JSLHR

Research Article

Electrophysiology of Perception and Processing of Phonological Information as Indices of Toddlers' Language Performance

1822

Vanessa Harwood,[a,b] Jonathan Preston,[b,c] Bernard Grela,[a] Dooti Roy,[a] Olivia Harold,[b] Jacqueline Turcios,[b,d] Kiyomi Andrada,[a] and Nicole Landi[a,b]

Purpose: The toddler years are a critical period for language development and growth. We investigated how event-related potentials (ERPs) to repeated and novel nonwords are associated with clinical assessments of language in young children. In addition, nonword repetition (NWR) was used to measure phonological working memory to determine the unique and collective contribution of ERP measures of phonemic discrimination and NWR as predictors of language ability.
Method: Forty children between the ages of 24–48 months participated in an ERP experiment to determine phonemic discrimination to repeated and novel nonwords in an *old/new* design. Participants also completed a NWR task to explore the contribution of phonological working memory in predicting language.
Results: ERP analyses revealed that faster responses to novel stimuli correlated with higher language performance on clinical assessments of language. Regression analyses revealed that an earlier component was associated with lower level phonemic sensitivity, and a later component was indexing phonological working memory skills similar to NWR.
Conclusion: Our findings suggest that passive ERP responses indexing phonological discrimination and phonological working memory are strongly related to behavioral measures of language.

During toddlerhood, children experience synergistic and rapid language growth. Decades of research have focused on identifying the underlying mechanisms associated with these significant gains in language. Some theories suggest that phonological processing lays the foundation for further lexical and syntactic development (Joanisse & Seidenberg, 2003). Several studies have provided evidence that young children who demonstrate keen phonological processing abilities display more robust language skills than children who are poor phonological processors (Benasich & Tallal, 2002; Guttorm et al., 2005; Kuhl, Conboy, Padden, Rivera-Gaxiola, & Nelson, 2008; Rivera-Gaxiola, Klarman, Garcia-Sierra, & Kuhl, 2005). Understanding the relationship between phonological processing and early language development may aid in identifying early phonological processing deficits and provide information regarding language trajectories for toddlers.

There is reason to believe that phonemic discrimination and phonological working memory each play an essential role in early language acquisition and that deficits in these domains may lead to language impairment. According to the perceptual deficit theory (PDT), impairment in language stems from a perceptual phonological impairment (Joanisse & Seidenberg, 2003). The PDT predicts that degraded phonological perception may contribute to poor 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." It may include appropriate coding of phonemic information, storage, and organization of articulatory output. It is possible that phonological working memory acts as a catalyst to vocabulary growth within the first two years of life and that poor phonological working memory skills inhibit language production.

Several studies have indicated a relationship among phonological working memory skills and later language development. Specifically, these studies have suggested that deficits in phonological processing are linked to early

---

[a]University of Connecticut, Storrs
[b]Haskins Laboratories, New Haven, CT
[c]Syracuse University, NY
[d]Southern Connecticut State University, New Haven

Correspondence to Vanessa Harwood: Vanessa.harwood@uconn.edu

Editor: Sean Redmond
Associate Editor: Lisa Archibald

Received December 21, 2015
Revision received May 2, 2016
Accepted October 24, 2016
https://doi.org/10.1044/2016_JSLHR-L-15-0437

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

language delay and later language performance. Rescorla (2009) examined late-talking toddlers' language outcomes at age 17, finding that on average the late talkers performed similarly to typical peers on a number of general language and reading indices; however, the late-talker group performed significantly lower on subtests of vocabulary and grammar, as well as on tests of verbal (phonological) working memory. Further, Preston and Edwards (2010) reported that 8-year-olds with histories of late talking performed lower than children with histories of typical language acquisition on several measures of oral and written language. Functional brain imaging results from this study also indicated differences in processing of phonological information across a variety of cortical and subcortical regions. It is possible that early phonological processing abilities contribute to language acquisition in a bottom-up manner, laying the framework for expressive language development.

If phonemic discrimination and phonological working memory skills play a critical role in language development, it is essential for researchers to investigate links between early phonological processing and language skill. In a behavioral context, nonword repetition (NWR) tasks are particularly sensitive to the phonological aspects of word learning and working memory (Dollaghan & Campbell, 1998; Ebbels, Dockrell, & van de Lely, 2012). NWR tasks have recently been designed to assess the phonological working memory skills of toddlers and preschoolers (Clark, McRoberts, Van Dyke, Shankweiler, & Braze, 2012; Roy & Chiat, 2004; Thal, Miller, Carlson, & Moreno-Vega, 2005). Cognitive neuroscience techniques can be used to provide objective measurement of the neural response to speech using paradigms that do not require a behavioral response. In particular, event-related potentials (ERPs) might be utilized to provide a measure of the neural system's discrimination between changing phonemic stimuli (Kuhl, Conboy, Padden, Nelson, & Pruitt, 2005; Näätänen et al., 1997). In the current study, we employ both behavioral (NWR) and neurolinguistic (ERP) approaches in conjunction with standardized assessments in order to better understand the distinct phonological processing capacities that contribute to language production skills within toddlers.

ERPs and Language Acquisition

In a series of studies, Molfese and colleagues (Molfese, 1995, 2000; Molfese & Molfese, 1985, 1997) demonstrated that ERP responses in infancy strongly predicted preschool and school-age language and literacy ability. Molfese and Molfese (1997) reported that newborns' ERP responses to speech syllables could be used to classify children into high- and low-functioning language ability at age 5 years. In particular, 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- or low-functioning groups at age 5 years, based on standardized assessment with 80% accuracy. A subset of that same cohort was re-evaluated at age 8 years. N1 responses to syllables at birth discriminated between normal, poor, and dyslexic readers at age 8 years with 81.6% accuracy (Molfese, 2000). Faster latencies and larger N1 amplitudes characterized the effects for the control children. This evidence is also supported by other findings, which suggest that sensitivity to changes in phonological structures at birth within the N1 component differ in typically developing children and those with familial risk for impairment (Guttorm, Leppanen, Richardson, & Lyytinen, 2001).

These studies provide evidence that early sensitivity to phonemic and phonetic information, as measured by ERPs, is an important component of early language performance. Furthermore, the researchers suggest that there are differences in the neural substrates that underlie phonological processing when comparing children with and without language impairment. Therefore, ERPs could provide vital information regarding the perceptual skills in toddlers and perhaps identify individual differences in important linguistic processes. Such information might ultimately inform clinical and theoretical approaches to determining which children are at greatest risk for later language impairment.

