Developmental trends in auditory processing can provide early predictions of language acquisition in young infants

Weerasak Chonchaiya,1,2 Twila Tardif,1,3 Xiaoqin Mai,1,4 Lin Xu,5 Mingyan Li,5 Niko Kaciroti,1 Paul R. Kileny,1,6 Jie Shao5 and Betsy Lozoff1,7

1. Center for Human Growth and Development, University of Michigan, USA
2. Division of Growth and Development, Department of Pediatrics, Faculty of Medicine, King Chulalongkorn Memorial Hospital, Chulalongkorn University, Thailand
3. Department of Psychology, University of Michigan, USA
4. Department of Psychology, Renmin University, China
5. Children’s Hospital Zhejiang University School of Medicine, China
6. Department of Otorhinolaryngology, University of Michigan, USA
7. Department of Pediatrics and Communicable Diseases, University of Michigan, USA

Abstract

Auditory processing capabilities at the subcortical level have been hypothesized to impact an individual’s development of both language and reading abilities. The present study examined whether auditory processing capabilities relate to language development in healthy 9-month-old infants. Participants were 71 infants (31 boys and 40 girls) with both Auditory Brainstem Response (ABR) and language assessments. At 6 weeks and/or 9 months of age, the infants underwent ABR testing using both a standard hearing screening protocol with 30 dB clicks and a second protocol using click pairs separated by 8, 16, and 64-ms intervals presented at 80 dB. We evaluated the effects of interval duration on ABR latency and amplitude elicited by the second click. At 9 months, language development was assessed via parent report on the Chinese Communicative Development Inventory - Putonghua version (CCDI-P). Wave V latency z-scores of the 64-ms condition at 6 weeks showed strong direct relationships with Wave V latency in the same condition at 9 months. More importantly, shorter Wave V latencies at 9 months showed strong relationships with the CCDI-P composite consisting of phrases understood, gestures, and words produced. Likewise, infants who had greater decreases in Wave V latencies from 6 weeks to 9 months had higher CCDI-P composite scores. Females had higher language development scores and shorter Wave V latencies at both ages than males. Interestingly, when the ABR Wave V latencies at both ages were taken into account, the direct effects of gender on language disappeared. In conclusion, these results support the importance of low-level auditory processing capabilities for early language acquisition in a population of typically developing young infants. Moreover, the auditory brainstem response in this paradigm shows promise as an electrophysiological marker to predict individual differences in language development in young children.

Introduction

Development of the human auditory system, particularly the auditory cortex, is complex and requires more time to attain maturation when compared to other primary sensory systems (Moore & Linthicum, 2007). Moreover, compared to other sensory systems, the auditory system is temporally extremely precise both at peripheral sensory receptors and at the level of the auditory cortex (Phillips, 1993; Wang, 2007), and is thus highly sensitive to conditions that affect myelination (Algarin, Peirano, Garrido, Pizarro & Lozoff, 2003; Jiang, 1995). During the perinatal period (third trimester to 6 months of age postnatally), there is a rapid increase in axonal myelination, dendritic arborization, and synaptic organization in brainstem pathways from the proximal cochlear nerve to the thalamus. Although these auditory system pathways begin to function early, they do not reach maturity until

Address for correspondence: Twila Tardif, Center for Human Growth and Development, Department of Psychology, University of Michigan, 300 N. Ingalls, 10th Floor, Ann Arbor, MI 48109, USA; e-mail: twila@umich.edu.

© 2012 Blackwell Publishing Ltd, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.
roughly 2 years of age (Jacobson, 1985; Johnson, Nicol, Zecker & Kraus, 2008; Moore & Linthicum, 2007). The period from birth to age 2, in particular, represents the peak of brainstem input to the auditory cortex through the cortical marginal layer (Moore & Linthicum, 2007). It also corresponds to the period of initial and extremely rapid language development (9 to 24 months of age), which relies on efficient processing and organization of auditory information, and thus is an important period for assessing the potential relations between auditory processing and very early language development.

Normal language acquisition involves the development and complex coordination of multiple processes including sensory (both visual and auditory), motor (speech organs), and cognitive (memory, cross-modal mapping of sounds to visual stimuli and events) pathways in addition to the provision of an enriched language environment with responsive caregivers. In particular, the auditory functions required for speech language skills include auditory perception, discrimination, processing, and neural encoding of auditory information (Benasich & Tallal, 2002; Benasich, Thomas, Choudhury & Leppänen, 2002). Decreased abilities to perceive, process, and distinguish auditory stimuli may impact an individual’s ability to develop normal language acquisition (Benasich & Tallal, 2002; Benasich et al., 2002; Bishop & McArthur, 2004; Choudhury & Benasich, 2011; Kuhl, Conboy, Coffey-Corina, Padden, Rivera-Gaxiola & Nelson, 2008; Molfese, 1989; Wible, Nicol & Kraus, 2004). Moreover, a handful of studies have now shown that discrimination of speech stimuli measured by auditory evoked responses recorded from newborn infants predicted group differences in language development at 3 years of age (Molfese, 1989) and reading abilities much later (Espy, Molfese, Molfese & Modglin, 2004; Leppänen et al., 2010; Molfese, 2000).

