel blessed to have such an amazing family!

I would like to extend my sincere gratitude to my supervisors Prof. Stephanie Stokes and Prof. Thomas Klee. I couldn't have asked for better supervisors than the two of you. Thank you for being so patient with me and giving me the guidance and support that I needed throughout my Masters. Your knowledge and professionalism have been an inspiration to me and I feel fortunate that I had this opportunity to watch and learn from the best. Thank you!

A special thanks to my 'New Zealand family'- Sharimila, Daniela, Doreen, Asifa and Hamimah . Thank you for welcoming me into the Child Language Centre with such warmth and love. My experience at the CLC has been wonderful thanks to each one of you. I'll always cherish the moments spent with all of you. I am grateful for the words of support and guidance that I received from the wealth of knowledge and experience that you guys possess. You are my role models!

A big 'thank you' to Sajin Joseph for helping me find myself in New Zealand. Thank you for instilling the fear of my 'approaching deadline' and for your constant reminders to keep me on my toes. I owe much of the completion of my thesis to you!

I am grateful to all the children and their families who participated in my study. I am also grateful to all the clients I have worked with throughout my student and professional life. My experience working with them has taught me valuable lessons that cannot be learnt from books.

Last but definitely not the least I thank each and everyone who has been a part of this rollercoaster of a journey of mine. I may have missed a few names, but I will always be grateful for the help I received. Thank you!

Abstract

Aims: The main aims of the study were to examine the relative influence of phonotactic probability on nonword repetition accuracy using the Test of Early Nonword Repetition-Revised (TENR-R; Stokes & Klee, 2011) and to compare the sensitivity to these measures in Typically Developing (TD) and Late Talking (LT) preschoolers across a period of 18 months.

Method: The participants were enrolled in a long term project titled “Learning to talk: a research project on children’s early language development” (LTT; Klee, Stokes, & Moran, 2011-2015) and were recruited through early childhood specialists, by word of mouth from parents of the participants, and advertisement in nurseries. Data collection was carried out at three points in time: Time 1 (T1), Time 2 (T2) and Time 3 (T3). At T1 the participants were grouped into TD and LTs based on their language status assessed using the New Zealand adaption of the MacArthur-Bates Communicative Development Inventory: Words and Sentences (NZ CDI: WS; Reese & Read, 2000). A total of 134 participants between 24 and 30 months of age ($M= 27.38$, $SD= 1.64$) attempted the TENR-R at T1. At T3, a total of 138 participants between the age range of 42 and 49 months ($M= 45.36$, $SD= 1.97$) participated in the TENR-R. A computerised version of the TENR-R was administered at T1 and T3, and the participants were instructed to repeat the nonwords they heard.

Scoring and analysis: The audio and video recordings of the nonword repetitions on the TENR-R were transcribed using ‘Phon’ (Rose et al., 2006; Rose & MacWhinney, 2014) and scored for overall accuracy as well as percentage of consonant, vowel and phoneme accuracy. The phonotactic probabilities of the nonwords repeated accurately were calculated from the CELEX database and averaged for each child. Correlation and regression analyses were carried out to investigate the relationship between nonword repetition accuracy and phonotactic probability.

Results: The findings from correlation and regression analysis confirmed the presence of an association between phonotactic probability and nonword repetition accuracy. At T1, the TD participants were sensitive to Positional Segment Frequency ($r(91) = .18, p < .05$) while the LTs did not show sensitivity to any measure of phonotactic probability. At T3, both measures of phonotactic probability were found to correlate with nonword repetition accuracy in TD children (PSF average $r(105) = .34, p < .001$; BF average $r(105) = .51, p < .001$) and LTs (PSF average $r(33) = .63, p < .001$; BF average $r(33) = .39, p < .001$).

Conclusion: The results of the present study provide valuable insights into the nature of lexical representations in preschoolers. The presence of sensitivity to segmental information even at a very young age as well as in LTs at T3 along with the presence of the hypothesized sensitivity to Biphone Frequency at T3 is suggestive of the presence of an adult-like representational system that becomes fully adult-like with vocabulary growth and experience. There is a need for future studies to explore the nature of lexical representations from a continuity perspective and further expand the current findings.

Keywords: Biphone Frequency, nonword repetition, phonotactic probability, Positional Segment Frequency

1.0 Literature review

1.1 Introduction

Spoken language acquisition begins right from the time an infant hears its first words. From then on, the child begins to form associations from auditory input, by identifying the syntactic and phonological regularities occurring in the ambient language, such as phonological neighborhood density and phonotactic probability. Identification of the regularities is a sub-conscious process, and once the regularities are detected, rules and constraints underlying the structure of the native language emerge in the lexicon. These regularities are stored in the growing mental lexicon and are believed to underpin further language learning (Coady & Aslin, 2004; Storkel, 2001). The child has to first build up sufficient lexical and phonological representations in the mental lexicon in order to develop adult-like word recognition (Edwards, Beckman, & Munson, 2004; Jusczyk, Luce, & Charles-Luce, 1994). There are several statistical regularities that have been identified as lexical and sublexical features of words and among these, two in particular have emerged as influential in a child's language learning process: phonological neighborhood density (a lexical feature) and phonotactic probability (a sublexical feature). Phonological neighborhood density refers to the number of words that are phonologically similar in nature while differing by only one segment that is either substituted, added or deleted, for example *cat*, *fat*, *can*, *cot* (Luce & Pisoni, 1998). As the current research focuses on sublexical, rather than lexical characteristics of *nonword stimuli*, only phonotactic probability is considered further. Phonotactic probability refers to the frequency with which segmental and biphone regularities occur within words in a language (Jusczyk et al., 1994). A significant amount of research has been dedicated to understanding the role of these input regularities in language processing and language learning, and in particular the impact of input regularities on the development of children's lexical and phonological representations in the developing mental lexicon. Prior

research has generated two contrasting theories of the nature of early lexical representations.

1.2 Two accounts of the nature of children's lexical representations

One account, dubbed the emergent account, views infants' lexical representations to be holistic in nature, becoming more refined with development. This view developed from observations of differences in lexical processing between children and adults (Coady & Aslin, 2004). Some speech perception and production studies have found that young children have smaller vocabularies with very few similar sounding words (sparse phonological neighborhood density). Further, children's word production reflects sensitivities to syllables rather than individual segments. For example, Treiman and colleagues (as cited in Coady & Aslin, 2004) found that, when asked to group words, 4.4 year old children were observed to group the words based on similarities in gross word shapes rather than on similar segmental features. While some researchers associated vocabulary growth at 19 months of age with refinement of lexical representations (Ferguson, 1986, as cited in Coady & Aslin, 2004; Locke, 1988), others regarded the emergence of literacy skills, such as the ability to read, to be contributing to the emergence of adult-like representations from more holistic word forms (Fowler, 1991, as cited in Bowey, 2001; Walley, 1993). Findings from these and similar studies led to the conclusion that young children are sensitive to global lexical properties, such as syllables, rather than fine-grained properties, such as segmental properties, and that their representations are holistic and remain so until the early school years (Walley, 1993).

A second view of young children's lexical representations, dubbed the accessibility account, is that they resemble those of adults. In this view, young children possess adult-like fine-grained lexical and phonological representations and are capable of processing and expressing words in a manner similar to that of adults. It is children's constrained pre-linguistic and cognitive abilities that limit their performance on the various lexical decision,

word recognition and repetition tasks (Coady & Aslin, 2004). In a novel word learning task, Storkel (2001) taught children between the age range of 3.2 to 6.3 years unfamiliar monosyllabic words varying on two measures of phonotactic probability, Biphone Frequency and Positional Segment Frequency, across syllable positions. It was observed that novel word acquisition was facilitated by the influence of high phonotactic probability. In another study by Storkel and Rogers (2000) bisyllabic novel word acquisition was observed across three age groups (7 years, 10 years, & 13 years of age). The findings revealed that high phonotactic probability influenced novel bisyllabic word acquisition in older children between 10 and 13 years of age but not the younger children aged 7 years. This difference was attributed to limitations in either memory or expressive language. Evaluation of the expressive vocabularies of children ranging in age from 10 to 21 months revealed that approximately 84% of words in their lexicon contained at least one close phonological neighbor and hence the children would be required to use their knowledge of segmental features in order to distinguish among them (Dollaghan, 1994). In addition to this, Coady and Aslin (2003) found that given the smaller vocabulary size in children, the similar sounding words in their lexicon have a relatively greater neighborhood density than the adult lexicon. Based on their findings, Coady and Aslin (2003) concluded that children build up their vocabularies based on specific patterns or regularities common to words. Recent advances in research have shown that children's early representations, although not adult-like, are more fine-grained than once thought (Coady & Aslin, 2004). Children do show sensitivity to segmental aspects of language.

Both these accounts on the nature of lexical representations in young children describe a *discontinuity* in the development of the measures of phonotactic probability. The proponents of the emergent account (e.g., Ferguson, 1986, as cited in Coady & Aslin, 2004; Fowler, 1991, as cited in Bowey, 2001; Locke, 1988; Walley, 1993) believe that young

children store the words in their vocabulary holistically at first and these holistic lexical representations undergo a restructuring and become fine-grained with development. Conversely, proponents of the accessibility account (e.g., Coady & Aslin, 2003; Dollaghan, 1994) believe that the early lexical representations in children, although not adult-like, are fine-grained, with the level of storage and processing being larger than segments but more fine-grained than whole syllables (CVC). However, recent studies are now beginning to examine the nature of early lexical representations from a *continuity* standpoint, assuming a parallel course of development with both measures developing alongside each other (Coady & Aslin, 2004). However, there is still no consensus on how early lexical representations are stored, and how these representations change with development. Are children more sensitive to global/holistic word characteristics, or are they more sensitive to fine-grained characteristics such as phoneme frequency? Researchers have examined the relative effects of two types of phonotactic probability, biphone and segment frequencies to answer these questions.

1.3 Phonotactic probability

Every language is made up of patterns or regularities that are thought to play an important role in language processing (Gathercole, Frankish, Pickering, & Peaker, 1999; Storkel, 2001). When confronted with an unfamiliar word, both children and adults have been found to rely on sublexical correlates like phonotactic probability to decode and respond to the word. There are two measures of phonotactic probability commonly in use: Positional Segment Frequency and Biphone Frequency. Positional Segment Frequency refers to the frequency with which a phoneme occurs in a given position in a word. For example, in the nonword /bæk/, the Positional Segment Frequency for the phoneme /k/ would be the frequency with which the phoneme /k/ occurs in word-final position when compared with all

words in English that have any segment in the word-final position. Biphone frequency refers to the frequency with which a pair of phonemes exists in words in a language (Jusczyk et al., 1994). For example, in the nonword /dɛ:fi/, the Biphone Frequency for the phoneme sequence /dɛ:/ would be the frequency with which the phoneme sequence /dɛ:/ occurs in word-initial position when compared with all words in English that have any biphone in the word-initial position.

The sublexical locus of phonotactic probability was first hypothesized by Vitevitch and Luce (1998). The study demonstrated the contradictory effects of neighborhood density and phonotactic probability on the speed of verbal recall. The findings from performance on a repetition task showed that nonwords with high phonotactic probability and dense neighbours were at an advantage in naming as compared to nonwords with low phonotactic probability and sparse neighbours. Real words with high phonotactic probability and high density on the other hand, were responded to slower than real words with low phonotactic probability and low density. This led them to conclude that neighborhood density possesses a dominant lexical locus while phonotactic probability possesses a dominant sublexical locus (Vitevitch & Luce, 1998). Investigators have attempted to understand the organisation of the mental lexicon in children by observing the influence of neighborhood density which requires a considerable vocabulary size in order to be studied. Phonotactic probability has paved the way for analysing the structure and organisation of the mental lexicon even with a limited vocabulary size thereby providing an opportunity to study the nature of lexical representations and their associated sensitivity to sublexical characteristics in very young children (Coady & Aslin, 2004).

There is a considerable body of research that has attempted to investigate the role of phonotactic probability on language processing and production in both children and adults. These studies have focused on different aspects of language and linguistic processing such as

spoken word recognition (Frisch, Large, & Pisoni, 2000; Vitevitch & Luce, 1999), speech production (Goldrick & Larson, 2008), retention of suprasegmental patterns in nonwords (Tanida, Ueno, Lambon Ralph, & Saito, 2015), recognition of words in a foreign accent (Chan, 2012), electrophysiological studies focusing on effects from lexical and sublexical interactions and locus of activation during sublexical processing (Hunter, 2013; Rossi et al., 2011), influence of sublexical correlates in acquired speech and language disorders such as dysarthria and aphasia (Miller & Lallini, 2011; Reilly & Spencer, 2013), brand preference among consumers (Vitevitch & Donoso, 2012) and interaction between sublexical correlates and auditory-perceptual characteristics of words (Newman & Janse, 2013) in adults. The outcome of studies with adult participants provides evidence of the role of phonotactic probability in language acquisition and word learning.

In children, studies have focused on the influence of phonotactic probability on phonological memory and nonword repetition in monolinguals and bilinguals (Coady & Aslin, 2004; Freedman & Barlow, 2012; 2011; Gathercole et al., 1999; Lee & Gorman, 2013; Messer, Leseman, Boom, & Mayo, 2010; Munson, Kurtz, & Windsor, 2005; Storkel & Rogers, 2000; Zamuner, 2009), segmentation of speech (Adriaans & Kager, 2010; Blanchard, Heinz, & Golinkoff, 2010; Jusczyk et al., 1994; Mattys & Jusczyk, 2001), lexical acquisition (Hoover, Storkel, & Hogan, 2010; Storkel, 2001; Storkel, Bontempo, Aschenbrenner, Maekawa, & Lee, 2013; Storkel & Hoover, 2011; Storkel & Rogers, 2000), acquisition of later developing phonemes (Zamuner, Gerken, & Hammond, 2004), interaction with prosodic features during novel word acquisition (Estes & Bowen, 2013), lexical and sublexical factors influencing production variability of known words (Sosa & Stoel-Gammon, 2012), and word recognition (MacRoy-Higgins, Shafer, Schwartz, & Marton, 2014). The studies have also been extended into the clinical population involving children with Specific Language Impairment (SLI; Burke & Coady, 2015; Coady, Evans, & Kluender, 2010a; Coady, Evans,

& Kluender, 2010b; Gray, Pittman, & Weinhold, 2014; McKean, Letts, & Howard, 2013; Munson et al., 2005; Plante, Bahl, Vance, & Gerken, 2011), Phonological Delay (Munson et al., 2005; Storkel & Hoover, 2010; Storkel, Maekawa, & Hoover, 2010) and Late Talkers (LT; MacRoy-Higgins, Schwartz, Shafer, & Marton, 2013; Stokes, 2010). A novel word that consists of phoneme segments and sequences commonly found in a listener's vocabulary are easier to learn and retrieve. Nonwords with high phonotactic probability are more wordlike and have segments and sequences that are common, unlike words of infrequently occurring segments or biphone combinations (Munson, 2001). Hence they are easier to repeat and learn. These conclusions are supported by the findings that preschoolers between 2.2 to 2.8 years of age were able to perceive segmental contrasts when presented in a high phonotactic probability environment rather than in a low phonotactic probability environment (Zamuner, 2013).

Before considering how phonotactic probability affects early word and nonword repetition, the nature of the nonword repetition task is briefly described.

1.3.1 Phonotactic probability and nonword repetition. As mentioned above, research has explored the influence of phonotactic probability on children's ability to repeat nonwords (Storkel, 2001). Nonword repetition is a simple yet valuable task involving the repetition of a string of phonemes in nonwords that vary in syllable length. The most prevalent tools designed to understand language processing, such as the norm-referenced tests developed, rely heavily on the vocabulary knowledge and language experience of a user. The use of these knowledge dependent measures poses a problem in understanding the mechanism of language processing due to the variability in vocabulary and experience among language users. Poor performance on such tests could be attributed to lack of vocabulary knowledge or a language processing difficulty or both (Dollaghan & Campbell, 1998). To

overcome this, researchers began to develop processing dependent measures that are less biased and control variables such as vocabulary, age of acquisition and experience. Campbell, Dollaghan, Needleman, and Janosky (1997) evaluated the language performance of Typically Developing (TD) children from different cultural backgrounds using processing dependent measures that included nonword repetition and knowledge dependent measures. The children from the minority background scored significantly lower on the knowledge dependent measures whereas no difference in performance was observed between the two cultural groups on any of the processing dependent measures including the nonword repetition test. This led them to conclude that processing dependent measures such as nonword repetition tests are nonbiased culturally as compared to the knowledge dependent measures. A great deal of research over the years has led researchers to believe that a nonword repetition task is indeed a good tool for identifying phonological processing difficulties in very young children with language delay regardless of ethnicity and vocabulary knowledge (Campbell et al., 1997).

Despite the advantages of using nonword stimuli, there is a debate as to what nonword repetition actually measures. Researchers argue that nonword repetition does not assess phonological processing alone. Phonological memory also plays a role in the outcome of a nonword repetition task and this phonological memory component relies on the experience and long term knowledge of the language user (Bowey, 1997; Dollaghan, Biber, & Campbell, 1995; Snowling, Chiat, & Hulme, 1991). The consistent higher repetition accuracy of nonwords that are rated to be more wordlike suggests that the more language experience and vocabulary the participant has, the higher the accuracy would be (Gathercole, Willis, Emslie, & Baddeley, 1992). One must therefore be cautious while interpreting results based on these tests by controlling for the effects from inter-related factors such as memory, vocabulary, and articulation skills. Nevertheless, using nonwords rather than real words has

been found to be an effective research strategy in understanding the organization of language. The use of nonwords has proven to be advantageous as the phonotactic influence is believed to override the lexical influence for nonwords while the opposite is true for words (Storkel & Rogers, 2000; Vitevitch & Luce, 1999). The utility of nonwords also extends to the control of variables such as age of acquisition and word familiarity (Beckman, Munson, & Edwards, 2011; Stokes & Klee, 2009a). Studies using nonwords with high phonotactic probability have been found to have better participant reaction times to speeded same-different tasks (Vitevitch & Luce, 1999), and a higher identification accuracy (Zamuner, 2009). They are reported to be easier to learn with less exposure than nonwords with low phonotactic probability. The impact of phonotactic probability on learning new words is important from a developmental perspective because Gathercole et al. (1992) suggested that performance on nonword repetition can help predict lexical development.

1.3.2 Phonotactic probability and the onset of lexical development. Words and nonwords with high phonotactic probability are at an advantage in both comprehension and expression (Beckman et al., 2011; Zamuner, 2009). Infants as young as nine months of age are sensitive to these regularities and selectively give more attention to the words that contain frequently occurring patterns in their native language. Jusczyk et al., (1994) found that nine-month-old infants were more responsive to nonwords with high phonotactic probability than nonwords with low phonotactic probability in the English language. These findings have been repeatedly replicated and confirmed in several other cross-linguistic studies. Zamuner (2009) found a similar relationship between phonotactic probability and mean segment repetition accuracy in two to three year old monolingual Dutch children. Several other studies have shown that by one year of age infants possess skills that are required to acquire language and build the lexicon. In order for a child to be able to segment words, he/she should have

some knowledge about word features and boundaries (Jusczyk, 1999). In a recent study involving computer-based simulation of speech segmentation by infants, Adriaans and Kager (2010) confirmed that sublexical correlates enable segmentation of speech. In addition to this they also found that segmentation of speech is further supported by generalisations of the existing phonotactic probability knowledge over novel utterances. In the latter half of the first year, infants have also been found to show decreased sensitivity to components that are not part of the child's native language (Kuhl, 1983; Werker & Tees, 1984). Findings from these early studies led to a number of successive studies in the area attempting to identify the factors responsible for this decline by examining different aspects of language. The range of findings indicated that between six to twelve months of age, infants begin to build consonant and vowel categories (Kuhl et al., 1992; Pegg & Werker, 1997).

Sensitivity to stress patterns also begins to change from six to nine months of age. For example, by 7.5 months of age, infants are able to segment words that possess stress patterns similar to those of their native language. Jusczyk, Cutler, and Redanz (1993) observed that nine month old infants listened significantly longer to words with stress patterns commonly found in their native language. This difference was not seen in six month old infants. Similarly, when six and nine month old native Dutch and English learning infants were presented with unfamiliar Dutch and English words, it was once again observed that only nine month old infants showed a preference to the words of their respective native languages (Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993). The effects of phonotactic probability were also investigated in six and nine month old infants by Jusczyk et al. (1994). The findings revealed that nine month old infants listened significantly longer to CVC combinations containing high phonotactic probability sequences than those with low phonotactic probability sequences. By ten months of age, infants gain the knowledge of how different sounds are combined to form words in their native language (Friederici & Wessels,

1993; Jusczyk, Cutler et al., 1993; Jusczyk, Friederici et al., 1993) and the frequency with which these segments and biphone sequences occur (Jusczyk et al., 1994). Infants first begin to segment words by relying on suprasegmental and word frequency related cues and over time develop sufficient lexicon and knowledge about the sound structure to rely on segmental properties of words. These findings suggest that sublexical properties play an important role during the critical stages of language development and the second half of the first year in particular. Sublexical properties of the input play a vital role in building the early lexical representations and organizing the lexicon.

