

# Converging evidence for [coronal] underspecification in English-speaking adults



Alycia Cummings<sup>a, b, \*</sup>, John Madden<sup>a</sup>, Kathryn Hefta<sup>a</sup>

<sup>a</sup> University of North Dakota, United States

<sup>b</sup> Idaho State University, United States

## ARTICLE INFO

### Article history:

Received 27 December 2015

Received in revised form 14 May 2017

Accepted 19 May 2017

Available online 29 May 2017

### Keywords:

ERP

EEG

MMN

Phonological representations

Underspecification

LIMO

Single-subject

## ABSTRACT

The goal of this study was to test the predictions of the Featurally Underspecified Lexicon (FUL) theory by examining event-related potential (ERP) indices of phonological representation. Two English consonants differing in place of articulation were selected: [labial]/b/ and [coronal]/d/. It was assumed that the phonological representation of /d/ contained less distinctive feature information due to its [coronal] place of articulation, as compared to /b/. English-speaking adults were presented with two syllables, /ba/ and /da/, in an ERP oddball paradigm where both syllables served as the standard and deviant stimulus in opposite stimulus sets. Three types of analyses were conducted: traditional mean amplitude measurements, cluster-based permutation tests, and single-trial general linear model (GLM) analyses of group-level and single-subject data. The less specified /da/ deviant elicited a large MMN while no MMN was elicited by the more specified deviant /ba/. Additionally, the /da/ standard syllable elicited larger responses than did the /ba/ standard, while deviant syllables did not differ. This implies that the MMN was driven by responses elicited by the standards rather than the deviants. At the single-subject level, not all participants demonstrated significant MMN responses, though all had measurable differences between the standard syllables. Thus, to continue to propose that [coronal] underspecification is a language universal phenomenon, ERP indices other than the MMN should be examined.

© 2017 Elsevier Ltd. All rights reserved.

## 1. Introduction

Distinctive features of phonemes, or speech sounds, form the basis for the phonological representations that are accessed during lexical representation tasks. Each phoneme typically differs in terms of one or more distinctive features (Jakobson, Fant, & Halle, 1952; Ladefoged, 2007). Distinctive features permit categorization of speech sounds; sounds that belong to the same phonological category are considered to form a natural class. One major class feature addresses phonemes' place of articulation. For example, /b/ is a [labial] sound produced with the lips, /d/ is a [coronal] sound produced at the alveolar ridge behind the top teeth, while /g/ is a [dorsal] sound produced at the back of the soft palate.

While phonological representations as features have been a part of the phonology literature since the inception of generative phonology (Chomsky & Halle, 1968), the question of how contrastive phonemic information is stored in the lexicon has been widely debated. An important aspect of this debate is the degree of feature specification. Some approaches

\* Corresponding author. 290 Centennial Drive, Stop 8040, Grand Forks, ND 58202-8040, United States.

E-mail address: [alycia.cummings@und.edu](mailto:alycia.cummings@und.edu) (A. Cummings).

suggest that phonological representations must be completely faithful to what is heard, thus typically requiring the storage of very detailed and specific representations for each production variant a listener encounters (Bybee, 2001; Johnson, 1997, 2005; Pisoni, 1993; Rabbitt & Connine, 2007). For example, the extreme view of this approach suggests that all phonetic detail, as well as all production forms of each word, are stored (Johnson, 2006).

An alternative view is that a more abstract phonological representation is created, stored, and accessed during speech perception and production. In this view, phonemes are said to be underspecified, in the sense that only contrastive or not otherwise predictable phonological information (i.e., distinctive features) is stored for each phoneme. Thus, to make speech processing easier, adults are hypothesized to not store all aspects of all phonemes in their underlying representations within the mental lexicon. Only the phonemically contrastive distinctive features are stored; the features that are common and similar across phonemes are not stored because they are predictable. In other words, underspecification models suggest that phonological representations are maximally abstract so that not all features are stored. This results in an efficient organization of the lexicon.

One model of sparse or underspecified representations is the Featurally Underspecified Lexicon (FUL) (Lahiri & Marslen-Wilson, 1991; Lahiri & Reetz, 2002, 2010). This model proposes that the phonological features identified in the speech sample are mapped onto a phonologically underspecified mental representation. The underspecified features are not just those that are redundant in the language, but also those which are contrastive but inactive. FUL has clear predictions about how features are specified. For example, it predicts that [coronal] place of articulation exists but is underspecified in all word positions as compared to other place of articulation features.

### 1.1. Coronal underspecification

Many researchers assume that the [coronal] consonantal place of articulation feature is less specified than others (e.g., [labial] or [dorsal]) of articulation (Cornell, Lahiri, & Eulitz, 2011, 2013; Eulitz & Lahiri, 2004; Friedrich, Eulitz, & Lahiri, 2006; Gaskell & Marslen-Wilson, 1996, 1998; Kiparsky, 1985; Lahiri & Reetz, 2002, 2010; Paradis & Prunet, 1991; Snoeren, Gaskell, & Di Betta, 2009; Wheeldon & Waksler, 2004; Yip, 1991; Zimmerer, Reetz, & Lahiri, 2009). The hypothesis that coronal consonants are less specified than labial and dorsal consonants stems from different types of phonological observations and evidence. For example, coronal consonants are the most frequent consonant class across the world's languages (Maddieson, 1984). Consonants that are more common across languages are thought to be less specified than consonants that are present in few languages. Moreover, coronal consonants are more likely to assimilate to the place of articulation of a neighboring consonant than are labial and velar consonants (Kabak, 2007; Paradis & Prunet, 1991; Stemberger & Stoel-Gammon, 1991). That is, [coronal] place assimilation is much more common than [labial]. For example, English place assimilation often occurs across word boundaries: “that pen” is produced as “thap pen” and “good boy” turns into “goob boy”. In contrast, labials and velars do not assimilate to coronals: “plum tree” is not produced as “plun tree” and “thick tree” is not produced as “thit tree” (Gimson, 1989). If coronal consonants are underspecified for place, they are more vulnerable to acquiring the place of a neighboring segment.

### 1.2. MMN as a possible electrophysiological index of phonological underspecification

Recently, a variety of studies using event-related potentials (ERPs) have examined whether there is neural evidence of underspecification during the perception of speech sounds. Currently, one ERP peak has been used to identify phonological underspecification: the Mismatch Negativity (MMN) (Näätänen & Winkler, 1999; Näätänen, Paavilainen, Rinne, & Alho, 2007; Picton, Alain, Otten, Ritter, & Achim, 2000). The MMN is an attention-independent neurophysiological response elicited by an acoustically different (deviant) stimulus when presented in a series of homogeneous (standard) stimuli (the oddball paradigm) (Näätänen, 1995; Näätänen, Gaillard, & Mäntysalo, 1978). The MMN typically is observed 100–350 ms after stimulus onset (Csépe, 1995; Korpilahti, Krause, Holopainen, & Lang, 2001; Morr, Shafer, Kreuzer, & Kurtzberg, 2002; Näätänen, Paavilainen, & Reinikainen, 1989; Shafer, Morr, Kreuzer, & Kurtzberg, 2000; Čepionienė et al., 2002), suggesting that low-level neural mechanisms exist to distinguish between certain phonological contrasts.

