Research Article

Predicting Perception in Noise Using Cortical Auditory Evoked Potentials

CURTIS J. BILLINGS,^{1,2} GARNETT P. McMILLAN,¹ TINA M. PENMAN,¹ AND SUN MI GILLE¹

¹*National Center for Rehabilitative Auditory Research, Portland VA Medical Center, 3710 SW U.S. Veterans Hospital Rd (NCRAR), Portland, OR 97239, USA*

²*Department of Otolaryngology, Oregon Health & Science University, Portland, OR 97239, USA*

Received: 17 May 2013; Accepted: 21 August 2013; Online publication: 13 September 2013

ABSTRACT

Speech perception in background noise is a common challenge across individuals and health conditions (e.g., hearing impairment, aging, etc.). Both behavioral and physiological measures have been used to understand the important factors that contribute to perception-in-noise abilities. The addition of a physiological measure provides additional information about signal-in-noise encoding in the auditory system and may be useful in clarifying some of the variability in perception-in-noise abilities across individuals. Fifteen young normal-hearing individuals were tested using both electrophysiology and behavioral methods as a means to determine (1) the effects of signal-to-noise ratio (SNR) and signal level and (2) how well cortical auditory evoked potentials (CAEPs) can predict perception in noise. Three correlation/regression approaches were used to determine how well CAEPs predicted behavior. Main effects of SNR were found for both electrophysiology and speech perception measures, while signal level effects were found generally only for speech testing. These results demonstrate that when signals are presented in noise, sensitivity to SNR cues obscures any encoding of signal level cues. Electrophysiology and behavioral measures were strongly correlated. The best physiological predictors (e.g., latency, amplitude, and area of CAEP waves) of behavior (SNR at which 50 % of the sentence is understood) were N1 latency and N1 amplitude measures. In addition, behavior was best predicted by the 70-dB signal/5-dB SNR CAEP

condition. It will be important in future studies to determine the relationship of electrophysiology and behavior in populations who experience difficulty understanding speech in noise such as those with hearing impairment or age-related deficits.

Keywords: cortical auditory evoked potentials (CAEPs), event-related potentials (ERPs), signals in noise, signal-to-noise ratio (SNR), background noise, N1

Abbreviations: SPL – Sound pressure level; HL – Hearing level; dB – Decibel; SNR – Signal-to-noise ratio; CAEPs – Cortical auditory evoked potentials; ANOVA – Analysis of variance; LOOCV – Leave-one-out cross-validation; PLS – Partial least squares; RMSPE – Root-mean-square prediction error

INTRODUCTION

Understanding speech in background noise presents a challenge for many individuals. Clinically, one of the hallmarks of perception-in-noise testing is the variability in performance across individuals. In this study both physiological and behavioral measures were used to begin to understand better the performance variability across individuals and groups. A measure of cortical neural encoding provides an estimate of what neural information is present and potentially accessible by the listener for perception to occur. The underlying assumption is that accurate perception is dependent on the neural encoding of the auditory stimulus. By combining information about an intermediate stage of the perceptual process, such as cortical neural encoding, with the endpoints of the perception-in-noise process (i.e., stimulus characteris-

Correspondence to: Curtis J. Billings · National Center for Rehabilitative Auditory Research, Portland VA Medical Center · 3710 SW U.S. Veterans Hospital Rd (NCRAR), Portland, OR 97239, USA. Telephone: +1-503-2208262; fax: +1-503-4022824 email: curtis.billings2@va.gov

tics and then resulting perception), we can improve our understanding of perception-in-noise variability across individuals. That is, the addition of physiological measures may help determine whether deficits are due to abnormal neural encoding or difficulties with higher order cognitive processes needed for speech perception in noise. It is hoped that the combination of physiological and behavioral information may help to inform our understanding of the underlying mechanisms involved in perception in noise and improve diagnosis and/or prediction of perception-in-noise difficulties, leading to more targeted and individualized auditory rehabilitation.

Previous research demonstrates that signal-to-noise ratio (SNR) is a major contributor to the physiological processing of signals in noise (Phillips 1990; Billings et al. 2009; Whiting et al. 1998; Kaplan-Neeman et al. 2006). In addition, a long history of psychophysical studies demonstrates the importance of SNR to the perception of sounds (Hawkins and Stevens 1950; Stevens and Guirao 1967). Hawkins and Stevens (1950) demonstrated that masking produced by white noise was directly proportional to the level of the noise. In other words, masked thresholds closely tracked the level of background noise, and increments in masking level resulted in threshold shifts of similar magnitude. The auditory image model proposed by Patterson and colleagues (e.g., Akeroyd and Patterson 1995; Patterson 1994), which incorporates adaptive threshold tracking, or adaptation to the background temporal envelope level of an ongoing stimulus, also demonstrates the potential importance of SNR. Several psychophysical studies also indicate that signal level plays an important role in perception in noise by demonstrating that signal level has a small but significant effect on accuracy when SNR is held constant (Summers and Molis 2004; Studebaker et al. 1999; Hornsby et al. 2005). In contrast, signal level was not a significant contributor to cortical-level processing (Billings et al. 2009). Billings et al. (2009) recorded cortical auditory evoked potentials (CAEPs) to 1,000-Hz tones presented in continuous background noise and demonstrated that N1 and P2 waves are primarily dependent upon SNR rather than signal level. While N1 and P2 are thought to be obligatory neural responses that are dependent upon the acoustics of the stimulus, the effect of signal level was absent or greatly diminished when signals were presented in noise. An absent signal level effect was surprising given the robust effect of signal level in quiet (Adler and Adler 1989; Billings et al. 2007). This effect of context (i.e., distinct sensitivity to signal level in quiet compared to when background noise is present) is important to understand from both physiological and behavioral perspectives.