Nonword Repetition

The PDT (Joanisse & Seidenberg, 2003) emphasizes the link between poor phonological working memory and language impairment. Nonword repetition tasks are designed to measure phonological working memory skill in children. Many NWR tasks are created to increase in complexity from simple to more complex syllable structures. It has been found that more complex items (e.g., nonwords with more syllables) may discriminate between children with and without specific language impairment (Archibald & Gathercole, 2007; Dollaghan & Campbell, 1998).

Stokes and Klee (2009) investigated the diagnostic accuracy 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 phonemic inventory of very young children (24 months) with low wordlikeness, while increasing in length from 1–4 syllables. The investigators concluded the TENR could be used for successful identification of 2-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 fact that both atypical phonemic perception and phonological working memory deficits have been implicated in language impairment (Dollaghan & Campbell, 1998; Weber-Fox, Leonard, Hampton, & Tomblin, 2010) suggests that further investigation of the relationship among phonological discrimination, phonological working memory skills, and early language development is warranted.

Purpose and Hypothesis

The current study was designed to determine whether neural response to changing phonological stimuli (repeated disyllabic nonwords within an *old/new* paradigm) using ERPs is associated with common clinical assessments used to measure language performance in toddlers. Also, within the PDT framework, we examine whether neurolinguistic indices of phonemic discrimination and behavioral measures of phonological working memory uniquely predict language competence in toddlers. We hypothesize that ERP indices of phonemic discrimination would be positively correlated to language performance. Furthermore, we expect that phonological sensitivity (as measured by ERP) and phonological working memory (as measured by NWR) should both contribute to language competence, with each process explaining a unique portion of variance in language skill.

Method

Participants

Forty children (22 boys, 18 girls) were evaluated as part of a study at Haskins Laboratories that assessed neurobiological markers of speech perception and production. Children between the ages of 24 and 48 months were recruited from local university clinics, private practices, the Rhode Island Birth-to-Three system, and the Connecticut Birth-to-Three system. All children met the following criteria to be included in the study: (a) monolingual English speakers, with no significant exposure to any language other than English; (b) no known psychiatric or neurological deficits, per parent report; and (c) hearing within typical limits at the time of the study, per parent report. All children were reported to have passed newborn hearing screenings.

To obtain a representative sample of young children, approximately 10% of the sample included children demonstrating language delay (four participants; Rescorla, Roberts, & Dahlsgard, 1997). Standardized assessments were used to provide descriptive data regarding the participants' language abilities. Children with language delay were identified by demonstrating a standard score of < 85 on the expressive and/or receptive portion of the Preschool Language Scale–Fifth Edition (PLS-5; Zimmerman, Steiner, & Pond, 2011). Typically developing children had average receptive and expressive functioning on the PLS-5, as well as average visual reception on the Mullen Scales of Early Learning (MSEL; see Table 1). The typically developing children group ($n = 36$) included 31 Caucasian participants, two African American participants, and three Asian/Pacific Islander participants. The children with language delay group included three Caucasian participants and one African American participant. All four participants with language delay were male.

Procedures

Overview

Parents completed a background questionnaire regarding medical history as well as information on hearing, motor, and language developmental milestones. The children participated in one or two sessions totaling approximately 2.5 hr, which included the ERP task, administration of the TENR, and standardized measures of language and cognitive function. Children were provided with breaks and reinforcements (e.g., snacks, stickers, books) to help avoid fatigue. Participating families were compensated for their time financially and parents were provided with a research report summarizing their child's performance on language measures.

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. Parents mark words the child produces from a predetermined list of vocabulary outlined in the mCDI-2 form. Raw scores were used as a measure of vocabulary production (the age of some of the participants extended beyond that of the normative data, precluding the use of standardized scores).

Language Sampling

A language sample of approximately 100 utterances ($M = 105.5$, $SD = 41.4$) was collected for each participant (Heilmann, Nockerts, & Miller, 2010). Conventional language sampling procedures were used to gather a representative sample of the child's language. Research assistants trained in language sampling and analysis transcribed the language samples. Reliability checks of utterance production were performed on 20 randomly selected participants (50% of the participant pool) and were found to be 0.86. Computerized Profiling (Version 9.7; Long, 2008) was used to derive mean length of utterance (MLU), a behavioral measure of morphological development and utterance length, and percent consonants correct (PCC), a measure of phonological production skills.

Given the constraints of time and unfamiliarity with the lab setting, it was acknowledged that some children might not provide a robust representative sample during the experiment. Therefore, the Language Environmental Analysis System (LENA; LENA Research Foundation, 2014) was used with those children who demonstrated limited language skills in the laboratory setting to collect a representative sample of language within the child's naturalistic environment. The child was equipped with a digital language processor, 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. A random sampling of 5-min intervals was collected, transcribed, and analyzed similarly to that of the laboratory samples.

To ensure comparability of language samples obtained in the laboratory and home settings, 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 that collected at 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 ($M = 2.48$, $SD = 1.11$);

Table 1. Assessment scores for typically developing children with language delays.

	Full sample (N = 40)	Typically developing (n = 36)	Language delay (n = 4)
Parameter	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Age in months	34.35 (6.29)	34.06 (6.42)	37 (4.12)
mCDI-2-WP	528.28 (160.89)	556.2 (128.01)	284 (206.17)
PLS-AC	110.78 (12.17)	113 (10.51)	90.75 (7.15)
PLS-EC	109.03 (14.62)	112.03 (11.97)	81.50 (2.29)
PLS-T			
GFTA-2	108 (14.56)	110.61 (11.85)	76.67 (2.36)
MLU	2.79 (1.16)	2.90 (1.16)	1.79 (0.40)
PCC	81% (0.13)	83% (0.12)	65% (0.13)
MSEL_VR	60.51 (12.93)	62.49 (11.68)	43.25 (10.21)
TENR_T	102.53 (26.74)	104.94 (26.13)	80.75 (21.48)

Note. mCDI-2-WP = Raw count of words produced on the MacArthur–Bates Communication Development Inventories–Second Edition (mCDI-2); PLS-AC = auditory comprehension standard score on the Preschool Language Scale–Fifth Edition; PLS-EC = expressive communication standard score on the Preschool Language Scale–Fifth Edition; PLS-T = total language score on the Preschool Language Scale–Fifth Edition; GFTA-2 = standard score on the Goldman-Fristoe Test of Articulation–Second Edition; MLU = mean length of utterance; PCC = percent consonants correct; MSEL_VR = T-score of the visual reception subtest of the Mullen Scales of Early Learning; TENR_T = total score (syllables plus phonemes) on the Test of Early Nonword Repetition.