The assumptions of the FUL approach to underspecification can be easily applied to MMN interpretation (Cornell et al., 2013, 2011; Eulitz & Lahiri, 2004). Let us assume that the two stimuli in an oddball paradigm are /da/ and /ba/ (as is the case in of this study). Each of the two consonants can function as a standard or a deviant stimulus. The standard is presented many times and as such, establishes an expected auditory feature pattern, which consists of the underlying representation of the standard's feature pattern in the mental lexicon. This representation is assumed to represent long-term memory representations of single sound segments (Näätänen et al., 1997). On the other hand, the deviant appears only sporadically, which provides a more surface-faithful (input-driven) representation involving an acoustic and feature-based mismatch to the standard representation (Eulitz & Lahiri, 2004; Scharinger, Merickel, Riley, & Idsardi, 2011; Winkler et al., 1999).

FUL predicts that asymmetric MMN response patterns will occur under certain conditions, with the size of the MMN depending on the degree of specification of the standard and deviant stimuli. To clarify, FUL assumes that variation in speech is resolved by the listener in two steps: 1) the auditory system parses the acoustic signal into features and not sound

segments, and 2) a mapping process matches features extracted from the acoustic signal with those stored in the mental lexicon (Lahiri, 2007). FUL specifies three possible outcomes of the matching process. The match condition is clear. The no mismatch condition assumes that certain imperfect matches are tolerated due to underspecification. A true mismatch occurs when a feature extracted from the signal is in direct conflict with a feature in the underlying, long term memory representation. Thus, the presence of the [coronal] feature in a surface (deviant) [d] conflicts with the underlying (standard) [labial] specification of /b/. The opposite does not hold however, as the [labial] specification of a surface [b] does not conflict with underlying /d/, since underlying /d/ is not specified for place (Walter & Hacquard, 2004). Thus, within the FUL framework, the MMN can be modulated by the feature properties of the standard and deviant stimuli.

Much of the previous ERP work examining underspecification has focused on German speakers listening to German phonemes that differ in their place of articulation: isolated vowels (Eulitz & Lahiri, 2004), vowels in single syllables (Diesch & Luce, 1997), vowels in multi-syllabic words/nonwords (Cornell et al., 2011), consonants presented between vowels (Cornell et al., 2013), or consonants presented in word-final consonant clusters (Scharinger, Bendixen, Trujillo-Barreto, & Obleser, 2012). In all of these studies, larger and earlier MMN responses were observed when there was a true mismatch between the phonological representations of the standard and deviant sounds of the oddball paradigm, with conflicting features considered to be mutually exclusive, such as [dorsal] versus [coronal].

Evidence supporting the idea of underspecification in English is more mixed. For example, Scharinger, Monahan, and Idsardi (2012) examined the underspecification of vowel height in three English vowels. Consistent with the previous German studies, a larger MMN response was observed when the underspecified vowel /e/, as in “bet” was the deviant, as compared to a vowel with a more specified height, /æ/, as in “bat” (Scharinger, Monahan et al., 2012). In contrast, Scharinger et al. (2011) examined [coronal] underspecification in two English consonant manner classes: glides (/w/and/j/) and fricatives (/v/and/z/), with /w/and/v/ classified as [labial] and /j/and/z/ classified as [coronal]. Each consonant was presented in a vowel-consonant-vowel (VCV) combination (e.g., “awa”). While FUL predicts that the underspecified [coronal] consonant should elicit larger MMN responses than the [labial], Scharinger and colleagues observed the opposite. Thus, while neural evidence for underspecification has been consistently found in German, evidence for underspecification in English is inconclusive.

### 1.3. The present study

The purpose of the present study was to test the FUL theory of phonological underspecification by examining ERP indices of phonological representation. This was accomplished in three ways. First, while previous ERP underspecification work has primarily focused on vowels or non-English languages, the present study compared the [coronal]/d/ to the [labial]/b/ in English-speaking adults. These consonants were selected because they varied only in their place of articulation, while being of the same consonant manner class (stops/plosives), and being voiced (i.e., the vocal folds vibrated during their production). Thus, the difference between the two consonant sounds was much more clear and distinct than that of the English consonants used by Scharinger et al. (2011). Plosives are the ideal consonants to use in ERP studies because they have abrupt, clear onsets, which ensures that the time-locking of the neural responses to the onset of the stimulus is a straightforward process. In addition, one of the previous German studies compared the plosives /d/and/g/ (Cornell et al., 2013); the present study sought to replicate the underspecification of /d/with a new plosive, /b/, in English.

The syllables /ba/and/da/ served as both standard and deviant stimuli in a multiple stimulus set oddball paradigm. The MMN was examined in same-stimulus, identity difference waves for evidence of phonological underspecification in the underlying representation of the phoneme /d/. Based on the conceptual framework regarding [coronal] underspecification, it was predicted that a larger MMN response would be elicited when /a/da/deviant was presented within a standard stimulus context of the more specified /ba/ than when the opposite stimulus pairing was employed.

Second, since that the MMN is often most evident at fronto-central and central electrode sites (Näätänen et al., 2007), many previous ERP studies examined underspecification MMN effects at a single electrode, such as Fz (Cornell et al., 2011, 2013; Eulitz & Lahiri, 2004; Scharinger, Monahan et al., 2012). However, the present study examined responses across the entire scalp for a variety of reasons. Given that the neural organization of the brain is often quite individualized (Luck, 2005; Teplan, 2002; Woodman, 2010), this approach allowed for a more widespread view of underspecification effects, as it was likely that the maximal MMN response location would differ somewhat across participants. In addition, while the generators of the MMN are thought to be located bilaterally in the supratemporal cortices (see Näätänen et al., 2007 for a review), speech processing is thought to activate a dual-stream system involving both ventral and dorsal brain locations<sup>1</sup> (Hickok & Poeppel, 2004, 2007). Thus, it is possible that other areas of the brain (outside of the supratemporal cortices) are also involved in phonological processing. The large array of electrodes allowed for the opportunity to record evidence of underspecification in locations other than electrode Fz.

Finally, while the MMN has been the preferred index of underspecification, generating difference waves during the averaging process creates bias in the data due to the fact that the averages for the standard and deviant syllables originate from largely unequal sample sizes.<sup>2</sup> As a result, the average for deviant trials can be much more affected by noise than the

<sup>1</sup> The dorsal stream involves structures in the posterior frontal lobe and the posterior dorsal-most aspect of the temporal lobe and parietal operculum. The ventral stream involves structures in the superior and middle portions of the temporal lobe.

<sup>2</sup> We owe this suggestion to an anonymous reviewer.

average for standards, meaning that the principle of homoscedasticity is violated. This implies that while the group-level MMN can demonstrate stimulus type differences, the underlying mechanisms driving the MMN differences could be masked by the difference wave subtraction process. This is especially important since the FUL underspecification theory proposes that the MMN is generated by differences in deviant stimuli responses, with the expectation that the standard stimuli generate similar responses. However, it is possible that both the standard *and* deviant stimuli can impact the morphology of the MMN. Thus, this study examined the responses elicited by standard and deviant syllables in averaged ERPs, as well as in unaveraged electroencephalogram (EEG) epochs.

