

# Neural patterns to speech and vocabulary growth in American infants

Maritza Rivera-Gaxiola,<sup>CA</sup> Lindsay Klarman, Adrian Garcia-Sierra and Patricia K. Kuhl

Institute for Learning and Brain Sciences University of Washington, Seattle, WA 98195 USA

<sup>CA</sup>Corresponding Author: rivegaxi@u.washington.edu

Received 29 December 2004; accepted 28 January 2005

We report infant auditory event-related potentials to native and foreign contrasts. Foreign contrasts are discriminated at 11 months of age, showing significant differences between the standard and deviant over the positive (P150–250), or over the negative (N250–550) part of the waveform. The amplitudes of these deflections have different amplitude scalp distributions. Infants

were followed up longitudinally at 18, 22, 25, 27 and 30 months for word production. The infant speech discriminatory P150–250 and N250–550 are different components with different implications for later language development. *NeuroReport* 16:495–498  
© 2005 Lippincott Williams & Wilkins.

**Key words:** Event-related potentials; Native language neural commitment model; Neural patterns; Speech perception; Word production

## INTRODUCTION

Young infants discriminate phonetic contrasts from different languages of the world regardless of prior exposure. Between 6 and 12 months of age, they show a decline in the ability to perceive foreign-language contrasts [1,2]. It was argued that the behavior observed at 10–12 months of age and adulthood indicated an immutable loss of an initial universal language ability, though later work showed that adults still remain sensitive to foreign-language contrasts [3–5]. Tsao *et al.* [6] have shown in behavioral studies that infants' speech perception skills measured at 6 months predicts later language abilities. Moreover, Kuhl [7,8] has argued that sensitivity to native contrasts increases between 6 and 12 months, and that in the second half of the first year of life, infants neurally code the basic acoustic patterns of their native language, proposing the native language neural commitment (NLNC) model.

Recording event-related potentials (ERPs) allows one to study neural responses with precise time resolution, and has been used for a long time [9]. Various ERP components related to speech discrimination and language processing have been investigated [10–16]. Näätänen *et al.* [12] reported a mismatch negativity (MMN) in response to native vowel contrasts, but not to foreign ones in Finnish adults. The MMN is elicited by random, occasional changes in unattended sounds [13]. Rivera-Gaxiola and colleagues [14] reported an MMN and a late positive deflection during a passive task when English-speaking adults listened to a native contrast. When the same participants were presented with a difficult, foreign (Hindi) contrast, they only displayed an MMN. Rivera-Gaxiola and colleagues suggested that there was no permanent loss of the initial perceptual abilities that humans have as infants. Cheour *et al.* [15] showed an MMN to native and foreign vowel contrasts at 6 months of age. At 12 months, an MMN effect was observed

only in response to infants' native vowel contrast. However, language-specificity has been evidenced by 6 months of age for tasks that tap infants' perceptual organization of vowel speech categories [17] and performance on foreign vowel contrasts remains high throughout life [18].

The manner in which language experience heightens relevant phonetic contrasts and de-emphasizes foreign contrasts remains elusive. A recent ERP developmental study by Rivera-Gaxiola *et al.* [16] showed discriminatory capacity to native and foreign syllabic contrasts at 7 months, and differences to the native contrast only at 11 months of age when participants were pooled in a single group. However, they *also* reported two subgroups of infants: those who responded over the P150–250 time window of the ERP complex ('P' responders) and those who responded over the N250–550 time window ('N' responders) (P: positive; N: negative) to the foreign deviant stimulus, showing that at 11 months of age, infants remain capable of discriminating the foreign contrast at a neural level. Interindividual variability in ERP responses to the detection of changes in speech have been documented in the literature [15,16,19–23], however, the explanation as to why some groups show differences in positive components and others in negative components remains controversial.

The purpose of the present study was to explore the nature of the 'P- and the 'N' responders. We examined the implications of these neural measures on word production at 18, 22, 25, 27 and 30 months of age and also analyzed the scalp distribution of the amplitudes of each type of response. We hypothesized that (1) 11-month-old-infant ERPs will show discrimination of foreign speech sounds, and (2) if the P150–250 and the N250–550 are different components electrophysiologically and behaviorally, they will have a different impact on later language scores.

## MATERIALS AND METHODS

**Participants:** Fifty monolingual English infants (25 girls) aged 11 months (mean age=10.7 months) were recruited through the Infant Studies Subject Pool at the University of Washington. Infants were full term ( $\pm 14$  days from due date) and their mothers had normal pregnancies and deliveries. Parents signed approved consent forms, were informed of the procedures and were paid \$15 for ERP recording and \$10 each time for the Communicative Development Inventories at different ages. ERPs from 28 participants (15 girls/13 boys) were considered suitable for processing and analysis.

