

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/303811386>

Neural indices of phonemic discrimination and sentence-level speech intelligibility in quiet and noise: A mismatch...

Article in *Hearing Research* · June 2016

DOI: 10.1016/j.heares.2016.06.001

CITATIONS

5

READS

254

5 authors, including:



Tess K. Koerner

University of Minnesota Twin Cities

15 PUBLICATIONS 56 CITATIONS

SEE PROFILE



Yang Zhang

University of Minnesota Twin Cities

70 PUBLICATIONS 712 CITATIONS

SEE PROFILE



Peggy Nelson

University of Minnesota Twin Cities

53 PUBLICATIONS 584 CITATIONS

SEE PROFILE



Boxiang Wang

University of Minnesota Twin Cities

6 PUBLICATIONS 11 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Cracking the speech code: A cross-linguistic neurobehavioral approach to language learning in typical and atypical populations. [View project](#)



Investigating neural mechanisms for speech and voice processing deficits and plasticity in children with autism [View project](#)



Research paper

Neural indices of phonemic discrimination and sentence-level speech intelligibility in quiet and noise: A mismatch negativity study



Tess K. Koerner^a, Yang Zhang^{a, b, c, *}, Peggy B. Nelson^{a, c}, Boxiang Wang^d, Hui Zou^d

^a Department of Speech-Language-Hearing Sciences, University of Minnesota, Minneapolis, MN 55455, USA

^b Center for Neurobehavioral Development, University of Minnesota, Minneapolis, MN 55455, USA

^c Center for Applied Translational Sensory Science, University of Minnesota, Minneapolis, MN 55455, USA

^d School of Statistics, University of Minnesota, Minneapolis, MN 55455, USA

ARTICLE INFO

Article history:

Received 27 January 2016

Received in revised form

16 May 2016

Accepted 2 June 2016

Available online 4 June 2016

Keywords:

MMN

Time-frequency analysis

Theta band

Speech-in-noise perception

Linear mixed effects model

ABSTRACT

Successful speech communication requires the extraction of important acoustic cues from irrelevant background noise. In order to better understand this process, this study examined the effects of background noise on mismatch negativity (MMN) latency, amplitude, and spectral power measures as well as behavioral speech intelligibility tasks. Auditory event-related potentials (AERPs) were obtained from 15 normal-hearing participants to determine whether pre-attentive MMN measures recorded in response to a consonant (from /ba/ to /bu/) and vowel change (from /ba/ to /da/) in a double-oddball paradigm can predict sentence-level speech perception. The results showed that background noise increased MMN latencies and decreased MMN amplitudes with a reduction in the theta frequency band power. Differential noise-induced effects were observed for the pre-attentive processing of consonant and vowel changes due to different degrees of signal degradation by noise. Linear mixed-effects models further revealed significant correlations between the MMN measures and speech intelligibility scores across conditions and stimuli. These results confirm the utility of MMN as an objective neural marker for understanding noise-induced variations as well as individual differences in speech perception, which has important implications for potential clinical applications.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Speech communication often takes place in the presence of background noise, which can be difficult for hard of hearing listeners as well as many listeners with normal hearing. In recent years, there has been a surge of interest investigating noise-induced modulatory effects on cortical/subcortical responses to examine the neural networks and brain mechanisms supporting higher-level cognitive and linguistic skills (Anderson et al., 2010a; Billings et al., 2013; Du et al., 2014; Koerner and Zhang, 2015; Mesgarani et al., 2014; Vaden et al., 2015; Wong et al., 2008). Cortical auditory event-related potentials (AERPs) are one

representative method of measuring the neural coding of speech sounds in various listening conditions. In particular, the auditory mismatch negativity (MMN) response provides an objective electrophysiological measure of the neural timing and strength of pre-attentive auditory discrimination. It peaks at approximately 100–250 ms post-stimulus onset, which is typically generated when a participant's sensory memory trace of a "standard" stimulus detects a change by a less frequently occurring "deviant" stimulus in the absence of attention or any overt behavioral response (Näätänen et al., 2007). The present study attempts to address whether the MMN response is a good predictor of speech perception performance at both segmental and sentence levels in quiet and noise.

The pre-attentive cortical MMN response has been linked with behavioral speech perception in a number of studies. Representative topics include language learning and development in children (Kraus et al., 1996; Kraus and Cheour, 2000), native (Aaltonen et al., 1987; Christmann et al., 2014) and non-native speech perception in adults (Bidelman and Dexter, 2015; Brunellière et al., 2011; Näätänen et al., 1997; Winkler et al., 1999; Zhang et al., 2009), the

Abbreviations: EEG, electroencephalography; AERP, auditory event-related potential; MMN, mismatch negativity; SNR, signal-to-noise ratio; ANOVA, analysis of variance; SE, standard error; CV, consonant-vowel

* Corresponding author. 115 Shevlin Hall, Department of Speech-Language-Hearing Sciences, 164 Pillsbury Drive SE, University of Minnesota, Minneapolis, MN 55455, USA.

E-mail address: zhanglab@umn.edu (Y. Zhang).

<http://dx.doi.org/10.1016/j.heares.2016.06.001>

0378-5955/© 2016 Elsevier B.V. All rights reserved.

effects of hearing loss (Kraus et al., 1995b; Oates et al., 2002) and cochlear implantation (Groenen et al., 1996; Kraus et al., 1993, 1995b), and neural plasticity in auditory training (Kraus et al., 1995a; Tremblay et al., 1997; Tremblay et al., 1998; Zhang et al., 2009). Studies have shown that the MMN responses for phonetic discrimination in quiet can predict first- and second-language attainment (Garcia-Sierra et al., 2011; Jansson-Verkasalo et al., 2004; Jakoby et al., 2011; Kuhl et al., 2005; Molfese and Molfese, 1997). There is also evidence that pre-attentive speech perception in noise results in MMN amplitude decreases and latency increases when compared to quiet conditions (Kozou et al., 2005; Martin and Stapells, 2005; Martin et al., 1999; Muller-Gass et al., 2001). These noise-induced changes in the MMN response for detecting phonemic changes are associated with decrements in behavioral measures of discriminatory accuracy and increases in reaction time (Martin and Stapells, 2005; Martin et al., 1999; Muller-Gass et al., 2001). However, to our knowledge, no adult studies have examined brain-behavior relationships between changes in the pre-attentive MMN at the segmental level and performance on sentence-level word recognition performance across quiet and noise conditions.

In addition to the conventional ERP latency and amplitude measures, a recent trend in neurophysiological studies is the development of sophisticated time-frequency analyses to examine the role of various neural oscillation frequency bands of the EEG signal in the generation of AERP waveforms. These cortical oscillations are thought to modulate neural excitability and timing, which enables information exchange between cortical processes that are responsible for sensory and cognitive events (Başar et al., 1999; Klimesch et al., 2007; Koerner and Zhang, 2015; Makeig et al., 2004; Sauseng et al., 2007; Zhang et al., 2011). In particular, several studies have revealed the contribution of the theta frequency band (4–8 Hz) in driving the neuronal generation of the MMN in frontal and temporal areas (Bishop and Hardiman, 2010; Choi et al., 2013; Fuentemilla et al., 2008; Hsiao et al., 2009; Ko et al., 2012). Collectively, these studies show that neural generation of the MMN response are accompanied by phase alignment and power modulation of theta band activity. In the literature, the theta activity is proposed to be associated with several other cognitive functions including memory encoding, retrieval, and maintenance (Klimesch et al., 2008; Ward, 2003). Although previous studies have revealed prolonged latency and reduced amplitude in the MMN response due to the presence of background noise (Kozou et al., 2005; Martin and Stapells, 2005; Martin et al., 1999; Muller-Gass et al., 2001), it remains unknown how noise may modulate MMN spectral power in the theta band.

The present study aimed to investigate the effects of speech-babble background noise on the pre-attentive cortical processing of consonant and vowel changes by analyzing MMN latency, amplitude, and EEG spectral power measures. It is well established that the MMN responses can show high inter- and intra-subject variability in amplitude and latency (Kurtzberg et al., 1995; Lang et al., 1995; Martin et al., 1999, 2008; Näätänen et al., 2007; Näätänen, 1995; Stapells, 2002). This study was designed to take individual variability into account and investigate whether the objective neurophysiological measures in response to consonant and vowel phonemic contrasts in a double-oddball paradigm (Xi et al., 2010) can predict sentence-level speech intelligibility performance across quiet and noise listening conditions. The double oddball paradigm is a modified version of the conventional MMN protocol, in which the presentation trials for the single deviant stimulus are shared by two deviants (e.g., a consonant contrast and a vowel contrast) at equal probability of occurrence. Thus this paradigm allows the investigation of two MMN responses, one for each deviant stimulus, during the same recording session. Animal and human studies examining the neural processing of speech in

noise have revealed differential effects of noise on consonant and vowel stimuli, such that the neural responses to steady-state vowel stimuli are more robust in noise than those to more transient, aperiodic consonant stimuli (Cunningham et al., 2002; Russo et al., 2004; Shetake et al., 2011; Song et al., 2011). Additionally, behavioral and neurophysiological research suggests that consonant and vowel stimuli may be processed by separate neural mechanisms in the auditory cortex (Caramazza et al., 2000; Carreiras et al., 2009; Fogerty and Humes, 2012; Fogerty et al., 2012; Kewley-Port et al., 2007; Liberman and Mattingly, 1985; Miceli et al., 2004). Thus the use of a double-oddball paradigm would allow us to test the differences in neural sensitivity to vowel and consonant changes across the quiet and noise conditions and their relative contributions to higher-level behavioral performance in sentence recognition.

