Electrophysiological Correlates of Speech Perception in Young Children: Associations Among ERP, Nonword Repetition and Language

Vanessa Marie Harwood
University of Connecticut - Storrs, vanessa.harwood@uconn.edu

Follow this and additional works at: http://digitalcommons.uconn.edu/dissertations

Recommended Citation
http://digitalcommons.uconn.edu/dissertations/733
Event Related Potentials (ERP) recorded during infancy and early childhood have been used to predict future language outcomes in children. Furthermore, there is recent evidence that nonword repetition (NWR) can be used to identify language delay in toddlers. This investigation assesses the relationships among ERP markers of sensitivity to phonemic stimuli, nonword repetition, and language to determine if the aforementioned methodologies could improve diagnostic measures for young children. Forty children between the ages of 24 to 48 months participated in a series of behavioral speech and language measures including the mCDI-2, the PLS-5, the GFTA-2 and conventional language sampling. ERPs were recorded during an “old-new” paradigm to examine sensitivity to phonological changes. A nonword repetition task was also administered as a compliment to the ERP recordings to determine the independent and combined contribution of phonological working memory in predicting language ability. Results reveal that ERP markers of phonemic processing are strongly correlated with clinical assessments and are able to predict language skill independently from nonword repetition. These findings suggest that phonological sensitivity as measured by ERP and phonological working memory as measured by nonword repetition have a fundamental yet distinct relationship to general language ability in young children. Both clinical implications and fundamental questions regarding the underlying mechanisms of language disorders are addressed.
Electrophysiological Correlates of Speech Perception in Young Children:
Associations among ERP, Nonword Repetition and Language

Vanessa M. Harwood

B.A., Providence College, 2002
M.S., University of Rhode Island, 2005

A Dissertation
Submitted in Partial Fulfillment of the
Requirements for the
Degree of Doctor of Philosophy
at the
University of Connecticut
2015
APPROVAL PAGE

Doctor of Philosophy Dissertation

Electrophysiological Correlates of Speech Perception in Young Children: Associations Among ERP, Nonword Repetition and Language

Presented by

Vanessa M. Harwood, B.A., M.S.

Major Advisor ____________________________________________________________

Bernard Grela

Associate Advisor _________________________________________________________

Emily Myers

Associate Advisor _________________________________________________________

Tammie Spaulding

Associate Advisor _________________________________________________________

Nicole Landi

Associate Advisor _________________________________________________________

Jonathan Preston

University of Connecticut

2015
Acknowledgements & Dedication

Thank you to Dr. Bernard Grela who has been a very supportive, kind and encouraging advisor. I also thank my dissertation committee: Dr. Emily Myers, Dr. Tammie Spaulding, and Dr. Nicole Landi, who have been not only an inspiration to me as women in science, but also have prepared me for my next steps. Finally, I am forever indebted to Dr. Jonathan Preston who has guided me in the idea, design and preparation for this dissertation. It has been a privilege to work with you and I am forever indebted.

I am grateful to Miss Dootie Roy and Dr. Michael Harwood who cheerfully provided wonderful statistical support for this study: Dootie much more cheerfully than Michael. Butch- I could not have done this without you. You have supported me every step of the way. You think an MD would be enough- but now you had to get yourself a Ph.D. as well. Thank you for believing in me. I love you.

I am grateful to Miss Nina Gumkowski & Miss Olivia Harold who provided ERP technical support for this study.

I am forever grateful for all the undergraduate, graduate and professional research assistants who provided data analysis or in some way supported this study.

Thank you to my moms Margaret Ravenelle & Anne Harwood who deserve all the good things that accompany this degree. I love you and could not have done this without you. Sam and Jakie, you are ok too.

To my Bradley colleagues: Stephanie Lancellotti, OTRL, Jessica Drinkwater, M.Ed., Maggie LeClerc, COTA, and Dr. Greta Francis- you continued to remind me over the last five years to 1) laugh because somehow- it’s always funny. And 2) even though this research has been slow...and difficult…the children are worth it. Thank you for letting me watch true genius at work.

Dana and Sayako: I couldn’t have asked for better lab mates, colleagues and friends. I will always treasure our times together and I wish you all the best in your future careers. I have learned so much from the both of you and I am so grateful for all you have shared with me.

Thank you to the American Speech Language and Hearing Foundation for bestowing the 2012 Student Research Grant in Early Childhood Development and also to the College of Liberal Arts and Sciences (CLAS) for awarding the 2014 Graduate Fellowship.

Lastly, thank you to all the wonderful children (with beautiful brains) and families who took part in the LEARN study. Because of you, we know more and for that we are grateful.

This dissertation is dedicated to my first mentor, esteemed colleague, and dear friend Rod A. Gragg, Ph.D. of the Bradley Schools who always encouraged me to pay attention to the little details while inspiring me to enjoy the big picture.
TABLE OF CONTENTS

ABSTRACT..............................................................................................................................................

TITLE PAGE...........................................................................................................................................i
APPROVAL PAGE......................................................................................................................................ii
ACKNOWLEDGEMENTS......................................................................................................................iii
TABLE OF CONTENTS..........................................................................................................................iv
LIST OF TABLES......................................................................................................................................vi
LIST FIGURES..........................................................................................................................................vii

CHAPTER: INTRODUCTION ....................................................................................................................1
Language Assessment for Young Children
Early Markers of Phonological Acquisition
Event Related Potentials
Nonword Repetition
Purpose and Hypothesis

CHAPTER 2: METHODS ..........................................................................................................................15
Participants
Procedures
Behavioral Language Assessment
  Parent Report
  Language Sampling
  Standardized Assessment
  Factor Analysis
Nonword Repetition & ERP
  Nonword Repetition
  ERP Procedures
  ERP Task
Data Analysis
  ERP Processing
  ERP Analysis Method

CHAPTER 3: RESULTS ............................................................................................................................29
Analysis of Language Factor Score
N1/P2
Principal Components Analysis
  Temporal factor 4 (136-220ms)
  Temporal factor 2 (248-360ms)
ERP Time Windows and the LFS
  N1/P2 and TF2_SF1
  TF4_SF2 and TF2_SF1
  N1/P2 and TF4_SF2
CHAPTER 4: DISCUSSION

Summary
Theoretical Implications
Clinical Implications
Limitations
General Conclusion

Appendices
Appendix A: Test of Early Nonword Repetition Word List
Appendix B: TENR Scoring Protocol

References
List of Tables

Table 1  Participant Behavioral Scores  Page 22

Table 2  Factor Loadings for Language Factor Score  Page 31

Table 3  Correlations Table for Language Measures and ERP  Page 42

Table 4  Regression Table for ERP, NWR and Age Predicting Language  Page 43
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Electrode Montage for N1/P2 complex</td>
<td>32</td>
</tr>
<tr>
<td>Figure 2</td>
<td>ERP Waveforms for N1/P2 complex</td>
<td>33</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Electrode Montage for TF4_SF2 (136-220ms)</td>
<td>35</td>
</tr>
<tr>
<td>Figure 4</td>
<td>ERP Waveforms for TF4_SF2</td>
<td>36</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Electrode Montage for TF2_SF1 (136-220ms)</td>
<td>38</td>
</tr>
<tr>
<td>Figure 6</td>
<td>ERP Waveforms for TF4_SF2</td>
<td>39</td>
</tr>
</tbody>
</table>
Electrophysiological Correlates of Speech Perception in Young Children: Associations among ERP, Nonword Repetition and Language

Early diagnosis and intervention play a critical role in advancing language development for toddlers with language delay (Girolametto, Wiigs, Smyth, Weitzman, & Pearce, 2001; Guralnick, 1997; Reynolds, Temple, Robertson, & Mann, 2000; Roberts & Kaiser, 2011). However, late talking toddlers are a perplexing population for clinicians due to the vast variability in language performance within the group. This paper explores the use of ERPs as a measure of speech perception and a novel nonword repetition task as a measure of phonological working memory to determine their efficacy in improving diagnostic procedures in young children. Furthermore, theoretical implications on the relationships among sensitivity to phonological information and phonological working memory are explored.

**Literature Review**

Approximately 10-15% of the toddler population demonstrate delays in language acquisition despite intact sensory and motor development (Law, Boyle, Harris, Harkness, & Nye, 2000) A deferral in diagnosis of language impairment is problematic in the face of consistent evidence that early and intensive language intervention provides the best means for improvement of skills (Ramey-Landesman & Ramey, 1999). By determining which toddlers are most at risk for future language problems, it is possible to capitalize on early learning and cultivate meaningful language gains.

Many of the assessments currently used to capture language ability for young children have limitations in their ability to sensitively detect impairment. It is possible that these current assessments are not driven to test underlying mechanisms which impact language development, but rather assess language performance. Investigations in neuroscience provide evidence that
early speech perception, as measured by event related potentials (ERPs), can adequately predict language skills further along the developmental trajectory (Kuhl, Conboy, Padden, Rivera-Gaxiola, & Nelson, 2008; Molfese & Molfese, 1997). Furthermore, deficient nonword repetition skills have been associated with language impairment in school-aged children and recent studies have investigated the use of nonword repetition to detect language delay in toddlers (Stokes & Klee, 2009; Dollaghan & Campbell, 1998). To date, there are no studies that have investigated ERPs within the toddler population to determine how phonological sensitivity relates to language output at this critical point in development. It is possible that by determining the relationships among phonological sensitivity, production of nonwords and general language ability, we can improve upon diagnostic procedures for young children while providing evidence for mechanisms impacting language delay.

Language Assessment for Young Children

In a recent review, Crais (2011) outlined methods and strategies for assessment of toddlers and young children. The Crais review is a condensed version of the American Speech and Hearing Association (ASHA) document, *Roles and responsibilities of speech-language pathologists in early intervention: Guidelines* (ASHA, 2008). The author reports that a variety of instruments such as criterion referenced probes, play based-dynamic and authentic assessments, parent report, clinical observations and clinical judgment are appropriate methodologies to measure the language abilities within the toddler population. Two methods of interest, language sampling and standardized assessment, were also suggested as a means to assess language competence within toddlers and young children.

Many clinicians utilize language sampling to provide a fine-grained analysis of expressive language for young children. Language sampling is a recommended procedure that
captures a child’s language production in the absence of prompted speech (Bernstein &
Tiegerman-Faber, 1997; Leadholm & Miller, 1992). Language sampling also provides the
ability to calculate mean length of utterance (MLU). A child’s MLU is used to examine the
structural changes in children’s productions on the basis of increased utterance length. Loeb,
Kinsler, and Bookbinder (2000) surveyed preschool SLPs and found that over 90% of therapists
reported MLU as being their primary language sampling measure.

Eisenberg, Fersko, and Lundgren (2001) used criteria similar to that used by McCauly
and Swisher (1984), including: clearly stated-purpose of the test, specified normative data,
appropriate reference data, and evidence of reliability and validity, to evaluate the usefulness of
MLU in diagnosing a language disorder in preschool children. Based on their evaluation, the
authors claimed that MLU should be indicative of utterance length only, and not syntactic
complexity. Longer utterances are not necessarily more syntactically complex than shorter ones
(e.g., “want more cookies Mommy” vs. “I want to go home”). Furthermore, MLU can be used to
identify some, but not all, preschool children with language impairment. The authors determined
that by using a – 1.5 standard deviation cutoff, the efficiency of MLU in identifying a child as
truly impaired, or test sensitivity, is approximately 63%. By defining a cutoff score, the authors
concluded that MLU could be useful in defining specific children as unimpaired; however,
having a score above the cutoff does not guarantee that a child is free of impairment and
typically developing. The authors suggest that the use of MLU may be effective in supporting a
diagnosis of a language disorder, but should not be utilized as the sole criteria in doing so.

Assessment practices for toddlers and young children also include standardized testing as
one particular method to capture language performance. Clinicians serving young children often
rely heavily on results of standardized assessments in their decision to qualify a child as having a
language impairment (Roulstone, Peters, Glogowska, & Enderby, 2008). However, standardized assessments for young children are not without limitations. First, there are few standardized language assessments normed for children as young as 18-36 months that have been validated by third party research regarding sensitivity and specificity. This gives SLPs limited ability in choosing an assessment of good quality. Next, standardized assessments used for the toddler population are lacking in their ability to accurately determine the nature and severity of impairment as well as demonstrate adequate predictive validity, which provides information on how well the child will perform in the future (Friberg, 2010; Spaulding, Szulga, & Figueroa, 2012).