**Stimuli:** The three consonant-vowel syllables used by Rivera-Gaxiola *et al.* [16] were employed in this study (Table 1). Stimuli were naturally produced by a female Spanish/English bilingual speaker and manipulated using Praat and SoundForge 4.0 to obtain a match in duration ( $229.65 \pm 0.3$  ms), intensity and average RMS power. The average fundamental frequency was 180 Hz. Pilot studies confirmed that native English speakers behaviorally discriminated the English but not the Spanish contrast (Table 1).

**Design:** A double-oddball paradigm was used. The phoneme common to both languages was used as the standard, and the two language-specific sounds served as deviants in an 80/10/10 presentation proportion (Table 1). Deviants appeared in a semirandom fashion with at least three standards between deviants. The interstimulus interval was 700 ms and a total of 1000 stimuli were presented. A 1 min silent period was inserted every 2 min of stimulation to allow interaction with the baby. Stimuli were delivered to two loudspeakers placed 1 m in front of the child at 69 dB SPL.

**Procedure:** Infants were awake and were tested inside a sound attenuated booth. The parent sat next to the child. In front of them, a research assistant entertained the child while a muted movie played on a TV. The research assistant and the parent wore headphones with masking music while the stimuli were played.

**Electroencephalogram recording:** The electroencephalogram (EEG) was recorded using electrocaps with preinserted tin Ag/AgCl electrodes referenced to the left mastoid from Fp1, Fp2, F3, F4, C3, C4, P3, P4 O1, O2, F7, F8, T3, T4, T5, T6, Fz, Cz, Pz, and the active right mastoid of the 10/20 International System. The vertical electrooculogram was recorded from 1 infraorbital electrode placed on the infant's left cheek. The amplifier bandwidth was set between 0.1 and 40 Hz. Electrode impedances were kept below 5 k $\Omega$ . Signals were amplified with a gain of 20 000. EEG was sampled every 4 ms. Segments with electrical activity  $+/-150 \mu$ V at any electrode site were rejected. EEG segments of 700 ms with a prestimulus baseline time of 100 ms were selected and averaged offline to obtain the ERPs. Baseline correction

was performed in relationship to the prestimulus time. Further low-pass filtering was set at 15 Hz.

**Data analyses:** ERP data were accepted when clear auditory P-N complexes within the first 600 ms were displayed (60–80 artifact-free trials for each standard and deviant types required). ERPs were explored within two time windows: The first positive deflection between 150–250 ms after stimulus onset (P150–250) and the following negative deflection 250–550 ms after stimulus onset (N250–550). The whole group was analyzed first and then subdivided into 'P' and 'N' responders according to the time window and polarity in which the largest amplitude differences between the standard and each deviant type were observed [16]. Peak amplitude values for standards preceding a deviant, and deviant stimuli within each time window for each child and condition were identified and used to calculate three independent three-way repeated measures ANOVAs. The factors included were: condition (native and foreign contrasts), lateral electrode position (left and right) and anterior-posterior location (frontal-polar, frontal, central, parietal, frontal-lateral and temporal). Further topographical analyses using normalized peak amplitude values within participants for each component in each condition and electrode site were carried out [24].

**MacArthur Communicative Development Inventories [25]:** We followed-up each participant's vocabulary development at 18, 22, 25, 27 and 30 months of age and compared 'P' and 'N' responders at each age.

## RESULTS

**Electrophysiology:** Normally developing 11-month-old American infants learning English displayed distributed P-N complexes flattening over T5 and T6. The P150–250 and the following N250–550 as described above and in [16] are reported.

**Group data:** Significantly larger N250–550 amplitudes to the deviant with respect to the standard can be observed for the native contrast in all participants [ $F(1,27)=38.57$ ,  $p < 0.01$ ,  $\eta^2 p = 0.588$ , observed power=1.00]. No significant differences were observed in either ERP time-window for the foreign contrast in the group data analyses [ $F(1,27)=2.69$ ,  $p = 0.113$ ] (Fig. 1).