In addition, individual differences in auditory ERPs to native vs. non-native contrasts in infancy have also been found to predict spoken language abilities at ages 2 and 3 (Kuhl et al., 2008; Rivera-Gaxiola, Silva-Pereya & Kuhl, 2005). Moreover, infants with heightened sensitivity to non-native contrasts at earlier processing stages (P150-250) showed greater vocabulary growth in the second and third year of life than those with heightened sensitivity in a later ERP component (N250-550). Finally, deficits in auditory processing to either basic non-linguistic or speech stimuli have been reported in children with specific language impairment (SLI) (Tomblin, Records, Buckwalter, Zhang, Smith & O’Brien, 1997) and/or a positive family history, and thus elevated risk for, language-related learning impairments, as demonstrated by various behavioral (Benasich & Tallal, 2002; Benasich et al., 2002) and electrophysiological studies including both cortical event related potentials (ERP) (Bishop & McArthur, 2004; Choudhury & Benasich, 2011; Cunningham, Nicol, Zecker, Bradlow & Kraus, 2001; Kuhl et al., 2008; Leppänen et al., 2010), and the subcortical auditory brainstem response (ABR) (Basu, Krishnan & Weber-Fox, 2010; King, Warrier, Hayes & Kraus, 2002; Song, Banai & Kraus, 2008; Wible, Nicol & Kraus, 2005).

Nonetheless, it has been challenging to demonstrate that individual differences in low-level auditory processing are related to language development using behavioral or neuroimaging paradigms (functional magnetic resonance imaging (fMRI), etc.) in samples of typically developing infants. Auditory processing results derived from behavioral measures sometimes can be difficult to interpret and are not highly reliable at the individual level, since they may rely on other factors, such as attention, working memory, task demands, learning processes, temperament, and cognitive ability (Choudhury, Leppänen, Leevers & Benasich, 2007). Neuroimaging measures also present challenges. Functional MRI measures, for instance, are inherently difficult because it is difficult to examine auditory responses in fMRI machines given the noise involved in the magnet’s operation, and it is not clear which tasks would be reliable enough to discriminate amongst populations at the individual level. Recent developments in both cortical ERP and fNIRS technologies and techniques have made imaging more possible for young infants, and yet neither technique is yet ready for diagnostic purposes in a typically developing population.

Compared to complex behavioral or neuroimaging measures, the ABR is a simple, sensitive, reliable, relatively robust, non-invasive, and practical electrophysiological tool for examining both the integrity and the efficiency of low-level auditory processing, particularly in infants. The ABR, both for simple click and speech stimuli, is a highly reliable measure that assesses auditory pathways necessary for auditory processing from the periphery to the brainstem (subcortical auditory function) (Despland & Galambos, 1980; Jacobson, 1985; Song, Nicol & Kraus, 2011). The ABR to click stimuli has been used both as a hearing screening tool and a neurodiagnostic measure in many conditions, including conductive hearing loss, cochlear lesions, tumors of the auditory nerve, and brainstem lesions (Despland & Galambos, 1980; El-Kashlan, Eisenmann & Kileny, 2000; Jacobson, 1985; Van Riper & Kileny, 1999). The ABR components in newborns and young infants typically consist of three well-defined wave forms, Wave I (representing activity at the cochlear nerve), Wave III (activation at the cochlear nucleus), and Wave V (reflecting activity at the lateral lemniscus) (Jacobson,
discrete stimuli, the cochlear nerve increases stimulus rate and very short intervals between stimuli. Of interest is the fact that in both cases of system to recover and prepare for the next pair of presentation rate in that there is a sufficiently long Burkard, 1999). It is different from simply varying the temporal processing and frequency discrimination of the adult and infants for the purpose of investigating the components in adults, albeit with longer latencies because of incomplete myelination of the auditory system in infants, relative to adults and older children. Wave V, which occurs roughly 6 msecs after stimulus onset, is frequently used as an indicator both of auditory processing and of the neurological integrity of the auditory system (Hecox & Galambos, 1974).

The present study uses an ABR protocol that involves two successive stimuli separated by differing intervals (8, 16, and 64 ms) to determine infants’ capabilities to discriminate rapidly presented auditory information in subcortical pathways (Lasky, 1991; Lasky & Rupert, 1982). Although this is similar to the effects obtained by varying click-stimulus presentation rates in older children (e.g. Basu et al., 2010), it is not entirely identical. In the paradigm used in the present study, we are particularly interested in examining the effects on the response elicited by the second click stimulus, as a function of the timing of the preceding click. This is also known as a ‘forward masking’ paradigm (e.g. Lasky & Rupert, 1982). As with varying the stimulus presentation rate more generally, varying the intervals between pairs of stimuli (potentially creating a ‘masking’ effect of one stimulus by another) has been used as a tool in both adults and infants for the purpose of investigating the temporal processing and frequency discrimination of the auditory system (Lasky, 1991; Walton, Orlando & Burkard, 1999). It is different from simply varying the presentation rate in that there is a sufficiently long interval after the second stimulus to allow the nervous system to recover and prepare for the next pair of stimuli. Of interest is the fact that in both cases of increased stimulus rate and very short intervals between discrete stimuli, the cochlear nerve’s action potential response is reduced in the presence of a competing acoustic signal. In such paradigms, as long as the signals are spaced closely in time (e.g. at 70 ms or less), both signals activate auditory nerve fibers, and nerve fibers responding to the first stimulus have not recovered to become available to respond to the second stimulus. The shorter the interval between stimuli, the longer it takes for the auditory nerve fibers to recover and become available to respond (Abbas & Gorga, 1981).

