



Published in final edited form as:

Ear Hear. 2011 February ; 32(1): 53–60. doi:10.1097/AUD.0b013e3181ec5c46.

CORTICAL ENCODING OF SIGNALS IN NOISE: EFFECTS OF STIMULUS TYPE AND RECORDING PARADIGM

Curtis J. Billings, Keri O. Bennett, Michelle R. Molis, and Marjorie R. Leek

National Center for Rehabilitative Auditory Research, Portland VA Medical Center, 3710 SW US Veterans Hospital Road, 97239, Portland, OR, USA

Abstract

Objectives—Perception-in-noise deficits have been demonstrated across many populations and listening conditions. Many factors contribute to successful perception of auditory stimuli in noise, including neural encoding in the central auditory system. Physiological measures such as cortical auditory evoked potentials can provide a view of neural encoding at the level of the cortex that may inform our understanding of listeners' abilities to perceive signals in the presence of background noise. In order to understand signal-in-noise neural encoding better, we set out to determine the effect of signal type, noise type, and evoking paradigm on the P1-N1-P2 complex.

Design—Tones and speech stimuli were presented to nine individuals in quiet, and in three background noise types: continuous speech spectrum noise, interrupted speech spectrum noise, and four-talker babble at a signal-to-noise ratio of -3 dB. In separate sessions, cortical auditory evoked potentials were evoked by a passive homogenous paradigm (single repeating stimulus) and an active oddball paradigm.

Results—The results for the N1 component indicated significant effects of signal type, noise type, and evoking paradigm. While components P1 and P2 also had significant main effects of these variables, only P2 demonstrated significant interactions among these variables.

Conclusions—Signal type, noise type, and evoking paradigm all must be carefully considered when interpreting signal-in-noise evoked potentials. Furthermore, these data confirm the possible usefulness of CAEPs as an aid to understanding perception-in-noise deficits.

Keywords

Cortical auditory evoked potentials (CAEPs); Event-related potentials (ERPs); signals in noise; background noise; N1; N100

1. Introduction

Successful communication in difficult listening environments is dependent upon how well the auditory system is able to extract signals of interest from other competing acoustic information. Perception-in-noise difficulties have been studied extensively in the behavioral domain; however, physiological data are more limited. Some animal data indicate that the

*Corresponding Author: Curtis J. Billings, Ph.D., National Center for Rehabilitative Auditory Research, Portland VA Medical Center, 3710 SW U.S. Veterans Hospital Rd (NCRAR), Portland, OR 97239 USA, Phone: 00-1-503-220-8262 ext.54574, Fax: 503-402-2824, curtis.billings2@va.gov.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

central auditory system, especially the cortex, is an important contributor to signal-in-noise encoding (Phillips, 1985, 1986, 1990; Phillips & Kelly, 1992). A measure of human cortical encoding would be beneficial in showing how the cortex processes signals in noise and may help explain performance variability across individuals. One approach to studying how signal-to-noise ratio (SNR) is encoded in the human central auditory system is to use cortical auditory evoked potentials (CAEPs), which can provide valuable information about the temporal encoding of large populations of cortical neurons recorded at the scalp. Traditionally, CAEP waveform morphology has been described in terms of amplitude and latency; amplitude refers to the strength of the response (measured in microvolts) and latency refers to the time of the response after stimulus onset (measured in milliseconds). These voltage changes over time result from post-synaptic potentials within the brain and are influenced by the number of recruited neurons, extent of neuronal activation, and synchrony of the neural responses (for a review, see Eggermont, 2007).

The CAEP consists of several waves, including the P1-N1-P2 complex. The P1-N1-P2 complex is generated by structures in the thalamo-cortical segment of the central auditory system, and is sensitive to the acoustics of the evoking stimuli (Vaughan & Ritter, 1970; Wolpaw & Penry, 1975; Näätänen & Picton, 1987; Hyde, 1997). The P1-N1-P2 is particularly useful with speech-in-noise processing because it can be evoked using a variety of speech stimuli (for a review, see Martin et al., 2008). Furthermore, abnormal P1-N1-P2 responses have been associated with subject populations who experience perception difficulties, such as elderly people and those with hearing impairment (Kraus et al., 2000; Oates et al., 2002; Tremblay et al., 2003, 2004). The P1-N1-P2 complex is sensitive to stimulus characteristics that contribute to temporal processing abilities (e.g., Papanicolaou et al., 1984; Ostroff et al., 2003; Michalewski et al., 2005; Tremblay et al., 2003), and temporal processing abilities have been implicated in the extraction of signals from noise (e.g., Stuart & Phillips, 1996; Bacon et al., 1998). However, for CAEPs to be useful, it must be determined what variables are driving cortical signal-in-noise encoding. Relatively few studies have systematically examined signal-in-noise encoding using the P1-N1-P2 complex (Whiting et al., 1998; Martin et al., 2005; Androulidakis et al., 2006; Kaplan-Neeman et al., 2006; Hiraumi et al., 2008; Billings et al., 2009; Michalewski et al., 2009;). These studies indicate that characteristics of both the signal and the noise must be considered in the interpretation of this response.

The neural encoding of signals in noise has been studied using primarily broadband white noise as the background masker. No studies that we are aware of recorded CAEPs in speech backgrounds, and only two recorded the P1-N1-P2 complex to signals presented in modulated background noise (Androulidakis & Jones, 2006; Hiraumi et al., 2008). In both cases, normal-hearing individuals participated and pure tones served as the signal. These studies reported that tones in modulated noise evoked more robust responses than those evoked by tones in continuous noise. Presumably, gaps in the background noise allowed for more efficient encoding of the signal, resulting in amplitude increases and latency decreases. This agrees with behavioral perception-in-noise data which demonstrate improved perception in modulated and interrupted noise relative to continuous noise (Miller & Licklider, 1950; Festen & Plomp, 1990; Stuart & Phillips, 1995).

The interpretation of these CAEP studies is limited because only tonal signals were used. It is important to use speech as test stimuli if results are to be applied to speech-in-noise abilities. Whiting et al. (1998) and Kaplan-Neeman (2006) recorded electrophysiological responses to multiple speech stimuli. Whiting and colleagues used an active oddball paradigm in which participants were asked to identify a target stimulus presented randomly in a sequence of nontarget stimuli. While this paradigm is optimally designed to evoke the discriminative P3 response, N1 values were also analyzed and reported. However, we are

interested in the obligatory encoding that is associated with the passively elicited P1–N1–P2 complex (i.e., participants ignore a train of a repeating stimulus—homogenous paradigm). It is difficult to compare the passive homogenous paradigm results that were evoked with tones and several noise types (e.g., Androulidakis et al., 2006; Hiraumi et al., 2008; Billings et al., 2009) with active oddball paradigm results in which speech contrasts were used (Whiting et al., 1998; Kaplan-Neeman et al., 2006). The comparison is difficult because the larger P3 response evoked in an oddball paradigm task often overlaps, and greatly affects, earlier CAEP components such as the N1, P2, and N2 (Tremblay & Burkard, 2007). Oddball paradigms also introduce cognitive variables such as attention and stimulus context effects, which complicate the interpretation of earlier CAEP components. In this study we manipulated these effects by recording responses in both active and passive paradigms.

The purpose of this study was to consider three important variables that affect signal-in-noise encoding: signal type, noise type, and CAEP paradigm. Testing these variables in one experiment allowed for a better understanding of the interactions that exist among them.