It was predicted that the less specified/da/syllable would elicit larger standard and deviant responses than/ba/. If underspecification is a language-universal phenomenon, it should be evidenced in every response (standard and deviant) elicited by a [coronal] consonant, as compared a [labial] consonant. To explain, when distinctive features are not explicitly stated in a phonological representation, any number of neuronal populations could respond since they were not coded to *not* respond. That is, with no place of articulation specified, no one set of neurons is established as the target neuronal population, which could result in many neuronal populations increasing their activation levels. Thus, if a feature is not specified, any number of neurons have the option to respond. This would lead to an underspecified sound eliciting a larger response than a more specified sound.

## 2. Methods

### 2.1. Participants

Fifteen native speakers of (American) English (five male; mean age: 21.16 years, range: 19–23 years) who were undergraduate students participated in the study. All of them had normal or corrected-to-normal vision, and none had a history of speech, language, and/or hearing impairment. This study was approved by the university internal review board and each participant signed informed consent in accordance with the university human research protection program.

### 2.2. Stimuli

Syllables (consonant + /a/) were pronounced by a male North American English speaker. The syllables were digitally recorded in a sound isolated room (Industrial Acoustics Company, Inc., Winchester, UK) using a Beyer Dynamic (Heilbronn, Germany) Soundstar MK II unidirectional dynamic microphone and Behringer (Willich, Germany) Eurorack MX602A mixer. All syllables were digitized with a 16-bit AD converter at a 44.1 kHz sampling rate. The average intensity of all the syllable stimuli was normalized to 65 dB SPL.

The adults heard two oddball stimulus sets, each containing the same four English speech consonant-vowel (CV) syllables: “ba” (/ba/), “da” (/da/), “pa” (/pa/), and “ga” (/ga/). In one stimulus set, /ba/ served as the standard syllable, with the other three CV syllables serving as deviants. In the second stimulus set, /da/ served as the standard syllable, with the other three syllables being deviants. Only responses to the /ba/ and /da/ syllables will be addressed further since they served as both standard and deviant stimuli, which allowed for the creation of same-stimulus identity difference waves. Since /pa/ and /ga/ deviants were incorporated to prevent MMN habituation, they were not examined. As initially recorded, the syllables varied slightly in duration, due to the individual phonetic make-up of each consonant. Syllable duration was minimally modified (by shortening the vowel duration) so that all syllables were 375 ms in length. Each syllable token used in the study was correctly identified by at least 15 adult listeners. Formant frequency measurements of /ba/ and /da/ are presented in Table 1.

### 2.3. Stimulus presentation

The stimuli were presented in blocks containing 237 standard stimuli and 63 deviant stimuli (21 per deviant), with five blocks being presented to each participant. Each block lasted approximately 6 min and the participants were given a break between blocks when necessary, which was typically 2–3 breaks per session. Within the block, the four stimuli were presented using an oddball paradigm in which the three deviant stimuli (probability = 7% for each) were presented in a series of standard stimuli (probability = 79%). Stimuli were presented in a pseudorandom sequence and the onset-to-onset inter-stimulus interval varied randomly between 600 and 800 ms. The syllables were delivered by stimulus presentation software (Presentation software, [www.neurobs.com](http://www.neurobs.com)). The syllable sounds were played via two loudspeakers situated 30° to the right

**Table 1**

Consonant-vowel (CV) syllable formant frequencies (in hertz). Two frequency values are given for the formants of each syllable. The first value specifies the formant onset frequency at the beginning of the syllable and the second represents the steady-state (vowel) frequency attained after the formant transition toward the center of the syllable.

	/ba/		/da/	
F1 (Hz)	486	818	410	863
F2 (Hz)	998	1158	1792	1180
F3 (Hz)	2575	2731	2812	2722

and left from the midline 120 cm in front of a participant, which allowed the sounds to be perceived as emanating from the midline space. The participants sat in a sound-treated room and watched a silent cartoon video of their choice. The recording of the ERPs took approximately 1 h.

#### 2.4. EEG recording and averaging

Sixty-six channels of continuous EEG (DC–128 Hz) were recorded using an ActiveTwo data acquisition system (Biosemi, Inc, Amsterdam, Netherlands) at a sampling rate of 256 Hz. This system provides “active” EEG amplification at the scalp that substantially minimizes movement artifacts. The amplifier gain on this system is fixed, allowing ample input range (–264 to 264 mV) on a wide dynamic range (110 dB) Delta-Sigma ( $\Delta\Sigma$ ) 24-bit AD converter. Sixty-four channel scalp data were recorded using electrodes mounted in a stretchy cap according to the International 10–20 system. Two additional electrodes were placed on the right and left mastoids. Eye movements were monitored using FP1/FP2 (blinks) and F7/F8 channels (lateral movements, saccades). During data acquisition, all channels were referred to the system's internal loop (CMS/DRL sensors located in the centro-parietal region), which drives the average potential of a subject (the Common Mode voltage) as close as possible to the Analog-Digital Converter reference voltage (the amplifier “zero”). The DC offsets were kept below 25  $\mu$ V at all channels. Off-line, data were re-referenced to the average of the left and right mastoid tracings.

Prior to data averaging, sporadic artifact rejection of the continuous EEG was completed using EEGLAB (Delorme & Makeig, 2004). This involved marking and rejecting the time periods during which sporadic artifacts occurred. Sources of the artifacts included random head movements, muscle movements related to speaking, excessive electrode activation stemming from pressing the head into the back of the chair, etc. After sporadic artifact rejection, independent-component analysis (ICA) (Jung et al., 2000) was completed so that the experimenters could identify eye blink and saccade components. These blink and saccade components were then deleted from the continuous EEG. The remaining artifactual trials due to amplifier blocking as well as muscle and overall body movements were rejected from further analyses using the simple voltage threshold measure in ERPLAB (Luck & Lopez-Calderon, 2012). Voltage limits were set at –100 to 100  $\mu$ V.

Epochs containing 100 ms pre-auditory stimulus and 800 ms stimulus times were baseline-corrected with respect to the pre-stimulus interval and averaged by stimulus type: standard syllable and deviant syllable. The data were low-pass filtered at 30 Hz and high-pass filtered at 0.05 Hz, using 2-way least squares FIR filters. On average, the remaining individual data contained 735 (SD = 95)/ba/standard syllable trials, 752 (SD = 90)/da/standard syllable trials, 90 (SD = 11)/ba/deviant syllable trials, and 88 (SD = 11)/da/deviant syllable trials.

#### 2.5. ERP and EEG measurements

Three different data analysis strategies were used in the present study: 1) traditional mean amplitude repeated measure ANOVA analyses using averaged data, 2) cluster-based permutation analyses of averaged data (Bullmore et al., 1999; Groppe, Urbach, & Kutas, 2011), and 3) general linear modeling of epoched (i.e., unaveraged) data (Pernet, Chauveau, Gaspar, & Rousselet, 2011).