1.3.3 Phonotactic probability and short term memory. Short term memory is commonly measured by recall tasks including word repetition, nonword repetition, and serial and immediate recall. During the recall of real/familiar words, information is drawn from the phonological and lexical representations of previously encountered words in the child's repertoire, resulting in an advantage of words over nonwords in recall accuracy. For example, in a serial recall experiment, seven and eight year olds were asked to repeat words and nonwords of varying phonotactic probability. While word recall was at an advantage over nonword recall, they found that short term recall was facilitated by both lexical and phonological knowledge stored in long term memory (Gathercole et al., 1999). That is, word recall recruits a lexical-phonological route for word processing.

When encountered with a nonword or an unfamiliar word, the sublexical route comes into play as the lexical and phonological representations of familiar words in the long-term memory do not match the pattern and structure of the unfamiliar word. As a result, the repetition accuracy is lower for nonwords and unfamiliar words as compared to real words (Gathercole et al., 1999; Hulme, Maughan, & Brown, 1991; Hulme et al., 1997). However, knowledge of Positional segment and Biphone frequencies provides an advantage for the

nonwords and unfamiliar words that are composed of frequently occurring segments and biphones when the lexical-phonological route fails (Edwards et al., 2004; Vitevitch & Luce, 1999). For example, repetition accuracy in seven and eight year olds was found to be higher for nonwords that contained only high probability biphoneme sequences (Gathercole et al., 1999).

Phonotactic probability has been found to aid short term memory functioning (Gathercole et al., 1999). Vitevitch and Luce (2005) found that phonotactic probability has a facilitatory effect on nonword repetition latency. Nonwords with high phonotactic probability were found to be associated with reduced processing duration. Adults are able to identify and recall nonwords containing high probability biphonemes more accurately than low probability biphonemes (Frisch et al., 2000). Children have been found to repeat nonwords that contain embedded real words in their native language more accurately than those that do not (Dollaghan, Biber, & Campbell, 1993). Gathercole and Baddeley (1989) found that children repeated simple nonwords with singleton consonants more accurately than nonwords that contained consonant clusters. In a further study, they observed that children repeated nonwords with high wordlikeness ratings more accurately than the nonwords that were rated to be less wordlike by adults.

1.3.4 Phonotactic probability, vocabulary size and change over time. The effect of phonotactic probability on phoneme repetition is observed to change with developing age (Munson et al., 2005). Munson (2001) investigated the effects of phonotactic probability on nonword repetition in three to four year olds, seven to eight year olds and adults. The nonwords used were bisyllabic with CVCCVC structure wherein the CC sequence contained either a high probability or low probability biphoneme sequence. The three groups of participants were assessed in terms of biphoneme sequence accuracy and fluency. The adults

were able to repeat both high and low probability biphone sequences accurately and fluently while both groups of children had better accuracy and fluency for high probability biphone sequences. The influence of phonotactic probability was greater for three to four year olds when compared with the older children as evidenced by the magnitude of difference in performance between the high and low probability sequences in the two groups. These findings led Munson (2001) to conclude that as vocabulary increases with age, the influence of phonotactic probability decreases and this explains the absence of the effect in adults. Edwards et al. (2004) tested this hypothesis by administering a nonword repetition test across three groups of participants ranging in age from three to four years, five to six years and adults. The results demonstrated findings similar to that of Munson (2001). Accuracy of nonword repetition was related to phonotactic probability and this influence decreased with age across the three groups. Also, outcomes from regression analyses revealed that a significant proportion of variance in the frequency effect was accounted for by vocabulary size (9.9%) and expressive vocabulary (10.1%) in particular. In addition to being correlated with phonotactic probability, vocabulary has also been found to correlate with nonword repetition accuracy with approximately 36% of the variance in vocabulary size being accounted for by nonword repetition in two year olds (Stokes & Klee, 2009b).

Seeking to identify the mechanisms underlying these phenomena, Munson et al. (2005) pointed out that as vocabulary size increases, the child has sufficient exposure to low probability combinations, thereby minimizing the frequency effect. For a word to be stored in the expressive vocabulary of a child, in addition to the auditory-acoustic representation required for storage in the receptive vocabulary, the child must also have a robust articulatory-motor representation for the word (Edwards et al., 2004). Repeated exposure to words develops a strong association between the semantic, articulatory, phonological and acoustic representations of the words in the lexicon. This allows for an ease in manipulating

and producing different segmental and sequential combinations of phonemes when required to imitate a nonword (Edwards et al., 2004). Thus nonword repetition accuracy is influenced by both lexical and phonological knowledge that develops with an increase in vocabulary size. Children with a small vocabulary size have limited lexical and frequency information to rely on during verbal memory tasks as a result of which they could be more dependent on their short term memory to reconstruct the word that was presented thereby leading to a decrease in repetition accuracy. Gathercole et al. (1999) found significantly higher repetition accuracy for nonword recall in participants with larger vocabularies. As a further confirmation of the association between sublexical correlates and vocabulary size, Rispens, Baker, and Duinmeijer (2015) reported that older TD participants ranging from seven to eight years of age had good repetition accuracy for nonwords with low phonotactic probability. This led the authors to conclude that an increase in vocabulary size provided the older participants with an ease of access to store lexical and phonological representations that are robust and flexible, aiding recall, confirming Edwards et al.'s (2004) hypotheses. Finally, variation in vocabulary size in adults even impacted on adult ratings of nonwords as being of low or high wordlikeness. Frisch (as cited in Edwards et al., 2004) found that adults with larger vocabulary sizes had different rating points for low biphone frequency nonwords with varying frequencies. Whereas adults with a smaller vocabulary size gave a uniform low rating to all sequences with different low biphone frequencies. The range of findings highlights the importance of vocabulary size, age and their collective influence on the effects of phonotactic probability on nonword recall. To date, research on these relationships has been mostly limited to cross sectional studies and children aged 4 and over, even though variability in vocabulary size emerges from the onset of word production.

1.4 Variability in development of the spoken lexicon

Nonword repetition, in association with standardised tests of language, has been a useful tool in detecting phonological difficulties in late talkers (LTs) and a great deal of research has shown differential performance in such verbal recall tasks by children with early language delay. Previously labels such as ‘Specific Expressive Language Delay’ (SELD), ‘Late Language Emergence’ (LLE) and ‘Early Language Delay’ (ELD) were used to refer to this group of children. Language delay is usually identified between 18 to 30 months of age and is characterised by a small early word inventory progressing into a slow or deviant growth in vocabulary followed by an absence or delay in the emergence of word combinations. There are several parent-report questionnaires developed to identify LTs and the standards set by these measures provide the criteria for defining LTs. The definitions are predominantly based on the expressive vocabulary of the children. For example, the Language Development Survey (LDS; Rescorla, 1989) defines LTs as the children who either have an expressive vocabulary of less than 50 words or an absence of word combinations by approximately 2 years of age (e.g. Rescorla, 1989). The Ages and Stages Questionnaire Communication Scale (ASQ; Bricker & Squires, 1999) defines LTs as those whose combined score on the ASQ lies 1 SD below the mean (Zubrick et al., 2007). According to the NZ CDI: WS (Reese & Read, 2000), LTs are the children who score below the 10th percentile on the expressive vocabulary section of the questionnaire (e.g. Reilly et al., 2007). Based on these measures, approximately 10 to 20% of children are being identified as LTs. LTs are extensively studied because they are at risk of developing language disorders (Olswang et al., 1998). Poor performance in nonword repetition tasks by children with SLI has been linked to issues in phonological processing (Chiat, 2001). Therefore, identification of crucial developmental factors can help focus intervention approaches and highlight areas that could

be focused on during critical stages of language development to help evade problems in these areas (Desmarais, Sylvestre, Meyer, Bairati, & Rouleau, 2008). Although it is believed that most children are able to outgrow this delay and achieve age appropriate language skills by preschool age (Dale, Price, Bishop, & Plomin, 2003), a considerable number of LTs move on to develop a clinically diagnosed language disorder such as SLI (Leonard, 1998). The children who recover from the delay by four to five years of age have been observed to demonstrate delayed syntactic development. While there is a 50 to 70% recovery rate for LTs (Leonard, 1998), the chances of recovery are highly limited for children with SLI who are usually identified at age four. SLI refers to a language impairment that manifests itself in the absence of any co-existing medical, neurological or cognitive impairments (Leonard, 1998). Children with SLI exhibit linguistic, academic, behavioural and social deficits (Bishop & Snowling, 2004; Tomblin, Zhang, Buckwalter, & Catts, 2000). Stothard et al. (1998) conducted a longitudinal study on children identified as having SLI at age four. The study involved observation of linguistic characteristics in children who showed recovery at age 5.6 and those who did not. While the children who continued to have SLI at age 5.6 had significant impairments across all areas of spoken and written language, the children who recovered demonstrated academic and phonological impairments when compared with their age-matched peers at 15 to 16 years of age. Even with significant advancement in research and innovative technology it has not yet been possible to ascertain which of the LTs will go on to develop a language disorder (Dale et al., 2003; Eernisse, 2010) and which of them are simply 'late bloomers' (Stokes, 2010). Hence it becomes necessary to identify and manage the predictor variables earlier on, which leads us to the observation of language development in LTs.

Several studies have identified that LTs as a group perform significantly lower than their age-matched peers on a number of tasks related to different linguistic components such

as vocabulary, verbal recall, expressive language, fast mapping and phonology that persist even later in life (Paul, 1996; Rescorla, 2002, 2009; Weismer & Evans, 2002). These differences have been observed right from infancy to be associated with reduced canonical babbling and moving on to a restricted phonetic repertoire with simplified utterances observed in speech (MacRoy-Higgins et al., 2013). Tomblin and Samulson (as cited in Desmarais et al., 2008) observed that children identified as LTs developed problems in pragmatic and academic issues as well. Rescorla, Alley, and Christine (2001) noted that the nature of words in the vocabulary of LTs is similar to their TD peers and the only differentiating factor is the size of the vocabulary. Although both groups of children acquire similar words, a number of studies have identified key elements that affect the performance of LTs suggesting that there is more to language delay than just the vocabulary size. Assessment of language and literacy skills in six to nine year olds identified as LTs at 24 to 30 months of age revealed that LTs had a consistent poor performance across most of the linguistic measures. Differences in syntactic structure, phonology and vocabulary were observed at six years of age. Differences in vocabulary alone were observed at age seven, followed by lower performance on auditory comprehension, reading and syntactic skills in addition to limitations in vocabulary at age eight. At nine years of age only differences in reading ability were observed (Rescorla, 2002).

Novel words are also learnt at different rates by LTs and their age-matched TD peers. Weismer and Evans (2002) observed novel word acquisition in LTs and age-matched TD children. The stimuli used consisted of two familiar words and two nonwords associated with referents and presented in the form of a story using puppets. The findings revealed similar receptive and expressive language scores across the two groups for familiar words while for nonword stimuli the receptive scores were significantly poor in the LTs group. Overall both groups performed poorly on the expressive language task. They concluded that LTs do indeed

have difficulty creating and storing new lexical and phonological representations when compared with their TD peers. MacRoy-Higgins et al. (2013) found that LTs demonstrated lower levels of performance across the reception, expression and detection of mispronunciation tasks assessing the influence of sublexical correlates on lexical acquisition in toddlers.

The relationship between phonology and expressive vocabulary has been well documented in both TD children and LTs. A larger vocabulary is associated with a diverse phonetic repertoire whereas vocabularies that are smaller in size are found to have a restricted phonetic inventory (MacRoy-Higgins & Schwartz, 2013). This relationship was observed in LTs by Thal, Oroz, and McCaw (1995) where toddlers with a vocabulary size of less than 10 words exhibited basic phonological skills while those with a vocabulary size of more than 10 words demonstrated superior phonological skills as evidenced by the complexity and intelligibility of the utterances produced. However not all studies converge on the nature of the phonological differences. While some studies have found a delay in the acquisition of phonological characteristics with similarities to younger expressive vocabulary matched controls, others have found LTs to have atypical phonology. Paul and Jennings (1992) found similarities in syllable structure and consonant categories between LTs and younger vocabulary-matched TD children. Similarly, Thal et al. (1995) found that TD children and LTs had similar phonetic repertoires and syllable structures. Williams and Elbert (2003) observed the phonological and linguistic characteristics in five children identified as LTs and reported that while three of the children were able to outgrow the delay and catch up with their TD peers, two children continued to show a persistent delay. Detailed analysis of their phonological and linguistic characteristics at three years of age revealed differential phonological errors, deviant phonological characteristics and restricted language development. The analysis of phonological skills in LTs has revealed a restricted consonant

and vowel inventory with an inability to use complex syllabic structures. These findings suggest that the mechanisms underlying language acquisition are different in LTs. During the course of typical language acquisition different communication variables are known to develop in parallel. The existence of a phonological deficit in the presence of a language delay further supports this fact.

1.5 Late Talkers, Language Impairment and sensitivity to phonotactic probability

It has been well established that TD children are sensitive to phonotactic probability and they access this knowledge during the course of development. However, studies related to phonotactic probability have shown atypical results in LTs suggesting that they may not be as sensitive to these regularities as their TD peers (Stokes, 2010). Eernisse (2010) tested unfamiliar word recognition ability in LTs and TD children between the age range of 24 and 42 months. Both groups of children were presented with novel words that varied only in phonotactic probability. There were no differences in performance for both high and low probability nonwords for the TD children. LTs on the other hand had better scores for the low probability nonwords. The ease of recognizing low probability target stimuli in LTs was attributed to more clear-cut representation of the target in the lexicon due to lesser interference from related words already in the vocabulary. Other researchers have suggested that low probability words may be better recognized by children as they differ from the commonly occurring words in the lexicon and hence gain greater attention from the child (Gierut & Dale, 2007).

Phonotactic probability influences have also been evident in children with language disorders including SLI. Children with SLI exhibit difficulties in nonword repetition. Graf-Estes, Evans, and Else-Quest (2007) found the performance of children with SLI on a nonword repetition task to be 1.27 standard deviations below that of their TD peers. While

the influence of phonotactic probability on nonword repetition accuracy has been reported to be similar to their age matched TD peers, as the complexity is increased by varying the phonotactic probability across multiple positions in the nonword, the frequency effect has been found to be enhanced in children with SLI. In addition to this, analysis of substitution errors in both groups of children has identified greater errors on vowels than consonants with children substituting a low probability segment or biphone with a high probability counterpart thereby increasing the overall phonotactic probability of the nonword (Burke & Coady, 2015). Munson et al. (2005) found that children with SLI had better repetition accuracy for nonwords with high phonotactic probability. The phonotactic probability influences in children with SLI were found to be similar to younger TD vocabulary matched peers. Similarly, children with Phonological Delays have also been found to exhibit comparable effects of phonotactic probability with TD peers but with a greater frequency effect (Storkel & Hoover, 2010). The continuing influence of phonotactic probability in children with language disorders and the qualitative differences in sensitivity observed between LTs and age-matched TD peers warrants further research in understanding the mechanisms operating in LTs.

1.6 Rationale for the current study

As described earlier, both measures of phonotactic probability (Positional Segment Frequency and Biphone Frequency) have been used to study language processing in children. Most studies have employed either a combination of Positional Segment and Biphone Frequency measures or a single measure in isolation to understand the sensitivity to probabilistic phonotactics. The current study differs from previous work in the way phonotactic probability effects are examined. While most studies on phonotactic probability have developed stimuli by manipulating syllable structure to account for the two measures of

phonotactic probability, this study will examine children's sensitivity to the two measures of phonotactic probability during nonword repetition, using the Test of Early Nonword Repetition - Revised (TENR-R; Stokes & Klee, 2011). It is well established that children are sensitive to regularities and patterns in their native language by the second half of the first year. However, there is little information on how early this sensitivity is reflected in their speech production. This study will examine the influence of phonotactic probability in production by calculating each of the two measures of phonotactic probability as a function of repetition accuracy. The first research question will explore the impact of phonotactic probability on the accuracy of responses on the TENR-R.

Coady and Aslin (2004) reported one of the few studies that looked into the influence of each phonotactic probability measure on nonword repetition in 2.5 and 3.5 year olds. They hypothesized that if children stored words holistically, they should be sensitive to Biphone Frequency and if the lexical representations are more fine-grained, they should be sensitive to Positional Segment Frequency (Coady & Aslin, 2004). In their study, both groups of children demonstrated sensitivity to Positional Segment Frequency cues when manipulated across the entire nonwords. This sensitivity was observed to undergo refinement with age as evidenced by age-related improvement in performance when the manipulation of Positional Segment Frequency was limited to the syllable-initial position alone. If sensitivity to Positional Segment Frequency improves with age, then the early lexical representations of children must be holistic in nature. Hence Coady and Aslin (2004) defined the early representations to be smaller than a whole syllable unit (CVC) but larger than individual phonemes - biphones. Their findings revealed that sensitivity to Biphone Frequency too improves with age. These results were attributed to a possible floor effect due to the overall lower phonotactic probability of the nonwords for the biphone frequency task and the low repetition accuracy observed in the younger children (Coady & Aslin, 2004). Through the current study we seek

to further explore young children's language organization using a larger sample size, with the hypothesis that lexical representations in very young children are discontinuous and more holistic and therefore younger children must be sensitive to Biphone Frequency cues. Hence the second research question will attempt to identify the measure of phonotactic probability that is most closely associated with accuracy on the TENR-R.

Although there is a growing body of research on the effects of sublexical correlates like phonotactic probability on language systems, evidence from longitudinal studies is sparse. The study conducted by Bowey (2001) is one of the few longitudinal studies that has attempted to view the relationship between nonword repetition and receptive vocabulary at three, four and five years of age. As noted earlier, phonotactic probability is influenced by expressive vocabulary size (Munson, 2001), and children with smaller vocabularies are more sensitive to frequency than those with larger vocabularies (Edwards et al., 2004). Similarly, LTs have been found to show differential sensitivity to phonotactic probability than their TD peers. While LTs are predominantly known to have an expressive language deficit, studies by MacRoy-Higgins and Schwartz (2013), Thal et al. (1995) and Williams and Elbert (2003) have shown the influence of phonology in the performance of LTs characterised by a limited vocabulary size. An understanding of the qualitative differences in LTs can have important clinical implications. Neighborhood density and phonotactic probability have been found to influence generalisation skills in children receiving clinical intervention (Morrisette & Gierut, 2002). Keeping all this in mind, there is a need for longitudinal studies, commencing with younger children that take these factors into account. Hence the current study investigates the relationship between phonotactic probability and repetition accuracy on the TENR-R in two groups of participants classified based on their expressive vocabulary size into LTs and TD children. The study aims to shed light on age-related changes on the sensitivity to sublexical correlates in nonword processing among TD and LTs.

Finally, there are two issues to be briefly mentioned before the research questions are presented. First, there are two main ways to score nonword repetition accuracy. The first type of scoring is by word item and second is by phoneme (Percentage of Phonemes Correct). Gathercole et al. (1999) found no significant difference between scoring methods in a study of the influence of phonotactic probability on nonword repetition for low and high vocabulary groups. Stokes and Klee (2009b) used PPC and reported a moderate significant correlation between CDI scores and Percentage of Phonemes Correct (PPC) for the TENR. As a replication of the Stokes and Klee's (2009b) findings the current study also examines the association between vocabulary size and phoneme repetition accuracy on the TENR-R with the hypothesis that phoneme repetition accuracy improves with an increase in vocabulary size.

Second, as nonword length increases, so does phonetic complexity. One would expect the influence of phonotactic probability to be affected by an increase in phonetic complexity. However Goldrick and Larson (2008) found that the influence of phonotactic probability on the processes involved in speech production is independent of phonetic complexity. Hence the final research question will examine the relationship between the two measures of phonotactic probability and nonword length of the TENR-R stimuli.

1.7 Aims of the study

The main aims of the study were to investigate the age-related sensitivity to sublexical correlates across a period of 18 months in preschoolers and to compare the sensitivity to these phonotactic probability measures among TD and LTs using the TENR-R.

1.8 Research questions

1. What is the impact of phonotactic probability on the accuracy of response on the TENR-R?

2. To what extent does vocabulary size play a role on the accuracy of responses and how does the role of sublexical properties change with age?
3. Which measure of phonotactic probability at Time 3 (T3) is most closely associated with T3 TENR-R scores?
4. Is there a significant difference in the mean values of the descriptive variables (proportion of TENR-R nonwords correct, Positional Segment Frequency average and Biphone Frequency average) between TD children and LTs at T3?
5. Is there an association between vocabulary size and phoneme repetition accuracy on the TENR-R?
6. Are the two measures of phonotactic probability associated with nonword length on the TENR-R?

2.0 Method

2.1 Ethical Considerations

Ethics approval was obtained from the University of Canterbury Research Ethics Committee for all members on the research team for “Learning to talk: a research project on children’s early language development” (LTT: Klee, Stokes, & Moran, 2011-2015).

2.2 Participants

The participants for this study were enrolled in a long term project titled “Learning to Talk: A research project on children’s early language development” (Klee et al., 2011-2015). The participants were recruited through early childhood specialists, by word of mouth from parents of the participants, and advertisement in nurseries. Data collection was carried out at three points in time: Time 1 (T1; 24-30 months), Time 2 (T2; 27-33 months) and Time 3 (T3; 42-49 months). T2 results are not discussed in the current study.

A total of 168 participants between 24 and 30 months of age ($M= 27.44$, $SD= 1.71$) participated in the study at T1. The ‘Test of Early Nonword Repetition-Revised’ (TENR-R; Stokes & Klee, 2011) was administered at T1 and T3 (see Appendix A for the TENR-R form). The participants were predominantly monolingual English speakers. The sample included 22 participants at T1 and 17 participants at T3 who had been exposed to more than one language. A summary of additional languages the participants were exposed to at T1 and T3 is presented in Table 1.