The ability to process auditory stimuli presented in rapid succession or in noise also continues to develop through early childhood (Choudhury & Benasich, 2011; Johnson et al., 2008; Lasky, 1991). Moreover, because individual differences in rapid auditory processing have been found to be related to language and reading difficulties (Tallal, 2004), we hypothesized that children with longer Wave V latencies in a paradigm that presents a temporal challenge to the auditory system will subsequently show slower language acquisition, even, and perhaps especially, at the very earliest stages of language acquisition.

The choice of the ABR for this study was based on several response characteristics and diagnostic sensitivity. First, this is a response that has been extremely well documented in the pediatric population, and while there are latency and amplitude changes with neuro-maturational, those changes are minimal, when compared to the more central auditory evoked potentials. The infant response is very similar in configuration to the adult response and becomes adult-like by about 2 years of age. Furthermore, the ABR, as opposed to the later auditory potentials, is impervious to subject or patient state – it remains stable during natural sleep, and is only slightly affected by sedatives and anesthetics. There is a large body of work related to ABR use in determining hearing thresholds in newborns and infants, as well as the effects of brainstem lesions on the ABR. There is also literature indicating that older children with language-based learning problems show longer wave V latencies in ABR paradigms (Banai, Nicol, Zecker & Kraus, 2005; Basu et al., 2010; Cunningham et al., 2001; Marler & Champlin, 2005; Wible et al., 2004). Overall, the ABR is a very reliable response in infants through adults, and is sensitive to changes associated with a variety of conditions affecting auditory perception, as well as higher level processing differences.

Nonetheless, associations between rapid auditory processing capabilities at the subcortical level and very early language development have not been considered in healthy infants, nor have they been examined in a typically developing, rather than at-risk or affected (for language/literacy impairment), population. Recent findings have suggested that individual differences in infants’ cortical auditory ERP responses, for rapidly presented auditory stimuli, pitch discriminations, and even native vs. non-native speech contrasts, are related to language skills in the preschool years and beyond (Choudhury & Benasich, 2011; Kuhl et al., 2008; Leppänen et al., 2010; Rivera-Gaxiola et al., 2005). However, it is not clear whether these differences are a result of solely cortical processes, whether they derive from subcortical processes, or involve interactions between both subcortical and cortical processes in development. Moreover, although the auditory ERP responses show good correlations with language skills at numerous ages from 6 months to 4 years of age, it is not clear whether these correlations are simply stable over age or whether there is any compounded effect from differing rates of development over this period, and when these differences might emerge. Thus, we examined relations between auditory processing capabilities at 6 weeks and at 9 months of age.
and the very earliest stages of language development at 9 months of age in a sample of healthy typically developing infants who passed ABR screening criteria for normal hearing.

**Method**

**Participants**

This study was conducted in conjunction with the first of three periods of enrollment in an ongoing project on Brain and Behavior effects of Early Iron Deficiency – a collaboration between the Children’s Hospital of Zhejiang University, Hangzhou, China, and the University of Michigan, Ann Arbor, USA. Data for this study were obtained during January 2009 to March 2010. All healthy infants who had ABR at 6 weeks and/or 9 months and a language assessment at 9 months were included. The following criteria were used to define ‘healthy’ for this analysis: singleton full-term birth (37–42 weeks gestation) weighing > 2500 grams; no prenatal complications or congenital malformations; no general undernutrition (< 10th percentile for weight or length); no acute or chronic illness, no multiple or prolonged hospitalizations (> 5 days); and not iron-deficient anemic at 9 months. Infants with iron deficiency anemia (low hemoglobin (< 110 g/L) and two or more abnormal iron measures (Centers for Disease Control and Prevention, 1998, 2001; Looker, Dallman, Carroll, Gunter & Johnson, 1997) were not included, because previous studies show longer latencies in Wave I, III, and V, and also prolongation of Wave I–V, and Wave III–V interpeak latencies with iron deficiency anemia in infancy, suggesting delayed myelination and ABR maturation (Algarin et al., 2003; Wang, 2007). Seventy-one infants (31 boys and 40 girls) met inclusion criteria.

The mean age at the 9-month assessment was 279 days (SD = 7). Of 71 participants, 53 (74.6%) lived in rural areas, whereas 18 (25.4%) lived in more urban areas. Of 66 fathers who reported their occupation, 12 (18.2%) worked as managers in a middle-to-large-sized or private industry/company, 16 (24.2%) worked as professional personnel, 16 (24.2%) worked as industrial or commercial entrepreneurs, 18 (27.3%) worked as service staff or workers in rural or urban factory settings, and 4 (6%) were unemployed. Of 71 mothers, 3 (4.2%) worked as managers in a middle-to-large-sized or private industry/company, 6 (8.5%) worked as educational or professional personnel, 8 (11.3%) worked as industrial or commercial entrepreneurs, 22 (31.0%) worked as service staff or workers in rural or urban factory settings, and 32 (45.1%) had no employment outside the home. As such, our participants were generally middle class, as defined in a prosperous rural area of southeastern China. Additional demographic details of study participants are listed in Table 1. The project was approved by the Institutional Review Boards of the University of Michigan and the Children’s Hospital of Zhejiang University. Parents provided written informed consent.