##### 2.5.1. Mean amplitude measurements of averaged data

In an oddball paradigm, the MMN is typically examined by subtracting the standard ERP response from the deviant response in difference waves. The dual stimulus set nature of the present study allowed for the creation of “same-stimulus”, or identity, difference waveforms. These difference waves were created by subtracting the ERP response of a stimulus serving as the standard from that of the same stimulus serving as the deviant, across stimulus sets. For example, the ERP response for/ba/as the deviant was subtracted from the ERP response for/ba/as the standard (of the reversed stimulus set) (Cornell et al., 2013, 2011; Eulitz & Lahiri, 2004). The creation of identity difference waveforms eliminates the potential confound that variations in ERP morphology may result from acoustic stimulus differences, since the same stimulus is used to elicit both the standard and deviant responses. The waveforms were visually inspected from 0 to 300 ms, with the MMN appearing between approximately 75 and 250 ms post-syllable onset.

Peak measurement of MMN was a multi-step process. Given that the MMN is typically maximal over fronto-central midline electrode sites (e.g., Fz, FCz, and Cz) (Näätänen, Teder, Alho, & Lavikainen, 1992), these electrodes were selected for the mean amplitude analyses. Each maximal peak latency was first measured in the grand averaged waveforms using the ERPLAB Toolbox (Luck & Lopez-Calderon, 2012) across the time window referred above: 0–300 ms. Peak latencies elicited by the/ba/and/da/stimuli were averaged across the three electrodes to determine the “center” latency of each peak. The “center” latency was then used to align a 50 ms window (25 ms on either side) to measure the mean amplitudes for all electrodes and stimulus types. This meant the MMN mean amplitude window was 110–160 ms. The/ba/syllable did not elicit an obvious MMN peak; therefore, the/ba/identity difference wave was measured from the latency window obtained from the/da/stimulus. Phonological underspecification in these difference waves was analyzed using a Syllable Type (/ba/,/da/) x Electrode (Fz, FCz, Cz) repeated measure ANOVA.

Given that the difference waves were generated from the standard and deviant syllable ERPs, the mean amplitude measurements of the standard and deviant waveforms were taken from the same time window as that of the MMN: 110–160



ms post-syllable onset. In terms of ERP waveform morphology, this measurement approximately captured the time period between the peak of the auditory N1 and the peak of the auditory P2. Phonological underspecification in these ERPs was analyzed using a Syllable Type (/ba/, /da/) x Trial Type (Standard, Deviant) x Electrode (Fz, FCz, Cz) repeated measure ANOVA. Partial eta squared ( $\eta^2$ ) effect sizes are also reported for all significant effects and interactions. When applicable, Geiser-Greenhouse corrected p-values are reported.

### 2.5.2. Cluster mass permutation tests of averaged data

The ERPs were also submitted to repeated measures two-tailed cluster-based permutation tests (Bullmore et al., 1999; Groppe et al., 2011) with a family-wise alpha level of 0.05. These permutation test analyses provide better spatial and temporal resolution than conventional ANOVAs while maintaining weak control of the family-wise alpha level (i.e., it corrects for the large number of comparisons). To estimate the distribution of the null hypothesis, 2500 permutations were used, which was more than twice the number recommended for a family-wise alpha level of 0.05 (Manly, 2006). These analyses enabled identification of differences between the underspecified/da/ and the more specified/ba/. Thus, the high temporal resolution of this analysis could be used to identify a specific time period during which indices of underspecification were present.

Five different tests were conducted: 1) /ba/ vs. /da/ identity MMN difference waveforms, 2) /ba/ vs. /da/ standard ERPs, 3) /ba/ vs. /da/ deviant ERPs, 4) /ba/ standard vs. deviant ERPs, and 5) /da/ standard vs. deviant ERPs. Each test included 28 different electrodes that encompassed four different anterior-posterior levels (Frontal, Frontal-Central, Central, Central-Parietal) and seven different laterality measures: F5/F6, F3/F4, F1/F2, Fz, FC5/FC6, FC3/FC4, FC1/FC2, FCz, C5/C6, C3/C4, C1/C2, Cz, CP5/CP6, CP3/CP4, CP1/CP2, CPz. All of the time points (measured every 4 ms; 78 total time points) between 0 and 300 ms at the 28 scalp electrodes were included in the test (i.e., 2184 total comparisons).

T-tests were performed for each comparison using the original data and 2500 random within-participant permutations of the data. For each permutation, all t-scores corresponding to uncorrected p-values of 0.05 or less were formed into clusters. Electrodes within about 5.44 cm of one another were considered spatial neighbors, and adjacent time points were considered temporal neighbors. The sum of the t-scores in each cluster was the “mass” of that cluster. The most extreme cluster mass in each of the 2501 sets of tests was recorded and used to estimate the distribution of the null hypothesis (i.e., no difference between conditions). The permutation cluster mass percentile ranking of each cluster from the observed data was used to derive p-values assigned to each member of the cluster. T-scores that were not included in a cluster were given a p-value of 1.

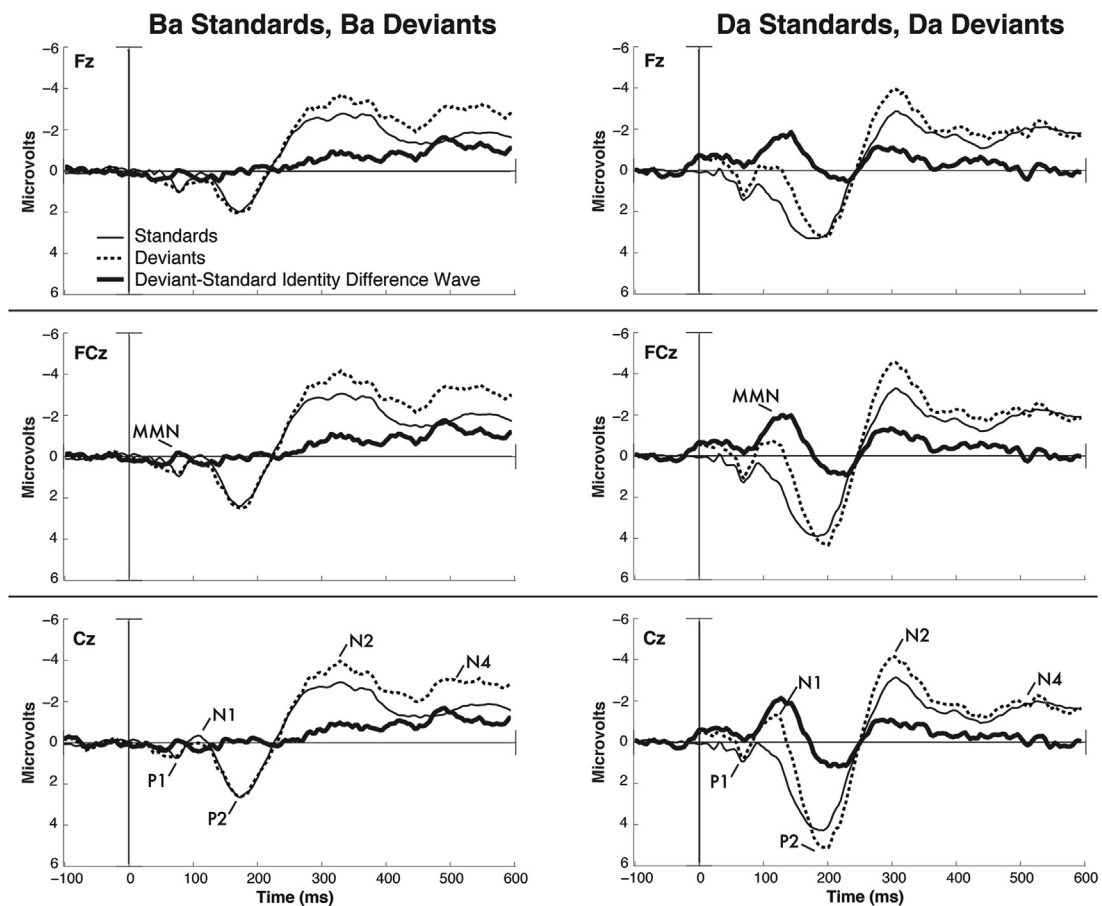
### 2.5.3. General Linear Modeling (GLM) of epoched data

