Effects of Consonant-Vowel Transitions in Speech Stimuli on Cortical Auditory Evoked Potentials in Adults

Michael Doellinger*[,a], Martin Burger[a], Ulrich Hoppe[b] and Ulrich Eysholdt[a]

[a]*Department of Phoniatrics and Pediatric Audiology, University Hospital Erlangen, Bohlenplatz 21, 91054 Erlangen, Germany*

[b]*ENT Department, University Hospital Erlangen, Waldstraße 1, 91054 Erlangen, Germany*

Abstract: We examined the neural activation to consonant-vowel transitions by cortical auditory evoked potentials (AEPs). The aim was to show whether cortical response patterns to speech stimuli contain components due to one of the temporal features, the voice-onset time (VOT). In seven normal-hearing adults, the cortical responses to four different monosyllabic words were opposed to the cortical responses to noise stimuli with the same temporal envelope as the speech stimuli. Significant hemispheric asymmetries were found for speech but not in noise evoked potentials. The difference signals between the AEPs to speech and corresponding noise stimuli revealed a significant negative component, which correlated with the VOT. The hemispheric asymmetries can be referred to rapid spectral changes. The correlation with the VOT indicates that the significant component in the difference signal reflects the perception of the acoustic change within the consonant-vowel transition. Thus, at the level of automatic processing, the characteristics of speech evoked potentials appear to be determined primarily by temporal aspects of the eliciting stimuli.

Keywords: Auditory evoked potentials, voice-onset time, speech stimuli, N1/P2 complex.

1. INTRODUCTION

During the auditory perception of speech, the central hearing has to process complex acoustic structures in real-time. Adequate processing of acoustic information is the requirement for speech and language development [1]. Particularly, disorders in the temporal processing are supposed to be responsible for deficient auditory discrimination abilities, which can result in an impaired speech perception [2]. Cortical auditory evoked potentials (AEPs) provide an objective method with high temporal resolution for the investigation of auditory speech processing.

In the absence of attention, AEPs represent the automatic cortical processing on auditory sensory level. In this case, AEPs are dominated by obligatory components. In adults, these components are particularly the N1 and P2, also called N1/P2 complex [3]. The N1/P2 complex is thought to reflect the synchronous neural activity of structures in the thalamic-cortical segment of the central auditory system. The N1, a negative peak occurring approximately 100 ms after stimulus onset, is suggested to represent sound detection functions [4], since it is sensitive to onset sound features, such as intensity and interstimulus interval. For example, it could be demonstrated that the N1 indicates segmentation during the perception of continuous speech [5]. P2, a positive peak occurring approximately 200 ms after stimulus onset, is assumed to reflect sound content properties like acoustic or phonetic structures [6].

Several investigators have examined the possibility of using AEPs to determine the individual discrimination ability of phonetic structures in speech sounds [7-10]. Of particular interest are stimulus features constituting minimal pairs, such as formant gradients of consonant vocal-transitions and the voice onset-time (VOT). VOT is defined as the interval between the release from stop closure and the onset of laryngeal pulsing [11]. Sharma and colleagues (1999) investigated the morphology of speech evoked potentials in absence of attention depending on the subjective perception of stop consonants. Analyzing the cortical responses to a /da/-/ta/ continuum they found that stimuli with short VOTs (0–30 ms) evoked a single N1, whereas stimuli with long VOTs (50–80 ms) evoked two distinct negative components (N1' and N1). This discontinuity in AEP morphology was in line with the subjective perception of voiceless sounds and thus, it was suggested to represent an electrophysiological correlate of categorical perception [7]. However, in another study using a /ga/-/ka/ continuum as stimuli, the change in N1 morphology to a double-peaked component did not signal subjective perception of a voiceless sound. This result indicated that the minimum VOT value of 40 ms required for the temporal separation of N1 depends on acoustic properties of the stimulus rather than the perceptual categorization [8].

Martin *et al.* investigated cortical responses to syllables with different formant transitions embedded in masking noise [12]. The stimuli evoked an N1 which systematically changed with the stimulus energy (which is a sound onset feature), but not with the individual subjective discrimination.

Ostroff *et al.* compared the cortical response to the syllable /sei/ with the cortical responses to the sibilant /s/ and to

*Address correspondence to this author at the Department of Phoniatrics and Pediatric Audiology, University Hospital Erlangen, Bohlenplatz 21, 91054 Erlangen, Germany; Tel: +49 9131 8533814; Fax: +49 9131 8532687; E-mail: michael.doellinger@uk-erlangen.de

the vowel /ei/. They ascertained that the response to /sei/ was a combination of the AEPs to the onsets of the two constituent phonemes /s/ and /ei/ [13]. These overlapping AEPs within one stimulus response have been termed the acoustic change complex (ACC) [14]. The ACC is supposed to be composed of different N1/P2 complexes reflecting the acoustic changes across the entire stimulus [15].

Overall, the aforementioned studies indicate that speech evoked potentials primarily reflect the phonetic composition of the stimuli. The aim of the presented study was to investigate whether cortical responses to speech stimuli contain AEP components referring to the VOT. For this purpose, AEPs to four natural monosyllabic words were compared with AEPs to noise stimuli showing the same temporal envelope as the word stimuli [16]. For a systematic analysis of the cortical responses, the speech stimuli were chosen with respect to varying VOTs.

2. MATERIALS AND METHODOLOGY

2.1. Participants

We examined 7 normal-hearing, right-handed monolingual native German speakers (3 female, 4 male). The age range was 22 to 27 years. None was on medication at that time. All participants signed an informed consent.