**Subgroup data:** Two groups were formed, both of which showed larger N250–550 to the native deviant with respect to the standard [ $F(1,16)=33.15$ ,  $p < 0.001$ ,  $\eta^2 p = 0.547$ , observed power=0.985;  $F(1,10)=19.71$ ,  $p = 0.001$ ,  $\eta^2 p = 0.663$ , observed power=0.978]. The groups basically differed in how infants responded to the foreign contrast. The first group, named the 'P' responders, included 13 infants who displayed significantly larger (more positive) P150–250 amplitudes to the foreign deviant with respect to the

**Table 1.** Stimuli used and main physical attributes.

Syllable used	Voicing	VOT	Aspiration	Phonemic use	Event-related potential used
Spanish /da/	Prevoiced	–24 ms	Unaspirated	Spanish	Foreign deviant
Sp /ta/-Eng /da/	Voiceless	12 ms	Unaspirated	Spanish and English	Standard
English [t <sup>h</sup> a]	Voiced	46 ms	Aspirated	English only	Native deviant

standard [F(1,12)=26.41,  $p < 0.001$ ,  $\eta^2 p = 0.706$ , observed power=0.966] (Fig. 2).

The second group, named the 'N' responders included 11 participants, all of whom displayed larger (more negative) N250–550 amplitudes to the foreign deviant with respect to the standard [F(1,10)=16.44,  $p = 0.001$ ,  $\eta^2 p = 0.558$ ] (Fig. 2).

**Event-related potentials to the standard syllable:** 'P' and 'N' responders did not differ in either P150–250 or N250–550 ERP amplitudes to the standard stimulus ( $p < 0.05$ ).

**Topographic analyses:** The topographical distribution across the scalp of the P150–250 and the N250–550 normalized peak amplitudes within participants for each condition showed important component  $\times$  electrode site interactions [native F(1.24,3.023)=40.75,  $p < 0.001$ ; foreign F(1.005,2.72)=28.53,  $p < 0.001$ ]. In both conditions, P150–250 amplitudes show a frontocentral distribution, while the N250–550 largest amplitudes are parietal.

**Word production:** We calculated an independent samples' *t*-test to differences between the number of words produced

by the 'P' and by the 'N' responders. All infants produced a well within normal range of words at each age [25]. However, 'P' responders produced more words at each age than 'N' peers at the same ages (Table 2). 'P' responders produced a mean of  $466.5 \pm 39.12$  (SE) and 'N' responders produced a mean of  $343.25 \pm 34.91$  (SE) words across all ages. The difference was significant ( $p = 0.015$ ,  $t = -2.257$ ). Both distribution curves show a similar kurtosis ( $-0.107$  and  $-0.103$ , respectively). All infants had a vocabulary spurt around 18–25 months, but the size of this spurt was consistently larger for 'P' responders than for 'N' responders at all ages.

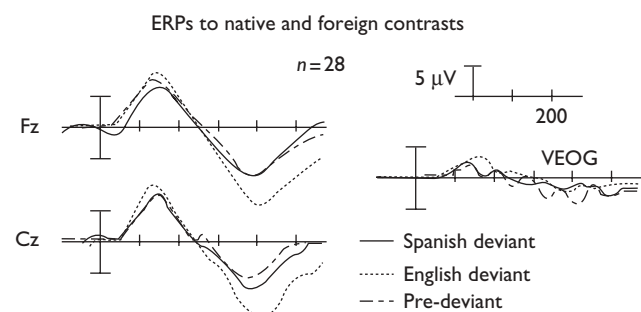
**DISCUSSION**

We evaluate the implications of the different neural patterns seen in 11-month-old American infants when they discriminate native and foreign contrasts on their communicative abilities at 18–30 months of age and characterize further the 'P' and 'N' responders described by Rivera-Gaxiola and colleagues in their developmental study [16]. In the present study, we (1) confirmed that 11-month-old infants remain capable of foreign contrast discrimination at a neural level, (2) found that the amplitudes of the P150–250 and the N250–550 have different scalp distributions and, very importantly, (3) found that responding to the foreign contrast at the P150–250 as opposed to the N250–550 level at 11 months of age, results in higher scores for word production at 18, 22, 25, 27 and 30 months of age.

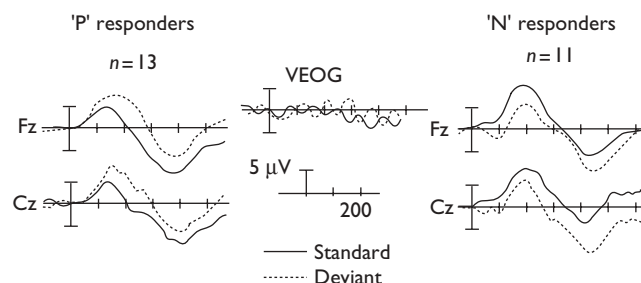
In the present study, all infants showed the largest N250–550 amplitudes to the native deviant. When infants were sorted into subgroups on the basis of whether the amplitude differences between the standard and deviant were in the P150–250 or in the N250–550 time window, we found evidence of discrimination for the foreign contrast: some infants displayed more positive P150–250s while other infants displayed more negative N250–550s to the foreign deviant with respect to the standard.