GLM analyses were used to help account for the correlation in time and space dimensions found in EEG data, and to provide an alternate analysis technique to the repeated measure ANOVAs commonly used in ERP data analysis. They also helped to account for the trial number imbalance found in the standard and deviant syllable data generated by the oddball paradigm.

Following the protocol described in previous studies (Rousselet, Gaspar, Wiczorek, & Pernet, 2011; Salvia et al., 2014), subjects' epoched data were modeled using LIMO EEG, an open source Matlab toolbox for hierarchical GLM, compatible with EEGLAB: [https://gforge.dcn.ed.ac.uk/gf/project/limo\\_eeg/](https://gforge.dcn.ed.ac.uk/gf/project/limo_eeg/) (Pernet et al., 2011). The general linear model was used to examine single-trial ERP amplitudes, in microvolts, independently at each time point and each electrode. Parameters ( $\beta$ -values) were estimated at each electrode and time point independently, yielding a matrix of 64 (electrodes) x 78 (time points, from 0 to 300 ms post-stimulus in 4 ms steps) for each regressor. Similar electrode x time point matrices were computed for  $R^2$ , F, and p-values for both the overall models and for each regressor (partial F-values). Probability values were determined using a permutation approach for which trial labels were permuted 1000 times using a bootstrap-t technique (Wilcox, 2005). To examine underspecification of /da/ as compared to /ba/ in standard and deviant trials, four GLM analyses of epoched data were conducted at both the group level and single-subject level: 1) /ba/ vs. /da/ standards, 2) /ba/ vs. /da/ deviants, 3) /ba/ standards vs. deviants, and 4) /da/ standards vs. deviants.

**2.5.3.1. Group level analyses.** For each analysis, bootstrap paired t-tests were computed between the contrasts of interest at all time points across the entire scalp. Multiple comparisons were controlled for using bootstrap and the clustering technique as implemented in the Matlab Fieldtrip toolbox, with a minimum of two neighboring channels per cluster (Maris & Oostenveld, 2007). Spatio-temporal clustering functions were then used in both space and time to correct for multiple comparisons ( $p \leq 0.05$ ): first, independently at each sensor and for each permutation, F-values (for each variable) reaching the  $p \leq 0.05$  threshold, were clustered in time. The sum of F-values inside each spatio-temporal cluster was computed and the maximum sums were kept. Then, the maximum sums across electrodes were sorted to obtain a 95th percentile threshold to which actual F-values were compared (Bieniek, Pernet, & Rousselet, 2012; Pernet et al., 2011; Rousselet & Pernet, 2011; Salvia et al., 2014).

**2.5.3.2. Single-subject analyses.** For each analysis, bootstrap paired t-tests were computed between the contrasts of interest at all time points across the entire scalp. Due to the small number of deviant trials and hence a low signal-to-noise ratio, most individual participants' analyses were not significant when controlled for multiple comparisons. Thus, for the purpose of examining phonological underspecification at a single-subject level, uncorrected data are reported.



**Fig. 1.** ERP waveforms elicited by the/ba/(left side) and/da/(right side) syllables in the ERP study. The deviant waveforms represent the neural responses when the deviant syllable was presented within a stream of the opposite syllable standards. Subtracting the standard syllable response from the deviant syllable response resulted in the identity difference waves. Note that negative is plotted up in all waveforms in this figure.

In an effort to quantify and compare the individual results, two analyses were conducted. First, using the full-scalp uncorrected comparison analyses, the data of each participant were examined for a significant effect of at least 20 continuous milliseconds at electrode Fz during the 0–300 ms time window.<sup>3</sup> The second analysis involved identifying a significant continuous 20 ms effect in at least five separate electrode sites; this analysis did not have to include electrode Fz, though this electrode was included in some cases.

### 3. Results

For both the/ba/and/da/syllables, the ERP waveforms elicited by the standard and deviant stimuli consisted of P1 at ca. 75 ms, N1 at ca. 100 ms, P2 at ca. 180 ms, N2 at ca. 320 ms, and N4 at ca. 500 ms (Fig. 1). In the same-stimulus identity difference wave, the/ba/identity waveform showed virtually no evidence of a MMN, while the/da/identity waveform demonstrated a MMN response at ca. 130 ms (Fig. 1).

#### 3.1. Averaged ERP results

##### 3.1.1. MMN mean amplitude analysis

The/da/identity difference waveform demonstrated a significantly larger MMN response than did the/ba/identity waveform ( $F(1,14) = 9.838$ ,  $p < 0.008$ ,  $\eta^2 = 0.413$ ) (Table 2). In order to better characterize the group-level MMN results, scatterplots

<sup>3</sup> Electrode Fz was chosen as it is the electrode most commonly discussed in the MMN and ERP underspecification literature.

**Table 2**

Mean amplitude measurements in microvolts taken from three midline electrodes (SEM) during the 110–160 ms post-syllable onset time window. The MMN was measured from the identity difference waveforms.