Table 1: *Summary of Languages Exposed to in Addition to English at T1 (N=22) and T3 (N=17)*

Languages Exposed to	No. of Participants	
	T1	T3
Maori	9	3
Sign Language	4	1
German	2	1
Samoan	2	3
Cantonese	1	1
Italian	1	1
Spanish	1	2
Japanese	-	1
Afrikaans	1	1
French	1	1
Dutch	-	1
Welsh	-	1

Among the 168 children invited to participate at T1, 134 participants between 24 and 30 months of age ($M= 27.38$, $SD= 1.64$) attempted the TENR-R at T1. The remaining 34 participants were excluded due to the following reasons: technical problems such as an inaudible audio track in the recorded clip or problems with the computer system presenting the stimuli (3), family relocation (1), no video (17), TENR-R not administered due to poor compliance towards preceding tasks (13). A total of 138 children between the age range of 42

and 49 months ($M= 45.36$, $SD= 1.97$) participated in the TENR-R at T3. The remaining 30 participants were excluded due to the following reasons: technical issues such as an inaudible audio track in the recorded clip or problems with the system presenting the stimuli (4), missing video (1), family relocation (1), TENR-R not administered due to poor compliance towards preceding tasks (5) and attrition from the study (19). Overall, a total of 123 participants attempted the TENR-R at both T1 and T3. The sample of participants included 68 boys (55%) and 55 girls (45%). The Preschool Language Scale, Fourth Edition (PLS4; Zimmerman, Steiner, & Pond, 2002) and the New Zealand adaptation of the MacArthur-Bates Communicative Development Inventory: Words and Sentences (NZ CDI: WS; Reese & Read, 2000) were administered as part of a battery of tests in the LTT project (Klee et al., 2011-2015) to document the language abilities of the participants. The children were classified into two groups based on the responses to the NZ CDI: WS questionnaire—Typically Developing (TD) and Late Talkers (LT), with a ratio of 2:1. The children with age appropriate scores were included in the TD group, while the children with CDI scores at or below the 10th percentile for age and sex, referenced against the USA norms (Fenson et al., 2007) were classified as LTs, for early identification of risk for language disorders. Of the 134 participants at T1, 33 were classified as LTs and 101 were classified as TD. Four more TD children attempted the TENR-R at T3, so the total sample size at T3 was 138.

2.3 Materials/Tests

Interested families received a mailed pack containing a consent form, a parent questionnaire and NZ CDI: WS (Reese & Read, 2000). The parent questionnaire focused on demographic information of the children, including the child's birth history, medical history and family history. The NZ CDI: WS (Reese & Read, 2000) was developed to adapt the American version of the CDI: WS (Fenson et al., 1993) to the New Zealand dialect. The New

Zealand English adaptation of the CDI: WS retains the structural organization of the ‘Sentences and Grammar’ section (Part II) found in the original inventory. The only modification is in the ‘Words Children Use’ section (Part I) of the inventory where forty one words have been replaced to match the New Zealand English dialect (Reese & Read, 2000). The NZ CDI: WS consists of two parts: Words children use (Part I) and Sentences and Grammar (Part II). Part I includes a total of 676 words, used in everyday life, grouped across 22 categories. Part II assesses a child’s knowledge and use of different sentence types and grammatical structures using a word form inventory and a three point rating scale. The child’s spoken vocabulary size is measured by calculating the total words in the inventory that are marked by the parent in the form provided.

On receipt of the signed consent form and the questionnaires, the participants meeting the selection criteria were shortlisted for the project. The TENR-R (Stokes & Klee, 2011) was administered as part of a battery of tests designed for the LTT project. The nonwords in the TENR-R are designed such that the constituent consonants are generally within the phonetic repertoire of TD two year olds (Stokes & Klee, 2009a). The TENR-R consisted of a list of 20 nonwords ranging from one to five syllables in length. [The original TENR-R included nonwords from one to four syllables in length, Stokes & Klee, 2009a.]

2.4 Procedure

The children were individually tested by one of four research assistants at T1 and three other research assistants at T3. One research assistant was common to both T1 and T3. The research assistants were trained to carry out all the tests included in the project. The testing was carried out at the Child Language Centre (CLC), University of Canterbury, New Zealand. Each session was audio and video recorded using Beyerdynamic MPC22 boundary mics and Pelco SD423-SMB-1-X wall mounted cameras. A computer based version of the

TENR-R (Stokes & Klee, 2011), recorded by a female speaker, was administered at T1 and T3. Live voice was occasionally used when a child was noncompliant to the computer based task or was inattentive when the target was presented. The children were asked to repeat the nonwords after hearing a standard set of instructions given by the research assistant. The test involved one practice trial where the child was instructed to repeat the word ‘Teddy’. The pictures following the stimuli served as reinforcements. Additional reinforcements such as rolling cars down a track and feeding the dinosaur were provided when pictures alone were insufficient for a child to perform the task.

2.5 Transcription

Children’s speech productions on the TENR-R were transcribed using the Phon program (Rose et al., 2006; Rose & MacWhinney, 2014). Phon is a software program developed to aid in the study of speech through transcription and phonological analysis of acoustic information from the recordings. Phon has additional features that allow for an in-depth phonological analysis using queries such as features, syllables, epenthesis, metathesis, etc. The nonwords were transcribed using Phon version 1.6 or 2.0.6 and the latter version was used for scoring and analysis. The IPA for New Zealand English (Bauer et al., 2007) was used to transcribe the nonword repetitions. The audio and video recordings were available during the transcription task and each transcriber used them while transcribing. The nonword imitations of the participants were segmented from the audio-video session recording and waveforms were generated for each segment that could be replayed as many times as necessary before transcribing the nonword. An extensive set of guidelines was developed to maintain uniformity in transcription between two transcribers.

2.5.1 Transcription guidelines

1. Listen to each production four times before transcribing.
2. In cases with spontaneous child multiple repetitions of the nonword, the best production is taken into account.
3. Voiced/voiceless substitutions (t/d, k/g, p/b) are scored as correct because this distinction is difficult to make while transcribing (Shriberg & Lohmeier, 2008).
4. Addition of phonemes is ignored because the production continues to represent the original nonword without any loss of information (Dollaghan & Campbell, 1998).
5. A real word substitution for a nonword is scored 0 for overall accuracy but scored correct for every phoneme correctly produced. For example, if a child substitutes the word /kɛ:ɪ/ for the nonword /mɛ:d/, the overall nonword is scored as incorrect. However, the child is credited for every phoneme produced that matched with the target. In the case of this example, the child is scored for the phoneme /ɛ:/. The nonword presented could have activated the real word in the child's memory trace and hence the corresponding phonemes correctly produced should receive credit.
6. A preceding nonword repeated continuously for a subsequent item is scored 0.
7. If the test is discontinued (TX; see Table 2), the child is scored 0 for all items discontinued.
8. If a nonword is presented but not attempted by the child (CX; see Table 2), score 0.
9. If a nonword is presented but not attempted by the child because he was distracted when the word was presented, the item is excluded from the total for that participant.
10. A child who fails to imitate the practice word due to poor imitation skills is included and scored 0 for all items on the TENR-R.

11. Change in order of presentation or later re-representation of a nonword stimulus is acceptable if it is evident that the child isn't attending to the task or hasn't understood the task.
12. Use of live voice to repeat a stimulus if the child did not hear the word is acceptable and is scored.
13. Inaudible and unintelligible responses are represented by the same symbol used for unintelligibility (X; see table 2) in Phon (Rose et al., 2006; Rose & MacWhinney, 2014).
14. Inaudible and unintelligible sounds and segments are scored 0.
15. Productions with transpositions (metathesis) or epenthesis (addition/insertion) are scored 0 for whole word accuracy but are credited for each phoneme correctly attempted during the percentage of consonants correct (PCC) and percentage of vowels correct (PVC) calculations.
16. Consistent misarticulation of the target phoneme is scored as correct.

Table 2: *Additional Notes Added to Phon during Transcription:*

Code	Indicates
IC	Correct imitation of the nonword but on repetition after the research assistant following the computer-based stimulus presentation (Score 0)
IX	Incorrect imitation of the nonword even on repetition after the research assistant following the computer-based stimulus presentation (Score 0)
CX	No response to the presented nonword stimulus (score 0)
RX	Target nonword skipped
TX	Test discontinued due to poor performance or non compliance
MX	The child's production cannot be transcribed confidently due to objects obstructing the mouth
VX	Child's production cannot be transcribed confidently due to soft voice, background noise
X	Unintelligible phoneme or vegetative sounds that cannot be transcribed confidently even after four listening attempts
ATU	Alternative Target Used

2.6 Inter-transcriber agreement

Two trained transcribers independently and blindly transcribed a random sample of 15% to 20% of the TENR-R recordings from Time 1 (N=29) and Time 3 (N=22). Functional and near functional equivalence criteria (Shriberg & Kent, 2013) were used to measure the percentage of agreement between the two transcribers. In this criterion, phonemes that matched in terms of place and manner of articulation between the two transcribers for a target nonword were considered to be in agreement and scored. The percentage of agreement was calculated separately for consonants, vowels and both combined (overall). Each transcribed segment that was found to be in agreement was given a score of 1 and a consonant, vowel

and overall agreement total was obtained for T1 and T3. This total was divided by the number of segments (consonant, vowel and overall) present in the transcription of the nonwords to obtain the percentage of agreement between the two transcribers. If there was a discrepancy in the number of segments present in the transcription by the two transcribers, the transcription with higher number of segments was taken into consideration. The following percentages of agreement were obtained for Time 1: consonants (87.12%), vowels (88.06%), overall (87.53%). For Time 3 the following percentages of agreement were obtained: consonants (80%), vowels (83.24%) and overall (81.45%).

2.7 Scoring

Scoring was carried out by a trained Speech-Language Pathologist blind to the participants' demographic data, assessments and language status. Phon 2.0.6 (Rose et al., 2006; Rose & MacWhinney, 2014) program was used to generate the scores for the TENR-R (Stokes & Klee, 2011). Three types of scores were generated from the program:

2.7.1 Word match. Nonword repetition was scored as correct or incorrect based on the overall repetition accuracy for each nonword using the 'Word match' query in Phon (Rose et al., 2006; Rose & MacWhinney, 2014). The program automatically matched the child's production with the IPA target. Correct responses were represented as TRUE and incorrect responses were represented as FALSE by the query output. These scores were then converted to numerals with 1 representing a TRUE score and 0 representing a FALSE score using the 'If' conditional in Apache Open Office 4.1.1. An overall total score on the TENR-R was obtained for each child and this was converted to a proportion score by dividing the total number of nonwords correctly imitated with the total number of nonwords in the TENR-R.

2.7.2 PCC/PVC (Standard). Percentage of Consonants Correct (PCC) and Percentage of Vowels Correct (PVC) for each nonword were calculated using the PCC/PVC (standard) query in Phon. This method scored consonants and vowels as correct irrespective of their position in the child's production when compared with the target. For example, if a child repeated the nonword /bɛləkɒn/ as /ləbəkɒn/, the PCC/PVC standard query would give a 100% score for PCC even though the consonants /b/ and /l/ have been transposed. The output included the raw scores for consonants and vowels in the target, number of consonants and vowels attempted by the child, number of consonants and vowels correct and the PCC and PVC for each nonword. In addition to these, the output sheet also provided raw scores and percentages for phonemes deleted and epenthesized in each nonword. The raw scores from Phon were further used to calculate the Percentage of Phonemes Correct (PCC) for each nonword followed by an overall PCC, PVC and PCC for each child at T1 and T3 using Apache Open Office 4.1.1.

2.7.3 APCC/ APVC (Aligned). The Aligned Percentage of Consonants Correct (APCC) and Aligned Percentage of Vowels Correct (APVC) were calculated using the PCC/PVC (Aligned) query in Phon. In this method of scoring, phonemes are scored as correct only if they occur in the same position as the target. In the same example as above, a child repeating the nonword /bɛləkɒn/ as /ləbəkɒn/ will be given an APCC score of 50%. The raw scores were then used to generate the Aligned Percentage of Phonemes Correct (APPC) followed by the overall APCC, APVC and the APPC for each child using Apache Open Office 4.1.1 at both T1 and T3.

2.8 Data analysis

The phonotactic probabilities of the TENR-R nonwords (Stokes & Klee, 2011) were calculated from an R-script (LaShell, 2014) using data in the CELEX database. The CELEX is a database of lexicons for English, German and Dutch. The CELEX British-English database (Baayen, Piepenbrock, & Gulikers, 1995) was used to calculate the phonotactic probabilities of the nonwords in the TENR-R. This database has approximately 17.9 million tokens from which the word frequencies and probabilities are calculated. CELEX only approximates spoken values since 90% of the database is derived from written corpora. Several studies, including that of Pisoni and Garber (1990) have found no significant difference between measures derived from spoken and written language use and hence it was suggested that the findings from one modality could be generalised to the other modality with some degree of confidence. Therefore it seems reasonable to derive measures from a database that is based on written language. Although the database includes alternative phonologies for many words, for the purpose of this study the phonotactic probabilities were derived from the base phonology. The ‘R project for statistical computing’ was used to calculate both measures of phonotactic probability from the CELEX database by LaShell (2014). The two measures of phonotactic probability- Positional Segment Frequency and Biphone Frequency- were calculated for the 20 nonwords using a method described by Vitevitch and Luce (2004) in their study to develop a web-based program that calculates phonotactic probability for words and nonwords for American English. The script used to calculate the two measures of phonotactic probability can be found in Appendix B. The phonotactic probability of each segment and biphone in a nonword were then averaged to obtain an overall Positional Segment Frequency average and Biphone Frequency average respectively for each of the 20 nonwords on the TENR-R (see Appendix C for the average phonotactic probability of each

nonword in the TENR-R). The method used to calculate the two measures of phonotactic probability is described below.

2.8.1 Positional Segment Frequency. The Positional Segment Frequency of a particular phoneme was calculated by finding all the words in the CELEX database, irrespective of word length, that had the target phoneme in the same position as the target nonword. The sum of the log (base 10) values of the frequencies with which those words occurred in the database was then divided by the sum of the log (base 10) frequency of all words in the database that had any segment in that word position. For example, the Positional Segment Frequency of /f/ in the nonword /fʌ:pəm/ can be calculated by dividing the sum of the log (base 10) frequencies of all words in the CELEX database that contain the phoneme /f/ in the initial position with the sum of the log (base 10) frequency of all words in the database that have an initial consonant.

2.8.2 Biphone Frequency. The Biphone Frequency was calculated by finding all the words in the CELEX database, irrespective of word length, that had the target biphone in the same position as the target nonword. The sum of the log (base 10) values of the frequencies with which those words occurred in the database was then divided by the sum of the log (base 10) frequency of all words in the database that had any biphone combination in that word position. For example, the Biphone Frequency of /lɛ/ in the nonword /wʊgələmæk/ can be calculated by dividing the sum of the log (base 10) frequencies of all words in the CELEX database that contain the biphone /lɛ/ in the same position as the target nonword with the sum of the log (base 10) frequency of all words in the database that had any biphone in that word position.

These two phonotactic probability values were assigned for each nonword that a child repeated correctly. A Microsoft Access database has been created for the LTT project (Klee et al., 2011-2015). This database stores information relating to the demographic details of the participants such as age, gender, languages exposed and scores on all the tests administered as part of the LTT project. The scores on the TENR-R (Stokes & Klee, 2011) and the phonotactic probability values (Positional Segment and Biphone Frequency) for each nonword correctly repeated were also entered into this database. Using this master database queries were generated to compute an overall average Positional Segment and Biphone Frequency for each child (Anderson & Byrd, 2008; Guo, McGregor & Spencer, 2015). For example, if a child repeated /mɜ:d/, /kɛʊə/, /dɜ:fi/ and /fi:sɛmɒt/ correctly, the phonotactic probability values for each of those nonwords would be averaged to generate an average Positional Segment Frequency and an average Biphone Frequency for the child. This data along with raw TENR-R scores (word match, A/PPC, A/PVC and A/PCC) and PLS scores at T1 and T3, CDI scores at T1 and information on participants' age and gender were compiled and saved into a single Microsoft Excel spreadsheet using the queries generated in the Microsoft Access master database.

3.0 Results

All statistical analyses were carried out using IBM SPSS Statistics 20.0 (IBM Corporation, 2011). The language profile of the participants is summarised below. The raw scores of the participants on the two subscales (auditory comprehension and expressive communication) of the PLS4 are summarised in Table 3 for both T1 and T3.

Table 3: *Language Profile of Participants based on PLS4 (N=123) Raw Scores*

	N	Mean	Std. Deviation	Range
T1 AC	123	34.84	5.90	32
T1 EC	123	37.54	6.63	30
T3 AC	123	52.42	4.72	27
T3 EC	123	56.54	5.30	30

Note. AC= Auditory comprehension, EC= Expressive communication, T1= Time 1, T3= Time 3

3.1 Descriptive statistics

The means, standard deviations, standard errors, minimum and maximum values for the variables proportion of nonwords correctly imitated, Positional Segment Frequency average (PSF average) and Biphone Frequency average (BF average) at T1 and T3 are presented below (Table 4 and Table 5). The sample size is lower for PSF average and BF average in both groups at T1 because there were participants who scored 0 on the TENR-R

and hence did not have a corresponding Positional Segment Frequency and Biphone Frequency value for any nonword.

Table 4: *Descriptive Statistics for Typically Developing Children at T1 and T3*

	N	Minimum	Maximum	Mean	Std. Error	Std. Deviation
T1 Proportion correct	101	.00	.55	.20	.01	.13
T1 PSF average	91	.0327	.0530	.0427	.0004	.0041
T1 BF average	91	.0013	.0041	.0022	.0000	.0005
T3 Proportion correct	105	.10	.80	.40	.01	.15
T3 PSF average	105	.0327	.0534	.0431	.0003	.0031
T3 BF average	105	.0014	.0034	.0023	.0000	.0004

Note. Numerals are reported to four decimal places because the values are small. T1= Time 1, T3= Time 3, Proportion Correct= Proportion of nonwords correctly imitated on the TENR-R, PSF average= Positional Segment Frequency average, BF average= Biphone Frequency average.

Table 5: *Descriptive Statistics for Late Talkers at T1 and T3*

	N	Minimum	Maximum	Mean	Std. Error	Std. Deviation
T1 Proportion correct	33	.00	.15	.05	.01	.05
T1 PSF average	17	.0343	.0592	.0434	.0015	.0063
T1 BF average	17	.0013	.0026	.0021	.0001	.0004
T3 Proportion correct	33	.05	.45	.23	.02	.11
T3 PSF average	33	.0305	.0534	.0418	.0009	.0050
T3 BF average	33	.0009	.0044	.0021	.0001	.0006

Note. Numerals are reported to four decimal places because the values are small. T1= Time 1, T3= Time 3, Proportion Correct= Proportion of nonwords correctly imitated on the TENR-R, PSF average= Positional Segment Frequency average, BF average= Biphone Frequency average.

The descriptive statistics have been presented separately for TD children (Table 4) and LTs (Table 5). NWR performance of LTs at both T1 and T3 was lower than that of the TD participants. Overall T3 participants had better scores than T1 participants. In addition to these findings, the descriptive statistics also revealed that the LTs at T3 had all the three mean scores very similar to that of the TD group from T1.

This difference in performance between LTs and TD children at T1 and T3 was further explored using statistical measures to answer the research questions.

Research questions 1 and 2

*What is the impact of phonotactic probability on the accuracy of response on the TENR-R?
To what extent does vocabulary size play a role on the accuracy of responses and how does the role of sublexical properties change with age?*

Both measures of phonotactic probability have been known to influence nonword repetition with high probability nonwords being repeated more accurately than low probability ones. Findings from prior studies have led us to believe that the early lexical representations of young children are holistic in nature, with sensitivity to Positional Segment Frequency developing with age. Hence, first we attempted to understand the relative influence of both measures of phonotactic probability on the accuracy of nonword repetition on the TENR-R at T1 and T3, and second, by taking age and vocabulary into account. Based on this understanding, we expected participants at both T1 and T3 to be more sensitive to Biphone Frequency than Positional Segment Frequency, with sensitivity to Positional Segment Frequency getting stronger with age and increasing with vocabulary size. We also expected the LTs at T3 to show findings similar to TD children from T1 due to the similarity in scores across all the three variables of interest. To test these hypotheses a correlation analysis was carried out.

3.2 Correlation Analysis. Bivariate correlation analysis was carried out to establish the relationship among the proportion of nonwords correctly imitated and the phonotactic probability measures using Pearson's product-moment correlation (denoted by 'r') at T1 and T3. Since this is one of the few exploratory studies reported, a Bonferroni correction was not

used to reduce the risk of Type I errors.

At T1, there was a significant weak positive correlation between the proportion of nonwords correctly imitated and Positional Segment Frequency ($r(91) = .18, p < .05$) for the TD children while the correlation between proportion correct and Biphone Frequency was not significant ($r(91) = .08, p = .23$). There was no significant correlation between the proportion of nonwords imitated correctly and Positional Segment Frequency ($r(17) = -.10, p = .35$) or Biphone Frequency ($r(17) = -.39, p = .06$) at T1 for the LTs.

