



Published in final edited form as:

*Clin Neurophysiol.* 2008 December ; 119(12): 2855–2861. doi:10.1016/j.clinph.2008.09.015.

## The representation of voice onset time in the cortical auditory evoked potentials of young children

Katrina King<sup>1</sup>, Julia Campbell<sup>2</sup>, Anu Sharma<sup>2</sup>, Kathryn Martin<sup>1</sup>, Michael Dorman<sup>3</sup>, and Justin Langran<sup>1</sup>

<sup>1</sup>Callier Advanced Hearing Research Center, Department of Brain and Behavioral Sciences, University of Texas at Dallas, Texas, USA

<sup>2</sup>Brain and Behavior Laboratory, Department of Speech Language and Hearing Sciences, University of Colorado at Boulder, USA

<sup>3</sup>Department of Speech and Hearing Science, Arizona State University, Tempe, Arizona, USA

### Abstract

**Objective**—The aim of this study was to determine whether, in young children, a cortical neurophysiologic representation of the voicing status of a stop consonant could be found in the absence of the N1 component in the cortical auditory evoked potential (CAEP). In adults, a ‘double-on’ N1 response is often considered a cortical marker of VOT representation.

**Methods**—Scalp-recorded CAEPs were measured from 6 electrode sites in 10 children aged 4–8 years in response to a /da/ - /ta/ continuum in which voice onset times (VOTs) varied from 0 – 60 ms. CAEPs were also recorded from C<sub>z</sub> in a group of 8 children aged 2–3 years in response to stimuli with VOTs of 0 and 60 ms.

**Results**—Cortical evoked responses elicited by stimuli with VOT values ranging from 0 to 60 ms (i.e., by stimuli perceived as /da/ and /ta/) were similar in morphology. There was no evidence of a ‘double-on’ morphology for stimuli with long VOTs. However, latency changes in the P1 and N2 components were observed as a function of VOT changes.

**Conclusion**—Our results demonstrate that a representation of VOT, as recorded by scalp electrodes, exists in the developing cortical evoked response, but that representation is different than that in the adult response. The results describe the developmental changes in cortical representation of VOT in children ages 2–8 years.

**Significance**—The child’s CAEP reflects physiologic processes, which are involved in the cortical encoding of VOT. Overall, cortical representation of VOT in children ages 2–8 is different than in adults.

### Keywords

Voice onset time (VOT); Cortical auditory evoked potentials (CAEP); Maturation; Development; Refractory period; Children; P1; N2; N1; Double-on

## A. Introduction

In this study we examined the neurophysiologic representation of voice onset time in the cortical auditory evoked response (CAEP) of young children. Voice onset time (VOT) is an acoustic cue that is critical for speech perception. VOT is defined as the interval between the release of a stop consonant (the burst) and the beginning of vocal fold vibration or voicing onset (Lisker and Abramson, 1964). Several languages contain relevant temporal cues for VOT perception. In American English, voiced, syllable-initial stop consonants (such as /ba/, /da/, /ga/) have a short time interval (generally 0 to 20 ms) between the release of the consonant and the onset of voicing. In contrast, voiceless stop consonants (such as /pa/, /ta/, and /ka/) have a longer interval (generally 30 – 80 ms) between consonant release and voicing onset.

Listeners respond to stimuli that vary in VOT in a manner termed ‘categorical’. In the paradigm experiment, Abramson and Lisker (1970) created a set of synthetic speech sounds that varied in VOT from –150 ms to +150 ms in physically equal steps. When the stimuli were presented to native speakers of American English for identification, stimuli with VOTs of less than 25 ms were identified as /da/ and stimuli with VOTs greater than 25ms were identified as /ta/. The identification function showed a sharp discontinuity between about 20 and 30 ms of VOT – the phoneme boundary between the /da/ and /ta/ categories. Listeners were good at discriminating stimuli drawn from opposite sides of the phoneme boundary but were poor at discriminating stimuli drawn from within the same voicing category. The good discrimination of stimuli drawn from different phonetic categories and the poor discrimination of stimuli drawn from the same phonetic category led to the term ‘categorical’ perception’ (see Liberman et al., 1957).

In a series of elegant experiments, Steinschneider et al. (1982, 1990, 1994, 1995b, 2003b, 2005) have explored the cortical representation of VOT in monkeys and humans. Recording auditory evoked potentials intracranially from Heschel’s Gyrus in humans (1999), and from the primary auditory cortex in monkeys (2005), Steinschneider et al. reported that the cortical response to stimuli with long VOTs, which were identified by human listeners as voiceless stops, was characterized by an initial response peak followed by a second peak (‘double-on’ response) which was time-locked to the onset of voicing. For stimuli with a short VOT, which were identified as voiced stops, the cortical response was characterized by a single peak. These outcomes are consistent with other intracranial studies in humans (Liegeois-Chauvel et al., 1999), and document that the stimulus features critical for identifying syllable-initial voiced and voiceless stop consonants are represented by synchronized responses in cortical neural populations.

Given the outcome that intracortical recordings reflect differences in VOT, there has been great interest in determining if a similar cortical marker for VOT appears in scalp recordings from humans. Such a biological marker could have clinical significance as it would allow for a non-invasive and non-behavioral examination of one aspect of speech perception in at-risk infants and young children.