To better understand the physiological processes associated with listening to speech in noise, a situation that often results in communication problems, this study examined the perceptual and physiological processes associated with speech processing in a speech-spectrum background noise as a function of signal level and SNR. It was also determined whether an individual's CAEPs predicted behavioral results. It is hypothesized that the robust SNR effects seen for both perception and neural encoding of signals in noise will lead to significant correlations between the measures and good predictions of behavior by physiology.

METHODS

Using a repeated measures design, electrophysiology and behavior were measured under stimulus conditions where absolute signal level and SNR were varied. Electrophysiology outcome measures were amplitudes and latencies for evoked responses P1, N1, P2, and N2; in addition, rectified area measures were computed. Participants were tested behaviorally using sentences presented in background noise.

Participants

Fifteen young normal-hearing listeners participated in this study (mean age = 27.6 years, age range = 23–34; seven males and eight females). All participants were right-handed by self-report and had normal hearing from 250 to 8,000 Hz (i.e., thresholds ≤ 25 dB HL) with normal tympanometric measures (single admittance peak between ± 50 daPa to a 226-Hz tone). All participants reported good general health and no history of otologic disorders. Participants provided informed consent and research was completed with approval from the pertinent institutional review boards.

Stimuli

Stimuli consisted of signals and noises presented to the right ear using an Etymotic ER-2 insert earphone. The electrophysiological signal was the speech syllable /ba/ that was a naturally produced female exemplar from the UCLA Nonsense Syllable Test (Dubno and Schaefer 1992) shortened to 450 ms by windowing the steady vowel offset of the stimulus. While previous CAEP research focusing on signal level and SNR effects used 1,000-Hz pure tones (Billings et al. 2009), a short consonant–vowel speech token provided an ideal CAEP stimulus while providing better face validity when comparing to behavioral speech perception measures. The behavioral signal consisted of

Institute of Electrical and Electronic Engineers (IEEE) sentences (IEEE 1969) produced by a female talker. Stimuli were presented at four intensity levels: 50, 60, 70, and 80 dBC (broadband C-weighted) sound pressure level (SPL). Speech spectrum continuous noise was added to the background to create varying SNRs. For electrophysiology testing, noise was present continuously for each testing condition/block. For behavioral testing, the noise was gated on and off for the presentation of each sentence (gated on 2 s prior and remained on until 2 s after the sentence). Spectral shaping of the continuous noise matched the long-term spectrum of speech. Recordings of the 720 IEEE sentences (IEEE 1969) were concatenated and submitted to a fast Fourier transform (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. Both the /ba/ signal and the background noise were low-pass filtered at 4,000 Hz. SNRs ranged from -10 to 35 dB. A total of 14 conditions were tested for electrophysiology and 22 for behavior (because of restrictions on test time, SNRs of -10 and 0 dB were tested for behavior only; see Table 1). Condition presentation order was randomized for each participant to avoid order effects associated with fatigue. Calibration of these signals was completed by measuring the overall RMS level of 10 s of a concatenated version of each signal.

Electrophysiology

A PC-based system controlled the timing of stimulus presentation and delivered an external trigger to the acquisition system (Compumedics Neuroscan Scan 4.5, Charlotte, NC, USA). A passive homogeneous paradigm for each condition was used to evoke the P1, N1, P2, and N2 waves. Participants were asked to ignore the stimuli and watch a silent closed-captioned movie of their choice. A single stimulus was repeated for a total of at least 150 trials. Evoked potential activity was recorded using a 64-channel tin electrode

cap (Electro-Cap International, Inc., Eaton, OH, USA). The ground electrode was located on the forehead during CAEP acquisition, and Cz was the reference electrode. Data were baseline corrected, linear detrended, and re-referenced off-line to an average reference. The recording window consisted of 100 ms pre-stimulus and 700 ms post-stimulus periods. Evoked responses were low-pass filtered online at 100 Hz (12 dB/octave roll-off), amplified with a gain of 10 times, and converted using an analog-to-digital sampling rate of 1,000 Hz. Horizontal and vertical eye movement was monitored with electrodes located inferiorly and at the outer canthi of both eyes. Trials with eye-blink artifacts were corrected off-line, using Neuroscan software (Neuroscan, Inc 2007). This blink reduction procedure calculates the amount of covariation between each evoked potential channel and a vertical eye channel using a spatial singular value decomposition and removes the vertical blink activity from each electrode on a point-by-point basis to the degree that the evoked potential and blink activity covaried. After blink correction, trials containing artifacts exceeding $\pm 70 \mu\text{V}$ were rejected from averaging. After 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).

Behavioral Testing

During a separate session, participants completed a behavioral sentence-in-noise identification task. Signals were taken from the IEEE sentence lists spoken by a female talker (IEEE 1969). Each sentence contains five keywords and was scored as the number of keywords that were correctly identified. Ten sentences were presented for each condition (i.e., total possible per condition = 50 points). Participants were asked to repeat each sentence, and two judges scored correct responses. Prior to testing, approximately 10 total sentences presented at a signal level of 80 dB and at varying SNRs (some at a good SNR and some at a poor SNR) were presented to all the