**Study protocol**

**ABR testing**

At 6 weeks and/or 9 months, participants underwent ABR testing with a standard hearing screening protocol at 30 dB and a second protocol which varied the intervals between pairs of clicks by 8, 16, or 64 msecs at 80 dB. ABR testing was administered during a spontaneous nap without sedative medications using a Biologic Navigator (Bio-Logic Systems Corp., Mundelein, IL)/Traveler evoked potential system. ABR responses were recorded with surface silver/silver chloride electrodes attached to participant’s foreheads using adhesive tabs, in the midline below the hairline (non-inverting) and at the mastoid on each side (ipsilateral as...
Auditory brainstem responses for 64-ms (top trace), 16-ms conditions (middle trace), and the result of the subtraction of the 64-ms condition from the 16-ms condition (bottom trace).

Language development

Language development was measured at 9 months of age via parent report on the Communicative Development Inventory - Putonghua version (CCDI-P) Words and Gestures (infant) Short Form, a brief screening instrument that evaluates typically developing children’s early vocabularies between 8 and 16 months of age. This instrument was adapted and normed by Tardif et al. (2008) (Tardif, Fletcher, Zhang, Liang & Zuo, 2008) in a Beijing population sample and is based on the original MacArthur-Bates Communicative Development Inventories in English (CDI) (Fenson et al., 1993). The CDI is a simple, sensitive, validated, reliable, clinically relevant, and age-normed parent report instrument, widely used in many countries for assessment of early language development in very young children, 8–30 months of age. Putonghua, the ‘common language’, is the term used for Mandarin Chinese, the official language in mainland China. The CCDI-P Words and Gestures Short Form consists of a total of 222 items including five short phrases commonly used with young children (phrases understood), five communicative gestures (gestures), and 106 vocabulary words used with young children that are scored ‘able to comprehend’ (words understood) or to produce (words child ‘can say’). Each item of the

The elimination of the first stimulus in the subtracted trace indicates that the first stimulus was not affected by the interval conditions, and hence is considered a valid ‘masker’ for the rest of our analyses. ABR data were analyzed off-line. For each condition (8, 16, and 64 ms), two replications were added yielding a grand-average waveform representing 2600 accepted sweeps per ear that was used in data analysis. With the aid of cursors, the individual waves of each condition were identified and marked by trained technicians. Wave V latency and amplitude of responses elicited by the second stimulus (the ‘probe’) were measured for each condition. Wave V latency was identified as the peak latency. Wave V amplitude was identified as the peak amplitude of Wave V minus the amplitude of the negative trough following Wave V, standardized at a latency of 10 ms after the onset of the second stimulus (i.e. 18, 26, and 74 ms after the onset of the first stimulus, respectively). Per usual practice, the latency and amplitude values obtained for the right and left ears were averaged so that each participant was represented by one value, derived from a total of 5200 sweeps per condition, in any given group mean. Of 71 participants whose language developmental milestones were assessed at 9 months of age, the ABR for this forward masking protocol was obtained in 54 and 60 participants at 6 weeks and 9 months of age, respectively. There were 43 participants who received the 64-ms forward masking protocols at both ages.

Auditory processing and prediction of language acquisition
Wave V amplitudes of the 64-ms forward masking condition 0.93
Wave V amplitudes of the 16-ms forward masking condition 0.68
Wave V amplitudes of the 8-ms forward masking condition 0.46
Wave V latencies of the 64-ms forward masking condition 6.56
Wave V latencies of the 16-ms forward masking condition 6.62
Wave V latencies of the 8-ms forward masking condition 6.69

Words child for each section of the CCDI-P Words and Gestures Short Form, and CCDI-P composite is presented in parentheses. Cronbach’s alpha (α) for each section of the CCDI-P Words and Gestures Short Form is presented in parentheses. 4CCDI-P composite included phrases understood, gestures, and words produced.

CCDI-P Words and Gestures Short Form was scored 0 or 1 depending upon participants’ performance, as documented by parent report. The exception was the Gestures section, in which items were scored 0, 1, or 2 if reported as ‘never’ having occurred, ‘sometimes’, or ‘often’ occurring, respectively. Total score for the CCDI-P Words and Gestures Short Form is 227 and for each of the subscales is 5, 10, 106, and 106, respectively. Because of high intercorrelations among subscales, and the reliability benefits of pooling across subscales, a CCDI-P composite was calculated, consisting of phrases understood, gestures, and words produced. Words understood was not included in this composite score because it was not as highly correlated with the other subscales or with the composite score, as shown in Table 3.

Statistical analyses

Statistical analyses were performed with SPSS version 18.0 (IBM, Inc., Somers, NY, USA) for Windows and was used to first compare ABR variables at 6 weeks and 9 months by paired t-tests. Cronbach’s alphas (α) for each major section of the CCDI-P Words and Gestures Short Form and the CCDI-P composite were also computed to examine internal consistency, and are reported in Table 2. In addition, Pearson correlations among various potential predictors and the CCDI-P outcomes were computed and are displayed in Table 4. Finally, a path analysis using Amos version 18.0 was performed to account for multiple correlations and explore effects for both the direct and indirect pathways. P-values reported are two-tailed with p < .05 as the significance level.