2. Materials and Methods

Using a repeated measures design, participants were tested under eight stimulus conditions that included two signals and four noise conditions. These eight conditions were presented in two different recording paradigms: an active oddball presentation and a passive homogenous presentation. Amplitude and latency values for evoked responses P1, N1, and P2 were determined and analyzed for all conditions tested.

2.1. Participants

Nine young normal-hearing individuals participated in this study (mean age = 25.7 years, range = 19–31 years; 4 males and 5 females; all right-handed). Participants had normal hearing bilaterally from 250 to 8000 Hz (<20 dBHL) and normal tympanometry measures (single admittance peak between ± 50 daPa to a 226Hz tone). All participants were in good general health with no report of significant history of otologic or neurologic disorders. All participants provided informed consent and research was completed with approval from the pertinent institutional review boards.

2.2. Stimuli

Stimuli consisted of signals and noises presented to the right ear using an Etymotic ER2 insert earphone. For the active oddball paradigm, a 1000-500 Hz tone contrast or a /ba-/da/ speech contrast was used. For the passive homogenous paradigm, signals were the 1000 Hz tone or the speech syllable /ba/. The rise/fall times of all tones were 9 ms, included in total durations of 150 ms. Speech contrasts were the naturally produced female exemplars from the UCLA Nonsense Syllable Test (Dubno & Schaefer, 1992) shortened to 150 ms by windowing the steady vowel offset of the stimulus. All signals were presented at 65 dB SPL. Calibration of these signals was completed by measuring the overall RMS level of 10 s of a concatenated version of each signal. Signals were presented in four background conditions: quiet, continuous noise, interrupted noise, and four-talker babble. These three noise types were selected because they would allow two comparisons: (1) continuous vs. interrupted noise, which would reveal the physiological release from masking that one would expect based on behavioral studies, and (2) continuous vs. four-talker babble noise, which might reveal the neural correlates of informational masking that would be present in a speech on speech masking condition (i.e., babble masker of a speech syllable). For the quiet condition, no background noise was added to the signal. Figure 1 shows two-second waveforms of the noise types that were used. Spectral shaping of the continuous noise matched the long term spectrum of speech. Recordings of the 720 IEEE sentences (IEEE, 1969) produced by a

female talker were concatenated and submitted to an FFT; the phases of the spectrum frequencies were randomized and an inverse FFT was calculated, resulting in a continuous noise with the same long-term spectral average as the original speech sample (Molis & Summers, 2003). The interrupted noise had the same spectrum as the continuous noise but its temporal characteristics matched the interrupted noise used by Phillips and colleagues (Phillips et al., 1994; Rappaport et al., 1994; Stuart et al., 1995; Stuart et al., 1996; Stuart et al., 1997). Briefly, it was generated by interrupting the continuous speech spectrum noise such that periods of noise and periods of silence between them were independently randomized with durations ranging from 5 to 95 ms, reflecting the duration of important elements in speech. The random interruption used by Phillips and colleagues prevented any problems associated with the CAEP time-locking to the noise onsets/offsets, thereby evoking responses to the noise rather than to the signal. The four-talker babble consisted of two female and two male talkers reading printed prose for 10 minutes. The amplitude of each speech signal was maximized to a point just below peak clipping and the four recordings were mixed into a single channel (Lilly et al., submitted). Three versions of each noise type were used pseudo-randomly (i.e., approximately equivalent total number of uses) across subjects to reduce possible effects of amplitude fluctuations that would be present in only one version of the noise. All background noises were presented at an overall RMS level of 68 dB SPL, resulting in a signal-to-noise ratio (SNR) of -3 dB. This SNR represented an ecologically valid SNR (Plomp, 1977) and was selected based on the behavioral and evoked potential signal-in-noise literature as a level at which noise type is an important factor (Androulidakis et al., 2006; Festen & Plomp, 1990; Hiraumi et al., 2008; Miller & Licklider, 1950; Wilson & Carhart, 1969).

2.5. Electrophysiology

A PC-based system controlled the timing of stimulus presentation and delivered an external trigger to the Compumedics Neuroscan system (Charlotte, NC). Two paradigms were used to evoke P1, N1, and P2 waves¹. In the passive homogenous paradigm participants were asked to ignore the stimuli and watch a silent close-captioned movie of their choice. A single stimulus was repeated for a total of 300 trials. The /ba/ was presented in one run, and the 1000 Hz tone was presented in a separate run. In the active oddball paradigm participants were asked to identify the deviant stimulus with a button press. The oddball paradigm consisted of either a /ba-/da/ contrast or a 1000-500 Hz contrast presented in separate runs. Probability of presentation was 0.80 for standards (/ba/ and 1000 Hz) and 0.20 for deviants (/da/ and 500 Hz). Two blocks of 250 trials were presented for a total of 400 standards and 100 deviants. The 100 standards that directly followed a deviant were not included in averaging to minimize any higher order switching effects that might be a factor (Näätänen & Picton, 1987), resulting in a total of 300 possible standard trials. By design, this matched the total trials in the homogenous paradigm. The same inter-stimulus interval of 1100 ms (onset to onset) was used in both paradigms. Stimulus condition order within signal type condition (tone or speech) was randomized across subjects. Testing occurred over two days consisting of about two hours of testing each day. Homogenous and oddball paradigms for a given signal (i.e., tone or speech) were presented one day and for the other signal on the second day with signal order randomized across subjects. Participants were asked to minimize head and body movement.

Evoked potential activity was recorded using an elastic cap (Electro-Cap International, Inc) which housed 64 tin electrodes. The ground electrode was located on the forehead during CAEP collection, and Cz was the reference electrode. Data were then re-referenced offline

¹Data were also collected on the P3 evoked potential as well as behavioral target identification. These data are reported in a separate article focused on informational masking and cognitive encoding (Bennett et al., in preparation)

to an average reference. Horizontal and vertical eye movement was monitored with electrodes located inferiorly and at the outer canthi of both eyes. An 800 ms recording window was used (100 ms pre-stimulus period and a 700 ms post-stimulus period). Evoked responses were analog band-pass filtered on-line from 0 to 100 Hz (12 dB/octave roll off). All channels were amplified with a gain \times 500, and converted using an analog-to-digital sampling rate of 1000 Hz. Trials with eye-blink artifacts were corrected offline using Neuroscan software. This blink reduction procedure calculates the amount of covariation between each EEG channel and a vertical eye channel using a spatial, singular value decomposition and removes the vertical blink activity from each EEG electrode on a point-by-point basis to the degree that the EEG and blink activity covaried (Neuroscan, 2007). Following blink correction, trials containing artifacts exceeding \pm 70 microvolts were rejected from averaging. For all individuals and conditions, 60% or more of the collected trials were available for averaging after artifact rejection. Following artifact rejection, the remaining sweeps were averaged and filtered off-line from 1 Hz (high-pass filter, 24 dB/octave) to 30 Hz (low-pass filter, 12 dB/octave). While some researchers have suggested that optimal filter settings for P1 should not include frequencies below 10 Hz (e.g., Jerger et al., 1992), we used broader filter settings to facilitate comparison with portions of the existing literature (e.g., Billings et al., 2009; Roth et al., 1984).

2.6. Data Analysis and Interpretation

Waves P1, N1, and P2 were analyzed at electrode site Cz. Peak amplitudes were calculated relative to baseline, and peak latencies were calculated relative to stimulus onset (i.e., 0 ms). Latency and amplitude values of each wave were determined by agreement of two judges. Each judge used temporal electrode inversion, global field power traces, and grand averages to determine peaks for a given condition.