$t(9) = 1.29$, $p = .23$. These results suggest no significant difference in scores based on the environment of the sample. Therefore, we concluded that the transcript data for home and lab transcripts was substantially similar. If both a lab sample and home sample were collected, the laboratory sample was used to preserve consistency. The total number of children for whom the home sample was used was seven.

Standardized Assessment

The PLS-5 is an individually administered standardized language assessment designed for children from birth to age 7 years to assess language skill. Both the auditory comprehension and expressive communication portions were administered. The Goldman-Fristoe Test of Articulation–Second Edition (GFTA-2; Goldman & Fristoe, 2000) was also administered to evaluate speech sound production skills. GFTA-2 provides information regarding articulatory errors, which were accounted for when scoring the TENR.[1] The visual reception portion of the Mullen Scales of Early Learning (Mullen, 1995) was also administered to provide information regarding the participant's nonverbal cognitive skill.

Nonword Repetition

The TENR (Stokes & Klee, 2009) was used to measure individual participants' phonological working memory skills behaviorally. The TENR is designed to include phonemes that are typically present in the inventories of 2-year-old children. There were a total of 16 nonwords comprised of 90 phonemes. Modifications to particular phonemes and stress patterns were made to ensure the stimuli were consistent with American English (see Appendix A). All stimuli were recorded by a female speaker and were presented on a computer in a PowerPoint presentation. Each slide depicted a cartoon alien character with an (nonword) alien name. Children were given the following directions: "Let's play a game. Listen carefully and say just what I say." The children repeated the alien names following the presentation of the recorded production. Children were awarded one point for each syllable produced, and one point for each vowel and consonant produced correctly. A total score was calculated by adding the total number of syllables correct and total phonemes produced correctly. This scoring procedure was adopted to prevent floor effects and to capture the children's syllable as well as segmental accuracy. By providing credit for syllable preservation, children with inaccurate speech sounds could demonstrate memory for word parts. Speech sound substitution errors on the TENR that were consistent with errors on the GFTA-2 were 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 NWR scoring in young populations (Stokes & Klee, 2009). A second scorer trained in transcription and scoring of NWR also scored the TERN for each participant. Reliability was determined by dividing the total number of phonemes and syllables scored similarly by the total number of possible points. The reliability for the TENR task was .81.

ERP Procedures

Children were fitted with a 128-sponge Ag/AgCl electrode high-density sensor array net (Electrical Geodesics, Inc. [EGI], Eugene, OR) that was used to acquire electrophysiological data. Prior to placement, the net was soaked for 10 min in a warm KCl solution to improve conductance. The net was placed on the head using standard procedures outlined by the manufacturer (Dien, 2010). Electroencephalogram (EEG) data were recorded at a sample rate of 500 Hz using Net Station 4.5 software (EGI) with an

---

[1] One participant did not complete testing with the GFTA-2. A phonological analysis was performed using his language sample data to account for substitutions produced on the TENR.

EGI Net Amps 3 high impedance amplifier. All electrode impedances remained under 40k ohms as indicated by impedance measures made immediately before and after the test sessions. ERP data were filtered to retain signal frequencies between 1–30 Hz. The child sat on a parent's or caregiver's lap in a comfortable chair. In front of the child was a portable DVD player that played a silent movie during ERP data collection (clips of puppets from the *Yo Gabba Gabba!* television show) that facilitated compliance and provided nonauditory stimulation.

ERP Task

Participants were presented with two rhyming nonword tokens of speech, /bidu/ and /gibu/, in an old/new design. This design was chosen to allow for examination of both lower level phonemic discrimination between similar nonwords not linked to semantic representations and higher level phonological working memory skills. A similar task has been used to study language acquisition in infants (Molfese, Morse, & Peters, 1990) and adolescents exposed to cocaine in utero (Landi, Crowley, Wu, Bailey, & Mayes, 2012). Further, equiprobable designs as the one used here do not require as many stimulus presentations, which is advantageous for use with toddlers.

Stimuli were recorded by a female native English speaker. The stimulus duration for each token was 595 ms with a varied ISI of 1,800 or 2,800 ms to avoid habituation. The auditory stimulus was presented via an overhead speaker positioned above the participant presented at 85 dB SPL. The first block was a sensitization block, which consists of one token /gibu/ repeated for 50 trials. The second block was a mixed block where the tokens /bidu/ and /gibu/ were randomly presented in equal proportions for 100 total trials. 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/) acted as the "new" stimulus.

ERP Processing

ERP data were segmented into epochs including 100-ms prestimulus baseline and a 600-ms poststimulus interval. The data were visually inspected to identify bad electrodes (e.g., displacement caused by large ballistic movements). Automated routines were used to further detect bad electrodes and eye movement and blink artifacts (bad electrode > 200 μV, eye blink/eye movement > 150 μV). If an electrode was bad for more than 40% of the segments, 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 rereferenced to the average reference (vertex reference, Cz, was used during recording) and baseline corrected to 100 ms prestimulus presentation (Junghofer, Elbert, Tucker, & Braun, 1999). Finally, artifact-free segments from within Block 2 only (containing both old and new tokens) were averaged within the old and new conditions and used for statistical analysis. A criteria of at least 20 preserved trials for each condition was used to include subjects in the ERP analysis. There was no significant difference in number of preserved trials between the 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 (blink slope threshold = 14 μV/ms; Gratton, Coles, & Donchin, 1983) was conducted on six participants due to fewer than 20 blink- or artifact-free trials per condition prior to ocular artifact correction. All 40 participants were included in the ERP analysis.

ERP Analysis

EEG data were submitted to a temporal/spatial principal components analysis (PCA) to identify time frames and electrode montages of interest using the ERP PCA Toolkit (Dien, 2010). The purpose of the PCA was to identify systematic variance in the ERP signal within the temporal and scalp topographic domains. This data-driven approach is particularly useful for studying toddlers, given the limited literature on ERP components in this population. Further, this approach facilitates comparisons of ERP data across different developmental populations and laboratories (Molfese, Nunez, Seibert, & Ramanaiah, 1976).

Our data were first subjected to a temporal PCA, conducted with promax 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). This analysis identified four temporal factors that accounted for ≥ 5% each of the total variance (62% total), and five additional factors contributed < 5% of the variance and were not explored further. Temporal Factor 1 accounted for 27% of the variance and encompassed a time window from 544–700 ms; Temporal Factor 2 accounted for 17% of the variance and encompassed a time window from 248–360 ms; Temporal Factor 3 accounted for 13% of the variance and encompassed a time window from 404–500 ms; and Temporal Factor 4 encompassed a time window from 136–220 ms and accounted for 5% of the variance.

