UC Irvine
UC Irvine Previously Published Works

Title

Auditory cortical activity in normal hearing subjects to consonant vowels presented in quiet and in noise

Permalink

https://escholarship.org/uc/item/8990c5nc

Journal

Clinical Neurophysiology, 124(6)

ISSN

1388-2457

Authors

Dimitrijevic, Andrew
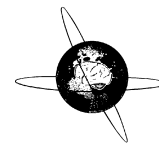Pratt, Hillel
Starr, Arnold

Publication Date

2013-06-01

DOI

10.1016/j.clinph.2012.11.014

Copyright Information

Peer reviewed

Auditory cortical activity in normal hearing subjects to consonant vowels presented in quiet and in noise

Andrew Dimitrijevic [a,*], Hillel Pratt [c], Arnold Starr [b]

[a] Communication Sciences Research Center, Cincinnati Children's Hospital Medical Center, Department of Otolaryngology, Head and Neck Surgery, University of Cincinnati, USA
[b] Department of Neurology, University of California, Irvine, USA
[c] Evoked Potentials Laboratory, Technion – Israel Institute of Technology, Haifa, Israel

HIGHLIGHTS

- Cortical evoked potentials (N100) to speech can be recorded as "onset responses" when the consonant-vowel is in first position and as "change responses" when it follows a vowel.
- N100 latency was prolonged in consonant-vowels with long voice onset times only in change responses whereas only N100 amplitude was affected by voice onset time in onset responses.
- Source estimations indicated similar regions of cortical activation during N100 of change and onset responses to speech, whereas the sources of the subsequent slow cortical potential changed across time.

ABSTRACT

Objective: Compare brain potentials to consonant vowels (CVs) as a function of both voice onset times (VOTs) and consonant position; initial (CV) versus second (VCV).

Methods: Auditory cortical potentials (N100, P200, N200, and a late slow negativity, (SN) were recorded from scalp electrodes in twelve normal hearing subjects to consonant vowels in initial position (CVs: /du/ and /tu/), in second position (VCVs: /udu/ and /utu/), and to vowels alone (V: /u/) and paired (VVs: /uu/) separated in time to simulate consonant voice onset times (VOTs).

Results: CVs evoked "acoustic onset" N100s of similar latency but larger amplitudes to /du/ than /tu/. CVs preceded by a vowel (VCVs) evoked "acoustic change" N100s with longer latencies to /utu/ than /udu/. Their absolute latency difference was less than the corresponding VOT difference. The SN following N100 to VCVs was larger to /utu/ than /udu/. Paired vowels (/uu/) separated by intervals corresponding to consonant VOTs evoked N100s with latency differences equal to the simulated VOT differences and SNs of similar amplitudes. Noise masking resulted in VCV N100 latency differences that were now equal to consonant VOT differences. Brain activations by CVs, VCVs, and VVs were maximal in right temporal lobe.

Conclusion: Auditory cortical activities to CVs are sensitive to: (1) position of the CV in the utterance; (2) VOTs of consonants; and (3) noise masking.

Significance: VOTs of stop consonants affect auditory cortical activities differently as a function of the position of the consonant in the utterance.

© 2012 International Federation of Clinical Neurophysiology. Published by Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Stop consonant-vowel (CV) utterances (e.g., /du/ and /tu/), are distinguished by acoustic differences of the time between onset of the consonant ("stop release burst") and onset of the following vowel ("voicing"). The difference in ms between the onsets of the consonant and the vowel is known as "voice onset time" or VOT (Lisker and Abramson, 1964). The consonant /d/ is identified when

VOTs are relatively brief (<25 ms in English) whereas /t/ is identified when VOTs are relatively long (>45 ms in English). Consequently, discrimination between boundaries (i.e., /tu/ vs. /du/) is better than discrimination within boundaries ("short" /tu/ vs. "long" /tu"). VOTs that fall within these boundaries, are more variably classified as /t/ or /d/ when compared to short or long duration VOTs. The temporal boundaries described above are typical for English speakers (Lisker and Abramson, 1964) but can differ depending on the native language of the subject (e.g., Williams, 1977 for Spanish; Laufer, 1998 and Horev et al., 2007 for Hebrew).

There has been considerable interest in characterizing the central auditory processes underlying categorical characterization of stop consonants. Sharma and Dorman (1999, 2002) showed that CV syllables with short VOTs evoked a single N100 response to vowel-onset while syllables with long VOTs evoked a double N100 response: the first peak coincided with the onset of the consonant and the second with the onset of the vowel (Sharma and Dorman, 1999, 2000). However, these N100 changes were not systematically related to the temporal boundaries of category specific perception of stop consonants (Sharma and Dorman, 2000; Sharma et al., 2000; Eggermont and Ponton, 2002). Using synthetic speech stimuli, longer VOTs have resulted in delayed N100s (Frye et al., 2007; Tremblay et al., 2003). N100 amplitudes have also been described as being larger with short VOTs compared to long VOTs (Toscano et al., 2010) without any accompanying changes of latencies (Tremblay et al., 2003b). Toscano's study demonstrated that subjects engaged in an active task of classifying words based on VOT differences of the initial stop consonant (e.g., "beach" or "peach") showed change in amplitude of both N100 (sensitive to "sensory" features of the stimulus) and P300 (sensitive to "perceptual" features of stimulus) with changes of VOTs of the target stop consonant. Interestingly, there were no abrupt changes of amplitudes of either N100 or P300 corresponding to category boundaries.

Understanding the normal neural processing of speech, and particularly the role of temporal cues such as VOT, would provide insight into potential rehabilitative strategies of clinical populations with temporal processing disorders such as in auditory neuropathy (AN). AN is a disorder of speech comprehension with impaired auditory nerve function (AN; Starr et al., 1996). The magnitude of their speech comprehension deficits is larger than expected from the patients' degree of audibility changes, and reflects impaired processing of auditory temporal cues (Starr et al., 1996; Zeng et al., 2005). Specific etiologies for the disorder affect both pre-synaptic sites (e.g., impaired function of ribbon synapses of inner hair cells; Varga et al., 2006; Marlin et al., 2010) and post-synaptic sites (e.g., impaired function of auditory nerve fibers; Starr et al., 2003). Detailed analyses of the word processing errors in a group of AN subjects with Freidreich's ataxia, a post-synaptic disorder of neural transmission, showed abnormal identification of stop consonants distinguished by VOT (/tu//du/, /ba/ /pa/, /ka/ /ga/) but normal classification of sibilant consonants containing high frequencies (e.g., /s/ vs. /f/). In contrast, patients with a sensory hearing loss typically have an opposite pattern of speech comprehension deficits with impaired discriminations of high frequency fricatives and normal discrimination of stop consonants (Rance et al., 2008).

In this paper we report results of scalp-recorded brain potentials from normal-hearing subjects to spoken stop consonant vowel combinations presented with the consonant at the initial position of the phoneme (CV: /du/ and /tu/), or at second position following an initial vowel (VCV: /udu/ and /utu/). Acoustic features located at the initial position of a stimulus sequence evoke "acoustic onset responses" (N100/P200) of large amplitude and short latency. The same acoustic features located at later positions of the stimulus sequence evoke "acoustic change responses" (N100) that are both reduced in amplitude and delayed in latency compared to these measures in initial position (Hari et al., 1988; Jones and Perez, 2001; Martin and Boothroyd, 1999, 2000). The present study examined the effect of consonant position in the utterance on auditory cortical potentials. We measured latencies and amplitudes of scalp potentials, their global field power (GFP), and estimated their brain sources as a function of: (1) consonant position in the phoneme; (2) VOTs; and (3) noise masking. We hypothesized that auditory cortical activities would show amplitude and latency differences when comparing responses to CVs with different VOTs, in initial and second positions in the utterance (onset and change responses). Furthermore, we expected noise masking to alter the relative contribution of vowel and consonant acoustic cues to the evoked brain activities.

2. Methods

2.1. Subjects

Twelve (4 males, 8 females) subjects (mean age: 20 years, all self-reported right-handed) with normal pure tone thresholds (500–8000 Hz) and no history of neurological illnesses participated in the study. All subjects were tested in quiet and 6 of these subjects were tested with noise masking as well. All subjects gave informed consent prior to testing. All except one subject were tested with left ear stimulus presentation.

2.2. Stimuli

Speech sounds were recorded by a native English speaking male. Speech utterances included /u/, /tu/, /du/, /utu/, /udu/ and /uu/ (see Fig. 1). The /uu/ included two stimulus conditions (/uu$_{15ms}$/, /uu$_{110ms}$/) with temporal separations between the vowels that corresponded to the VOT separations in /udu/ and /utu/. The /uu/ conditions allowed the examination of brain activity to two vowels without an intervening stop consonant. The comparison of /tu/-/du/ and /utu/-/udu/ was used to test effects of consonant position in the bisyllabic utterances by evoking "acoustic onset responses" (/tu/ and /du/) and "acoustic change responses" (/udu/ and /utu/). The stimuli were created by manipulating a real speech recording of /u/ and /tu/ in CoolEdit. In this talker, a naturally produced /tu/ had a 110 ms noise burst (aspiration) before the vowel (periodic voicing) began; this condition will be referred to as the 110 ms voice onset time (VOT) condition. Removing the first 85 ms of aspiration in the /tu/ changed the perception from /tu/ to /du/; in this condition (15 ms VOT), there were only 15 ms of aspiration left from the initial 110 ms before the vowel began. The total duration of the /u/ stimulus was 160 ms, the /tu/ stimulus was 380 ms (110 ms VOT) and 285 ms for the /du/ stimulus (15 ms VOT). The /utu/ and /udu/ stimuli were created by concatenating the /tu/ or /du/ stimulus at the offset of the /u/ with a 5 ms gap. This gap was chosen as it approximates a "natural" sounding /utu/ or /udu/. The /uu/ condition was created by concatenating a more intense (6 dB) and time shifted version of the initial /u/ with an RMS value for the second /u/ equivalent to that of the /u/ in /tu/. Their acoustic waveforms are shown in Fig. 1. All speech stimuli were presented at 90 dB SPL. The rise time of /tu/ and /du/ was less than 5 ms; and the RMS of the initial 50 ms of /tu/ and /du/ was .0076 and .0040, respectively (arbitrary units of the sound file). Speech-shaped-noise maskers were continuously presented for the entire duration of the recording at 85 dB SPL. Both noise and speech sounds were calibrated using a continuous peak SPL measured on F setting. All sound stimuli were presented through Etymotic © ER3 insert earphones.