Child temperament and issues of test validity can also limit the integrity of an assessment to capture language skill. Young children are limited in their ability to attend for long periods of time, which can compromise a clinician’s ability to make a valid judgment of a child’s knowledge of a particular language skill. Many of the standardized measures used to assess receptive and expressive language in toddlers encompass a wide variety of skills including phonological productions, semantic understanding and production, use and processing of grammatical morphology, and pragmatics. It is possible that the skills measured using standardized assessment are too broad in scope and include a range of language and cognitive processes which makes it difficult to clearly define children at risk for impairment. Furthermore, standardized assessments may be inadequate in measuring language skill for very young children, due to their failure to test underlying processes essential for language learning.

**Early Markers of Phonological Acquisition**

Infants demonstrate the ability to perceive phonemic differences within their native language, but also demonstrate the ability to perceive non-native phonemic differences as well.
Around the age of 9 months, children appear to lose the ability to perceive non-native contrasts as the child is immersed within their native language (Kuhl, Conboy, Padden, Rivera-Gaxiola, & Nelson, 2008). This change in perception may stem from the infant’s reliance on the acoustic properties of phonemic information to detect changes within speech during the first few months of life; however, as the infant is bathed in the native language, they home in on the salient features of native phonemes to support word learning. Categorical perception, which allows for the detection of changes in phonemes even in the presence of acoustic variance, is a hallmark of early language learning. Strong categorical perception may link to the infant’s formation of phonemic representations.

There is theoretical basis to believe that some forms of language impairment stem from degraded phonological representations. According to the perceptual deficit theory (PDT) (Joanisse & Seidenberg, 2003), impairment in language, specifically poor grammatical morphology, stems from a perceptual phonological impairment in which degraded perceptual skills affect phonological working memory leading to weaknesses in a child’s ability to form stable linguistic representations. Phonological working memory is an active memory process in which phonological information is stored for a short period of time so that it can be “manipulated”. Because phonological working memory is required to establish critical relationships among sentence parts (Just & Carpenter, 1992; Waters & Caplan, 1996; MacDonald & Christiansen, 2002), poor phonological working memory leads to poor comprehension such that syntactic relationships are neither forged nor maintained. There is evidence that certain clinical populations demonstrate poor phonological working memory such as children with SLI (Montgomery, 1995), children with reading disabilities (Mann, Shankweiler, & Smith, 1984) and adults with conduction aphasia (Gvion & Friedmann, 2012).
These aforementioned groups perform poorly on specific parameters of sentence comprehension, particularly those that require reactivation of phonological information given increased sentence length.

Joanisse and Seidenberg (2003) provide support for the PDT by devising a connectionist model. They provide two simulations which involve several aspects of syntax, including: pronoun resolution, the recognition of word meanings in sentence context, the acquisition of abstract phrase structure, and use of syntactic structure to resolve long distance syntactic complexities. First, a “typical” model was created to demonstrate how adequate speech perception enables syntactic learning through distributed neural networks. Within the second simulation, a perceptual deficit was introduced by adding “noise” to the phonological input preventing the model from developing consistent phonological representations. The noisy phonological input was used within the same networks to determine if sentence comprehension problems would occur given inconsistent phonological forms. The authors predicted that disrupted phonological input would lead to a decline in the model’s ability to maintain words in memory, in essence affecting phonological working memory. Results showed that the unimpaired simulation correctly recognized 93% of the sentences in the training set whereas the impaired network recognized 74%. The impaired network performed significantly worse in computing (or recognizing) grammatical from ungrammatical sentences. In terms of pronoun and reflexive resolution, the impaired network performed worse than the unimpaired network on identifying the correct pronoun referent for both regular pronouns and reflexives. Interestingly, a “gender” set was devised in which the gender information was useful in helping to resolve anaphors. For example, in the sentence *Bob thinks Sally likes him*, the use of the female name allows for greater information to determine that *him* refers to *Bob* versus the sentence, *Bob thinks
Stan likes him. The impaired network performed similarly to the unimpaired network on the gender set, suggesting that the perceptual deficit did not lead to a “wholesale degradation in performance,” but to a specific deficit in utilizing syntactic information. In conclusion, the model suggests degraded phonological representations were responsible for deficits in comprehension of grammatical morphology, and that phonology and working memory are “in fact inseparable and indistinct components of cognitive processing” (p.54).

The PDT suggests that perceptual abilities provide a solid foundation for further language competence. There is a considerable body of research that suggests early perceptual abilities predict later language development (Benasich & Tallal, 2002; Kuhl, Conboy, Padden, Rivera-Gaxiola, & Nelson, 2008; Molfe, & Molfe, 1985; 1997; Rivera-Gaxiola, Klarman, Garcia-Sierra, & Kuhl, 2005). Furthermore, compromised or atypical perceptual skills have been found in a variety of developmental disorders including SLI (Ceponiene, Cunnings, Wulfeck, Ballantyne, & Townsend, 2009; Stevens, Paulsen, Yasen, Mitsunaga, & Neville, 2012; Weber-Fox, Leonard, Hampton, & Tomblin, 2010), speech sound disorders (Rvachew & Growburg, 2006), dyslexia (Guttorm, et al., 2005; Leppanen, et al., 2003, 2011) and autism (Kuhl, Coffey-Corina, Padden, & Dawson, 2005; Roth, Muchnik, Shabta, Hildesheimer, & Henkin, 2011). The aforementioned studies collectively suggest that children who are able to perceive subtle, yet distinct, changes in auditory stimuli may fare better at language and language related skills, including literacy. By examining the perceptual characteristics of young children, researchers may not only provide further evidence for the neural networks critical to language learning, but may also improve upon diagnostic procedures for young children by objectively testing underlying mechanisms which give rise to global language competence.

Event Related Potentials
Advances in neuroscience now allow for sensitive measurement of the neural response to speech through use of ERPs. The ERP method requires an experimenter to record voltage changes on the human scalp resulting from electrical activity generated by neurons within the brain. The electroencephalogram (EEG) measures summed postsynaptic potentials, which are produced when neurotransmitters bind to receptors on the membrane of the postsynaptic cell, causing ion electrodes to open or close resulting in a graded change in potential across the cell membrane (Wood & Allison, 1981). Postsynaptic potentials can last hundreds of milliseconds that allow voltage to summate and be recorded on the scalp using electrodes (Luck, 2005). ERP is a time-locked analysis of the ongoing electroencephalogram, which can reflect precise temporal changes in neural activity when provided with a stimulus such as speech.

The use of ERPs offers distinct advantages in studying language. This noninvasive technique is excellent for studying the human perception of fine-grained phonological input. Extensive measures have been taken to ensure that electrophysiological techniques are feasible and safe to use with infants and young children. Electrophysiological techniques can be passive in nature, not requiring an overt behavioral response. ERPs therefore become an attractive tool to use with children who are too young to provide an overt response and in clinical populations in which behavioral and attention issues impede valid test results (Naatanen, 2003; deRegnier, 2005).

Several studies have explored the correlates of ERPs recorded early in infancy to language skills within toddlerhood and school-aged years. In a series of studies, Molfese and his colleagues demonstrated that ERP responses taken at infancy were able to strongly predict language and literacy ability for preschool and school age children (Molfese & Molfese, 1985; 1997; Molfese, 1995; 2000). Molfese and Molfese (1997) demonstrated that classification into
high functioning and low-functioning language ability at age five was possible based on ERP responses to speech syllables as newborns. The group differences in ERP components at birth were reflected in the large initial negative peak (N220) recorded over the left hemisphere and a second negative peak (N630), which occurred over both hemispheres. A discriminant function analysis predicted classification into either the high-functioning or low-functioning groups at age five based on standardized assessment with 80% accuracy. A subset of that same cohort was re-examined at age eight. N1 responses to syllables at birth discriminated between normal, poor and dyslexic readers at age 8 with 81.6% accuracy (Molfese, 2000). This evidence is also supported by other findings, which suggest that sensitivity to changes in phonological structures at birth differ in typically developing children and those with familial risk for impairment (Guttorm, Leppanen, Richardson, & Lyytinen, 2001).

There is evidence that between 35-60% of children who demonstrate slow emergence of language will eventually present normal expressive and receptive abilities by 3-4 years of age (Rescorla, Roberts, & Dahlgard, 1997; Thal & Tobias, 1992). This suggests that approximately 50% of late talkers will demonstrate persistent language deficits. Many of these children identified as late talkers are later classified as having SLI. There is an extensive body of research that has investigated the neural substrates of auditory processing using both speech and non-speech stimuli in school aged children with SLI which demonstrate abnormal processing in the SLI groups (Archibald & Joanisse, 2012; Ceponiene, Cunnings, Wulfeck, Ballantyne, & Townsend, 2009; McArthur & Bishop, 2004; Weber-Fox, Leonard, Hampton, & Tomblin, 2010). A recent ERP study by Archibald and Joanisse (2012) provides support for the PDT by examining the neural response to speech in co-articulation and lexical match/mismatch conditions in school aged children with SLI. Fifteen children (mean age, 8 years) with SLI and
15 typical peers were measured along four conditions of a picture word-matching task. Stimuli consisted of 60 CV or CVC words in which the initial sound was spliced to contain either valid or invalid co-articulatory information (e.g., initial /h/ in /hat/ contained /h/ spliced from another token of /hat/ for a valid co-articulation match, or contained /h/ spliced from /hot/ for an invalid co-articulation match). Therefore, the four conditions contained 1) lexical match/co-articulatory match, 2) lexical match/co-articulatory mismatch, 3) lexical mismatch/co-articulatory match, 4) lexical mismatch/co-articulatory mismatch.

The results showed that the SLI group demonstrated different patterns of ERP response when compared to typical children for the processing of co-articulatory, but not lexical information. The children with SLI showed atypical responses within the N1 component in which the initial sound of the word contained mismatching co-articulatory information despite being a lexical match to the target picture. Furthermore, a phonological mapping negativity (PMN) was only present within the SLI group when a lexical mismatch was present; however, the typical group demonstrated PMN for two conditions that presented mismatching co-articulatory information. Similar N400 responses to mismatch lexical information were found in the SLI and typical group.

The findings suggest that the children with SLI were sensitive to subtle changes in co-articulatory stimuli: the neural signature differed from that of typical peers. Unlike the control group, the children with SLI consistently showed a modulation of N1 to unexpected co-articulatory information. Furthermore, the inconsistent patterns of PMNs to co-articulatory mismatches were evidenced for the SLI group. It is possible that the increased sensitivity to acoustic variation in the N1 response within the SLI group could possibly detract from perceiving cues that are relevant to phonemic distinctions within their language. This suggestion
is supported by other evidence claiming that infants who are better able to home in on relevant
details of their native language while losing the ability to make distinctions between non-native
or irrelevant phonemic categories fare better in language skills within the toddler years (Kuhl,
Conboy, Padden, Rivera-Gaxiola, & Nelson, 2008; Kuhl, et al., 2006; Rivera-Gaxioloa, Silva-Pereyra, & Kuhl, 2005). Also, lack of PMN could suggest that children with SLI grapple with
use of sub-phonemic information in the speech stream to support rapid encoding of linguistic
information. It is also possible that children with SLI struggle with the mapping of acoustic
inputs onto phonological categories.

The aforementioned studies provide evidence that phonological processing as measured
by ERP can predict language performance along the developmental trajectory. These studies
also suggest that there are differences in the neural substrates that underlie phonological
processing when comparing typical children to children with language impairment. ERP
components are used to measure a particular neurophysiological response that reflects processing
of the experimental stimuli. The N1/P2 component has been cited as a measure of phonological
sensitivity (Hillard, Hink, Schwent, & Picton, 1973; Kutas, Van Petten, & Kluender, 2006). In a
paper from Dehane Lambertz (1997) the author showed that changes in P2 were evident only
across phonological boundary changes but not within category discriminations in adults.
Furthermore, Landi et al. (2012) used the N1/P2 component to record the neural response to
changing phonological stimuli in a large sample of 11-year-old children exposed to cocaine in
utero and typical control group. The typical group demonstrated greater amplitudes and faster
response times when compared to the cocaine exposed group. The N1/P2 component therefore,
becomes a quintessential tool when measuring the neural response to changes in phonological
stimuli.
Nonword Repetition

The computational model utilized by Joanisse and Seidenberg (2003) provides support for theories that link poor phonological working memory to language impairment. Nonword repetition (NWR) tasks are the closest researchers have come to developing a “gold standard” in capturing phonological working memory deficits in children with language impairments above the age of three. Nonword repetition tasks can vary in terms of syllable length, articulatory complexity, prosodic features and wordlikeness, and therefore are cited as measuring a variety of cognitive processes (Snowling, Chiat, & Hulme, 1991). However, nonword repetition tasks which vary in syllable length, are aimed at measuring phonological working memory and there is evidence that children with SLI perform poorly on this discrete parameter of nonword repetition (Archibald & Gathercole, 2007; Dollaghan & Campbell, 1998; Gathercole & Baddeley, 1990). There is also recent evidence suggesting that difficulties in nonword repetition among children with SLI stem from poor phonological representations (Ebbels, Dockrell, & van der Lely, 2012). Although underlying mechanisms leading to nonword repetition deficits continue to be debated, in general, deficits in nonword repetition remain a clinical marker for children with SLI. There is consistent evidence that nonword repetition has the ability to sensitively identify children with language impairment, independent of intelligence and socioeconomic status (Dollaghan & Campbell, 1998; Conti-Ramsden, Botting, & Faragher, 2001).