At T3, there was a significant relationship between proportion of nonwords correctly imitated and both the phonotactic probability measures in the TD group. The strength of the relationship ranged from weak positive between proportion correct and PSF average ($r(105) = .34, p < .001$) to a moderate positive association between proportion correct and BF average ($r(105) = .51, p < .001$). There was also a significant correlation observed between proportion of nonwords correctly imitated and both measures of phonotactic probability in the LT group. There was a moderate positive relationship between proportion of nonwords correctly imitated and PSF average ($r(33) = .63, p < .001$) and a weak positive relationship between proportion correct and BF average ($r(33) = .39, p < .001$).

The scatterplots displaying the relationship between proportion of nonwords correctly imitated on the TENR-R and the phonotactic probability measures, in TD and LTs, are presented in figures 1, 2, 3 and 4 for T1 and in figures 5, 6, 7 and 8 for T3. The figures clearly show a stronger relationship between proportion correct and BF average for TD children and proportion correct and PSF average for LTs at T3. At T1, there is an association between proportion correct and PSF average alone in the TD participants (Figure 1). LTs at T1 do not show any relationship between the proportion of nonwords correctly imitated and the phonotactic probability measures (Figures 2 & 4).

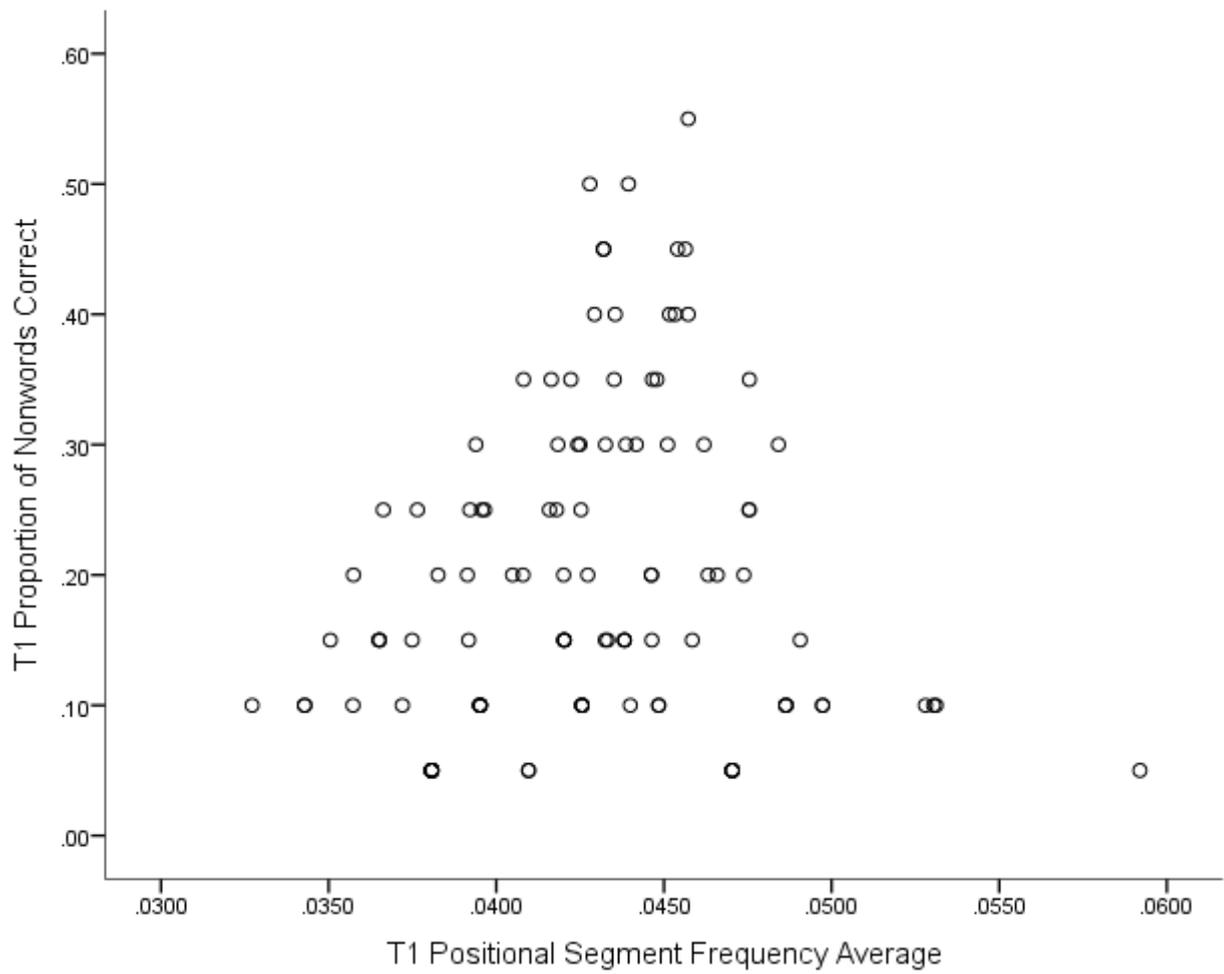


Figure 1: Scatterplot of the Relationship between Proportion of Nonwords Correctly Imitated at T1 by Positional Segment Frequency Average for Typically Developing Children ($r(91) = .18, p < .05$)

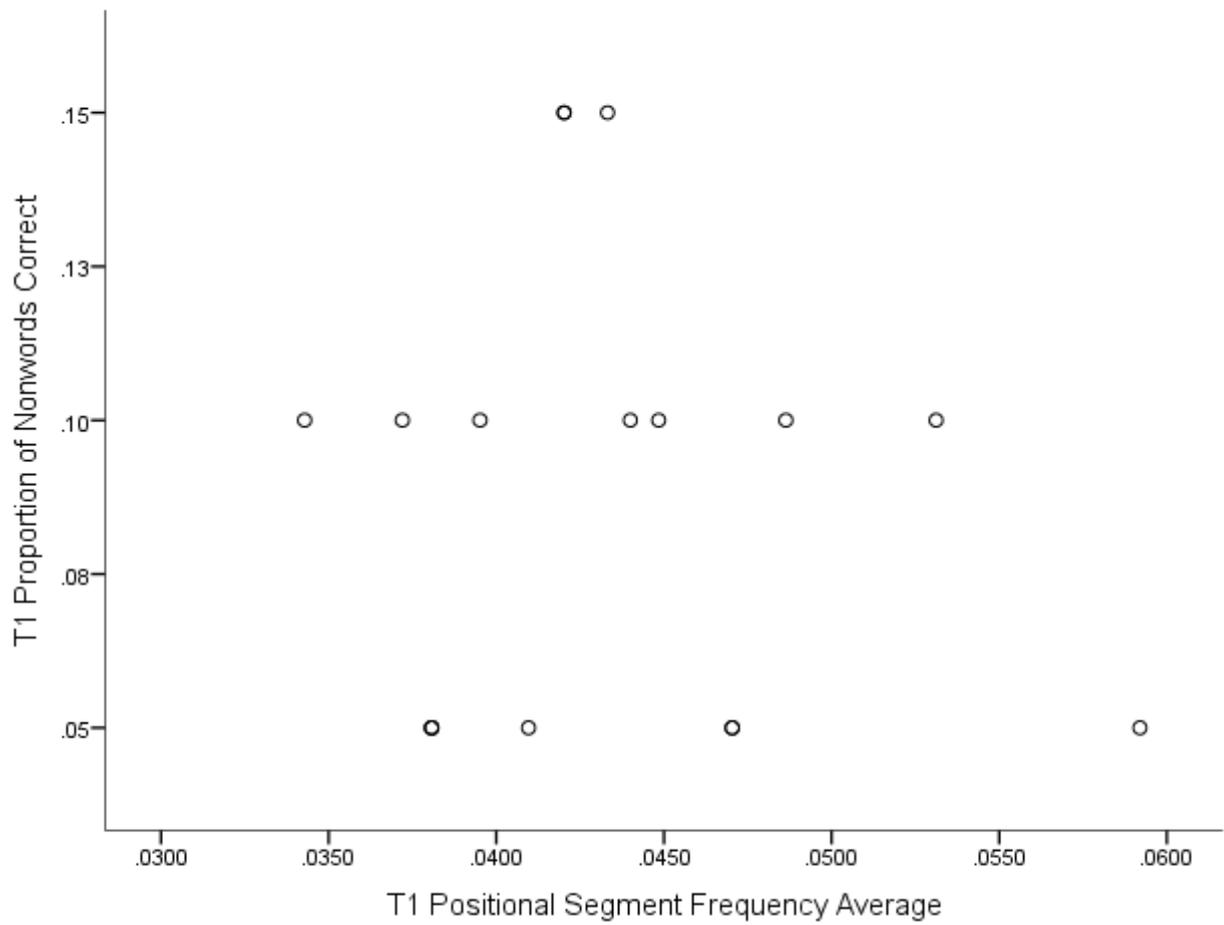


Figure 2: Scatterplot of the Relationship between Proportion of Nonwords Correctly Imitated at T1 by Positional Segment Frequency Average for Late Talkers ($r(17) = -.10, p = .35$)

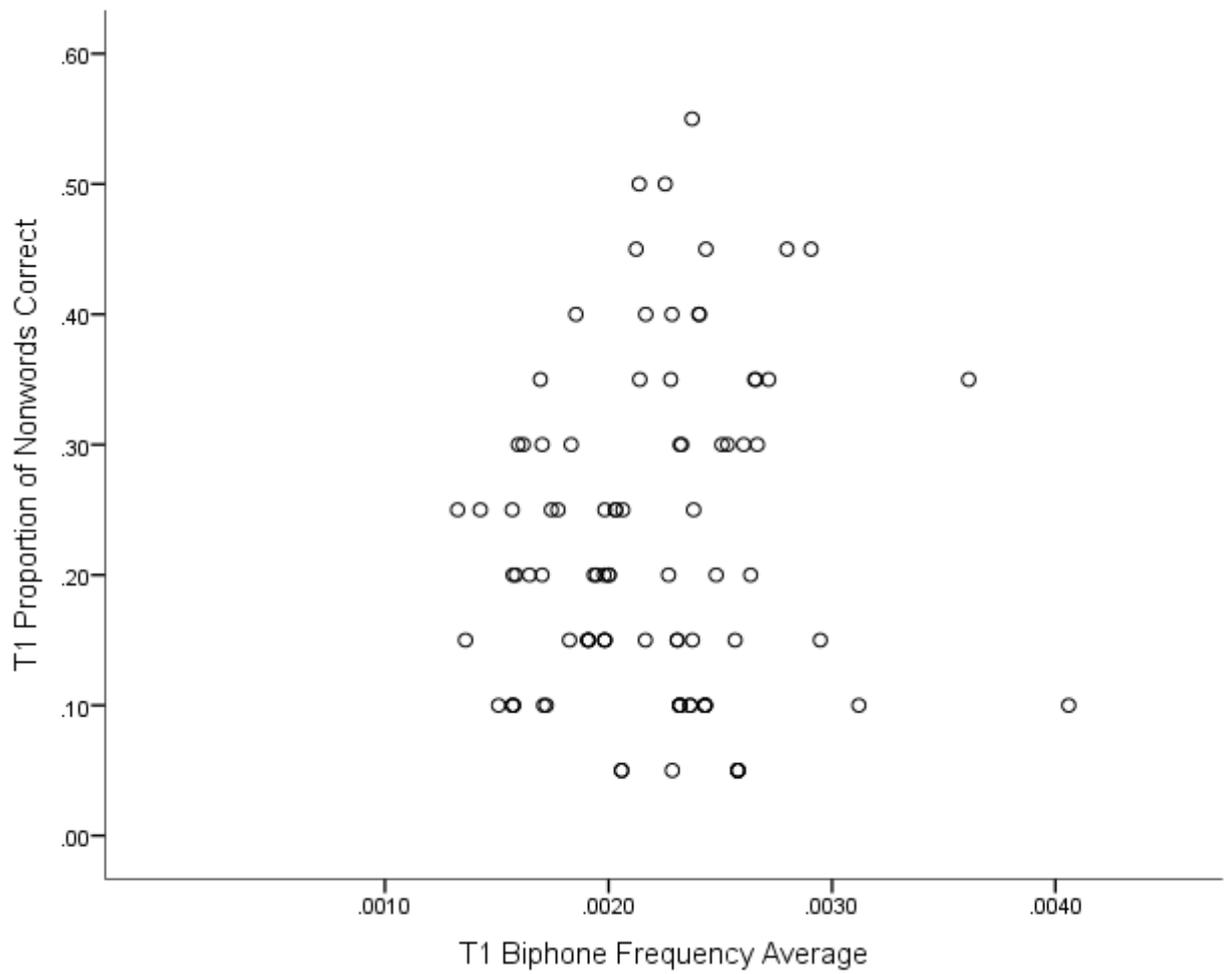


Figure 3: Scatterplot of the Relationship between Proportion of Nonwords Correctly Imitated at T1 by Biphone Frequency Average for Typically Developing Children ($r(91) = .08, p = .23$)

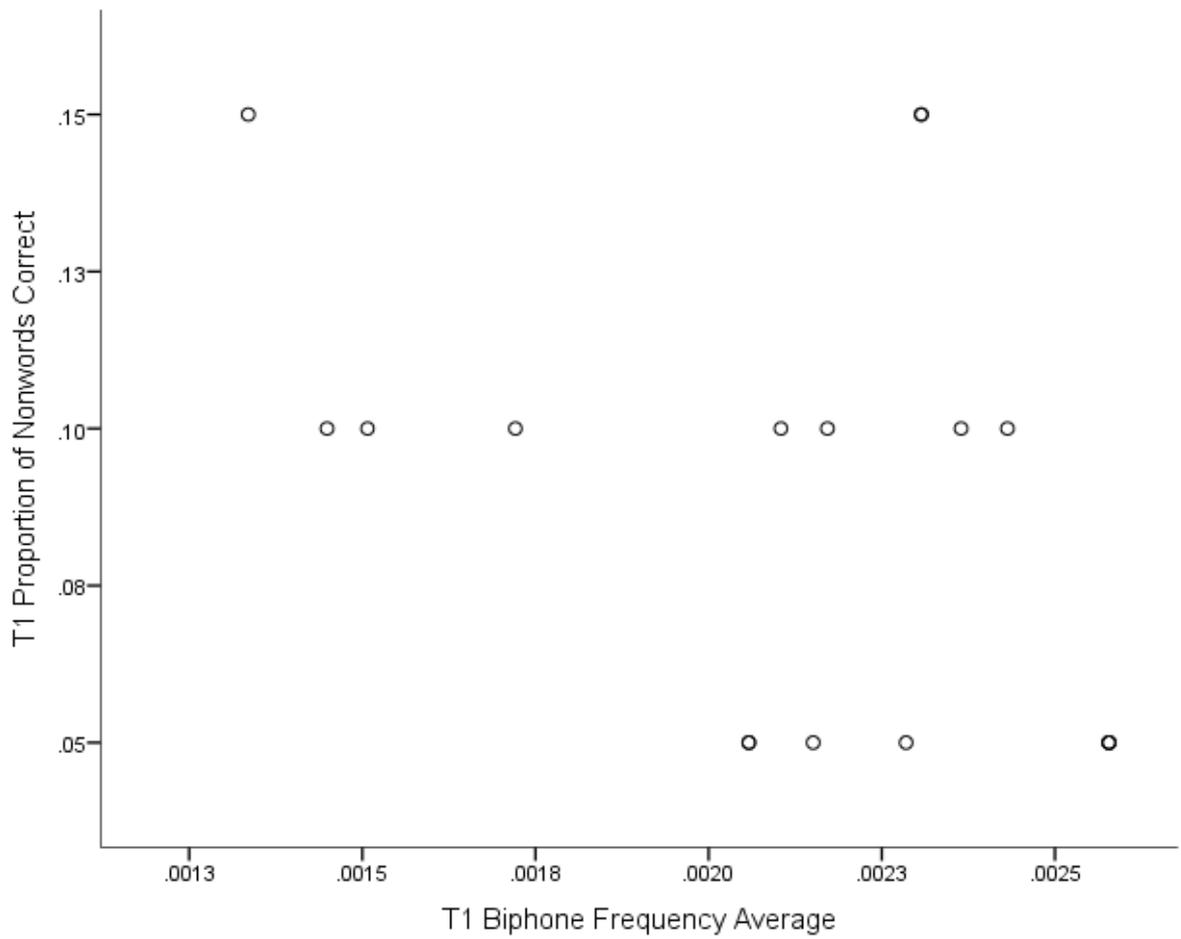


Figure 4: Scatterplot of the Relationship between Proportion of Nonwords Correctly Imitated at T1 by Biphone Frequency Average for Late Talkers ($r(17) = -.39, p = .06$)

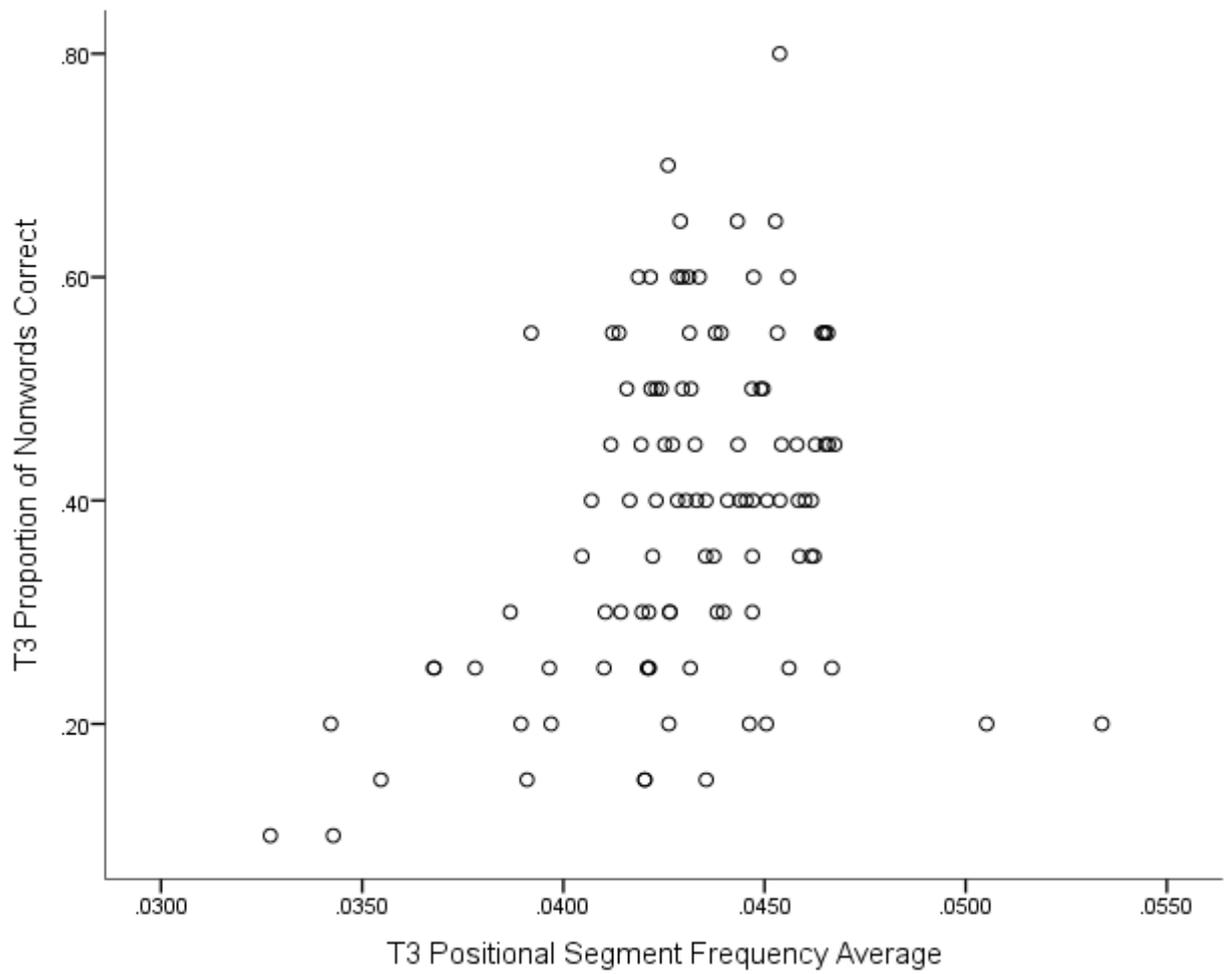


Figure 5: Scatterplot of the Relationship between Proportion of Nonwords Correctly Imitated at T3 by Positional Segment Frequency Average for Typically Developing Children ($r(105) = .34, p < .001$)