TABLE 1

Behavioral and cortical auditory evoked potential conditions tested							
Signal level (dBC SPL)	Signal-to-noise ratio (dB)						
	-10	-5	0	+5	+15	+25	+35
50	Beh	Beh/CAEP	Beh	Beh/CAEP	*	*	*
60	Beh	Beh/CAEP	Beh	Beh/CAEP	Beh/CAEP	*	*
70	Beh	Beh/CAEP	Beh	Beh/CAEP	Beh/CAEP	Beh/CAEP	*
80	Beh	Beh/CAEP	Beh	Beh/CAEP	Beh/CAEP	Beh/CAEP	Beh/CAEP

*Higher SNR conditions for the 50-, 60-, and 70-dB signal level conditions were not tested due to noise floor constraints

Beh Behavioral, CAEP cortical auditory evoked potential

participants to allow acclimation to the task. The IEEE sentences were presented in the same signal and background noise conditions as those used for electrophysiological testing with the addition of 0- and -10-dB SNR conditions for each level. Overall RMS of the concatenated sentences was used for sentence calibration. Sentences and noise were presented monaurally to the participants' right ear using ER-2 insert earphones.

Data Analysis and Interpretation

Waves P1, N1, P2, and N2 were analyzed at electrode site Cz and using the global field power (GFP) waveform, which quantifies simultaneous activity from all electrode sites (Skrandies 1989). Because GFP is the standard deviation across channels as a function of time, the GFP waveform is always positive. Robust activity of opposite polarities resulting from generating dipoles leads to greater standard deviations and results in positive peaks in the GFP waveform that generally correspond with latencies of N1, P2, and N2 waves. It should be noted that P1 is generally not present in the GFP waveform because it is usually very close to baseline (i.e., its standard deviation across channels is very low). Peak amplitudes were calculated relative to baseline, and peak latencies were calculated relative to stimulus onset (i.e., 0 ms). A general rectified area amplitude was also used as an overall measure of waveform robustness; area calculations were made with a time window of 50 to 500 ms, comprising the entire CAEP complex. Latency and amplitude values for each wave were determined by agreement of two judges. To determine peaks, each judge used temporal electrode inversion, global field power traces, and two subaverages made up of even and odd trials (to demonstrate replication) for a given condition.

Mean amplitudes and latencies were modeled using a linear mixed model to determine effects of signal level and SNR on physiological measures. This model has the twofold advantage of (1) being fit using maximum likelihood so that all observations are included in the analysis and not just observations for subjects with complete data and (2) allowing a general model of the data covariance structure that may vary across conditions and physiological measures. This general-purpose modeling framework reduces to the conventional repeated measures analysis of variance (ANOVA) model if one accepts a constant covariance across all observations taken on a subject.

Additionally, the accuracy with which the electrophysiology measures predicted behavior was assessed by (1) selecting empirically promising candidates using Pearson's correlational analysis, (2) using leave-one-out cross-validation (LOOCV) with stepwise linear

regression to generate nearly unbiased estimates of the prediction error for CAEP measures, and (3) partial least squares (PLS) regression to determine the optimal CAEP condition for predicting behavior.

RESULTS

Electrophysiology: Effects of SNR and Signal Level

Grand averaged Cz waveforms/growth functions are shown in Figure 1. Grand averaged waveforms are displayed (top panel) as a function of SNR with overlaid signal level conditions. Mean N1 and P2 latency and amplitude growth functions are also shown (bottom panel), overlaid on individual subject data. Response waveforms and growth functions for GFP are shown in Figure 2. These figures illustrate that as SNR increases, amplitudes get larger and latencies get shorter. It should be noted that GFP waveforms are always positive (as explained above) and therefore show a positive N1 peak getting more positive, whereas the Cz N1 is negative and gets more negative with increases in SNR. Statistical results are shown in Table 2 and indicate that amplitudes increased and latencies decreased significantly as SNR increased; this was true for N1, P2, and N2 waves. P1 latency also decreased with SNR; however, P1 amplitude did not significantly change with SNR. An overall rectified area measure demonstrated significant effects of SNR as well, such that the area increased as SNR increased. Potential main effects of signal level can be seen for GFP N1 amplitude and GFP rectified area. However, it should be noted that no correction has been made for multiple comparisons (e.g., Bonferroni, etc.). No significant SNR $\times$ signal level interactions were found.

Behavior: Effects of SNR and Signal Level

Sentence perception results are shown in Figure 3; average psychometric curves and individual data points for each signal level are shown with overlaid signal level curves below. As with CAEPs, increases in SNR result in improved performance. Signal level effects are also evident in Figure 3 such that at smaller SNRs, performance is worse at higher signal levels. A logistic regression model with categorical variables (SNR, signal level, and SNR $\times$ signal level interaction) was fit by maximum likelihood. Correlations among repeated measures on a subject were modeled with random intercepts. This model is analogous to a two-way repeated measures ANOVA model fit to binomial data. Analysis results demonstrated main effects of both SNR and signal level ($F(6,270)=535.5$, $p<0.0001$ and $F(3,270)=5.10$, $p<0.0019$, respectively) and a significant SNR $\times$ signal level interaction ($F(12,270)=$

7.26, $p < 0.0001$). Additional analysis was completed to determine the causes of the interaction term. Figure 4 shows the individual comparisons as a function of SNR (separate panels) and signal level (abscissa). Asterisks indicate statistically significant differences between that specific signal level and others at that given SNR. Brackets indicate significant differences for a specific contrast. Generally, the 80-dB signal level resulted in poorer performance than lower signal levels, particularly at lower SNRs (i.e., SNRs of -10, -5, and 0 dB). In the literature, the poorer performance that was found at higher signal levels is

often referred to as “rollover” and will be discussed further below.