Recently, researchers have attempted to design nonword repetition tasks for children as young as two years-of-age (Clark, McRoberts, Van Dyke, Shankweiler, & Braze, 2012; Roy & Chiat, 2004). Given the poor diagnostic accuracy of standardized tests in identifying very young children who are at risk for language impairment, there is a need for more sensitive measures of language. Stokes and Klee (2009) investigated the sensitivity, specificity, positive/negative
likelihood ratios and diagnostic odds of a new Test of Early Nonword Repetition (TENR) on a sample of 232 British-English speaking children aged 27 (±3) months. The words were designed to include sounds within the phonetic inventory of very young children and also demonstrate low wordlikeness, while increasing in length from 1-4 syllables. The investigators concluded the TENR could be used for successful identification of two-year-old children at risk for language impairment as it demonstrated high correlations to parent report of vocabulary development and other standardized measures of vocabulary. The 1–4 syllable version of the TENR produced a positive likelihood ratio of 14.88; 95% (CI = 6.1–36.2) and a negative likelihood ratio of 0.13; 95% (CI = 0.02–0.83). Test sensitivity was 88% and specificity was 94%. The authors suggest the use of nonword repetition has promise in the identification of language impairment for very young children. In summary, given that both atypical perception and phonological working memory deficits have been implicated in language impairment, perhaps measurement of both these skills should be considered when identifying children at risk for impairment.

**Purpose and Hypothesis**

This study will explore the use of ERP and its association to language in young children. Prior to determining the ability for ERP to be used clinically within the toddler population in identifying impaired from unimpaired children, critical steps should be taken to investigate how phonological sensitivity skills relate to language outcomes in young children. By taking the initial step in determining if a relationship between phonological sensitivity as measured by ERP and language skills exist, we therefore lay the foundation for further investigation of ERP to be used as a clinical tool which can provide support in diagnosing impairment for children demonstrating language difficulties.
There is mounting evidence that ERPs have the potential to predict language skills in children at a later point of development (Kuhl, et al., 2006; Molfese & Molfese, 1985; 1997). Furthermore, there is consistent evidence suggesting that children with language impairment demonstrate atypical ERPs when compared to typical peers (Archibald & Joanisse, 2012; McArthur & Bishop, 2004; Weber-Fox, Leonard, Hampton, & Tomblin, 2010). The current project attempts to investigate how auditory sensitivity to phonological changes relates to language competence within a representative sample of children in the understudied toddler population.

In addition to use of ERP, nonword repetition will be utilized as a behavioral measure of phonological working memory. The use of nonword repetition provides a strong complement to the ERP work. There is preliminary evidence suggesting nonword repetition can be used to identify language delay in toddlers (Stokes & Klee, 2009). By examining the relationship between perceptual sensitivity measured by ERP and phonological working memory measured by nonword repetition, we can determine the collective usefulness of ERP and nonword repetition in identifying children with language impairment. We can also investigate the validity of the connectionist account of phonological working memory put forth by Joanisse and Seidenberg (2003), which claims that phonological representations and phonological working memory are indistinct cognitive processes. By determining if ERP response, as a measure of phonological sensitivity, has an independent contribution to predicting language over and above that of phonological working memory measured we can test whether predictions from the connectionist account of phonological processing are supported by data from young children learning language. This investigation will explore the following aims:
1) The first aim is to determine if the neural response to repeated spoken disyllabic speech tokens of nonwords within an ERP task modeled after Molfese, Morse, and Peters (1990) and Landi, Crowley, Wu, Bailey, and Mayes (2012) is a robust indicator of language competence within toddlers and young children. Two particular components of interest that will be examined, are represented in the N1/P2 complex. The N1/P2 complex is associated with “lower level” auditory sensitivity to changes in acoustic parameters and has also been cited as a measure of phonological sensitivity and rhyme detection (Hillard, Hink, Schwent, & Picton, 1973; Kutas, Van Petten, & Kluender, 2006). Based on ERP research in infants, it is predicted that ERP measurements of phonemic sensitivity will have a significant relationship with language skill measured by clinical assessments.

2) The second aim is twofold. First the perceptual deficit theory will be investigated by determining if indeed phonological sensitivity and phonological working memory predict language competence in the young child population. Secondly, this investigator will attempt to test the computational framework by determining if phonological sensitivity, as measured by ERP and phonological working memory as measured by NWR, will each uniquely contribute to language skill, or explain a significant amount of the variance in language ability separately. It is predicted that both phonological sensitivity and phonological working memory will contribute to language competence but each process will be able to explain a unique portion of variance in language skill separately.

**Methods**

**Participants**

Participants were a subgroup of 80 total children from the Haskins laboratories pilot study of the Language and Early Assessment Research Network (LEARN), which assessed
neurobiological markers of speech perception and production. For the current study, a sample of forty children between the ages of 24 to 48 months (22 male) were recruited from local university clinics, private practices, the Rhode Island Birth-to-Three system and the Connecticut Birth-to-Three system. This subgroup was chosen based on completion of the ERP experiment plus nonword repetition task. All children met the following criteria to be included in the study: 1) monolingual English speakers 2) no known psychiatric or neurological deficits per parent report 3) hearing was within normal limits at the time of the study per parent report. All children were reported to have passed newborn hearing screenings. A distortion product otoacoustic emission-screening test (DPOAEs) was performed on a subgroup at the time of ERP recordings (N= 15). One child failed the screen and was seen for a follow up audiological evaluation, which reported normal hearing acuity. An independent sample t-test was conducted to compare ERP response (specifically N1/P2 amplitude differences in the visually inspected large electrode cluster) between the DPOAE and parent report matched group to rule out hearing acuity effects on the ERP results. There was no significant difference between the N1/P2 amplitude differences between the DPOAE (M=2.32, SD=2.52) and the parent report group (M=1.60, SD=1.64); t(28) =0.93, p=0.36. Therefore, we assume that the children results of the ERP recording for both the otoacoustic emissions group and the parent report group are commensurate. ERP results between the 2 groups will be combined for ERP analyses.

Under the assumption that language skill is a continuous construct, the data were treated as such. To account for a representative sample of young children, 10% of the sample included children demonstrating language delay (4 participants) (Rescorla, Roberts, & Dahlsgard, 1997). By accounting for children with language delay, we treated the data as a continuous variable; the variance within the language abilities of the participants was preserved, as it is within the
population. Standardized assessments were used to provide descriptive data regarding the participants’ language abilities. Children demonstrating language delay were considered by demonstrating a standard score of < 85 on the expressive and/or receptive portion of the Preschool Language Scale-Fifth Edition (PLS-5: sensitivity = .93, specificity = .78 for ages 0-3.11) (Zimmerman, Steiner, & Pond, 2011). Two children with language delay also demonstrated below average scores on the visual reception subtest of the Mullen Scales of Early Learning (Mullen, 1995). Typically developing children demonstrated no history of speech and language services and met all developmental milestones within the average range as indicated by parent report. Typically developing children were considered as having average receptive and expressive functioning on the PLS-5 as well as average visual reception on the Mullen Scales of Early Learning (see table 1 for mean standard scores on behavioral assessments).

The typically developing children (n=36) included 31 Caucasian participants, 2 African American participants, and 3 Asian/Pacific Islander participants. Children with language delay included 3 Caucasian participants and 1 African American participant. All four participants with language delay were male.

**Procedures**

Parents completed a background questionnaire regarding medical history (including audiological history) as well as information on motor and language developmental milestones. The children participated in 1 – 2, 120 minute sessions. Aforementioned standardized measures and conventional language sampling procedures were used to determine language functioning. Next, the child participated in the ERP task and the administration of the TENR. Children were provided with breaks and reinforcements (e.g., small edibles, stickers, books) as needed. Upon completion of the experiment, a cohort of children wore the digital language processor (DLP) of
the LENA system within the home environment as explained in the previous section. The investigator provided the parent with a specialized vest, which included a protective pocket for the processor. The parents returned the processor within two weeks of the testing session. Parents were provided with a research report regarding performance on language measures. All participating families were compensated $20 per hour for their time and provided with travel expense money. Additional compensation ($20 total) was provided for families that took part in the LENA home recordings.

**Behavioral Language Measurement**

**Parent report.** The MacArthur-Bates Communication Development Inventories- Second Edition (mCDI-2; Fenson, et al., 2007) was utilized as a parent report of vocabulary development. The mCDI-2 compares responses regarding the child’s language skills to information gathered from a large sample of children learning English throughout the United States. The mCDI-2 sections yield percentile ranks and percentages of affirmative answers based on the child’s age given the responses of the parent or caregiver. Parents measure vocabulary by marking a set of words from a listed pre-determined set of vocabulary outlined in the mCDI-2 form. Raw scores were used as a measure of vocabulary production based on the age of some of the participants extending beyond that of the normative data.

**Language sampling.** A language sample of approximately 50 utterances was collected for each participant (Heilmann, Nockerts, & Miller, 2010). The examiner used conventional language sampling procedures within a play-based communicative exchange to gather a representative sample of the child’s language. Graduate and undergraduate students trained in language sampling and analysis transcribed the language samples. Reliability checks were performed on 20 randomly selected participants (50% of the participant pool) and were found to
be 0.86. Computerized Profiling v9.7 (Long, 2008) was used to analyze the transcriptions. A Language Assessment Remediation and Screening Procedure (LARSP: Crystal, Fletcher, & Garman, 1989) provided a mean length of utterance in morphemes (MLU), which was used as a behavioral measure of morphological development and utterance length. The Profile of Phonology (PROPH; Crystal, 1982) analysis was also used to provide the percent consonants correct (PCC) within the language sample. The PCC was used as one of the language variables given the strong correlations between phonology and language among young children.

Given the constraints of time and unfamiliarity associated with the lab setting, it was acknowledged that some children, especially children with language delay, might not provide a robust representative sample during the experiment. Therefore, the Language Environmental Analysis System (LENA) (LENA Foundation, 2014) was used with those children who demonstrated limited language skills within the laboratory setting to collect a representative sample of language within the child’s naturalistic environment. The child was equipped with a (DLP), which collected data as the child interacted with a caregiver within the home during a play period. The LENA software allowed the examiner to view child vocal output throughout a given time period within the day. A random sampling of 5-minute intervals was collected, transcribed and analyzed similarly to that of the laboratory samples.

To account for experimental confounds which may occur by providing LENA to only a particular cohort of children within the sample (meaning that the language sample taken at the laboratory may be in essence different than that taken within the home) a group of children (n=10) provided both a laboratory language sample and a LENA home sample. A paired samples t-test was run to determine if there were significant differences between the MLU collected in the lab compared to home. There was no significant difference in MLU scores.
collected from the home (M: 2.72, SD: 0.90) compared to the MLU collected at the lab (M: 2.48, SD: 1.11); t(9) = 1.29, p < 0.23. These results suggest no significant difference in scores based on the environment of the sample. Therefore, we assumed that the transcript data among home and lab transcripts was similar and representative of true language ability. The total number of children providing a home sample is 7. If both a lab sample and home sample were collected, the laboratory sample was used to preserve consistency.

**Standardized assessment.** The PLS-5 (Zimmerman, Steiner, & Pond, 2011) is an individually administered standardized language assessment designed for children from birth to age 7;11 to assess language skill. The PLS-5 was utilized to provide information on global language functioning. Both the auditory comprehension and expressive communication portions were administered. The PLS-5 provided a broad measure of language functioning in phonologic, semantic, morpho-syntactic, and pragmatic domains.