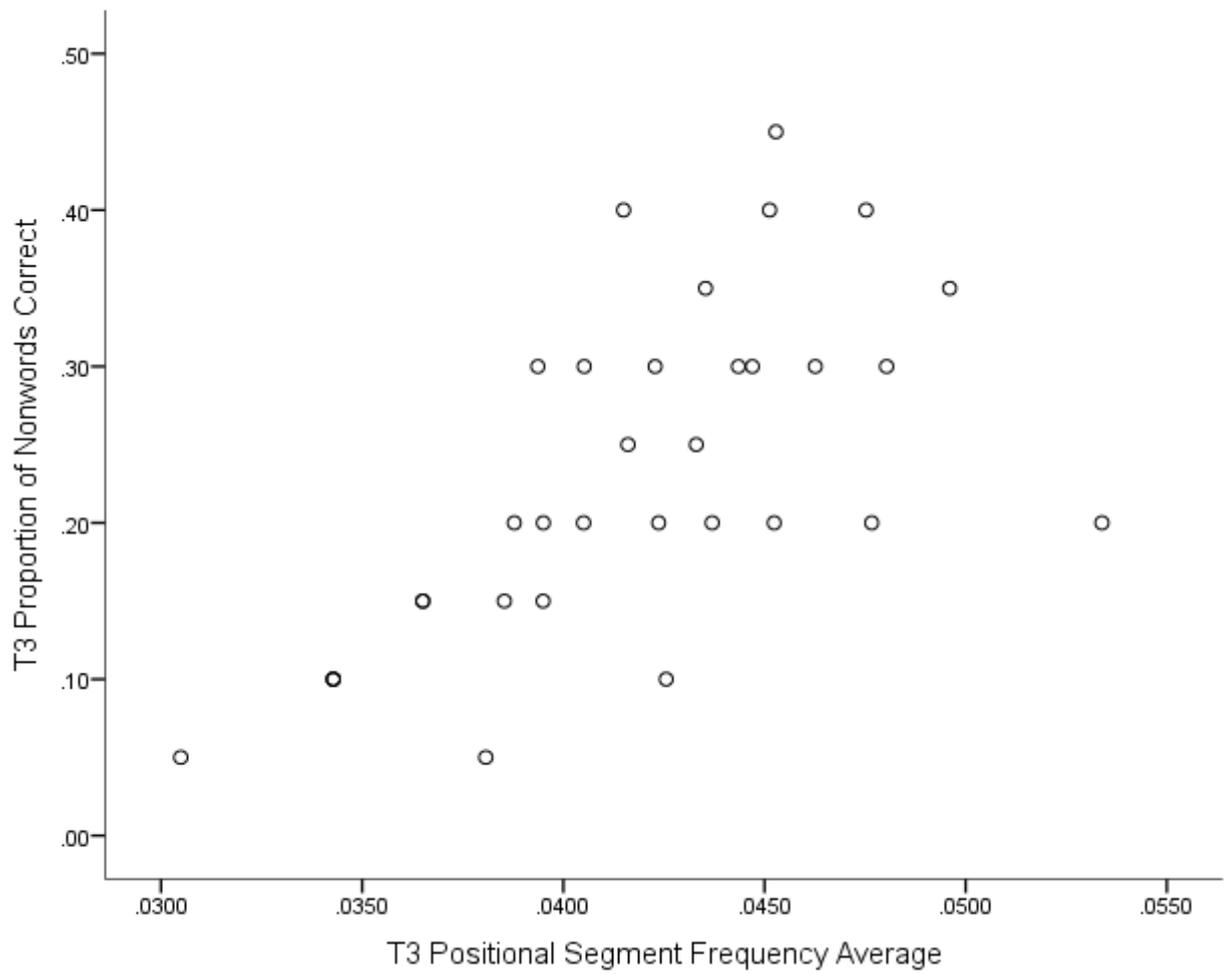


Figure 6: Scatterplot of the Relationship between Proportion of Nonwords Correctly Imitated at T3 by Positional Segment Frequency Average for Late Talkers ($r(33) = .63, p < .001$)

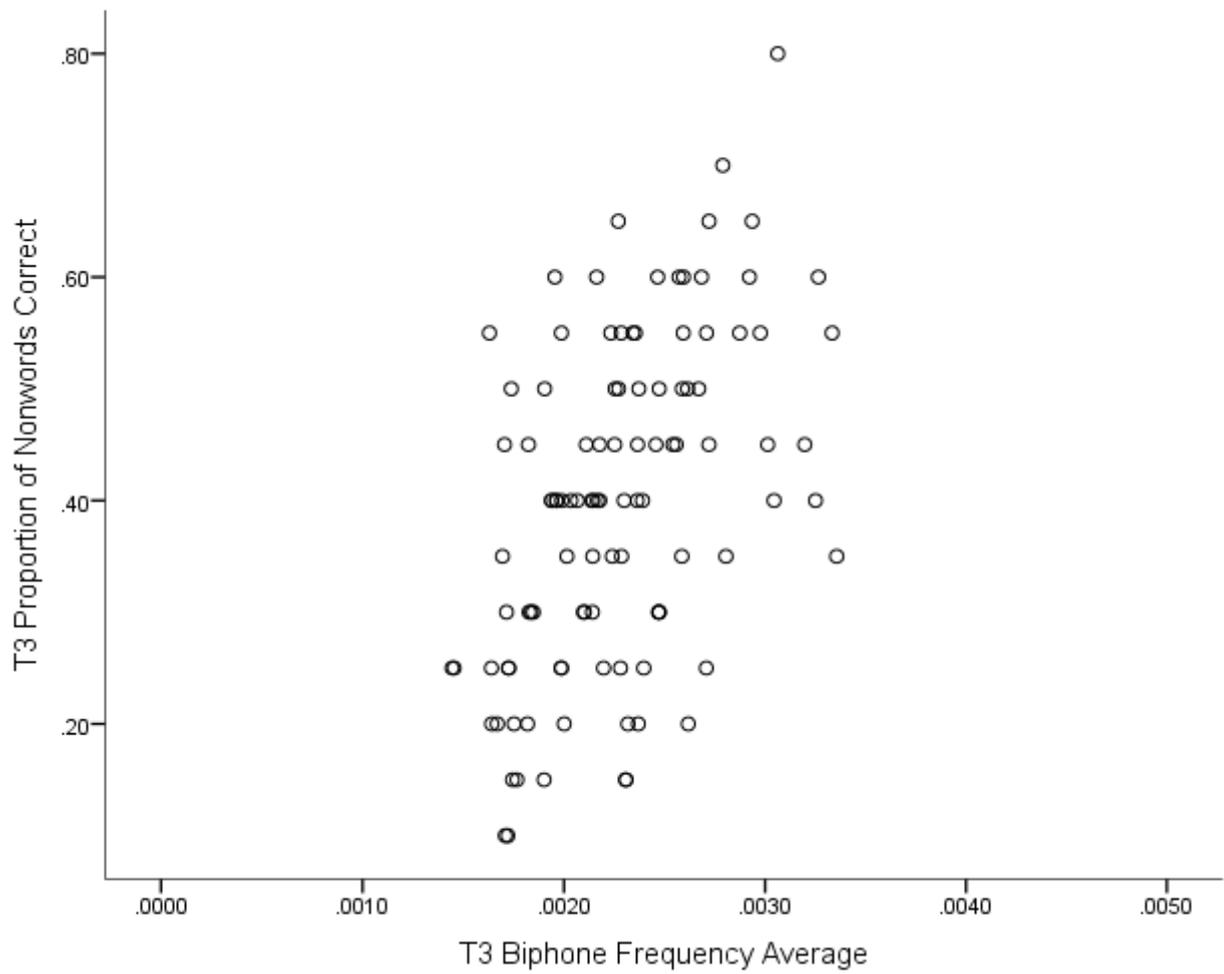


Figure 7: Scatterplot of the Relationship between Proportion of Nonwords Correctly Imitated at T3 by Biphone Frequency Average for Typically Developing Children ($r(105)= .51$, $p < .001$)

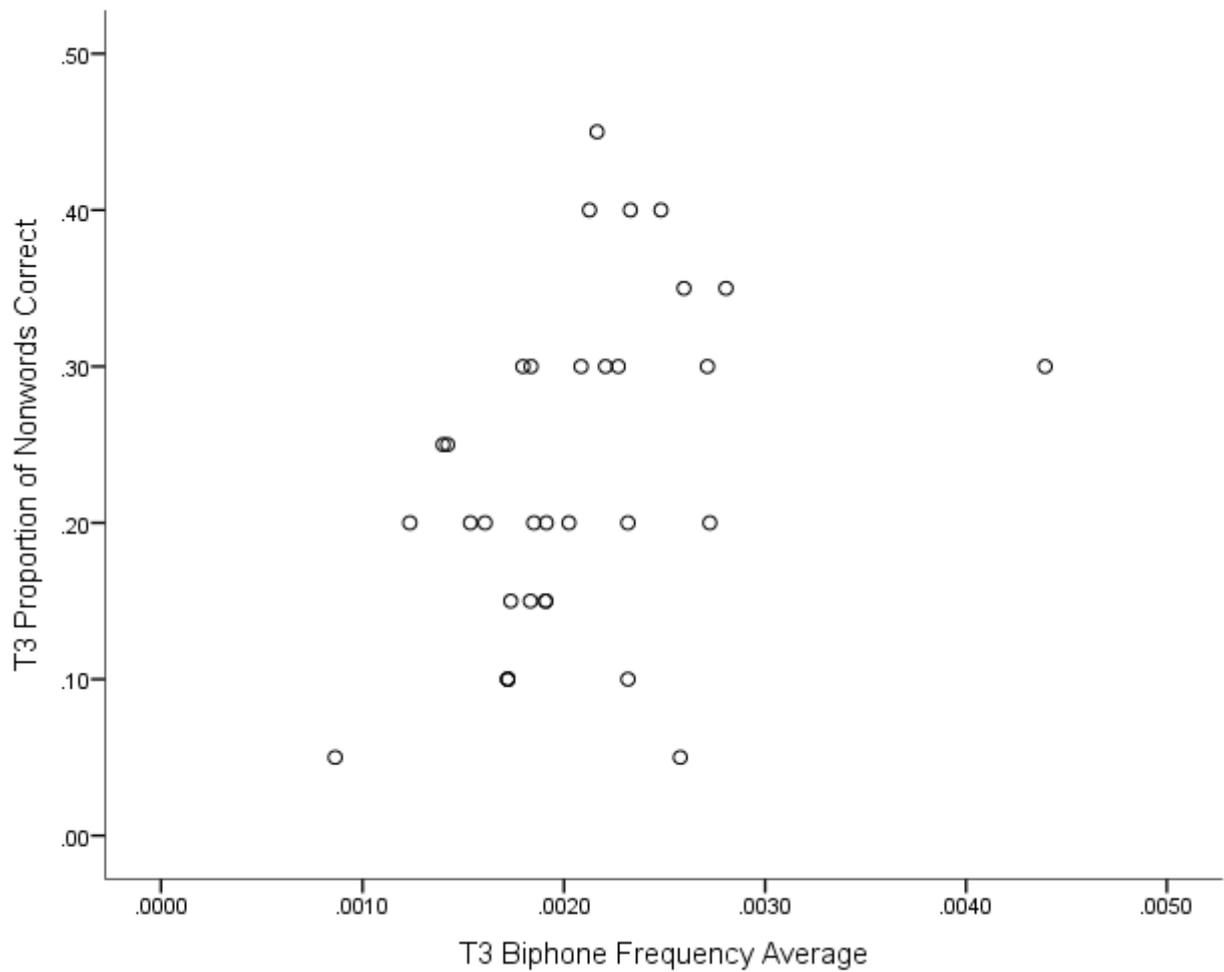


Figure 8: Scatterplot of the Relationship between Proportion of Nonwords Correctly Imitated at T3 by Biphone Frequency Average for Late Talkers ($r(33)= .39, p < .001$)

Research question 3

Which measure of phonotactic probability at T3 is most closely associated with the T3 TENR-R scores?

The relationship between the two measures of phonotactic probability and nonword repetition accuracy were further explored using regression analysis to establish the measure of phonotactic probability that maximally accounts for the accuracy of response on the TENR-R. We had predicted that Biphone Frequency would account for a significant

proportion of the variance for both groups at T3.

3.3 Regression Analysis. Linear regression analysis was carried out with PSF average and BF average as the predictor variables and proportion of nonwords correctly imitated as the outcome variable (Tables 6 & 7). The analysis was carried out using the ‘Forward’ command in SPSS to determine the amount of variance in the outcome variable that is accounted for by each of the two phonotactic probability measures.

Table 6: *Regression Analysis Summary for Phonotactic Probability Variables Predicting Percentage of Nonwords Repeated Accurately in Typically Developing Children*

Model	B	Std. Error	Beta	t	Sig.	95% Confidence Interval for B	
						Lower Bound	Upper Bound
(Constant)	.01	.07		.18	.86	-.12	.14
T3 BF average	169.57	28.80	.50	5.89	.00	112.46	226.68

Note. BF average= Biphone Frequency average, T3= Time 3.

The model was significant with BF average accounting for 25.2% of the unique variance in the proportion of nonwords correctly imitated ($F(1,103)=34.68$, $p < 0.001$) while Positional Segment Frequency was not a predictor ($B = 0.13$, $t = 1.30$, $p = 0.20$).

Table 7: *Regression Analysis Summary for Phonotactic Probability Variables Predicting Percentage of Nonwords Repeated Accurately in Late Talkers at T3*

Model	B	Std. Error	Beta	t	Sig.	95% Confidence Interval for B	
						Lower Bound	Upper Bound
(Constant)	-.33	.13		-2.56	.02	-.60	-.07
T3 PSF average	13.48	3.09	.62	4.36	.00	7.17	19.79

Note. T3= Time 3, PSF average= Positional Segment Frequency average.

For LTs, Positional Segment Frequency accounted for 38.8% of the unique variance in the proportion of nonwords correctly imitated ($F(1, 30)= 19.02, p < .001$) while Biphone Frequency was not a significant predictor ($B= 0.11, t= 0.62, p= .54$).

3.4 Partial Correlations. Partial correlations were carried out to understand the independent influence of each phonotactic probability measure when one of the two measures was controlled. The findings from partial correlations with proportion of nonwords correctly imitated when one of the two measures of phonotactic probability is controlled are displayed in Table 8. First Biphone Frequency was controlled for both groups of participants followed by Positional Segment Frequency.

Table 8: *Partial Correlation among Proportion of Nonwords Correctly Imitated and Phonotactic Probability Measures when one of the two measures is controlled across both groups at T3*

Control Variables		T3 proportion correct	
		Typically Developing	Late Talkers
T3 BF average	T3 PSF average	.13	.55**
T3 PSF average	T3 BF average	.41**	.09

** . Correlation is significant at the 0.01 level (1-tailed). Note. T3= Time 3, PSF average= Positional Segment Frequency average, BF average= Biphone Frequency average, Proportion correct= Proportion of nonwords correctly imitated.

The correlation between proportion of nonwords correctly imitated and PSF average for TD children was $r(102)=.13$ when the influence of BF average was controlled. However, the correlation between the two variables was not significant ($p > .05$). The correlation between proportion of nonwords correctly imitated and PSF average for LTs was significant with $r(30)= .55$ and $p < .001$ when the influence of BF average was controlled. There was a moderate positive relationship between the two variables. There was also a moderate positive correlation ($r(102)= .41, p < .001$) between proportion of nonwords correctly imitated and BF average when the influence of PSF average was controlled. However, the correlation between proportion of nonwords correctly imitated and BF average for LTs was not significant with $r(30)= .09$ and $p > .05$ when the influence of PSF average was controlled. These findings

suggest that at T3, TD children rely on Biphone Frequency cues in order to accurately imitate the nonwords while LTs rely on Positional Segment Frequency cues while imitating the nonwords in the TENR-R.

Given these findings, correlation and regression analyses were carried out for the entire sample of participants at T3, irrespective of language status, to further explore the above findings. The descriptive statistics for the variables proportion correct, PSF average and BF average for all participants at T3 are presented in Table 9.

Table 9: *Descriptive Statistics for all Participants in the Study at T3*

	Mean	Std. Deviation	N
T3 Proportion correct	.36	.16	18
T3 PSF average	.0426	.0036	138
T3 BF average	.0022	.0005	138

Note. Numerals are reported to four decimal places because the values are small. T3= Time 3, Proportion Correct= Proportion of nonwords correctly imitated, PSF average= Positional Segment Frequency average, BF average= Biphone Frequency average.

The mean for the proportion of nonwords correctly imitated for the sample inclusive of all participants in the study was much higher than that of the LTs group at T3 ($M= 0.23$, $SD= 0.11$). As a result, the corresponding PSF average and BF average means were also higher. The values were very close to that of TD children at T3 ($M= 0.40$, $SD= 0.15$). Hence

we would expect the correlation and regression analysis findings to be similar to that of TD children at T3.

Results from correlation analysis between the three variables proportion correct, PSF average and BF average are presented in Table 10.

Table 10: *Bivariate Correlations among Proportion of Nonwords Correctly Imitated, Positional Segment Frequency Average and Biphone Frequency Average for the Entire Sample at T3*

	Proportion correct	T3 PSF average	T3 BF average
T3 Proportion correct		.41**	
T3 PSF average			.51
T3 BF average	.48**		

** . Correlation is significant at the 0.01 level (1-tailed). Note. T3= Time 3, Proportion Correct= Proportion of nonwords correctly imitated, PSF average= Positional Segment Frequency average, BF average= Biphone Frequency average.

The proportion of nonwords correctly imitated variable significantly correlated with both measures of phonotactic probability. There was a moderate positive relationship between proportion correct and PSF average ($r(138) = .41, p < .001$) as well as proportion correct and BF average ($r(138) = .48, p < .001$) at T3.

Visual representations of the relationships between proportion of nonwords correctly imitated with PSF average and BF average for all participants at T3 are displayed in Figures 9 and 10 respectively.

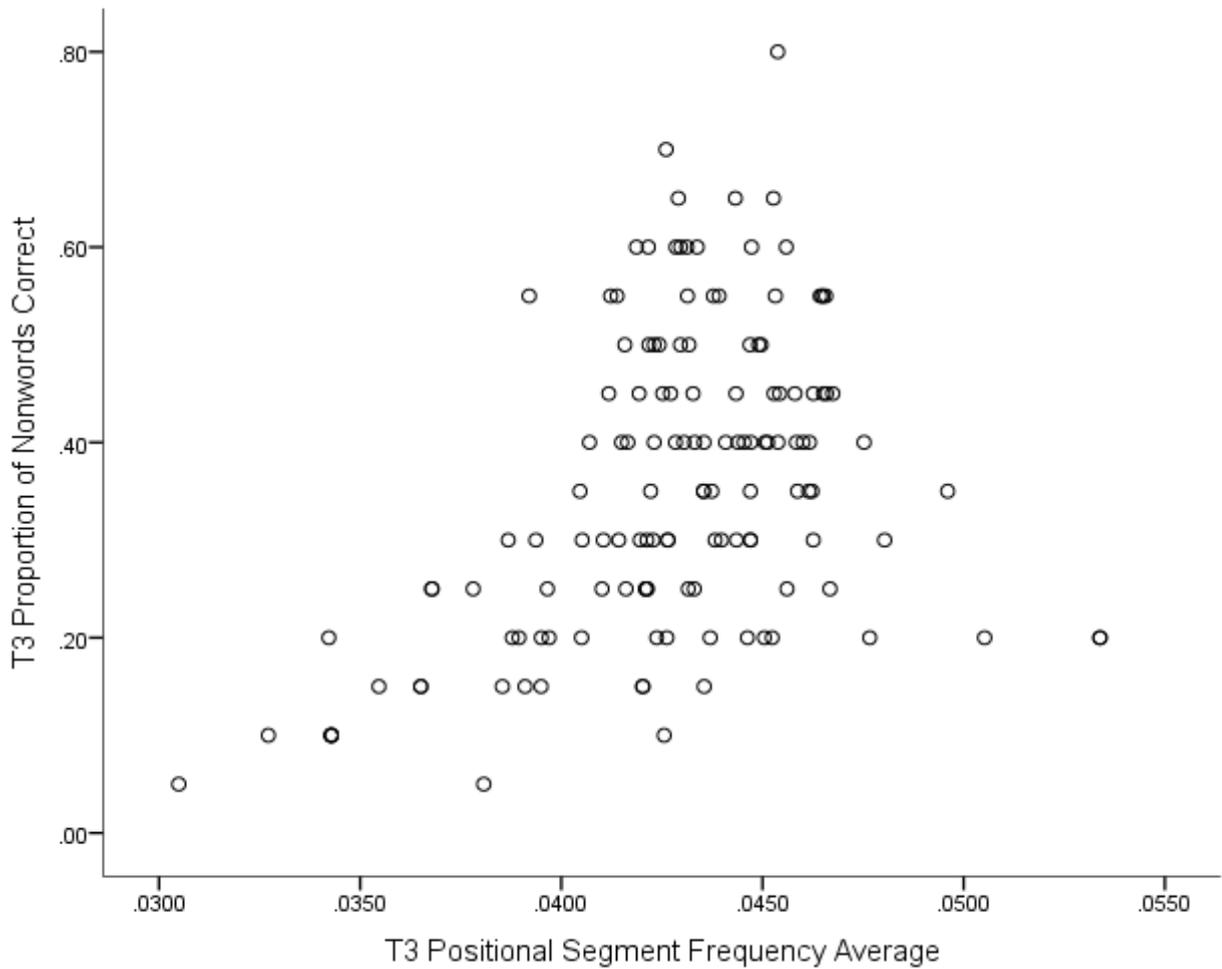


Figure 9: Scatterplot of the Relationship between Proportion of Nonwords Correctly Imitated at T3 by Positional Segment Frequency Average for all Participants ($r(138)= .41, p < .001$)