Electrophysiology as a Predictor of Behavior

Many approaches could be taken to determine the relationship between electrophysiology and behavior. Three methods were used to determine what CAEP measures best predicted behavior. The behavioral measure that was to be predicted was the SNR at which 50 % of the sentences were understood (SNR50); the SNR50 is a common

FIG. 1. Grand averaged ($n=15$) Cz waveforms and SNR growth functions. The *top panel* shows waveforms as a function of SNR; signal level waveforms are overlaid for each SNR. The *bottom panel* shows N1/P2 amplitude and latency modeled growth functions on top of individual data points. Each *column* represents growth functions at a given signal level. Overlaid signal level functions are displayed in the *last column*. A robust effect of SNR and minimal effect of signal level can be seen.

clinical measure of sentence-in-noise perception. All CAEP measures (i.e., amplitudes, latencies, and areas) were treated as possible predictors of SNR50 at each of the four behavioral signal levels (50, 60, 70, and 80 dB).

In the first method, predictors were selected based on ranked Pearson's correlations between the CAEP measure and SNR50 at a given signal level. The five best predictors at each level are shown in the top portion of Table 3 with the best predictor in *italics*. Good Pearson's correlation coefficients ranging from 0.62 to 0.77 were found, with root-mean-square prediction errors (RMSPE) of about 1 dB (see

Table 3, column labeled "Biased RMSPE"). That is, for this sample, the SNR50 prediction will be off by approximately 1 dB on average. There are some disadvantages of the Pearson correlation approach. With the Pearson approach, only one CAEP measure is used to predict behavioral results; however, it may be that a combination of CAEP measures better predicts the SNR50. In addition, an estimate of the average prediction error based on Pearson's correlation is almost certainly optimistic (i.e., biased) since the error is evaluated from the data that are used to compute the correlation. Rankings based on this biased approach may select suboptimal CAEP mea-

FIG. 2. Grand averaged ($n=15$) GFP waveforms and SNR growth functions. The *top panel* shows waveforms as a function of SNR; signal level waveforms are overlaid for each SNR. The *bottom panel* shows N1/P2 amplitude and latency modeled growth functions on top of individual data points. Each *column* represents growth functions at a given signal level. Overlaid signal level functions are displayed in the *last column*. A robust effect of SNR and minimal effect of signal level can be seen.

TABLE 2

Statistical analysis of SNR and signal level on amplitudes, latencies, and area for Cz and GFP waveforms						
	<i>Effect of SNR</i>		<i>Effect of tone level</i>		<i>SNR $\times$ tone interaction</i>	
	<i>F statistic (df)</i>	<i>p value</i>	<i>F statistic (df)</i>	<i>p value</i>	<i>F statistic (df)</i>	<i>p value</i>
Cz						
Latency						
P1	45.5 (4.0, 85.4)	<0.001	0.9 (3.0, 108.3)	0.431	1.2 (6.0, 108.1)	0.323
N1	229.0 (4.0, 100.1)	<0.001	1.2 (3.0, 73.0)	0.310	0.6 (6.0, 105.2)	0.725
P2	87.7 (4.0, 82.4)	<0.001	0.7 (3.0, 98.0)	0.574	0.4 (6.0, 104.8)	0.865
N2	16.1 (4.0, 72.8)	<0.001	1.7 (3.0, 116.0)	0.176	0.2 (6.0, 110.1)	0.973
Amplitude						
P1	0.8 (4.0, 76.0)	0.513	0.4 (3.0, 180.0)	0.751	1.0 (6.0, 180.0)	0.450
N1	23.7 (4.0, 28.6)	<0.001	1.6 (3.0, 139.9)	0.182	0.2 (6.0, 139.9)	0.971
P2	19.1 (4.0, 30.6)	<0.001	1.8 (3.0, 145.6)	0.144	0.3 (6.0, 145.6)	0.952
N2	17.2 (4.0, 44.4)	<0.001	0.2 (3.0, 143.7)	0.874	1.1 (6.0, 143.8)	0.384
Area (50–550 ms)	32.8 (4.0, 32.1)	<0.001	0.9 (3.0, 143.1)	0.431	0.8 (6.0, 143.1)	0.551
GFP						
Latency						
N1	42.6 (4.0, 104.0)	<0.001	0.2 (3.0, 71.0)	0.913	0.3 (6.0, 103.9)	0.949
P2	40.4 (4.0, 91.1)	<0.001	0.3 (3.0, 96.3)	0.851	0.9 (6.0, 107.6)	0.469
N2	5.1 (4.0, 72.3)	<0.001	1.0 (3.0, 115.8)	0.380	0.9 (6.0, 110.7)	0.528
Amplitude						
N1	16.2 (4.0, 26.0)	<0.001	3.2 (3.0, 135.6)	0.026	1.2 (6.0, 135.7)	0.299
P2	20.8 (4.0, 35.3)	<0.001	1.8 (3.0, 181.0)	0.149	0.1 (6.0, 181.0)	0.998
N2	27.2 (4.0, 40.8)	<0.001	0.5 (3.0, 151.0)	0.710	0.7 (6.0, 151.0)	0.667
Area (50–550 ms)	50.5 (4.0, 34.9)	<0.001	3.4 (3.0, 147.5)	0.019	1.2 (6.0, 147.5)	0.287

asures, depending on the degree of bias. An unbiased estimate can be derived from a leave-one-out cross-validation (see the second method below) with a regression model including only the predictor selected by ranking Pearson's correlations (Hastie et al. 2009). This procedure allowed for more unbiased estimation of the RMSPE. The unbiased estimation of the RMSPE would likely be more representative of groups or individuals outside this study. This procedure resulted in unbiased RMSPE values (i.e., 0.85 to 1.3 dB) that are slightly larger than biased values.