The GFTA-2 (Goldman & Fristoe, 2000) was also administered to gain information regarding the child’s articulation abilities. Since speech skill and language are highly correlated within early years (Paul & Jennings, 1992), the GFTA-2 raw score will be considered as one variable within the behavioral language tests, along with PCC. The GFTA-2 also provides information regarding articulatory errors, which will be accounted for when scoring the TENR. The visual reception portion of the Mullen Scales of Early Learning (Mullen, 1995) was also administered to provide information regarding the participant’s non-verbal cognitive skill. The visual reception scale measures nonverbal skills as they pertain to patterns, memory and sequencing. Standardized procedures were followed for all assessments as indicated by the manual.

1 One participant did not complete testing with the GFTA-2. A phonological analysis was performed using his PCC data to account for substitutions produced on the TENR
**Factor analysis.** Given that most behavioral measurements for young children are limited in sensitivity, the factor analysis provides a variable that encompasses similarities among multiple language and speech measures. Factor analysis is a statistical method used to derive the shared variance among multiple variables and reduce them to a lower number of variables, termed factors. The language factor served as the dependent variable in a multiple regression analysis in which ERP, nonword repetition and age act as the independent variables or predictors. The following measures were included in the factor analysis: 1) total number of words reported by the parent of the mCDI-2, 2) raw score of the auditory comprehension portion of the PLS-5, 3) raw score of the expressive portion of the PLS-5, 4) the number of errors produced on the GFTA-2, 5) MLU, and 6) PCC. Raw scores on the PLS-5 were used due to other non-standardized variables in the analysis such as total words reported on the mCDI-2, MLU, and PCC.
Table 1: Assessment scores for typically developing and language delayed children

<table>
<thead>
<tr>
<th></th>
<th>Full Sample (n=40)</th>
<th>Typically Developing (n=36)</th>
<th>Language Delay (n=4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in months</td>
<td>34.35 (6.29)</td>
<td>34.06(6.42)</td>
<td>37 (4.12)</td>
</tr>
<tr>
<td>mCDI-2-WP</td>
<td>528.28(160.89)</td>
<td>556.2 (128.01)</td>
<td>284 (206.17)</td>
</tr>
<tr>
<td>PLS-AC</td>
<td>110.78 (12.17)</td>
<td>113 (10.51)</td>
<td>90.75 (7.15)</td>
</tr>
<tr>
<td>PLS-EC</td>
<td>109.03 (14.62)</td>
<td>112.03 (11.97)</td>
<td>81.50 (2.29)</td>
</tr>
<tr>
<td>GFTA-2</td>
<td>108 (14.56)</td>
<td>110.61 (11.85)</td>
<td>76.67 (2.36)</td>
</tr>
<tr>
<td>MLU</td>
<td>2.79 (1.16)</td>
<td>2.90 (1.16)</td>
<td>1.79(0.40)</td>
</tr>
<tr>
<td>PCC</td>
<td>81% (0.13)</td>
<td>83% (0.12)</td>
<td>65% (0.13)</td>
</tr>
<tr>
<td>MSEL_VR</td>
<td>60.51(12.93)</td>
<td>62.49 (11.68)</td>
<td>43.25(10.21)</td>
</tr>
<tr>
<td>TENR_T</td>
<td>102.53 (26.74)</td>
<td>104.94(26.13)</td>
<td>80.75 (21.48)</td>
</tr>
</tbody>
</table>

Nonword Repetition and ERP

Nonword repetition. The TENR (Stokes & Klee, 2009) was administered to measure phonological working memory skills within the participants. The TENR is designed to include phonemes that are typically included in the inventories of 2 year-old children. The assessment contains 1, 2, 3 and 4 syllable nonwords (4 tokens of each syllable type) that are consistent with British-English trochaic stress and wordlikeness. There were a total of 16 nonwords comprised of 90 phonemes for the entire test. Modifications to particular phonemes and stress patterns were made to ensure the stimulus is consistent with American English (see appendix 1). All stimuli were recorded and presented at the maximum volume level of the Dell computer (approximately 60dB) within a computerized PowerPoint presentation to ensure consistency within the stimuli. Each power point slide depicted a friendly alien character with an (nonword) alien name. Children were given the following simple directions; “Let’s play a game. Listen carefully and say just what I say”. The children were to repeat the alien names following the voice in the power point slide. A practice item was administered so that the examiner could provide feedback regarding directions. The examiner repeated the nonword verbally if the child did not respond within 5 seconds. This is standard practice for nonword repetition tasks for this age group. Participant productions were recorded by a (Sony) digital audio recorder with an internal microphone. Children were awarded one point for each syllable produced, and one point for each vowel and consonant produced correctly. Then a total score was calculated by adding the total number of syllables and total phonemes produced correctly. This scoring procedure was adopted to prevent floor effects and provide a more comprehensive scale for scoring. By providing credit for syllable preservation, children with significant articulation difficulties may still be able to demonstrate memory for word parts. Prior to scoring the TENR, a phonological analysis of the
child’s speech was performed using the Goldman Fristoe Test of Articulation – Second Edition (GFTA-2; Goldman & Fristoe, 2000). Each consistent substitution error produced on the GFTA_2 was accounted for and given credit on the TENR. If a phoneme was deleted on the TENR, it was counted as an error. This analysis is consistent practice for nonword repetition scoring in young populations (Stokes & Klee, 2009). Reliability measures on the final scoring of the nonword repetition task were found to be 0.81.

**ERP procedures.** Children were fitted with a 128-sponge Ag/AgCl electrode high-density sensor array net (EGI, Inc.) that was used to acquire electrophysiological data. Prior to placement, the net was soaked for 10 minutes in a warm potassium and chloride (KCl) solution to improve conductance. The net was placed on the head using standard procedures outlined by EGI Inc. (Dien, 2010). EEG data were recorded using Netstation v. 4.5 software (EGI Inc.) with an EGI Net Amps 3 high impedance amplifier, at a sample rate of 500Hz. All electrode impedances remained under 40kohms as indicated by impedance measures made immediately before and after the test sessions. The child sat on the parent/caregiver’s lap in a comfortable chair. In front of the child was a computer screen and next to it a small portable DVD player located 50 inches from the child. The DVD player displayed a silent movie (clips of Yo Gabba Gabba puppets) that facilitated compliance and provided non-auditory stimulation.

**ERP task.** Participants were presented with two rhyming nonword tokens of speech, /bidu/ and /gibu/, in an old/new design. This task has been used previously to examine speech perception ability and nonword learning in adolescents exposed to cocaine in-utero (Landi, Crowley, Wu, Bailey, & Mayes, 2012) and infants (Molfese, Morse, & Peters, 1990). The auditory stimulus is presented via an overhead speaker positioned above the participant (distance from the floor to the speaker 190 cm) presented at 85dB SPL. The first block is a sensitization
block, which consists of one token (/gibu/), repeated for 50 trials. The second block is a mixed block where the tokens, (/bidu/), and (/gibu/) are randomly presented in equal proportions. There were 100 trials (50 /bidu/ and 50 /gibu/) within the second block. There was a 20 second rest-delay between the first and second block. The stimuli were designed so that the sensitization block stimulus (/gibu/) acted as the “old” stimulus in block 2 and the second stimulus in block 2 (/bidu/) acts as the “new” stimulus. The stimulus duration for each token is 595ms with a varied ISI of 1800 or 2800ms to avoid habituation. E-prime v.2.0 (PST, Inc.) was used to control stimulus presentation and time lock the stimulus to Netstation softwarecollected. Once the cap was prepared, the experiment took approximately 10 minutes. Children were rewarded with a small prize for their participation following the experiment.

**Data Analysis**

**ERP processing.** Data were filtered to retain signal frequencies between 1 and 30Hz. ERP Data were segmented into 700ms epochs including 100ms pre-stimulus baseline and a 600ms post-stimulus interval. After filtering and segmentation, data were visually inspected to identify poor electrodes. Automated routines were used to further detect bad electrodes and eye movement/blink artifact (bad electrode > 200µV, eye blink/eye movement > 150µV). If an electrode was bad for more than 40% of the segments then it was marked bad for the entire file. If a segment contained more than 10 bad electrodes then the segment was marked as bad. Bad electrodes were replaced using spherical spline interpolation (Perrin, Pernier, Bertrand, & Echallier, 1989). The data were re-referenced to the average reference (vertex reference, Cz, was used during recording) and baseline corrected to 100ms pre-stimulus presentation (Junghofer, Elbert, Tucker, & Braun, 1999). Finally, artifact free segments were averaged within the old and new conditions. A criteria of at least 20 preserved trials for each condition was used to include
subjects in the ERP analysis (Landi, Crowley, Wu, Bailey, & Mayes, 2012). There was no significant difference within the averaged new condition (M: 33.40, SD: 6.71) compared to the old condition (M: 32.25, SD: 6.67); t(39) = 1.58, p< 0.12. Ocular Artifact Correction (OAR) (Blink Slope Threshold = 14µV/ms) (Gratton, Coles, & Donchin, 1983) was conducted on 6 participants due to less than 20 blink or artifact free trials per condition prior to OAC. All forty participants are included in the ERP analysis.

**ERP analysis method.** Two sets of analyses were conducted. First, data were visually inspected for peak identification of the N1/P2 complex. Electrodes and time windows of interest were chosen based on previous literature (cf. Landi et al. 2011; Wunderlich & Cone-Wesson, 2006). The N1/P2 complex was identified visually in a cluster of electrodes in the medial-parietal cortical region (see Figure 1 for electrode montage). Within this cluster, peaks were identified as the most negative peak occurring between 50-150ms post stimulus onset (N1) and the most positive peak occurring from the next150-300ms (P2) (see Figure 2 for ERP waveforms).

The combined amplitude of the N1/P2 complex was taken by subtracting the amplitude of N1 from the amplitude of P2. The amplitude difference effect of the old relative to the new condition was then derived by subtracting the N1/P2 amplitude of the new condition from the old condition (i.e. new-old) for each participant. The average combined amplitude for the N1/P2 complex within the new condition for the visually inspected cluster was 5.5µV. The average combined amplitude of the N1/P2 complex within the old condition was 6.4µV. These differences in amplitude between the new condition (M: 5.5µV, SD: 2.5) and the old condition (M: 6.4, SD: 3.2) were statically significant: t (39)= 2.32, p< .03. This suggests that the combined amplitude for the old condition was greater than that for the new condition.
The combined latency for the N1/P2 complex was derived similarly to that of the combined amplitude. The latency of N1 was subtracted from the latency of P2 within the new and old conditions. The average combined latency of the N1/P2 complex within the new condition was 111.39ms. The average combined latency of the N1/P2 complex within the old condition was 125.51ms. The differences in latency of the N1/P2 complex between the new condition (M: 111.39, SD: 36.14) and the old condition (M: 125.51, SD: 35.16) were statically significant: t(39) = 2.53, p<0.02. This suggests that the latency of the new condition was faster or occurring within an earlier time frame for the new condition relative to the old condition.

Therefore, to preserve directionality within the analysis and adhere to the parameters of the latency time frame, the latency difference effect of the old relative to the new condition was then derived by subtracting the N1/P2 latency of the old condition from the new condition (i.e. old-new) for each participant.

For the second set of analyses, EEG data were submitted to a temporal/spatial principal components analysis (PCA) to identify temporal and spatial factors of interest using the ERP PCA Toolkit (Dien, 2010). The purpose of the PCA was to identify systematic variance within the temporal domain in the absence of stimulus condition. The PCA divides the ERP into a smaller number of uncorrelated components while accounting for the maximum level of variance. This data driven approach also facilitates comparisons of ERP data across different developmental populations (Molfese, Nunez, Seibert, & Ramanaiah, 1976). Given the limited literature on ERP within the toddler population, the PCA was used to extract significant time factors above and beyond that of conventional ERP components described in the literature, which may provide different insight into the neural mechanisms associated with an emerging language system.
First a temporal PCA was conducted with promax (oblique) rotation to identify time windows of interest. Although PCA temporal factors are active over the course of the entire ERP average, a loading criterion of 0.6 was used to identify time windows when the factors were most active (Dien, 2010). Ten temporal factors were extracted from the PCA using a scree test (Cattell, 1966) which accounted for 97% of the total variance within the ERP signal. Following the temporal PCA, a spatial PCA with infomax rotation was then run on each temporal factor to identify electrodes that loaded strongly within each time window. These spatial factors were used based on the amount of variance explained in each (above 5%) as well as their orientation on the scalp which coincided with the general parameters for recording the auditory evoked response.