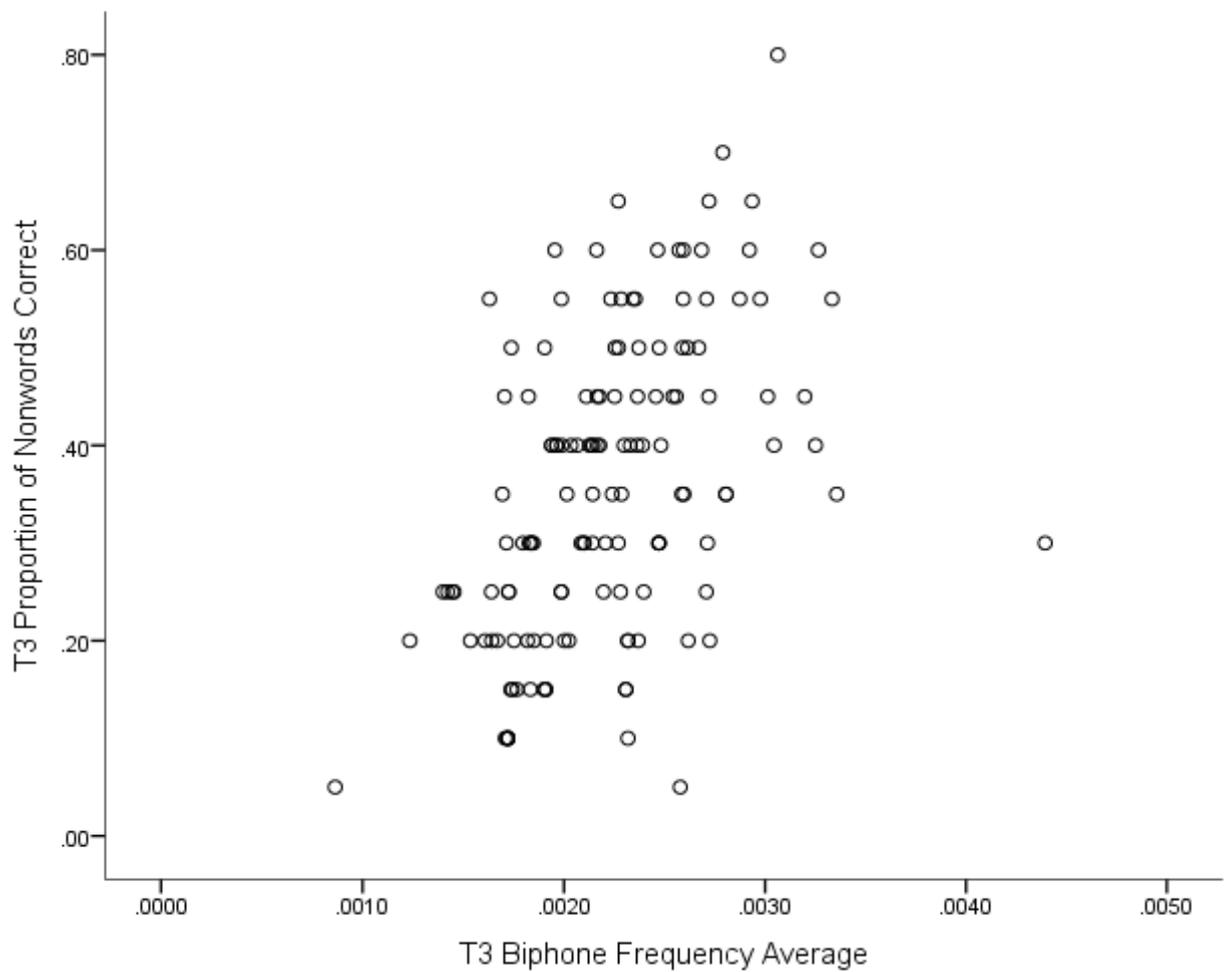


Figure 10: Scatterplot of the Relationship between Proportion of Nonwords Correctly Imitated at T3 by Biphone Frequency Average for all Participants ($r(138) = .48, p < .001$)

Regression analysis and partial correlation findings for the variables proportion correct, PSF average and BF average are summarised in Tables 11 and 12.

Table 11: *Regression Analysis Summary for Biphone Frequency and Positional Segment Frequency Predicting the Proportion of Nonwords Imitated Correctly for all Participants in the Study at T3*

	Model	Beta	t	Sig.	95% Confidence Interval for B	
					Lower Bound	Upper Bound
1	(Constant)	.018	.054	.327	.744	-.090 .125
	T3 BF average	152.863	23.839	.482	6.412	.000 105.721 200.005
2	(Constant)	-.317	.137	-2.311	.022	-.588 -.046
	T3 BF average	116.676	27.040	.368	4.315	.000 63.198 170.153
	T3 PSF average	9.703	3.666	.226	2.647	.009 2.453 16.954

Note. T3= Time 3, Proportion Correct= Proportion of nonwords correctly imitated, PSF average= Positional Segment Frequency average, BF average= Biphone Frequency average.

The model was significant with BF average accounting for 23.2% of the unique variance in the proportion of nonwords correctly imitated ($F(1, 136)= 41.12, p < .001$) while

Positional Segment Frequency accounted for 3.8% of the unique variance in the proportion correct ($F(1, 135)= 7.01, p < .001$).

Table 12: *Partial Correlation among Proportion of Nonwords Correctly Imitated and Phonotactic Probability Measures When One of the Two Measures is Controlled for all Participants at T3*

Control Variables		T3 BF average	T3 PSF average
T3 PSF average	T3 Proportion correct	.348**	-
T3 BF average	T3 Proportion correct	-	.222**

** . Correlation is significant at the 0.01 level (1-tailed). Note. PSF average= Positional Segment Frequency average, BF average= Biphone Frequency average, Proportion correct= Proportion of nonwords correctly imitated, T3= Time 3.

The correlation between proportion of nonwords correctly imitated and BF average was significant at $r(135)=.35, p < .001$ when the influence of PSF average was controlled. The correlation between proportion of nonwords correctly imitated and PSF average was also significant at $r(135)= .22, p < .001$ when the influence of BF average was controlled.

The findings computed for the whole sample were similar to that of TD children at T3 with a major portion of unique variance accounted for by Biphone Frequency. This suggests that higher T3 TENR-R scores are achieved, as the nonword become more wordlike (which is what a high biphone frequency indicates). Similarly, at T3 higher TENR-R scores are

achieved when the nonwords contain segments that are frequently occurring in the ambient language.

Research question 4

Is there a significant difference in mean values of the variables proportion correct, PSF average and BF average between TD children and LTs at T3?

An Independent samples t-test was carried out to examine if there was a significant difference between TD children and LTs across any of the three variables of interest (proportion correct, PSF average and BF average). The test revealed that the two groups of participants significantly differed in terms of proportion correct ($t= 7.01$, $df= 74.56$, $p< .001$) and BF average ($t= 2.14$, $df= 136$, $p< .05$). TD participants had significantly higher proportions of nonword repetition accuracy ($M= 0.40$, $SD= 0.15$) than LTs ($M= 0.23$, $SD= 0.11$). The TD participants also had BF average scores ($M= 0.0023$, $SD= 0.0004$) that were significantly higher than the BF average scores of the LTs ($M= 0.0021$, $SD= 0.0006$). The two groups of participants did not show any significant group differences in terms of the PSF average ($t= 1.43$, $df= 40.00$, $p> .05$).

Research question 5

Is there an association between vocabulary size and phoneme repetition accuracy on the TENR-R?

Stokes and Klee (2009b) reported a moderate significant positive correlation between CDI scores and PPC for the TENR ($r= .60$, $p< .001$). As a replication of the Stokes and Klee's (2009b) findings a correlation analysis was carried out between CDI scores and APPC scores at T1. The correlation between CDI and APPC scores for the whole sample was found to be significant with $r(133)= .64$ and $p< .001$.

Research question 6

Are the two measures of phonotactic probability associated with nonword length on the TENR-R?

Lastly, correlation analysis was also carried out to see if there was an association between nonword length and the two phonotactic probability measures using Spearman's rank order correlation. There was no significant correlation between nonword length (in phonemes or syllables) and PSF average or BF average. The correlation among number of phonemes by PSF average was $r_s(20) = -.30$, $p = .20$ and by BF average was $r_s(20) = -.08$, $p = .73$. The correlation among number of syllables by PSF average was $r_s(20) = -.35$, $p = .13$ and by BF average was $r_s(20) = -.10$, $p = .68$.

3.5 Summary

In summary, phonotactic probability demonstrated an influence on the accuracy of nonword repetition on the TENR-R. While Positional Segment Frequency significantly correlated with nonword repetition accuracy for TD children at T1, both measures of phonotactic probability, Positional Segment Frequency and Biphone Frequency significantly correlated with the accuracy of nonword repetition across both groups of participants at T3. However, interesting group differences were observed. For the TD children, Biphone Frequency was found to be a predictor of accuracy while for the LTs group Positional Segment Frequency was found to predict accuracy on the TENR-R. Analysis of the entire sample regardless of group membership revealed findings similar to that of the TD group at T3 with Biphone Frequency accounting for a major proportion of variance in the nonword repetition accuracy while Positional Segment Frequency accounted for a very small proportion of variance. Another interesting finding is that the LTs at T3 had scores resembling the TD participants at T1. At T1, the TD TENR-R scores significantly correlated

with Positional Segment Frequency and at T3, the LT TENR-R scores significantly correlated with Positional Segment Frequency. But by T3, the TD TENR-R scores were significantly correlated with wordlikenes (Biphone Frequency).

4.0 Discussion

The main aims of the study were to investigate the age-related sensitivity to sublexical correlates during a nonword repetition task in preschoolers and to compare TD and LT children's sensitivity to sublexical characteristics of nonwords. The TENR-R (Stokes & Klee, 2011) was used to observe the role of phonotactic probability on nonword repetition accuracy in two groups of participants classified as TD and LTs based on their expressive vocabulary size at 27 months. An overall comparison of nonword repetition performance and sensitivity to phonotactic probability was also made in the sample as a whole at T3, when the children were 45 months old.

Although the literature is rife with studies on children's sensitivity to phonotactic probability in word and nonword processing tasks, and their association with various interdependent factors such as the neighborhood density of task stimuli, and working memory constraints, so far we are aware of only one study (Coady & Aslin, 2004) that has investigated the relative effects of Positional Segment Frequency and Biphone Frequency on nonword repetition accuracy. However, this study was cross-sectional. Several cross-sectional studies have provided valuable information about the nature of lexical representations in young children and differences have been observed in children with language difficulties. In addition to this, vocabulary size has been repeatedly linked with the phonotactic probability of words in influencing performance on various processing and production tasks (Stokes & Klee, 2009b). However, there is no longitudinal study to date that has investigated the influence and interaction of sublexical correlates, vocabulary size and age on a nonword processing task in preschoolers.

The two accounts on the nature of lexical representations in young children, dubbed the emergent and accessibility accounts have been most prevalent in the literature. These two theories assume a discontinuity in the development of sublexical correlates with the lexical

representations being either holistic or fine-grained to the level of biphones at first, and undergoing refinement and restructuring during the later course of development. However Coady and Aslin (2004) highlighted a pattern of continuity in the development of Positional Segment and Biphone Frequency measures with both measures demonstrating similar trends in development. Taking the two prevalent accounts of early lexical representations into consideration, this study was conducted with the hypothesis that there is a discontinuity in the development of sublexical correlates, with young children being more sensitive to Biphone Frequency. This sensitivity was further hypothesized to become more fine-grained (segmental) with age and increase in vocabulary size. Considering the long standing debate on how early lexical representations are stored, retrieved and produced, and their variability in children with language delay and clinically diagnosed disorders of phonology, a study of this nature can provide useful insights on the nature of children's early stored lexical representations and how they develop over time.

The study addressed the following questions. Do young children rely on cues from their knowledge of phonotactic probability for nonword repetition accuracy? If so, which measure of phonotactic probability are they most sensitive to during the two time periods? Is the nature of sensitivity during the two time periods identical or different? Does vocabulary size influence the access to these cues? Do LTs demonstrate a different pattern of sensitivity than that of their TD peers?

From a statistical viewpoint, the following research questions were addressed:

1. What is the impact of phonotactic probability on the accuracy of response on the TENR-R?
2. To what extent does vocabulary size play a role on the accuracy of responses and how does the role of sublexical properties change with age?

3. Which measure of phonotactic probability (positional segment frequency or biphone frequency) at T3 is most closely associated with the T3 TENR-R scores?
4. Is there a significant difference in the mean values of the study variables between TD children and LTs at T3?
5. Is there an association between vocabulary size and phoneme repetition accuracy on the TENR-R?
6. Are the two measures of phonotactic probability associated with nonword length on the TENR-R?

The specific research questions along with the findings from the study will be discussed in the sections below.

4.1 Nonword repetition accuracy

Nonword repetition tests have been found to be highly sensitive in identifying difficulties in linguistic processing. By controlling for variables such as vocabulary size, experience, culture and age, tasks such as nonword repetition are thought to provide a non-biased means for assessing the processing and production abilities of participants (Campbell et al., 1997). Likewise, in the present study, the TENR-R provided valuable information about the nature of sensitivity to phonotactic cues during the course of development. The test also highlighted variability in sensitivity to phonotactic cues between TD and LTs. Overall, the nonword repetition accuracy of participants improved from T1 to T3. In terms of group comparisons, the raw scores for nonword repetition accuracy and their corresponding phonotactic probability measures improved over the period of 18 months across both groups. These findings are consonant with previous cross-sectional studies that have reported a significant effect of age on nonword repetition accuracy (Coady and Aslin, 2004).

Children with language delay perform worse than their TD peers on nonword repetition tasks and LTs have been known to perform poorly on verbal recall and novel word learning tasks (Weismer & Evans, 2002). Similar findings have been observed in the present study. The TD participants at T1 and the LTs at T3 had comparable raw scores and the highest scores were achieved by the TD participants from T3. These findings suggest that phonotactic probability does have an impact on nonword repetition accuracy as a corresponding increase in phonotactic probability scores was observed with an increase in repetition accuracy.

4.2 Sensitivity to sublexical correlates

A wealth of literature has repeatedly confirmed that phonotactic probability influences nonword repetition. This study adds to existing research by identifying the measure of phonotactic probability that influences nonword repetition accuracy across the 18 month period. The analysis firstly confirmed that phonotactic probability influences nonword repetition in preschool children. Nonword repetition accuracy at T1 was found to be significantly correlated with Positional Segment Frequency in the TD group with no measure showing any correlation with the performance of LTs at T1. The latter finding suggests that LTs are not sensitive to phonotactic probability at 27 months. The nonword repetition accuracy at T3 was found to be significantly correlated with Positional Segment and Biphone Frequency overall and also when grouped into TD and LTs. Our hypothesis was that younger children undertake a more holistic approach to processing and production and hence would be sensitive to Biphone frequency. While this hypothesis is confirmed in TD children at T3, where 25.2% of the variance in nonword repetition accuracy was accounted for by Biphone Frequency alone, the nonword repetition accuracy in LTs at T3 was accounted for by Positional Segment Frequency alone (38.8%). These findings were further confirmed by

partial correlations which showed that older TD children may rely on cues from Biphone Frequency and LTs may rely on Positional Segment Frequency cues while repeating nonwords. Interestingly, the sensitivity of LTs at T3 was similar to that of younger TD participants from T1 with both groups relying on cues from Positional Segment Frequency for nonword repetition accuracy. Further, partial correlations for all participants from T3 taken together, both measures of phonotactic probability were found to account for the variance in repetition accuracy on the TENR-R, with PSF average accounting for 12.11% of the variance and BF average accounting for 4.92 % of the variance in nonword repetition accuracy.

4.2.1 Sensitivity to Positional Segment Frequency. Recent studies on children just beginning to talk have found that children as young as two years of age are sensitive to Positional Segment Frequency cues in nonwords. Zamuner (2013) found that two year olds are sensitive to Positional Segment Frequency contrasts presented in both high and low phonotactic probability environments, with greater accuracy for high probability nonwords. Similarly, Coady and Aslin (2004) found both younger and older children between 2.5 and 3.5 years of age to be sensitive to Positional Segment Frequency cues. However, it was observed that the nature of this sensitivity improved with age. The younger children were sensitive to Positional Segment Frequency only when it was maintained across the entire nonword and not when limited to a particular position in the nonword. This outcome suggests that children possess sensitivity to fine-grained segmental frequency knowledge even at a very young age. However, this sensitivity is not adult-like and is susceptible to development with age and experience. Although it has been observed that young children are unable to distinguish minimal phonetic contrasts presented in novel words (Stager & Werker, 1997), they are able to distinguish both members of the pairs when presented as part of a

conversation (Barton, 1976). This is suggestive of a limitation in the creation of novel lexical representations due to cognitive or pre-linguistic deficits or a combination of these rather than an underspecified representational system (Coady & Aslin, 2004). Furthermore, when three and four year olds were presented with a set of real and nonword pairs that differed by two phonetic features in the same position and another set that differed by two phonetic features in different positions, both groups of children were found to confuse a real word with the nonword pair when the phonetic features were manipulated in the same position. If children stored words holistically, then the performance for both sets of stimuli should have been the same (Gerken, Murphy, & Aslin, 1995). These findings are supportive of the presence of sensitivity to fine-grained cues related to Positional Segment Frequency similar to the outcome seen in TD children at T1 and LTs at T3. As previously described, the vocabulary of young children is composed of a relatively dense neighborhood (Coady & Aslin, 2003) with approximately 84% of words in the lexicon having at least one near phonological neighbour, as observed in children between 10 and 21 months of age by Dollaghan (1994). In such a case, the children would be required to use fine-grained phonetic detail in differentiating and assigning meaning to the words in the lexicon. Hence it could be assumed that children build their lexicon based on their knowledge of phonological neighborhood density, wordlikeness and phonotactic probability related information. It must be noted that the TD children at T1 were between two to three years of age and it has been suggested that this being approximately one year after the vocabulary spurt, the children would have developed sufficient knowledge of phonotactic probability (Coady & Aslin, 2004).

4.2.2 Sensitivity to Biphone Frequency. At T3, the children have undergone an additional year and a half of vocabulary growth and have a larger lexicon. Rispens, Baker and Duinmeijer (2015) have recently reported that vocabulary growth promotes an ease of access

to stored lexical representations, providing greater flexibility to manipulate and combine different segments and sequences. The current results at T3 corroborate the findings from Coady and Aslin's (2004) study where only the older children aged 3.5 years were found to be sensitive to Biphone Frequency cues. Even in the current study, only the TD participants from T3 demonstrated sensitivity to Biphone Frequency. Biphone Frequency has been found to strongly correlate with wordlikeness (Frisch, 2001, as cited in Edwards et al., 2004). Nonwords with high Biphone Frequency are rated to be more wordlike by adults. The TD participants at T3 have a larger lexicon and greater experience in language use. This advantage has been reported to improve nonword recall (Gathercole et al., 1999). Given the known association between phonotactic probability and expressive vocabulary size, and the role of both lexical and phonological knowledge in nonword repetition, the participants at T3 could be relying on their pre-existing lexical knowledge in addition to phonological knowledge to gain better repetition accuracy. Though the present study identified Biphone Frequency to be a predictor of nonword repetition accuracy at T3, both measures of phonotactic probability showed significant correlations with nonword repetition accuracy. This is suggestive of a parallel course of development of both phonotactic probability measures.

4.3 Developmental trends in phonotactic probability- Continuity versus discontinuity

Continuity and discontinuity are two terms that have been used in the literature to describe the developmental trends in Positional Segment Frequency and Biphone Frequency measures. The view on continuity suggests that the two measures of phonotactic probability develop in parallel. Young children possess both fine-grained and holistic knowledge related to their stored lexicon. These skills although present are not adult-like, as phonological development is known to continue until approximately nine years of age (Templin, 1953).

The stand on discontinuity suggests that young children initially possess holistic lexical representations. With an increase in productive vocabulary, these representations are restructured by incorporating fine-grained segmental information into the lexicon (Walley, 1993). The current study draws more towards the continuity standpoint since the TD participants at T1 and the LTs at T3 demonstrated sensitivity to Positional Segment Frequency thereby providing evidence of the existence of fine-grained segmental knowledge even at a very young age. In addition to this, the TD children at T3 relied more on cues from Biphone Frequency while the overall performance of all children at T3 indicated sensitivities to both measures of phonotactic probability. The absence of sensitivity to Biphone Frequency cues for all participants at T1 and LTs at T3 could be attributed to several factors such as a possible floor effect on the TENR-R since approximately 52% of the participants at T1 were unable to complete the task while approximately 54% of the LTs at T3 scored below 25% on the TENR-R. Studies have reported that as the nonword length increases, there is a greater demand placed on the child's short term memory due to an increase in phonetic complexity. Phonotactic probability has been found to be independent of phonetic complexity (Goldrick & Larson, 2008). As the current results have also shown that both measures of phonotactic probability did not correlate with nonword length on the TENR-R, the lack of task completion and the poor scores could be related to factors such as short-term memory and articulatory complexity. There was a greater influence of nonword length on the performance of participants at T1, while approximately 98% of TD children at T3 were able to complete the task. Other factors such as smaller vocabulary size and limitations in cognitive and pre-linguistic skills also need to be taken into consideration before arriving at a conclusion for these groups of participants.

4.4 Performance variability in Late Talkers

With growing awareness about the different developmental language disorders and with greater focus being placed on early intervention, LTs have received more attention in recent years. Researchers have dedicated their knowledge and resources to understanding the mechanisms operating in LTs. This group has been of interest because of the continuing difficulties faced by them even after having outgrown the vocabulary delay. The difficulties observed in some children with clinically diagnosed language disorders such as SLI can be traced back to their early years of development. Differences in babbling patterns, expressive language, phonological processing, fast mapping, pragmatic skills and academic performance have been reported (MacRoy-Higgins et al., 2013; Paul, 1996; Rescorla, 2002, 2009; Weismer & Evans, 2002). While some authors believe LTs to have a language delay (e.g. Paul & Jennings, 1992), others have reported deviant linguistic characteristics (e.g. Williams & Elbert, 2003). Therefore, identification of key elements that are deviant or lacking in LTs could help in the early identification and management of the children who may not outgrow the delay and are therefore at risk of developing language disorders such as SLI. In order to address these issues and gain some additional insights regarding the nature of lexical representations in LTs, the performance on TENR-R was examined with the participants grouped into TD and LTs.