We hypothesized that the introduction of background noise would result in increases in MMN latency and decreases in MMN amplitude, which would be accompanied by reduced spectral power in the theta band. In addition, noise would differentially affect the cortical processing of the consonant and vowel changes, such that the pre-attentive detection of the consonant change would be more vulnerable to disruption in noise than the vowel change. We further hypothesized that at least some of the MMN measures would be able to predict higher-level behavioral sentence recognition.

2. Methods

2.1. Subjects

The participants in the study were 15 individuals (mean age = 22.6 years, age range = 19–32 years, 5 males, 10 females) with normal hearing (as shown in standard audiological assessment with hearing thresholds <25 dB HL for pure tones from 0.25 to 8 kHz) and no history for speech, language, or cognitive difficulties. All participants were right handed and were native speakers of American English. The Human Research Protection Program at the University of Minnesota approved the research protocol and all participants provided informed consent prior to beginning the study.

2.2. Stimuli

2.2.1. Stimuli for ERP measures

The consonant-vowel (CV) syllables, /ba/, /da/, and /bu/, were synthesized with the Hlsyn software program (Sensimetrics Corporation, USA) using a 10 kHz sampling rate (Koerner and Zhang, 2015). All the syllables were 170 ms in duration with a steady fundamental frequency of 100 Hz and a steady F4 at 3300 Hz. The Hlsyn software generated formant transitions in the first 50 ms of the CV syllables with onset frequencies at 328 Hz, 1071 Hz, and 2298 Hz respectively for F1, F2, and F3 of the /ba/ sound. For /da/, the F1, F2, and F3 onset frequencies were 362 Hz, 1832 Hz, and 2540 Hz, and for /bu/, the formant onset frequencies were at 230 Hz, 900 Hz, and 2480 Hz. The steady center F1, F2, and F3 frequencies for the vowel portion (50–170 ms) of the /ba/ and /da/ syllables were 674 Hz, 1140 Hz, and 2350 Hz. The steady center F1, F2, and F3 frequencies for the vowel portion of /bu/ were 320 Hz, 860 Hz, and 2620 Hz, respectively. The background noise used in this study was a four-talker speech babble noise that was adopted from the Quick Speech In Noise Test (Quick-SIN) (Niquette et al., 2001). All of the CV syllables and the noise stimuli were resampled at 44.1 kHz and were normalized to create a –3 dB SNR using Sony SoundForge 9.0 (Sony Creative Software, USA) (Koerner and Zhang, 2015).

2.2.2. Stimuli for behavioral measures

The CV syllables were utilized in an active listening condition (described below) in order to obtain percent correct phoneme change-discrimination as well as response reaction time. Participants also listened to IEEF sentences (IEEF, 1969) in both the quiet and four-talker speech babble masker conditions to obtain sentence recognition scores.

2.3. Procedure

All testing was conducted in an electrically and acoustically treated booth (ETS-Lindgren Acoustic Systems). The reported EEG data were taken from a larger scale study with a two-hour recording session including both passive and active listening conditions (Koerner and Zhang, 2015; Koerner et al., 2013). The CV syllable stimuli were presented using Evoke software (ANT Inc., Netherlands) via bilateral Etymotic ER-2 insert headphones. The speech signal was presented at 60 dB SL relative to the individual participants' hearing thresholds at 1 kHz (Koerner and Zhang, 2015). Participants were presented with two different conditions: signals in quiet and signals in a four-talker speech babble noise at a -3 dB SNR in both EEG and behavioral tests.

Stimulus presentation order in the double-oddball paradigm was pseudo-randomized so that no two deviants were presented in succession and that no blocks began with a deviant stimulus. Two speech contrasts were included in the double-oddball paradigm: a vowel change (from /ba/ to /bu/) and a consonant change (from /ba/ to /da/). The vowels /a/ and /u/ are both back vowels, which differ primarily in F1 in the steady vowel portion (Ladefoged, 2006). On the other hand, the /b/ and /d/ consonants, represent transient differences in place of articulation cued by second and third formant frequency transitions, which are in a frequency range commonly affected by hearing impairment (Ladefoged, 2006; Miller and Nicely, 1955). The standard stimulus /ba/ had a probability of occurrence of 0.75 and the two deviant stimuli, /da/ and /bu/, each had a probability of occurrence of 0.125. Each condition consisted of 10 blocks for a total of 832 trials for standard stimuli and 104 trials for each deviant stimulus. Both /da/ and /bu/ were presented as standard stimuli in 4 separate alternating blocks of 30 repetitions for a total of 120 repetitions of each stimulus. Neural responses to these "standard" presentations of /da/ and /bu/ were subtracted from the same syllables when they were presented as deviant stimuli in the double-oddball paradigm. Kraus et al. (1995a,b) advocated the use of a "deviant alone" recording session when there might be large acoustic confounds (in our case, consonant vs. vowel contrasts) in the interpreting the MMN data. This method of ERP subtraction is used to obtain the "identity MMN," which controls for potential acoustic confounds from simple deviant-minus-standard subtraction (Zhang et al., 2005; Pulvermüller and Shtyrov, 2006). The interstimulus interval (ISI) was randomized from trial to trial within the range of 600–700 ms. The inter-block interval was 5 s. The quiet and noise conditions were counter-balanced across subjects to reduce potential test order effects. During the MMN recording session, participants were instructed to relax, minimize excessive movements (including blinking), and stay awake while watching a muted movie of their choice with subtitles and ignoring the auditory stimuli played through the headphones.

Behavioral responses for syllable detection were recorded from a separate discrimination task using the same trials and presentation levels as in the double-oddball paradigm. Unlike the MMN recording session, no muted movie was presented. Instead, the participants were asked to focus their attention on detecting sound changes and make a button-press response using a desktop keyboard whenever they heard a deviant stimulus. This was

followed by an additional 30-min behavioral session where speech intelligibility responses to randomized lists of IEEF sentences presented through TDH-39 headphones were recorded. During this speech intelligibility test session, participants were instructed to repeat out loud word by word, as best as they could, the IEEF sentences they heard. Their word-by-word responses were then evaluated by a certified audiologist for recognition accuracy.

2.4. Data analysis

2.4.1. ERP measures

Continuous EEG data were recorded using the Advanced Neuro Technology EEG System (Advanced Source Analysis version 4.7) and a 64-channel Ag AgCl electrode WaveGuard cap with a REFA-72 amplifier (TMS International BV) (bandwidth = 0.016–200 Hz, sampling rate = 512 Hz). The average impedance of electrodes was below 5 kohms. ERP waveform analysis was completed offline in BESA (Version 6.0, MEGIS Software GmbH, Germany). The EEG data were bandpassed at 0.5–40 Hz. The ERP epoch length consisted of a 100 ms prestimulus baseline and a 700 ms poststimulus interval. Automatic artifact rejection criteria were set at ± 50 μ V. The MMN was analyzed with a common average reference at the Cz electrode. Based on the grand average waveforms in the quiet and noise conditions, MMN peak latency, relative to the pre-stimulus interval, was assessed within the time window of 100–300 ms. The MMN amplitude quantification used an integration (averaging) window of 10 ms centered at peak (Rao et al., 2010).

Spectral power analysis was completed using the subtracted MMN waveform at electrode Cz with the *newtimef* function in EEGLAB (Delorme and Makeig, 2004). A short-term Fourier Transform (STFT) with Hanning window tapering (Koerner and Zhang, 2015), which is recommended for the analysis of low frequency activities, was adopted to extract the logarithmic spectra power for the theta band. The modified STFT method in EEGLAB used overlapping sliding windows that are adapted to the target frequency bins to overcome limitations due to the use of fixed windows in conventional analysis. Zero-padding was applied for short epochs with insufficient sample points for the Fourier transform. The time window used for our time-frequency analysis represented the entire analysis epoch, including the pre-stimulus baseline from -100 to 700 ms, and estimated frequencies were from 0.5 to 40 Hz with a step interval of 0.5 Hz (Koerner and Zhang, 2015).