In adults, cortical auditory evoked potential (CAEP) recordings from scalp electrodes have demonstrated differences in morphology as a function of the VOT of a stimulus. Sharma and Dorman (1999; 2000) found a single negative peak ( $N1'$ ) was present for short VOT values of 0 to 30 ms. For longer VOT values (i.e., 50 to 80 ms), a double negative peak response ( $N1'$  and  $N1$ ) emerged. This outcome is consistent with results from Giraud et al. (2005) and Liegeois-Chauvel et al. (1999), who used pre-voiced stimuli, and from animal studies (Eggermont, 1995; Steinschnieder et al. 1994, 1995b, 2003b, 2005).

The relationship of the single *vs.* double-peak response of the N1 in adults to behavioral perception of signals as voiced or voiceless is not clearly understood. Some studies (e.g., Sharma et al., 2000) suggest that the single *vs.* double-peaked response is independent of how a stimulus is identified. Other studies infer a causal link between single *vs.* double peaks and the identification of stimuli as voiced *vs.* voiceless (e.g., Simos et al., 1998).

In adults, the N1 component is a dominant feature of the obligatory cortical auditory evoked response. The development of the N1 is age and stimulus rate dependent (Ceponiene et al., 1998; Gilley et al., 2005; Sussman et al., 2008) and typically emerges in an inconsistent fashion from 3–8 years of age. The presence of the N1 typically becomes more reliable around 9 to 11 years (Sharma et al., 1997; Gilley et al., 2005). In young children the dominant features of the cortical response are the P1, which varies in latency as a function of age and the N200 or N2 response (sometimes also called the N250), a negativity following the P1 at about 200 or 250 ms (Gilley et al., 2005; Sussman et al., 2008).

In the absence of a reliable N1 in children under age 9 years, it is possible that another component of the CAEP complex, such as the P1 or N2, functions as a cortical marker of VOT in young children. Another possibility is that the developing nervous system may have a completely different cortical representation of VOT, which is not reflected in the CAEP. To find out, we recorded CAEPs in response to stimuli differing along a /da/-/ta/ VOT continuum in children aged 2–8 years. Our goal was to examine whether the P1 and/or N2 components of the CAEP would reflect changes in VOT in a manner similar to the adult N1 single *vs.* double-peaked response.

## B. Methods

### 1. Subjects

Two groups of normal-hearing children who were native speakers of American English were tested. An older group consisted of ten children, 6 females and 4 males, ranging in age from 4.4 to 8.4 years (mean age = 6.2 years, standard deviation = 1.6 years). A younger group consisted of 8 children, 4 females and 4 males, ranging in age from 2.3 to 3.9 years (mean age = 3.2 years, standard deviation = 0.7 years). All children had normal hearing as evidenced by pure tone air and bone conduction thresholds less than 15 dB HL at octave frequencies from 250–8000 Hz, with no history of speech or hearing problems. The study had approval from the University of Texas at Dallas Institutional Review Board and subjects were consented according to UTD IRB procedures.

### 2. Stimuli

The stimuli were created with the Klatt (1980) speech synthesizer and consisted of VOT values of 0, 20, 40, and 60 ms. The tokens were identical to those used by Sinex et al. (1991) in animal studies and Sharma et al. (1999, 2000) in human studies. For each syllable, a brief 10 ms burst of friction noise was present at the onset. From syllable onset to the first formant onset (F1), the higher formants (F2 to F5) were excited by using aspiration noise. This interval corresponded to the VOT which was varied from 0 to 60 ms in 20 ms steps. The center frequency of F1 was 155 Hz for the first 10 ms of the syllable, increased to 270 Hz at 15 ms, and then increased linearly to the steady state value of 770 Hz at 65 ms. F2 and F3 decreased from 1550 to 1200 Hz and 3700 to 2200 Hz over a transition duration of 65 ms. Each syllable had a fundamental frequency of 114 Hz and overall duration of 200 ms.

### 3. Procedure

**Identification Experiment**—A forced-choice identification experiment was performed before evoked potential testing with the children in the 4–8 year age group. Children in the

younger age group were not able to perform the task accurately. Each subject was asked to listen to stimuli with VOT values ranging from 0 to 60 ms and decide whether the syllable was a /da/ or a /ta/ sound. In order to achieve a balanced design, five subjects were asked to indicate their response by raising one finger if the syllable was classified as /da/ and two fingers if the syllable was /ta/. Similarly, the remaining five subjects were asked to raise two fingers if the syllable was classified as a /da/ and one finger if the syllable was classified as a /ta/. Ten repetitions of each stimulus in the continuum were presented to the subject in random order. Stimuli were presented via a loudspeaker at 75 dB SPL (sound pressure level) at a distance of 1 meter. The loudspeaker was placed at 45 degrees azimuth to the right of 5 of the subjects and at 45 degrees azimuth to the left of 5 of the subjects.

**Electrophysiological Experiment**—CAEPs were recorded while participants were seated comfortably in a reclining chair in a sound-treated booth. To control for arousal state during testing, participants watched a movie on a screen located at 0 degrees azimuth. Background audio levels for the movie were turned off during testing. A NeuroScan STIM and SCAN (version 4.3) evoked potential system was used for stimulus generation and CAEP recording. Silver/silver-chloride electrodes were placed at  $C_z$ ,  $F_z$ ,  $C_3$ ,  $C_4$ ,  $T_3$ , and  $T_4$  for the ten older children (4–8 years). A ground electrode was placed at  $FP_z$  and a reference electrode was placed on the nose for the 10 older children (4–8 years). For the younger children (2–3 years), only an active  $C_z$  electrode was tolerated. A reference electrode was placed on the mastoid ipsilateral to the speaker for the younger children who would not tolerate a nose electrode. A ground electrode was placed at  $FP_z$ . Eye blinks were monitored via electrodes placed at right lateral canthus and right supra-orbital sites. Electrode impedances were maintained below 5 kOhms. A recording window of 600 ms was used, which included a 100 ms pre-stimulus period. The onset-to-onset inter-stimulus interval was 800 ms.