Three-way repeated measures analyses of variance (ANOVA) were completed on amplitude and latency measures of each component of the evoked response (P1, N1, and P2). The $2 \times 3 \times 2$ analysis included the factors of signal type (tone and speech), noise type (continuous, interrupted, and four-talker babble) and paradigm (homogenous passive and oddball). Greenhouse-Geisser corrections (Greenhouse & Geisser, 1959) were used where an assumption of sphericity was not appropriate.

3. Results

Latency and amplitude means and standard error of the means are shown in Table 1 for P1, N1 and P2 components of the CAEP. All three independent variables (i.e., signal type, noise type, and paradigm) affected the latency of the evoked response, and amplitude was affected by both noise type and paradigm, but not signal type. Specific effects are described below.

3.1. Effect of Signal Type

Repeated measures ANOVAs completed on P1, N1, and P2 indicated main effects of signal type for all latency measures [P1: $F(1,8) = 37.8$, $p < .001$; N1: $F(1,8) = 226.0$, $p < .001$; and P2: $F(1,7) = 67.1$, $p < .001$] but not for amplitude measures [P1: $F(1,8) = .17$, $p = .69$; N1: $F(1,8) = 2.0$, $p = .19$; and P2: $F(1,7) = 3.0$, $p = .12$]. Speech evoked peaks generally occurred later for all components than tone evoked peaks. Figure 2 shows grand averaged waveforms for (a.) the passive homogenous recording paradigm and (b.) the active oddball recording paradigm. In addition, P1, N1, and P2 means for both homogenous and oddball paradigms can be seen in Table 1.

3.2. Effect of Noise Type

For latency measures, significant main effects for noise type were found for P1 and N1 peaks but not for P2 [P1: $F(2,16) = 7.4, p = .005$; N1: $F(2,16) = 8.5, p = .003$; and P2: $F(1,2,8,1) = 3.8, p = .082$]. Significant amplitude effects were seen only for N1 [P1: $F(2,16) = 2.3, p = .13$; N1: $F(2,16) = 19.2, p < .001$; and P2: $F(2,14) = 1.4, p = .28$]. In general, this main effect of noise type is demonstrated by mean N1 amplitude and latency values presented in Figure 3; peaks were generally earliest and largest for the interrupted noise and latest and smallest for the four-talker babble with the continuous noise condition results usually falling somewhere in between interrupted and babble noise.

Existing data demonstrated the possible effect of a physiological release from masking when comparing interrupted and continuous noise (Androulidakis et al., 2006; Hiruami et al., 2008). We were interested a priori in this same comparison, especially for N1 because of its sensitivity to stimulus acoustics. Therefore, we completed paired comparisons of N1 amplitude and latency in the interrupted and continuous noise conditions. No significant differences were found in the passive homogenous paradigm for speech [latency: $t(8) = .39, p = .70$; amplitude: $t(8) = 1.6, p = .14$] or tones [latency: $t(8) = 2.4, p = .05$; amplitude: $t(8) = 1.7, p = .12$] when a Bonferroni-corrected alpha level of .0125 was used. Similarly, in the active oddball paradigm, paired comparisons for speech [latency: $t(8) = 3.2, p = .013$; amplitude: $t(8) = .61, p = .56$] and tones [latency: $t(8) = 2.8, p = .021$; amplitude: $t(8) = 1.8, p = .12$] were also not significant, although latency differences approached significance.

The comparison between continuous noise and four-talker babble allowed for the characterization of informational masking effects that occur in addition to energetic masking effects. Energetic masking refers to a target and masker that overlap in time and frequency resulting in portions of the target that are inaudible; whereas, informational masking occurs when the target and masker are both audible, but the listener has a reduced ability to identify the target (Brungart, et al., 2001). The continuous noise would be representative of energetic masking while the babble would contain both energetic and informational masking components. Paired comparisons were completed using a Bonferroni-corrected alpha level of .0125. In the passive homogenous condition, only N1 amplitude for the speech condition was significant, with the four-talker babble resulting in longer latencies [speech latency: $t(8) = -.07, p = .95$; speech amplitude: $t(8) = 3.3, p = .011$; tone latency: $t(8) = .91, p = .39$; tone amplitude: $t(8) = .42, p = .69$]. None of the comparisons reached significance for the active oddball paradigm [speech latency: $t(8) = .05, p = .96$; speech amplitude: $t(8) = 2.2, p = .057$; tones latency: $t(8) = .57, p = .59$; tone amplitude: $t(8) = 1.3, p = .24$].

3.3. Effect of Paradigm & Interactions

Main effects of paradigm were found for N1 latency and amplitude [$F(1,8) = 6.2, p = .037$; $F(1,8) = 7.0, p = .030$], and for P2 latency but not amplitude [$F(1,7) = 17.0, p = .004$; $F(1,7) = 3.1, p = .12$]. No significant paradigm effects were found for P1 latency or amplitude [$F(1,8) = .44, p = .52$; $F(1,8) = .83, p = .40$]. Effects of paradigm on N1 amplitude can be seen in Figure 4, which displays robust differences in waveform morphology in the oddball paradigm for the tone signal relative to the speech signal.

Significant interactions were found only for P2 latency [paradigm \times noise: $F(2,14) = 6.9, p = .008$] and P2 amplitude [paradigm \times signal: $F(1,7) = 32.4, p = .001$; signal \times noise: $F(2,14) = 9.6, p = .002$]. No significant 3-way interactions were found.

4. Discussion & Conclusions

4.1 Signal Type

Speech signals resulted in delayed peak latencies relative to tonal signals. This result is consistent with previous literature that demonstrates the importance of onset and spectral characteristics of the stimulus. For example, longer latencies have been demonstrated when rise times are slower (Onishi & Davis, 1968; Milner, 1969; Ruhm & Jansen, 1969; Kodera et al., 1979). Lower frequency content in the evoking stimulus also results in prolonged latencies (Jacobson et al., 1992). Both of these effects may have been present in the current comparison between /ba/ and a 1k Hz tone. We also would have expected comparable signal effects on amplitude (i.e., smaller amplitudes for the speech signal); however it may be that the inherent variability in amplitude measures contributed to these signal effects. Another explanation for the absence of a signal type effect on amplitude might have to do with the contribution of noise type. Figure 3 demonstrates how amplitudes (right side of figure) differ for tone and speech signals across the four background noise conditions. For the quiet and four-talker babble conditions, mean amplitudes are equivalent or smaller for speech signals than for tone signals. In contrast, for interrupted and continuous noise conditions, amplitudes are larger for speech signals. The reason for these results may be the specific spectral and rise characteristics of the syllable interacting with the noise type.

4.2 Noise Type: Release from Masking & Informational Masking

As seen in Figure 2, the general effect of noise type on peak waves was that four-talker babble resulted in longest latencies and smallest amplitudes, and that interrupted noise resulted in shortest latencies and largest amplitudes with continuous noise results usually falling between the two other noise types. Interestingly, differences between interrupted and continuous noise were not robust. Based on previous work (Androulidakis et al., 2006), we expected a physiological release from masking to occur when signals were presented in interrupted noise as compared to continuous noise. However, planned paired comparisons of N1 latency and amplitude demonstrated no significant amplitude effects and only a non-significant trend for latency effects. The lacking significant latency effect may be a result, in part, of insufficient statistical power given that only nine subjects were tested. In addition, we hypothesized that due to the random nature of interruptions in the noise, improved encoding during silent gaps was obscured by averaging trials that occurred during silent gaps with those that occurred during noise segments. To verify this hypothesis, we further processed the interrupted noise data by separating trials where the signal was presented in noise from those where the signal was presented in the silent gaps between the noise segments. These gap trials and noise trials were then averaged separately. Signals were considered gap trials if at least the first 30 ms during signal presentation were in a silent interval. If noise was present during the first 30 ms or more of the signal, then the response was considered a noise trial. When noise or gap segments at signal onset were less than 30 ms, the trial was not included in averaging. This time period was chosen based on previous literature that indicates the first 30 ms of the signals determines the morphology of the CAEP (Onishi & Davis, 1968). Approximately 90 trials were available for both gap trial and noise trial averages.