2.3. Categorical perception of CVs and VCVs

The identification of the CVs and VCVs was performed using a 2-alternative forced choice identification paradigm. CVs consisted of a /du/-/tu/ continuum with VOTs of 15, 25, 45, and 65 ms. Similarly, for VCVs an /udu/-/utu/ continuum with VOTs of 15, 25, 45, and 65 ms was used. After an initial training period, 20 tokens at each of the VOTs were randomly presented at 90 dB SPL. CVs and VCVs and the effect of speech-shaped noise masker were evaluated in separate runs. Subjects were instructed to indicate via button
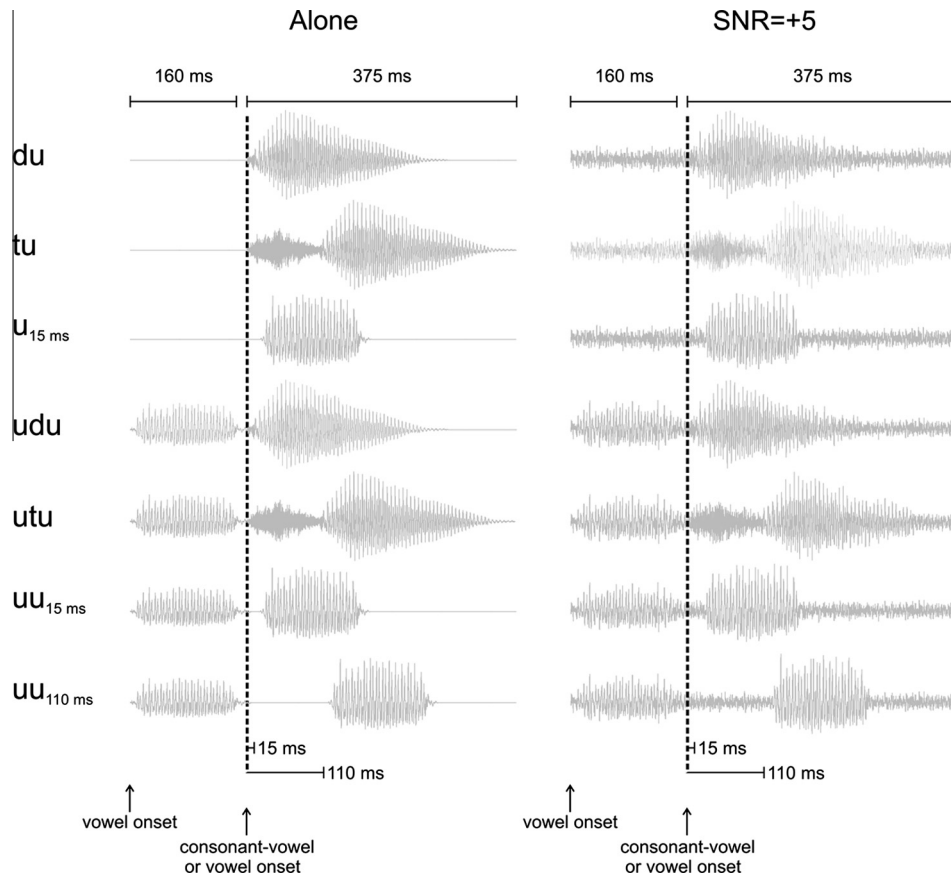
Fig. 1. Acoustic stimuli consisted of consonant-vowels (CVs) or vowels (Vs) in initial position to elicit "onset" responses (top three rows) and in second position following/u/ to elicit "change" responses (fourth and fifth rows). The consonants had VOTs of 15 for /d/ and 110 ms for /t/. Paired vowels /uu/were also separated by 15 and 110 ms (bottom two rows) to define cortical activities in the absence of consonants. These utterances were presented alone at 90 dB SPL (left column) and with speech-shaped noise masker at 85 dB SPL (SNR = +5), right column.

press which token was heard (/du/ or /tu/ for CVs and /udu/ or /utu/ for VCVs).

CV and VCV identification were analyzed using logistic regression (McCullagh and Nedler, 1983). The probability of identifying /tu/ or /utu/ was examined as a function of VOT. Data were fitted using customized software written in Matlab using the following equation:

$$f(VOT) = 1/e^{-z} \quad \text{where} z = \beta_0 + \beta_1 \times VOT$$

This function represents the probability of identifying /tu/ or /utu/ as a function of VOT (in ms). For each subject, the parameters β_0 and β_1 were allowed to vary. If fits were significant at the $p < 0.05$ level, the β_0 and β_1 values were used to create a continuous function that allowed estimations of the VOT value (in ms) that represented the 50%, or midway point, where there was equal probability of hearing /tu/ (or /utu/) and /du/ (or /udu/). This point was defined as the category boundary.

The behavioral identification task was performed by all 12 subjects for the alone condition and 6 of these subjects participated in the noise condition.

2.4. Recordings

A 64-channel Neuroscan Synamps[2]© recording system was used to collect electrophysiological data. Electrode placements included the standard 10–20 locations and intermediate sites. Impedances were kept below 10 kΩ. Vertical and lateral eye movements were monitored using two bipolar electrodes above and below the right eye and two bipolar electrodes on the left and right

outer canthi, for defining the vertical and horizontal electro-oculogram (EOG), respectively. Signals were digitized at 1000 Hz, amplified by a factor of 2010, and band-pass filtered (cutoffs at 0.05 and 200 Hz). Epochs were extracted using a −200 to 1900 ms window relative to the onset of the speech utterance. Offline analysis included re-referencing the recordings to an average reference (excluding the EOGs). Eye movement effects on scalp potentials were removed offline in the continuous recording in each subject using a singular value decomposition-based spatial filter utilizing principal component analysis of averaged eye blinks for each subject (Ille et al., 2002). All electrophysiological recordings were performed while subjects passively watched a muted, closed-captioned movie of their choice.

One hundred repetitions of each stimulus were presented to the subject. The inter-stimulus interval was set to 2.2 seconds.

Cortical potentials in the quiet condition were recorded to: CVs from all 12 subjects, VCVs from 11 subjects, and Vs and VVs from 6 subjects. Cortical potentials in the noise masking condition were recorded to: CVs, VCVs, Vs and VVs from 6 subjects.

2.5. Waveform analysis

N100 and P200 peaks were identified in the FCz channel and global field power (GFP) was calculated as the variance for each time point across all 64 channels. FCz was chosen because typically this channel yields the largest auditory cortical evoked potentials. GFP was chosen because it gives an indication of the strength of all the components across the scalp. N100 was defined at the most negative peak in the 80–250 ms following the consonant or vowel

onset. The N100 was confirmed present by visual inspection if all three conditions were met: (1) peak within the 80–250 ms time window; (2) a similar latency peak in the GFP (latency differences up to 10 ms were accepted), and (3) a polarity inversion of the peak at the mastoids compared to FCz (latency differences of up to 5 ms were accepted). N100 amplitudes were quantified as the mean voltage of a 50 ms window centered on the N100 latency based on the grand mean waveform. The nomenclature adopted for the potentials included the stimulus-evoking event-related component. For example, when the stimulus was a consonant vowel (CV) the N100 was designated as $N100_{CV}$. A low amplitude positivity at 50 ms (P50) was not consistently observed and therefore not studied further. A slow negativity (SN) was seen as a prolonged negativity with acoustic change stimuli (i.e., /udu/, /utu/, and /uu/). This potential was quantified by applying a low pass filter (5 Hz) and measured as the average voltage over a 500 ms interval beginning 150 ms after the $N100_{CV}$. Both the time delay and filtering attenuated the contribution of the immediately preceding consonants' N100–P200 potentials to the CV. The time interval used for the baseline was 200 ms before the onset of the initial vowel in VCVs (or VVs), even when the stimulus was a CV or a V. Fig. 1 illustrates the temporal relationship of the stimuli. For the purposes of clarity, waveforms in all figures are plotted with a vertical line at "0 ms" that represents the onset of the CV stimulus, but in all cases, the baseline used for baseline correction was 200 before the initial vowel even in the absence of an initial vowel (i.e., Vs and CVs).

2.6. Source analysis

Brain source analysis was performed on grand averaged data from five subjects using BESA (Brain Electrical Source Analysis 5.3) software in response to left ear stimulation (i.e., /du/, /tu/, /u/, /udu/, /utu/, /uu_{15ms}/ and /uu_{110ms}/). Two types of source analysis were performed, a distributed source model: CLARA (Classical LORETA recursively applied) and single equivalent dipoles. A combined approach was employed where CLARA was applied for each point over the entire 1900 ms waveform. Maximal activation CLARA sources were examined at the peak of the GFP representing N100 and localized to both temporal lobes. Single equivalent dipoles were placed at left and right temporal regions previously identified by the CLARA maximal activations. Dipole fits, constrained in location by CLARA maxima were evaluated using a 40 ms window centered on N100 peak latency measured at GFP. All dipole results yielded goodness of fits greater than 85%. In the CLARA procedure, regularization constants affect the distribution of the source estimates (BESA manual). We utilized a 0.5 regularization parameter for the first iteration and 0.001 for the second iteration.

2.7. Statistical analysis

Repeated measures analysis of variance procedures (ANOVA) were used to examine the effects of experimental conditions on evoked potential data. Post-hoc comparisons were made using Tukey Honestly Significant Difference test. The ANOVA factors used depended on the context of the results and are detailed in the results section.