In a second method, LOOCV was used to generate nearly unbiased estimates of the prediction error for each CAEP measure (Hastie et al. 2009). LOOCV is widely used in developing prediction models, whereby a data set composed of n observations is replicated n times and each observation response is iteratively set to “missing” in each replicate. The prediction model is developed separately for each iteration, and prediction error is computed on the individual response that was set to missing. The result is a data set of n predictions made on observations that were not used in developing the prediction model. This procedure eliminates bias in the prediction errors ordinarily obtained when conducting standard regression analysis. Models including multiple CAEP measures were

evaluated using stepwise selection linear regression. Various stepwise selection stopping rules were considered to achieve a reasonable balance between prediction error (which is reduced with more predictors) and generalizability (which is increased with fewer predictors). In the end, a $p=0.025$ variable entry and exclusion criterion was used, which consistently yielded no more than two important predictors at each behavioral test level. The LOOCV predictors are shown in the middle section of Table 3 with their associated RMSPE values. The LOOCV method resulted in slightly smaller or equivalent prediction errors. Correlations between the best predictors and SNR50 are shown in Figure 5. A drawback of these first two approaches is that different CAEP conditions were found to best predict the SNR50 at a given behavioral signal level. For example, a prediction of the 50-dB SNR50 would require testing the 70-dB signal/5-dB SNR CAEP condition, whereas predicting the 80-dB SNR50 would require testing the 60-dB signal/−5-dB SNR CAEP condition.

A third method was used in an attempt to minimize the potential need to test multiple CAEP conditions and to make the prediction of the SNR50 more intuitive. What CAEP condition, with all of its accompanying measures, best predicts the SNR50 at

FIG. 3. Behavioral SNR growth functions. Percent correct as a function of SNR is plotted for each signal level. Each panel contains the model fit curve plotted on top of individual data points for the 15 participants. A robust effect of SNR and a small but significant signal level effect are shown (seen in the lower panel of overlaid curves).

all behavioral levels? That is, if only one CAEP condition could be run, which one would best predict the SNR50 of all four signal levels tested, and how good is the prediction? PLS regression was used; the response vector (behavioral SNR50 at 50-, 60-, 70-, and 80-dB signal levels) was regressed on a linear combination of the predictors weighted by their covariance with the response vector (Hastie et al. 2009). PLS regression has the distinct advantage of allowing one to predict multivariate responses using several correlated predictors, which can be challenging with standard methodology. The model development was restricted to one factor due to limited sample size.

The CAEP condition that offered the smallest overall average RMSPE was a signal level of 70 dB presented at 5-dB SNR. As shown in Table 3, the 70-dB signal/5-dB SNR condition gave an RMSPE of 0.74 dB averaged across all behavioral levels. That is, CAEP measures from the 70-dB signal/5-dB SNR CAEP condition predict the behavioral SNR50 to within 1 dB on average. Figure 6 shows the fitted regression coefficients for each CAEP measure as a function of the four behavioral signal levels (separate lines). Larger absolute value coefficients represent CAEP measures that are more important to the prediction than those that are close to zero (these include Cz N1 amplitude and GFP N1 amplitude). Another feature of Figure 6 is the common shape of the four curves indicating that the same features best predict all of the behavioral outcomes.

DISCUSSION

The results demonstrate a robust SNR effect for both electrophysiology and behavior. In contrast, significant signal level effects are only present for behavioral outcomes. The best condition-specific electrophysiology predictors of behavior were N1 amplitude and latency values taken from Cz and/or GFP waveforms; furthermore, the best overall CAEP condition for predicting the behavioral SNR50 of any level is the 70-dB signal/5-dB SNR condition.

Electrophysiology

Existing data demonstrate that SNR, rather than tonal signal level, is a key factor affecting CAEPs recorded in noise (Billings et al. 2009). The current study extends those findings to speech stimuli presented at a larger range of signal levels. SNR had a significant effect on latencies and amplitudes of N1, P2, and N2 CAEP waves, as well as a rectified area measure of the entire complex.

Some level effects were found for GFP N1 amplitude and GFP area amplitude measures, although it should be noted that no multiple-comparison corrections have been applied making these effects less noteworthy. There is evidence in the literature for a potential signal level effect in the presence of noise. Whiting et al. (1998) varied SNR and signal level in young normal-hearing individuals. While SNRs were not equated across signal level (i.e., for each signal level, background noise levels resulted in slightly different SNRs making comparisons across level more difficult), they showed some qualitative grouping by signal level of N1 latencies and N2 amplitudes at SNRs greater than 5 dB. Given some subtle indications of a level effect in this study and the literature, it may be

that level encoding is present, but that the stronger SNR effect conceals any level coding. The strong effect of SNR seems to come at the expense of the signal level sensitivity that is present without background noise (e.g., Adler and Adler 1989; Billings et al. 2007). Interestingly, near-field animal studies demonstrate that signal level, in the presence of background noise, is encoded to varying degrees at different levels in the central auditory pathway; these effects are seen in systematic shifts in the neural firing rate-level functions (Phillips 1990; Costalupes et al. 1984; Gibson et al. 1985; Rees and Palmer 1988). Therefore, it is likely that signal level is encoded, but that the effect remains unseen in human CAEPs. This may be due to the specific firing characteristics of cortical neurons. For example, while cortical neurons demonstrate some sensitivity to sustained stimuli (Wang et al. 2005), the cortical population response measured through electroencephalography as post-synaptic potentials is most sensitive to stimulus changes such as onsets and offsets; in contrast, auditory nerve fibers fire continuously to ongoing background noise (Gibson et al. 1985; Phillips and Hall 1986). Interestingly, recent data demonstrate that offset CAEPs may be sensitive to signal levels in noise in ways that the onset CAEPs are not (Baltzell and Billings 2013). It will be important to determine if the dominance of SNR encoding over signal level is modified in different groups of individuals, such as older people or people with hearing impairment.