Results

Comparison of the ABR variables at 6 weeks and 9 months of age

Paired t-tests were used to compare the change on ABR variables from 6 weeks to 9 months. As shown in Table 2, Wave V latencies decreased for all conditions from 6 weeks to 9 months as myelination of the auditory pathways matured (all p-values < .001). In addition, prolongation of the Wave V latencies and reduction in the Wave V amplitudes of the ABR were demonstrated when masker-probe intervals were decreased (i.e. the temporal challenge to the nervous system was increased).

Intercorrelations among the components of the CCDI-P Words and Gestures Short Form

Table 2 lists descriptive statistics of the CCDI-P Words and Gestures Short Form at 9 months of age. There were moderate correlations (rs = 0.28 to 0.51) between some components of the CCDI-P Words and Gesture Short Form and relatively high correlations (rs = 0.73 to 0.84) between other components, as shown in Table 3. Relations between gestures and vocabulary production were highly significant, even after chronological age, gender, gestational age, and birth weight were controlled.

Path analysis of language outcomes based on potential predictors

Table 4 lists potential predictors of early language development such as gestational age, gender, and Wave V latency z-scores of the 64-ms condition at 6 weeks and...
Table 3  Intercorrelations (simple correlations above diagonal; controlling for chronological age, gestational age, gender, and birth weight, shaded below diagonal) between major sections of the CCDI-P Words and Gestures Short Form (N = 71)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Phrases understood</th>
<th>Gestures</th>
<th>Words understood</th>
<th>Words child ‘can say’</th>
<th>CCDI-P composite</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phrases understood</td>
<td>–</td>
<td>0.37**</td>
<td>0.51**</td>
<td>0.41**</td>
<td>0.73**</td>
</tr>
<tr>
<td>Gestures</td>
<td>0.30*</td>
<td>–</td>
<td>0.28*</td>
<td>0.50**</td>
<td>–</td>
</tr>
<tr>
<td>Words understood</td>
<td>0.49**</td>
<td>0.24†</td>
<td>–</td>
<td>0.35**</td>
<td>0.84**</td>
</tr>
<tr>
<td>Words child ‘can say’</td>
<td>0.42**</td>
<td>0.50**</td>
<td>0.34**</td>
<td>–</td>
<td>0.78**</td>
</tr>
<tr>
<td>CCDI-P composite</td>
<td>0.69**</td>
<td>0.83**</td>
<td>0.42**</td>
<td>0.81**</td>
<td>–</td>
</tr>
</tbody>
</table>

**p < .01; *p < .05; †p < .1.

Table 4  Pearson correlations (simple correlations above diagonal; controlling for gestational age and gender, shaded below diagonal) among potential predictors on the CCDI-P composite

<table>
<thead>
<tr>
<th>Variables</th>
<th>CCDI-P composite</th>
<th>Wave V latency z-scores of the 64-ms forward masking condition at 6 weeks</th>
<th>Wave V latency z-scores of the 64-ms forward masking condition at 9 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gestational age</td>
<td>0.16 (N = 69)</td>
<td>– 0.30* (N = 54)</td>
<td>– 0.42** (N = 58)</td>
</tr>
<tr>
<td>Gender</td>
<td>0.24* (N = 71)</td>
<td>– 0.44* (N = 54)</td>
<td>– 0.58*** (N = 60)</td>
</tr>
<tr>
<td>CCDI-P composite</td>
<td>–</td>
<td>– 0.13 (N = 54)</td>
<td>– 0.40** (N = 60)</td>
</tr>
<tr>
<td>Wave V latency z-scores of the 64-ms forward masking condition at 6 weeks</td>
<td>– 0.02 (N = 39)</td>
<td>–</td>
<td>0.80** (N = 43)</td>
</tr>
<tr>
<td>Wave V latency z-scores of the 64-ms forward masking condition at 9 months</td>
<td>– 0.26 (N = 39)</td>
<td>0.70** (N = 39)</td>
<td>–</td>
</tr>
</tbody>
</table>

***p < .001; **p < .01; *p < .05.

9 months of age (calculated separately at each age group from the mean latencies for each ear for each child in this condition). We used the CCDI-P composite and Wave V latency z-scores of the 64-ms condition in the final model since this condition included the most participants (many infants woke up or had movement artifacts for the later-occurring 8- and 16-msec conditions) and thus provided the most robust results. Importantly, infants were clearly able to resolve all of the masking conditions, as evidenced by the appearance of clear Wave V peaks with stable latencies and amplitudes. Correlations between the latencies for the 8- and 16-msec conditions with the 64-msec, moreover, were above .9 for both the 6-weeks and 9-month testing periods. In addition, it should be noted that the stimulus presentation of the 64-msec condition with an overall presentation rate of 11.7/sec was approximately comparable to the longest stimulus interval in Basu et al.’s study with older children and thus provides a good basis for comparison across studies (Basu et al., 2010). Simple correlations among potential predictors in Table 4 show that shorter Wave V latency z-scores in the 64-ms condition at 9 months and gender (girls) were both correlated with better CCDI-P composite scores for our participants. In addition, girls and those with more mature gestational age tended to have shorter Wave V latency z-scores in the 64-ms condition at both 6 weeks and 9 months. Using Pearson correlations only, it appeared that the only remaining predictors after controlling for gestational age (in days) and gender, as demonstrated below the diagonal, were Wave V latency z-scores at 6 weeks with Wave V latency scores at 9 months. However, given the numerous correlations within and across variables and ages, it is more appropriate to analyze these data using path analyses or other methods (e.g. structural equation modeling) to account for the multiple correlations among factors and to best examine the relative contributions of gender, gestational age, and ABR latencies to 9-month language scores.