We also hypothesized that if the infant P150–250 and N250–550 were electrophysiologically and behaviorally different components, they would differ in (1) scalp distribution and (2) impact in later language development. By comparing the within-participant normalized amplitudes of each deflection across the scalp [24], we provided evidence that these two responses are electrophysiologically different components. By comparing the infants' later language scores, we confirmed that each ERP response has a different behavioral impact in development.

Our previous ERP results with 7 and 11-month-old infants [16] show, that at 11 months infants display stronger and consistent N250–550s to the native deviant than at 7 months. In the present study, all infants showed a more negative N250–550 to the native deviant with respect to the standard. The behavioral literature shows that at 10–12 months infants show a strong phonetic response to the native contrast; the ERP literature shows that a larger N2a or an MMN to the



**Fig. 1.** Grand average event-related potentials (ERPs) to the standard syllable Spanish /ta/-English /da/ and the deviants aspirated English [t<sup>h</sup>a] (native contrast) and Spanish voiced /da/ (foreign contrast). When infants are pooled together, they show neural discrimination to the native contrast only. Positive is plotted up.



**Fig. 2.** Event-related potential patterns to the foreign contrast: the 'P' responders displayed a more positive P150–250 peak and the 'N' responders showed a more negative N250–550 peak to the deviant than to the standard.

**Table 2.** Number of words produced by each group at each age studied.

	18 m (n=13/7 girls)	22 m (n=23/13 girls)	25 m (n=26/14 girls)	27 m (n=24/13 girls)	30 m (n=22/13 girls)
P responders	90	286	461	495	610
N responders	63	214	350	428	529

n=Number of reports returned.

deviant with respect to the standard is the response expected [11–15,19,21–23]. We therefore believe that the time window of phonetic response was reflected by the N250–550 in our study. Acoustic processing, on the other hand, taking place early after the onset of stimulus differences, could be reflected by the P150–250 amplitude differences. However, only studies that tap both ERPs and behavior concomitantly will inform us which response triggers behavioral attention in infants.

Very importantly, we provide here, in a prospective study, the impact of each type of neural response in language development: when exploring the productive vocabulary of each group of infants ('P' or 'N' responders) from 18 to 30 months of age, we found that 'P' responders produced more words at all ages than 'N' responders. Behavioral measures of infants at 7.5 months also show that better native perception predicts advanced language skills whereas better foreign perception predicts slower language growth [8]. According to Kuhl [7,8], an infant's successful language learning requires neural commitment to native-language speech patterns, while a strong sensitivity to the foreign contrast at 11 months of age will reflect poorer native language scores at later ages. In this way, if responding over the P150–250 or the N250–550 time window yields different linguistic scores months later, then these ERP responses are a reflection of differential commitment to the infant's native language. In this study, the 'P' responders are more committed to their native language than the 'N' responders.

In our view, 'N' responders, who show larger N250–550 amplitudes to both native and foreign deviants, are not attending differentially to native language acoustic patterns. They may be using the same type of resources to process differences, regardless of phonemic status. The 'P' responders are responding differentially: they show larger N250–550 amplitudes to the native deviant, but larger P150–250 amplitudes to the foreign deviant. Our analyses showed that the impact of such differential processing is an advantage on the number of words produced at later stages. Perhaps the P150–250 is a response that does not trigger further processing or attention and this is a benefit, given that the foreign contrast is not relevant to monolingual language learning. 'N' responders may be at an advantage for second or bilingual language learning.

## CONCLUSIONS

ERPs reflect important individual differences in linguistic processing. The infant auditory P150–250 and N250–550 differ in polarity, latency, behavioral implications and the distribution of their amplitudes across the scalp. Classifying participants as 'P' and 'N' responders is a useful tool to evaluate later language outcomes: Allocating similar resources to what is and is not phonemic negatively impacts later vocabulary scores in monolingual first language acquisition.