The dominance of SNR cues as demonstrated in this study has implications for aided CAEPs, where background noise is often present (e.g., amplified ambient noise, hearing aid circuit noise). These data demonstrate that when background noise is audible, care must be taken when examining aided CAEPs for assumed level effects; it may be that waveform morphology changes, or the lack thereof, are due to SNR cues rather than absolute signal level cues. The key factor to consider is whether the noise is audible or not; inaudible noise results in expected signal-level effects, whereas audible noise results in dominance of SNR (Billings et al. 2012).

Behavior

Speech perception-in-noise testing revealed significant effects of both SNR and signal level. These results contrast with CAEP results, where generally only SNR was a significant factor. However, a main effect of signal level is consistent with the perception-in-noise literature. Sentence, word, and syllable perception has been shown to get worse at high signal levels (clinically referred to as rollover) in quiet and in background noise, and the magnitude of the performance decrement at high signal levels is modulated by other factors, such as SNR, noise spectrum, and noise bandwidth (Molis and Summers 2003; Hornsby

et al. 2005; Studebaker et al. 1999; Summers and Molis 2004). In the current study, the signal level effects found in the literature were replicated; decrements in this study occurred above 60 dB, primarily at the 70- and 80-dB signal levels. Furthermore, rollover effects were dependent upon SNR, such that significant decrements due to signal level occurred at or below a 5-dB SNR (see Fig. 4). This result is in agreement with the results of Studebaker et al. (1999) and conforms with their hypothesized mechanisms of rollover; when higher levels of noise are present (i.e., poorer SNRs), more spread of masking and distortion will result, leading to rollover effects.

Differences in Signal Level Encoding Between Electrophysiology and Behavior

The behavioral rollover effect (i.e., decrements in performance at higher presentation levels) has been attributed to the spread of excitation and masking effects at the level of the cochlea that result in distorted encoding throughout the auditory system (Studebaker et al. 1999; Summers and Molis 2004). It is significant, then, that these cochlear effects of signal level were not seen at the level of the cortex as demonstrated by CAEPs. This is not to say that signal level is not encoded at the level of the cortex; it may only be a function of the differential sensitivity that CAEPs, specifically, have to SNR cues rather than absolute signal level cues. CAEPs are largely obligatory responses that show sensitivity to stimulus acoustics such as signal level in quiet; however, the presence of background noise modifies how signal level is encoded. It may be that the salience of the SNR cue masks any underlying encoding of signal level. Given that subjects were not actively attending to stimuli, it appears that the sensitivity to SNR rather than signal level is not driven by attention effects related to active gain control mechanisms (e.g., Hillyard et al. 1973; Hillyard et al. 1998). The sensitivity to SNR may result from a mechanism like the automatic normalization process that occurs in the visual system (Schwartz and Simoncelli 2001), or may simply be a manifestation of the change-dependent characteristics of cortical neurons that respond to acoustic changes rather than continuous unchanging stimuli (Goldstein et al. 1968; Naatanen and Picton 1987; Phillips and Hall 1986). Regardless of the mechanism, this study demonstrates that CAEPs are sensitive to relative change within the acoustic environment (i.e., SNR) rather than to absolute signal magnitude.

Electrophysiology as a Predictor of Behavior

A purpose of this study was to determine what CAEP measure (or combination of measures) would best predict the behavioral SNR50. Some studies have used

FIG. 4. Behavioral interactions between SNR and signal level. Individual comparisons are shown as a function of SNR (separate panels). Asterisks indicate significant differences with all other levels at that SNR. Brackets identify specific comparisons that were significant at that SNR. The figure reveals that generally the 80-dB signal level demonstrates poorer accuracy scores, particularly at lower SNRs (i.e., SNRs of -10, -5, and 0 dB).

Pearson’s correlation between electrophysiological and behavioral variables of interest (Anderson et al. 2010; Bennett et al. 2012; Parbery-Clark et al. 2011); however, the Pearson approach may not be ideal for

TABLE 3
Predictions based on ranking absolute Pearson’s correlation and on LOOCV stepwise regression

Behavioral signal level	CAEP signal level	CAEP SNR	Best predictors	Correlation coefficient (<i>r</i>)	Biased RMSPE	Unbiased RMSPE
Absolute Pearson’s correlation						
50	70	5	<i>CZ P1 latency</i>	0.726	1.06	1.25
	70	5	<i>CZ N1 amplitude</i>	0.725	–	–
	60	15	<i>CZ P1 amplitude</i>	0.722	–	–
	70	5	<i>GFP N1 amplitude</i>	–0.700	–	–
	60	5	<i>CZ N1 amplitude</i>	0.677	–	–
60	70	5	<i>CZ N1 amplitude</i>	0.770	0.51	0.57
	80	35	<i>CZ P1 amplitude</i>	0.743	–	–
	60	–5	<i>GFP N1 amplitude</i>	–0.705	–	–
	60	–5	<i>CZ N1 amplitude</i>	0.688	–	–
	80	5	<i>GFP area amplitude</i>	–0.632	–	–
70	80	5	<i>CZ N1 latency</i>	0.627	1.19	1.39
	50	–5	<i>GFP P2 latency</i>	0.600	–	–
	60	–5	<i>CZ P2 latency</i>	0.591	–	–
	60	–5	<i>GFP area amplitude</i>	–0.577	–	–
	70	–5	<i>GFP N1 latency</i>	0.569	–	–
80	60	–5	<i>GFP N1 amplitude</i>	–0.699	0.73	0.85
	60	–5	<i>GFP area amplitude</i>	–0.687	–	–
	60	–5	<i>CZ N1 amplitude</i>	0.669	–	–
	70	5	<i>CZ P1 latency</i>	0.654	–	–
	70	5	<i>CZ N1 amplitude</i>	0.642	–	–
LOOCV stepwise regression						
50	70	5	<i>CZ N1 amplitude</i>			1.22
60	70	5	<i>CZ N1 amplitude and GFP N1 amplitude</i>			0.48
70	80	5	<i>CZ N1 latency</i>			1.39
80	60	–5	<i>GFP N1 amplitude</i>			0.85
Partial least squares regression						
50	70	5				0.85
60	70	5				0.69
70	70	5				0.65
80	70	5				0.72
Average	70	5				0.74