Stimuli were presented via a loudspeaker at 75 dB SPL at a distance of 1 meter from the subjects and at a 45 degree azimuth. For the older children, evoked responses were recorded in response to VOT stimuli at 0, 20, 40, and 60 ms. For the younger children, evoked responses were recorded to stimuli with 0 and 60 ms VOT values only. These values corresponded to the end points of the /da-/ta/ continuum. The limited attention span of the younger children prompted the reduction in the number of stimuli presented.

#### 4. Data Analysis

In order to compute the CAEP response components, the EEG signals were filtered online (0.1–100 Hz, 24 dB/octave slope). The EEG signals were band-pass filtered offline at 4–30 Hz to allow an N1 response to emerge, if present (Ceponiene et al., 2002; Gilley et al., 2005; Sharma et al., 1997). Artifact rejection was used to exclude responses exceeding  $\pm 100 \mu V$  and eye blinks. Baseline correction was applied for each response. Individual subjects had two averaged CAEP waveforms, resulting in a grand average waveform of 300–400 sweeps for each stimulus presentation.

P1 amplitude and latency was measured for all subjects. The latency was measured at the most positive point of the peak, or halfway on the peak if it was broad. This positive component usually occurred between 75 and 150 ms after presentation of the stimulus. The absolute amplitude was measured at the peak of the P1 response in relation to the average baseline. The N2 was measured as the immediate negative component occurring after the P1, usually present around 250 ms after stimulus presentation. Absolute amplitude and latency were chosen at the most negative point or halfway point of a broad peak. The group average waveforms were filtered at 0.1–30 Hz with a 12 dB/octave slope.

## C. Results

### I. Older children (4–8 year old age group)

**1. Identification Functions**—Figure 1 shows the mean identification scores for the children in the older age group ( $n=10$ ). Stimuli with VOTs of 0 and 20 ms were identified as /da/. Stimuli with longer VOTs, 40 and 60 ms, were identified as /ta/. The boundary between the /da/ and /ta/ categories, i.e., the point of 50% identification, occurred at a VOT of 32.5 ms. These results in children are generally consistent with results for adult subjects whose category boundary occurred at a VOT of 40 ms for the identical stimulus continuum (Sharma and Dorman, 1999, 2000).

**2. CAEP morphology**—Grand average CAEP responses across the /da-/ta/ continuum for the 4–8 year old age group are shown in Figure 2. Visual inspection revealed a robust P1 component at about 100 ms and a N2 component at about 250 ms. As expected, in this age group there was no evidence of an N1 component in the majority of subjects and in the grand average, although a small emerging N1-like component was present in a few children. Overall, the CAEP responses elicited by stimuli with VOT values of 0, 20, 40, and 60 ms were similar in morphology. There was no evidence of a ‘double-on’ or ‘double-peak’ morphology for stimuli with long VOTs. However, latency changes in both the P1 and N2 components were observed as a function of VOT. These changes are described in the sections that follow.

**3. P1 Latency and amplitude**—A two-way repeated measures ANOVA was performed to determine the effect of VOT and electrode site on P1 latency. There was a significant main effect of VOT ( $F(3, 216) = 9.19, p < 0.01$ ) on P1 latency but not of electrode site ( $F(5, 216) = 0.27, p > 0.05$ ), and there was no interaction between the two variables ( $F(15, 216) = 0.32, p > 0.05$ ). Mean latencies for all electrodes and VOT stimuli are shown in Figure 3. Mean latencies and amplitudes for all subjects in all conditions are shown in Table 1. A post hoc pair-wise comparison of means (using the Bonferroni correction) revealed that P1 latencies at 0 and 20 ms differed significantly from latencies at 40 ms and 60 ms across all electrode sites.

P1 absolute amplitude was analyzed using a two-way repeated measures ANOVA, which showed a significant effect of electrode site upon amplitude values ( $F(5, 216) = 3.32, p < 0.01$ ), but not of VOT ( $F(3, 216) = 0.69, p > 0.05$ ). There was no interaction between VOT and electrode site shown ( $F(15, 216) = 0.08, p > 0.05$ ). Post hoc comparison analyses using the Bonferroni correction confirmed that the electrode T<sub>3</sub> differed from C<sub>3</sub> and C<sub>z</sub> across all VOTs. There were no differences in amplitudes as a function of VOT at any electrode sites.

**4. N2 Latency and amplitude**—A two-way repeated measures ANOVA was performed to determine the effect of VOT and electrode site on N2 latency. The analysis revealed a significant main effect of both VOT ( $F(3, 216) = 2.93, p < 0.05$ ) and electrode site ( $F(5, 216) = 7.23, p < 0.01$ ) on N2 latency. There was no interaction between the two variables ( $F(15, 216) = 0.21, p > 0.05$ ). Changes in N2 latencies as a function of VOT and electrode sites are shown in Figure 4 and Table 1. Post hoc comparison of means (using the Bonferroni correction) revealed that N2 latencies at 0 ms and 40 ms VOT differed significantly across all electrodes. Post hoc analyses also showed that, across all VOTs, responses recorded from the temporal (T<sub>3</sub>) electrode had shorter latencies compared to the frontal (F<sub>z</sub>) and central (C<sub>z</sub>, C<sub>3</sub>, and C<sub>4</sub>) electrode locations. N2 latencies at the temporal (T<sub>4</sub>) electrode were also significantly shorter than at C<sub>3</sub>.