A path analysis, shown in Figure 2, for predicting the CCDI-P composite at 9 months was therefore conducted and included Wave V latency z-scores of the 64-ms forward masking condition at 6 weeks and 9 months of age, gestational, and gender. This model showed excellent goodness of fit indicators, including a Chi-square of 0.102 (df 3), p = .992, Comparative Fit Index (CFI) of 1.000, Root Mean Square Error of Approximation (RMSEA) of less than 0.001, and Normed Fit Index (NFI) of 0.999. As was shown in the preliminary correlational analyses, Wave V latency z-scores of the 64-ms condition at 6 weeks were highly correlated with and had strong direct relationships with the 9-month Wave V latency z-scores. The Wave V latency z-scores of the
Wave V latencies at 9 months (i.e. relatively faster auditory processing) at 6 weeks and 9 months (longer/shorter)' processors in this ABR paradigm. As with the model including separate 6-week and 9-month ABRs (Figure 2), this second path analysis, shown in Figure 3, also showed excellent goodness of fit, with a chi-square of 0.102 (df 2), $p = .950$, CFI of 1.000, RMSEA of less than 0.001, and NFI of 0.991. Importantly, even when the four participants who were in the ‘faster’ group at 6 weeks but ‘slower’ at 9 months were excluded from the path analysis, all path coefficients showed the same trends, although the goodness of fit indicators were reduced as would be expected from both a decrease in the extreme cases contributing to this trend and in the number of participants. As shown in Figure 4, in this more parsimonious model of the full sample, the only significant predictor of the 9-month CCDI-P composite was change in Wave V latencies over the
6-week to 9-month period. Specifically, infants who had greater reductions (i.e. showed greater improvements in speed) in Wave V latency z-scores of the 64-ms condition from 6 weeks to 9 months had higher CCDI-P composite scores at 9 months, with both the gestational age and gender effects on the CCDI-P dropping out when ABR latencies and the multiple correlations between measures were taken into account.

**Discussion**

This study demonstrated an association between auditory processing capabilities at the subcortical level and subsequent language development in a population of young healthy infants. The path analysis demonstrated that infants who showed shorter Wave V latency z-scores in an ABR forward masking paradigm with pairs of stimuli spaced 64 ms apart at 9 months and greater improvements in Wave V latencies from 6 weeks to 9 months of age had better language development (CCDI-P composite) scores at 9 months. This finding supports the importance of low-level auditory processing capabilities for normal language acquisition in young infants (Benasich & Tallal, 2002; Choudhury & Benasich, 2011). Deficits in low-level auditory processing of stimuli that present temporal challenges to the nervous system have been thought to be an underlying mechanism in language-based learning problems (Boets, Vandermoten, Poelmans, Luts, Wouters & Ghesquiere, 2011; Hornickel, Chandrasekaran, Zecker & Kraus, 2011; King et al., 2002; Wible et al., 2004; Leppänen et al., 2010), but no previous studies have examined normal variations in ABR latencies and language development in a sample of typically developing infants. The pattern of longer Wave V latencies or relatively slower ABR to the click stimuli presented here are a marker of developmental variation in the auditory processing capabilities of an individual. Importantly, our results show that these patterns, which have been shown to be related to language and reading impairments in older children, are also related to slower early language development in a typically developing population.

On the surface, our result appears to present a different picture from that found in a study which compared ABR latencies in typically developing children to those with language-based learning difficulties. In that study, ABR responses to click stimuli did not differ between groups, whereas ABR responses to more complex speech syllables did show differences (King et al., 2002). However, we believe that the results from these studies converge despite the superficial appearance of differences. In fact, we believe that our forward masking paradigm is more complex than simple click stimuli and this alone could account for the apparent differences in findings. In addition, there are numerous differences between the King et al. (2002) study and the present study with regard to stimulus presentation rate (31.1/sec vs. 11.7/sec) and intensity of click stimuli (68 dB vs. 80 dB), as well as with age (8–12 years vs. 6-week- and 9-month-olds), sample size (N = 33 vs. 71), and populations examined (group comparisons of children with known language delays compared to children with no language delays vs. individual differences within an assumed typically developing population) that warrant further exploration. Many of these differences across paradigms could affect the results – for instance, one might expect differences in the sensitivity of the ABR as a measure in populations for whom the ABR is actively maturing (infants) vs. those with a more stable ABR response (2-year-olds, older children, and adults). Moreover, as indicated by the very earliest studies comparing infant and adult responses, suprathreshold measures which present challenges to the nervous system’s ability to recover from a preceding response may be more sensitive to measures than those that are closer to hearing thresholds (Lasky, 1991). Indeed, a more recent paper also found differences in both the latencies and amplitude of ABR responses to rapidly presented clicks for SLI vs. typically developing children, with greater differences between the groups shown for the most challenging and rapidly presented stimuli (Basu et al., 2010). Taken together, we believe the previous studies and our own suggest that the relationships between subcortical auditory processes and language are most evident when the nervous system is challenged, whether it is through the presentation of speech syllables in isolation or click stimuli that tax the auditory system’s capacity to respond (i.e. with a rapid overall presentation rate or with short intervals between pairs of stimuli). This makes sense, given the rapid frequency changes and
temporal cues present in speech, although the fact that these differences are evident at subcortical, pre-attentive levels of processing just 6 to 10 msecs after the onset of a stimulus is still quite astonishing.