4.4.1 Nonword repetition performance in Late Talkers. The findings of the present study highlighted the growing need to understand the factors operating in LTs. Although LTs from T1 showed an improvement in repetition accuracy during the 18 month period, they continued to score lower than their TD peers at T3, with their performance being more analogous to the younger TD group. These findings are similar to several studies that have found that LTs exhibit difficulties in nonword repetition and other tasks of phonology

requiring the creation of novel lexical representations and fast mapping (e.g. Weismer & Evans, 2002; MacRoy-Higgins et al., 2013). Even at T3, the LTs can be seen to have continuing difficulties in nonword repetition. These results are consistent with previous studies that have reported that although LTs appear to catch up with their peers, as evidenced by their performance on standardised tests of language, they continue to face difficulties in certain areas of language. In terms of the nature of phonological representations in LTs being delayed or deviant, the current study is suggestive of LTs demonstrating a phonological delay with their performance resembling that of younger TD peers thereby providing support to the proponents of delayed phonological characteristics in LTs such as Paul and Jennings (1992) and Thal et al., (1995).

4.4.2 Late Talkers' sensitivity to phonotactic probability. The differential performance on the nonword repetition task by LTs was further explored by observing the nature of sensitivity to phonotactic probability. At T1, the LTs did not show sensitivity to any measure of phonotactic probability. This lack of sensitivity could be attributed to possible floor effects on the TENR-R as all the LTs at T1 scored below 25% of the total score on the TENR-R and nearly 75% of LTs were unable to complete the nonword repetition task at T1. With an improvement in nonword repetition performance in LTs at T3, sensitivity to phonotactic probability cues emerged. Both measures of phonotactic probability were found to correlate with nonword repetition accuracy at T3. Further statistical analysis revealed that LTs rely on Positional Segment Frequency cues while repeating the nonwords on the TENR-R. Their nature of sensitivity and raw scores at T3 resembled that of younger TD peers from T1. The T3 LTs were found to be both qualitatively (sensitivity to Positional Segment Frequency) and quantitatively (raw scores) similar in nature to the TD participants from T1. These findings further support the presence of a delay rather than a deviancy in this group of

participants. The presence of sensitivity to fine-grained regularities even in LTs who are known to have a constrained linguistic system confirms the presence of an adult-like, but immature, representational system in young children.

4.5 Phonotactic probability and short term memory

Short term memory plays an important role during tasks such as nonword repetition. The presentation of an unfamiliar word or nonword activates the sublexical route that utilises the existing knowledge of phonotactic probability (Gathercole et al., 1999). Sensitivity to phonotactic probability aids short term memory functioning by associating the stimuli presented with information stored in long term memory thereby helping in restructuring the stimuli for recall. Hence nonwords that are composed of frequently occurring segments and biphones have greater repetition accuracy than those that do not (Frisch et al., 2000; Gathercole & Baddeley, 1989). In the present study, sensitivity to phonotactic probability cues appears to influence nonword repetition. At T1, nonword repetition accuracy is higher among the TD participants who have been found to be sensitive to Positional Segment Frequency cues. The LTs do not show sensitivity to cues from phonotactic probability of the nonwords on the TENR-R and have the lowest repetition accuracy. However, the nonword repetition accuracy of LTs shows an improvement at T3 with their repetition accuracy being found to be influenced by cues from Positional Segment Frequency. The repetition accuracy of the TD participants is further enhanced at T3, with sensitivity to Biphone Frequency cues accounting for the nonword repetition accuracy in this group of participants. These results highlight two things about nonword repetition. First, nonword repetition accuracy improves if the nonword contains frequently occurring segments. Second, the repetition accuracy is higher if the nonword is more wordlike, as represented by a high Biphone Frequency. This confirms the role of phonotactic probability in aiding short term memory.

4.6 Vocabulary, age and their association with phonotactic probability

Vocabulary size and age are two factors that have been found to be associated with phonotactic probability in literature. The frequency effect, as evidenced by sensitivity to high versus low probability stimuli is influenced by age and age-related development in vocabulary (Edwards et al., 2004; Munson, 2001). This frequency effect has been observed to be high during the early stages of language development and declines as the lexicon becomes more adult-like. It has been suggested that sufficient exposure to low probability combinations through vocabulary development and language use minimizes the frequency effect in older children and adults (Munson, 2005). In terms of vocabulary, expressive vocabulary has been particularly found to influence phonotactic probability (Munson, 2001). Gathercole et al., (1999) found nonword repetition accuracy to be influenced by vocabulary size, with repetition accuracy being higher in participants with larger vocabularies. The presence of a word in expressive vocabulary indicates that it possesses a strong articulatory-motor representation. This makes the word easily accessible for manipulation and production of varied and novel combinations (Edwards et al., 2004).

Similar to the findings above, the present study also observed the influence of age and vocabulary size on phonotactic probability. When grouped into TD and LTs based on their expressive vocabulary size at T1, phonotactic probability was found to be influenced by vocabulary size. The TD children, identified as having a larger vocabulary, demonstrated higher nonword repetition accuracy and a correspondingly high PSF and BF average in production than their LT counterparts. The effects of age were also apparent with development from T1 to T3. Nonword repetition accuracy and phonotactic probability scores improved with age in both TD and LTs at T3. This age related improvement in nonword repetition performance associated with phonotactic probability has been linked to vocabulary

(Gathercole et al., 1999; Rispen et al., 2015). Hence, even in the present study, the improvement in performance could be related to an increase in vocabulary size and experience, as the children have had an additional year of learning and experience with the language before the testing phase at T3. A record of the expressive language status at T3 could provide some evidence to confirm the role of vocabulary in phonotactic probability and nonword repetition.

4.7 Reliability between the two forms of nonword repetition scoring using TENR-R

Different studies have adopted different means of scoring for nonword repetition tasks. Apart from differences in nonword stimuli developed for studies, differences in scoring have led to difficulty in generalising the findings to other studies of a similar nature. In the current study the TENR-R was used. To maintain a consistency and assess the reproducibility with the original TENR, a correlation analysis was carried out between the APPC scores and the CDI scores at T1. Similar to the word-match findings, the APPC scores too showed a significant correlation with the expressive vocabulary size of the participants at T1. These findings are consistent with Stokes and Klee's (2009b) findings.

5.0 Study Limitations

The present study, being one of the first longitudinal studies of its kind, has several limitations. Firstly, most studies on phonotactic probability develop carefully controlled nonword stimuli in high and low phonotactic probability environments to study their effects. The TENR-R had a collection of stimuli that were not uniformly grouped as being of high or low phonotactic probability. Secondly, 52% of the participants at T1 did not complete the task suggesting a possible floor effect on the TENR-R. Since the test did not contain nonwords uniformly grouped as being of high or low phonotactic probability, the comparisons between T1 and T3 must be made with caution. There have been recent reports recommending the inclusion of receptive language status in addition to expressive language while identifying LTs. The present study identified LTs only based on their expressive vocabulary size and hence there could be some heterogeneity among the LTs. The improvement in vocabulary and expressive language skills at T3 haven't been reported. The comparison of these measures with nonword repetition and phonotactic probability could also provide further information regarding the nature of phonotactic sensitivity with the rate of vocabulary and expressive language development.

6.0 Future directions and conclusion

There is a great scope for further studies to look into the nature of phonotactic sensitivity in very young children. Both longitudinal and cross-sectional studies could track the development across a wider age range to confirm the differential sensitivity to phonotactic probability in TD, LTs and children with a clinically diagnosed language disorder. Extensions to this study could further explore the relative influence of the phonotactic probability measures in high and low probability environments to gain greater clarity on the continuity versus discontinuity debate. Future studies could also explore the

nature of errors made during nonword repetition while looking for a possible influence of phonotactic probability as a compensatory mechanism in the production of the mispronounced nonword.

The present study has provided some initial insights into the relative influence of phonotactic probability during the course of development. The findings are in support of continuity in the development of the two measures of phonotactic probability rather than a discontinuity that is frequently debated. The presence of sensitivity to Positional Segment Frequency cues in TD children at T1 and LTs at T3 along with the presence of sensitivity to cues from Biphone Frequency at T3 suggests that young children possess a nascent adult-like lexical and phonological representational system that undergoes further refinement with growth in the lexicon and experience with language. The factors involved in language development are complex and inter-related. Yet the skill of communicating using language is being acquired by TD children with such ease. It is the comparisons between TD individuals and those with various types of language difficulties that give rise to the need to study the influence of the lexical and sublexical correlates in linguistic competence and performance. Studies of this nature can provide vital insights into these issues and help improve the approaches used in the management of the difficulties arising from innate and acquired limitations.

7.0 References

- Adriaans, F., & Kager, R. (2010). Adding generalization to statistical learning: The induction of phonotactics from continuous speech. *Journal of Memory and Language*, 62(3), 311-331. doi: 10.1016/j.jml.2009.11.007
- Anderson, J. D., & Byrd, C. T. (2008). Phonotactic probability effects in children who stutter. *Journal of Speech, Language, and Hearing Research*, 51(4), 851-866.
- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database [CD-ROM]. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.
- Barton, D. (1976). Phonemic discrimination and the knowledge of words in children under 3 years. *Papers and Reports on Child Language Development*, 11, 61-68.
- Beckman, M. E., Munson, B., & Edwards, J. (2011). *Methodological issues in the analysis of phonotactic probability effects in nonwords*. Paper presented at the XVIIth International Congress of Phonetic Sciences, Hong Kong.
- Bishop, D. V. M., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological bulletin*, 130(6), 858-886. doi: 10.1037/0033-2909.130.6.858
- Blanchard, D., Heinz, J., & Golinkoff, R. (2010). Modeling the contribution of phonotactic cues to the problem of word segmentation. *Child development*, 37(3), 487-511. doi: 10.1017/S030500090999050X
- Bowey, J. A. (1997). What does nonword repetition measure? A reply to Gathercole and Baddeley. *Journal of experimental child psychology*, 67(2), 295-301. doi: 10.1006/jecp.1997.2408
- Bowey, J. A. (2001). Nonword repetition and young children's receptive vocabulary: A longitudinal study. *Applied Psycholinguistics*, 22(3), 441-469. doi: 10.1017/S0142716401003083

- Bricker, D., & Squires, J. (1999). *Ages & Stages Questionnaires®: A parent-completed child monitoring system (2nd ed.)*. Baltimore, MD: Paul H. Brookes Publishing Co.
- Burke, H. L., & Coady, J. A. (2015). Nonword repetition errors of children with and without specific language impairments (SLI). *International Journal of Language & Communication Disorders, 50*(3), 337-346. doi: 10.1111/1460-6984.12136
- Campbell, T., Dollaghan, C., Needleman, H., & Janosky, J. (1997). Reducing bias in language assessment: Processing-dependent measures. *Journal of Speech, Language, and Hearing Research, 40*(3), 519-525.
- Chan, K. Y. (2012). *Where is the locus of difficulty in recognizing foreign-accented words? Neighborhood density and phonotactic probability effects on the recognition of foreign-accented words by native English listeners*. (Dissertation/Thesis), ProQuest, UMI Dissertations Publishing. Retrieved from http://kuscholarworks.ku.edu/bitstream/handle/1808/10325/Chan_ku_0099D_12154_DATA_1.pdf?sequence=1
- Chiat, S. (2001). Mapping theories of developmental language impairment: Premises, predictions and evidence. *Language and Cognitive Processes, 16*(2), 113-142. doi: 10.1080/01690960042000012
- Coady, J. A., & Aslin, R. N. (2003). Phonological neighbourhoods in the developing lexicon. *Journal of child language, 30*(2), 441-469. doi: 10.1017/S0305000903005579
- Coady, J. A., & Aslin, R. N. (2004). Young children's sensitivity to probabilistic phonotactics in the developing lexicon. *Journal of experimental child psychology, 89*(3), 183-213. doi: 10.1016/j.jecp.2004.07.004
- Coady, J., Evans, J. L., & Kluender, K. R. (2010a). Role of phonotactic frequency in nonword repetition by children with specific language impairments. *International Journal of Language & Communication Disorders, 45*(4), 494-509. doi:

10.3109/13682820903222783

- Coady, J. A., Evans, J. L., & Kluender, K. R. (2010b). The role of phonotactic frequency in sentence repetition by children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 53*(5), 1401-1415. doi: 10.1044/1092-4388(2010/07-0264)
- Dale, P. S., Price, T. S., Bishop, D. V. M., & Plomin, R. (2003). Outcomes of early language delay: 1. Predicting persistent and transient language difficulties at 3 and 4 years. *Journal of Speech, Language, and Hearing Research, 46*(3), 544-560. doi: 10.1044/1092-4388(2003/044)
- 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 & Communication Disorders, 43*(4), 361-389. doi: 10.1080/13682820701546854
- Dollaghan, C., Biber, M., & Campbell, T. (1993). Constituent syllable effects in a nonsense-word repetition task. *Journal of Speech and Hearing Research, 36*, 1051-1054.
- Dollaghan, C. A. (1994). Children's phonological neighbourhoods: half empty or half full? *Journal of child language, 21*(2), 257-271. doi: 10.1017/S0305000900009260
- Dollaghan, C. A., Biber, M. E., & Campbell, T. F. (1995). Lexical influences on nonword repetition. *Applied Psycholinguistics, 16*(2), 211-222. doi: 10.1017/S0142716400007098
- Dollaghan, C. A., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41*(5), 1136.
- Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research, 47*(2), 421-

436. doi: 10.1044/1092-4388(2004/034)

Eernisse, E. R. (2010). *Vocabulary comprehension in late-talking toddlers* (Order No. 3437268). Available from ProQuest Dissertations & Theses A&I. (857089557). Retrieved from <http://search.proquest.com.ezproxy.canterbury.ac.nz/docview/857089557?accountid=14499>

Estes, K. G., & Bowen, S. (2013). Learning about sounds contributes to learning about words: Effects of prosody and phonotactics on infant word learning. *Journal of experimental child psychology*, *114*(3), 405. doi: 10.1016/j.jecp.2012.10.002

Fenson, L., Dale, P., Reznick, J. S., Thal, D., Bates, E., Hartung, J., . . . Reilly, J. (1993). *The MacArthur Communicative Development Inventories: User's guide and technical manual*. San Diego: Singular Publishing Group.

Fenson, L., Marchman, V. A., Thal, D., Dale, P. S., Reznick, J. S. & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories: User's guide and technical manual* (2nd ed.) Baltimore, MD: Brookes.

Freedman, S. E., & Barlow, J. A. (2012; 2011). Using whole-word production measures to determine the influence of phonotactic probability and neighborhood density on bilingual speech production. *International Journal of Bilingualism*, *16*(4), 369-387. doi: 10.1177/1367006911425815

Friederici, A. D., & Wessels, J. M. I. (1993). Phonotactic knowledge of word boundaries and its use in infant speech perception. *Perception & psychophysics*, *54*(3), 287-295. doi: 10.3758/BF03205263

Frisch, S. A., Large, N. R., & Pisoni, D. B. (2000). Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language*, *42*(4), 481-496. doi: 10.1006/jmla.1999.2692

- Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, 28(2), 200-213. doi: 10.1016/0749-596X(89)90044-2
- Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental psychology*, 28(5), 887. doi: 10.1037/0012-1649.28.5.887
- Gathercole, S. E., Frankish, C. R., Pickering, S. J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(1), 84-95. doi: 10.1037/0278-7393.25.1.84
- Gerken, L. A., Murphy, W. D., & Aslin, R. N. (1995). Three- and four-year-olds' perceptual confusions for spoken words. *Perception and Psychophysics*, 57(4), 475-486. doi: 10.3758/BF03213073
- Gierut, J. A., & Dale, R. A. (2007). Comparability of lexical corpora: Word frequency in phonological generalization. *Clinical Linguistics & Phonetics*, 21(6), 423-433. doi: 10.1080/02699200701299891
- Goldrick, M., & Larson, M. (2008). Phonotactic probability influences speech production. *Cognition*, 107(3), 1155-1164. doi: 10.1016/j.cognition.2007.11.009
- Graf-Estes, K., Evans, J. L., & Else-Quest, N. M. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, 50(1), 177-195. doi: 10.1044/1092-4388(2007/015)
- Gray, S., Pittman, A., & Weinhold, J. (2014). Effect of phonotactic probability and neighborhood density on word-learning configuration by preschoolers with typical development and specific language impairment. *Journal of Speech, Language, and Hearing Research*, 57(3), 1011. doi: 10.1044/2014_JSLHR-L-12-0282

- Guo, L.-Y., McGregor, K. K., & Spencer, L. J. (2015). Are young children with cochlear implants sensitive to statistics of words in the ambient spoken language? *Journal of Speech, Language, and Hearing Research*. doi:10.1044/2015_JSLHR-H-14-0135
- Hoover, J. R., Storkel, H. L., & Hogan, T. P. (2010). A cross-sectional comparison of the effects of phonotactic probability and neighborhood density on word learning by preschool children. *Journal of Memory and Language*, 63(1), 100-116. doi: 10.1016/j.jml.2010.02.003
- Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, 30(6), 685-701. doi: 10.1016/0749-596X(91)90032-F
- Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D. A., Martin, S., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(5), 1217-1232. doi: 10.1037/0278-7393.23.5.1217
- Hunter, C. R. (2013). Early effects of neighborhood density and phonotactic probability of spoken words on event-related potentials. *Brain and language*, 127(3), 463-474. doi: 10.1016/j.bandl.2013.09.006
- Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' Preference for the Predominant Stress Patterns of English Words. *Child development*, 64(3), 675. doi: 10.2307/1131210
- Jusczyk, P. W., Friederici, A. D., Wessels, J. M. I., Svenkerud, V. Y., & Jusczyk, A. M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, 32(3), 402-420. doi: 10.1006/jmla.1993.1022
- Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic

- patterns in the native language. *Journal of Memory and Language*, 33(5), 630-645.
doi: 10.1006/jmla.1994.1030
- Jusczyk, P. W. (1999). Narrowing the distance to language: One step at a time. *Journal of communication disorders*, 32(4), 207-222. doi: 10.1016/S0021-9924(99)00014-3
- Klee, T., Stokes, S. F., & Moran, C. (2011-2014). *Early factors in childhood communication disorders*. Marsden Fund, The Royal Society of New Zealand.
- Kuhl, P. K. (1983). Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior and Development*, 6(2), 263-285. doi: 10.1016/S0163-6383(83)80036-8
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., Lindblom, B., Institutionen för, l., . . . Humanistiska, f. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255(5044), 606-608. doi: 10.1126/science.1736364
- LaShell, P. (2014). Personal communication, 8th August, 2014.
- Lee, S. A. S., & Gorman, B. K. (2013). Nonword repetition performance and related factors in children representing four linguistic groups. *International Journal of Bilingualism*, 17(4), 479-495. doi: 10.1177/1367006912438303
- Leonard, L. B. (1998). *Children with specific language impairment*. Cambridge, Mass: The MIT Press.
- Locke, J. L. (1988). The sound shape of early lexical representations. In M. D. Smith & J. L. Locke (Eds.), *The emergent lexicon: The child's development of a linguistic vocabulary* (pp. 3-22). New York: Academic Press.
- Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear & Hearing*, 19(1), 1-36. doi: 10.1097/00003446-199802000-00001
- MacRoy-Higgins, M., & Schwartz, R. G. (2013). Phonology and the lexicon in late talkers. In

- L. Rescorla & P. S. Dale (Eds.), *Late talkers* (pp. 113-128). Baltimore, Maryland: Paul H. Brookes Publishing Co.
- 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 & Communication Disorders, 48*(2), 188. doi: 10.1111/j.1460-6984.2012.00198.x
- 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. doi: 10.1177/0265659013487534
- Mattys, S. L., & Jusczyk, P. W. (2001). Phonotactic cues for segmentation of fluent speech by infants. *Cognition, 78*(2), 91-121. doi: 10.1016/S0010-0277(00)00109-8
- McKean, C., Letts, C., & Howard, D. (2013). Developmental change is key to understanding primary language impairment: the case of phonotactic probability and nonword repetition. *Journal of Speech, Language, and Hearing Research, 56*(5), 1579-1594. doi: 10.1044/1092-4388(2013/12-0066)
- Messer, M. H., Leseman, P. P. M., Boom, J., & Mayo, A. Y. (2010). Phonotactic probability effect in nonword recall and its relationship with vocabulary in monolingual and bilingual preschoolers. *Journal of experimental child psychology, 105*(4), 306-323. doi: 10.1016/j.jecp.2009.12.006
- Miller, N., & Lallini, N. (2011). Do phonological neighbourhood density and phonotactic probability influence speech output accuracy in acquired speech impairment? *Aphasiology, 25*(2), 176-190. doi: 10.1080/02687031003798247
- Morrisette, M. L., & Gierut, J. A. (2002). Lexical organization and phonological change in treatment. *Journal of Speech, Language, and Hearing Research, 45*(1), 143.
- Munson, B. (2001). Phonological pattern frequency and speech production in adults and