All statistical analyses from both the AERP and behavioral portions of the study were analyzed in R (R Core Team, 2014). A repeated measures analysis of variance (RM-ANOVA), with $\alpha = 0.05$, was conducted to examine the statistical significance of stimulus type (/da/ or /bu/), condition (quiet or noise), and any potential interactions (stimulus \times condition) on MMN latency, amplitude, and spectral power recorded at electrode Cz. Where significant interaction effects were observed, tests of simple main effects and post hoc two-tailed t-tests for selected factors of interest were also conducted to evaluate how consonant and vowel changes were processed differently in quiet and in noise.

2.4.2. Behavioral measures

In the phoneme discrimination test, the percent correct scores and reaction time for the detection of consonant and vowel changes in the double-oddball paradigm were obtained from the button-press responses recorded during the quiet and noise conditions. A RM-ANOVA was completed to investigate the statistical significance of stimulus type (/da/ or /bu/), condition (quiet or noise), and any potential interactions (stimulus \times condition) on behavioral discrimination accuracy and reaction time.

In the speech intelligibility test, the IEEF lists consisted of 10 low context sentences with 5 key words in each sentence. For both the

quiet and noise conditions, two sentence lists of 50 key words were presented from one female and one male talker for a total of 100 key words per listening condition. Participants repeated key words from 4 randomized lists of 10 behavioral IEEE sentences (IEEE, 1969) and percent correct scores for key words were recorded. An additional RM-ANOVA was carried out to examine the significance of listening condition (quiet vs. noise) on IEEE sentence recognition.

2.4.3. Linear mixed-effects model analysis

Using R (R Core Team, 2014) and the *nlme* package (Pinheiro et al., 2016), linear mixed-effects models were developed to examine whether the MMN measures were predictive of behavioral speech perception at both the syllable- and sentence-level. LME models are extension of linear regression models, which are particularly useful in settings involving repeated measures with coefficients that can vary with respect to one or more grouping variables. Regression analyses have previously been used to examine whether combinations of ERP data collected from individual subjects are viable predictors of behavioral speech perception across different listening conditions (Billings et al., 2013; Billings et al., 2015). Data transformations included re-scaling MMN latency and behavioral reaction time values. Additionally, the behavioral percent correct phoneme detection and sentence recognition variables were log transformed to account for skewness in the data that was observed during visual analysis.

Participants were used as a “by-subject” random effect in each linear mixed-effect model. Listening condition (quiet vs. noise) and stimulus (/bu/ vs. /da/) were included as blocking variables. We then added MMN latency, amplitude, and theta power as fixed effects in order to predict percent correct phoneme detection and reaction time. Similar models were developed to evaluate whether the MMN was able to predict sentence-level perception. In implementing the LME models, we first transformed the original data to alleviate large skewness and difference in magnitudes. The significance of each fixed effect in predicting each behavioral outcome measure was assessed with $\alpha = 0.05$.

3. Results

Analysis of behavioral data revealed noise-induced increases in reaction time for phoneme discrimination as well as reductions in percent correct phoneme discrimination and sentence recognition (Table 1). In the ERP analysis, the presence of speech-babble background noise increased MMN latency and decreased MMN amplitude (Fig. 1 and Table 2), which was accompanied by decreased MMN spectral power in the theta frequency band (Table 2). Linear mixed-effects regression analysis revealed several significant neural predictors of behavioral perception at the phoneme and sentence-level (Table 3).

Table 1

Means (standard error) for behavioral percent correct detection of vowel (/bu/) and consonant (/da/) changes, behavioral reaction time for vowel (/bu/) and consonant (/da/) changes, as well as percent correct behavioral sentence recognition performance.

Behavioral measure	Listening condition	
	Quiet	Noise
/bu/ detection (%)	98.20 (0.6)	91.79 (2.22)
/da/ detection (%)	92.68 (1.63)	34.05 (4.03)
Sentence recognition (%)	99.27 (0.37)	67.53 (3.77)
/bu/ response time (ms)	468.04 (10.01)	551.51 (16.83)
/da/ response time (ms)	510.98 (12.61)	621.91 (12.10)

3.1. Behavioral phoneme discrimination and sentence recognition

RM-ANOVA results revealed that the presence of background noise significantly prolonged behavioral reaction time during phoneme discrimination ($F(1,14) = 41.4$, $p < 0.001$) (Table 1). Additionally, reaction time was significantly shorter during behavioral discrimination of the vowel change (from /ba/ to /bu/) compared to the consonant change (from /ba/ to /da/) ($F(1,14) = 24.95$, $p < 0.001$). Performance accuracy in noise was significantly poorer compared to the quiet listening condition ($F(1,14) = 179.1$, $p < 0.001$). As expected, there was also a significant stimulus effect with better accuracy for detecting the vowel change ($F(1,14) = 232.3$, $p < 0.001$). A significant interaction between listening condition and stimulus indicated a differential effect of noise on the two deviant stimuli ($F(1,14) = 223.2$, $p < 0.001$) with a larger noise-induced decline in behavioral performance for detecting the consonant change. Post hoc comparisons of the two speech syllables revealed that the introduction of background noise significantly decreased percent correct detection of /da/ ($t(14) = 15.50$, $p < 0.001$) and had a smaller, but still significant impact on the detection of /bu/ ($t(14) = 3.37$, $p < 0.01$). Noise-induced reduction in performance was also seen in behavioral sentence recognition ($F(1,14) = 72.15$, $p < 0.001$).

3.2. Amplitude and latency measures for averaged MMN response

RM-ANOVA results showed significant effects of speech babble background noise on MMN latency ($F(1,14) = 29.43$, $p < 0.001$) and amplitude ($F(1,14) = 32.52$, $p < 0.001$) (Table 2 and Fig. 1). The presence of noise led to a significant increase in the MMN latency and decrease in amplitude during passive speech perception. Additionally, there were significant differences in latency ($F(1,14) = 17.84$, $p < 0.001$) between the two deviant syllables /da/ and /bu/ across the quiet and noise conditions. The MMN for /da/ peaked later than that for /bu/. A significant interaction effect was found between stimulus and condition for the MMN amplitude measure ($F(1,14) = 18.77$, $p < 0.001$), indicating that noise had a differential effect on the neural processing of the consonant and vowel change. Further t-tests showed that a large reduction in MMN amplitude occurred for the AERP recorded in response to the more salient CV syllable /bu/ ($t(14) = -6.30$, $p < 0.001$) but not in response to /da/ ($t(14) = -0.65$, $p = 0.528$). Conversely, significant increases in MMN latency occurred for both the CV syllable /bu/ ($t(14) = -3.18$, $p < 0.01$) and /da/ ($t(14) = -4.97$, $p < 0.001$). In addition, MMN responses to /da/ and /bu/ in quiet were significantly different in terms of amplitude ($t(14) = 3.61$, $p < 0.01$) and latency ($t(14) = 3.92$, $p < 0.01$).

3.3. MMN spectral power

As predicted, MMN spectral power in the theta band was significantly reduced in background noise compared to the quiet condition across the two deviant CV stimuli ($F(1,14) = 19.37$, $p < 0.001$) (Table 2). There was also a main effect of stimulus with /da/ showing significantly lower theta power than /bu/ ($F(1,14) = 11.21$, $p < 0.01$). A significant stimulus by condition interaction was found ($F(1,14) = 10$, $p < 0.01$). Further t-tests showed a large reduction in response to the CV syllable /bu/ in noise ($t(14) = 5.60$, $p < 0.001$) but not in response to /da/ ($t(14) = 0.45$, $p = 0.661$). In addition, theta power in response to /da/ was significantly different from that of /bu/ in quiet ($t(14) = -4.33$, $p < 0.001$).