Further research with both infants and older children is clearly needed to understand the roles of stimulus type, stimulus presentation rate, and child’s age in predicting later language or reading development, and the relationship between subcortical and cortical levels of auditory and speech processing more generally. Nonetheless, the potential for drawing these links seems close at hand. Consistent with the results of the present study, children who performed better on language tasks at 3 years of age appeared to show better lateralized discrimination of both non-speech and speech sounds in an auditory evoked potential study conducted when the children were newborns and again at 1 year of age (Molfese, 1989). Differences in the N1 (at 174 ms) and N2 (at 458 ms post-stimulus) peak latencies for speech stimuli in both the left and the right hemispheres, measured in newborns, reliably discriminated between children who became ‘typical’, ‘poor’, and ‘dyslexic’ readers at age 8 (Molfese, 2000), as well as predicted performance on a number of language-related tasks in the preschool and early school years (Espy et al., 2004). In addition, ERPs in infants with a family risk for dyslexia, regardless of their later learning outcomes, were significantly different from infants with no family history and who later had typical reading development (Leppänen et al., 2010). These studies provide converging evidence that early differences in even subcortical processing of auditory stimuli by children with normal hearing and development in other domains can be potentially predictive of later phonological processing and reading skills.

Taken together, our results and those of others suggest that the ability to perceive, process, and discriminate both speech and non-speech stimuli contributes both to the rate of language development in a typically developing population and to whether a child may be at risk for language or reading delays. The more efficiently an individual is able to process both non-speech and speech sounds in the environment, the more likely the individual can use such information to distinguish sound differences critical for attaining normal language development and literacy skills years later (Boets et al., 2011; Hornickel et al., 2011; Leppänen et al., 2010; Molfese, 1989).

Differences in several low-level processes may explain why infants process auditory stimuli differently and in ways that predict language and reading difficulties (Boets et al., 2011; Choudhury & Benasich, 2011; Hornickel et al., 2011; Jacobson, 1985; Johnson et al., 2008; Leppänen et al., 2010; Moore & Linthicum, 2007; Wible et al., 2004). In particular, there may be individual differences in axonal myelination, dendritic arborization, conduction velocity along the dendritic arbors and/or axonal projections, and synaptic organization in brainstem pathways. When these differences occur during the peak of brainstem input to the auditory cortex (i.e. in the birth to 2-year period), they set in motion a larger set of differences in auditory processing of language-relevant stimuli, despite the fact that linguistic discriminations are on the time-scale of 10s of milliseconds for phonemes and hundreds of milliseconds for syllables, and not the 100ths of millisecond scales measured by the auditory brainstem response.

Specifically, we propose that these differences become relevant during the central nervous system slowing down of faster auditory signals from the periphery and its conversion into categorical, non-synchronized responses. It is this latter stage, of conversion from the preservation of extremely precise signals into perceptual categories in the cortex, that can be relevant to speech contrasts (Wang, 2007). Imprecise processing of rapid temporal signals can lead to imprecision in category boundaries and increased slippage and time for processing of speech contrasts and integration with other sensory modalities necessary for language and reading development (Wang, 2007). We therefore hypothesize that infants with less mature neuronal morphology and functioning in the auditory system – in particular, incomplete myelination and reduced synaptic organization – will have limited abilities to sample the auditory stimulus stream appropriately, resulting in incomplete and/or distorted auditory information. This, in turn, can lead to increased times to appropriately categorize meaningful speech distinctions (Werker & Tees, 1984; Rivera-Gaxiola et al., 2005) and subsequently slower language acquisition.

Consistent with this overall picture are the gender differences that are typically found on language tasks and which also appeared in both our ABR and CDI data. On average, girls tend to have slightly more accelerated language development than boys, but gender typically accounts for only 1–2% of the variance and results in, at best, a month or two advantage at 2 to 3 years of age (Fenson, Dale, Reznick, Bates, Thal, and Pethick, 1994; Huttenlocher, Haight, Bryk, Seltzer & Lyons, 1991; Tardif et al., 2008; Tardif, Fletcher, Liang & Kaciroti, 2009). In our study, gender was moderately correlated with Wave V latencies at both 6 weeks and 9 months, and was slightly correlated with the CCDI-P composite (Table 4), but this correlation dropped out once the ABR latencies were included (Figure 2). Thus, it appears that a significant portion of the association between gender and language development, at least at these very earliest stages of language development, may
be explained by gender differences in low-level rapid auditory processing capabilities.