Following the temporal PCA, a spatial PCA with infomax rotation (Dien, Khoe, & Mangun, 2007) was then run on each temporal factor to identify electrodes that loaded strongly within each time window. Temporal Factors 1 through 3 produced four spatial factors and Temporal Factor 4 produced three spatial factors. A total of 15 temporal/spatial factor combinations accounted for at least 5% of the variance each and were thus further explored for analyses.

Following the temporal/spatial PCA, adaptive mean amplitude and peak latency for each condition (old/new) were derived for every time window and electrode montage. The amplitude difference effect of the old relative to the new condition was derived by subtracting the mean amplitude between the two conditions (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 total language score of the PLS-5 to avoid a type I error. If a significant relationship was found with the total language score of the PLS-5 and ERP measurements, additional correlations with individual language assessments were explored.[2] Regression analyses used the PLS-5 total language score as the dependent variable (predicted variable).

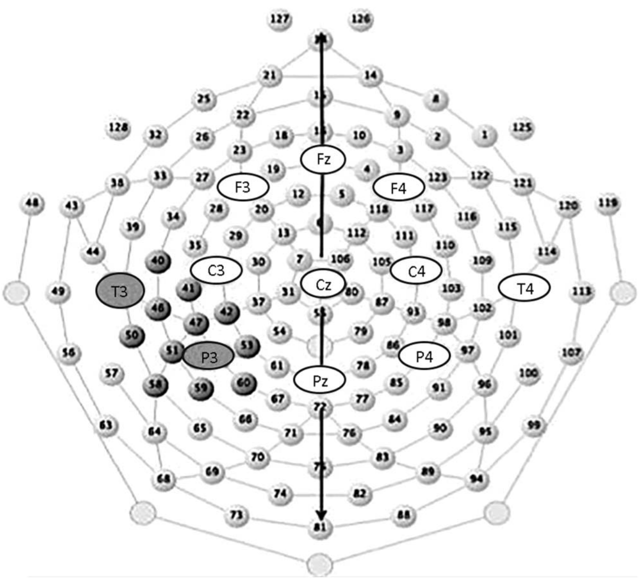
Results

Correlations were run among amplitude and latency differences within each of the 15 PCA temporal/spatial pairings and the PLS-5 total language score to determine if ERP responses were associated with behavioral assessments used to measure toddlers' language performance. There were no significant correlations among the amplitude difference between the old and new tokens for any of the 15 temporal/spatial factor pairings and the PLS-5 total language score (see Appendix B). Latency difference analyses revealed a significant early component within Factor 4 (136–220 ms) and a late component within Factor 2 (248–360 ms). These early and late components coincided with other ERP time frames identified in the literature as measuring perceptual and/or phonological abilities (Guttorm et al., 2005; Molfese & Molfese, 1997). Following correlational analyses, regression was used to determine if latency differences within early and late ERP components, when included in a model with NWR, explained unique variance within the PLS-5 total language score while accounting for age.

Early Component (136–220 ms: Left Temporal)

An early component loaded onto a cluster of 10 electrodes located over the left temporal region and encompassed a time frame between 136–220 ms (see Figure 1 for electrode montage and Figure 2 for waveforms). The average amplitude for the early component within the new condition was 1.64 μV ($SD = 2.44$) and within the old condition was 1.80 μV ($SD = 2.25$). The differences in amplitude between the new condition and the old condition were not statistically significant: $t(39) = -0.33$, $p = .75$. The average latency of the early component within the new condition was 182.05 ms ($SD = 21.81$) and within the old condition was 186.23 ms ($SD = 21.33$). The differences in latency between the new condition and the old condition were not statistically significant: $t(39) = -1.06$, $p < .30$ for the group average. There was a positive correlation among the latency difference between the new and old conditions and the PLS- 5 total language score as well as five out of the six individual language measures (see Table 2 for correlations and Figure 3 for graph). A correlation was also present with NWR. The direction of the correlation suggests that as language skills increased, so too did the difference between the old and new conditions such that the neural response recorded to the new tokens was faster than the response to the old tokens.

Figure 1. Electrode montage for early component (136–220 ms): left temporal variable.



To address the question of whether ERP latencies corresponding to changing phonemic stimuli explain a significant amount of variance within language when added to a model with NWR, a regression analysis was conducted to predict the PLS-5 total language score from the latency differences in the early ERP component, NWR, and Age (see Table 3, Number 1). The model including the three factors was significant. ERP latency differences within this early component, NWR, and Age accounted for 39% ($R^2 = .39$) of the variance in the PLS-5 total language score. ERP latency difference was significant when predicting the PLS-5 total language score ($p = .05$) as well as NWR ($p < .001$). Age approached significance when predicting the PLS-5 total language score ($p = .06$). As can be seen by the beta weights, NWR is the strongest predictor of language skill ($\beta = .45$), followed by the early component ERP latency difference ($\beta = .28$). This suggests that phonemic perception measured within the ERP early component is significant and explains a unique amount of variance in the PLS-5 total language score separately from NWR. Further, NWR explains a unique amount of variance in the PLS-5 total language score in addition to the ERP latency difference within the early component (136–220 ms).

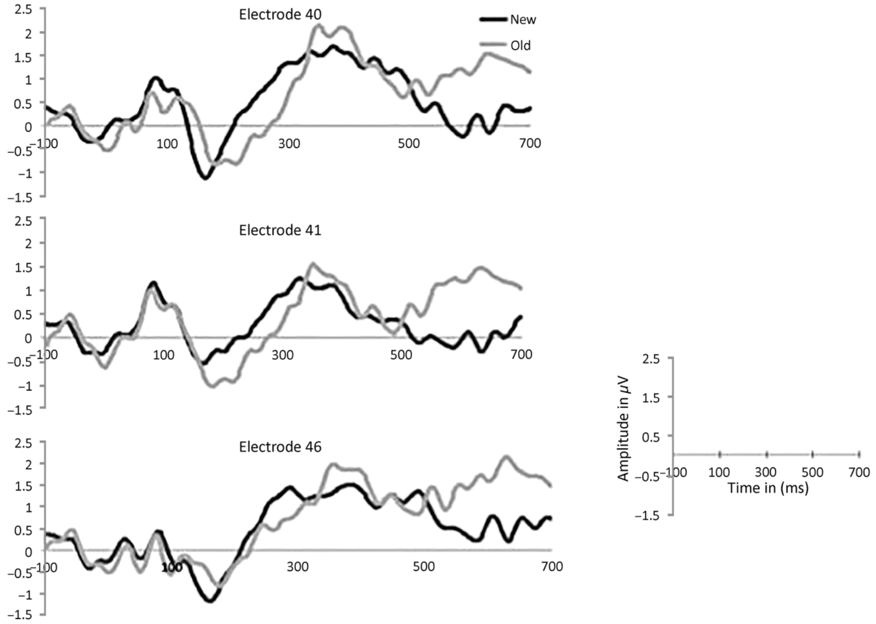
Late Component (248–360 ms: Midline Frontal)

A late component loaded onto a cluster of 35 electrodes located in the midline frontal cortical region and encompassed

---

[2] Two participants did not have complete data sets. Correlational analyses for ERP data and mCDI2_WP as well as the GFTA-2 include $n = 39$.