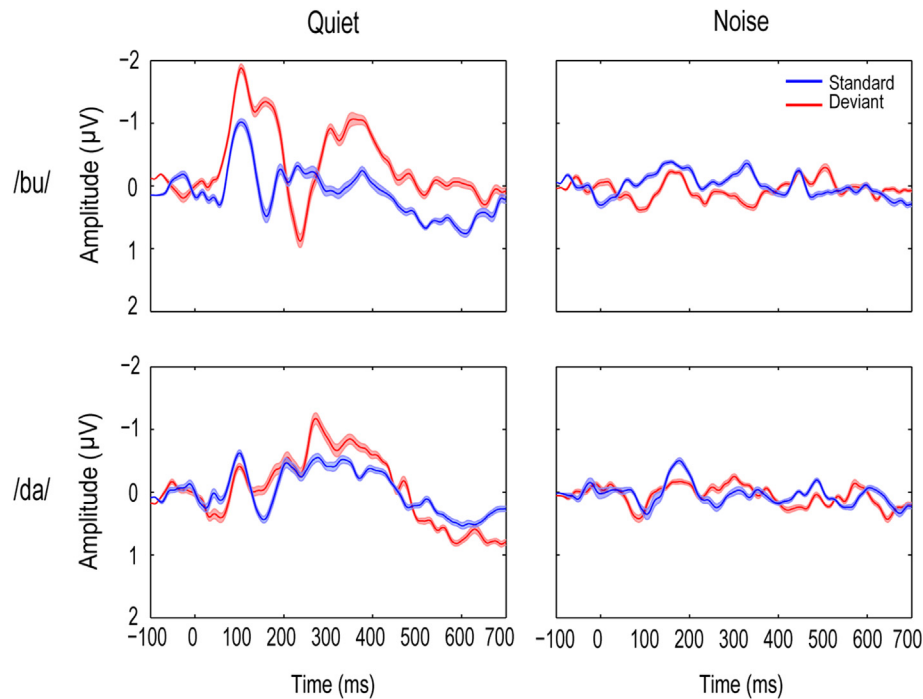


Fig. 1. Grand mean AERP waveforms averaged across participants in response to the standard (blue) and deviant (red) stimuli (top row: /bu/, bottom row: /da/) for quiet and noise conditions at electrode Cz with shaded standard error envelopes. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.4. Linear mixed-effects model results

Linear mixed-effects regression analysis revealed that MMN latency ($F(1,40) = 7.86$, $p < 0.01$) and theta power ($F(1,40) = 6.61$, $p < 0.05$) were significant predictors of behavioral phoneme detection accuracy across conditions and stimuli. Additionally, MMN amplitude for phoneme detection showed a trend of approaching significance as a predictor of behavioral accuracy ($F(1,40) = 3.10$, $p = 0.086$). Linear mixed-effects regression analysis also revealed that theta power was significantly correlated with MMN latency ($F(1,42) = 5.61$, $p < 0.05$) and amplitude ($F(1,42) = 11.28$, $p < 0.01$). In contrast, there was no significant correlation between any of the MMN measures and the behavioral reaction time data (MMN latency ($F(1,40) = 0.00$, $p = 0.996$),

amplitude ($F(1,40) = 0.00$, $p = 0.961$), and theta power ($F(1,40) = 0.37$, $p = 0.547$)). In the analysis for behavioral sentence-level scores, the MMN amplitude in response to /bu/ was the only significant predictor of behavioral performance ($F(1,11) = 7.21$, $p < 0.05$) (see Table 3 for a summary of regression model outputs for each behavioral outcome measure).

As the residual plots of all models showed no sign of any significant trend or heteroscedastic variance, we do not expect any potential improvement from the use of non-linear models. As all the residuals appear normally distributed, we also do not expect better results (in terms of finding more or stronger brain-behavior correlations in the current data set) from generalized linear models.

Table 2
MMN mean latency (ms), amplitude (μV), and power in the theta band (dB) (standard error) in response to the CV syllable vowel change (/bu/) and consonant change (/da/) at electrode Cz in quiet and in noise.

	Latency (ms)		Amplitude (μV)		Spectral power (dB)	
	Quiet	Noise	Quiet	Noise	Quiet	Noise
/bu/	171.88 (10.53)	236.98 (14.80)	-2.11 (0.30)	-0.40 (0.12)	-10.77 (1.30)	-19.53 (1.46)
/da/	228.26 (13.79)	293.10 (11.45)	-1.16 (0.20)	-1.02 (0.18)	-17.79 (1.28)	-18.57 (1.55)

Table 3
F-statistics for fixed effects from linear mixed-effects regression models for each behavioral measure.

Variable	Percent correct phoneme detection	Phoneme detection reaction time	Percent correct sentence recognition (/bu/)	Percent correct sentence recognition (/da/)
Intercept	161.51***	4199.98***	431.41***	335.12***
Condition	131.68***	61.92***	291.32***	247.69***
Stimulus	114.20***	21.05***	–	–
Latency	7.86**	0.000	1.24	0.44
Amplitude	3.10	0.002	7.21*	0.41
Theta power	6.61*	0.368	0.46	1.50

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, $p < 0.10$.

4. Discussion

This study was designed to examine the effects of background noise on the MMN recorded in response to a consonant versus vowel change and to determine whether noise-induced variations in these objective cortical measures are able to predict segmental- and sentence-level behavioral speech recognition in noise.

4.1. Noise-induced decreases in behavioral accuracy and increases in response time

Our overall results are consistent with previous studies that have examined the effects of noise on the behavioral perception of consonant and vowel stimuli (Miller and Nicely, 1955; Parikh and Loizou, 2005; Phatak and Allen, 2007; Pickett, 1957). Although all behavioral measures were negatively impacted by the introduction of background noise, our data showed consistently lower performance for detecting a consonant change across the two listening conditions. It is well known that the important perceptual cues contained in consonants differ from those in vowels (Ladefoged, 2006). Consonants tend to be aperiodic, weaker in acoustic energy, and have faster formant transitions that may induce more susceptibility to the deleterious effects of noise compared to more salient and sustained vowel cues. Parikh and Loizou (2005) reported that even at poor SNRs (signal-to-noise ratios), vowels still had relatively intact F1 cues and partial F2 cues available for vowel identification, and although performance was still high in noise, several features known to be associated with consonant place of articulation identification were significantly impacted.

This difference in acoustic susceptibility to noise masking implies that the relative importance of consonant and vowel categories to overall speech perception may shift in the presence of noise. It has been shown that low- and high-frequency acoustic landmarks in consonant sounds provide useful information for word boundary segmentation and accurate sentence recognition, and that disruption of these landmarks by background noise produces errors in the perception of voicing and place of articulation cues (Li and Loizou, 2008). In fact, when Owren and Cardillo (2006) replaced consonant or vowel segments with silence in a same/different word-meaning task, they found that listeners were better able to perceive words with consonant-only information compared to vowel-only stimuli. However, several studies have shown that vowels tend to contribute more than consonants to sentence intelligibility (Cole et al., 1996; Fogerty and Kewley-Port, 2009; Kewley-Port et al., 2007), potentially due to contributions from amplitude envelope and temporal fine structure cues contained within vowel segments in meaningful, sentence-level contexts compared to word recognition tasks (Fogerty and Humes, 2012). For instance, when consonant or vowel segments were removed from sentence stimuli and replaced with noise, both Cole et al. (1996) and Kewley-Port et al. (2007) found superior sentence recognition performance when listeners had access to vowel-only information compared to consonant-only information, but showed no differences in performance on a word intelligibility task (Fogerty and Humes, 2010). Although these studies did not compare quiet and noise listening conditions, they revealed that consonants and vowels might play different roles in speech understanding under different linguistic contexts. Examining how background noise affects the processing of consonants and vowels and how these differential effects are reflected at the cortical level is important for understanding underlying causes of variability and decreased performance during speech perception in noise.

4.2. Differential effects of noise on the neural processing of consonants and vowels

Consistent with the behavioral results, our neurophysiological data revealed that when stimuli were presented in a double-oddball paradigm in quiet, the MMN responses to the consonant change were weaker than those to the vowel change. It is known that measures of neural processing time and magnitude of cortical activation in response to a target stimulus are dependent on the magnitude of deviation from the standard stimulus (Pakarinen et al., 2013; Sams et al., 1985). Since our behavioral results showed significantly poorer performance for detection of the consonant change compared to the vowel change, a possible straightforward explanation for the same patterns in the behavioral data and MMN data in quiet is simply that the /ba/-/da/ contrast represented a smaller, or more subtle, acoustic change than the competing /ba/-/bu/ contrast.

However, the acoustics-based explanation has difficulty in accounting for noise-induced MMN changes, which appeared to show a different pattern from the behavioral data. While the overall pattern of noise-induced latency increase and amplitude decrease in MMN to both deviants in our study replicated previous findings (Kozou et al., 2005; Martin and Stapells, 2005; Martin et al., 1999; Muller-Gass et al., 2001), the background noise appeared to affect the detection of the vowel change more than the detection of the consonant change. Similar results were provided by Niemitalo-Haapola et al. (2015), who found a larger noise-induced MMN reduction in response to /e/-/i/ vowel changes, such that no significant MMN for a vowel change was recorded in noise, compared to a reduced, yet still present MMN to /p/-/k/ consonant changes in normally developing toddlers. Why would the background noise have a greater influence on the MMN for the more salient speech contrast? A possible explanation is that the larger noise-induced effects on pre-attentive neural processing of the more salient vowel cue could arise due to the need to internally resolve consonant information in background noise when consonant and vowel changes are juxtaposed in a double-oddball paradigm. This could imply that information contained in consonant segments might contribute more to speech perception in adverse listening environments. However, there are other important factors to consider if we extend the interpretation to sentence-level processing in noise as the relative importance of consonant and vowel cues may change depending on the level of lexical information available (Fogerty and Humes, 2010, 2012).