MMN Mean Amplitude in the 110–160 ms time window						
	Fz	FCz	Cz	Overall		
/ba/identity	0.309 (0.433)	0.242 (0.411)	0.148 (0.349)	0.233 (0.393)		
/da/identity	–1.547 (0.387)	–1.712 (0.420)	–1.738 (0.402)	–1.666 (0.394)		
Overall	–0.619 (0.262)	–0.735 (0.261)	–0.795 (0.248)			
Standard and Deviant ERP Mean Amplitudes in the 110–160 ms time window						
	Fz	FCz	Cz	Overall	Syllable Type Overall	Trial Type Overall
/ba/standards	0.795 (0.416)	0.878 (0.426)	0.776 (0.371)	0.817 (0.401)	/ba/	Standards
/ba/deviants	1.104 (0.602)	1.120 (0.601)	0.924 (0.486)	1.049 (0.558)	0.933 (0.444)	1.401 (0.442)
/da/standards	2.107 (0.530)	2.085 (0.534)	1.765 (0.490)	1.986 (0.514)		
/da/deviants	0.560 (0.487)	0.373 (0.518)	0.027 (0.508)	0.320 (0.496)	/da/	Deviants
Overall	1.142 (0.455)	1.114 (0.466)	0.873 (0.412)		1.153 (0.465)	0.685 (0.473)

of the individual participants' MMN mean amplitudes are presented in Fig. 2a. No effect of electrode location found ( $p > 0.17$ ). No syllable type  $\times$  electrode location interaction was found ( $p > 0.80$ ).

### 3.1.2. MMN cluster permutation analysis

Cluster-level mass permutation procedures encompassing the timeline of the MMN (0–300 ms) were applied to the data. Colored rectangles indicate electrodes/time points in which the ERPs to/ba/are significantly different from those to/da/. The color scale dictates the size of the  $t$ -test result, with dark red and blue colors being more significant. Gray areas indicate electrodes/time points at which no significant differences were found. Note that the electrodes are organized along the y-axis somewhat topographically. Electrodes on the left and right sides of the head are grouped on the figure's top and bottom, respectively; midline electrodes are shown in the middle. Within those three groupings, y-axis top-to-bottom corresponds to scalp anterior-to-posterior.

The results of this cluster permutation analysis are displayed in a raster diagram in Fig. 3a. One broadly distributed effect from 86 to 156 ms signified the time period during which a larger MMN was elicited by/da/, as compared to/ba/; the smallest significant  $t$ -score was:  $t(14) = -2.145$ ,  $p < 0.04$ , with a test-wise alpha level of 0.0499.

### 3.1.3. Standard and deviant ERP mean amplitude analysis

No main effect of syllable was found ( $p > 0.94$ ). Overall the/ba/and/da/syllables elicited similar sized responses during the 110–160 ms time window. Standard syllables elicited overall larger mean amplitude responses than did deviant syllables ( $F(1,14) = 8.107$ ,  $p < 0.02$ ,  $\eta^2 = 0.367$ ) (Table 2).<sup>4</sup> Scatterplots of the individual participants' standard and deviant responses to/ba/and/da/are presented in Fig. 2b and c. No effect of electrode location was observed ( $p > 0.06$ ).

A syllable type  $\times$  trial type interaction was found ( $F(1,14) = 9.834$ ,  $p < 0.008$ ,  $\eta^2 = 0.413$ ). Post-hoc ANOVAs revealed a large difference between the/da/standard and deviant mean amplitudes ( $F(1,14) = 17.829$ ,  $p < 0.002$ ,  $\eta^2 = 0.562$ ); no difference was observed for the/ba/standards and deviants ( $p > 0.56$ ). Moreover, the/ba/and/da/standard syllable mean amplitudes differed significantly ( $F(1,14) = 20.043$ ,  $p < 0.002$ ,  $\eta^2 = 0.589$ ) while the deviant trials elicited by the two syllables did not differ ( $p > 0.14$ ).

### 3.1.4. Standard and deviant cluster permutation analyses

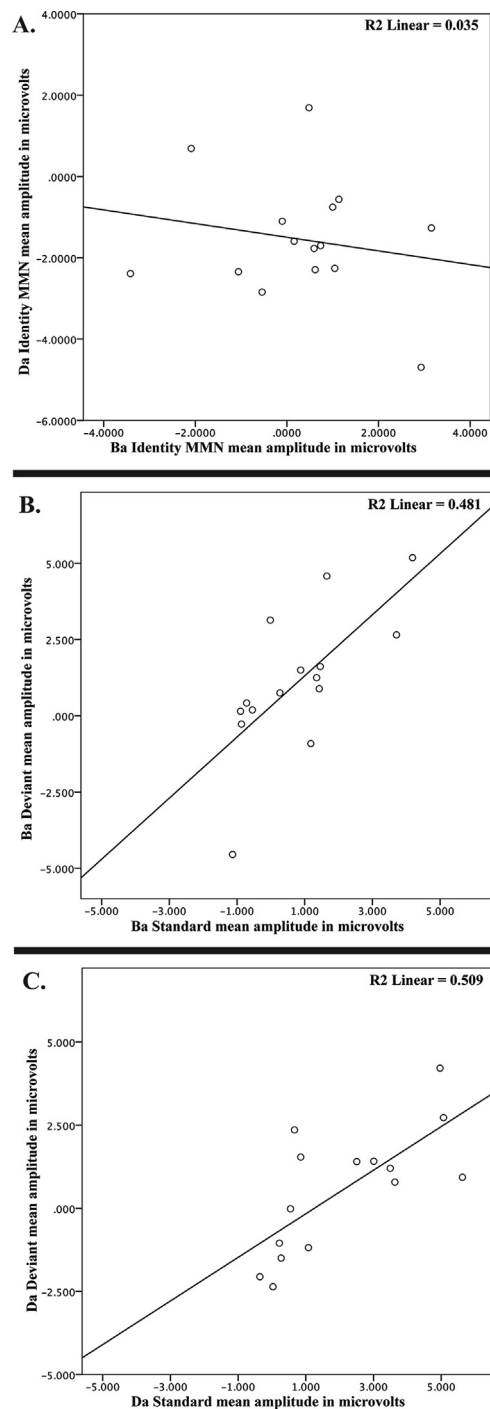
Four cluster-level mass permutation tests encompassing the timeline of the MMN (0–300 ms) were applied to the standard and deviant syllable data. The results of the tests are displayed in raster diagrams in Fig. 3b–d. All cluster permutation tests results were consistent with the ANOVA findings.

When first considering the trial type differences, no significant clusters were identified when examining the difference between/ba/standards and deviants. In contrast, a broadly distributed effect from 78 to 164 ms signified the time period during which the/da/deviants elicited more negative ERP responses than the/da/standards; the smallest significant  $t$ -score was:  $t(14) = 2.1459$ ,  $p < 0.02$ , with a test-wise alpha level of 0.0499. Since this time window primarily encompassed the time period of the N1 ERP response, this negative difference implies that the/da/deviants elicited a larger response than the standards.

When contrasting the syllable type differences, a broadly distributed effect from 101 to 281 ms signified the time period during which the/da/standards elicited more positive ERP responses than the/ba/standards; the smallest significant  $t$ -score was:  $t(14) = 2.1463$ ,  $p < 0.0001$ , with a test-wise alpha level of 0.0498. Since this time window primarily encompassed the

<sup>4</sup> Since this time window primarily encompassed the time period of the P2 ERP response, a smaller mean amplitude measurement is indicative of a more negative response, which is consistent with the fact that these data were used to generate the identity difference waves.