Figure 2. Averaged waveforms for early component (136–220 ms): left temporal variable.



a time frame between 248–360 ms (see Figure 4 for electrode montage and Figure 5 for waveforms). The average amplitude within this late component for the new condition was 1.86 μV (*SD* = 2.50) and for the old condition was 1.43 μV (*SD* = 3.36). These differences in amplitude between the new condition and the old condition were not statically significant, *t*(39) = 0.82, *p* = .42. The average latency of the late component within the new condition was 307.77 ms (*SD* = 24.32) and within the old condition was 316.77 ms (*SD* = 20.50). The differences in latency between the new condition and the old condition were statically significant, *t*(39) = −2.17, *p* = .04; therefore, the peak latency in response

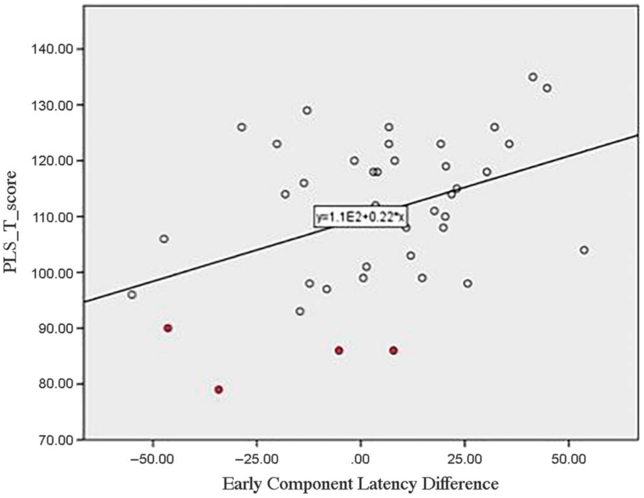
Table 2. Correlations among event-related potential latency difference and behavioral measures for principal components analysis temporal/spatial pairings.

ERP Pairing	PLS-T	mCDI-2-WP	PLS-AC	PLS-EC	GFTA-2	MLU	PCC	TENR_T
TF1_SF1	.22							
TF1_SF2	−.16							
TF1_SF3	.16							
TF1_SF4	.15							
TF2_SF1 (LC)	.41**	.49**	.40*	.39*	.36*	.36*	.41**	.51**
TF2_SF2	−.17							
TF2_SF3	−.12							
TF2_SF4	.02							
TF3_SF1	−.16							
TF3_SF2	−.15							
TF3_SF3	.10							
TF3_SF4	−.01							
TF4_SF1	−.11							
TF4_SF2 (EC)	.41**	.34*	.33*	.43**	.46**	.23	.32*	.31*
TF4_SF3	−.26							

Note. PLS-T = total language standard score on the Preschool Language Scale–Fifth Edition; mCDI-2-WP = raw count of words produced on the MacArthur–Bates Communicative Development Inventories–Second Edition; PLS-AC = auditory comprehension standard score on the Preschool Language Scale–Fifth Edition; PLS-EC = expressive communication standard score on the Preschool Language Scales–Fifth Edition; GFTA-2 = standard score on the Goldman-Fristoe Test of Articulation–Second Edition; MLU = mean length of utterance; PCC = percent consonants correct; TENR_T = total score (syllables plus phonemes) on the Test of Early Nonword Repetition; TF = temporal factor; SF = spatial factor; LC = late component; EC = early component.

**p* < .05, ***p* < .01.

Figure 3. Correlation between latency difference within the early event-related potential component (136–220 ms) and the Preschool Language Scale–Fifth Edition total language score.



to the new stimuli occurs "faster" than the old stimuli. There were positive correlations between the difference in latency in the new and old conditions with PLS-5 total language score as well as with all of the individual language measures (see Table 2 for correlations and Figure 6 for graph). This suggests that faster responses to new stimuli, relative to old stimuli, are associated with better language performance across all measured domains of language.

To determine if ERP latency difference independently predicted the PLS-5 total language score beyond that of NWR and Age, a regression analysis was conducted to predict the PLS-5 total language score from the late component latency difference, NWR, and Age (see Table 3, Number 2). Given the high multicolinearity among the predictors in this model, specifically between the ERP data and NWR ($r = .50$, $p < 0.01$), independent regressions were necessary to determine the amount of variance explained within the PLS-5 total language score for the late component latency difference separately from NWR.

When predicting the PLS-5 total language score from the ERP late component latency difference and Age, the model was significant (see Table 3, Number 2). The ERP latency difference and Age accounted for 22% ($R^2 = .22$) of the variance in the PLS-5 total language score. The ERP latency difference within the late time frame significantly predicted the PLS-5 total language score ($\beta = .43$, $p = .01$); however, Age was not a predictor when included in the model with ERP latency difference within the late component ($p = .14$). This suggests the ERP latency difference has a strong relationship to language skill regardless of age. A final regression was run predicting the PLS-5 total language score from NWR and Age. As seen in Table 3, Number 3, the model was significant. Both NWR significantly predicted the PLS-5 total language score ($\beta = .54$, $p < .001$); however, Age was not a significant predictor when included in the model with NWR ($p = .07$). This suggests that NWR is strongly associated with PLS-5 total language score regardless of age.