As we did not include nonspeech control stimuli in the current study, we could not rule out the possibility that the differences between vowel and consonant processing that we observed might reflect general auditory mechanisms of acoustic processing rather than phonemic processing. Moreover, as the current study tested only one consonant contrast and one vowel contrast, it remains to be tested whether the same phenomena would hold for other vowels/consonants. The advantage of the double oddball paradigm in the present study is that it allows us to directly compare how noise affects neural sensitivity to consonant and vowel contrasts at the pre-attentive level. Previous behavioral studies have shown the relative importance of consonant and vowel segments for speech intelligibility (Cole et al., 1996; Fogerty and Humes, 2010, 2012; Fogerty and Kewley-Port, 2009; Kewley-Port et al., 2007; Owren and Cardillo, 2006). Brain research further indicates that consonants and vowels are processed by distinct neural mechanisms (Caramazza et al., 2000; Carreiras et al., 2009; Carreiras and Price, 2008). Although differential effects of noise on speech sounds have been found in neurophysiological studies (Anderson et al., 2010b; Cunningham et al., 2002; Cunningham et al., 2001; Russo et al., 2004; Song et al., 2011), these studies only examined

responses to the consonant and vowel portion within a single CV syllable. For instance, while examining the brainstem frequency following response (FFR) to the CV syllable /da/, Russo et al. (2004) found that noise caused greater disruption of the transient portion of the neural response, which reflects coding of rapidly changing stimulus features that are characteristic of consonant sounds. Additionally, previous ERP studies that examined the effect of noise on speech processing only used one phonemic contrast in an oddball paradigm, representing either a single consonant change or a vowel change (Bennett et al., 2012; Kozou et al., 2005; Martin et al., 1997; Martin and Stapells, 2005; Martin et al., 1999; Whiting et al., 1998; Muller-Gass et al., 2001). While the use of new multi-feature recording paradigms with several juxtaposed “deviant” stimulus-changes allow for a comparison of MMN responses to deviant consonant and vowel changes in direct competition (Näätänen et al., 2004; Pakarinen et al., 2013; Pakarinen et al., 2007), these studies typically only tested stimuli in quiet. The results showed that consonant changes were more difficult to detect than vowel changes (Pakarinen et al., 2009), which were replicated in our quiet condition.

Overall, the MMN amplitude and latency results may suggest that consonants and vowels play different roles in speech processing, which has both theoretical and practical implications for understanding theories of speech perception and developing strategies for improving speech understanding in noise. For instance, performance may be improved by making some speech sound features more accessible through rehabilitation options such as amplification from hearing aids or by training listeners to attend to important cues contained in certain speech segments. These strategies could be tested using neural measures to examine whether reductions in noise-induced cortical effects are observed during speech perception in noise after rehabilitation or training.

4.3. *Theta power modulation and speech perception in noise*

Although it has already been established that event-related cortical oscillations in the theta frequency band are associated with linguistic processing of phonemic contrasts (Jin et al., 2014), little is known about the effects of noise on theta power modulation during pre-attentive speech discrimination. Our results are consistent with previous findings that reveal the importance of the theta frequency band in the pre-attentive neural processing of auditory deviant stimuli in quiet (Choi et al., 2013; Fuentemilla et al., 2008; Ko et al., 2012; Hsiao et al., 2009, 2010). Furthermore, we demonstrated that noise significantly decreased theta power in response to both speech stimuli.

Research examining neurocognitive linguistic development in infants has also shown that measures of EEG spectral power modulation are sensitive to stimulus features (Radicevic et al., 2008; Santesso et al., 2007; Zhang et al., 2011). For instance, Zhang et al. (2011) found that when infants passively listened to the vowel /i/ in alternating blocks of exaggerated infant-directed or unexaggerated adult-directed speech in quiet, ERPs were enhanced and EEG spectral power in the theta band was stronger in response to the more prominent vowel. Our theta power modulation data are consistent with these findings, showing stronger spectral power for the salient vowel change in quiet. In the meanwhile, the spectral power of theta band also showed that neural processing of the more prominent vowel change had larger noise-induced effects in comparison with the consonant change.

4.4. *The MMN as a neurophysiological marker of behavior*

We chose to use a -3 dB SNR in this study to purposely evoke a wide range of performance on the speech perception tasks in order

to examine brain-behavior correlations. Pilot data in our lab (Koerner et al., 2013) revealed that this noise level induced a range of performance in the detection of the easier vowel change and more difficult consonant change, as well as sentence recognition, without ceiling or floor effects. Based on previous studies showing correlations between speech-evoked neural responses and behavioral speech perception in noise (Anderson et al., 2010a, 2010b, 2011; Bennett et al., 2012; Billings et al., 2013, 2015; Cunningham et al., 2001; Martin and Stapells, 2005; Martin et al., 1999; Muller-Gass et al., 2001; Song et al., 2011), we presumed that this range of behavioral performance would be reflected in measures of pre-attentive cortical speech processing.

As expected, our results confirmed that the MMN could serve as a neurophysiological predictor of behavioral speech perception at both the syllable level and sentence level. Our data are consistent with previous studies that have examined relationships between the speech-evoked MMN and behavioral performance at the segmental level (Martin and Stapells, 2005; Martin et al., 1999; Muller-Gass et al., 2001). Using a /ba/-/da/ syllable contrast, Martin et al. (1999) examined the effects of decreased audibility on the MMN in normal hearing listeners using different high-pass noise filters and found significant correlations between noise-induced variations in behavioral phoneme-change detection sensitivity and MMN amplitude as well as behavioral reaction time and MMN latency. Our results add to this current body of knowledge by providing information about the ability of the MMN to reflect noise-induced variability in behavioral performance when measured in response to a both consonant and a vowel change, and more importantly, how it relates to sentence-level performance.

Although MMN latency and theta power were significant predictors of percent correct phoneme detection, none of these variables significantly predicted behavioral reaction time for phoneme detection. Martin et al. (1999) reported a significant correlation between MMN latency and behavioral reaction time; however, they reported that the correlation was weak, especially when compared to the relationship between MMN amplitude and behavioral sensitivity. These results imply that the MMN is accurate in assessing behavioral phoneme detection accuracy, but may not provide strong predictive information about timing of the conscious decision-making process. This may be explained by the fact that the MMN is an obligatory response that reflects pre-attentive auditory discrimination; it is expected that the presence of an MMN would be associated with the ability to behaviorally discriminate between stimuli. However, participants may differ in how conservative or liberal they are in making a response during an auditory discrimination task, which would greatly impact behavioral reaction time. Therefore, the active P3 response, which is thought to reflect completion in the decision making process (Picton, 1992), may be a better indicator of behavioral reaction time than the pre-attentive MMN (Martin et al., 1997; Whiting et al., 1998).

While MMN amplitude in response to /bu/ (i.e., the vowel change) was a significant predictor of sentence-level speech intelligibility scores, MMN latency, amplitude, and theta power in response to /da/ (i.e., the consonant change) were not. The inability of the MMN in response to the consonant change to predict behavioral sentence recognition could be due to the lack of significant differences between the quiet and noise listening conditions in MMN amplitude or theta power.

Our analysis revealed that theta power was a significant predictor of both MMN latency and amplitude. Similarly, significant correlations were found between theta power and MMN amplitude in studies examining MMN attenuation in clinical populations with schizophrenia using tonal stimulus contrasts (Hong et al., 2012; Kaser et al., 2013). These relationships suggest that our noise-

induced MMN changes were mediated by the strength of synchronization of event-related oscillations during pre-attentive discrimination of speech contrasts. Our data confirm that using time-frequency analysis to obtain a measure of theta power represents an additional tool for examining speech-evoked MMN and its susceptibility to noise.

4.5. Novelty, limitations and future directions

To our knowledge, this is the first study to examine whether noise-induced changes in the MMN reflect variability in higher-level speech perception across quiet and noise conditions. Previous studies have examined relationships between AERPs and behavioral sentence recognition in noise (Bennett et al., 2012; Bidelman and Howell, 2016; Billings et al., 2013, 2015; Parbery-Clark et al., 2011), but none have examined pre-attentive auditory change-detection as a neurophysiological predictor of behavior both in quiet and in noise. Novel results from linear mixed-effects models showed that both MMN latency and amplitude may reflect variability in behavioral performance. Additionally, this study provided novel information regarding the ability of theta power modulation across quiet and noise conditions to predict behavioral speech perception abilities, which suggests that phase alignment and/or spectral power modulation of the theta oscillation within a neural population for the generation of the MMN response can be reflected in behavioral performance. These results imply that the theta spectral power measure represents an additional tool for predicting the effects of noise on behavioral speech perception. The use of event-related cortical oscillations allows for a deeper examination of underlying neural processes that are implicated in noise, which possibly contribute to the wide range of variability seen in measures of speech perception in noise across listeners.