3. Results

3.1. Psychoacoustics

Classification of stop consonants as a function of VOT in quiet and in noise: With short duration VOTs (15 ms), subjects identified the phoneme as a "du" or "udu" whereas with longer duration VOTs (110 ms) subjects identified the phoneme as "tu" or "utu", mean values are shown Table 1. The identification of /du/ and /tu/ as a function of VOTs of 15 and 110 ms, both in quiet and in noise (SNR = + 5 dB) was not significantly different whether presented as CV or as VCV. Differences between 15 and 110 ms in both quiet and in noise (SNR = +5 dB) and as CV or VCV were all highly significant ($p < 0.001$ using a paired t-test). The output of the logistic regression showed that all regression fits were significant at $p = 0.05$. The 50% identification point showed no significant differences between CV or VCV nor between presentation in quiet or in noise.

3.2. Potentials to CVs and Vs in initial and second positions (onset vs. change)

Fig. 2, shows grand averaged potentials (FCz) and global field power (GFP) to CVs, VCVs, and VVs. The time scale is arranged so that the onset of the CVs occurs at 0 ms (vertical interrupted lines) and the onsets of vowels, when presented alone, were delayed by 15 or 110 ms, simulating the presence of consonants. When CVs were in initial position, i.e., onset (black traces for /du/, and /tu/), auditory cortical activities consisted of $N100_{CV}$, $P200_{CV}$, and $N200_{CV}$ (marked by black circles and only labeled for /du/). A low amplitude P50 component (not labeled) occurred immediately before N100. When CVs were in second position, i.e., change (red traces for /udu/ and /utu/), the initial low amplitude vowel /u/ elicited $N100_V$ and $P200_V$ components (−50 and 50 ms) succeeded by $N100_{CV}$ that was followed by a slow negativity (SN). The GFP had two peaks ($N100_{CV}$, $P200_{CV}$) for both /du/ and /tu/, but only a single peak corresponding to N100 for /udu/. For /utu/, the GFP was broad and prolonged, but the N100 peak was still evident. A low amplitude N200 component occurred with onset responses while a slow negativity (SN) occurred with change responses. Table 2 contains the mean peak latency and mean amplitude of N100 and SN to CVs VCVs, VVs and Vs. Onset-evoked potentials (CV and V) had consistent N100 and P200 components, whereas change-evoked potentials had low amplitude N100 (but no P200) response that was followed by an SN.

CV N100s to acoustic change (/udu/, /utu/, /uu_{15ms}/, and /uu_{110ms}/) were significantly delayed compared to onset responses for both CVs and Vs (all $p < 0.01$).

These results suggest that patterns of auditory cortical activity evoked by V and VV were of similar form to those evoked by CV and VCV, respectively (Fig. 2, bottom two traces). These results also

Table 1
Summary of behavioral identification of CVs and VCVs presented alone and with noise masking. Columns 3 to 6 indicate the mean (SD) percent of the trials (out of 20) where subjects reported hearing "tu" or "utu" as a function of VOT. The last column represents the mean (SD) VOT value for the estimated 50% performance, or equiprobable hearing of "du" and "tu" (or "udu" and "utu") estimated based on the logistic function (see Methods for calculation).

		Percent identified as "tu" or "utu"				50% point (ms)
		15 ms VOT	25 ms VOT	45 ms VOT	110 ms VOT	
Alone	/du-tu/	0 (0)	1 (2)	96 (5)	100 (2)	33 (6)
	/udu-utu/	5 (8)	22 (21)	94 (11)	99 (2)	34 (6)
SNR = +5	/du-tu/	1 (3)	8 (13)	96 (5)	100 (0)	33 (4)
	/udu-utu/	3 (8)	21 (22)	94 (11)	97 (7)	32 (3)
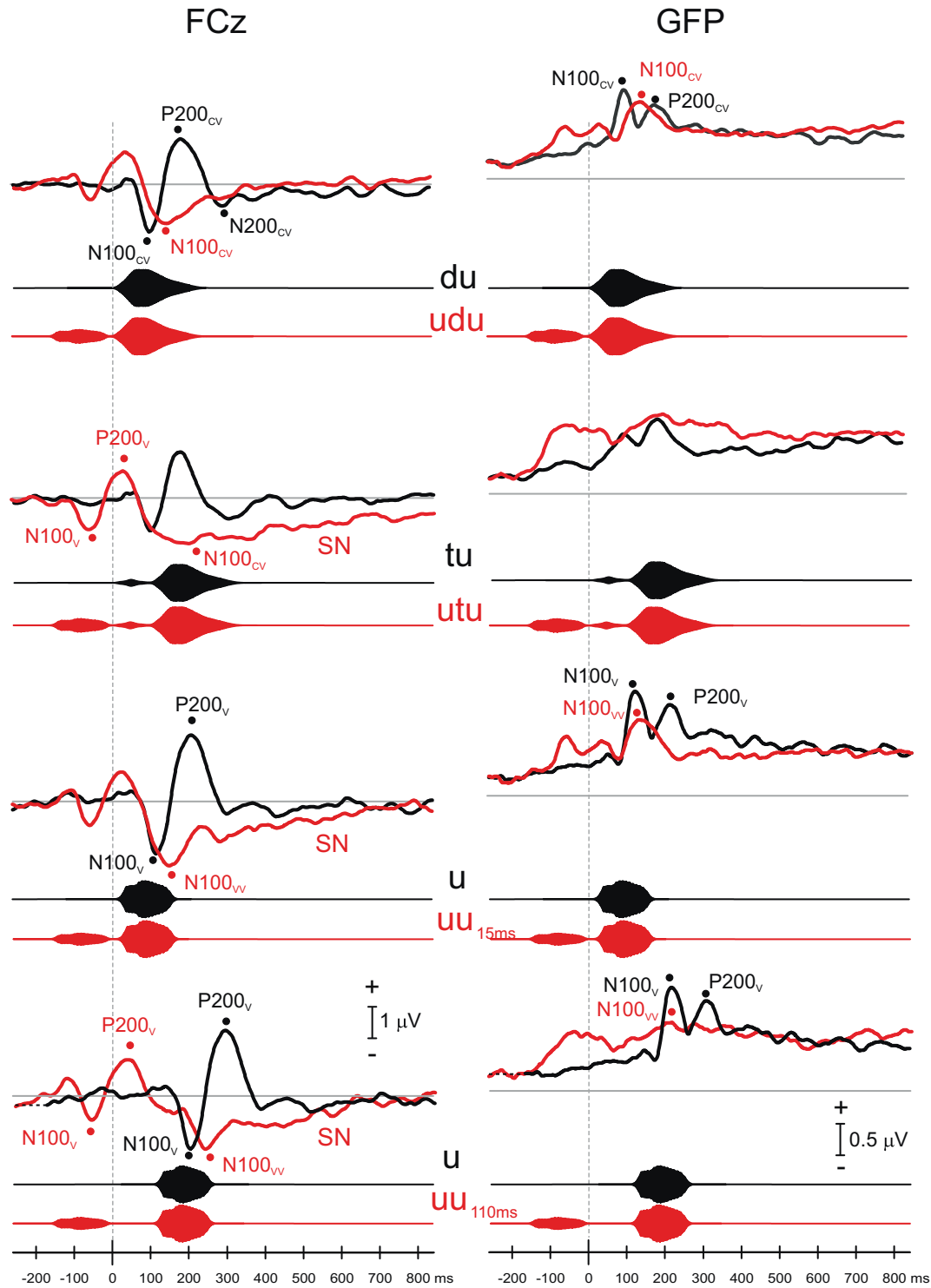
Fig. 2. Auditory cortical activities recorded at electrode FCz (left), and their Global Field Power (GFP; right). Consonant-vowels (CVs) in initial position (/du/ or /tu/, black traces) evoke "onset" responses that differ from CVs in second position (/udu/ or /utu/, red traces) that evoke "change" responses. The time scale has been adjusted so that the consonant onset for CVs presented both alone (/du/ and /tu/) and with a preceding vowel (/udu/ and /utu/) occurs at 0 ms. Also shown are the vowels (Vs) presented alone and for paired vowels (VVs; bottom two panels for vowel separations of 15 and 110 ms). The onset/$u_{110\ ms}$/response (bottom black trace) is identical to/$u_{15\ ms}$/except that it is shifted by 85 ms in order to correspond to the 110 ms difference of VOTs (indicated by dotted line). The corresponding acoustic stimuli are plotted below each evoked response. Note that for onset responses to CVs, the components identified are N100, P200, and N200. For change responses to CVs (VCVs), the initial vowel N100 is at −50 ms, P200 is at 50 ms followed by a large amplitude CV N100 to/udu/. To/utu/there is a prolonged slow negative potential (SN) that contains a low amplitude N100 at approximately 190 ms. Grand average waveforms are based on: /du/ and /tu/ $n = 12$; /udu/ and /utu/ $n = 11$; /u/,/uu_{15ms}/ and /uu_{110ms}/$n = 6$.

document differences of auditory cortical activity to CVs in initial position (acoustic onset) compared to CVs in second position when preceded by a vowel (acoustic change), similar to auditory cortical

responses to pure tone onset (acoustic onset) and to a frequency change in an ongoing pure tone (acoustic change) (see Pratt et al., 2009).

Table 2
Latency (SD) and mean amplitude (SD) of cortical potentials recorded at FCz and Global Field Power to VCVs, VVs, CVs, and Vs in quiet.