2.2. Stimuli and Procedure

In this study, two types of stimuli were used to elicit electrophysiologic responses: speech stimuli and noise sound stimuli. The speech stimuli consisted of four monosyllabic words naturally produced by a male speaker, taken from Freiburger speech discrimination test [17]. The stimuli were *Ei* /aːi/, *Bett* /bɛt/, *Dieb* /diːb/, and *Pult* /pult/, which mean egg, bed, thief and desk. The durations of the stimuli *Ei, Bett, Dieb*, and *Pult* were 544 ms, 430 ms, 501 ms, and 567 ms, respectively. The speech stimuli exhibited a bandwidth of 8 kHz and were digitalized with 14 bit at a sampling rate of 20,000 s^{-1}. Except *Ei*, all speech stimuli started with an initial stop consonant and were chosen with respect to their varying VOT. The VOTs were determined as VOT(*Ei*) = 0 ms, VOT(*Bett*) = 35 ms, VOT(*Dieb*) = 60 ms, and VOT(*Pult*) = 80 ms. The onset times of the closing consonants which succeed the vowels were determined as 215 ms with *Bett,* 390 ms with *Dieb*, and 395 ms with *Pult*.

The speech stimuli served as raw material for the synthesis of the noise stimuli. Noise stimuli were created by randomly multiplying the discrete samples of the speech stimuli with ± 1 and subsequent 8 kHz low pass filtering. Fig. (**1**) shows exemplarily the generation of the noise stimulus corresponding to *Dieb*. Thus, any phonetic information was removed from the speech stimuli, whereas the intensity was held constant. By this means, the synthesized noise stimuli Ei_N, $Bett_N$, $Dieb_N$, and $Pult_N$ were generated.

Spectrograms of all presented stimuli are depicted in Fig. (**2**). The speech stimuli mainly consisted of frequencies lower than 2 kHz, whereas the noise stimuli were distributed over the entire bandwidth of 8 kHz.

For equated presentation loudness, the applied stimuli were calibrated with respect to their root mean squares to a 1 kHz reference tone with an intensity level of 70 dB SPL (see [18]). Stimuli were presented in four blocks of 320 sounds each. Within the block, all stimuli were successively presented 40 times in the order of speech and corresponding noise stimulus (speech-noise pairs). Thus, each stimulus was presented 160 times. The interstimulus interval varied randomly between 1400 and 2100 ms. Subjects were tested in a soundproof booth. In order to minimize effects of the subjects' state of attention on the applied stimuli, they were asked to watch a silent movie presented on a TV screen.

2.3. EEG Recording and Data Processing

The EEG was derived with Ag/AgCl electrodes, which were integrated in a cap (Braincap, Brain Products GmbH, Gilching, Germany) with 30 fixed electrode positions (Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, P7, P8, Fz, FCz, Cz, Pz, FC1, FC2, CP1, CP2, FC5, FC6, CP5, CP6, TP9, TP10). The electrode impedances were kept below 5 kOhm. For eye artefact rejection, an electro-oculogram was recorded by an electrode placed under the right eye. The EEG was recorded with Brain Amp-MR (Brain Products GmbH, Gilching, Germany) at a sampling rate of 500 Hz and digitalization of 16 bit.

During data acquisition, all channels were referenced to FCz. Offline, data were re-referenced to the mean of TP9 and TP10. The EEG was 0.13 Hz − 20 Hz band pass filtered with a slope of 12 dB/octave. The recording window included 100 ms prestimulus and 700 ms poststimulus time.
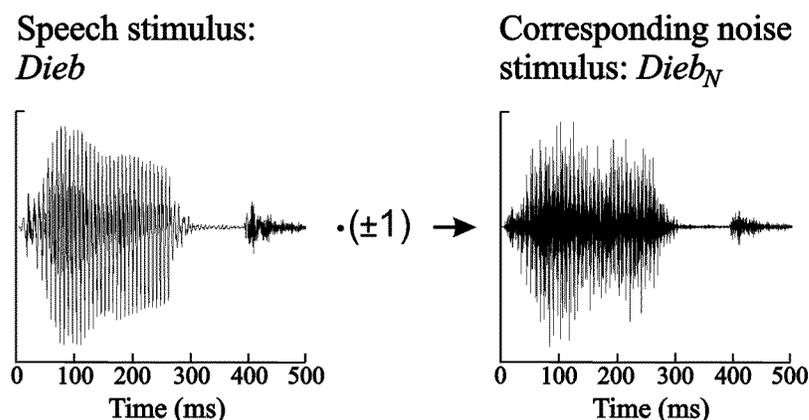
Speech stimulus: *Dieb*