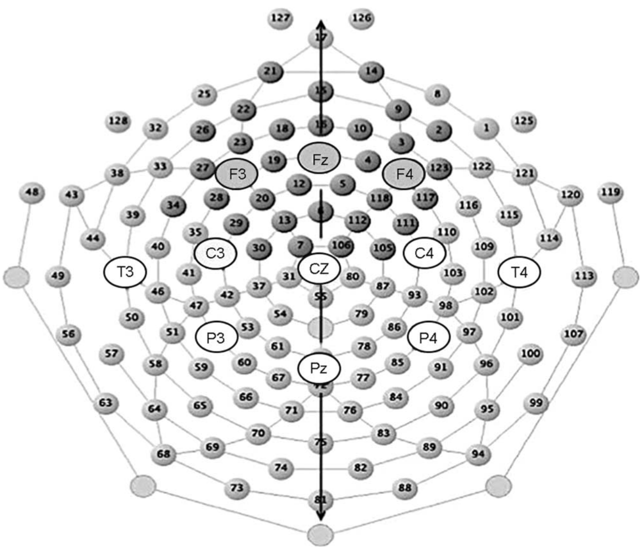
Discussion

The first aim of the current investigation was to determine if ERP indices of phonological sensitivity were associated with clinical assessments used in the field of speech and language pathology. The results suggested that ERP latency differences in response to changing phonological information at the sublexical level were strongly associated with language skills measured by clinical assessments. The PCA analysis yielded two time frames of interest, an early component located in the left temporal region (136–220 ms) and a later component (248–360 ms) located in the midline frontal region. For both components, differences in latency between old and new stimuli were associated with language performance such that, as language performance increased, the neural response to new nonword stimuli was faster than the neural response to the old nonword stimuli. This finding is consistent with other studies demonstrating that ERP measures indexing phonemic discrimination are linked to higher language skill in children (Kuhl et al., 2005; Molfese, 2000; Torkildsen et al., 2009). Specifically, Kuhl and colleagues (2005) found that infants who demonstrated keen discrimination between native phonemic contrasts presented higher language abilities at 2 years of age, compared to children who demonstrated discrimination of both native and nonnative contrasts. Children who are high language performers may be more proficient at processing

Table 3. Multiple regressions predicting the language factor score from event-related potential data, nonword repetition (NWR), and age.

Model				Independent variables				
Number	R^2	F	p	Variable function	Variable	β	t	p
1	.39	7.65	.00	Control	Age	–.26	–1.97	.06
				Phono working memory	NWR	.45	3.28	.00
				Phono sensitivity	Early Component, Latency Difference	.28	2.02	.05
2	.22	5.16	.01	Control	Age	–.22	–1.53	.14
				Phono sensitivity	Late Component, Latency Difference	.43	2.95	.01
3	.32	8.71	.00	Control	Age	–.26	–1.89	.07
				Phono working memory	NWR	.54	3.95	.00

Figure 4. Electrode montage for late component (248–360 ms): midline frontal variable.



changes in phonemic information within their native language when compared to children with lower language ability.

The second aim of the study was to investigate whether perceptual sensitivity to phonemic changes measured by ERP could uniquely predict language skill independently from phonological working memory (measured by NWR).

We hypothesize that both ERP measures of phonological sensitivity and phonological working memory would represent a distinct portion of variance, demonstrating that these skills are distinct and each uniquely contribute to language performance. The results from regression analyses of latency differences between new and old stimuli within early component (136–220 ms) suggested rapid phonemic discrimination was significant when predicting language skill separate from NWR and Age (see regression Table 3, Number 1). One interpretation of this finding is that these early discrimination differences are indexing the neural encoding of distinct speech features (i.e., the place of articulation of stop consonants, voicing features, and vowel features) or integrating those features to support discrimination of phonemic differences. 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).

The findings associated with our later component (248–360 ms) seem to capture a level of phonological processing that is highly associated with language development in young children, based on significant correlations between faster responses for the new phonemic stimuli and all of our language assessments. When latency differences between old and new stimuli in this late window were included in a regression model with NWR and Age, ERP measures of

Figure 5. Averaged event-related potential waveforms for late component (248–360 ms): midline frontal variable.

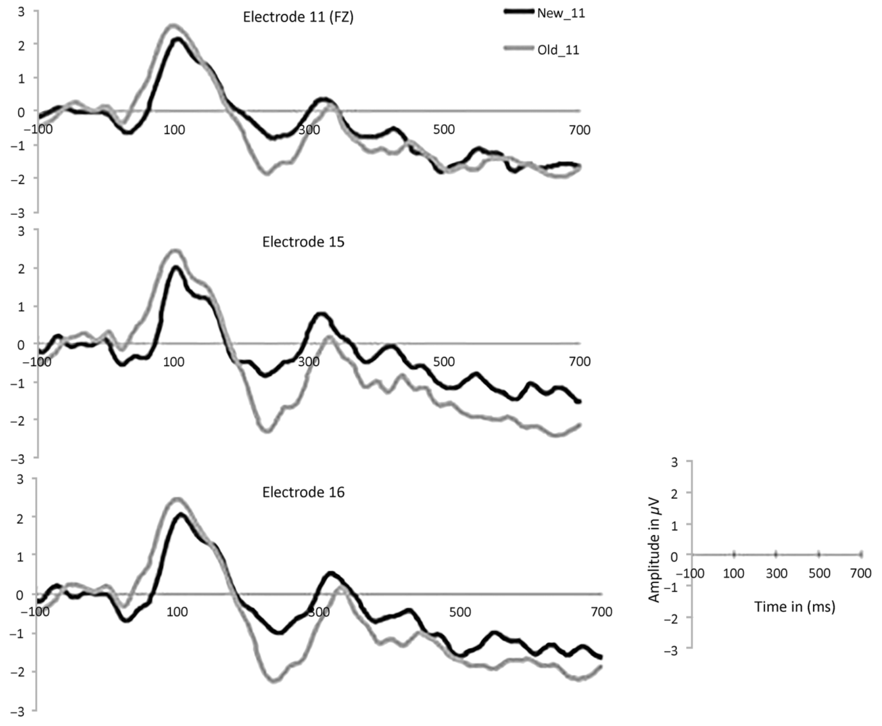
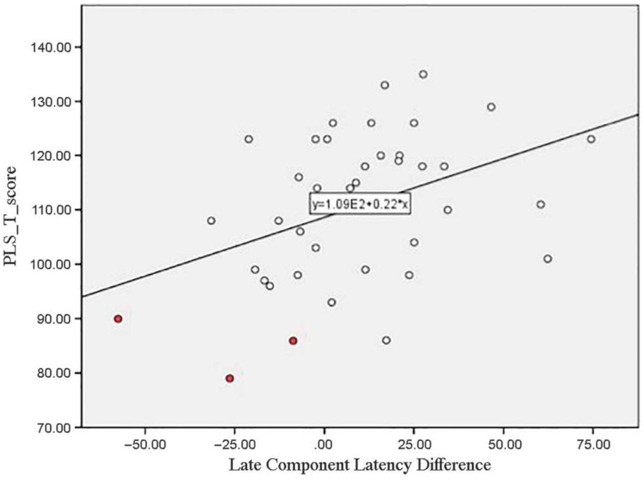


Figure 6. Correlation between the latency difference within the late event-related potential component (248–360 ms) and the Preschool Language Scale-5 total language score.



perceptual sensitivity did not significantly predict the PLS-5 score independently of NWR, despite this late component being highly correlated with all behavioral measures of language. The lack of independent prediction was due to the high correlation between our ERP measure and NWR. Therefore, we suggest that the nature of the processing captured during this time window may be a neural signature of the same processes that underlie NWR, specifically phonological working memory.