Condition		N100		SN
	Quiet	Latency (ms)	Amplitude (μV)	Amplitude (μV)
FCz	Onset – /du$_{15 ms}$/	99 (9)	−1.26 (0.87)	–
	Onset – /tu$_{110 ms}$/	98 (10)	−0.77 (0.69)	–
	Onset – /u$_{15 ms}$/	120 (12)	−1.22 (1.11)	–
	Change – /udu$_{15 ms}$/	142 (17)	−1.16 (0.52)	0.02 (0.44)
	Change – /utu$_{110 ms}$/	187 (24)	−1.21 (0.64)	−0.50 (0.88)
	Change – /uu$_{15 ms}$/	151 (17)	−1.80 (0.49)	−0.40 (0.38)
	Change – /uu$_{110 ms}$/	243 (10)	−1.33 (0.56)	−0.25 (0.11)
GFP	Onset – /du$_{15 ms}$/	101 (9)	1.21 (0.55)	–
	Onset – /tu$_{110 ms}$/	97 (15)	0.92 (0.53)	–
	Onset – /u$_{15 ms}$/	123 (10)	1.45 (0.91)	–
	Change – /udu$_{15 ms}$/	141 (24)	1.12 (0.40)	0.56 (0.25)
	Change – /utu$_{110 ms}$/	194 (30)	1.14 (0.45)	0.82 (0.32)
	Change – /uu$_{15 ms}$/	153 (15)	1.65 (0.66)	0.59 (0.17)
	Change – /uu$_{110 ms}$/	245 (11)	1.46 (0.66)	0.57 (0.09)

3.3. Potentials to CV with different VOTs (15 vs. 110 ms)

Fig. 3 contains grand averaged cortical potentials to both CVs ("onset" responses, top set of traces) and VCVs ("change" responses, middle set of traces) to consonants with "short" (/d/, VOT 15 ms) and "long" (/t/, VOT 110 ms). N100$_{CV}$ amplitudes to CVs were larger to /du/ than /tu/ whereas their latencies did not differ. However, for VCVs, the N100$_{CV}$ was of longer latency for /utu/ than /udu/. The slow negativity (SN) that follows N100$_{CV}$ was significantly larger for /utu/ than /udu/.

For "acoustic onset" responses (/du/ and /tu/), a repeated measures ANOVA showed that N100$_{CV}$ were significantly larger for /du/ than /tu/ (FCz: −1.26 and −0.77 μV respectively, $p = 0.009$; GFP: 1.21 and 0.92 μV respectively, $p = 0.034$). N100$_{CV}$ latencies were not significantly different (FCz: 99 and 98 ms and GFP peak: 101 and 97 ms).

For "acoustic change" responses (i.e., /udu/ vs. /utu/), the amplitude of N100$_{CV}$ did not significantly differ between /udu/ and /utu/ (FCz: −1.16 and −1.21 μV respectively; GFP: 1.12 and 1.14 μV respectively). In addition, no significant amplitude differences between /udu/ and /utu/ were observed when the "prestimulus baseline" was modified to 200 ms prior to the CV as opposed to 200 ms prior to the initial vowel. N100$_{CV}$ latency was significantly different (FCz: 142 vs. 187 ms, $p = 0.002$, respectively; GFP peak: 141 vs. 194 ms, $p < 0.001$, respectively). Moreover the extent of the N100$_{CV}$ latency difference (45 and 53 ms for FCz and GFP respectively) was smaller than the corresponding acoustic difference of the VOTs employed (95 ms).

A slow negativity (SN) following N100$_{CV}$ was significantly of greater amplitude to /utu/ compared to /udu/ (Fig. 4; FCz: −0.50 vs. 0.02 μV, $p = 0.014$; GFP: 0.82 vs. 0.56 μV, $p = 0.005$).

3.4. Potentials to paired vowels /uu/ with different temporal separations (15 ms vs. 110 ms)

Fig. 3 also illustrates cortical potentials to paired vowels (VV; /uu/) to help address the contributions of the vowels to acoustic change activities accompanying /udu/ and /utu/. The overall finding here was that N100$_{VV}$ differences between 15 and 110 ms VOTs was 92 ms, similar to the acoustic difference of 95 ms. The /uu/ related potentials without an intervening consonant and with temporal gaps corresponding to the VOTs compared in the CV conditions (/uu$_{15 ms}$/ and /uu$_{110 m}$/) are shown at the bottom of Fig. 3. The cortical potentials accompanying /uu/ had a similar appearance to those accompanying /udu/ and /utu/. N100$_V$ latencies but not amplitudes were significantly different for vowel sep-

arations of 15 and 110 ms (FCz: 151 ms vs. 243 ms; $p < 0.001$; GFP peak: 153 vs. 245 ms, $p < 0.001$) and their difference of 92 ms corresponded to the 95 ms difference between the two vowels. SN amplitude did not differ significantly between /uu$_{15 ms}$/ and /uu$_{110 ms}$/.

We also examined whether the cortical potentials differed between VV and VCVs. A repeated measures ANOVA showed a significant interaction for N100 latency between VOT and stimulus type (/uu/ vs. /utu/) for FCz (F(1,6) = 6.6; $p = 0.042$). Posthoc analysis showed that N100$_{VV}$ was prolonged compared to N100$_{CV}$ for 110 ms VV separation ($p < 0.001$) but not for 15 ms VV separation. Differences in SN amplitudes between VV and VCV did not reach significance.

3.5. Effects of noise masking on onset and change potentials responses

Noise masking (SNR = + 5 dB) affected latency of N100 VCVs (see Fig. 5 and Table 3). Most striking was that with noise masking, N100 latency differences (with VCVs at FCz) between VOTs of 15 and 110 ms were 93 ms, close to the acoustic difference of 95 ms, whereas in the quiet condition, this difference was only 45 ms. N100 amplitude of CVs presented in noise were significantly decreased compared to quiet (main effect of noise at FCz: $p < 0.001$ and GFP; $p = 0.022$). In contrast to the quiet condition, N100 amplitudes between /du/ and /tu/ were no longer significant. It should be noted that although still recognizable, N100 amplitudes were smaller with noise masking and therefore differences between /du/ and /tu/ may be harder to detect. Similar to VCVs N100 latencies were significantly prolonged in noise for both /du/ and /tu/ at FCz ($p = 0.037$).

Latencies, but not amplitudes, of N100 to VCVs, were affected by noise. N100$_{CV}$ latency was prolonged in noise compared to quiet, and posthoc analysis showed that only the response to /utu/ was delayed in noise compared to quiet (187 vs. 245 ms; $p = 0.011$) while /udu/ latency was unaffected. N100$_{CV}$ latency differences between 15 and 110 ms (/udu/ and /utu/) persisted and were significantly increased with noise masking ($p < 0.001$), to correspond to the respective VOT difference (95 ms). Although SN amplitude differences between 15 and 110 ms were evident without masking, noise masking obliterated this effect. No effects of noise masking were seen with VVs.

3.6. Source analysis of auditory cortical potentials

Scalp distributions of N100 to CV as well as to the vowels alone had a negativity that was most pronounced fronto-centrally with a slight bias to the right. The posterior and lateral electrodes were most positive with a slight bias to right posterior scalp. This distribution is compatible with a generator in the right temporal lobe, and source current density and equivalent dipole preliminary estimates on grand-averaged waveforms confirmed this putative location (Fig. 6). Fig. 6 shows that dipole and CLARA source analysis of N100 showed similar locations for both CVs and VCVs. Fig. 7 summarizes the time course of the SN source analysis. The lower left panel summarizes a progressive anterior to posterior shift in SN source with time.

4. Discussion

The present study of auditory cortical activities to stop consonants in normal-hearing subjects revealed three findings: (1) CVs in initial (/du/ and /tu/) and in second positions (/udu/ and /utu/) evoked different auditory cortical activities. In initial position cortical activities consisted of P50, N100, P200, and N200 components as in "acoustic onset" responses to tone bursts (Onishi and Davis,
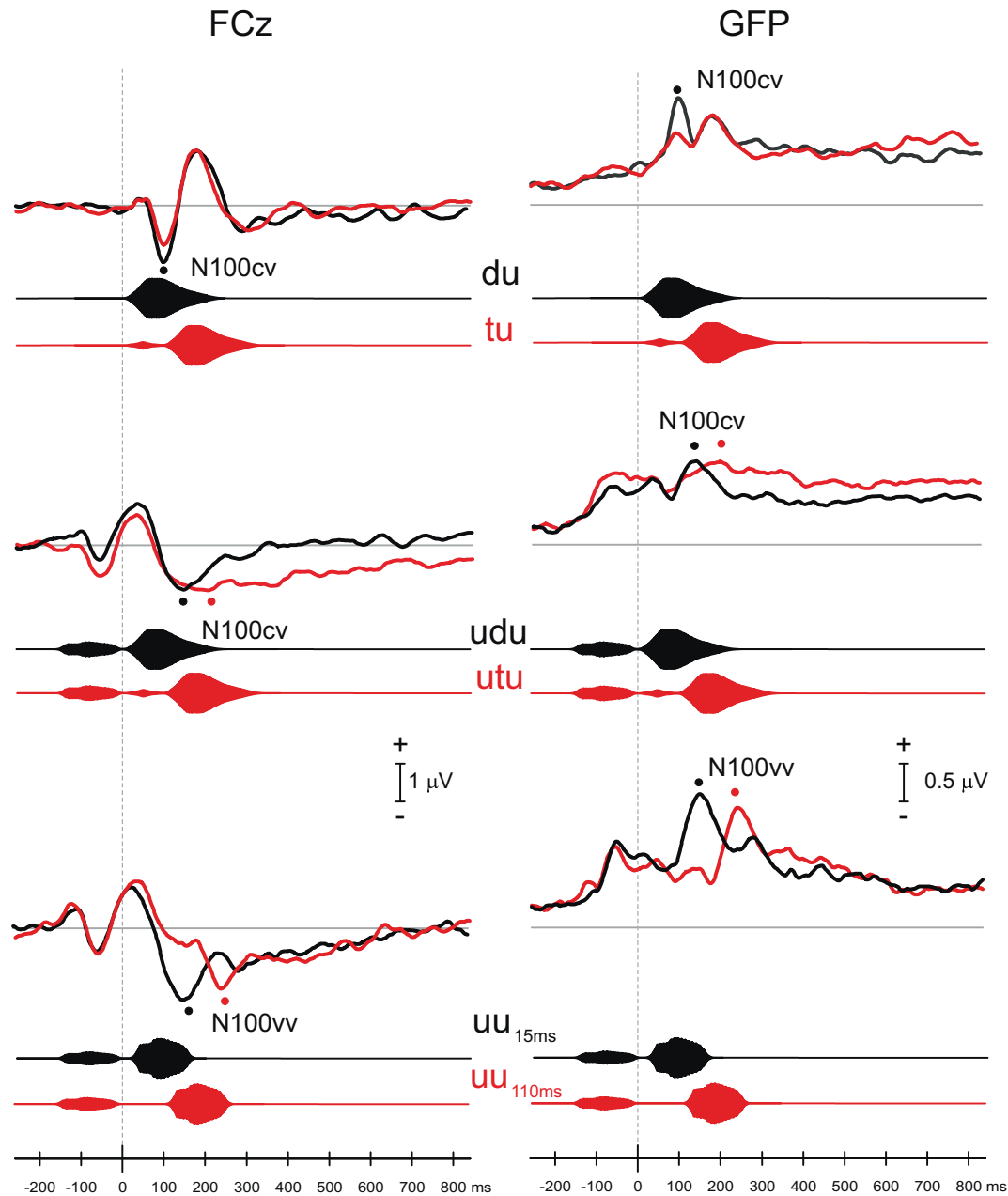
Fig. 3. Auditory cortical activities and GFP to CVs (top row) and VCVs (second row) as a function of VOTs (black traces are for 15 ms perceived as/d/;red traces are for 110 ms perceived as /t/). Note that for onset responses (top traces) the amplitude of N100 is larger for /d/ than /t/ (FCz and GFP) with no differences in P200. For VCV change responses (middle traces), N100s are separated by 45 ms, significantly less than the VOT differences. The SN that follows is of larger amplitude and longer duration for/utu/than/udu/. The bottom traces show activities to paired vowels/uu/ with time separations of 15 and 110 ms showing two distinct N100 components that are separated by 92 ms (close to the 95 ms separation of the two vowels). The SN that follows is of comparable amplitude for the two VV stimuli. Grand average waveforms are based on: /du/ and /tu/ $n = 12$;/ udu/ and /utu/ $n = 11$; /u/,/uu_{15ms}/ and /uu_{110ms}/$n = 6$.