N2 absolute amplitude was analyzed using a two-way repeated measures ANOVA. A significant effect of both electrode site ( $F(5, 216) = 3.98, p < 0.01$ ) and VOT ( $F(3, 216) = 7.60, p < 0.01$ ) was found. There was no interaction found between VOT and electrode site ( $F(15, 216) = 0.54, p > 0.05$ ). Post hoc comparisons of the means using the Bonferroni correction revealed the electrode site T<sub>4</sub> to differ from both C<sub>3</sub> and C<sub>z</sub> across VOTs. Comparisons of the means also showed that a VOT value of 60 ms elicited a smaller amplitude response compared to 0, 20, and 40 ms across all electrodes. Though not significant, a slight decrease in N2 amplitude was found as VOT increased (see Table 1).

## II. Younger Children (2–3 year old age group)

**1. CAEP morphology**—For children in the 2–3 year age group, we were able to measure responses from only the C<sub>z</sub> electrode. Grand average CAEP responses to end-points of the /da-/ta/ continuum for the 2–3 year old age group are shown in Figure 5. As was the case for the older children, the morphology of the CAEP response elicited by stimuli with VOT values of 0 ms and 60 ms was similar. And, as was the case for older children, there was an increase in P1 latency for stimuli with longer VOTs. There was no evidence of a double-on response to stimuli with long VOTs.

**2. P1 latency and amplitude**—A one-way repeated measures ANOVA revealed a significant effect of VOT on P1 latency ( $F(1, 14) = 6.30, p < 0.05$ ). Mean P1 latency values for the 0 and 60 ms VOT conditions were 133 and 156 ms respectively (Table 1). Both values were longer than the respective values for subjects in the older age group. These differences were expected given the known relationship between P1 latency and age (Ceponiene et al., 1998; Gilley et al., 2005; Sharma et al., 1997; Sussman et al., 2008).

For children in the 2–3 years age group, a one-way repeated measures ANOVA revealed no significant effect of VOT on absolute amplitude values ( $F(1, 14) = 0.07, p > 0.05$ ).

**3. N2 latency and Amplitude**—In contrast to children in the older age group, N2 responses in children in the 2–3 year old age group did not show a main effect of VOT on either N2 latency ( $F(1,14) = 3.72, p > 0.05$ ) or absolute amplitude ( $F(1, 14) = 1.87, p > 0.05$ ). Though not significant, a decrease in N2 amplitude was found as VOT increased.

## D. Discussion

We examined the representation of VOT in the CAEP waveform of young children in whom the N1 response was not yet developed. Our results demonstrate that a representation of VOT exists in the developing CAEP response, but that representation is different than that in the adult response. For adults, stimuli with a short VOT elicit a single peak (N1') and stimuli with a long VOT elicit a double peak (N1' and N1). In young children the CAEP waveform is dominated by the P1 and N2 responses and the latencies of both vary with VOT. Specifically, stimuli with short VOTs, i.e., 0 and 20 ms, elicit shorter peak latency CAEPs than stimuli with long VOTs, i.e., 40 and 60 ms. (Simos and Molfese, 1997).

The children in our study showed adult-like labeling of the VOT continuum, yet lacked the single vs. double-on response morphology characteristic of adult responses to voiced and voiceless stimuli. This result demonstrates that the link between single and double-on responses recorded from the scalp in humans and the perception of signals as voiced or voiceless is not causal. We have made a similar observation previously (Sharma et al., 2000). However, caution must be exercised when comparing our scalp-recorded results in humans to those from animal studies of intracortical recordings. The scalp-recorded CAEP is a composite wave that reflects activation of multiple auditory cortical fields, each with its own capacity to follow temporal features of complex sounds (Steinschneider et al., 1999;

Fishman et al., 2000b, 2001b). It is therefore difficult to compare results from scalp-recorded responses to the temporal response patterns recorded intracranially from the primary auditory cortex. That is, unlike scalp-recorded responses, intracranially recorded morphological patterns (described in the studies by Steinschneider and colleagues) may well provide biological markers for VOT representation.

The present results indicate that the scalp-recorded CAEP response of the developing CNS is not able to represent the burst and onset of voicing as two separate onset responses. One possible explanation for this is the slower refractoriness of the neuronal population of the generators underlying the immature CAEP response. Neuronal refractoriness is the time needed for a neural population to recover after generating a response to a stimulus before responding to a second stimulus. In the immature central auditory system (CAS) incomplete myelination and synaptogenesis result in longer neuronal refractory periods for CAEP components (Ceponiene et al., 1998; Gilley et al., 2005). Long neuronal refractory periods may explain the absence of a second 'on' response to the onset of voicing in our stimuli. On this view the CAEP response in children represents a merged response containing the representation of both the onset of the syllable (burst) and the onset of voicing as a single broad peak. It would be of interest to examine whether a double peaked CAEP response occurs in young children when the voicing lag is increased to a very long value, e.g., greater than 60 ms, which may allow neurons to recover sufficiently after responding to the syllable onset.