Responses to tones and speech were collapsed, resulting in grand averages for gap trials or noise trials displayed in Figure 5. The original interrupted noise response is also shown. Very different responses are generated for the signals presented in gaps compared with those presented in noise. Based on these results, we hypothesize that larger gaps in the interrupted noise would be necessary to demonstrate the release from masking effect when all trials are averaged together. It is noteworthy that Stuart and Phillips (1995) found behavioral differences for performance in this same interrupted noise relative to continuous noise.

There may be several reasons for this difference between behavioral and electrophysiological results, one of which may be higher order cognitive contributions that were active in the behavioral task but not in the passive listening condition used in this study. It is noteworthy that for the active oddball paradigm in which some cognitive factors were present, N1 latency differences between the interrupted and continuous conditions demonstrated a strong trend toward significance. Another explanation may be the differences in stimuli. Stuart & Phillips (1995) used Northwestern University Test single syllable word lists, whereas, we used the single nonsense syllable /ba/.

One of the most difficult listening situations is the perception of speech in speech background noise. Here, we present the first CAEP signal-in-noise data where a speech background is used. Psychophysical data suggest that informational masking effects are present in some speech on speech behavioral tasks (e.g., Brungart et al., 2001). We hypothesized that a neurophysiological correlate of the informational masking effect would be observed in the current data through a comparison of the continuous and four-talker babble noise types. However, comparisons between the continuous and babble noise conditions were generally not significant. Only the passive homogenous N1 amplitude for the speech signal demonstrated a significant effect of babble (i.e., smaller amplitude for babble compared to continuous noise). Therefore, generally, the four-talker babble did not provide any additional masking beyond that caused by the continuous noise. It may be that the informational masking effects are not as robust for these obligatory evoked potentials as they would be for the later, more cognitive evoked potentials such as the P3. Future studies using speech backgrounds may help to improve our understanding of how different levels of signal-in-noise neural encoding contribute to performance in noisy real-world listening environments.

4.3 Differences Between Passive Homogenous and Active Oddball Paradigms

Context is an important determinant of the latency and amplitude of the evoked response. Decades of literature have demonstrated that signal level is an important contributor to the morphology of evoked responses (e.g., Rapin et al., 1966; Picton et al., 1977; Adler and Adler, 1989; Martin and Boothroyd, 2000). However, previous work has demonstrated that when signals are presented in background noise, signal level becomes less important and signal-to-noise ratio becomes the dominant factor (Billings et al., 2009). In addition to signal and noise characteristics, the current study demonstrates that the evoking paradigm is important as well. Many researchers have illustrated the effect of attention on evoked potentials across a wide range of tasks and stimulus conditions (for reviews, see Alho, 1992; Näätänen, 1992). In general, CAEP amplitude increases in an attended compared to an unattended condition. However, it was important for us to determine the effects of the two current paradigms because they reflect the main paradigms present in the limited signal-in-noise CAEP literature. Figure 4 shows significant increases in N1 amplitude for the active oddball paradigm relative to the passive homogenous paradigm. These effects are consistent with existing literature (e.g., Alho et al., 1994; Fruhstorfer et al., 1970; Picton & Hillyard, 1974; Woldorff & Hillyard, 1991) with an exception: Michalewski and colleagues (2009) found no effect of paradigm on N1 latency or amplitude in an active and passive paradigm. However, passive and active conditions were somewhat different than the current study in that participants fixated on a spot and did not respond to signals (passive) or fixated on a spot and pushed a button when the signal was heard (active). In both cases a homogenous train of stimuli was presented. In contrast, we used a movie to draw attention away from the signal (passive) and presented stimuli in an oddball configuration with a button press to the deviant signal (active). These distinctions are not trivial and could explain the differences in paradigm effect results. It is interesting that the attention-related N1 amplitude effects are absent for the tone-in-quiet condition shown in Figure 6 for various electrode sites, possibly

an effect of suprathreshold signal presentation. Picton & Hillyard (1974) demonstrated larger attention effects near threshold, similar to the current signal-in noise conditions which were presented near masked threshold at -3 dB SNR.

Large paradigm effects were seen for tone-evoked responses relative to speech-evoked responses. Figure 4 shows a large positivity from 180–380 ms for the tone-in-noise evoked responses only in the active oddball paradigm. Scalp topography of this positivity is shown in Figure 6 for the tone-in-quiet condition. The pattern across these 11 scalp electrodes reveals the typical N1-P2 inversions at temporal electrodes (i.e., T7 and T8) relative to central electrodes (e.g., Cz) are present (solid arrows). In addition, another negative peak is present (dotted arrow) corresponding to the latency of the positivity seen at Cz. This activity may reflect a component of the CAEP that is sensitive to the acoustics of the stimuli. For example, it is possible that acoustic distance is an important factor, because only the tone signal produced the positivity. In other words, the activity may be due to the larger acoustic differences between the tone contrast (1000 versus 500 Hz) relative to the speech contrast (/ba/ versus /da/). These results indicate that one must be careful when comparing active oddball CAEPs to potentials recorded in a passive homogenous paradigm, because cognitive variables such as attention and task demands may affect the response. This is relevant given the current state of the CAEP signal-in-noise literature. With the exception of the current study, speech-evoked signal-in-noise CAEPs have been collected exclusively in active paradigms (Whiting et al., 1998; Martin et al., 2005; Kaplan-Neeman et al., 2006), while modulated noise backgrounds have been used in only passive paradigms with a tonal signal (Androulidakis et al., 2006; Hiraumi et al., 2008). Therefore, comparisons among these groups of studies should be made with care.

4.4 Conclusion

Tone and speech evoked CAEPs can be recorded in a variety of background noises, including speech backgrounds. Signal type and noise type are important factors that affect the latency and amplitude of passively elicited CAEP components and also CAEP components recorded in an active oddball paradigm. However, one must be careful in comparing CAEPs across these paradigms because interactions between stimuli and recording paradigm can occur. We conclude that CAEPs may be a useful tool in understanding the underlying causes of perception-in-noise deficits and possibly in understanding the variability in performance that is seen across individuals and groups of people. Further testing of impaired populations is necessary to begin to differentiate CAEP morphology across individuals and groups.

Abbreviations

| | |
|-------|-------------------------------------|
| CAS | central auditory system |
| SNR | signal-to-noise ratio |
| EEG | electroencephalography |
| CAEPs | cortical auditory evoked potentials |
| ANOVA | analysis of variance |

Acknowledgments

We wish to thank Kelly Tremblay, Robert Folmer, and two anonymous reviewers for comments on early versions of this manuscript. This work was supported by a grant from the National Institute on Deafness and Other Communication Disorders (R01 DC00626) and the VA Rehabilitation Research & Development Service as Career Development Awards (Billings; Molis), and a Senior Research Career Scientist award (Leek).