1968). In contrast, CVs in second position (/udu/ and /utu/) evoked cortical activities consisting of N100 followed by a slow negativity (SN) as in "acoustic change potentials" to changes of pitch or intensity of ongoing tones (Dimitrijevic et al., 2008; Pratt et al., 2009); (2) Consonant VOT duration had different effects on "onset" and "change potentials". For CVs in initial position (e.g., /du/ and /tu/) N100 amplitudes decreased as VOT increased whereas N100 latency was not affected. For CVs in second position (e.g., /udu/ and /utu/) N100 latency increased as VOT increased, N100 amplitudes were not affected, while SN amplitudes increased with VOT; (3) These patterns of "onset" and "change" evoked potentials were also found to vowels presented alone (/u/) or in pairs (/uu/),

separated by silent intervals. N100 latencies to paired vowels accurately reflected the time separation between the two vowels.

4.1. Consonants in initial position

We found N100 amplitude to decrease with increasing VOT using natural speech without changes of latency similar to other reports (Horev et al., 2007; Tremblay et al., 2003b). In contrast, other studies using synthetic speech described N100 latency to become delayed (Tremblay et al., 2003a; Frye et al., 2007) or double peaked as VOT increased (Elangovan and Stuart, 2011; Hoonhorst et al., 2009; Sharma et al., 2000). The effect of VOT on N100 latency
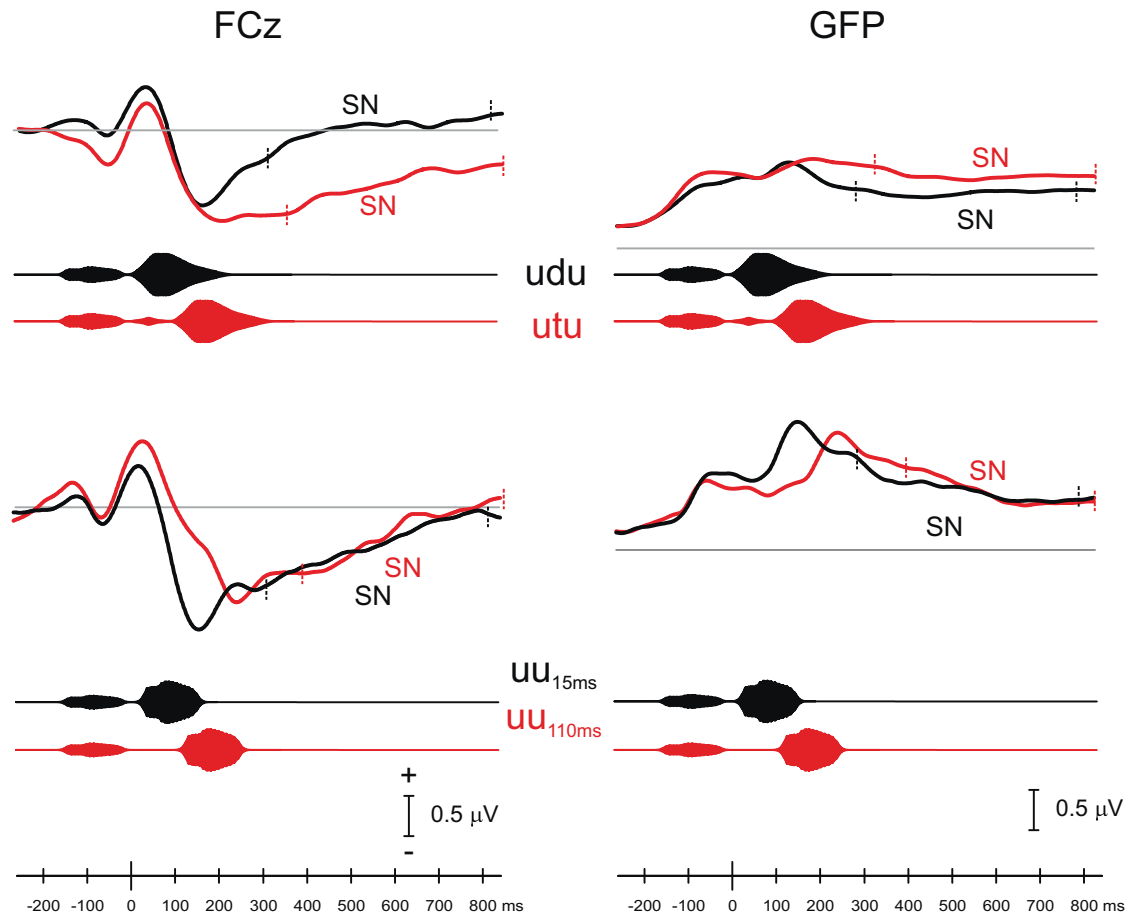
Fig. 4. Low pass filtered (5 Hz) waveforms used in the quantification of SN. SN time range was defined as 150 ms after the second N100 to 500 ms after (ranges shown in vertical dashed lines). Grand average waveforms are based on: /udu/ and /utu/ $n = 11$; /u/,/uu$_{15ms}$/ and /uu$_{110ms}$/$n = 6$.

was not accompanied by differences of perception of stop consonants as a function of VOT suggesting that auditory cortical processes contributing to perception are relatively indifferent to N100 latency differences (Sharma et al., 2000). The N100 amplitude differences that we observed between /du/ and /tu/ are consistent with previous results showing that N100 amplitude is inversely related to VOTs in both synthetic and natural speech (Toscano et al., 2010; Tremblay et al., 2003b). The cortical responses to CVs are similar to those obtained with tone bursts (Onishi and Davis, 1968; and Skinner and Jones, 1968). These studies showed that N100 latency is directly related to onset rise time of tone bursts, being longer as rise time increased. N100 amplitude was related to the peak amplitude of the tone burst. For CVs in the present study, the initial rise time of the consonant is abrupt and similar for /du/ and /tu/ resulting in similar N100 latencies. N100 amplitude to tones is directly related to the peak amplitude of the stimulus during the initial 30 ms of the stimulus. For CVs with short VOTs (/du/) the higher amplitude portions of the vowel are included within the initial 30 ms of the stimulus and are therefore associated with higher amplitudes of N100. In contrast for CVs with long VOTs (/tu/) the higher amplitude portion of the vowel occurs beyond the initial 30 ms of the CV resulting in N100 amplitudes being significantly larger to /du/ than /tu/. Aside from the temporal stimulus features, the spectral features of the stimulus will also affect the magnitude of the response. Picton et al. (1978), Dimitrijevic et al. (2008) and Pratt et al. (2009) using non-speech sounds demonstrated that lower frequencies elicit larger responses compared to higher frequencies. In this study, the

/du/ stimulus is lower in frequency than /tu/ given that voicing begins earlier. In summary, we believe that the processing of CVs is governed by the physical properties of the stimulus that include both rise time and spectral content. It remains unclear why different cortical evoked potential patterns are seen with synthetic versus natural speech. Similar onset rise times are seen for both types of stimuli, yet prolonged N100's (or double peaks) are seen with longer VOTs in synthetic speech (Elangovan and Stuart, 2011; Frye et al., 2007; Hoonhorst et al., 2009; Tremblay et al., 2003a; Sharma et al., 2000) and no latency changes with natural speech (current study and Tremblay et al., 2003b). Tremblay et al., 2003b suggested that these differences could be attributed to stimulus characteristic differences. Perhaps accentuated formant transitions in synthetic CVs result in a spectral change N100 occurring later compared to natural speech.