Corresponding noise stimulus: $Dieb_N$

Fig. (1). Speech stimulus Dieb and synthesized noise stimulus Dieb$_N$. Both stimuli exhibit the same temporal envelope but phonetic information is discarded in the noise stimulus.
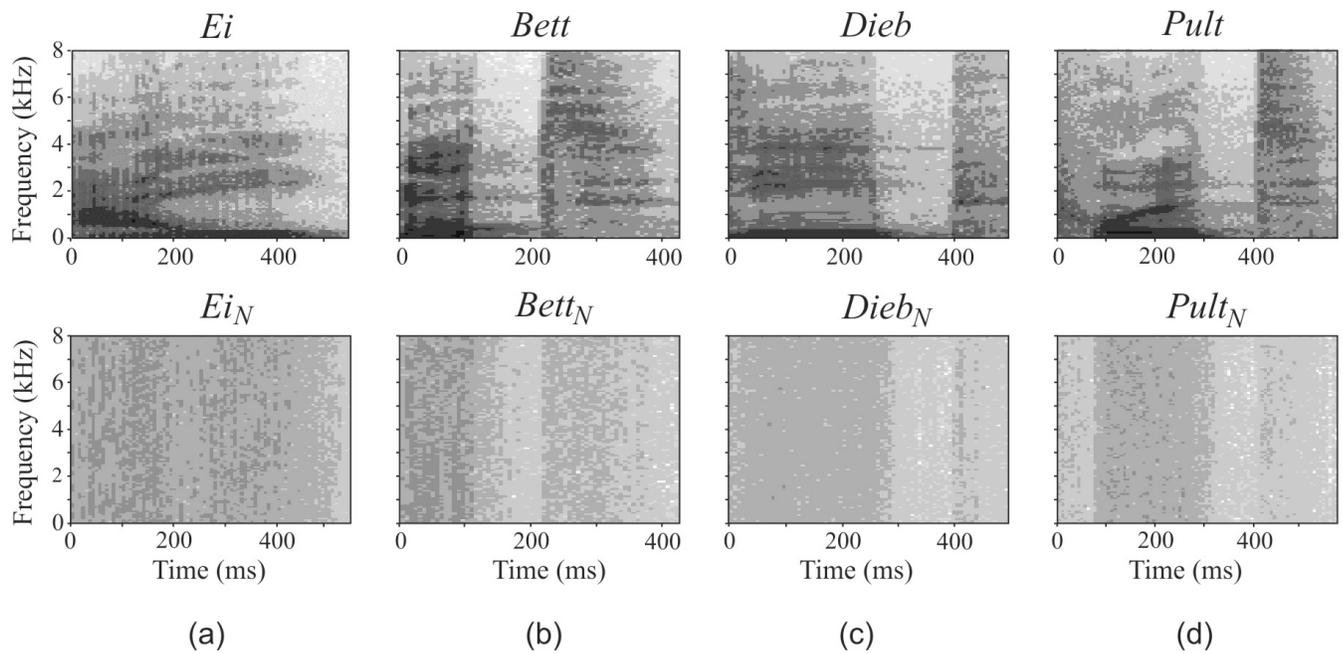
Fig. (2). Spectral view of the presented stimuli. Top: Natural spoken monosyllabic words. Bottom: Synthetic noise sounds, derived from the monosyllabic words. Bandwidth of the stimuli is 8 kHz.

Sweeps with artefacts measuring higher than 75 µV were rejected. The remaining sweeps were averaged separately for each stimulus and prestimulus baselined.

For each stimulus, grand mean waveforms were composed and grand mean latencies of the AEP components were visually identified. The individual peak latencies were determined in a time interval of the central latencies ± 24 ms. Amplitudes were calculated as a mean voltage at the 20 ms period centered at the individual peak latencies.

2.4. Data Analysis

In the first step, statistical comparative tests of the AEP characteristics were carried out for both stimulus types, i.e. speech and noise stimuli, separately. In the second step, cortical responses to speech and corresponding noise sounds were analysed and compared among each other.

2.4.1. Responses to Speech and Noise Stimuli

Preliminary one-way analyses of variance (ANOVA) were conducted to determine the scalp distribution of the amplitudes of N1 and P2. Since the largest responses for all stimuli were obtained at central positions, channel Cz was chosen for following one-way ANOVAs, which examined N1 and P2 latencies and amplitudes of the distinct speech and noise elicited AEPs. Newman-Keuls tests were carried out as post hoc analyses for the determination of equal-mean subsets. For the examination of hemispheric differences, N1-P2 interpeak amplitudes of the lateral scalp sites (T3, FC5, C3 vs. T4, FC6, C4) were compared *via* paired t-tests.

2.4.2. Speech-Noise Pairs

For each subject time intervals were determined, in which speech and corresponding noise responses statistically differed. For this purpose, the sampling values of the single recorded sweeps in response to each stimulus were pooled for every discrete time step independently, resulting in 350 samples of speech and noise responses each (700 ms post-stimulus interval x 500 s⁻¹ EEG sampling rate). Speech response samples of every time step were compared to the corresponding noise response sample by Student's two-tailed t-test. As level of significance $p < 0.01$ was selected. In order to avoid Bonferroni problems due to the large number of compared samples, a statistical difference between speech and noise response was accepted, when at least 12 consecutive sampling points (i.e., 24 ms interval) showed significance [19].

On average level, difference signals were created by subtracting the noise elicited AEPs from the corresponding speech elicited AEPs. Latencies and amplitudes of occurring components in the difference signals were determined visually. Amplitudes were tested for significant appearance *via* one-tailed t-test.