Although there are many limitations in the present study, including a relatively small sample size and the non-direct and very early parent report assessment of language development at 9 months of age, it is important to note that we found significant results between very low-level auditory processing and these measures in a typically developing population of infants. Follow-up studies which include infants at risk for SLI or later reading difficulties will also be informative if this method is used to identify infants for risk of SLI or reading delay even before language delay is recognized.

Although it is generally accepted that subcortical processing of auditory information is general to all sounds, including speech sounds from multiple languages as well as non-speech sounds, there is some evidence to suggest that there may be real differences in subcortical stages of speech processing (i.e. at 5–7 msecs post-stimulus, exactly the point at which Wave V ABR responses occur) which may be language-dependent (Krishnan & Gandour, 2009). Moreover, Kuhl et al. (2008) also showed that both native and non-native ERPs predicted future language development, but that better native discrimination predicted faster language growth whereas better non-native discrimination predicted slower language growth. Thus, it will be important to follow up this result with languages other than Chinese, and to understand whether rapid auditory processing these very early stages (< 10 msecs) of the auditory signal is as important for non-tonal languages as it is for tonal languages. Finally, long-term developmental follow-up of both the ABR variables and language outcomes, including direct testing at older ages, are also warranted to further examine the connections between low-level rapid auditory processing capabilities and language acquisition, including the issue of whether this measure can be used to predict language difficulties specifically, or whether it is simply a predictor of global processing capacities that are important for language as well as a host of other higher order cognitive abilities.

Despite these limitations, we believe that the ABR protocol used in the present study (forward masking) shows promise as an electrophysiological marker of rapid auditory processing that may be used to predict individual differences in language development. Such an approach may expand the utility of the ABR beyond its current use as both a hearing screening tool and a neurodiagnostic measure commonly used in clinical practice for assessing overall neural integrity. The ABR is not expensive and entails relatively simple diagnostic procedures with extremely precise and reliable latencies. Therefore, it may be feasible to use it to predict long-term language development in typically developing children. Use of the ABR with stimuli that provide temporal processing challenges to the nervous system (e.g. syllables, fast presentation rates for stimulus trains, or forward masking with short inter-stimulus intervals) might also be used to detect individuals who are at risk of developing SLI or language-based learning problems so that early interventions and treatment can be initiated. Finally, a strength of the approach is that low-level auditory processing in subcortical pathways can be measured directly. Recently, unrelated studies by Hornickel et al. (2011) and Boets et al. (2011) demonstrated that neural indices of auditory brainstem function necessary for auditory processing were significantly associated with reading ability and speech-in-noise perception in kindergarten and school-aged children even after letter knowledge, phonological awareness, and working memory were taken into consideration. As such, the variability in temporal processing abilities documented by the ABR measures presented here may provide potentially useful measures for predicting later language and reading skills in typically developing populations beginning in infancy. In addition, the current data are relevant to the issue of gender differences in early language development. Although gender differences have frequently been reported for early vocabulary and many other language measures (e.g. Bornstein, Hahn & Haynes, 2004; Huttenlocher et al., 1991; Lovas, 2011), there has been much less of an emphasis on the structural and physiological aspects to early brain development that might give rise to these differences, as opposed to various social and personality factors. Nonetheless, the current findings of gender differences for an indirect measure of myelination (the subcortical ABR response) in the first year of life joins with other recent reports of gender differences in brain structure and growth in the postnatal period that have examined the role of sex hormones and testosterone on the organization of language in the brain (Friederici, Pannekamp, Partsch, Ulmen, Oehler, Schmitzler & Hesse, 2008), as well as earlier work with other electrophysiological measures and EEG coherence (e.g. Hanlon, Thatcher & Cline, 1999; Shucard, Shucard & Thomas, 1987). Echoing the more suggestive findings in these earlier studies, our data show that relationships between gender and ABR latencies, as well as between ABR latencies and early language outcomes, completely mediate the gender effect on language in our study. This result is clearly provocative and worthy of future follow-up and replication.

In sum, it is not just a single point in developmental time that appears to be important. Rather, the
developmental trajectory of auditory processing capabilities, both at cortical and at subcortical levels, needs to be examined carefully. Considering change in the ABR at two time points, and particularly changes in which infants started out processing auditory information more rapidly than their peers, but showed a decrease in their relative standing over time, better predicted language acquisition in young healthy infants in this study than either time point by itself. As such, the current study clearly contributes to our understanding of how differences in relatively low-level processes in the brain can cascade into a number of higher order differences during development that may not be detectable until several years later.

Acknowledgements

We deeply thank all participants and families who participated in this study, as well as the research team from the Children’s Hospital, Zhejiang University and Fuyang Maternal and Child Health Hospital in China for their assistance with subject enrollment and data collection. This work was supported by P01 HD39386 from the National Institute of Child Health and Human Development, B. Lozoff, Principal Investigator and also the National Natural Science Foundation of China (#30671773), J. Shao, Principal Investigator. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Child Health and Human Development or the National Institutes of Health.

References


Received: 24 June 2011
Accepted: 25 August 2012