4.2. Consonants in second position

When the consonant was in second position (VCV), the N100 "change response" varied with VOT differently than the N100 "onset response" (CV). N100 latency to VCV was longer to /utu/ than /udu/ whereas the respective amplitudes did not differ. In contrast, the SN accompanying VCVs was affected by VOT, being larger to /utu/ compared to /udu/. The effects of VOTs on N100 and SN to CVs and VCVs are similar to those seen on the N100 and the SN to different magnitudes of change of tonal stimuli (Dimitrijevic et al., 2008, 2009). The SN identified in the present study to VCV is distinguished from the sustained potential appearing during
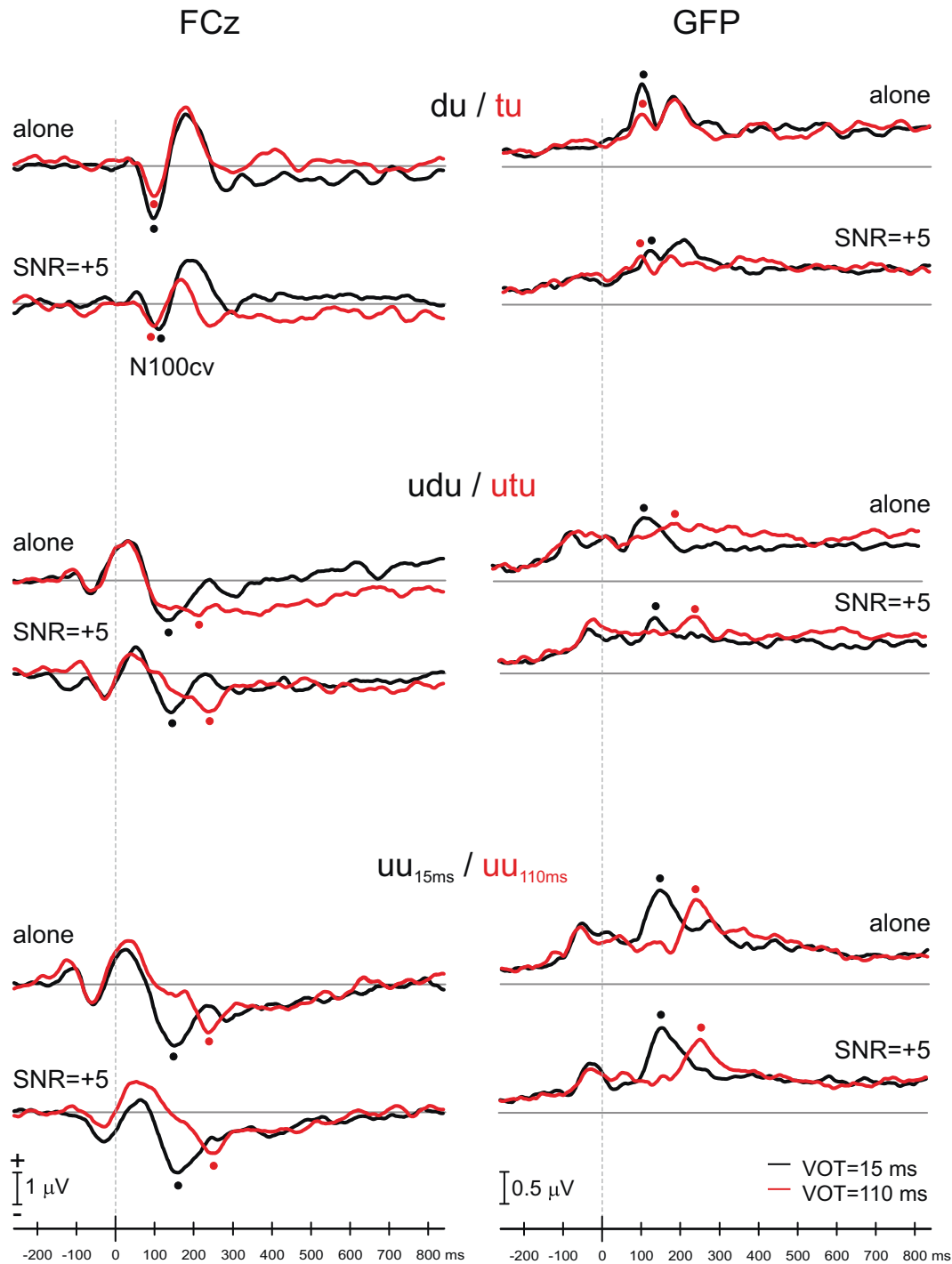
Fig. 5. Effects of noise on N100 and SN activities defined for CVs, VCVs, and VVs. For CVs (top traces), note that the amplitude difference for N100 to /du/ versus /tu/ is maintained but attenuated by noise masking. For VCVs (middle traces) the SN amplitude differences in quiet are lost with noise masking. Moreover, the N100 to 110 ms VOT (/utu/) that is small in quiet becomes more apparent with noise and N100 latency difference between/udu/ and /utu/ (93 ms) increased to correspond to the respective VOT difference of 95 ms. With noise masking, /udu/-/utu/ waveforms resemble /uu/ waveforms in quite. For paired vowels (bottom traces) noise masking was accompanied by small latency delays of N100 and attenuation of SN. Grand average waveforms are based on 6 subjects all of whom participated in both the alone and masking conditions.

continuous tones (Picton et al., 1978). The SN to VCVs (/utu/) persists beyond the VCV while the negativity to sustained tones terminates when the tone ends. The SN to VCV in the present study is similar to a negative potential that follows brief (100 ms) changes of pitch or intensity of continuous tones, a "negative change potential" (Dimitrijevic et al., 2009, 2008; Pratt et al., 2009) that was unrelated to the magnitude of the spectral or intensity change. However, the amplitude of the SN found in the present study to

VCVs is related to VOT duration of the consonant. Therefore these results suggest that the SN may relate to the timing of acoustic change responses and not necessarily the magnitude of the change. Further study is needed to define stimulus features that evoke a negative potential accompanying spectral or intensity changes of tones and of VCVs.

Both forward masking by the initial vowel and/or adaptation may contribute to these changes. Intracellular recordings of audi-

Table 3
Latency (SD) and mean amplitude (SD) of cortical potentials recorded at FCz and Global Field Power to VCVs, VVs, CVs, and Vs in noise (SNR = +5 dB).

Condition		N100$_{CV}$		SN
	SNR = +5	Latency (ms)	Amplitude (µV)	Amplitude (µV)
FCz	Onset – /du$_{15\ ms}$/	109 (14)	−0.97 (0.90)	–
	Onset – /tu$_{110\ ms}$/	106 (10)	−0.61 (0.76)	–
	Onset – /u$_{15\ ms}$/	152 (20)	−1.07 (0.75)	−0.33 (0.34)
	Change – /udu$_{15\ ms}$/	245 (15)	−1.21 (0.64)	−0.40 (0.57)
	Change – /utu$_{110\ ms}$/	143 (10)	−1.10 (0.65)	–
	Change – /uu$_{15\ ms}$/	159 (16)	−1.78 (1.02)	−0.30 (0.39)
	Change – /uu$_{110\ ms}$/	250 (16)	−1.19 (0.70)	−0.19 (0.42)
GFP	Onset – /du$_{15\ ms}$/	107 (12)	1.03 (0.32)	–
	Onset – /tu$_{110\ ms}$/	102 (13)	0.82 (0.41)	–
	Onset – /u$_{15\ ms}$/	151 (21)	0.95 (0.50)	0.50 (0.18)
	Change – /udu$_{15\ ms}$/	232 (32)	1.07 (0.25)	0.68 (0.25)
	Change – /utu$_{110\ ms}$/	144 (7)	1.15 (0.54)	–
	Change – /uu$_{15\ ms}$/	158 (12)	1.46 (0.79)	0.54 (0.21)
	Change – /uu$_{110\ ms}$/	254 (17)	1.26 (0.60)	0.52 (0.14)
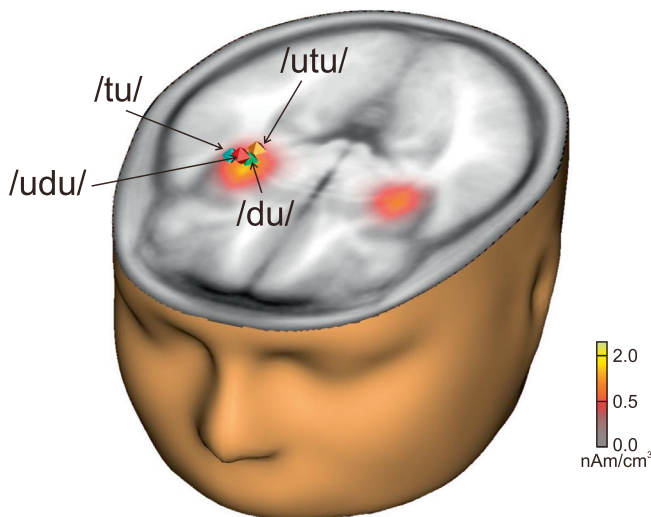
Fig. 6. CLARA and associated right-sided dipoles for onset and change N100 responses. Note the similar source locations for all N100 sources suggesting that CVs and VCVs likely arise from similar areas of cortex. Grand average based on 6 subjects that had 85% or higher dipole goodness of fits.

tory cortical neurons have shown that excitation by a brief acoustic stimulus is accompanied by forward masking that affects excitability of auditory neurons particularly those in auditory cortex (Alves-Pinto et al., 2010; Wehr and Zador, 2005). Moreover, cortical excitability changes accompanying sequential acoustic stimuli differ as a function of the neuron's sensitivity to acoustic transients and sustained sounds (Zheng and Escabí, 2008). In humans, the N100 of auditory cortical onset responses decreases in amplitude and increases in latency as inter-stimulus intervals shorten (ISI, Davis and Zerlin, 1966). However, paradoxically, with ISIs of less than 500 ms, an enhancement of the N100 is seen peaking near 200 ms ISI (Loveless et al., 1989). In this experiment, the "ISI" (i.e., initial vowel onset to CV or second V onset) was 165 ms and therefore in the range for enhancement and not suppression. Clearly the auditory system is able to generate N100 responses in close succession as evidenced by the two N100s in the VCV and VV stimuli. Moreover, the differences in N100 latency to the second "u" for /uu$_{15ms}$/ and /uu$_{110ms}$/ closely matched the acoustic separation. Additionally, the amplitudes of the N100 to the second "u" for /uu$_{15ms}$/ and /uu$_{110ms}$/ were similar. Therefore these factors suggest that decreases in neural excitability plays a minor role. The N100 latency differences between VCVs with 15 and 110 ms VOTs

suggest a greater contribution of the vowel following the consonant (/ud**u**/ and /ut**u**/) as the peak of the vowel is shifted by 85 ms (see Fig. 1). This would suggest that the N100 latency to the CVs in initial position is dominated by the onset (regardless of VOT) whereas the N100 to second position is dominated by the "acoustic change". Since the N100 latency difference between /udu/ and utu/ is 45 ms (less than the acoustic difference of 95 ms) there is an additional "consonant influence" albeit reduced.