3. RESULTS

3.1. Responses to Speech and Noise Stimuli

Preliminary one-way ANOVAs, which separately analysed N1 and P2 peak amplitudes, revealed significant amplitude differences between the electrode sites for both, speech and noise stimuli ($p < 0.001$, $F_{26,162} = 3.3 - 8.6$). Newman-Keuls post hoc analysis revealed that with all stimuli, the greatest amplitudes were elicited in the central area of the scalp (i.e., channels Cz3, Cz, Cz4) for both, N1 and P2 amplitude. Hence, the following illustrations and statistics are restricted to channel Cz. Fig. (**3a**) shows the grand mean waveforms of the cortical responses to the four different speech stimuli. Morphologies of the speech-evoked potentials differed clearly across the stimuli. AEPs in response to *Bett* and *Pult* appear to reveal multiple overlapping response patterns. In all subjects, the predominant components were N1 and P2 in all responses. All speech stimuli elicited an N1 component which peaked around 113 ± 8 ms, and a P2 com-
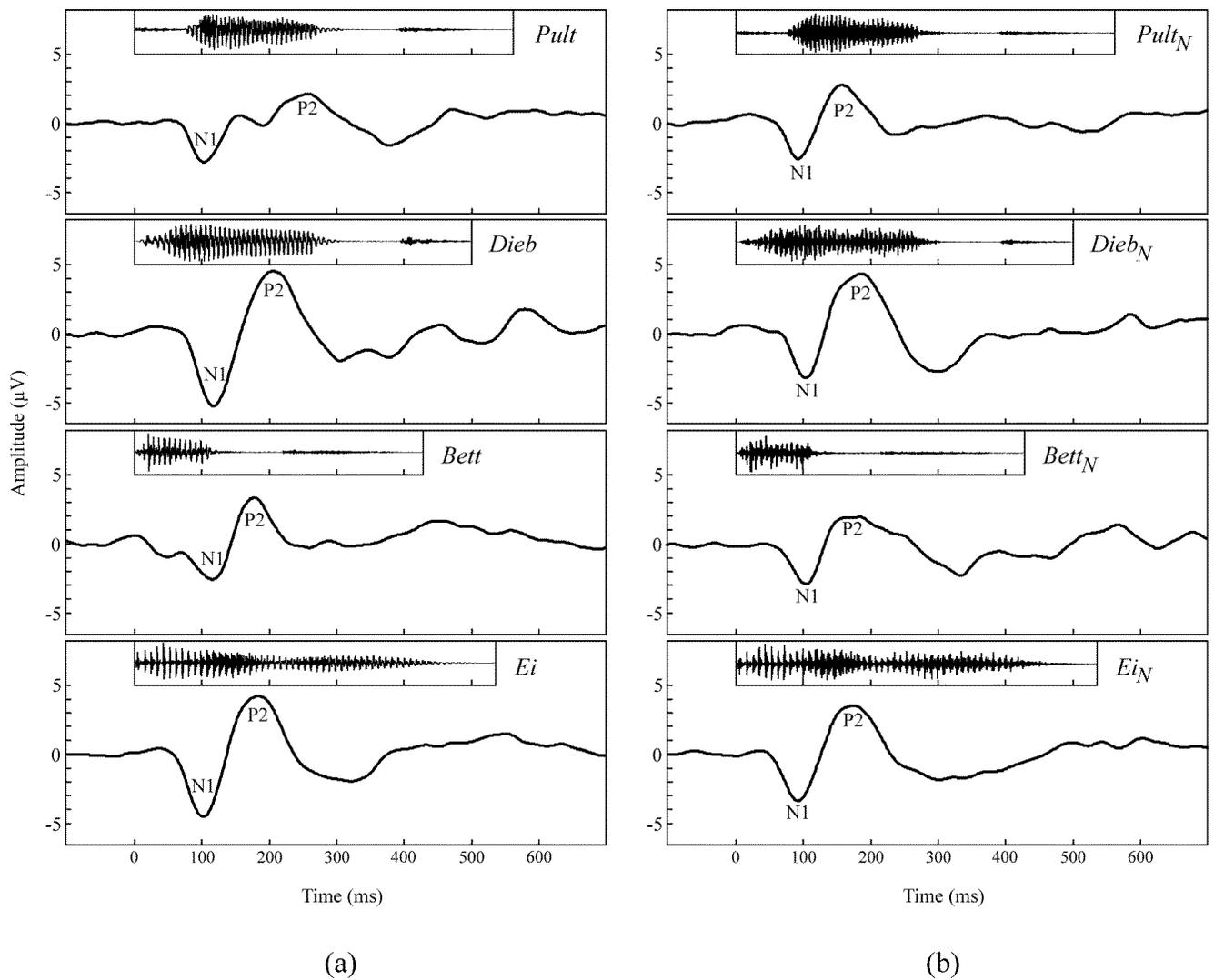
(a)   (b)

Fig. (3). (**a**) Grand mean waveforms of speech stimuli elicited AEPs in response to Ei, Bett, Dieb, and Pult at channel Cz. (**b**) Grand mean waveforms of noise-sound elicited AEPs in response to Ei_N, $Bett_N$, $Dieb_N$, and $Pult_N$ at channel Cz.

ponent that peaked in a range from 180 to 260 ms around 207 ± 36 ms.

Fig. (**3b**) shows the grand mean waveforms of the cortical responses to the four different noise stimuli. The waveforms of the different noise sound elicited AEPs were similar. The predominant components in all responses were N1 and P2. All noise stimuli elicited an N1 component which peaked around 103 ± 8 ms, and a P2 component that peaked around 175 ± 11 ms.

In Tables **1** and **2**, N1 and P2 latencies and amplitudes of both stimulus types are given. Table **3** shows F values of

one-way ANOVAs, conducted to reveal significant differences of these characteristics. Newman-Keuls post hoc tests did not show consistent correlations within a subset of both stimulus types. As it is apparent from the higher levels of significance, speech evoked AEP characteristics diverged stronger among each other than those evoked by noise sounds.

For the analysis of hemispheric asymmetries, a paired t-test was conducted with N1-P2 interpeak amplitudes of opposite channel pairs T3 – T4, FC5 – FC6, and C3 – C4, continuing from lateral to central. The mean N1-P2 interpeak

Table 1.   Mean Latencies of N1 and P2 in ms Elicited by the Speech/Noise Stimuli at Channel Cz

Peak	Stimulus Type	*Ei*	*Bett*	*Dieb*	*Pult*
N1	Speech	106 (4)	119 (8)	120 (5)	107 (7)
	Noise	95 (8)	111 (7)	108 (7)	97 (7)
P2	Speech	183 (12)	181 (8)	206 (12)	259 (10)
	Noise	177 (15)	181 (16)	185 (10)	160 (12)

Standard deviations are in parentheses.

Table 2.　Mean Amplitudes of N1 and P2 in µV Elicited by the Speech/Noise Stimuli at Channel Cz

Peak	Stimulus Type	*Ei*	*Bett*	*Dieb*	*Pult*
N1	Speech	-3.82 (1.29)	-2.36 (1.04)	-4.45 (1.54)	-2.36 (1.36)
	Noise	-3.01 (1.58)	-2.28 (1.48)	-2.58 (1.33)	-2.03 (1.04)
P2	Speech	4.00 (1.29)	2.75 (1.18)	4.40 (1.50)	1.99 (0.57)
	Noise	3.22 (1.01)	2.35 (0.85)	4.11 (1.46)	2.49 (0.90)

Standard deviations are in parentheses.