The best predictors are in italics
CAEP cortical auditory evoked potential, SNR signal-to-noise ratio, RMSPE root-mean-square prediction error, LOOCV leave-one-out cross-validation

FIG. 5. Relationship between behavior at a given signal level and the best predicting CAEP measure. These *plots* show the relationships between the behavioral results and the best predicting EP components for 50-, 60-, 70-, and 80-dB sentences in noise.

at least two practical theoretical reasons. First, the method explicitly allows for only one electrophysiology measure for predicting behavioral test results, which underutilizes the available data since each electrophysiology trace may include several measures (e.g., N1 amplitude, N1 latency, P2 amplitude, etc.). Second, an estimate of the average prediction error

based on Pearson's correlation is almost certainly optimistic since the error is evaluated from the data that are used to compute the correlation. Therefore, in addition to using the Pearson correlation approach, LOOCV with stepwise linear regression was used to generate nearly unbiased estimates of the prediction error for combinations of CAEP measures.

FIG. 6. Regression coefficients for each CAEP feature used in the PLS regression method of predicting behavioral SNR50s. Behavioral signal levels are indicated by each *line*. CAEP features with coefficients furthest from zero are more important than those features with coefficients close to zero. Parameter estimates are based on standardized predictors (i.e., CAEP measures) and responses (i.e., behavioral SNR50). *Lines* connecting the features have been added to highlight relative coefficient strength across features and signal levels.

The LOOCV approach determined that the N1 wave (both latency and amplitude depending on the behavioral signal level) was the best predictor of sentence-in-noise performance. Because the N1 wave is driven primarily by the acoustics of the evoking stimulus (Naatanen and Picton 1987; Hyde 1997), it may be the ideal wave to demonstrate the acoustically driven effects of SNR in these young normal-hearing participants. It is possible that the good prediction characteristics of N1 are population specific. For example, when a top-down processing approach is used, as is the case when hearing impairment is present, the somewhat more endogenous P2 or even the cognitive P3 may prove to be the better predictors of behavior. An improved understanding of the relationship between behavior and physiology (i.e., which evoked potentials best predict behavior) across different populations may enhance our ability to diagnose and rehabilitate speech perception-in-noise deficits. It is noteworthy that in this procedure, any CAEP measure at any level was allowed to predict the behavioral SNR50 measure and that the -5- or 5-dB SNR conditions resulted in the best predictors of behavior. Acoustically, the face validity of this result is good given that most behavioral SNR50s were in the same range (i.e., -5- to 5-dB SNR).

In an effort to simplify predictions, the CAEP condition (with all its measurements) that would best predict the behavioral SNR50 at all four signal levels was determined. That is, if only one CAEP condition could be tested, which one would best predict behavior? A PLS regression demonstrated that the 70-dB signal/5-dB SNR CAEP condition was the best predictor of the behavioral SNR50. It will be important to validate the results of these predictions on other additional subjects and populations. It may be that the CAEP measures or conditions that correlate best with behavior depend on the population that is being tested.

CONCLUSION

CAEPs are sensitive to SNR changes rather than absolute signal level when signals are presented in background speech noise. In contrast, sentence perception in noise revealed sensitivity to both SNR and signal level. The reasons for this discrepancy are not clear; however, it may be due to specific sensitivities of the neural populations that produce CAEP activity. Despite these differences, strong correlations between CAEPs and sentence perception support the idea that electrophysiology may be used to predict perception-in-noise difficulties experienced by different populations. It will be important to determine the relationship between electrophysiology and behavior

in populations who experience difficulty understanding speech in noise such as those with hearing impairment or age-related deficits.

ACKNOWLEDGMENTS

We wish to thank Drs. Marjorie Leek, Robert Burkard, and Kelly Tremblay for the comments on the design of this experiment and earlier versions of this manuscript. This work was supported by a grant from the National Institute on Deafness and Other Communication Disorders (R03DC10914) and career development awards from the VA Rehabilitation Research and Development Service (C4844C and C8006W).