- children. *Journal of Speech, Language, and Hearing Research*, 44(4), 778-792. doi: 10.1044/1092-4388(2001/061)
- Munson, B., Kurtz, B. A., & Windsor, J. (2005). The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research*, 48(5), 1033.
- Newman, R. S., & Janse, E. (2013). Identifying nonwords: Effects of lexical neighborhoods, phonotactic probability, and listener characteristics. *Language and Speech*, 56(4), 421-441. doi: 10.1177/0023830912447914
- Olswang, L. B., Rodriguez, B., & Timler, G. (1998). Recommending intervention for toddlers with specific language learning difficulties: We may not have all the answers, but we know a lot. *American Journal of Speech-Language Pathology*, 7(1), 23-32.
- Paul, R., & Jennings, P. (1992). Phonological behavior in toddlers with slow expressive language development. *Journal of speech and hearing research*, 35(1), 99-107.
- Paul, R. (1996). Clinical implications of the natural history of slow expressive language development. *American Journal of Speech-Language Pathology*, 5(2), 5-21. doi: 10.1044/1058-0360.0502.05
- Pegg, J. E., & Werker, J. F. (1997). Adult and infant perception of two English phones. *Journal of the Acoustical Society of America*, 102(6), 3742-3753. doi: 10.1121/1.420137
- Pisoni, D. B. & Garber, E. E. (1990). Lexical memory in visual and auditory modalities: The case for a common mental lexicon. *Proceedings of the 1990 International Conference on Spoken Language Processing* (pp. 401-404). Kobe: The Acoustical Society of Japan.
- Plante, E., Bahl, M., Vance, R., & Gerken, L. (2011). Beyond phonotactic frequency:

- Presentation frequency effects word productions in specific language impairment. *Journal of communication disorders*, 44(1), 91-102. doi: 10.1016/j.jcomdis.2010.07.005
- Reese, E., & Read, S. (2000). Predictive validity of the New Zealand MacArthur Communicative Development Inventory: Words and Sentences. *Journal of Child Language*, 27(02), 255-266.
- Reilly, S., Wake, M., Bavin, E. L., Prior, M., Williams, J., Bretherton, L., . . . Ukoumunne, O. C. (2007). Predicting language at 2 years of age: A prospective community study. *Pediatrics*, 120(6), e1441-e1449. doi: 10.1542/peds.2007-0045
- Reilly, K. J., & Spencer, K. A. (2013). Sequence complexity effects on speech production in healthy speakers and speakers with hypokinetic or ataxic dysarthria. *PLOS ONE*, 8(10), e77450. doi: 10.1371/journal.pone.0077450
- 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.
- 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. doi: 10.1044/1092-4388(2002/028)
- 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. doi: 10.1044/1092-4388(2008/07-0171)
- Rescorla, L., Alley, A., & Christine, J. B. (2001). Word frequencies in toddlers' lexicons. *Journal of Speech, Language, and Hearing Research*, 44(3), 598.
- Rispens, J., Baker, A., & Duinmeijer, I. (2015). Word recognition and nonword repetition in children with language disorders: the effects of neighborhood density, lexical frequency, and phonotactic probability. *Journal of Speech, Language, and Hearing*

Research, 58(1), 78. doi: 10.1044/2014_JSLHR-L-12-0393

Rose, Y., MacWhinney, B., Byrne, R., Hedlund, G., Maddocks, K., O'Brien, P., & Wareham, T. (2006). Introducing Phon: A software solution for the study of phonological acquisition. In David Bamman, Tatiana Magnitskaia & Colleen Zaller (eds.), *Proceedings of the 30th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press. 489-500.

Rose, Y., & MacWhinney, B. (2014). The PhonBank Project: Data and software-assisted methods for the study of phonology and phonological development. In Jacques Durand, Ulrike Gut & Gjert Kristoffersen (eds.), *The Oxford Handbook of Corpus Phonology*. Oxford: Oxford University Press. 308-401.

Rossi, S., Jürgenson, I. B., Hanulíková, A., Telkemeyer, S., Wartenburger, I., & Obrig, H. (2011). Implicit processing of phonotactic cues: Evidence from electrophysiological and vascular responses. *Journal of cognitive neuroscience*, 23(7), 1752-1764. doi: 10.1162/jocn.2010.21547

Shriberg, L. D., & Lohmeier, H. L. (2008). *The syllable repetition task (SRT)* [Technical report No. 14]. Madison, WI: Waisman Center.

Shriberg, L. D., & Kent, R. D. (2013). Procedures to calculate transcription reliability and research findings. In *Clinical Phonetics, 4th ed.* (pp. 371-379). Boston, MA: Pearson Education.

Snowling, M., Chiat, S., & Hulme, C. (1991). Words, nonwords, and phonological processes—some comments on Gathercole, Willis, Emslie, and Baddeley. *Applied Psycholinguistics*, 12(3), 369-373. doi: 10.1017/S0142716400009279

Sosa, A. V., & Stoel-Gammon, C. (2012). Lexical and phonological effects in early word production. *Journal of Speech, Language, and Hearing Research*, 55(2), 596-608. doi: 10.1044/1092-4388(2011/10-0113)

- Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, 388(6640), 381-382. doi: 10.1038/41102
- Stokes, S. F., & Klee, T. (2009a). The diagnostic accuracy of a new test of early nonword repetition for differentiating late talking and typically developing children. *Journal of Speech, Language, and Hearing Research*, 52(4), 872-882. doi: 10.1044/1092-4388(2009/08-0030)
- Stokes, S. F., & Klee, T. (2009b). Factors that influence vocabulary development in two-year-old children. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 50(4), 498.
- Stokes, S. F. (2010). Neighborhood density and word frequency predict vocabulary size in toddlers. *Journal of Speech, Language, and Hearing Research*, 53(3), 670-683. doi: 10.1044/1092-4388(2009/08-0254)
- Stokes, S. F., & Klee, T. (2011). *Verbal and visual-spatial short-term memory in preschool children*. Paper presented at the 12th International Congress for the Study of Child Language (IASCL), Montreal, Canada.
- Storkel, H. L., & Rogers, M. A. (2000). The effect of probabilistic phonotactics on lexical acquisition. *Clinical Linguistics & Phonetics*, 14(6), 407-425. doi: 10.1080/026992000415859
- Storkel, H. L. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research*, 44(6), 1321.
- Storkel, H. L., & Hoover, J. R. (2010). Word learning by children with phonological delays: Differentiating effects of phonotactic probability and neighborhood density. *Journal of communication disorders*, 43(2), 105-119. doi: 10.1016/j.jcomdis.2009.11.001
- Storkel, H. L., Maekawa, J., & Hoover, J. R. (2010). Differentiating the effects of phonotactic

- probability and neighborhood density on vocabulary comprehension and Production: A comparison of preschool children with versus without phonological delays. *Journal of Speech, Language, and Hearing Research*, 53(4), 933-949. doi: 10.1044/1092-4388(2009/09-0075)
- Storkel, H. L., & Hoover, J. R. (2011). The influence of part-word phonotactic probability/neighborhood density on word learning by preschool children varying in expressive vocabulary. *Journal of child language*, 38(3), 628-643. doi: 10.1017/S0305000910000176
- Storkel, H. L., Bontempo, D. E., Aschenbrenner, A. J., Maekawa, J., & Lee, S.-Y. (2013). The effect of incremental changes in phonotactic probability and neighborhood density on word learning by preschool children. *Journal of Speech, Language, and Hearing Research*, 56(5), 1689-1700. doi: 10.1044/1092-4388(2013/12-0245)
- Stothard, S. E., Snowling, M. J., Bishop, D. V. M., Chipchase, B. B. and Kaplan, C. A. (1998). Language impaired preschoolers: A follow up into adolescence. *Journal of Speech, Language and Hearing Research*, 41, 407–418.
- Tanida, Y., Ueno, T., Lambon Ralph, M. A., & Saito, S. (2015). The roles of long-term phonotactic and lexical prosodic knowledge in phonological short-term memory. *Memory & cognition*, 43(3), 500-519. doi: 10.3758/s13421-014-0482-2
- Templin, M. C. (1953). Speech development in the young child. III. The development of certain language skills in children. *The Journal of speech and hearing disorders*, 17(3), 280.
- Thal, D. J., Oroz, M., & McCaw, V. (1995). Phonological and lexical development in normal and late-talking toddlers. *Applied Psycholinguistics*, 16(4), 407-424.
- Tomblin, J. B., Zhang, X., Buckwalter, P., & Catts, H. (2000). The Association of reading disability, behavioral disorders, and language impairment among second-grade

- children. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 41(4), 473-482. doi: 10.1017/S002196300000559X
- Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science*, 9(4), 325-329. doi: 10.1111/1467-9280.00064
- Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, 40(3), 374-408. doi: 10.1006/jmla.1998.2618
- Vitevitch, M. S., & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, & Computers*, 36(3), 481-487. doi: 10.3758/BF03195594
- Vitevitch, M. S., & Luce, P. A. (2005). Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of Memory and Language*, 52(2), 193-204. doi: 10.1016/j.jml.2004.10.003
- Vitevitch, M. S., & Donoso, A. J. (2012). Phonotactic probability of brand names: I'd buy that. *Psychological research*, 76(6), 693-698. doi: 10.1007/s00426-011-0374-z
- Walley, A. C. (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review*, 13(3), 286-350.
- Weismer, S. E., & Evans, J. L. (2002). The role of processing limitations in early identification of specific language impairment. *Topics in Language Disorders*, 22(3), 15.
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7(1), 49-63. doi: 10.1016/S0163-6383(84)80022-3
- 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. doi: 10.1044/0161-1461(2003/012)
- Zamuner, T. S., Gerken, L., & Hammond, M. (2004). Phonotactic probabilities in young children's speech production. *Journal of child language*, 31(3), 515-536. doi: 10.1017/S0305000904006233
- Zamuner, T. S. (2009). Phonotactic probabilities at the onset of language development: Speech production and word position. *Journal of Speech, Language, and Hearing Research*, 52(1), 49-60. doi: 10.1044/1092-4388(2008/07-0138)
- Zamuner, T. S. (2013). Perceptual evidence for young children's developing knowledge of phonotactic probabilities. *Language Acquisition*, 20(3), 241.
- Zimmerman, I., Steiner, V., Pond, R. (2002). *Preschool Language Scale, Fourth Edition*. San Antonio, TX The Psychological Corporation
- Zubrick, S., Taylor, C., Rice, M., & Slegers, D. (2007). Late language emergence at 24 months: An epidemiological study of prevalence, predictors and covariates. *Journal of Speech, Language, and Hearing Research*, 50, 1562-1592.

Appendix A

TEST OF EARLY NONWORD REPETITION

Child's Number _____ Date _____

IPA Target	IPA child response	Score
mad		
neit		
paim		
bouk		
kougə		
dafi		
lɜpou		
fupɪm		
moukɜri		
doupɜlut		
bæɜɜkɜn		
fisaimɔt		
pɜduləmeip		
fɜnɜraisek		
wugɜləmɪk		
lɔdɜnætɪf		
gɪɜmafukou		
lɜteidikunei		
gɜlumɜfinai		
bafumouwudi		

© Stokes and Klee, 2011

Appendix B

R script used to calculate phonotactic probability values used in this study

```
#####  
# Phonotactic Probability Calculator for Celex using Vitevitch's method  
#####  
# Requires:  
  
# CELEX - Lexical Corpus  
# wordlist.txt - contains DISC phonology transcriptions for your words. 1 per line  
  
# Calculates Positional Segment Frequency and Position-specific Biphone Frequency  
# Note: Celex claims to have ~17.9 million tokens, but summed frequency is ~1 million  
higher.  
# Note: Celex is 90% a written Corpus (of secret Documents). It is therefore a rough  
approximation of verbal frequency  
# Note: Celex has alternative phonologies for many words. I only count the "base" phonology.  
#####  
  
#####  
#Notes and options for later development  
#####  
  
##Pick your corpus  
  
# eol English Orthography, Lemmas  
# epl English Phonology, Lemmas  
# eml English Morphology, Lemmas  
# esl English Syntax, Lemmas
```

```

#   efl   English Frequency, Lemmas
#
#   eow   English Orthography, Wordforms
#   epw   English Phonology, Wordforms
#   emw   English Morphology, Wordforms
#   efw   English Frequency, Wordforms
#
#   ect   English Corpus Types
#   efs   English Frequency, Syllables

```

```
#####
```

```
#Set the Source of your Celex datafile
```

```
#####
```

```
#setwd("C:/Documents and Settings/Patrick
LaShell/Desktop/nzilbb/CELEX/ENGLISH/EPL")
```

```
#setwd("C:/Documents and Settings/Patrick
LaShell/Desktop/nzilbb/CELEX/ENGLISH/EFL")
```

```
#setwd("C:/Documents and Settings/Patrick
LaShell/Desktop/nzilbb/CELEX/ENGLISH/EPL")
```

```
#setwd("C:/Documents and Settings/Patrick
LaShell/Desktop/nzilbb/CELEX/ENGLISH/EFW")
```

```
#setwd("C:/Documents and Settings/Patrick
LaShell/Desktop/nzilbb/CELEX/ENGLISH/EMW")
```

```
setwd("C:/Documents and Settings/Patrick
LaShell/Desktop/nzilbb/CELEX/ENGLISH/EPW") #English Phoneme Frequency for Word
```

Forms

```
#####
```

```
#Read the data from CELEX file
```

```
celex <- readLines(con = file("EPW.CD", open = "r"))
```

```
#The epw.cd file contains the following fields:
```

```
#
```

```
# 1. IdNum
```

```
# 2. Word
```

```
# 3. Cob
```

```
# 4. IdNumLemma
```

```
# 5. PronCnt
```

```
# 6. PronStatus
```

```
# 7. PhonStrsDISC
```

```
# 8. PhonCVBr
```

```
# 9. PhonSylBCLX
```

```
#
```

```
#Those words which appear with alternative pronunciations are assigned
```

```
#4 extra fields for each pronunciation. For instance, the columns
```

```
#
```

```
# 10. PronStatus
```

```
# 11. PhonStrsDISC
```

```
# 12. PhonCVBr
```

```
# 13. PhonSylBCLX
```

```

##
#COBUILD frequency 17.9m Cob
#COBUILD 95% confidence deviation 17.9m CobDev
#COBUILD all sources
#COBUILD frequency 1m CobMln
#COBUILD frequency , logarithmic CobLog
#COBUILD written frequency 16.6m CobW
#COBUILD written sources COBUILD written frequency 1m CobWMln
#COBUILD written frequency , logarithmic CobWLog
#COBUILD spoken frequency 1.3m CobS
#COBUILD spoken sources COBUILD spoken frequency 1m CobSMln
#COBUILD spoken frequency , logarithmic CobSLog

#load the plyr library which makes the script simpler
library(plyr)

#split the dataset and combine into a dataframe. All values are factors
epw <- ldply(strsplit(celex, "\\|"), rbind)

#Name the columns of our dataset
names(epw)[1:9] <- c("IdNum", "Word", "Cob", "IdNumLemma", "PronCnt", "PronStatus",
"PhonStrsDISC", "PhonCVBr", "PhonSylBCLX")

#Convert our Frequency column to numerics
#Two options here to deal with words with 0 frequencies.

```

```

# Option 1. Add 1 to every frequency

# Option 2. Remove all words with 0 Frequencies

# We are currently using Option 2.

epw$Word <- as.character(epw$Word)

epw$Cob <- as.numeric(as.character(epw$Cob))

epw <- epw[epw$Cob > 0,]

#Log frequency column

epw$lCob <- log10(epw$Cob)

#Strip out syllable separation and stress markers from the phonology

epw$Phon <- gsub("\\", "", epw$PhonStrsDISC)

epw$Phon <- gsub("\\'", "", epw$Phon)

epw$Phon <- gsub("-", "", epw$Phon)

#Get wordlist from file

setwd("C:/Documents and Settings/Patrick LaShell/Desktop/nzilbb/stephanie") #Source
directory for wordlist on my computer

#read list of words to find.

wordlist <- read.csv("wordlist.txt", header = FALSE, stringsAsFactors=FALSE)

names(wordlist) <- "word"

#####

#Create output columns in wordlist

#####

```

```

#calculate max number of phonemes

max.phonemes <- max(nchar(wordlist[,1]))

#create individual phonemes

phon.frame <- data.frame(matrix(0, nrow(wordlist), max.phonemes))

names(phon.frame)<- paste("phon.", 1:max.phonemes, sep= "")

#create biphones

biphon.frame <- data.frame(matrix(0, nrow(wordlist), max.phonemes - 1 ))

names(biphon.frame)<- paste("biphon.", 1:(max.phonemes-1), sep= "")

#create individual phoneme positional frequency

phon.freq.frame <- phon.frame

names(phon.freq.frame)<- paste("phon.freq.", 1:max.phonemes, sep= "")

#create biphone positional frequency

biphon.freq.frame <- biphon.frame

names(biphon.freq.frame)<- paste("biphon.freq.", 1:(max.phonemes-1), sep= "")

#glue everything together

wordlist <- data.frame(wordlist, phon.frame, phon.freq.frame, biphon.frame,
biphon.freq.frame)

#####

```

```

#Calculate sum freq of distribution of wordlengths.

#####

length.freq <- rep(0, max.phonemes)

for(i in 1:max.phonemes){

length.freq[i] <- sum(epw$Icob[nchar(epw$Phon) >= i])

}

#####

#####

#Calculate positional phoneme frequency

#####

#Iterate through each word and through each phoneme

for(i in 1:nrow(wordlist)){

for(j in 1:nchar(wordlist$word[i])){

#write the phoneme to the phon.position

wordlist[i, grep(paste("^phon.", j, "$", sep=""), names(wordlist))] <-

substr(wordlist$word[i],j,j)

#create the search pattern for each phoneme

search.string <- paste("^{", j - 1, "}", "[", substr(wordlist$word[i],j,j), "]", sep= "")

#sum the phoneme match frequency and divide by the number of words with j or more

phonemes

wordlist[i, grep(paste("^phon.freq.", j, "$", sep=""), names(wordlist))] <-

```

```

sum(epw$ICob[grep(search.string, epw$Phon)]/length.freq[j]
}
}

#####

#Calculate positional biphoneme frequency

#####

#Iterate through each word and through each phoneme
for(i in 1:nrow(wordlist)){
for(j in 1:(nchar(wordlist$word[i])-1)){

#write the phoneme to the phon.position
wordlist[i, grep(paste("^biphon.", j, "$", sep=""), names(wordlist))] <-
substr(wordlist$word[i],j,j+1)

#create the search pattern for each phoneme

#Note: [x] syntax is a character class, anything in it is valid, so we can't just lump both
phonemes together

#We have to use character classes because some of the DISC characters are special escape
characters.

search.string <- paste("^{", j - 1, "}", "[", substr(wordlist$word[i],j,j), "]",
substr(wordlist$word[i],j+1,j+1), "]",sep= "")

#sum the phoneme match frequency and divide by the number of words with j or more
phonemes

```

```

wordlist[i, grep(paste("^biphon.freq.", j, "$", sep=""), names(wordlist))] <-
sum(epw$ICob[grep(search.string, epw$Phon)]/length.freq[j+1]
}
}

#write to file

write.csv( wordlist,"output.txt", row.names = FALSE)

#####

### Code to check if the measures are correct

#####

###check 1st letter

##bob <- data.frame(phon =
unlist(strsplit("IE{VQU@i#u312456789cq0~pbtckgNmnlrfvTDszSZjxhwJ_CFHPR", "")))

##bob$freq <- 0

##

##for(i in 1:nrow(bob)){

##search.string <- paste("^", "[", bob$phon[i], "]", sep= "")

##bob$freq[i] <- sum(epw$ICob[grep(search.string, epw$Phon)])

##}

##bob$lfreq <- bob$freq/length.freq[1]

#

#

##table(substr(epw$Phon, 1,1))

##epw[grep("$U", epw$Phon),]

```

```

#
###check 2nd letter
#fred <- data.frame(phon =
unlist(strsplit("IE{VQU@i#u312456789cq0~pbtdkgNmnlrfvTDszSZjxhwJ_CFHPR", "")))
#fred$freq <- 0
#
#for(i in 1:nrow(fred)){
#search.string <- paste("^.", "[", fred$phon[i], "]", sep= "")
#fred$freq[i] <- sum(epw$Icob[grep(search.string, epw$Phon)])
#}
#fred$lfreq <- fred$freq/length.freq[2]
#
##check when biphon 7 is "mI"
#sum(epw$Icob[substr(epw$Phon, 7,8) == "mI" & nchar(epw$Phon) > 7])/length.freq[8]

```

Appendix C

Phonotactic probability averages for the nonwords on the TENR-R calculated using CELEX

IPA Target	Positional Segment Frequency average	Biphone Frequency average
mad	0.038072988	0.002578496
neit	0.040959915	0.002284826
paim	0.052427624	0.001084140
bouk	0.047035803	0.002058151
kougə	0.059198223	0.002151444
dafi	0.030489693	0.000863957
lɜpou	0.027361138	0.000840222
fupɪm	0.050713938	0.002262233
moukɜri	0.046888797	0.005969089
doupɜlut	0.057119267	0.003641382
bɛlɜkɒn	0.054884948	0.003978146
fɪsɪmɒt	0.046269249	0.001279070
pɜduləmeɪp	0.037278869	0.002093501
fɛnɜraɪsɛk	0.050517677	0.006120821
wugɜlɛmɪk	0.044690065	0.004367569
lɒdɜnætɪf	0.064041183	0.011024475
gɪlɜmafukou	0.029305779	0.001920124
lɜteɪdɪkuneɪ	0.032971937	0.001081469
gɒlumɜfɪnɪ	0.022810026	0.000778061
bafumouwudi	0.024141710	0.000712128