Table 3.　ANOVAs F Values of the Different AEP Characteristics

	F value, N1		F value, P2	
	Latency	**Amplitude**	**Latency**	**Amplitude**
Speech stimuli	9.9***	4.5*	84.9***	5.9**
Noise stimuli	8.4**	0.7[n.s.]	4.7**	3.9*

[n.s.] not significant, *$p<0.05$, **$p<0.01$, ***$p<0.001$

Table 4.　N1-P2 Interpeak Amplitudes (in µV) and t Values for the Comparison of Selected Electrode Pairs (T3-T4, FC5-FC6, C3-C4)

	T3	T4	*t*	FC5	FC6	*t*	C3	C4	*t*
Speech	2.29(1.18)	3.23(1.15)	-5.7***	4.80(1.13)	4.80(1.11)	0.0[n.s.]	5.19(1.35)	4.98(1.26)	2.9**
Noise	2.15(1.14)	2.24(1.29)	-0.5[n.s.]	3.89(1.09)	3.83(1.20)	0.7[n.s.]	3.98(1.21)	3.95(1.32)	0.5[n.s.]

Standard deviations are in parentheses. [n.s.] not significant, *$p<0.05$, **$p<0.01$, ***$p<0.001$.

amplitude of T3 was significantly smaller than that of T4 for the speech stimuli but not for the noise stimuli, see *t* values in Table **4**. While N1-P2 interpeak amplitudes at the positions FC5 and FC6 did not differ, an inversion from right to left hemispheric predominance occurred at the central positions C3, C4 for the speech stimuli (C3 > C4, see Table **4**). Exemplarily, cortical responses to *Dieb* and *Dieb*$_N$ across the entire scalp are depicted in Fig. (**4**), in order to illustrate the hemispheric asymmetries of the speech evoked potentials.

3.2. Speech-Noise Pairs

The AEP waveforms, in response to the speech sounds, were compared to those in response to the corresponding noise sounds. The time intervals, in which speech and corresponding noise responses significantly differed, are depicted in Fig. (**5**). Blocks of black colour mark temporal areas of significant differences ($p<0.01$) for each subject and each speech-noise pair. None of the speech-noise pair responses differs strongly during the initial 50 ms, as one can see from Figs. (**5a-5d**). The highest accumulations of areas of significant differences occur in the time window of 100 – 300 ms after stimulus onset.

Fig. (**6**) shows the grand mean AEP waveforms of the speech (dashed lines) and noise stimuli (dotted lines) as well as the difference signals (solid lines) speech – noise. In the difference signal, a negative component could be observed in a latency range of approximately 140 ± 15 ms. For each speech-noise pair, one-tailed t-tests (subjects x channels) showed that the occurring negative component was significantly smaller than zero. Latencies and amplitudes of this negative component are given in Table **5**. *Ei* elicits the smallest latency, latencies in response to *Bett* and *Dieb* are

similar and *Pult* elicits the longest latency. Fig. (**7**) shows the VOT of the four speech stimuli plotted against the peak latency of the negative component in the difference signal. A prominent correlation of VOT and peak latency is observable. Spearman rank correlation analysis revealed that the latencies were significant positively correlated with VOT ($r = 0.66$, $p < 0.01$).

4. DISCUSSION

Generally, different natural speech stimuli evoke distinct neural response patterns [20]. The aim of this study was to show whether cortical response patterns to speech stimuli contain components due to one of the temporal features, the voice-onset time. For this purpose, AEP characteristics in response to monosyllabic words were compared against each other and with cortical responses to noise stimuli with the same temporal envelope as the speech stimuli.

4.1. Responses to Speech and Noise Stimuli

In accordance with other studies (e.g. [6, 20, 21]), for both types of stimuli, the AEPs were maximal at central electrode sites. The different natural speech stimuli evoked distinct neural response patterns, partially showing multiple overlapping responses. These distinctions indicate an extended neural activity with the auditory processing of speech in contrast to noise sound processing. Considering previous studies, a plausible explanation are the variant VOTs of the presented speech stimuli which affect the N1 and P2 [8, 13, 22, 23].