References

- Adler G, Adler J. Influence of stimulus intensity on AEP components in the 80- to 200- millisecond latency range. *Audiology* 1989;28(6):316–324. [PubMed: 2597095]
- Alho K. Selective attention in auditory processing as reflected by event-related brain potentials. *Psychophysiology* 1992;29(3):247–263. [PubMed: 1626035]
- Alho K, Teder W, Lavikainen J, Näätänen R. Strongly focused attention and auditory event-related potentials. *Biol Psychol* 1994;38:73–90. [PubMed: 7999931]
- Androulidakis A, Jones S. Detection of signals in modulated and unmodulated noise observed using auditory evoked potentials. *Clinical Neurophys* 2006;117:1783–1793.
- Bacon S, Opie J, Montoya D. The effects of hearing loss and noise masking on the masking release of speech in temporally complex backgrounds. *J Sp Lang Hear Res* 1998;41:549–563.
- Bennett KO, Billings CJ, Molis MR, Leek MR. Neural encoding of speech signals in informational masking. Neuroreport. In Preparation.
- Billings CJ, Tremblay KL, Stecker GC, Tolin WM. Human evoked cortical activity to signal-to-noise ratio and absolute signal level. *Hearing Research* 2009;254(1–2):15–24. [PubMed: 19364526]
- Brungart DS, Simpson BD, Ericson MA, Scott KR. Informational and energetic masking effects in the perception of multiple simultaneous talkers. *J Acoust Soc Am* 2001;110(5):2527–2538. [PubMed: 11757942]
- Dubno JR, Schaefer AB. Comparison of frequency selectivity and consonant recognition among hearing-impaired and masked normal-hearing listeners. *J Acoust Soc Am* 1992;91:2110–2121. [PubMed: 1597602]
- Eggermont, JJ. Electric and magnetic fields of synchronous neural activity. In: Burkard, RF.; Don, M.; Eggermont, JJ., editors. *Auditory Evoked Potentials*. Philadelphia: Lippincott Williams & Wilkins; 2007. p. 3-21.
- Eisenberg L, Dirks D, Bell T. Speech recognition in amplitude-modulated noise of listeners with normal and listeners with impaired hearing. *J Sp Hear Res* 1995;38:222–233.
- Festen J, Plomp R. Effects of fluctuating noise and interfering speech on the speech reception threshold for impaired and normal hearing. *J Acoust Soc Am* 1990;88(4):1725–1736. [PubMed: 2262629]
- Fruhstorfer H, Soveri P, Jarvilehto T. Short-term habituation of the auditory evoked response in man. *Electroenceph Clin Neurophysiol* 1970;28:153–161. [PubMed: 4189933]
- Hiraumi H, Nagamine T, Morita T, Naito Y, Fukuyama H, Ito J. Effect of amplitude modulation of background noise on auditory-evoked magnetic fields. *Brain Research* 2008;1239:191–197. [PubMed: 18778694]
- Hyde M. The N1 response and its applications. *Audiology and Neuro-otology* 1997;2:281–307. [PubMed: 9390837]
- Institute of Electrical and Electronic Engineers. *IEEE Recommended practice for speech quality measures*. New York: IEEE; 1969.
- Jacobson GP, Lombardi DM, Gibbens ND, Ahmad BK, Newman CW. The effects of stimulus frequency and recording site on the amplitude and latency of multichannel cortical auditory evoked potential component N1. *Ear Hear* 1992;13(5):300–306. [PubMed: 1487089]
- Jerger K, Biggins C, Fein G. P50 suppression is not affected by attentional manipulations. *Biol Psychiatry* 1992;31:365–377. [PubMed: 1558899]
- Kaplan-Neeman R, Kishon-Rabin L, Henkin Y, Muchnik C. Identification of syllables in noise: electrophysiological and behavioral correlates. *J Acoust Soc Am* 2006;120(2):926–933. [PubMed: 16938980]
- Kodera K, Hink RF, Yamada O, Suzuki JI. Effects of rise time on simultaneously recorded auditory-evoked potentials from the early, middle and late ranges. *Audiology* 1979;18(5):395–402. [PubMed: 496722]
- Kraus N, Bradlow AR, Cheatham MA, Cunningham J, King CD, Koch DB. Consequences of neural asynchrony: a case of auditory neuropathy. *Journal of the Association for Research in Otolaryngology* 2000;1:33–45. [PubMed: 11548236]

- Lilly DJ, Hutter MM, Lewis MS, Folmer RL, Shannon JM, Wilmington DJ, Bourdette D, Fausti SA. A “virtual cocktail party” -- development of a sound-field system and materials for the measurement of speech intelligibility in multi-talker babble. *J Am Acad Audiol*. Submitted to.
- Martin BA, Boothroyd A. Cortical, auditory, evoked potentials in response to changes of spectrum and amplitude. *J Acoust Soc Am* 2000;107(4):2155–2161. [PubMed: 10790041]
- Martin BA, Stapells DR. Effects of low-pass noise masking on auditory event-related potentials to speech. *Ear Hear* 2005;26(2):195–213. [PubMed: 15809545]
- Michalewski HJ, Starr A, Nguyen TT, Kong YY, Zeng FG. Auditory temporal processes in normal-hearing individuals and in patients with auditory neuropathy. *Clin Neurophysiol* 2005;116:669–680. [PubMed: 15721081]
- Michalewski HJ, Starr A, Zeng FG, Dimitrijevic A. N100 cortical potentials accompanying disrupted auditory nerve activity in auditory neuropathy (AN): Effects of signal intensity and continuous noise. *Clinical Neurophysiology* 2009;120:1352–1363. [PubMed: 19535287]
- Miller G, Licklider J. The intelligibility of interrupted speech. *J Acoust Soc Am* 1950;22(1):167–173.
- Mollis MR, Summers V. Effects of high presentation levels on recognition of low- and high-frequency speech. *Acoustic Research Letters Online* 2003;4(4):124–128.
- Milner BA. Evaluation of auditory function by computer techniques. *Int J Audiol* 1969;(8):361–370.
- Näätänen, R. *Attention and Brain Function*. Hillsdale, NJ: Erlbaum; 1992.
- 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(4):375–425. [PubMed: 3615753]
- Neuroscan, Inc.. SCAN 4.4 – Vol II, Offline analysis of acquired data (Document number 2203, Revision E). Charlotte, NC: Compumedics Neuroscan; 2007. p. 141-148.
- Oates PA, Kurtzberg D, Stapells DR. Effects of sensorineural hearing loss on cortical event-related potential and behavioral measures of speech-sound processing. *Ear Hear* 2002;23(5):399–415. [PubMed: 12411773]
- Onishi S, Davis H. Effects of duration and rise time of tone bursts on evoked V potentials. *J Acoust Soc Am* 1968;44(2):582–591. [PubMed: 5665530]
- Ostroff JM, McDonald KL, Scheider BA, Alain C. Aging and the processing of sound in human auditory cortex. *Hear Res* 2003;181(1–2):1–7. [PubMed: 12855356]
- Papanicolaou AC, Loring DW, Eisenberg HM. Age-related differences in recovery cycle of auditory evoked potentials. *Neurobiol Aging* 1984;5(4):291–295. [PubMed: 6531067]
- Phillips D. Temporal response features of cat auditory cortex neurons contributing to sensitivity to tones delivered in the presence of continuous noise. *Hear Res* 1985;19(3):253–268. [PubMed: 4066524]
- Phillips DP, Hall SE. Spike-rate intensity functions of cat cortical neurons studied with combined tone-noise stimuli. *J Acoust Soc Am* 1986;80(1):177–187. [PubMed: 3745663]
- Phillips DP. Neural representation of sound amplitude in the auditory cortex: effects of noise masking. *Behavioural brain research* 1990;37(3):197–214. [PubMed: 2340096]
- Phillips DP, Kelly JB. Effects of continuous noise maskers on tone-evoked potentials in cat primary auditory cortex. *Cereb Cortex* 1992;2(2):134–140. [PubMed: 1633411]
- Phillips D, Rappaport J, Gulliver J. Impaired word recognition in noise by patients with noise-induced cochlear hearing loss: contribution of temporal resolution defect. *Am J Oto* 1994;15(5):679–686.
- Picton TW, Hillyard SA. Human auditory evoked potentials. II. Effects of attention. *Electroencephalogr Clin Neurophysiol* 1974;36(2):191–199. [PubMed: 4129631]
- Picton TW, Woods DL, Baribeau-Braun J, Healey TM. Evoked potential audiometry. *J Otolaryngol* 1977;6:90–119. [PubMed: 1030745]
- Plomp R. Acoustical aspects of cocktail parties. *Acustica* 1977;38:186–191.
- Rapin I, Schimmel H, Tourk LM, Krasnegor NA, Pollak C. Evoked responses to clicks and tones of varying intensity in waking adults. *Electroencephalography and clinical neurophysiology* 1966;21(4):335–344. [PubMed: 4162205]
- Rappaport J, Gulliver M, Phillips D, Van Dorpe R, Maxner C, Bhan V. Auditory temporal resolution in multiple sclerosis. *J Otolaryngology* 1994;23(5):307–324.