## REFERENCES

- Eimas PD, Siqueland ER, Jusczyk P, Vigorito J. Speech perception in infants. *Science* 1971; **171**:303–330.
- Werker JF, Tees RC. Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behav Dev* 1984; **7**:49–63.
- Werker JF, Tees RC. Phonemic and phonetic factors in adult cross-language speech perception. *J Acoust Soc Am* 1984; **75**: 1866–1878.
- Werker JF, Logan JS. Cross-language evidence for three factors in speech perception. *Percept Psychophys* 1985; **37**:35–44.
- Carney AE, Widin GP, Viemeister NF. Noncategorical perception of stop consonants differing in VOT. *J Acoust Soc Am* 1977; **62**:961–970.
- Tsao F, Liu HM, Kuhl PK. Speech perception in infancy predicts language development in the second year of life: a longitudinal study. *Child Dev* 2004; **75**:1067–1084.
- Kuhl PK. Early language acquisition: cracking the speech code. *Nat Rev Neurosci* 2004; **5**:831–843.
- Kuhl PK, Conboy B, Padden D, Nelson T, Pruitt J. Early speech perception and later language development: implications for the 'critical period'. *Lang Learning Dev* (in press).
- Hillyard S, Kutas M. Electrophysiology of cognitive processing. *Annu Rev Psychol* 1983; **34**:33–61.
- Molfese D. The phoneme and the engram: electrophysiological evidence for the acoustic invariant in stop consonants. *Brain Lang* 1980; **9**: 370–376.
- Maiste AC, Wiens A, Hunt MJ, Scherg M, Picton TW. Event-related potentials and the categorical perception of speech sounds. *Ear Hear* 1995; **16**:68–90.
- Näätänen R, Lehtokoski A, Lennes M, Cheour M, Huotilainen M, Iivonen A et al. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 1997; **385**:432–434.
- Näätänen R, Gaillard AWK, Mäntysalo S. Early selective attention effects on the evoked potential: a critical review and reinterpretation. *Biol Psychol* 1978; **8**:81–136.
- Rivera-Gaxiola M, Csibra G, Johnson MH, Karmiloff-Smith A. Electrophysiological correlates of cross-linguistic speech perception in native English speakers. *Behav Brain Res* 2000; **111**:13–23.
- Cheour M, Ceponiene R, Lehtokoski A, Luuk A, Allik J, Alho K et al. Development of language-specific phoneme representations in the infant brain. *Nat Neurosci* 1998; **1**:351–353.
- Rivera-Gaxiola M, Silva-Pereyra J, Kuhl PK. Brain potentials to native and nonnative contrast in 7 and 11 month old American infants. *Dev Sci* 2005; **8**:162–172.
- Kuhl PK, Williams KA, Lacerda F, Stevens KN, Lindblom B. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 1992; **255**:606–608.
- Stevens KN. The quantal nature of speech: evidence from articulatory-acoustic data. In: Denes PB, David EE Jr (eds). *Human Communication, A Unified View*. New York: McGraw-Hill; 1972. pp. 51–66.
- Cheour M, Alho K, Ceponiene R, Reinikainen K, Sainio K, Pohjavouri M et al. Maturation of mismatch negativity in infants. *Int J Psychophysiol* 1998; **29**:217–226.
- Dehaene-Lambertz G, Dehaene S. Speed and cerebral correlates of syllable discrimination in infants. *Nature* 1994; **370**:292–295.
- Pang EW, Edmonds GE, Desjardin R, Khan SC, Trainor LJ, Taylor MJ. Mismatch negativity to speech stimuli in 8-month-old infants and adults. *Int J Psychophysiol* 1998; **29**:227–236.
- Leppänen PHY, Pihko E, Eklund KM, Lyytinen H. Cortical responses of infants with and without a genetic risk for dyslexia: II group effects. *Neuroreport* 1999; **10**:969–973.
- Pihko E, Leppänen PH, Eklund KM, Cheour M, Guttorm TK, Lyytinen H. Cortical responses of infants with and without a genetic risk for dyslexia: I. age effects. *Neuroreport* 1999; **10**:901–905.
- McCarthy G, Wood CC. Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models. *Electroenceph Clin Neurophysiol* 1985; **62**:203–208.
- Fenson L, Dale P, Reznick JS, Thal D, Bates E, Hartung J et al. *MacArthur Communicative Development Inventories: User's Guide and Technical Manual*. San Diego, California: Singular Publishing Group; 1993.

Acknowledgements: This work is supported by grants from NIH (HD 37954), the UW's ILABS, and the Talaris Research Institute and the Apex Foundation (Bruce and Jolene McCaw). We acknowledge T. Imada, K. Schoolcraft, D. Paden, R. Cabanis, Juan Silva, B. Conboy and J. Sommerville.