Our results are consistent with studies which describe different developmental trajectories for the different components of the CAEP response (Gilley et al., 2005; Ponton et al., 2000; Sharma et al., 1997; Sussman et al., 2008). In the present study, the older children showed changes in both P1 and N2 latency as a function of VOT changes, while the younger children showed only an increase in P1 latency for a comparable change in VOT. That is, if VOT representation is considered an aspect of functional development of the CAEP response, then our data suggest that 2–3 year old children show a more immature representation of VOT in the CAEP compared to 4–8 year old children. It should be noted that only one active site at  $C_z$  was used to record the CAEPs for the younger children. Additional electrode sites may have yielded additional information regarding possible topographic differences in VOT encoding for older and younger children. In the present study, while the amplitude of the P1 and N2 components for the older children recorded at the temporal electrodes was smaller than that recorded at the frontal and central sites, we did not see clear evidence of the T-complex which has been reported in children at temporal sites (Ponton et al., 2000; Tonnquist-Uhlen et al., 2003). A possible explanation for this discrepancy may lie in methodological differences. Ponton et al., 2000 and Tonnquist-Uhlen et al., 2003 stimulated subjects using a click train (18.9 ms in duration), presented monaurally at a rate of 1.3 per second, and with an interstimulus interval of 2 ms. Our stimuli consisted of speech syllables (200 ms in duration) presented in freefield at a rate of 1.25 per second, and with an interstimulus interval of 800 ms. Further research is needed to examine the effects of stimulation differences on elicitation of the T-complex in children.

In summary, our results show that the double-on morphology of CAEPs recorded from adults in response to stimuli with long VOTs does not appear in young children, and that short vs. long VOTs are represented in the CAEP of young children by changes in P1 and N2 latency.

## Acknowledgments

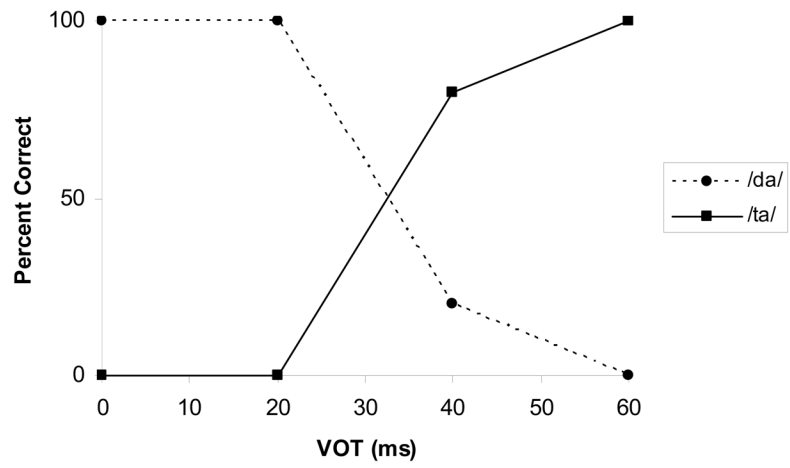
We would like to acknowledge the assistance of Dr. Phillip Gilley and Amy Nash with statistical analysis. We would also like to thank the anonymous reviewers for their comments and insights. This project was supported by grants from the U.S. National Institutes of Health (NIDCD R01DC04552 and R01DC0627) to author A.S.