Furthermore, the late component waveforms showed a positive peak recorded in the frontal midline region of the scalp with the average maximum peak occurring around 308 ms. It is possible that the late component is reflecting the P3a component. In a review of the P300 effect, Linden (2005) reported that both attention and working memory are measured within the P300 time window such that recognition of the deviant stimulus is supported by working memory, which maintains the features of the standard stimulus for comparison against the deviant within a passive listening condition. Bonala and Jansen (2012) devised a computational model that mimics the learning mechanisms associated with the P300 component. Their model supported the P300 effect being elicited from a working memory process.

ERP latency differences between old and new tokens may capture aspects of phonemic processing that are sensitive to differences in neuroanatomical structures between high and low language performers. This finding of greater speed for the processing of novel phonological stimuli may reflect effective signal conduction of the neural response within white matter myelin tracts (Eggermont, 1988). During early growth, myelination is rapidly increasing. Differences in the growth of white matter among other neural processes may contribute to the fine-grained linguistic differences present in toddlerhood and account for heterogeneity within the population.

Theoretical Implications

The results of this study support the general premise of the PDT. Both ERP measurements of phonemic sensitivity and NWR significantly predicted language skill within this toddler sample. The processes represented by an early and a late ERP component identified in this study are both broadly related to phonological processing. We suggest that each time frame measures auditory processing along a time continuum from basic discrimination of auditory features (early) to more complex process of phonological working memory (later), which entails a series of processes requiring discrimination, storage, and encoding of phonological features.

Clinical Implications

At the present time, the toddler population presents significant clinical challenges due to limitations of behavioral measures of language. This study utilized two experimental methodologies, ERP and NWR, for measuring language in the young child population. NWR may provide insight into language learning and support clinical decision-making for young children when used in tandem with other standardized behavioral language assessments (Clark et al., 2012; Roy & Chiat, 2004; Stokes & Klee, 2009).

Advances in ERP technologies may also one day yield improvements in the technique not only to support clinical practice, but also to enhance understanding of the neural basis for perception and production of language. When considered together, ERPs and NWR show promise in providing insight into language functioning and may ultimately provide critical information to improve identification of language impairment in young children.

Our results support the theory that phonological processing has a significant relationship to language performance. These results support models of language that link higher level linguistic processing of speech features to speech and language production (Hickock, 2012). Many language interventions used with young children focus on whole-word approaches, yet few methods are designed to offer explicit training of phonological input. 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 (Paul & Jennings, 1992; Rvachew & Grawburg, 2006).

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 four of which demonstrated language impairment. More children, especially children demonstrating language delays, are needed to improve generalization of the results to impaired populations. Furthermore, the current study includes only one phonemic contrast in our ERP experiment to index phonological perception; inclusion of additional contrasts would provide a more robust measure of phonological perception and should be a target for future research examining the relationship between phonological perception abilities and language performance.

General Conclusions

Many studies have linked perceptual abilities measured by ERPs in infancy to later language and language-related skills, such as reading (Guttorm et al., 2005; Guttorm et al., 2001; Molfese, 1995, 2000); however, few studies have explored how ERPs to spoken language relate to behavioral language performance on a variety of clinical assessments within the toddler population. The study showed that pairing neurolinguistic methodology and psycholinguistic theory might have utility in understanding clinical aspects of language acquisition.

The current study suggests that phonological processing abilities as measured by ERPs are significantly correlated with behavioral measures of language such that, as language performance increased, the participants' response to novel stimuli was faster. Furthermore, phonological working memory, measured behaviorally or electrophysiologically, significantly predicted language performance within toddlers. Future studies are needed to explore how phonological processing abilities are related to early language development and later language performance.

Acknowledgments

This research is supported by a 2012 Student Research Grant in Early Childhood Language Development awarded to Vanessa Harwood by the American Speech-Language-Hearing Foundation, Rockville, MD, and a donation to Haskins Laboratories by a generous philanthropist who wishes to remain anonymous. Special thanks are given to Michael Harwood, Sayako Earle, Dana Arthur, Julia Irwin, and Peter Molfese who provided helpful suggestions about the article.

References

Archibald, L. M. D., & Gathercole, S. E. (2007). Nonword repetition in specific language impairment: More than a phonological short-term memory deficit. *Psychonomic Bulletin & Review, 14,* 919–924.

Benasich, A., & Tallal, P. (2002). Infant discrimination of rapid auditory cues predicts later language impairment. *Behavioral Brain Research, 136,* 31–49.

Bonala, B., & Jansen, B. (2012). A computational model for generation of the P300 evoked potential component. *Journal of Integrative Neuroscience, 11,* 277–294.

Clark, N., McRoberts, G., Van Dyke, J., Shankweiler, D., & Braze, D. (2012). Immediate memory for pseudoword and phonological awareness are associated in adults and pre-reading children. *Clinical Linguistics and Phonetics, 26,* 577–596.

Dien, J. (2010). The ERP PCA Toolkit: An open source program for advanced statistical analysis of event related potential data. *Journal of Neuroscience, 187,* 138–145.

Dien, J., Khoe, W., & Mangun, G. (2007). Evaluation of PCA and ICA of simulated ERPs: Promax and infomax rotations. *Human Brain Mapping, 28,* 742–763.

Dollaghan, C., & Campbell, T. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41,* 1136–1146.

Ebbels, S., Dockrell, J., & van de Lely, H. (2012). Nonword repetition in adolescents with specific language impairment. *International Journal of Language and Communication Disorders, 47,* 257–273.

Eggermont, J. (1988). On the rate of maturation of sensory evoked potentials. *Acta Otolaryngologica, 70,* 293–305.

Fenson, C., Marchman, V., Thal, D., Dale, P., Reznick, J., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories, Users Guide and Technical Manual, Second Edition.* Baltimore, MD: Brookes.

Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation.* Circle Pines, MN: AGS.

Gratton, G., Coles, M., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology, 55,* 468–484.

Guttorm, T., Leppanen, P., Poikkeus, A.-M., Eklund, K., Lyytinen, P., & Lyytinen, H. (2005). Brain event-related potentials (ERPs) measured at birth predict later language development in children with and without familial risk for dyslexia. *Cortex, 41,* 291–303.