These novel results have implications for the clinical utility of the MMN as an obligatory cortical measure of more “real world” speech perception abilities, as it appears as if pre-attentive cortical measures can be used to predict sentence-level behavioral performance across stimuli and conditions. The pre-attentive MMN and measures of event-related cortical oscillations could prove useful in examining speech-processing abilities in clinical populations who are unable to provide consistent or reliable behavioral responses, such as adults with cognitive impairments or infants. For instance, pediatric hearing aid fittings may employ a measure of pre-attentive cortical auditory discrimination to examine pre- to post-fitting improvements in speech processing at the phonemic level. Additionally, pre-attentive cortical responses at the segmental level could be compared pre- and post-auditory training not only to assess the success of the program, but also to predict relative improvements in higher-level speech-in-noise perception in adults. In order to determine the reliability of the MMN as a neurophysiological correlate of behavioral perception at the individual level and further the practical field, additional stimuli, noise conditions, and participant populations, such as those with hearing loss, cochlear implants, or auditory processing disorders, should be tested to examine whether these significant effects generalize to other listening situations. Moreover, since our data showed lower variance for MMN amplitude and theta power values to the consonant change relative to the vowel change, further assessments should evaluate the predictive power of the MMN for a consonant change that shows greater variance in noise. This would add to our theoretical knowledge by allowing for an evaluation of whether neural responses to the consonant or vowel change are better predictors of individual speech performance, which may contribute to a better understanding of the relative contributions of consonant and vowel

information in speech perception across different listening contexts.

A current barrier to using the MMN for any clinical assessment of speech processing is the wide range of variability in individual responses even when behavioral performance is at a fixed level (Kurtzberg et al., 1995; Lang et al., 1995; Martin et al., 1999, 2008; Näätänen et al., 2007; Näätänen, 1995; Stapells, 2002). Future research may evaluate whether the spectral power analysis for targeted cortical rhythms would provide a more robust measure than MMN latency or amplitude, which can have strong practical implications for assessing event-related cortical oscillations in clinical populations with speech perception difficulties and potentially monitoring intervention outcomes.

Since our behavioral auditory change-discrimination data showed differences in the effects of noise on consonant and vowel perception compared to the pre-attentive cortical responses, future research should examine how noise impacts change detection in an active listening condition in the double-oddball paradigm by analyzing the P3 responses. This would help determine the roles that attention and listening context play in the neural processing of consonants and vowels in adverse listening conditions. Additionally, the present study only used three CV speech stimuli, representing only one consonant and one vowel change. Due to time constraints and study design, only one noise condition was tested. It is necessary to test additional stimuli and noise levels to determine how these two speech sound classes are affected by background noise at different SNRs. It is possible that the noise level may cause shifts in differential neural processing of consonant and vowel stimuli depending on attentional demand. Additionally, non-speech acoustic control stimuli (e.g., Zhang et al., 2005) should be examined to determine whether the MMN data as reported in the current study truly reflected consonant and vowel processing or general auditory mechanisms of detecting distinct acoustic changes. Furthermore, data from hearing-impaired listeners should be evaluated, as these listeners most often have reports of difficulty listening to speech in noise even after audibility is improved via amplification from hearing aids.

5. Conclusion

This study aimed to determine whether variations in cortical speech-evoked MMN latency, amplitude, and spectral power measures in response to a consonant and vowel change could predict behavioral speech perception abilities at both syllable and sentence levels across quiet and noise conditions. Results were consistent with our predictions that the introduction of background noise would increase MMN latencies as well as decrease MMN amplitude and EEG power in the theta frequency band. As expected, the speech-babble background noise had a differential effect on the neural processing of the consonant and vowel changes. On average, MMN responses to the CV syllable /da/ had longer latencies, smaller amplitudes, and less power in the theta frequency band across quiet and noise conditions compared to /bu/. However, it was unexpected that the addition of background noise would have a greater effect on the neural processing of the more salient vowel change when compared to the quiet condition. Finally, consistent with our expectations, the objective MMN measure in a double-oddball paradigm was a significant predictor of variations in behavioral percent correct detection of segmental-level speech stimuli as well as higher-level sentence recognition. The relevance and utility of the reported measures for potential clinical applications require further studies with different populations, stimuli, and experimental conditions.

Acknowledgments

This research project was supported in part by the Charles E. Speaks Graduate Fellowship (TKK), the Bryng Bryngelson Research Fund (TKK and YZ), the Capita Foundation (YZ), the Brain Imaging Research Project award (YZ) from the College of Liberal Arts, University of Minnesota. Participant payment was partially supported by NIDCD R01-DC008306 (PBN). We would like to thank Drs. Edward Carney, Sharon Miller, Yingjiu Nie, and Adam Svec for their assistance. Author contributions: YZ and TKK designed and performed research; TKK, YZ, and BW analyzed data; PBN and HZ provided consulting; TKK and YZ wrote the paper.