There were four temporal factors with variance above 5%. Temporal factor 1 accounted for 27% of the variance and encompassed a time from 544-700ms post stimulus onset. Temporal factor 2 accounted for 17% of the variance and included the time window of 248-360ms. Temporal factor 3 accounted for 13% of the variance and encompassed between 404-500ms. Finally the fourth temporal factor ranged from 136-220ms and counted for 5% of the variance. The adaptive mean amplitudes and peak latency from the electrodes in the first through forth-spatial factor were submitted for statistical analysis for each temporal/spatial factor pairing.

To summarize the PCA analysis, 4 temporal factors were extracted based on a threshold variance of 5%. For each of the 4 temporal factors, 4 spatial factors were retained which coincided with parameters for recording the auditory evoked response. For each temporal/spatial factor pairing, adaptive mean amplitude and peak latency data within the new and old conditions were taken. Analyses similar to those used for the large cluster average were performed, such that amplitude difference effect of the old relative to the new condition was derived by subtracting the mean amplitude between the two conditions (i.e. new-old). The latency
difference effect was derived by subtracting the latency of the new condition from the old condition to preserve a positive difference within a latency time frame (old-new). Given the high number of variables extracted from the PCA (total of 15 temporal spatial pairings each comparing both amplitude and latency differences for a total of 30 variables) correlations were conducted first exploring relationships with the LFS to avoid a type one error. If a significant relationship was found with the LFS, additional correlations with individual language assessments were explored.

Results

Analysis of the Language Factor Score

A factor analysis was conducted to summarize the behavioral language variables (mCDI-2-WP, PLS-AC, PLS-EC, GFTA, MLU, PCC). GFTA-2 results were inverted (added a negative sign to each raw score) to maintain similar interpretability and similar direction of the other responses. Oblique rotation was used as correlated factors were expected. A scree plot suggested the presence of one latent factor with an Eigen value above 1. The sum of all prior communality estimates is 3.87, which is the estimate of the common variance among all subtests. This initial estimate of the common variance constitutes approximately 65% of the total variance present within the first latent factor. The residual correlation matrix and partial correlation matrix were both less than 0.05 suggesting the factors were justified in explaining the data (see table 2 for factor loadings). The variables for factor 1 were saved for each subject and the variables
constitute the Language Factor Score (LFS) for each participant. The LFS became the dependent variable for all further analyses.

Two children did not complete a full battery of testing, therefore the LFS is derived from 38 out of the 40 participants.
Table 2 Factor Pattern for Language Factor Score (LFS)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Factor Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLS-AC</td>
<td>0.90</td>
</tr>
<tr>
<td>PLS-EC</td>
<td>0.89</td>
</tr>
<tr>
<td>MLU</td>
<td>0.84</td>
</tr>
<tr>
<td>GFTA-2</td>
<td>0.77</td>
</tr>
<tr>
<td>PCC</td>
<td>0.73</td>
</tr>
<tr>
<td>mCDI-2-WP</td>
<td>0.67</td>
</tr>
</tbody>
</table>

**N1/P2**

Partial correlation analyses, controlling for age, were performed to determine the relationship among the N1/P2 complex, specifically the amplitude and latency differences between the new and old conditions and the LFS (see Table 3). There was a positive partial correlation between the N1/P2 amplitude difference and the LFS suggesting that as language scores of the participants increased, the amplitude in response to the new stimulus was increasing with respect to the old. N1/P2 amplitude difference explained approximately 14% of the variance within the LFS when controlling for age ($r^2 = 0.14$). Additional analyses were also conducted to establish the relationship between amplitude difference and the individual language measures (see Table 3 for results). No significant correlations were present between latency differences of the N1/P2 complex and language scores within this visually inspected electrode cluster.
To address research aim 2 and determine if amplitude differences between new and old conditions within the N1/P2 complex predicted language ability separately from NWR, a regression analysis was then conducted to predict the LFS from N1/P2 amplitude difference, NWR and Age (see Table 4, No. 1). The model including the three factors was significant. N1/P2 amplitude difference, NWR and Age accounted for 67% ($R^2 = 0.67$) of the variance in the LFS. N1/P2 amplitude difference approached significance when predicting the LFS. Even though the p value for the N1/P2 predictor is 0.06, given the amount of children in the sample, the amount of predictors in the model and the degrees of freedom used for the analysis, it is determined that the alpha level of 0.06 is large enough to interpret a significant relationship. NWR and age significantly predicted the LFS. When examining predictors, Age was the strongest predictor in the model ($\beta = 0.52$), followed by NWR ($\beta = 0.50$) and finally N1/P2 amplitude difference ($\beta = 0.20$). This suggests that N1/P2 explains a unique portion of variance in the LFS over and above that of NWR and Age.

Figure 1: Electrode montage for N1/P2
Figure 2: Waveforms of N1/P2 at electrode Pz (62)

Figure 2: Averaged ERP waveforms for both old and new tokens of 40 visually inspected large cluster with Principal Components Analysis

The principal components analysis yielded two significant temporal which coincided within our theoretically driven timeframes of interest (i.e. perceptual sensitivity to phonological changes and memory) and were located in areas that are associated with recording of auditory ERPs.

**Principal Components Analysis**

The principal components analysis yielded two significant temporal-spatial pairings, which coincided within our theoretically driven timeframes of interest (i.e. perceptual sensitivity to phonological changes and memory) and were located in areas that are associated with recording of auditory ERPs.

**Temporal factor 4 (136-220ms).** Temporal factor 4_Spatial factor 2 (TF4_SF2) loaded onto a cluster of 10 electrodes located within the left temporal region and encompassed a time frame between 136-220ms (see Figure 3 for electrode montage and Figure 4 for waveforms). This timeframe is similar to that of the N1/P2 complex. The average amplitude for TF4_SF2 within the new condition was 1.64 µV. The average amplitude for TF4_SF2 within the old
condition was 1.80µV. The differences in amplitude between the new condition (M: 1.64µV, SD: 2.44) and the old condition (M: 1.80, SD: 2.25) were not statistically significant: t (39) = -0.33, p<0.75. The average latency of TF4_SF2 within the new condition was 182.05ms and within the old condition was 186.23ms. The differences in latency in TF4_SF2 between the new condition (M: 182.05, SD: 21.81) and the old condition (M: 186.23, SD: 21.33) were not statistically significant: t(39) = -1.06, p<0.30. There was a positive correlation, controlling for age, between the differences in latency within the new and old condition and the LFS. (see Table 3 for correlations). These correlations suggest that as language skills increased, so too did the difference between the old and new condition such that the response recorded for the new condition was “faster” than the response for the old.

To address the question of whether ERP measures of phonemic sensitivity to changing phonemic stimuli explain a significant amount of variance within language separate for NWR, a regression analysis was conducted to predict the LFS from the latency differences in TF4_SF2, NWR and Age (see Table 4, No 5). The model including the three factors was significant. Latency differences in TF4_SF2, NWR and Age accounted for 67% (R² = 0.67) of the variance in the LFS. TF4_SF2 latency difference approached significance when predicting the LFS (p=0.06). NWR and Age significantly predicted LFS. As can be seen by the beta weights, Age is the strongest predictor of language skills (β = 0.52), followed by NWR_T (β = 0.48) and finally TF4_SF2 (β =0.20). This suggests that phonemic perception being measured within TF4_SF2 explains a unique amount of variance in the LFS separately from NWR and Age.
Figure 3: Electrode Montage for TF4_SF2 (136-220ms)
Temporal factor 2 (248-360ms). Temporal factor 2_Spatial Factor 1 (TF2_SF1) loaded onto a cluster of 35 electrodes located in the midline-frontal cortical region and encompassed a time frame between 248-360ms (see Figure 5 for electrode montage and Figure 6 for waveforms). The average amplitude for TF2_SF1 within the new condition was 1.86 µV. The average amplitude for TF2_SF1 within the old condition was 1.43µV. These differences in amplitude between the new condition (M: 1.86µV, SD: 2.50) and the old condition (M: 1.43, SD: 3.36) were not statically significant; t (39)= 0.82, p< 0.42. There was no significant correlation between amplitude difference and the LFS within TF2_SF1.

The average latency of TF2_SF1 within the new condition was 308ms and within the old condition was 317ms. The differences in latency between the new condition (M: 307.77, SD: 24.32) and the old condition (M: 316.77, SD: 20.50) were statically significant: t (39) = -2.17, p<0.04; therefore, the peak latency in response to the new stimuli occurs “faster” than the old
stimuli. There were positive correlations, controlling for age, between the difference in latency between the new and old condition with the LFS, as well as all of the individual language measures (see correlations Table 3). This suggests that faster responses for the new stimuli are associated with greater language ability.

To determine if the latency independently predicted the LFS beyond the effects of from NWR and Age, a regression analysis was conducted to predict the LFS from the TF2_SF1 latency difference, NWR and Age (see Table 4, No.2). Given the high multicolinearity among the predictors in this model, specifically, between the ERP data and NWR (r = 0.50, p=0.00), the model could not be interpreted adequately. Therefore, independent regressions were run to determine the amount of variance explained within the LFS for both TF2_SF1 latency difference separately from NWR.

When predicting the LFS from TF2_SF1 latency difference and Age, the model was significant (see Table 4, No. 3). TF2_SF1 latency difference and Age accounted for 50% (R^2 = 0.49) of the variance in the LFS. Both TF2_SF1 latency difference and Age significantly predicted the LFS. When examining each of the individual predictors, Age was a stronger predictor (β =0.55) than TF2_SF1 latency difference (β = 0.40). Another regression was then run predicting the LFS from NWR and Age. As seen in Table 4, No 4, the model was significant. Both NWR (β=0.54) and Age (β= 0.52) significantly predicted the LFS.
Figure 5: Electrode Montage for TF2_SF1 (248-360ms)
ERP, NWR & LANGUAGE

Figure 6: ERP Waveforms for TF2_SF1 at electrode Fz (11)

Figure 4: Averaged ERP waveforms for both old and new tokens within TF2_SF1

**ERP Time Windows and the LFS**

It was hypothesized that the perception of phonological features measured by ERP would uniquely predict a significant proportion of variance separate from phonological working memory measured by NWR. Our results thus far suggest that ERP recordings within the N1/P2 complex and TF4_SF2 (136-220ms) do indeed predict language skill separately from NWR. However, ERP recordings within TF2_SF1 (248-360ms) when included in a model with NWR and Age were unable to explain a significant proportion of variance due to the high correlations among the factors. It is possible that the time window for TF2_SF1 is measuring an aspect of phonological working memory similar to that of NWR. Therefore, two additional regression analyses were performed to pit the ERP data from the N1/P2 and TF4_SF2 against that of
TF2_SF1 to determine if in fact; these different time windows (early verses late) were actually measuring different language processes. It is possible that the earlier time windows of the N1/P2 complex and TF4_SF1 are measuring phonological differences among old and new stimuli whereas ERPs within TF2_SF1 capture phonological working memory for word stimuli.

**N1/P2 & TF2_SF1.** A regression analysis was conducted to predict the LFS from the amplitude differences in N1/P2, the latency difference within TF2_SF1 and Age to address the question if these separate ERP time windows are able to uniquely explain a significant proportion of variance within the LFS (see Table 4, No 6). When N1/P2 amplitude difference was included in the model with TF2_SF1 latency difference, the N1/P2 amplitude difference was not significant when predicting the LFS. TF2_SF1 latency difference and Age significantly predicted LFS. When examining predictors, Age is the strongest predictor of language skills ($\beta = 0.57$), followed by TF2_SF1 latency difference ($\beta = .33$). There was a strong correlation between N1/P2 amplitude difference and TF2_SF1 latency difference ($r =.34$, $p<.04$). N1/P2 amplitude difference was unable to explain a significant proportion of variance independent from TF2_SF1.

**TF4_SF2 (136-220ms) & TF2_SF1 (248-360ms).** A regression analysis was conducted to predict the LFS from the latency differences in TF2_SF1, TF4_SF2 and Age to determine if the ERP recordings represented different neural processes (see Table 4, No 7). TF4_SF2 latency difference approached significance when predicting the LFS ($p< 0.066$). TF2_SF1 latency difference and Age significantly predicted LFS. When examining predictors, Age was the strongest predictor of language skills ($\beta = 0.55$), followed by TF2_SF1Ldif ($\beta = 0.31$) and finally TF4_SF2 ($\beta = 0.24$). These results suggest that TF2_SF1 and TF4_SF2 are each representing a unique portion of variance in explaining the LFS and therefore contributing differently to
language skill. It is possible that the processing represented in the earlier timeframe of TF4_SF2 (136-220ms) is sensitive to the changes in the phonetic features of the stimuli whereas the later timeframe of TF2_SF1 (248-360ms) is measuring memory processes.