The amplitude of the N100 for VCVs showed no differences between /udu/ and /utu/ even though there were differences when presented in isolation (i.e., /du/ and /tu/). Both the initial vowel and following CV (or vowel) will elicit N100/P200 waveforms which will overlap and interact. The identification of such overlaps would be possible using the Adjar technique (Woldorf, 1993), which was not possible in the present study given the fixed vowel and consonant-vowel separation. Another explanation for an absence of an N100$_{CV}$ amplitude difference between /udu/ and /utu/ could be a baseline shift arising from the previous stimulus SN. Examination of Fig. 3 shows that the initial vowel N100 is more negative for /utu/ than /udu/ reflecting the persistence of the SN shift to the previous stimulus that is larger to /utu/ than /udu/. Therefore, given the relatively short interstimulus time in our experimental paradigm, we were unable to correct for the baseline amplitude differences between the initial vowels of /udu/ and /utu/, we suggest that the use of long interstimulus times would resolve this uncertainty.

4.3. Effects of Noise masking

In response to VCVs, N100$_{CV}$ latency differences between /udu/ and /utu/ persisted and increased after masking to correspond to the 95 ms difference in VOT. However, the SN differences that were apparent in quiet were no longer significantly different between /udu/ and /utu/. This finding may bear on the observation that the classification of stop consonants in the presence of noise is affected to a greater extent when in second than first positions (Woods David et al., 2010). The absence of an SN difference is likely related to masking by the noise of /t/ in /utu/ leaving only the vowel component as evidenced by the 93 ms N100$_{CV}$ latency difference between /udu/ and /utu/ (see Table 3) under noise. This difference corresponds to both the /uu$_{15ms}$/ and /uu$_{110ms}$/ difference of 92 ms (alone) and 91 ms (with masking).

4.4. Generators of N100$_{CV}$ and SN

The scalp distribution of N100$_{CV}$ is compatible with a predominant generator in the right temporal lobe, and source current density and equivalent dipole preliminary estimates on grand-averaged waveforms confirmed this putative location (Figs. 6 and 7). Although this right hemisphere laterality may be attributed to stimulation of the contralateral left ear, earlier studies have indicated that hemispheric prominence can vary widely with speech elements and subject groups. Cortical processing of temporal cues of speech (VOT) has been reported to have left hemisphere prominence and this lateralization was not found with non-speech stimuli (Liegeois-Chauvel et al., 1999; Papanicolaou et al., 2003; Shtyrov et al., 2000). These findings have been confirmed by depth recordings from primary auditory cortex and from depth and scalp recordings from the same subjects (Trébuchon-Da Fonseca et al., 2005; Zaehle et al., 2007). In contrast, developmental dyslexic subjects display reversed or inconsistent asymmetry, depending on the severity of reading impairments (Giraud et al., 2008). However, whereas rapid temporal features (20–50 Hz) such as VOT lateralize to the left hemisphere, strong right-hemisphere dominance has been shown for coding the speech envelope, which represents syllable patterns (Abrams et al., 2008). Only careful control of the
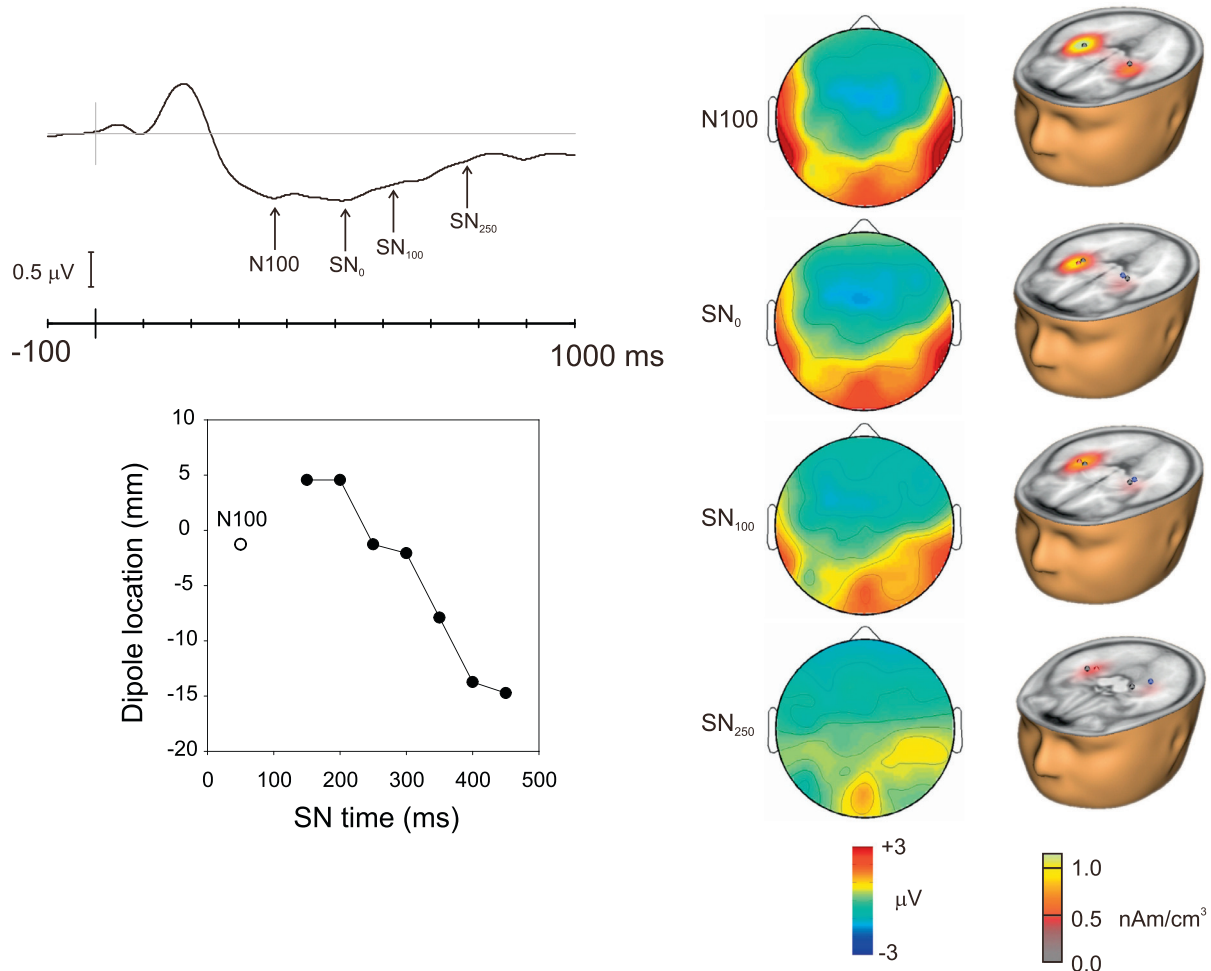
Fig. 7. Source and scalp topographies of the SN and N100 are shown for /utu/. The top left panel illustrates the grand averaged time waveform for /utu/. The arrows indicate time points for which the scalp topography and CLARA (right panel) were derived. The far right column shows the N100 (grey dipole) and successive CLARA and dipole estimates for the SN. The beginning of the SN (SN$_0$) was defined as 150 ms after the N100. The bottom left panel summarizes the changes in the right temporal lobe dipole location changes (positive to negative indicates anterior to posterior direction). Units are given in mm Cartesian coordinates. Grand average based on 6 subjects that had 85% or higher dipole goodness of fits.

specific acoustic features in the complex of VCV utterances can help resolve these apparently contradictory findings.

Earlier studies (Dimitrijevic et al., 2008; Pratt et al., 2009) on the sources of the N100 component localized them to the right temporal lobe in response to acoustic change, similar to this study, whereas to low-frequency change the source was on the left (Pratt et al., 2009). Thus, assuming homology of N100$_{CV}$ and the N100 component to frequency change, N100$_{CV}$ is evoked by the high-frequency changes associated with the consonant.

The similarity of the source estimates of onset and change responses suggest that both types of responses are mediated by similar cortical regions. However, the different effects of stimulus conditions on SN suggest that the scalp distributions of SN are different under different stimulus conditions, i.e., the relative contributions of multiple generators vary with stimulus type and as a function of time. The scalp distribution maps show the greatest degree of polarity reversal on the right. This distribution could be due to a larger source on the right superior temporal surface or due to a left source with some radially oriented dipole activity such that the lateral negativity cancels the polarity inversion from the left tangential dipole. Source current density analysis indicated that the main sources of SN are in the right temporal lobe, moving initially anteriorly and then posteriorly as a function of time: Initially, the area in the vicinity of the primary auditory cortex is involved, mov-

ing slightly anteriorly after 50–100 ms and then progressively and substantially more posteriorly with time (Fig. 7, for /utu/). The functional significance of the change in SN source distribution is difficult to interpret. If the distribution of SN were constant over time and simply diminished in amplitude, the changes would be consistent with a prolonged duration N100. This was clearly not the case and the SN likely represents a separate component distinct from the N100 change response that is maximally activated by long duration VOTs.