## References

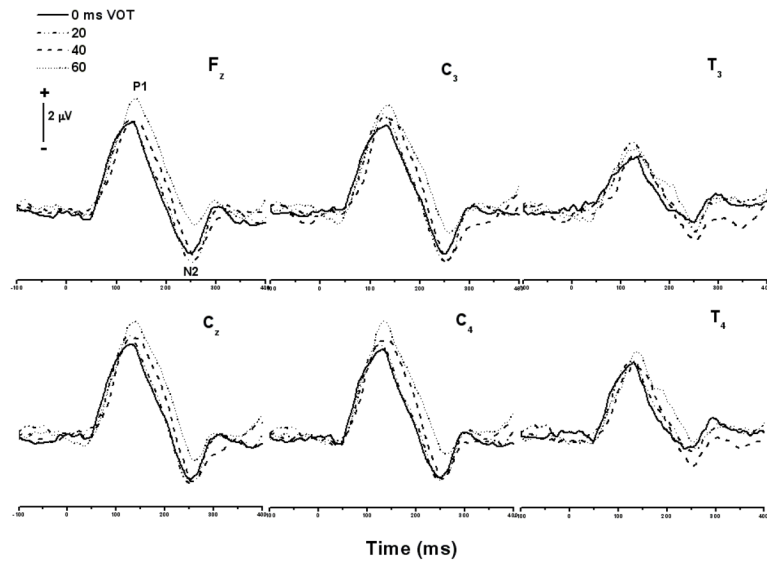
- Abramson, A.; Lisker, L. Discriminability along the voicing continuum: cross-language tests. In: Hala, B.; Romportl, M.; Janota, P., editors. *Proceedings of the Sixth International Congress of Phonetic Sciences*. Prague: Academia; 1970. p. 569-573.
- Ceponiene R, Cheour M, Naatanen R. Interstimulus interval and auditory event-related potentials in children: evidence for multiple generators. *Electroencephalogr Clin Neurophysiol*. 1998; 108:345–354. [PubMed: 9714376]
- Ceponiene R, Rinne T, Naatanen R. Maturation of cortical sound processing as indexed by event-related potentials. *Clin Neurophysiol*. 2002; 113:870–882. [PubMed: 12048046]
- Fishman YI, Reser DH, Steinschneider M. Complex tone processing in primary auditory cortex of the awake monkey. II. Pitch versus critical band representation. *J Acoust Soc Am*. 2000b; 108:247–262. [PubMed: 10923889]
- Fishman YI, Volkov IO, Noh MD, Garell PC, Bakken H, Arezzo JC, Howard MA, Steinschneider M. Consonance and dissonance of musical chords: neural correlates in auditory cortex of monkeys and humans. *J Neurophysiol*. 2001b; 86:2761–2788. [PubMed: 11731536]
- Gilley PM, Sharma A, Dorman M, Martin K. Developmental changes in refactoriness of the cortical auditory evoked potential. *Clin Neurophysiol*. 2005; 116:648–657. [PubMed: 15721079]
- Giraud K, Demonet JF, Habib M, Marquis P, Chauvel P, Liegeois-Chauvel C. Auditory evoked potential patterns to voiced and voiceless speech sounds in adult developmental dyslexics with persistent deficits. *Cereb Cortex*. 2005; 15:1524–1534. [PubMed: 15689520]
- Klatt D. Software for a cascade/parallel formant synthesizer. *J Acoust Soc Am*. 1980; 67:971–995.
- Lieberman AM, Harris KS, Hoffman HS, Griffith BC. The discrimination of speech sounds within and across phoneme boundaries. *J Exp Psychol*. 1957; 54:358–368. [PubMed: 13481283]
- Liegeois-Chauvel C, de Graaf JB, Laguitton V, Chauvel P. Specialization of left auditory cortex for speech perception in man depends on temporal coding. *Cereb Cortex*. 1999; 9:484–496. [PubMed: 10450893]
- Lisker L, Abramson AS. A cross-language study of voicing in initial stops: acoustical measurements. *Word*. 1964; 20:384–422.
- Ponton CW, Eggermont JJ, Kwong B, Don M. Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. *Clin Neurophysiol*. 2000; 111:220–236. [PubMed: 10680557]
- Sharma A, Dorman M. Cortical auditory evoked potential correlates of categorical perception of voice-onset time. *J Acoust Soc Am*. 1999; 106:1078–1083. [PubMed: 10462812]
- Sharma A, Dorman M. Neurophysiologic correlates of cross-language phonetic perception. *J Acoust Soc Am*. 2000; 107:2697–2703. [PubMed: 10830391]
- Sharma A, Kraus N, McGee TJ, Nicol TG. Developmental changes in P1 and N1 central auditory responses elicited by consonant-vowel syllables. *Electroencephalogr Clin Neurophysiol*. 1997; 104:540–545. [PubMed: 9402896]
- Sharma A, Marsh CM, Dorman M. Relationship between N1 evoked potential morphology and the perception of voicing. *J Acoust Soc Am*. 2000; 108:3030–3035. [PubMed: 11144595]
- Simos PG, Diehl RL, Breier JI, Molis MR, Zouridakis G, Papanicolaou AC. MEG correlates of categorical perception of a voice onset time continuum in humans. *Brain Res Cogn Brain Res*. 1998; 7:215–219. [PubMed: 9774735]
- Simos PG, Molfese DL. Electrophysiological responses from a temporal order continuum in the newborn infant. *Neuropsychologia*. 1997; 35:89–98. [PubMed: 8981381]
- Sinex D, McDonald L, Mott JB. Neural correlates of nonmonotonic temporal acuity for voice-onset time. *J Acoust Soc Am*. 1991; 90:2441–2449. [PubMed: 1774413]
- Steinschneider M, Arezzo J, Vaughan HG Jr. Speech evoked activity in the auditory radiations and cortex of the awake monkey. *Brain Res*. 1982; 252:353–365. [PubMed: 7150957]
- Steinschneider M, Arezzo JC, Vaughan HG Jr. Tonotopic features of speech-evoked activity in primate auditory cortex. *Brain Res*. 1990; 519:158–168. [PubMed: 2397403]