**N1/P2 & TF4_SF2 (136-220ms).** A regression analysis was performed to predict the LFS from the amplitude difference in in N1/P2 and the latency difference within TF4_SF2 while controlling for Age to determine if each of these predictors are able to explain a significant portion of variance within the LFS independently. The model including the three factors was significant (see Table 4, No 8). TF4_SF2 latency difference and Age were able to explain a significant portion of variance; however, N1/P2 amplitude difference was not significant. N1/P2 and TF4_SF2 were highly correlated predictors and (p<.04) and therefore may not be able to explain the LFS independently. TF4_SF2 was the strongest predictor in the model ($\beta = .273$) followed by age ($\beta =0.58$).
Table 3: Partial correlations among language measures and ERP

<table>
<thead>
<tr>
<th>Measure</th>
<th>NIP2</th>
<th>N1/P2</th>
<th>TF2_SF1</th>
<th>TF4_SF1</th>
<th>TF4_SF2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adif</td>
<td>Ldif</td>
<td>Adif</td>
<td>Ldif</td>
<td>Adif</td>
</tr>
<tr>
<td>mCDI-2-WP</td>
<td>0.34*</td>
<td>0.01</td>
<td>-0.05</td>
<td>0.45**</td>
<td>0.05</td>
</tr>
<tr>
<td>PLS-AC</td>
<td>0.27</td>
<td>-0.11</td>
<td>-0.04</td>
<td>0.41**</td>
<td>-0.03</td>
</tr>
<tr>
<td>PLS-EC</td>
<td>0.29</td>
<td>-0.18</td>
<td>-0.10</td>
<td>0.39*</td>
<td>0.02</td>
</tr>
<tr>
<td>GFTA-2</td>
<td>0.41*</td>
<td>-0.12</td>
<td>-0.37*</td>
<td>0.37*</td>
<td>0.13</td>
</tr>
<tr>
<td>MLU</td>
<td>0.25</td>
<td>-0.21</td>
<td>-0.17</td>
<td>0.35*</td>
<td>0.07</td>
</tr>
<tr>
<td>PCC</td>
<td>0.25</td>
<td>-0.25</td>
<td>-0.24</td>
<td>0.33*</td>
<td>0.18</td>
</tr>
<tr>
<td>LFS</td>
<td>0.37*</td>
<td>-0.18</td>
<td>-0.21</td>
<td>0.49**</td>
<td>0.94</td>
</tr>
</tbody>
</table>

*p<0.05,  **p<.01: Adif = amplitude difference between old and new tokens, Ldif = latency difference between old and new tokens
Table 4: Multiple regressions predicting the LFS from ERP data, NWR and Age

<table>
<thead>
<tr>
<th>No.</th>
<th>$R^2$</th>
<th>$F$</th>
<th>$p$</th>
<th>Model function</th>
<th>Independent Variables</th>
<th>Variable</th>
<th>$\beta$</th>
<th>$t$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>.67</td>
<td>22.53</td>
<td>.00</td>
<td>Control</td>
<td>Age</td>
<td>.54</td>
<td>5.37</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phono Working memory</td>
<td>NWR</td>
<td>.50</td>
<td>4.87</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phono Sensitivity</td>
<td>N1/P2 Amplitude dif</td>
<td>.20</td>
<td>1.93</td>
<td>.06</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>.65</td>
<td>21.28</td>
<td>.00</td>
<td>Control</td>
<td>Age</td>
<td>.52</td>
<td>5.11</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phono working memory</td>
<td>NWR</td>
<td>.46</td>
<td>3.96</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phono sensitivity</td>
<td>TF2_SF1 Latency dif</td>
<td>.18</td>
<td>1.51</td>
<td>.14</td>
<td></td>
</tr>
<tr>
<td>No.</td>
<td>$R^2$</td>
<td>$F$</td>
<td>$p$</td>
<td>Variable</td>
<td>$\beta$</td>
<td>$t$</td>
<td>$p$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-------</td>
<td>------</td>
<td>------</td>
<td>---------------------</td>
<td>---------</td>
<td>-------</td>
<td>------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>.49</td>
<td>16.97</td>
<td>.00</td>
<td>Control</td>
<td>.55</td>
<td>4.59</td>
<td>.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phono Sensitivity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TF2_SF1 Latency dif</td>
<td>.40</td>
<td>3.28</td>
<td>.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>.63</td>
<td>29.68</td>
<td>.00</td>
<td>Control</td>
<td>.52</td>
<td>5.04</td>
<td>.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phono Working Memory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NWR</td>
<td>.54</td>
<td>5.25</td>
<td>.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>.67</td>
<td>22.66</td>
<td>.00</td>
<td>Control</td>
<td>.52</td>
<td>5.20</td>
<td>.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phono working memory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NWR</td>
<td>.48</td>
<td>4.60</td>
<td>.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phono Sensitivity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TF4_SF2 Latency dif</td>
<td>.20</td>
<td>1.96</td>
<td>.059</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No.</td>
<td>$R^2$</td>
<td>$F$</td>
<td>$p$</td>
<td>Variable</td>
<td>$\beta$</td>
<td>$t$</td>
<td>$p$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-------</td>
<td>------</td>
<td>------</td>
<td>---------------------------</td>
<td>---------</td>
<td>------</td>
<td>------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>.53</td>
<td>12.54</td>
<td>.00</td>
<td>Control</td>
<td>.57</td>
<td>4.80</td>
<td>.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ERP measure of Phono Working Memory</td>
<td></td>
<td></td>
<td>TF2_SF1 Latency Dif</td>
<td>.33</td>
<td>2.58</td>
<td>.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ERP measure of Phono Sensitivity</td>
<td></td>
<td></td>
<td>N1/P2 Amplitude Dif</td>
<td>.19</td>
<td>1.54</td>
<td>.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No.</td>
<td>$R^2$</td>
<td>$F$</td>
<td>$p$</td>
<td>Variable</td>
<td>$\beta$</td>
<td>$t$</td>
<td>$p$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>-------------------------------</td>
<td>---------</td>
<td>------</td>
<td>------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>.54</td>
<td>13.35</td>
<td>.00</td>
<td>Control Age</td>
<td>.55</td>
<td>4.70</td>
<td>.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ERP Measure of Phono Working Memory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TF2_SF1 Latency Dif</td>
<td>.31</td>
<td>2.47</td>
<td>.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ERP Measure of Phono Sensitivity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TF4_SF1 Latency Dif</td>
<td>.24</td>
<td>1.90</td>
<td>.066</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>.50</td>
<td>11.18</td>
<td>.00</td>
<td>Control Age</td>
<td>.58</td>
<td>4.74</td>
<td>.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ERP Measure of Phono Sensitivity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N1/P2 amplitude difference</td>
<td>.21</td>
<td>1.60</td>
<td>.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ERP Measure of Phono Sensitivity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TF4_SF2</td>
<td>.27</td>
<td>2.1</td>
<td>.04</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Discussion

This study attempted to investigate two critical questions within the speech perception literature. The first aim was to determine if ERP indices of phonological sensitivity were associated with language competence in the toddler/young child population. The results suggested that ERPs recorded in response to changing phonological information at the word level were strongly associated with language skills measured by clinical assessments commonly used in the field. Differences within amplitude between the new and old conditions in the N1/P2 complex were positively associated with language performance, meaning that as children increased in language ability, their amplitude within the new condition increased relative to the old condition. Furthermore, differences within the latency domain for TF2_SF1 (248-360ms) and TF4_SF2 (136-220ms) were associated with language performance such that as language performance increased, the response within the new condition was faster in relation to the old response. Results of the regression analyses suggest that ERP measures significantly predict language skill within the young child population separately than that captured by performance on a behavioral language assessment of phonological working memory (NWR) and age. These results are consistent with other studies, which found ERPs recorded in infancy were significant predictors of language development during the toddler and school age years (Molfese & Molfese, 1985, 1997; Guttorm, Leppanen, Poikkeus, Eklund, Lyytinen, & Lyytinen, 2005; Kuhl, Conboy, Padden, Rivera-Gaxiola, & Nelson, 2008; Rivera-Gaxiola, Klarman, Garcia-Sierra, & Kuhl, 2005)

The second aim of the study was to investigate if perceptual sensitivity to phonemic changes measured by ERP distinctly contributed to language independently from phonological working memory measured by NWR. Joanisse and Seidenberg (2003) propose a connectionist
account of phonological working memory in line with that of MacDonald and Christiansen (2002). Joanisse and Seidenberg state “The conceptualization of working memory and phonology as two closely related mechanisms reflects the theory that the two are, in fact, are inseparable and indistinct components of cognitive processing” (p. 54). It was hypothesized that both ERP measures of phonological sensitivity and phonological working memory measured by NWR would represent a distinct portion of variance, essentially demonstrating that those skills independently contribute to language performance.

The results from regression analyses within the N1/P2 timeframe and within TF4_SF2 (136-220ms) suggested that perception of speech measured by ERP accounted for a unique amount of variance in predicting language skill separate of NWR and age (see regression Table 4 No. 1 & No 5). These early time frames appear to explain a significant portion of variance separate from phonological working memory and age. On the other hand, when latency differences within TF2_SF1 (248-360ms) were included in a regression model with NWR and Age, ERP measures of perceptual sensitivity were unable to predict the LFS separate from NWR (see Table 4, No.2). In essence, the predictors in the model were codependent and therefore the processing captured within TF2_SF1 (248-360ms) may be less associated with general sensitivity to phonemic changes and more associated with phonological working memory, also captured by NWR.

To determine if the differences in time windows of the ERP data were in fact measuring different processes (i.e. N1/P2 and TF4_SF2 measuring perceptual sensitivity to phonemic changes and TF2_SF1 measuring phonological working memory) three additional regression analyses were conducted to pit the ERP time windows against each other and determine if each accounted for unique variance within the LFS. N1/P2 amplitude difference did not account for a
significant portion of variance when included in the model with TF2_SF1; however, TF4_SF2 (136-220ms) predicted the LFS separate from that of TF2_SF1 (248-360ms). It appears that that the linguistic processes most relevant to language acquisition are more reflected by the peak latency difference in the 136-220 ms time window than the processes that are reflected by amplitude in the N1/P2 complex.

The PCA was used to extract independent factors, which do not overlap in time and explain a unique portion of variance within the entire EEG data set. Perhaps this is why TF4_SF2 was able to represent unique variance in the LFS separate from TF2_SF1. It is possible that the processing of phonemic information captured within the latency domain of TF4_SF2 is a more sensitive indicator of phonemic changes than that of the amplitude difference of the N1/P2 complex. It is also possible that the processing capacities responsible for detection of phonemic changes captured within TF4_SF2 are somewhat different than those measured within the N1/P2 complex. Perhaps the processing reflected in TF4_SF2 are those, which are more associated with the neural encoding of distinct speech features (i.e. the place of articulation of stop consonants, vowel space perception and voiced/voiceless distinctions) or integrating those features to support discrimination of phonemic change. This claim is supported by other studies, which have reported increased latencies recorded within left temporal regions to capture perceptual properties associated with phonemic discrimination (Korczak & Stapells, 2010; Tremblay, Inoue, & Bernhard, 2010).

Furthermore, the N1/P2 complex encompasses a timeframe that was determined by visual inspection of the data. It included a wide time window to allow for variability present in the grand average ERP data. The time windows designated for analyses based on visual inspection may be less sensitive than the PCA data to account for common variance detecting EEG data
change. Due to the slight overlap in timing between N1/P2 and TF2_SF1, the N1/P2 complex did not account for a unique portion of variance when included within the regression model with TF2_SF1.

Therefore, given that 1) the difference in latency in TF4_SF2 (136-220ms) as a measure of phonemic sensitivity predicted language skill separately from TF2_SF1 (248-360ms) and 2) the overlap in duration within N1/P2 (50-300ms) and TF2_SF1(248-360ms) may be responsible for N1/P2 to significantly predict language skill separate form TF2_SF1; it is concluded that processing captured within an earlier timeframe (prior to that of 248ms), may be more indicative of phonological sensitivity.