Guttorm, T., Leppanen, P., Richardson, U., & Lyytinen, H. (2001). Event-related potentials and consonant differentiation in newborns with familial risk for dyslexia. *Journal of Language Disorders, 34,* 534–544.

Heilmann, J., Nockerts, A., & Miller, J. (2010). Language sampling: Does the length of the transcript matter? *Language, Speech and Hearing Services in Schools, 41,* 393–404.

Hickock, G. (2012). Computational neuroanatomy of speech productions. *Nature Reviews Neuroscience, 13,* 135–145.

Joanisse, M., & Seidenberg, M. (2003). Phonology and syntax in specific language impairment: Evidence from a connectionist model. *Brain and Language, 86,* 40–56.

Junghofer, M., Elbert, T., Tucker, D., & Braun, C. (1999). The polar average reference effect: A bias in estimating the head surface integral in EEG recording. *Clinical Neurophysiology, 110,* 1149–1155.

Korczak, P., & Stapells, D. (2010). Effects of various articulatory features of speech on cortical event-related potentials and behavioral measures of speech-sound processing. *Ear Hearing, 31,* 491–504.

Kuhl, P., Conboy, B., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: Implications for the "critical period." *Language Learning and Development, 1,* 237–264.

Kuhl, P., Conboy, B., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and the native language magnet theory expanded. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363,* 979–1000.

Landi, N., Crowley, M., Wu, J., Bailey, C., & Mayes, L. (2012). Deviant ERP response to spoken non-words among adolescents exposed to cocaine in utero. *Brain and Language, 120,* 209–216.

LENA Research Foundation [Website]. (2014). Retrieved from http://www.lenafoundation.org

Linden, D. (2005). The P300: Where in the brain is it produced and what does it tell us? *The Neuroscientist, 11,* 563–576.

Long, S. (2008). *Computerized Profiling, Version 9.7.* Retrieved from http://www.computerizedprofiling.org

Molfese, D. (1995). Electrophysiological responses obtained during infancy and their relation to later language development: Further findings. In M. Tramontana & S. R. Hooper (Eds.), *Advances in child neuropsychology* (pp. 1–11). New York, NY: Springer-Verlag.

Molfese, D. (2000). Predicting dyslexia at 8 years of age using neonatal brain responses. *Brain and Language, 72,* 238–245.

Molfese, D., & Molfese, V. (1985). Electrophysiological indicies of auditory discrimination in newborn infants: The basis of predicting later language development. *Infant Behavior and Development, 8,* 197–211.

Molfese, D., & Molfese, V. (1997). Discrimination of language skills at five years of age using event related potentials recorded at birth. *Developmental Neuropsychology, 13,* 135–156.

Molfese, D., Morse, P., & Peters, C. (1990). Auditory evoked responses to names for different objects: Cross-modal processing as a basis for infant language acquisition. *Developmental Psychology, 26,* 780–795.

Molfese,, D., Nunez, V., Seibert, S., & Ramanaiah, N. (1976). Cerebral asymmetry: Changes in factors affecting its development. *Origins and Evolution of Language and Speech, 280,* 821–833.

Mullen, E. (1995). *Mullen Scales of Early Learning.* San Antonio, TX: NCS Pearson.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., . . . Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature, 385,* 432–434.

Paul, R., & Jennings, P. (1992). Phonological behavior in toddlers with slow expressive language development. *Journal of Speech and Hearing Research, 37,* 99–107.

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology, 72,* 184–187.

Preston, J., & Edwards, M. L. (2010). Phonological awareness and types of sound errors in preschoolers with speech sound disorders. *Journal of Speech, Language, and Hearing Research, 53,* 44–60.

Rescorla, L. (2009). Language and reading outcomes in late-talking toddlers: Support for dimensional perspective on language delay. *Journal of Speech, Language, and Hearing Research, 52,* 16–30.

Rescorla, L., Roberts, J., & Dahlsgard, K. (1997). Late talkers at 2: Outcomes at age 3. *Journal of Speech, Language, and Hearing Research, 40,* 556–566.

Rivera-Gaxiola, M., Klarman, L., Garcia-Sierra, M., & Kuhl, P. (2005). Neural patterns to speech and vocabulary growth in American infants. *NeurorReport, 16,* 495–498.

Roy, P., & Chiat, S. (2004). A prosodically controlled word and nonword repetition task for 2–4 year olds: Evidence from typically developing children. *Journal of Speech, Language, and Hearing Research, 47,* 223–234.

Rvachew, S., & Grawburg, M. (2006). Correlates of phonological awareness in preschoolers with speech sound disorders. *Journal of Speech, Language, and Hearing Research, 49,* 74–87.

Stokes, S., & Klee, T. (2009). The diagnostic accuracy of a new test of early non-word repetition for differentiating late talking and typically developing children. *Journal of Speech, Language, and Hearing Research, 52,* 872–882.

Thal, D., Miller, S., Carlson, J., & Moreno-Vega, M. (2005). Nonword repetition and language development in 4-year-old children with and without a history of language delay. *Journal of Speech, Language, and Hearing Research, 48,* 1481–1495.

Torkildsen, J. V. K., Hansen, H. F., Svangstu, J. M., Smith, L., Simonsen, H. G., Moen, I., & Lindgren, M. (2009). Brain dynamics of word familiarization in 20-month-olds: Effects of productive vocabulary size. *Brain and Language, 108,* 73–88.

Weber-Fox, C., Leonard, L., Hampton, W., & Tomblin, B. (2010). Electrophysiological correlates of rapid auditory and linguistic processing in adolescents with SLI. *Brain and Language, 115,* 162–181.

Zimmerman, I., Steiner, V., & Pond, R. (2011). *Preschool Language Scale– Fifth Edition: Examinier's Manual.* San Antonio, TX: The Psychological Corporation.

Appendix A

List of Nonwords for Test of Nonword Repetition task

/mɑd/
/neɪd/
/paɪm/
/boʊz/
/koʊgə/
/dɑfi/
/leɪpoʊ/
/moʊkɚi/
/doʊpəlut/
/bæləkɔn/
/fisaɪmɔt/
/pɛduləmeɪp/
/fɛnɚalsɛkʰ/
/wugɛləmɛkʰ/

Appendix B

Correlations Between ERP Amplitude Difference and the PLS-5 Total Language Score

Spatial factor	Temporal factor			
	1	2	3	4
1	.15	−.04	−.12	−.02
2	.03	−.07	.10	−.12
3	−.06	.14	−.03	.04
4	.14	.01	.07	

*p < .05, **p < .01.