Winkler *et al*. demonstrated that N1 latencies evoked by spectral-pitch and missing-fundamental tones do not differ [24]. Hence, the different pitches of the presented vowels are
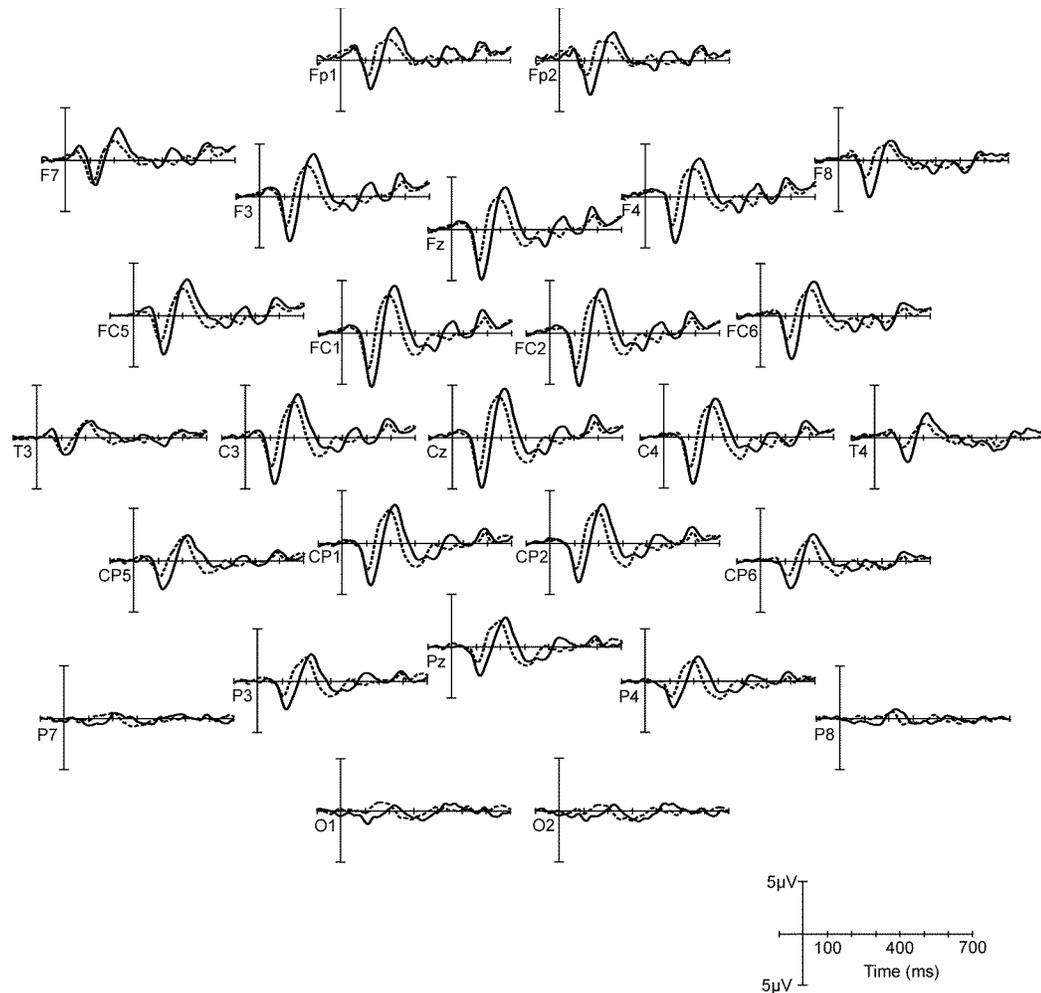
Fig. (4). Grand mean AEPs to Dieb (solid) and Dieb$_N$ (dashed) over the entire scalp.
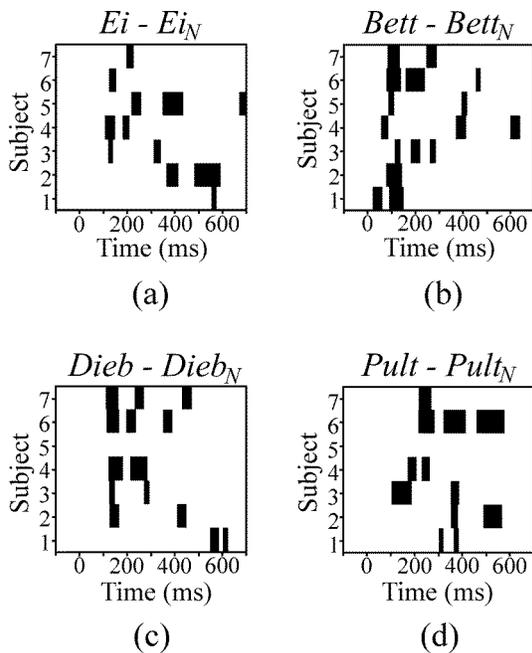

Fig. (5). Temporal areas of significant differences between the responses to speech and noise stimuli at channel Cz for each subject (black areas). Level of significance is p<0.01.

probably not the reason for the found variability in the speech evoked potentials, at least not regarding the N1. An influence of the cortical response to the closing consonant on the P2 is not expectable due to its relatively late onset time.

4.2. Hemispheric Asymmetries

In contrast to noise elicited AEPs, hemispheric differences occurred with speech evoked AEPs in the present study. Hemispheric left-overbalanced asymmetries with speech perception are well known [25-27]. Left-overbalance is primarily reported in studies with attentive designs. However, focused auditory attention can selectively modulate automatic processing in auditory cortex and thus, affect AEPs [28].