The phonemic processing capacities present within TF4_SF2 within the latency domain may be a more sensitive indicator of phonemic changes present in the stimuli when compared to the amplitude difference of N1/P2. The methodology of ERP is one that is extremely sensitive to the temporal domain. Latency may be capturing differences in phonemic processing differently than amplitude due to differences in neuroanatomical structures responsible for latency and amplitude measurements. The latency domain may be measuring signal conduction of the neural response within white matter myelin tracts (Eggermont,1988, 1992). The amplitude domain may be capturing synchronous neural activity from dense synaptic clusters within the auditory cortex as well as general brain noise and spectral power (Harris, Vaden, & Dubno, 2014). Differences along the recording parameters of the latency and amplitude domains may reflect nuances in the neural signature of networks underlying phonemic processing.

Furthermore, the processing capacities reflected in the amplitude and latency domains may be indicative of the heterogeneity in the neural networks enabling language and therefore
possibly contributing to differences in behavioral language performance within the participants. Differences in latency and amplitude measurements may be indicative of maturational changes within the cortex. Children with high language skill may demonstrate increased signal conduction of the neural response due to greater myelination of the white matter cortical tracts as well increased synaptic density within the auditory cortex. This synaptic density is reflective of synchronous neural networks for processing of phonological stimuli. Poorer language skill may indicate decreased myelination and decreased synchronous neural networks due to more dispersed synaptic firing within cortical structures responsible for processing phonological information.

There is an extensive body of ERP research, which supports the claim of distinct processing represented within different ERP timeframes. Early ERP components such as the N1/P2 complex have been associated with the ability to perceive phonemic contrasts that are present within the native language. Dehaene-Lambertz (1997) found that changes to phonemic stimuli within the native language were reflected within the P2 and mismatch negativity components in a sample of 16 French-speaking adults. The authors report that the ERP indices of phonemic change were highly accurate and generated an early specific evoked response around 200ms. More importantly, the P2 complex was not sensitive to changes in acoustic parameters of the stimulus that were irrelevant for phonemic detection.

Temporal factor 2 encompassed a timeframe from 248-360ms. Its ERP waveforms showed a positive peak recorded in the frontal midline region of the scalp with the average maximum peak occurring around 308ms. It is possible that TF2_SF1 is reflecting the P3a component. The P3a component occurs when a “distractor” or oddball stimulus is played among other frequently occurring stimuli (Fonaryova- Key, Dove, & Maguire, 2005) within a passive
listening condition. Despite numerous studies on the P3a effect, there have been diverse interpretations regarding its cognitive representations including memory updating and (Donchin & Coles, 1988) and stimulus discrimination (Verleger, 1988). In a review of the P300 effect, Linden (2005) reports that both attention and working memory are measured within the P300 time window such that attention or recognition of the deviant stimulus is supported by working memory which maintains the features of the standard stimulus for comparison against the deviant. Furthermore, Donchin and Coles proposed that infrequent novel stimuli elicit relatively large P300 amplitudes because the memory trace for the prior similar target decays and the presentation of the new target refreshes neural activity to a greater degree. Conversely, a more frequent novel stimulus will have a greater representation in memory and therefore the new target generates less activity compared to that of a more infrequent paradigm.

In a recent paper by Bonala and Jensen (2012), the authors devised a computational model that mimics the learning mechanisms associated with the P300 component to determine if working memory was responsible for the P300 effect. The model of working memory was based off of Baddley’s (2000) account which consists of a phonological loop and visuospatial sketchpad which act as short term memory storage for content domain areas and an episodic buffer which links information between the two. Simulation results of this model were such that a larger P300 amplitude was present for infrequent stimuli compared to that of more frequent. The computational model did mimic the Baddley model and therefore supported the P300 effect being elicited from a working memory process. In the current study, it is possible that TF2_SF1 is accounting for a P300 effect and measuring phonological working memory similar to that of the NWR task. The ERP response may be capturing the participant’s ability to
store the phonemic properties of the new stimulus at a more lexical level and compare the new stimulus to the phonemic properties of the old.

**Distinct Features of N1/P2 in the Toddler Population**

The N1/P2 complex has been cited as a measure of early processing of the phonemic features of speech (see Jerger, Martin, & Fitzharris, 2014 for review). Within this study, the N1/P2 complex was observed within the central-parietal region of the scalp. The combined amplitude of the “old” stimulus demonstrated larger amplitude than that of the “new” stimulus. This result is in contrast to the results of Landi et al. (2010) who found that within a sample of typically developing 11-year-olds (n=41) the amplitude for the “new” stimulus was substantially larger than that of the “old” when using the same stimuli and similar recording parameters.

Within the current study, the positive correlation between the LFS and the amplitude difference between the conditions within the N1/P2 complex suggests that as the language skills of the participants increased the difference between the new and old stimuli became more positive. In essence, the difference between the new amplitude and the old amplitude decreased due to the amplitude for the new condition showing increased voltage relative to that of the old. When comparing the top 10 language performers to the bottom 10 language performers (based on the LFS), the children with the higher language scores had larger amplitude for the new condition when compared to the old condition. The children who produced the top 10 language factor scores demonstrated new amplitude that was on average 85% as high as their old amplitude (i.e. average new amp = 4.3µV / average old amp = 5.1µV). The bottom 10 language
performers produced amplitude for new 74% as high as their old amplitude (i.e. average new amp = 6.2µV / average old amp = 8.5µV).

It is concluded that the children with greater language skill demonstrate increased synchronous neural activity when responding to a novel stimuli than the children with lower language performance. The children with lower language performance recruited less neural activation when provided with a novel stimulus. This result is consistent with a myriad of ERP investigations demonstrating increased neural activity to phonemic changes is related to better language outcomes (see Molfese, Molfese, & Pratt, 2007 for review). In light of the results of Landi et al, it is possible that the young population under investigation is presenting a maturational effect. Perhaps if this same cohort was tested further along the developmental trajectory, they would present similarly to that the 11-year-old subjects within the Landi et al. study with higher amplitudes for the new stimulus.

Furthermore, the children demonstrating lower language ability appear to recruit less neural activity for the new stimuli compared to old. It is also possible that the children with lower language ability demonstrate a maturational lag. In a study by Bosseler et al. (2013), infants 6-months-of-age showed greater cognitive effort in response to familiar stimuli as measured by magnetoencephalography (MEG) compared to adults who demonstrated greater neural activity for a novel stimulus. A sample of 12 month olds was also tested and they demonstrated a pattern of neural activity transitioning from that of the 6 month olds but more similar to that of the adults. The authors speculate that the 6 month olds’ responses to a familiar stimulus were consistent with statistical learning patterns in which the young infants were calculating the amount of frequent input in their environment to gain knowledge of the categorical patterns of the language. Once categorical learning is stable, attention is shifted to
novel stimuli, which elicits increased cognitive effort and attention. The children within lower language ability within this cohort may be presenting a similar maturational pattern in which cognitive effort is still greater for familiar phonemic stimuli as they continue to allocate resources to determine the statistical properties of the language. Children with greater language skill demonstrate more stable phonological representations and are stimulated therefore by the novel target.

Theoretical Implications

ERPs are a powerful tool in that they measure the temporal aspects of phonemic processing. What is still debated in the literature is the exact nature of how strong phonological encoding supports general language skill. It is possible that early perceptual abilities and salience of phonological representations enable language learning in a bottom up fashion, such that parsing the speech stream as an infant, word learning and eventually morphological learning stem in part from appropriate categorization of phonological units (Jusczyk & Bertoncini, 1988). It is also possible that deficits in phonological perception or goodness of phonological representations negatively affect language learning in a similar fashion. This is the general premise of the Perceptual Deficit Theory put forth by Joanisse and Seidenberg (2003).

The results of this study suggest that there is a connection between sensitivity to changing phonological stimuli measured by ERP within the N1/P2 complex and TF4_SF2 (136-220ms) and general language ability. Analyses revealed that ERP recordings within TF2_SF1 (248-360ms) may be reflecting phonological working memory processes. TF2_SF1 was highly correlated with all language measures and was highly predictive of language over and above that of age. It is concluded that a stronger ability in discrimination of phonological information and
phonological working memory skill are significantly related to general language aptitude. The data reveals that the opposite is also true. Children with lower language abilities were less sensitive to changes in phonological information at the word level as measured by ERP. Children with lower language abilities showed diminished amplitudes to the new stimuli when compared to the old condition in the N1/P2 complex. They also demonstrated reduced latencies for the new stimulus in relation to the old when compared children with greater language ability. Therefore this data supports the general premise of the Perceptual Deficit Theory.

What is not supported by the results of this study is the general theoretical framework, which suggests that the goodness of phonological representations and general phonological working memory capacities are inseparable cognitive mechanisms. What was shown within this investigation was that the ability to perceive changes in phonemic stimuli measured by ERP within the N1/P2 complex and TF4_SF2 (136-220ms) represented a unique portion of variance within language skill separate from phonological working memory. However, TF2_SF1 (248-360ms) explained a significant portion of variance separate from phonological working memory. Additional regressions were then conducted to determine if 1) TF2_SF1 was indeed measuring phonological working memory similarly to that of NWR or if TF2_SF1 was in essence capturing phonemic change and therefore supporting Joannises’s claim that phonemic representations enable phonological working memory so that they contribute similarly to the language system.

TF4_SF2 (136-220ms) explained a significant portion of variance within the language factor score separate from TF2_SF1 (248-360ms); therefore, it was concluded that what was being measured in TF4_SF2 was sensitivity to phonemic properties of the stimulus, and T2_SF1 was capturing working memory. Further evidence supporting TF4_SF2 as detecting phonemic change is the high correlation with the N1/P2 (see Table 4, No 8) complex which has been cited
within numerous studies as an indicator of phonological sensitivity (Dehaene-Lambertz, 1997; Jerger, Martin, & Fitzharris, 2014) and it’s strong correlations to language ability measured by the behavioral assessments. Furthermore, TF2_SF1 revealed a positive peak in the frontal-midline areas at approximately 300ms. This is interpreted to be a P300 effect reflecting working memory skills. This claim is supported by other studies demonstrating the P300 effect as a measure of working memory (Donchin & Coles, 1988; Bonala & Jansen, 2012; Kok, 2001).

**Clinical Implications**

This study aimed to investigate relationships between ERP measures of phonological sensitivity and their relationship to clinical assessments of language skill. Within the field of speech-language pathology, a clinician assumes a great responsibility in providing an appropriate diagnosis of impairment. At the present time, the toddler population presents significant challenges due to limitations in sensitivity and specificity of our behavioral measures of language. It is critical that the research community take steps to improve diagnostic procedures for young children. Inherent in that, is gaining a greater understanding into the skills that enable language and therefore contribute to language learning ability. If phonological sensitivity underlies general language skill, then assessment of perception could be added to a diagnostic battery to improve our understanding of the general clinical profile for a child presenting a language delay.

Electrophysiological measures such as the auditory brainstem response ABR and the middle latency response MLR have had a significant impact on the field of audiology in their ability to measure hearing acuity. These measures have been used clinically to assess hearing within newborns and clinical populations, which may be difficult to test behaviorally. In the
fields of psychology and speech & language pathology, there have been movements to utilize electrophysiological measures of speech perception such as the complex/speech ABR (Hornickel, Knowles, & Kraus, 2012) and the MMN (Naatanen, 2003). Studies of these electrophysiological measures of perception show promise in identifying distinct neural signatures for a variety of developmental disorders including specific language impairment (Basu, Krishnan, & Weber-Fox, 2010), (C)APD (Rocha-Muniz, Befi-Lopes, & Schochat, 2012), dyslexia (Billiet & Bellis, 2011; Banai, Hornickel, Skoe, Nicol, Zecher, & Kraus, 2009) and attention deficit hyperactivity disorders (ADHD) (Azzan & Hussan, 2010).