References

- Aaltonen, O., Niemi, P., Nyrke, T., Tuhkanen, M., 1987. Event-related brain potentials and the perception of a phonetic continuum. *Biol. Psychol.* 24, 197–207.
- Anderson, S., Chandrasekaran, B., Yi, H.G., Kraus, N., 2010a. Cortical-evoked potentials reflect speech-in-noise perception in children. *Eur. J. Neurosci.* 32 (8), 1407–1413.
- Anderson, S., Parbery-Clark, A., Yi, H., Kraus, N., 2011. A neural basis of speech-in-noise perception in older adults. *Ear Hear.* 32 (6), 750–757.
- Anderson, S., Skoe, E., Chandrasekaran, B., Zecker, S., Kraus, N., 2010b. Brainstem correlates of speech-in-noise perception in children. *Hear. Res.* 270 (1–2), 151–157.
- Başar, E., Demiralp, T., Schürmann, M., Başar-Eroglu, C., Ademoglu, A., 1999. Oscillatory brain dynamics, wavelet analysis, and cognition. *Brain Lang.* 66, 146–183.
- Bennett, K.O., Billings, C.J., Molis, M.R., Leek, M.R., 2012. Neural encoding and perception of speech signals in informational masking. *Ear Hear.* 33, 231–238.
- Bidelman, G.M., Dexter, L., 2015. Bilinguals at the “cocktail party”: dissociable neural activity in auditory-linguistic brain regions reveals neurobiological basis for nonnative listeners’ speech-in-noise recognition deficits. *Brain Lang.* 143, 32–41.
- Bidelman, G.M., Howell, M., 2016. Functional changes in inter- and intra-hemispheric cortical processing underlying degraded speech perception. *NeuroImage* 124, 581–590.
- Billings, C.J., Mcmillan, G.P., Penman, T.M., Gille, S.M., 2013. Predicting perception in noise using cortical auditory evoked potentials. *J. Assoc. Res. Otolaryngol.* 14, 891–903.
- Billings, C.J., Penman, T.M., Mcmillan, G.P., Ellis, E.M., 2015. Electrophysiology and perception of speech in noise in older listeners: effects of hearing impairment and age. *Ear Hear.* 36, 710–722.
- Bishop, D.V.M., Hardiman, M.J., 2010. Measurement of mismatch negativity in individuals: a study using single-trial analysis. *Psychophysiology* 47 (4), 697–705.
- Brunellière, A., Dufour, S., Nguyen, N., 2011. Regional differences in the listener’s phonemic inventory affect semantic processing: a mismatch negativity (MMN) study. *Brain Lang.* 117, 45–51.
- Caramazza, A., Papagno, C., Rumel, W., 2000. The selective impairment of phonological processing in speech production. *Brain Lang.* 75, 428–450.
- Carreiras, M., Price, C.J., 2008. Brain activation for consonants and vowels. *Cereb. Cortex* 18, 1727–1735.
- Carreiras, M., Vergara, M., Perea, M., 2009. ERP correlates of transposed-letter priming effects: the role of vowels versus consonants. *Psychophysiology* 46, 34–42.
- Choi, J.W., Lee, J.K., Ko, D., Lee, G.T., Jung, K.Y., Kim, K.H., 2013. Fronto-temporal interactions in the theta-band during auditory deviant processing. *Neurosci. Lett.* 548, 120–125.
- Christmann, C.A., Berti, S., Steinbrink, C., Lachmann, T., 2014. Differences in sensory processing of German vowels and physically matched non-speech sounds as revealed by the mismatch negativity (MMN) of the human event-related brain potential (ERP). *Brain Lang.* 136, 8–18.
- Cole, R.A., Yan, Y., Mak, B., Fanty, M., Bailey, T., 1996. The contribution of consonants versus vowels to word recognition in fluent speech. *Proc. ICASSP* 96, 853–856.
- Cunningham, J., Nicol, T., King, C., Zecker, S.G., Kraus, N., 2002. Effects of noise and cue enhancement on neural responses to speech in auditory midbrain, thalamus and cortex. *Hear. Res.* 169, 97–111.
- Cunningham, J., Nicol, T., Zecker, S.G., Bradlow, A., Kraus, N., 2001. Neurobiological responses to speech in noise in children with learning problems: deficits and strategies for improvement. *Clin. Neurophysiol.* 112 (5), 758–767.
- Delorme, Z., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *J. Neurosci. Methods* 134, 9–21.
- Du, Y., Buchsbaum, B.R., Grady, C.L., Alain, C., 2014. Noise differentially impacts phoneme representations in the auditory and speech motor systems. *PNAS* 111 (19), 7126–7131.
- Fogerty, D., Humes, L.E., 2010. Perceptual contributions to monosyllabic word intelligibility: segmental, lexical, and noise replacement factors. *J. Acoust. Soc. Am.* 128 (5), 3114–3125.
- Fogerty, D., Humes, L.E., 2012. The role of vowel and consonant fundamental frequency, envelope, and temporal fine structure cues to the intelligibility of words and sentences. *J. Acoust. Soc. Am.* 131 (2), 1490.
- Fogerty, D., Kewley-Port, D., 2009. Perceptual contributions of the consonant-vowel boundary to sentence intelligibility. *J. Acoust. Soc. Am.* 126 (2), 847–857.
- Fogerty, D., Kewley-Port, D., Humes, L.E., 2012. The relative importance of consonant and vowel segments to the recognition of words and sentences: effects of age and hearing loss. *J. Acoust. Soc. Am.* 132 (September), 1667.
- Fuentemilla, L., Marco-Pallarés, J., Münte, T.F., Grau, C., 2008. Theta EEG oscillatory activity and auditory change detection. *Brain Res.* 1220, 93–101.
- García-Sierra, A., Rivera-Gaxiola, M., Percaccio, C.R., Conboy, B.T., Romo, H., Klarman, L., et al., 2011. Bilingual language learning: an ERP study relating early brain responses to speech, language input, and later word production. *J. Phonetics* 1–12.
- Groenen, P., Snik, A., van den Broek, P., 1996. On the clinical relevance of mismatch negativity: results from subjects with normal hearing and cochlear implant users. *Audiol. Neurotol.* 1, 112–124.
- Hong, L.E., Moran, L.V., Du, X., O’Donnell, P., Summerfelt, A., 2012. Mismatch negativity and low frequency oscillations in schizophrenia families. *Clin. Neurophysiol.* 123 (10), 1980–1988.
- Hsiao, F., Cheng, C., Liao, K., Lin, Y., 2010. Cortico-cortical phase synchrony in auditory mismatch processing. *Biolog. Psychol.* 84, 336–345.
- Hsiao, F.J., Wu, Z.A., Ho, L.T., Lin, Y.Y., 2009. Theta oscillation during auditory change detection: a MEG study. *Biol. Psychol.* 81 (1), 58–66.
- Institute of Electrical and Electronic Engineers, 1969. *IEEE Recommended Practice for Speech Quality Measures*. IEEE, New York.
- Jakoby, H., Goldstein, A., Faust, M., 2011. Electrophysiological correlates of speech perception mechanisms and individual differences in second language attainment. *Psychophysiology* 48 (11), 1517–1531.
- Jansson-Verkasalo, E., Korpilahti, P., Jantti, V., Valkama, M., Vainionpa, L., Alku, P., et al., 2004. Neurophysiologic correlates of deficient phonological representations and object naming in prematurely born children. *Clin. Neurophysiol.* 115, 179–187.
- Jin, Y., Diaz, B., Colomer, M., Sebastian-Galles, N., 2014. Oscillation encoding of individual differences in speech perception. *PLoS One* 9 (7), 1–7.
- Kaser, M., Soltesz, F., Lawrence, P., Miller, S., Dodds, C., Croft, R., et al., 2013. Oscillatory underpinnings of mismatch negativity and their relationship with cognitive function in patients with schizophrenia. *PLoS One* 8 (12), 1–11.
- Kewley-Port, D., Burkle, T.Z., Lee, J.H., 2007. Contribution of consonant versus vowel information to sentence intelligibility for young normal-hearing and elderly hearing-impaired listeners. *J. Acoust. Soc. Am.* 122, 2365–2375.
- Klimesch, W., Freunberger, R., Sauseng, P., Gruber, W., 2008. A short review of slow phase synchronization and memory: evidence for control processes in different memory systems? *Brain Res.* 1235, 31–44.
- Klimesch, W., Sauseng, P., Hanslmayr, S., Gruber, W., Freunberger, R., 2007. Event-related phase reorganization may explain evoked neural dynamics. *Neurosci. Biobehav. Rev.* 31 (7), 1003–1016.
- Ko, D., Kwon, S., Lee, G., Im, H., Kim, H., Jung, K., 2012. Theta oscillation related to the auditory discrimination process in mismatch negativity: oddball versus control paradigm. *J. Clin. Neuro.* 8, 35–42.
- Koerner, T.K., Zhang, Y., 2015. Effects of background noise on inter-trial phase coherence and auditory N1-P2 responses to speech. *Hear. Res.* 328, 113–119.
- Koerner, T., Zhang, Y., Nelson, P., 2013. Links between mismatch negativity responses and speech intelligibility in noise. *J. Acoust. Soc. Am.* 132, 2049.
- Kozou, H., Kujala, T., Shtyrov, Y., Toppila, E., Starck, J., Alku, P., Näätänen, R., 2005. The effect of different noise types on the speech and non-speech elicited mismatch negativity. *Hear. Res.* 199 (1–2), 31–39.
- Kraus, N., Cheour, M., 2000. Speech sound representation in the brain. *Audiol. Neurotol.* 5 (3–4), 140–150.
- Kraus, N., Mcgee, T., Carrell, T.D., King, C., Tremblay, K., Nicol, T., 1995a. Central auditory system plasticity associated with speech discrimination training. *J. Cognitive Neurosci.* 7 (1), 25–32.
- Kraus, N., Mcgee, T., Carrell, T.D., Sharma, A., 1995b. Neurophysiologic bases of speech discrimination. *Ear Hear.* 16 (1), 19–37.
- Kraus, N., Mcgee, T.J., Carrell, T.D., Zecker, S.G., Nicol, T.G., Koch, D.B., 1996. Auditory neurophysiologic responses and discrimination deficits in children with learning problems. *Science* 273, 971–973.
- Kraus, N., Mcgee, T., Micco, A., Sharma, A., Carrell, T., Nicol, T., 1993. Mismatch negativity in school-age children to speech stimuli that are just perceptibly different. *Electroencephalogr. Clin. Neurophysiol.* 88 (2), 123–130.
- Kuhl, P.K., Conboy, B.T., Padden, D., Nelson, T., Pruitt, J., 2005. Early speech perception and later language development: implications for the “Critical Period.” *Lang. Learn. Dev.* 1, 237–264.
- Kurtzberg, D., Vaughan, H.G., Kreuzer, J.A., Fliegler, K.Z., 1995. Developmental studies and clinical application of mismatch negativity: problems and prospects. *Ear Hear.* 16, 105–117.
- Ladefoged, P., 2006. *A Course in Phonetics*, fifth ed. Thompson, Boston.
- Lang, A.H., Eerola, O., Korpilahti, P., Holopainen, I., Salo, S., Aaltonen, O., 1995. Practical issues in the clinical application of mismatch negativity. *Ear Hear.* 16 (1), 118–130.
- Li, N., Loizou, P.C., 2008. The contribution of obstruent consonants and acoustic landmarks to speech recognition in noise. *J. Acoust. Soc. Am.* 124 (6), 3947–.
- Liberman, A.M., Mattingly, I.G., 1985. The motor theory of speech perception revised. *Cognition* 21, 1–36.
- Makeig, S., Debener, S., Onton, J., Delorme, A., 2004. Mining event-related brain dynamics. *Trends Cognitive Sci.* 8 (5), 204–210.
- Martin, B.A., Tremblay, K.L., Korcak, P., 2008. Speech evoked potentials: from the laboratory to the clinic. *Ear Hear.* 29 (3), 285–313.