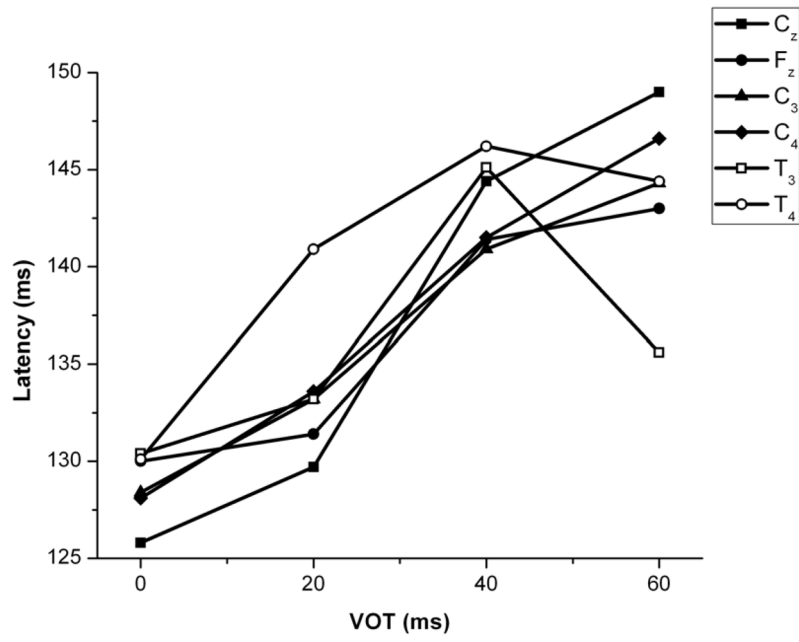
- Steinschneider M, Fishman YI, Arezzo JC. Representation of the voice onset time (VOT) speech parameter in population responses within primary auditory cortex of the awake monkey. *J Acoust Soc Am.* 2003; 114:307–321. [PubMed: 12880043]
- Steinschneider M, Schroeder CE, Arezzo JC, Vaughan HG Jr. Speech-evoked activity in primary auditory cortex: effects of voice onset time. *Electroencephalogr Clin Neurophysiol.* 1994; 92:30–43. [PubMed: 7508851]
- Steinschneider M, Schroeder CE, Arezzo JC, Vaughan HG Jr. Physiologic correlates of the voice onset time boundary in primary auditory cortex (A1) of the awake monkey: temporal response patterns. *Brain Lang.* 1995b; 48:326–340. [PubMed: 7757450]
- Steinschneider M, Volkov IO, Fishman YI, Oya H, Arezzo JC, Howard MA 3rd. Intracortical responses in human and monkey primary auditory cortex support a temporal processing mechanism for encoding of the voice onset time phonetic parameter. *Cereb Cortex.* 2005; 15:170–186. [PubMed: 15238437]
- Steinschneider M, Volkov IO, Noh MD, Garell PC, Howard MA 3rd. Temporal encoding of the voice onset time phonetic parameter by field potentials recorded directly from human auditory cortex. *J Neurophysiol.* 1999; 82:2346–2357. [PubMed: 10561410]
- Sussman E, Steinschneider M, Gumenyuk V, Grushko J, Lawson K. The maturation of human evoked brain potentials to sounds presented at different stimulus rates. *Hear Res.* 2008; 236:61–79. [PubMed: 18207681]
- Tonnquist-Uhlen I, Ponton CW, Eggermont JJ, Kwong B, Don M. Maturation of human central auditory system activity: the T-complex. *Clin Neurophysiol.* 2003; 114:685–701. [PubMed: 12686277]



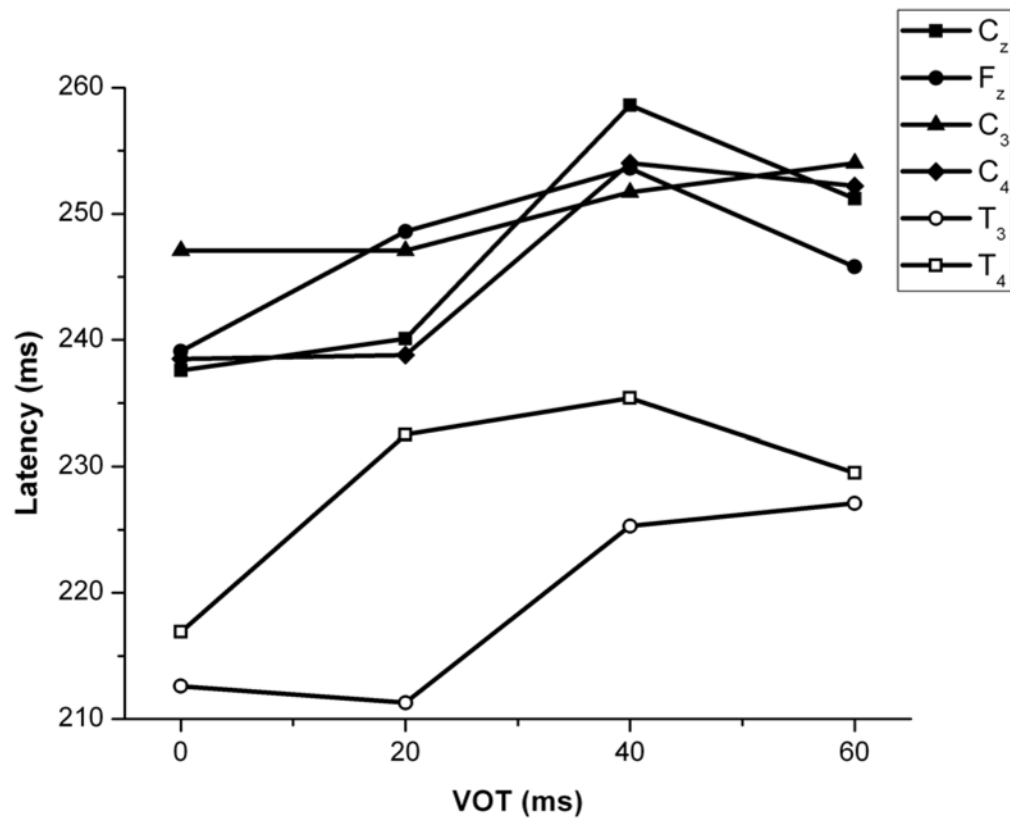
**Figure 1.** Mean Identification scores for the /da/ to /ta/ stimulus continuum for the 4–8 year old age group.



**Figure 2.** Grand average CAEP responses at different electrode sites across the /da-/ta/ continuum for the 4–8 year old age group. The 0 and 20 ms VOT stimuli were perceived as /da/ and the 40 and 60 ms VOT stimuli were perceived as /ta/.