Despite mounting evidence supporting the clinical utility of electrophysiological measures of speech perception, general poor reliability and confounds to testing parameters are evident. Reliability of electrophysiological indices of speech perception are poor due to individual variation in the ERP response and limited information regarding test-retest information on children with communication disorders. Given the complex, multi-dimensional properties of speech, heterogeneity of effects which stem from differences in language experience, and inconsistencies in the recording parameters for ERPs, it is possible that electrophysiological measures of perception may never reach reliability indices as stable as behavioral assessments (McFarland & Cacace 2012). The neural generators of ERP components and the underlying exogenous and endogenous mechanisms they reflect continue to be debated in the literature. Great strides must be taken to assess ERP methodology as a clinical tool. Future studies should be conducted to provide more normative data on the parameters of ERP components of typically developing children and how those parameters may differ in clinical populations.
In the current study, NWR became a compliment to the ERP task as a means to investigate its usefulness in measuring phonological working memory alongside perception. Over the course of the last 10 years, many studies have focused their efforts on designing NWR tasks for younger populations due to it’s usefulness in identifying language impairment within school, aged children (Dollaghan & Campbell, 1998). Nonword repetition is a complex task, which taps many language related skills. In fact, the underlying processes that contribute to NWR are still highly debated in the literature. Some scholars suggest that deficits in NWR are caused by poor phonological representations while others suggest deficits in phonological working memory or storage as cause for poor NWR performance (Ebbels, Dockrell, & van der Lely, 2012).

This study attempted to test very young children using a NWR task. Inherent in young children is variability of language skill, and more specifically general emergence of certain skills. It was evident when testing young children on a NWR task, that general speech production skills affected the level of success. Even when speech sound distortions or substitutions were accounted for, omissions of a speech sounds presented as a challenge. It was difficult to judge whether an omission of a phoneme on the NWR task was due to difficulties in speech sound production, or was perhaps due to deficits in phonological working memory. Given the emerging and unstable language system in young children, in addition to varied compliance with testing parameters, the results of the GFTA_2 or a PCC were not always indicative of speech sound production at the conversational level and therefore did not always provide a valid profile of articulation abilities. Therefore, difficulties in speech sound production ads yet another limitation to using NWR in testing young populations.
Studies that have investigated the diagnostic accuracy for identifying young children at risk for language impairment and present statistical evidence that NWR has the capacity to support the decision making process in referring young children (before the age of 5) for more comprehensive language testing (Stokes & Klee, 2009; Roy & Chiat, 2004). However, NWR may be insufficient as the primary source for referral given its limitations in accounting for speech sound errors in this population (Deevy, Wiseman Weil, Leonard, & Goffman, 2010). The results of this study are consistent with these findings. Prior to any clinical use, NWR tasks for young children must provide sensitive scoring parameters to account for speech sound productions, while preserving its ability to account for general phonological working memory skill. Further research is needed on scoring parameters as well as diagnostic accuracy.

Although the current study provides evidence for divergence between ERP measures of phonological sensitivity and NWR, these two tasks both measure perceptual linguistic abilities. All the regression analyses conducted for this study which pitted ERP measures against the NWR task, resulted in NWR being a stronger predictor of language skill above and beyond that of ERP. NWR is a task, which involves both perceptual abilities of phonological characteristics, as well as phonological working memory and speech production. Certainly, the NWR task taps numerous language related skills and therefore may provide a more encompassing perspective of general linguistic skill when compared to ERP.

However, a persistent problem in the diagnosis of late talking toddlers is the mere fact that some children who produce little to no speech at age 24 months and beyond are not amenable to behavioral testing. Therefore, a nonword repetition task becomes yet another measure for the self-directed toddler to avoid. The clinician is therefore again forced to rely on observation of non-linguistic skills such as gestures, social engagement and general parent report
to provide clinical recommendations regarding the presence of a developmental delay and furthermore whether the child should receive treatment. If advances are made to improve the reliability and validity of ERP measurement, it is possible that ERP can become an advantageous compliment to behavioral language assessment, including nonword repetition, to support general decisions regarding the diagnosis and treatment of young children with language delay.

The results of this paper support the theory that phonological sensitivity has a strong relationship to language performance. These results support models of language which link higher level linguistic processing of speech features to speech and language production (Hickock, 2012). If a link between perceptual abilities and language output exists, interventions such reflect these scientific findings. Many of the language interventions used with young children focus on whole word approaches and expanding general expressive language skill. Little attention is given to providing direct stimulation of phonological stimuli. If phonological perception is important to a language learning system, then interventions focused on perception may not only increase general language ability, but also bolster an emerging system to the extent that future academic deficits, particularly in the area of reading, may be prevented. Input focused interventions such as those described by Rvachew & Brosseau-Lapre (2010) may be a way to orient young children to the salient features of speech, improve phonological representations, increase phonological awareness and support general language ability.

**Limitations**

These findings should be considered in the context of several limitations. First, the entire sample size for the ERP data included 40 children, only 4 of which demonstrated language impairment. Two of the children did not provide full data sets, therefore, statistics are provided for 38 children. More children, especially children demonstrating language delays are needed to
improve generalization of the results and enhance statistical effects. It is acknowledged that the alpha levels for regression analyses are above that of 0.05 and therefore, the results should be interpreted with caution.

The literature on ERPs in the toddler/young child population is scant. Therefore, there is limited data on ERP indices of speech perception and their general recording parameters. Greater studies are necessary to provide guidelines on ERP measures of speech perception and how ERP measures change with development. Also, given the toddler/young child population, a significant amount of data loss is reported in the study due to contamination of movement and noise artifact. Furthermore, diploe modeling and source localization was not utilized for EEG measures; therefore conclusions regarding the neural generators of the ERP data cannot be made. EEG data recorded at the scalp does not necessarily reflect neural activity directly below the electrodes. This makes generalization of the current results limited to other neuroimaging studies of speech perception. Future research could pair temporally sensitive ERP data with source localization techniques such as Near Infrared Spectroscopy (NIRS) to provide insight regarding the neural architecture of speech processing networks.

Lastly, despite evidence suggesting the separation of robust phonological representations having and independent contribution to language performance when compared to phonological working memory, there is no doubt that these results are certainly an oversimplification of the intricate and complex neuro-linguistic processes associated with perception and production. The role of the P300 response in its representation of phonological working memory remains speculative. Therefore, a healthy level of speculation must exist regarding how the measures used for this study, namely ERP and NWR, are actually capturing the cognitive processes claimed by the author. It is acknowledged that the computational framework cited by Joanisse
and Seidenberg (2003) may still provide insight into general language learning mechanisms and
that the mere measurement approaches utilized in this study are incomprehensive. These
measures may not fully explain the nature of phonological representations and their instantiation
in memory.

**General Conclusions**

Many studies have linked perceptual abilities measured by ERPs in infancy to later language
and language related skills, such as reading (Molfese, 2000; Guttorm, Leppanen, Richardson, &
Lyytinen, 2001; Guttorm, Leppanen, Poikkeus, Eklund, Lyytinen, & Lyytinen, 2005; Molfese,
1995). This is the only study to date that provides individual difference data, concurrently
correlating ERPs to changing phonological stimuli with behavioral language performance on a
variety of clinical assessments within the toddler population. The results indicate that as
language skills of young children increase, the recruitment of synchronous neural activity to a
novel stimulus also increases. The children who performed more poorly on language
assessments did not show the same pattern of neural activation for the new stimulus,
demonstrating less neural sensitivity to changes in phonological information.

Moreover, phonological discrimination within early windows of ERP processing (136-220ms) and phonological working memory measured by nonword repetition demonstrated an
independent contribution to language performance within the young child group. Interestingly,
later ERP time windows (248-360ms) did not account for significant variance in language
performance. It is possible that these later ERP timeframes may be capturing phonological
working memory capacities similar to that of nonword repetition.
This study utilized two experimental methodologies for measuring language in the young child population. First, a NWR task was employed to measure phonological working memory. Although still in its infancy for young children, NWR may provide insight into developmental trends for young children and support clinical decision-making when used in tandem with other standardized behavioral language assessments (Roy & Chiat, 2004; Stokes & Klee, 2009; Clark, McRoberts, Van Dyke, Shankweiler, & Braze, 2012). The second experimental measure was the use of ERP to capture phonemic sensitivity to changing word stimuli. Although researchers continue to make advances in the clinical utility of electrophysiology it is not currently suited for clinical use due to limited construct validity and poor inter-subject reliability. Advances in technology may one day provide improvements for the use of this technique not only to support clinical practice, but also to provide increased understanding on the basic neural basis for perception and production. When considered together, ERP and nonword repetition show promise in providing insight into language functioning and may possibly one day provide critical information to improve identification of language impairment in young children.

Lines of infant ERP literature suggest that discrimination of phonological information is highly predictive of language skills later on in the developmental trajectory. Perhaps what’s most important in the infant brain is analyzing the frequency of the input to form stable phonological representations which supports parsing the speech stream into meaningful units. Once categorical perception is established, the brain then focuses on processing longer amounts of phonological information to support imitation and word production.

Phonological working memory enables children to hold phonological units in memory, prepare the motor speech code and produce the word. The current study suggests that high functioning language participants were able to excel at phonological working memory tasks
whereas lower language participants demonstrated weaker phonological working memory skills. It is possible that the toddler brain utilizes phonological working memory to enable the vocabulary burst and support word learning. Furthermore, deficits in phonological working memory may be implicated in late talking or language delay. This data strongly supports the instantiation of phonological working memory as being highly correlated and highly predictive of language skills within the toddler young child population. Phonological working memory was the greatest predictor of language whether measured by ERP or NWR. More studies are warranted to determine the role of phonological working memory in word learning and language delay in the toddler young child population and beyond.
Appendix A: List of Nonwords for TENR task

| /mɑd/       | /moɑkɪ/       |
| /neId/      | /dɔpəlut/     |
| /paIm/      | /bælʊkn/     |
| /boz/       | /fisaImt/     |
| /koɡə/      | /pɑ dúlməIp/  |
| /dəfi/      | /fənəlskʰ/    |
| /leIpɔ/      | /wugləmɑkʰ/   |
Appendix B: Scoring Protocol

Non-Word Repetition – Stokes and Klee 2009

Scoring: If correct, put (+) below the transcription. If incorrect, broadly transcribe participant’s production (first production if more than one occurred) under the transcription. Then, write the number of phonemes correct for the item (maximum given in parentheses)

<table>
<thead>
<tr>
<th>TARGET</th>
<th>TOTAL P:</th>
<th>TOTAL S</th>
</tr>
</thead>
<tbody>
<tr>
<td>/mæd/</td>
<td>m d</td>
<td>(____/3)</td>
</tr>
<tr>
<td>/neld/</td>
<td>n el d</td>
<td>(____/3)</td>
</tr>
<tr>
<td>/palm/</td>
<td>p al m</td>
<td>(____/3)</td>
</tr>
<tr>
<td>/boʊz/</td>
<td>b oʊ z</td>
<td>(____/3)</td>
</tr>
<tr>
<td>/koʊgə/</td>
<td>k o g o</td>
<td>(____/4)</td>
</tr>
<tr>
<td>/dɑfɪ/</td>
<td>d f i</td>
<td>(____/4)</td>
</tr>
<tr>
<td>/lɛlpə/</td>
<td>l el p o</td>
<td>(____/4)</td>
</tr>
<tr>
<td>/fʊpɪm/</td>
<td>f u p i m</td>
<td>(____/5)</td>
</tr>
<tr>
<td>TARGET</td>
<td>TOTAL P:</td>
<td>TOTAL S</td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>/moʊkəi/</td>
<td>m o k i</td>
<td>(___/5)</td>
</tr>
<tr>
<td>/doʊpəlut/</td>
<td>d o p a l u t</td>
<td>(___/7)</td>
</tr>
<tr>
<td>/bæləkən/</td>
<td>b æ l a k n</td>
<td>(___/7)</td>
</tr>
<tr>
<td>/fɪsalmt/</td>
<td>f i s a l m t</td>
<td>(___/7)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TARGET</th>
<th>TOTAL P:</th>
<th>TOTAL S</th>
</tr>
</thead>
<tbody>
<tr>
<td>/pædʌlemlp/</td>
<td>p æ d u l æ m l p</td>
<td>(___/9)</td>
</tr>
<tr>
<td>/fɪnąlsk/</td>
<td>f n a l s k</td>
<td>(___/8)</td>
</tr>
<tr>
<td>/wʊglæmk/</td>
<td>w u g l æ m k</td>
<td>(___/9)</td>
</tr>
<tr>
<td>/lædɪnatl/</td>
<td>l æ d æ n æ t l</td>
<td>(___/9)</td>
</tr>
</tbody>
</table>

Grand Total Syllables: (___/40)

Grand Total Phonemes: (___/90)  
Total Score ____________
References


Foundation, L. R. (2014). *lena foundation*. Retrieved from lena foundation.org:

[www.lenafoundation.org](http://www.lenafoundation.org)