REFERENCES

- ADLER G, ADLER J (1989) Influence of stimulus intensity on AEP components in the 80- to 200-ms latency range. *Audiology* 28:316–324
- AKERÖYD MA, PATTERSON RD (1995) Discrimination of wideband noises modulated by a temporally asymmetric function. *J Acoust Soc Am* 98:2466–2474
- ANDERSON S, SKOE E, CHANDRASEKARAN B, KRAUS N (2010) Neural timing is linking to speech perception in noise. *J Neurosci* 30:4922–4926
- BALTZELL LS, BILLINGS CJ (2013) Sensitivity of offset and onset cortical auditory evoked potentials to signals in noise. *Clin Neurophys*. doi:10.1016/j.clinph.2013.08.003
- BENNETT K, BILLINGS CJ, MOLIS MR, LEEK MR (2012) Neural encoding and perception of speech signals in informational masking. *Ear Hear* 32:1–8
- BILLINGS CJ, TREMBLAY KL, SOUZA PE, BINNS MA (2007) Effects of hearing aid amplification and stimulus intensity on cortical auditory evoked potentials. *Audiol Neuro-Otol* 12:234–246
- BILLINGS CJ, TREMBLAY KL, STECKER C, TOLIN WM (2009) Human evoked cortical activity to signal-to-noise ratio and absolute signal level. *Hear Res* 254:15–24
- BILLINGS CJ, PAPESH MA, PENMAN TM, BALTZELL LS, GALLUN FJ (2012) Clinical use of aided cortical auditory evoked potentials as a measure of physiological detection or physiological discrimination. *Int J Otolaryngol* 2012:1–14
- COSTALUPES JA, YOUNG ED, GIBSON DJ (1984) Effects of continuous noise backgrounds on rate response of auditory nerve fibers in cat. *J Neurophysiol* 51:1326–1344
- DUBNO JR, SCHAEFER AB (1992) Comparison of frequency selectivity and consonant recognition among hearing-impaired and masked normal-hearing listeners. *J Acoust Soc Am* 91:2110–2121
- GIBSON DJ, YOUNG ED, COSTALUPES JA (1985) Similarity of dynamic range adjustment in auditory nerve and cochlear nuclei. *J Neurophysiol* 53:940–958
- GOLDSTEIN MH, HALL JL II, BUTTERFIELD BO (1968) Single-unit activity in the primary auditory cortex of unanesthetized cats. *J Acoust Soc Am* 43:444–456
- HASTIE T, TIBSHIRANI R, FRIEDMAN J (2009) The elements of statistical learning, 2nd edn. Springer Series in Statistics. Springer, New York
- HAWKINS JE, STEVENS SS (1950) The masking of pure tones and of speech by white noise. *J Acoust Soc Am* 22:6–13

- HILLYARD SA, HINK RF, SCHWENT VL, PICTON TW (1973) Electrical signs of selective attention in the human brain. *Science* 182:177–180
- HILLYARD SA, VOGEL EK, LUCK SJ (1998) Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Phil Trans R Soc Lond* 353:1257–1270
- HORNSBY BWY, TRINE TD, OHDE RN (2005) The effects of high presentation levels on consonant feature transmission. *J Acoust Soc Am* 118:1719–1729
- HYDE M (1997) The N1 response and its applications. *Audiol Neurotol* 2:281–307
- INSTITUTE OF ELECTRICAL AND ELECTRONIC ENGINEERS (1969) *IEEE Recommended Practice for Speech Quality Measures*. New York: IEEE.
- KAPLAN-NEEMAN R, KISHON-RABIN L, HENKIN Y, MUCHNIK C (2006) Identification of syllable in noise: electrophysiological and behavioral correlates. *J Acoust Soc Am* 120:926–933
- MOLIS MR, SUMMERS V (2003) Effects of high presentation levels on recognition of low- and high-frequency speech. *Acoust Res Lett Onl* 4:124–128
- NAATANEN R, PICTON T (1987) The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24:375–425
- NEUROSCAN, INC. (2007) *SCAN 4.4—Vol II, Edit 4.4: offline analysis of acquired data* (Document number 2203, Revision E, pp. 141–148). Compumedics Neuroscan, Charlotte
- PARBERY-CLARK A, MARMEL P, BAIR J, KRAUS N (2011) What subcortical–cortical relationships tell us about processing speech in noise. *Eur J Neurosci* 33:549–557
- PATTERSON RD (1994) The sound of a sinusoid: time-interval models. *J Acoust Soc Am* 96:1419–1428
- PHILLIPS DP (1990) Neural representation of sound amplitude in the auditory cortex: effects of noise masking. *Behav Brain Res* 37:197–214
- PHILLIPS DP, HALL SE (1986) Spike-rate intensity functions of cat cortical neurons studied with combined tone–noise stimuli. *J Acoust Soc Am* 80:177–187
- REES A, PALMER AR (1988) Rate-intensity functions and their modification by broadband noise for neurons in the guinea pig inferior colliculus. *J Acoust Soc Am* 83:1488–1498
- SCHWARTZ O, SIMONCELLI EP (2001) Natural signal statistics and sensory gain control. *Nat Neurosci* 4:819–825
- SKRANDIES W (1989) Data reduction of multichannel fields: global field power and principal component analysis. *Brain Topogr* 2:73–80
- STEVENS SS, GUIRAO M (1967) Loudness functions under inhibition. *Percept Psychophys* 2:459–465
- STUDEBAKER GA, SHERBECOE RL, MCDANIEL DM, GWALTNEY CA (1999) Monosyllabic word recognition at higher-than-normal speech and noise levels. *J Acoust Soc Am* 105:2431–2444
- SUMMERS V, MOLIS MR (2004) Speech recognition in fluctuating and continuous maskers: effects of hearing loss and presentation level. *J Speech Lang Hear R* 47:245–256
- WANG X, THOMAS L, SNIDER RK, LIANG L (2005) Sustained firing in auditory cortex evoked by preferred stimuli. *Nature* 435:341–346
- WHITING KA, MARTIN BA, STAPELLS DR (1998) The effects of broadband noise masking on cortical event-related potentials to speech sounds /ba/ and /da/. *Ear Hear* 19:218–231