- Ruhm H, Jansen J. Rate of stimulus change and the evoked response: 1. Signal rise-time. *The Journal of Auditory Research* 1969;(3):211–216.
- Stuart A, Phillips D, Green W. Word recognition performance in continuous and interrupted broadband noise by normal-hearing and simulated hearing-impaired listeners. *Am J Otolaryngology* 1995;16(5): 658–663. [PubMed: 8588673]
- Stuart A, Phillips D. Word recognition in continuous and interrupted broadband noise by young normal-hearing, older normal-hearing, and presbycusis listeners. *Ear Hear* 1996;17:478–489. [PubMed: 8979036]
- Stuart A, Phillips D. Word recognition in continuous noise, interrupted noise, and in quiet by normal-hearing listeners at two sensation levels. *Scand Audiol* 1997;26:112–116. [PubMed: 9187004]
- Tremblay KL, Billings C, Rohila N. Speech evoked cortical potentials: effects of age and stimulus presentation rate. *J Am Acad Audiol* 2004;15(3):226–237. quiz 264. [PubMed: 15119463]
- Tremblay, KL.; Burkard, R. Aging and auditory evoked potentials. In: Burkard, R.; Don, M.; Eggermont, J., editors. *Auditory Evoked Potentials: Scientific Bases to Clinical Application*. 2007. p. 411
- Tremblay KL, Piskosz M, Souza P. Aging alters the neural representation of speech cues. *Neuroreport* 2002;13(15):1865–1870. [PubMed: 12395081]
- 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(7):1332–1343. [PubMed: 12842732]
- Vaughan HG Jr, Ritter W. The sources of auditory evoked responses recorded from the human scalp. *Electroencephalogr Clin Neurophysiol* 1970;28(4):360–367. [PubMed: 4191187]
- Whiting KA, Martin BA, Stapells DR. The effects of broadband noise masking on cortical event-related potentials to speech sounds /ba/ and /da. *Ear Hear* 1998;19(3):218–231. [PubMed: 9657596]
- Wilson R, Carhart R. Influence of pulsed masking on the threshold for spondees. *J Acoust Soc Am* 1969;46(4):998–1010. [PubMed: 5824043]
- Woldorff MG, Hillyard SA. Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalogr Clin Neurophysiol* 1991;79(3):170–191. [PubMed: 1714809]
- Wolpaw JR, Penry JK. A temporal component of the auditory evoked response. *Electroencephalogr Clin Neurophysiol* 1975;39(6):609–620. [PubMed: 53139]

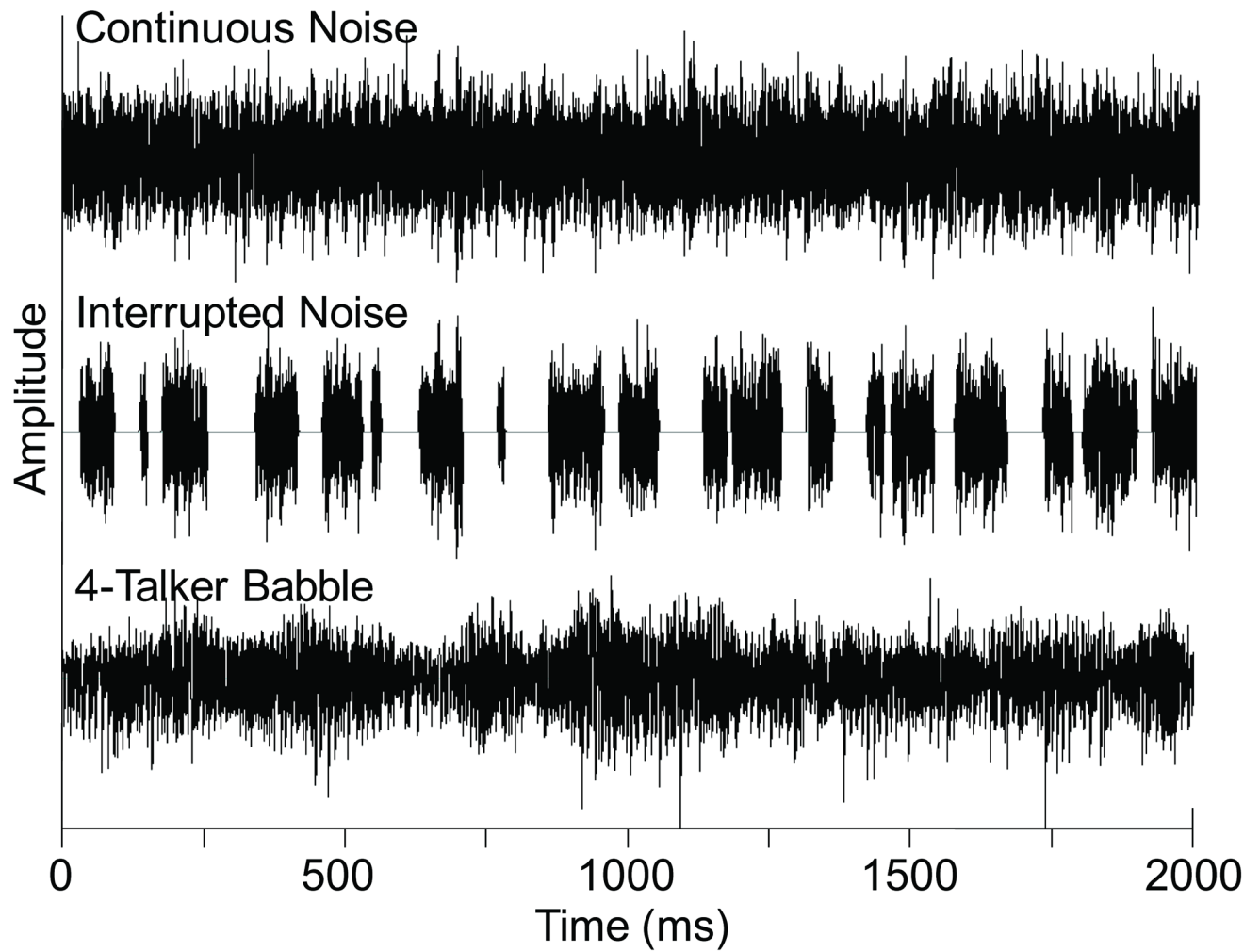


Figure 1. Two-second segments of the time waveforms for the three noise types are shown: continuous speech spectrum noise (top), interrupted speech spectrum noise (middle), and four-talker babble (bottom).