Dehaene-Lambertz *et al.* demonstrated in a discrimination task that the left-hemispheric predominance with phoneme perception also can be achieved with appropriate non-speech stimuli [29]. Thus, specialization of left auditory cortex is not speech specific but depends on rapid spectro-temporal changes. However, in the present study, a right-hemispheric predominance was observable at temporal sites which inverted to a left-overbalance at central sites. This could be due to deviant locations in the activated auditory areas of both sides, resulting in different dipole orientations which underlie AEP derivations [30]. A more plausible
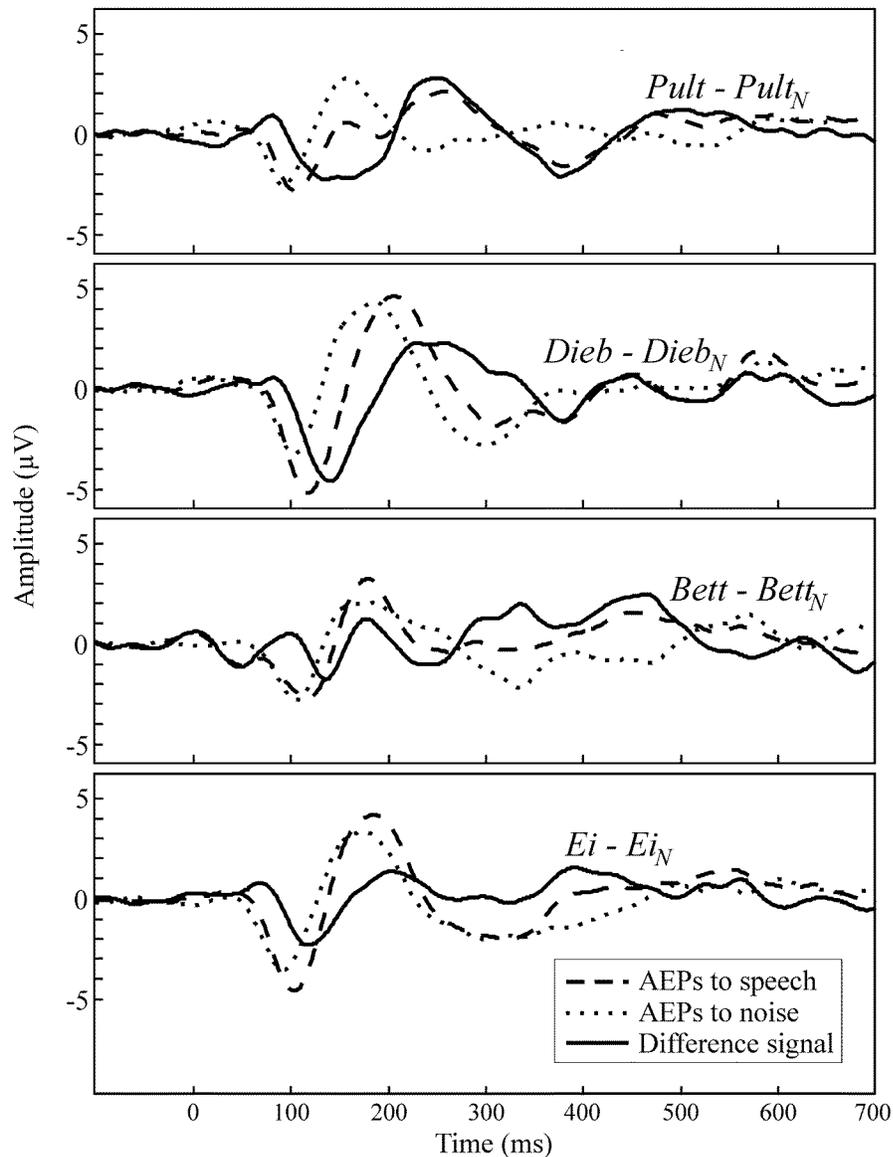
Fig. (6). Grand mean waveforms at channel Cz of speech evoked responses (dashed), noise evoked responses (dotted) and the difference signal (solid).

Table 5.   Mean Latencies and Amplitudes of the Negative Component in the Difference Signal out of the Speech-Noise Pairs Averaged over all Channels

Stimulus Type	*Ei*	*Bett*	*Dieb*	*Pult*
Latency (ms)	123 (12)	138 (12)	142 (11)	160 (17)
Amplitude (μV)	-1.36 (1.08)***	-0.78 (1.12)***	-2.64 (1.44)***	-1.81 (1.11)***

Standard deviations are in parentheses. Results of one-tailed *t* tests ***$p < 0.001$.

explanation is a functional asymmetry of the auditory cortices. PET findings indicated complementary specializations of left and right auditory cortex' belt areas [31]. Responses to temporal features were weighted towards the left, while responses to spectral features were weighted towards the right hemisphere.

Thus, the spectral changes of the consonant-vowel transitions may be the reason for the observed right hemispheric overbalance. In contrast, noise stimuli did not evoke asymmetric responses.

4.3. Speech-Noise Pairs

Although phonetic changes were present in the speech stimuli during the entire propagation of approximately 500 ms, AEPs of speech-noise pairs differed predominantly in the time window of the N1/P2 complex between 100 and 300 ms (Fig. **5**), reflecting distinct onset processing.

Subtracting noise from speech elicited AEPs facilitated a direct comparison of the speech-noise pairs. Thus, the presented paradigm allows observing the neural processing of
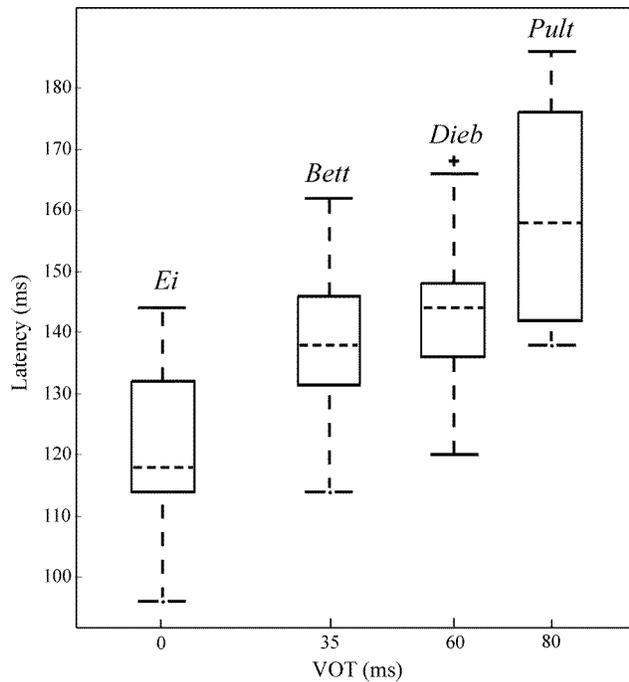
Fig. (7). Box and whisker plot of VOTs of the diverse speech stimuli against the latencies of the negative components in the difference signal of the cortical responses. Dashed lines within the boxes display the median, the edges of the boxes display the quartiles. Whiskers indicate maximal and minimal latencies unless outliers '+' occur (whiskers span 1.5 times the inter-quartile range).

spectral-acoustic and phonetic information, since differences in stimulus duration and amplitude are eliminated. The comparison revealed a significant negative component in the difference signal that seems to correlate with the VOT of the presented stimulus (see Steinschneider *et al.* who found that synchronized activity in the auditory cortex is time locked with consonant release and voicing onset [32]). Since the speech evoked N1 and P2 appear somewhat later and heightened in contrast to the noise evoked, the present study suggests that the observed negative component denotes a second N1/P2 complex which is merged with the onset response. Therefore it is not directly visible. This second N1/P2 complex is suggested to reflect the onset of the vowel as an acoustic event.