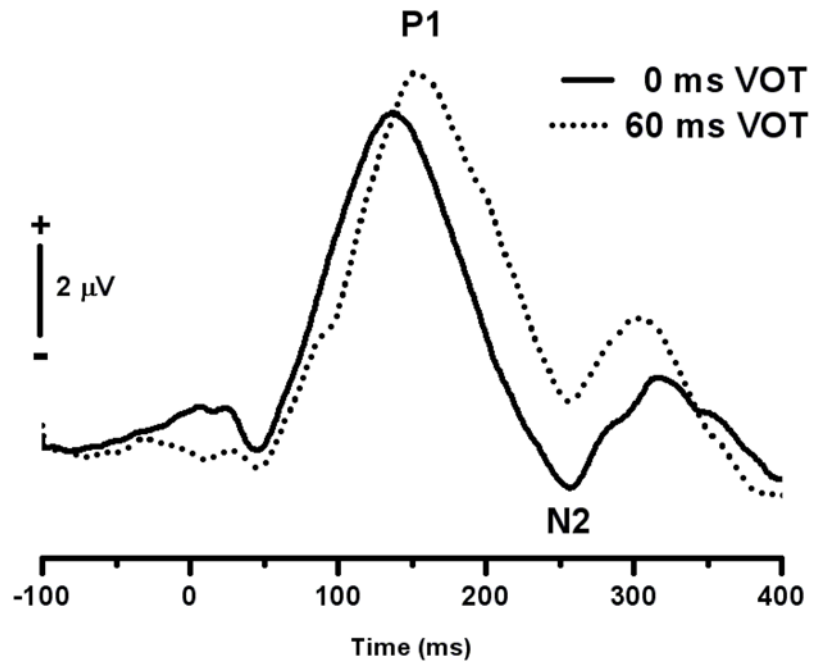


**Figure 3.**  
Mean P1 latencies for the 4–8 year age group as a function of electrode sites and VOT.



**Figure 4.** Mean N2 latencies for the 4–8 year age group as a function of electrode site and VOT.





**Figure 5.**  
Grand average CAEP responses at  $C_z$  for the endpoints of the /da-/ta/ stimulus continuum for the 2–3 year group.

**Table 1**

Mean latency and amplitude values. Standard deviation values are included in parentheses beside the mean values.

Ages 4-8									
P1 Latency (ms)		C <sub>Z</sub>	F <sub>Z</sub>	C <sub>3</sub>	C <sub>4</sub>	T <sub>3</sub>	T <sub>4</sub>		
VOT									
0		126(15)	130(16)	129(19)	128(14)	130(30)	130(14)		
20		130(16)	131(19)	133(17)	134(18)	133(20)	141(28)		
40		144(14)	141(12)	141(12)	141(14)	145(21)	146(19)		
60		149(13)	143(10)	144(11)	147(13)	136(40)	144(16)		
P1 Absolute Amplitude (µV)		C <sub>Z</sub>	F <sub>Z</sub>	C <sub>3</sub>	C <sub>4</sub>	T <sub>3</sub>	T <sub>4</sub>		
0		4.9(2.7)	4.8(3.2)	4.9(2.1)	4.6(2.5)	3.0(2.7)	4.2(2.6)		
20		5.5(3.0)	4.9(2.4)	5.5(2.4)	5.2(2.3)	3.7(2.5)	4.6(2.1)		
40		5.0(3.0)	5.0(2.2)	5.0(2.7)	5.0(2.5)	2.8(2.9)	3.5(2.8)		
60		5.8(3.5)	5.6(3.3)	5.2(2.6)	5.5(3.4)	3.4(3.6)	4.3(4.0)		
N2 Latency (ms)									
0		238(31)	239(19)	247(46)	238(32)	213(39)	217(29)		
20		240(32)	247(11)	247(14)	239(33)	211(33)	232(39)		
40		259(22)	254(13)	251(12)	254(14)	225(33)	235(27)		
60		251(25)	246(18)	254(15)	252(22)	227(50)	229(46)		
N2 Absolute Amplitude (µV)		C <sub>Z</sub>	F <sub>Z</sub>	C <sub>3</sub>	C <sub>4</sub>	T <sub>3</sub>	T <sub>4</sub>		
0		-2.4(1.6)	-2.6(2.2)	-3.3(1.8)	-2.4(1.4)	-2.1(1.9)	-1.5(1.3)		
20		-3.5(2.1)	-3.5(1.9)	-3.7(2.1)	-3.3(1.9)	-1.7(1.7)	-1.8(1.7)		
40		-3.3(1.0)	-2.8(0.9)	-3.3(1.4)	-2.7(1.1)	-2.6(1.5)	-1.6(2.0)		
60		-2.0(2.0)	-1.3(2.1)	-1.5(1.7)	-1.7(1.7)	-1.6(1.6)	-1.1(1.7)		
Ages 2-3									
P1 Latency (ms)		C <sub>Z</sub>							
0		133(18)							
60		156(18)							
P1 Absolute Amplitude (µV)		C <sub>Z</sub>							

Ages 4-8		C <sub>z</sub>	F <sub>z</sub>	C <sub>3</sub>	C <sub>4</sub>	T <sub>3</sub>	T <sub>4</sub>
P1 Latency (ms)							
0		7.5(2.6)					
60		7.8(2.0)					
N2 Latency (ms)		C <sub>z</sub>					
0		242(17)					
60		255(8)					
N2 Absolute Amplitude (μV)		C <sub>z</sub>					
0		-1.19(2.8)					
60		0.97(3.5)					