- Martin, B.A., Kurtzberg, D., Stapells, D.R., 1999. The effects of decreased audibility produced by high-pass noise masking on N1 and the mismatch negativity to speech sounds/ba/and/da/. *J. Speech Lang. Hear. Res.* 42, 271–287.
- Martin, B.A., Sigal, A., Kurtzberg, D., Stapells, D.R., 1997. The effects of decreased audibility produced by high-pass noise masking on cortical event-related potentials to speech sounds/ba/and/da/. *J. Acoust. Soc. Am.* 101 (3), 1585–1599.
- Martin, B.A., Stapells, D.R., 2005. Effects of low-pass noise masking on auditory event-related potentials to speech. *Ear Hear.* 26 (2), 195–213.
- Mesgarani, N., David, S.V., Fritz, J.B., Shamma, S.A., 2014. Mechanisms of noise robust representation of speech in primary auditory cortex. *PNAS* 111 (18), 6792–6797.
- Miceli, G., Capasso, R., Benvegnù, B., Caramazza, A., 2004. The categorical distinction of vowel and consonant representations: evidence from dysgraphia. *Neurocase Case Stud. Neuropsychol. Neuropsychiatry Behav. Neurol.* 10 (2), 109–121.
- Miller, G., Nicely, P., 1955. An analysis of perceptual confusions among some English consonant. *J. Acoust. Soc. Am.* 27 (2), 338–352.
- Molfese, D.L., Molfese, V.J., 1997. Discrimination of language skills at five years of age using event-related potentials recorded at birth. *Dev. Neuropsychol.* 13 (2), 135–156.
- Muller-Gass, A., Marcoux, A., Logan, J., Campbell, K.B., 2001. The intensity of masking noise affects the mismatch negativity to speech sounds in human subjects. *Neurosci. Lett.* 299, 197–200.
- Näätänen, R., 1995. The mismatch negativity: a powerful tool for cognitive neuroscience. *Ear Hear.* 16 (1), 6–18.
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., et al., 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385, 432–434.
- Näätänen, R., Paavilainen, P., Rinne, T., Alho, K., 2007. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118 (12), 2544–2590. <http://dx.doi.org/10.1016/j.clinph.2007.04.026>.
- Näätänen, R., Pakarinen, S., Rinne, T., Takegata, R., 2004. The mismatch negativity (MMN): towards the optimal paradigm. *Clinical Neurophysiol.* 115, 140–144.
- Niemitalo-Haapola, E., Haapala, S., Jansson-Verkasalo, E., Kujala, T., 2015. Background noise degrades central auditory processing in toddlers. *Ear Hear.* 36, 342–351.
- Niquette, P., Gundmundsen, G., Killion, M., 2001. QuickSIN Speech-in-noise Test Version 1.3. Etymotic Research, Elk Grove Village, IL.
- Oates, P., Kurtzberg, D., Stapells, D.R., 2002. Effects of sensorineural hearing loss on cortical event-related potential and behavioral measures of speech-sound processing. *Ear Hear.* 23 (5), 399–415.
- Owren, M.J., Cardillo, G.C., 2006. The relative roles of vowels and consonants in discriminating talker identity versus word meaning. *J. Acoust. Soc. Am.* 119 (3), 1727–1739.
- Pakarinen, S., Lovio, R., Huotilainen, M., Alku, P., Näätänen, R., Kujala, T., 2009. Fast multi-feature paradigm for recording several mismatch negativities (MMNs) to phonetic and acoustic changes in speech sounds. *Biol. Psychol.* 82, 219–226.
- Pakarinen, S., Takegata, R., Rinne, T., Huotilainen, M., Näätänen, R., 2007. Measurement of extensive auditory discrimination profiles using the mismatch negativity (MMN) of the auditory event-related potential (ERP). *Clin. Neurophysiol.* 118, 177–185.
- Pakarinen, S., Teinonen, T., Shestakova, A., Soo Kwon, M., Kujala, T., Hämäläinen, H., et al., 2013. Fast parametric evaluation of central speech-sound processing with mismatch negativity (MMN). *Int. J. Psychophysiol.* 87, 103–110.
- Parbery-Clark, A., Marmel, F., Bair, J., Kraus, N., 2011. What subcortical-cortical relationships tell us about processing speech in noise. *Cogn. Neurosci.* 33, 549–557.
- Pariikh, G., Loizou, P.C., 2005. The influence of noise on vowel and consonant cues. *J. Acoust. Soc. Am.* 118 (6), 3874–3888.
- Phatak, S.A., Allen, J.B., 2007. Consonant and vowel confusions in speech-weighted noise. *J. Acoust. Soc. Am.* 121 (4), 2312–2326.
- Pickett, J.M., 1957. Perception of vowels heard in noises of various spectra. *J. Acoust. Soc. Am.* 29 (5), 613–620.
- Picton, T.W., 1992. The P300 wave of the human event-related potential. *J. Clin. Neurophysiol.* 9 (4), 465–479.
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., R Core Team, 2016. nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-127. Retrieved from. <http://CRAN.R-project.org/package=nlme>.
- Pulvermüller, F., Shtyrov, Y., 2006. Language outside the focus of attention: the mismatch negativity as a tool for studying higher cognitive processes. *Progr. Neurobiol.* 79, 49–71.
- R Core Team, 2014. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from. <http://www.r-project.org/>.
- Radicevic, Z., Vujovic, M., Jelicic, L., Sovilj, M., 2008. Comparative findings of voice and speech: language processing at an early ontogenetic age in quantitative EEG mapping. *Exp. Brain Res.* 184, 529–532.
- Rao, A., Zhang, Y., Miller, S., 2010. Selective listening of concurrent auditory stimuli: an event-related potential study. *Hear. Res.* 268 (1–2), 123–132.
- Russo, N., Nicol, T., Musacchia, G., Kraus, N., 2004. Brainstem responses to speech syllables. *Clin. Neurophysiol.* 115 (9), 2021–2030.
- Sams, M., Paavilainen, P., Alho, K., Näätänen, R., 1985. Auditory frequency discrimination and event-related potentials. *Electroencephalogr. Clin. Neurophysiol.* 62, 437–448.
- Santesso, D.L., Schmidt, L.A., Trainor, L.J., 2007. Frontal brain electrical activity (EEG) and heart rate in response to affective infant-directed (ID) speech in 9-month-old infants. *Brain Cognit.* 65, 14–21.
- Sauseng, P., Klimesch, W., Gruber, W.R., Hanslmayr, S., Freunberger, R., Doppelmayr, M., 2007. Are event-related potential components generated by phase resetting of brain oscillations? A critical discussion. *Neuroscience* 146 (4), 1435–1444.
- Shetake, J.A., Wolf, J.T., Cheung, R.J., Engineer, C.T., Ram, S.K., Kilgard, M.P., 2011. Cortical activity patterns predict robust speech discrimination ability in noise. *Eur. J. Neurosci.* 34 (11), 1823–1838.
- Song, J.H., Skoe, E., Banai, K., Kraus, N., 2011. Perception of speech in noise: neural correlates. *J. Cognitive Neurosci.* 23 (9), 2268–2279.
- Stapells, D.R., 2002. Cortical event-related potentials to auditory stimuli. In: Katz, J. (Ed.), *Handbook of Clinical Audiology*, sixth ed. Lippincott, Williams, and Williams, Baltimore, pp. 395–430.
- Tremblay, K., Kraus, N., Carrell, T.D., McGee, T., 1997. Central auditory system plasticity: generalization to novel stimuli following listening training. *J. Acoust. Soc. Am.* 102, 3762–3773.
- Tremblay, K., Kraus, N., McGee, T., 1998. The time course of auditory perceptual learning: neurophysiological changes during speech-sound training. *Neuroreport* 9 (16), 3557–3560.
- Vaden, K.I., Kuchinsky, S.E., Ahlstrom, J.B., Dubno, J.R., Eckert, M.A., 2015. Cortical activity predicts which older adults recognize speech in noise and when. *J. Neurosci.* 35 (9), 3929–3937.
- Ward, L.M., 2003. Synchronous neural oscillations and cognitive processes. *Trends Cognitive Sci.* 7 (12), 553–559.
- Whiting, K.A., Martin, B.A., Stapells, D.R., 1998. The effects of broadband noise masking on cortical event-related potentials to speech sounds/ba/and/da/. *Ear Hear.* 19 (3), 218–231.
- Winkler, I., Kujala, T., Tiitinen, H., Alku, P., Alku, P., Lehtokoski, A., et al., 1999. Brain responses reveal the learning of foreign language phonemes. *Psychophysiology* 36, 638–642.
- Wong, P.C.M., Uppunda, Ajith K., Parrish, T.B., Dhar, S., 2008. Cortical mechanisms of speech perception in noise. *J. Speech Lang. Hear. Res.* 51, 1026–1041.
- Xi, Jie, Zhang, Linjun, Shu, Hua, Zhang, Yang, Li, Ping, 2010. Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience* 170 (1), 223–231.
- Zhang, Y., Koerner, T., Miller, S., Grice-Patil, Z., Svec, A., Akbari, D., et al., 2011. Neural coding of formant-exaggerated speech in the infant brain. *Dev. Sci.* 14 (3), 566–581.
- Zhang, Y., Kuhl, P.K., Imada, T., Kotani, M., Tohkura, Y., 2005. Effects of language experience: neural commitment to language-specific auditory patterns. *Neuroimage* 26 (3), 703–720.
- Zhang, Y., Kuhl, P.K., Imada, T., Iverson, P., Pruitt, J., Stevens, E.B., et al., 2009. Neural signatures of phonetic learning in adulthood: a magnetoencephalography study. *NeuroImage* 46 (1), 226–240.