This study demonstrated that rapid spectral changes in the perception of speech can be assessed electrophysiologically by means of a time-saving design. In case of modified cortical responses in subjects with central auditory processing disorders, this paradigm might serve as diagnostic tool for phonetic discrimination tasks since it reveals the neural processing of time-critical speech structures.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Gallagher T. Treatment research in speech, language and swallowing: lessons from child language disorders. Folia Phoniatr 1998; 50: 165-82.

[2]   Tallal P, Miller SL, Bedi G, Byma G, *et al.* Language comprehension in language-learning, impaired children improved with acoustically modified speech. Science 1996; 271: 81-3.

[3]   Wolpaw JR, Penry JK. A temporal component of the auditory evoked response. Electroencephalogr Clin Neurophysiol 1975; 39, 609-920.

[4]   Näätänen R, Picton T. The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. Psychophysiology 1987; 24: 375-425.

[5]   Sanders LD, Newport EL, Neville HJ. Segmenting nonsense: An event related potential index of perceived onsets in continuous speech. Nat Neurosci 2002; 5: 700-3.

[6]   Ceponiene R, Alku P, Westerfield M, Torki M, Townsend J. ERPs differentiate syllable and nonphonetic sound processing in children and adults. Psychophysiology 2005; 42: 391-406.

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

[8]   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.

[9]   Cunningham J, Nicol T, Zecker S, Kraus N. Speech-evoked neurophysiologic responses in children with learning problems: Development and behavioural correlates of perception. Ear Hear 2000; 21: 554-68.

[10]   Khedr E, El-Nasser WA, Abdel Haleem EK, Bakr MS, Trakham MN. Evoked potentials and electroencephalography in stuttering. Folia Phoniatr 2000; 52: 178-86.

[11]   Lisker L, Abramson AS. 'The voicing dimension: some experiments in comparative phonetics.' Proceedings of the VI International Congress of Phonetic Sciences. Prague 1970.

[12]   Martin BA, Kurtzberg D, Stapells DR. The effects of decreased audibility produced by high-pass noise masking on N1 and the mismatch negativity to speech sounds /ba/ and /da/. J Speech Lang Hear Res 1999; 42: 271-86.

[13]   Ostroff JM, Martin BA, Boothroyd A. Cortical evoked response to acoustic change within a syllable. Ear Hear 1998; 19: 290-7.

[14]   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.

[15]   Tremblay KL, Kraus N, McGee T, Ponton CW, Otis B. Central auditory plasticity: Changes in the N1-P2 complex after speech-sound Training. Ear Hear 2001; 22: 79-90.

[16]   Kummer P, Burger M, Schuster M, Rosanowski F, Eysholdt U, Hoppe U. Cortical auditory evoked potentials to acoustic changes in speech stimuli in children. Folia Phoniatr 2007; 59: 273-80.

[17]   Keller F. Verschiedene Aufnahmen des Sprachverständlichkeitstest "Freiburger Test" (Different recordings of the speech intelligibility test "Freiburger Test"). Biomed Techn 1977; 22: 292-8.

[18]   ANSI S3.6 – 1996. American national standards specification for audiometers. New York: American National Standards Institute 1996.

[19]   Guthrie D, Buchwald JS. Significance testing of difference potentials. Psychophysiology 1991; 28: 240-4.

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

[21]   Ponton CW, Eggermont J, Kwong B, Dent M. Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. Clin Neurophysiol 2000; 111: 220-36.

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

[23]   Gage N, Poeppel D, Roberts T, Hickok G. Auditory evoked M100 reflects onset acoustics of speech sounds. Brain Res 1998; 814: 236-9.

[24]   Winkler I, Tervaniemi M, Näätänen R. Two separate codes for missing-fundamental pitch in the human auditory cortex. J Acoust Soc Am 1997; 102, 1072-82.

[25]   Penhune VB, Zatorre RJ, McDonald JD, Evans AC. Interhemispheric anatomical differences in human primary auditory cortex: probabilistic mapping and volume measurement from magnetic resonance scans. Cereb Cortex 1996; 6: 661-72.

[26]   Liegeois-Chauvel C, de Graaf JB, Laquitton V, Chauvel P. Specialization of left auditory cortex for speech perception in man depends on temporal coding. Cereb Cortex 1999; 9: 484-96.

[27] 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: 1524-34.

[28] Woldorff MG, Gallen CC, Hampson SA, *et al.* Modulation of early sensory processing in human auditory cortex during auditory selective attention. Proc Natl Acad Sci USA 1993; 90: 8722-26.

[29] Dehaene-Lambertz G, Pallier C, Serniclaes W, Sprenger-Charolles L, Jobert A, Deheane S. Neural correlates of switching from auditory to speech perception. Neuroimage 2005; 24: 21-33.

[30] Rosburg T, Kreitschmann-Andermahr I, Emmerich E, Nowak H, Sauer H. Hemispheric differences in frequency dependent dipole orientation of the human auditory evoked field component N100m. Neurosci lett 1998; 258: 105-8.

[31] Zatorre RJ, Belin P. Spectral and temporal processing in human auditory cortex: Cereb Cortex 2001; 11: 946-53.

[32] Steinschneider M, Volkov IO, Noh MD, Garell PC, Howard MA. Temporal encoding of the voice onset time phonetic parameter by field potentials recorded directly from human auditory cortex. J Neurophysiol 1999; 82: 2346-57.