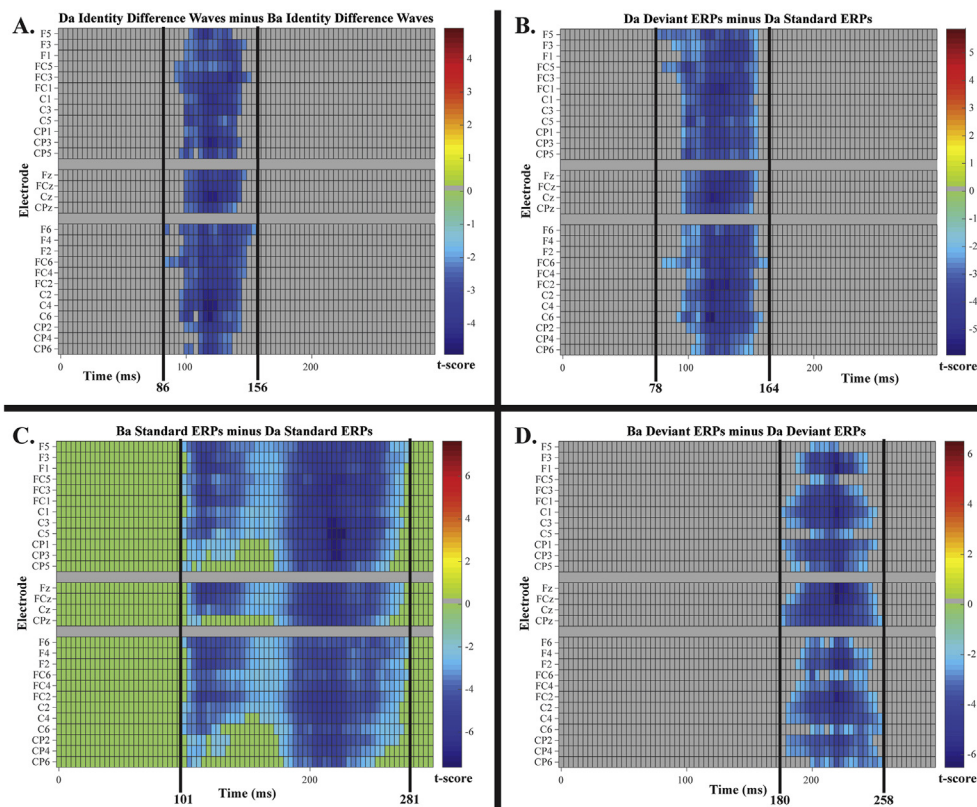




**Fig. 2.** Scatterplots highlighting the variation in individual participants' mean amplitudes of A)/ba/and/da/identity difference waves, B)/ba/standard and deviant ERPs, and C)/da/standard and deviant ERPs.

time period of the P2 ERP response, a positive difference implies that the/da/standards were larger than the/ba/standards. Note that within this broad time window, larger t-scores were observed across two separate time periods, one from approximately 101–150 ms and a second from 180 to 260 ms.

When the/ba/and/da/deviant syllables were contrasted, a broadly distributed effect from 180 to 258 ms signified the time period during which the/da/deviants elicited more positive (i.e., larger) ERP responses than the/ba/deviants; the smallest significant t-score was:  $t(14) = 2.1492$ ,  $p < 0.002$ , with a test-wise alpha level of 0.0459.



**Fig. 3.** A) The raster diagram illustrating differences between the/ba/and/da/identity difference waves is aligned by time below the waveforms. B) Raster diagram illustrating differences between the/da/deviants and standards. C) Raster diagram illustrating differences between/ba/standards and/da/standards. D) Raster diagram illustrating differences between/ba/deviants and/da/deviants. There were no significant clusters for the comparison of/ba/deviants and standards. See text for more detail.

### 3.2. Single-trial GLM results

#### 3.2.1. Group results

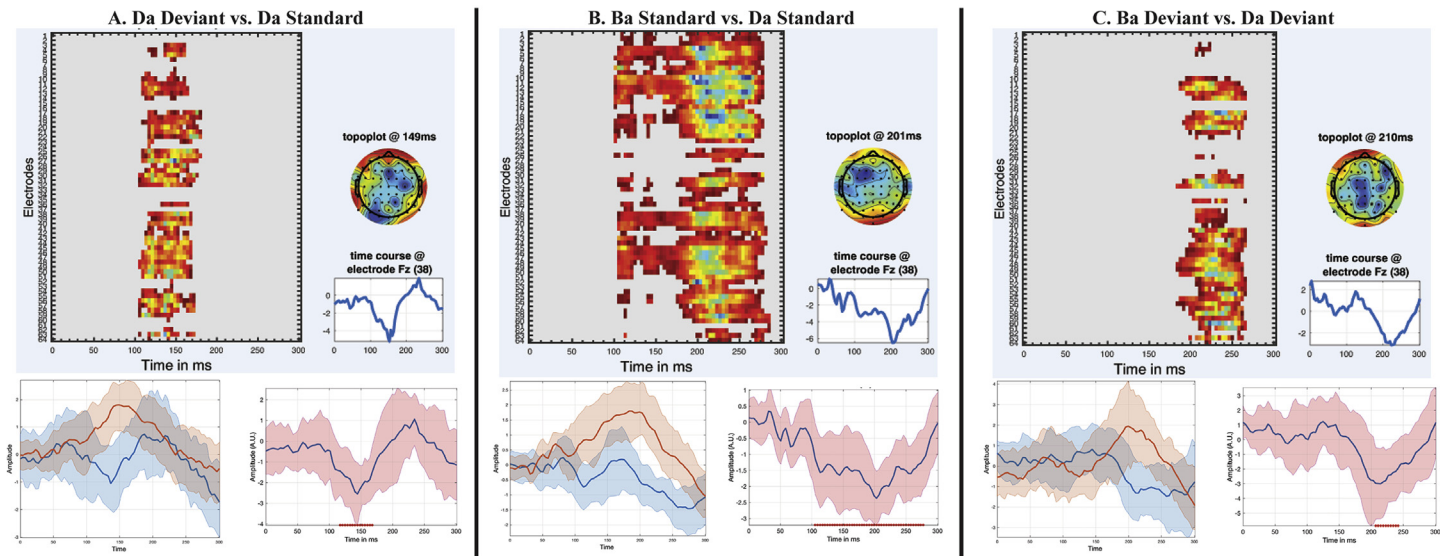
The single-trial group GLM results were consistent with the cluster-based permutation tests generated from the averaged ERPs, though some subtle latency differences were observed due to the nature of the data structure (averaged vs. unaveraged trials). These analyses are plotted in Fig. 4a–c. Each analysis figure has three parts. 1) In the top figure, statistically significant electrodes, based on the paired-*t* tests controlling for multiple comparisons across space and time, are represented by colored rectangles, with the color representing the size of the results. The electrodes are stacked up on the y-axis in the timeline from 0 to 300 ms post-syllable onset. The overall the time course of electrode Fz is represented on the lower right portion of the main figure; the topographic plot capturing Fz at one significant time point is displayed above. 2) The small left graph in each column plots the beta parameters of each comparison at Fz, with 20% trimmed means and a 95% confidence interval represented by the light shading of blue and red. In Figure A, the blue line represents the deviant syllable and the red line represents the standard. In Figures B and C, the blue line represents/ba/and the red line represents/da/. 3) The small right graph in each column shows the computed difference between groups at Fz, with a 95% confidence interval represented in pink; the red horizontal line at the bottom of the figure highlights the significant ( $p < 0.05$ ) time period of the effect.