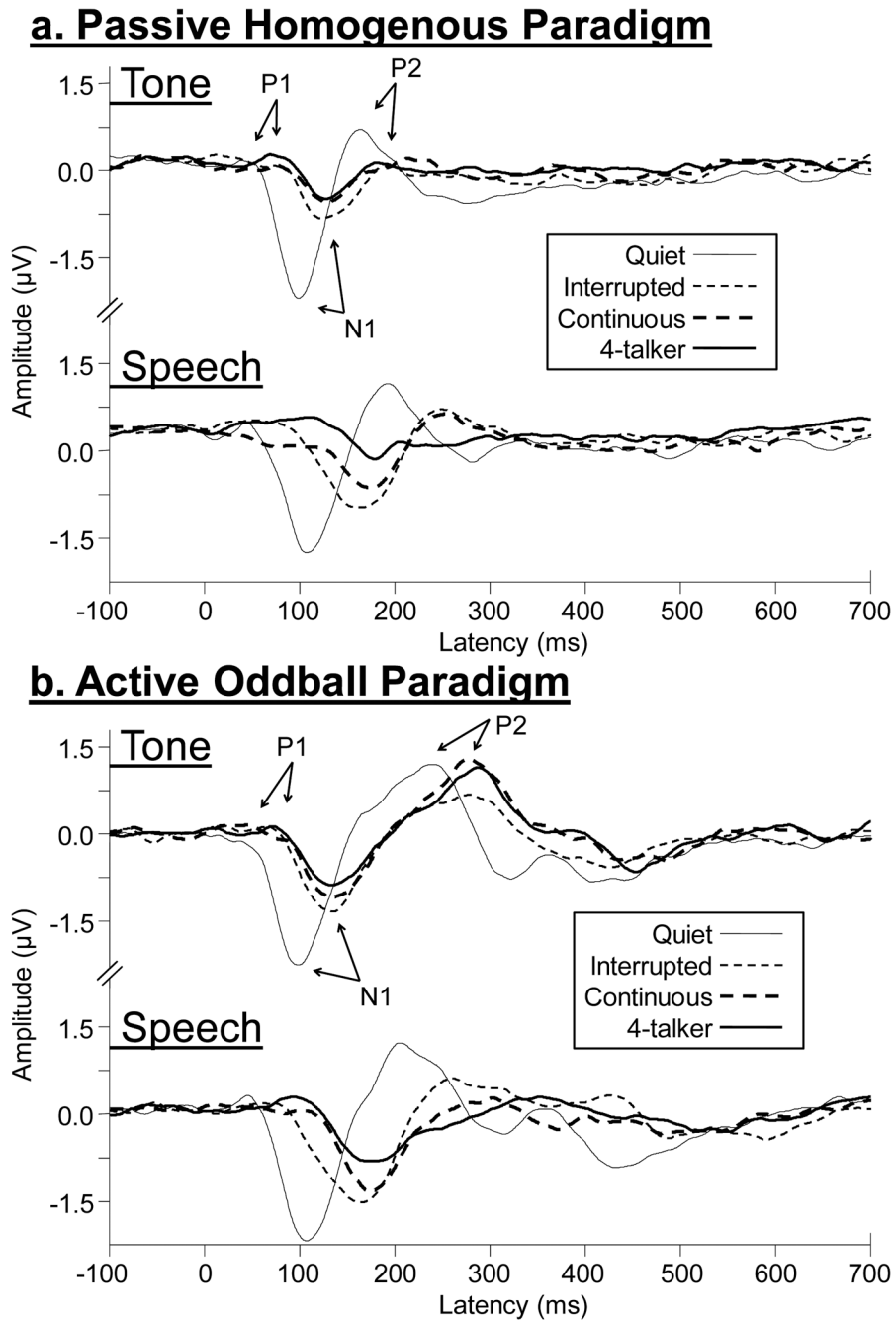
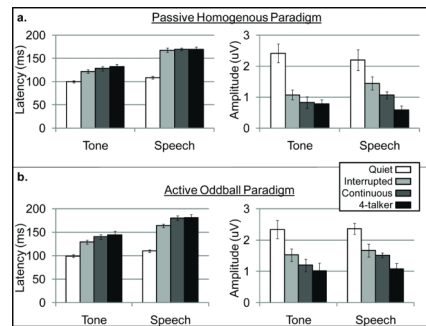


Figure 2.

Grand mean Cz waveforms ($n=9$) for the passive homogenous paradigm (a.) and active oddball paradigm (b.) are shown for the four different background conditions. Responses to the 1000 Hz tone the syllable /ba/ are displayed. Arrows mark approximate P1, N1, and P2 peaks for quiet and noise conditions. Some effects of noise type and signal type are apparent. For example, latencies for the speech syllable are prolonged relative to the tone. Also effects of noise type are seen for the speech syllable with some separation between noise types.

**Figure 3.**

Bar graphs display means and standard error of the means (error bars) for latency (left) and amplitude (right) of the N1 peak across all conditions tested. Results for the passive homogenous paradigm (a.) and the active oddball paradigm (b.) with noise type as the parameter display significant effects of signal type, noise type, and recording paradigm.

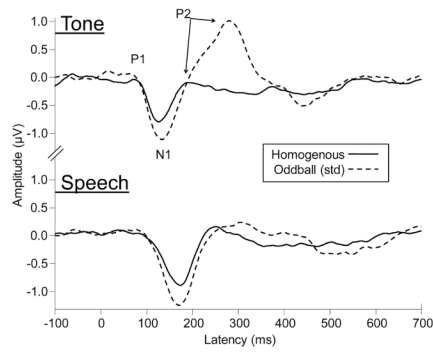


Figure 4. Grand mean Cz waveforms averaged across the nine participants for speech and tones as a function of recording paradigm. P1, N1, and P2 peaks are marked with arrows for the tone condition. These waveforms have been collapsed across the three noise types to clearly demonstrate the effect of paradigm on CAEPs. Increased N1 amplitudes for the standard trials of the oddball paradigm are apparent, as is a large positivity for the tone signal presented in the oddball paradigm.