4.5. Conclusions

The results of this study show that cortical activity related to VOT of the consonant changed as a function of consonant position in the utterance. When the consonant is in initial position (CV), an onset N100 response is observed whose amplitude, but not latency, varied inversely with VOT. The N100 latency is similar between /du/ and /tu/ reflecting that the stimulus rise time of the consonant is rapid and the same for both. The larger N100 amplitude seen for /du/ compared to /tu/ is likely related to differences of the timing of peak amplitude of the vowel in CV, being earlier for /du/ compared to /tu/. When the consonant is in second position preceded by a vowel (VCV) an acoustic change complex is observed in which the N100 latency does vary with VOT suggesting that stimulus factors

influencing N100 latency differ for consonants in initial (CV) or second position (VCV). A slow negativity (SN) appeared only to VCVs and its amplitude varied directly with VOT being larger to /utu/ than /udu/. When noise masking was present, the N100 latency difference between 15 and 85 ms VOTs was equal to the vowel onset difference suggesting a further reduced consonant influence and more dominance by the vowel. The quantification of auditory cortical activity in normal subjects in this paper may provide objective measures to define brain events accompanying abnormal consonant processing in patients with impaired speech processing.

Overall, our results suggest that multisyllabic stimuli evoke different patterns of neural activity compared to monosyllabic stimuli. Such multisyllabic stimuli are frequent in everyday listening situations in which speech is typically continuous rather than being separate short duration utterances. The differences in processing mono- and bi-syllabic words are in line with improved consonant identification of CVs when in initial position (CVs) than when in second position (CVCs) that was reported in previous studies by Woods and colleagues (2010).

Acknowledgements

References

Abrams DA, Nicol T, Zecker S, Kraus N. Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. J Neurosci 2008;28:3958–65.

Alves-Pinto A, Baudoux S, Palmer AR, Sumner CJ. Forward masking estimated by signal detection theory analysis of neuronal responses in primary auditory cortex. J Assoc Res Otolaryngol 2010;11:477–94.

Davis H, Zerlin S. Acoustic relations of the human vertex potential. J Acoust Soc Am 1966;39:109–16.

Dimitrijevic A, Lolli B, Zeng FG, Starr A. Intensity changes in a continuous tone: auditory cortical potentials comparison with frequency changes. Clin Neurophysiol 2009;120:374–83.

Dimitrijevic A, Zeng FG, Pratt H, Starr A. Frequency changes in a continuous tone: auditory cortical potentials. Clin Neurophysiol 2008;119:2111–24.

Eggermont JJ, Ponton CW. The neurophysiology of auditory perception: from single units to evoked potentials. Audiol Neurootol 2002;7:71–99.

Elangovan S, Stuart A. A cross-linguistic examination of cortical auditory evoked potentials for a categorical voicing contrast. Neurosci Lett 2011;490:140–4.

Frye RE, Fisher JM, Coty A, Zarella M, Liederman J, Halgren E. Linear coding of voice onset time. J Cogn Neurosci 2007;19:1476–87.

Giraud K, Trébuchon-DaFonseca A, Démonet JF, Habib M, Liégeois-Chauvel C. Asymmetry of voice onset time-processing in adult developmental dyslexics. Clin Neurophysiol 2008;119:1652–63.

Hari R, Pelizzone M, Makela JP, Huttunen J, Knuutila J. Neuromagnetic reponses from a deaf subject to stimuli presented through a multichannel cochlear prosthesis. Ear Hear 1988;9:148–52.

Hoonhorst I, Serniclaes W, Collet G, Colin C, Markessis E, Radeau M, et al. N1b and Na subcomponents of the N100 long latency auditory evoked-potential: neurophysiological correlates of voicing in French-speaking subjects. Clin Neurophysiol 2009;120:897–903.

Horev N, Most T, Pratt H. Categorical perception of speech (VOT) and analogous non-speech (FOT) signals: behavioral and electrophysiological correlates. Ear Hear 2007;28:111–28.

Ille N, Berg P, Scherg M. Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. J Clin Neurophysiol 2002;19:113–24.

Jones SJ, Perez N. The auditory "C-process": analyzing the spectral envelope of complex sounds. Clin Neurophysiol 2001;112:965–75.

Laufer A. Voicing in contemporary Hebrew in comparison with other languages. Hebrew Studies 1998;39:143–79.

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;9:484–96.

Lisker L, Abramson AS. A cross-language study of voicing in initial stops: acoustical measurements. Word 1964;20:384–422.

Loveless N, Hari R, Hämäläinen M, Tiihonen J. Evoked responses of human auditory cortex may be enhanced by preceding stimuli. Electroencephalogr Clin Neurophysiol 1989;74:217–27.

Marlin S, Feldmann D, Nguyen Y, Rouillon I, Loundon N, Jonard L, et al. Temperature-sensitive auditory neuropathy associated with an otoferlin mutation: deafening fever! Biochem Biophys Res Commun 2010;394:737–42.

Martin BA, Boothroyd A. Cortical, auditory, event-related potentials in response to periodic and aperiodic stimuli with the same spectral envelope. Ear Hear 1999;20:33–44.

Martin BA, Boothroyd A. Cortical, auditory, evoked potentials in response to changes of spectrum and amplitude. J Acoust Soc Am 2000;107:2155–61.

McCullagh P, Nedler JA. General Linear Models. London: Chapman & Hall; 1983.

Onishi S, Davis H. Effects of duration and rise time of tone bursts on evoked V potentials. J Acoust Soc Am 1968;44:582–91.

Papanicolaou AC, Castillo E, Breier JI, Davis RN, Simos PG, Diehl RL. Differential brain activation patterns during perception of voice and tone onset time series: a MEG study. Neuroimage 2003;18:448–59.

Picton TW, Woods DL, Proulx GB. Human auditory sustained potentials. II. Stimulus relationships. Electroencephalogr Clin Neurophysiol 1978;45:198–210.

Pratt H, Starr A, Michalewski HJ, Dimitrijevic A, Bleich N, Mittelman N. Auditory-evoked potentials to frequency increase and decrease of high- and low-frequency tones. Clin Neurophysiol 2009;120:360–73.

Rance G, Fava R, Baldock H, Chong A, Barker E, Corben L, et al. Speech perception ability in individuals with Friedreich ataxia. Brain 2008;131:2002–12.

Sharma A, Dorman MF. Cortical auditory evoked potential correlates of categorical perception of voice-onset time. J Acoust Soc Am 1999;106:1078–83.

Sharma A, Dorman MF. Neurophysiologic correlates of cross-language phonetic perception. J Acoust Soc Am 2000;107:2697–703.

Sharma A, Marsh CM, Dorman MF. Relationship between N1 evoked potential morphology and the perception of voicing. J Acoust Soc Am 2000;108:3030–5.

Shtyrov Y, Kujala T, Palva S, Ilmoniemi RJ, Naatanen R. Discrimination of speech and of complex nonspeech sounds of different temporal structure in the left and right cerebral hemispheres. Neuroimage 2000;12:657–63.

Skinner PH, Jones HC. Effects of signal duration and rise time on the auditory evoked potential. J Speech Hear Res. 1968;11(2):301–6.

Starr A, Michalewski HJ, Zeng FG, Fujikawa-Brooks S, Linthicum F, Kim CS, et al. Pathology and physiology of auditory neuropathy with a novel mutation in the MPZ gene (Tyr145->Ser). Brain 2003;126:1604–19.

Starr A, Picton TW, Sininger Y, Hood LJ, Berlin CI. Auditory neuropathy. Brain 1996;119:741–53.

Toscano JC, McMurray B, Dennhardt J, Luck SJ. Continuous perception and graded categorization: electrophysiological evidence for a linear relationship between the acoustic signal and perceptual encoding of speech. Psychol Sci 2010;21:1532–40.

Tremblay KL, Piskosz M, Souza P. Effects of age and age-related hearing loss on the neural representation of speech cues. Clin Neurophysiol 2003a;114:1332–43.

Tremblay KL, Friesen L, Martin BA, Wright R. Test-retest reliability of cortical evoked potentials using naturally produced speech sounds. Ear Hear 2003b;24:225–32.

Trébuchon-Da Fonseca A, Giraud Kimberly, Badier J-M, Chauvel Patrick, Liégeois-Chauvel Catherine. Hemispheric lateralization of voice onset time (VOT) comparison between depth and scalp EEG recordings. Neuroimage 2005;27:1–14.

Varga R, Avenarius MR, Kelley PM, Keats BJ, Berlin CI, Hood LJ, et al. OTOF mutations revealed by genetic analysis of hearing loss families including a potential temperature sensitive auditory neuropathy allele. J Med Genet 2006;43:576–81.

Wehr M, Zador AM. Synaptic mechanisms of forward suppression in rat auditory cortex. Neuron 2005;47:437–45.

Williams L. The perception of stop consonant voicing by Spanish-English bilinguals. Atten Percept Psychophys 1977;19:285–9.

Woldorff MG. Distortion of ERP averages due to overlap from temporally adjacent ERPs: analysis and correction. Psychophysiology 1993;30:98–119.

Woods David L, Yund EW, Herron TJ, Ua Cruadhlaoich MAI. Consonant identification in consonant-vowel-consonant syllables in speech-spectrum noise. J Acoust Soc Am 2010;127:1609–23.

Zaehle T, Jancke L, Meyer M. Electrical brain imaging evidences left auditory cortex involvement in speech and non-speech discrimination based on temporal features. Behav Brain Funct 2007;3:63.

Zeng F-G, Kong Y-Y, Michalewski Henry J, Starr Arnold. Perceptual consequences of disrupted auditory nerve activity. J Neurophysiol 2005;93:3050–63.

Zheng Y, Escabí MA. Distinct roles for onset and sustained activity in the neuronal code for temporal periodicity and acoustic envelope shape. J Neurosci 2008;28:14230–44.