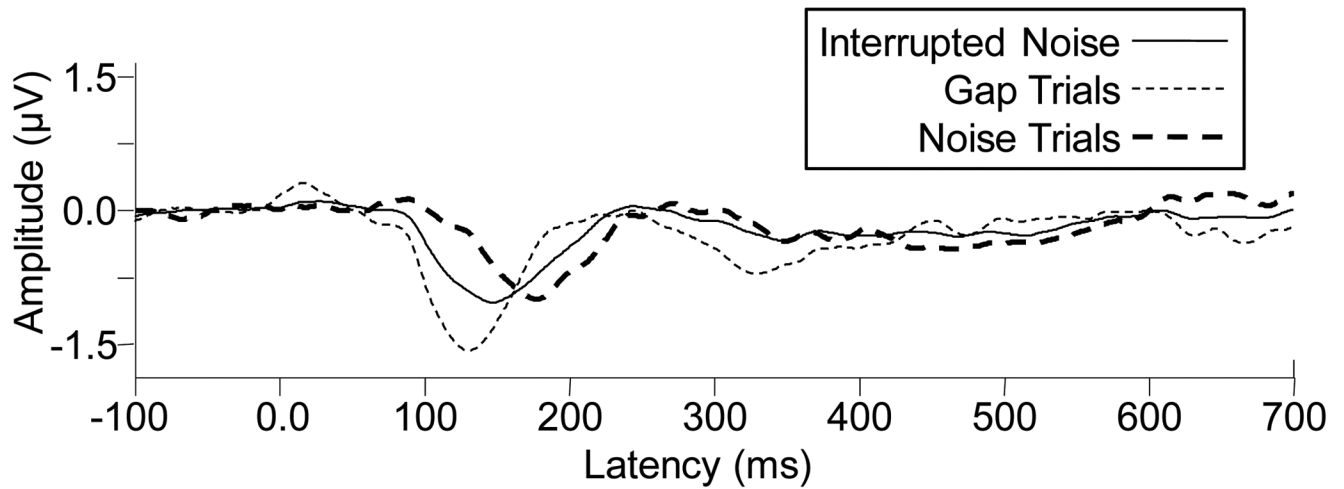


Figure 5.

Grand averages (n=9) for the interrupted speech spectrum noise condition (solid line) and for signals that occurred during silent gaps (thin dotted line) and those that occurred during noise portions (thick dotted line) of the interrupted noise. Increased amplitude and decreased latencies can be seen for the gap-trials waveform relative to the noise-trials waveform. It is not surprising that the response to the original interrupted noise falls between the two subgroups because it incorporates both gap trials and noise trials.

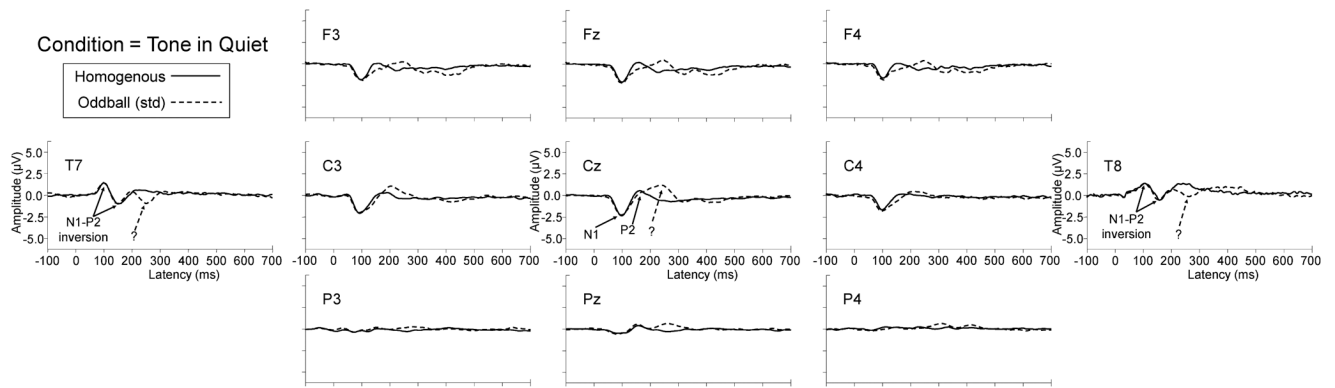


Figure 6.

Grand average waveforms for a subset of electrodes demonstrate scalp topography differences across paradigm for the tonal stimulus. The passive homogenous paradigm (solid line) and the standard stimulus of the active oddball paradigm (dashed line) demonstrate robust differences from 150 to 300 ms. Solid arrows identify N1 and P2 for Cz and corresponding polarity inversions at sites T7 and T8. The dotted arrows identify late activity, the cause of which is not clear. However, the fact that the activity is present only for the tonal stimulus suggests that it might be related to contrast characteristics such as greater acoustic distance between the 500 and 1000 Hz tone contrast than between the /ba/ and /da/ speech contrast.

Table 1

Mean CAEP Values

Mean P1, N1, and P2 latency and amplitude values are given for all conditions tested in this 2 (signals) \times 4 (noises) \times 2 (paradigms) repeated measures design. Standard errors of the mean are provided in parentheses.

| Signal: | Tone | | | | Speech | | | |
|--------------------|--------------|--------------|-------------|--------------|-------------|--------------|--------------|--------------|
| | Quiet | Interrupted | Continuous | 4-Talker | Quiet | Interrupted | Continuous | 4-Talker |
| Passive Homogenous | | | | | | | | |
| Latency | | | | | | | | |
| P1 | 46.33 (2.3) | 69.78 (5.6) | 72.67 (5.4) | 69.4 (3.4) | 45.33 (3.9) | 75.44 (4.2) | 97.33 (10) | 101.7 (8.4) |
| N1 | 99.89 (2.2) | 122.1 (4.1) | 128.0 (4.4) | 132.4 (4.8) | 108.9 (2.7) | 167.7 (4.6) | 169.67 (2.9) | 169.2 (5.4) |
| P2 | 166.4 (5.7) | 208.6 (12) | 210.9 (12) | 212.9 (17.5) | 192.3 (2.1) | 259.9 (13.9) | 259.8 (10.4) | 248.4 (14.1) |
| Amplitude | | | | | | | | |
| P1 | 0.08 (.08) | 0.07 (.12) | 0.02 (.10) | 0.17 (.09) | 0.19 (.12) | 0.21 (.13) | -0.08 (.08) | 0.32 (.07) |
| N1 | -2.42 (.30) | -1.08 (.16) | -0.83 (.20) | -0.79 (.14) | -2.20 (.34) | -1.45 (.22) | -1.06 (.11) | -0.58 (.14) |
| P2 | 0.68 (.29) | 0.17 (.13) | 0.29 (.10) | 0.22 (.13) | 0.85 (.24) | 0.66 (.12) | 0.45 (.13) | 0.20 (.10) |
| Active Oddball | | | | | | | | |
| Latency | | | | | | | | |
| P1 | 45.00 (3.1) | 64.22 (4.9) | 72.56 (5.5) | 70.78 (5.2) | 46.0 (2.2) | 73.89 (7.1) | 95.11 (5.0) | 95.33 (5.3) |
| N1 | 99.22 (3.1) | 129.2 (4.5) | 140.0 (4.8) | 144.4 (7.4) | 109.9 (2.7) | 163.6 (3.4) | 180.3 (5.1) | 180.8 (7.5) |
| P2 | 214.8 (11.5) | 243.9 (10.4) | 265.3 (8.9) | 272.4 (8.8) | 208.6 (5.2) | 265.2 (9.3) | 276.9 (10.2) | 324.9 (12.6) |
| Amplitude | | | | | | | | |
| P1 | -0.10 (.16) | 0.32 (.11) | 0.14 (.23) | 0.18 (.16) | 0.28 (.14) | 0.21 (.17) | 0.12 (.11) | 0.35 (.16) |
| N1 | -2.33 (.29) | -1.52 (.20) | -1.20 (.20) | -1.02 (.24) | -2.37 (.18) | -1.67 (.21) | -1.52 (.08) | -1.08 (.17) |
| P2 | 1.74 (.43) | 0.93 (.35) | 1.48 (.41) | 1.31 (.30) | 1.26 (.45) | 0.68 (.23) | 0.46 (.18) | 0.30 (.28) |