In comparing the trial types with the/ba/standards and deviants, no paired *t*-test comparisons remained significant at  $p < 0.05$  after controlling for spatio-temporal multiple comparisons. Alternatively, when the/da/standards and deviants were compared and spatio-temporal cluster comparisons were controlled, significant differences were widespread and extended between 104 and 181 ms post-syllable onset (Fig. 4a).

When the/ba/and/da/standards were compared, a large cluster was identified, extending from 100 to 287 ms post-syllable onset (Fig. 4b), with the largest significance values being observed between ca. 200 and 250 ms. Similarly, when the/ba/and/da/deviants were compared, a cluster ranging from 182 to 268 ms post-syllable onset remained after controlling for spatio-temporal multiple comparisons (Fig. 4c).

#### 3.2.2. Single-subject results

The single-subject data provided some interesting insight into how widespread the phonological phenomenon of [coronal] underspecification is across English-speaking adults. Individual participants' full scalp analyses of the four comparisons



**Fig. 4.** Significant group GLM results for A)/da/deviants and standards, B)/ba/standards and da/standards, and C)/ba/deviants and da/deviants. No significant GLM results were observed for the comparison of/ba/deviants and standards. Note that negative amplitudes are plotted down in all waveforms in this figure. See text for more detail.

are presented in Supplementary Files 1–4. While the vast majority of the participants' individual significant effects occurred during the time period of the MMN (approximately 100–200 ms post-syllable onset), some effects were observed both before and after this time window.

Table 3 provides an overview of the single-subject findings. In the/ba/standard and deviant analysis, only two participants had a significant effect at Fz, with 7/15 participants having a significant effect elsewhere across at least five electrodes. This means that less than half of the subjects elicited a significant/ba/identity difference wave. In the/da/standard and deviant analysis, again just two participants showed an effect at Fz, though 12/15 participants demonstrated a significant identity difference wave elsewhere.

When the standards of the two syllables were compared, 12/15 participants demonstrated a significant difference between/ba/and/da/at Fz, with all 15 participants demonstrating a syllable type difference across other scalp locations. The differences between the/ba/and/da/deviants were not quite as prevalent, as just three participants demonstrated an effect at Fz, and 12 participants showed a syllable type difference at other electrode sites.

## 4. Discussion

This study examined phonological underspecification in English-speaking adults by testing the predictions of the FUL theory (Lahiri & Marslen-Wilson, 1991; Lahiri & Reetz, 2002, 2010) with ERP/EEG data. The place of articulation of two early-acquired American English consonants was contrasted in simple consonant-vowel syllables: the [labial]/b/and the [coronal]/d/. Consistent with previous ERP studies of underspecification, a large MMN response was elicited by/da/while no obvious MMN response was elicited by/ba/. Cluster-based permutation tests using averaged ERP data and GLM analyses using unaveraged EEG data identified specific periods of time during which the neural responses elicited by the two syllables differed, thus providing a time course of underspecification. Finally, single-subject GLM analyses showed that most, but not all adults, demonstrated evidence of [coronal] underspecification.

### 4.1. Electrophysiological evidence for underspecification

#### 4.1.1. MMN evidence

Consistent with the prior underspecification studies using German speakers (Cornell et al., 2013, 2011;; Diesch & Luce, 1997; Eulitz & Lahiri, 2004; Scharinger, Bendixen et al., 2012), evidence supporting underspecification of consonantal place of articulation was found. A large and early MMN response was observed in the/da/identity difference wave, while no observable MMN was found in the/ba/identity difference wave. The approximately 70 ms window of significance identified by the cluster analysis indicates that there is a phonological difference between the two syllable types. Thus, the present findings provide the first evidence of [coronal] consonant underspecification in English.

#### 4.1.2. Standard and deviant ERP evidence

While the MMN has been the ERP index of underspecification, interpreting MMN data from a neuro-physiological standpoint can be complex given the nature of how the response is calculated via same-stimulus identity difference waves. It is more straightforward to directly examine the standard and deviant ERPs that are used to create the MMN. One of

**Table 3**

Single-subject results from the four LIMO GLM analyses. Each participant's data were examined for a significant (though uncorrected for multiple comparisons) effect of at least 20 continuous milliseconds at electrode Fz, and across at least five separate electrode sites (that did not have to include Fz). An "X" identifies participants who showed a significant effect for each measurement.

Adult	/ba/Standards vs./ba/ Deviants		/da/Standards vs./da/ Deviants		/ba/Standards vs./da/ Standards		/ba/Deviants vs./da/ Deviants	
	Fz	5 elect.	Fz	5 elect.	Fz	5 elect.	Fz	5 elect.
1		X			X	X		X
2	X	X	X	X	X	X	X	X
3						X		
4					X	X		
5		X		X		X	X	X
6		X		X	X	X		X
7				X	X	X		X
8				X	X	X		X
9		X		X	X	X		X
10	X	X		X	X	X		X
11		X		X	X	X		X
12				X		X		
13				X	X	X	X	X
14				X	X	X		X
15			X	X	X	X		X

the more interesting finding coming out of the standard and deviant analyses of the present study was that the/da/standards elicited significantly larger responses than did the/ba/standards. The FUL MMN theory has focused nearly exclusively on how the deviant syllables affect MMN amplitudes, but the present results suggest that the mean amplitude of the standard syllable also affects MMN generation.

This point is supported by the fact that the mean amplitudes of the/ba/and/da/deviants did not differ during the timeline of the MMN (110–160 ms). The same general underspecification interpretation can be used when discussing the differences in standard syllable amplitudes: The heightened neural response elicited by the/da/standards, in comparison to the/ba/standards, could represent the fact that a wide variety of place of articulation neuronal populations were responding because no place of articulation was specified. Thus, these results suggest that the FUL theory might need to be revised and expanded to not only address the MMN, but the responses of both the standard and deviant stimuli. Evidence for underspecification might be found not only in the difference wave MMN, but also in the responses elicited by the standard syllables.

Both the cluster permutation tests and the GLM analyses showed that the/ba/and/da/deviants differed between approximately 180 and 260 ms. These differences, in conjunction with continued differences between the standard syllables, resulted in P3a-type responses seen in the/da/identity difference waves around 200 ms. Since the P3a is often thought to reflect the process of involuntary attention shifting (Dien, Spencer, & Donchin, 2004; Horváth, Winkler, & Bendixen, 2008; Rinne, Särkkä, Degerman, Schröger, & Alho, 2006), it is not unexpected that its presence was seen in the/da/waveforms and not in the/ba/. The/ba/deviants, presented within the stream of/da/standards, did not attract any additional attention since they essentially met the place of articulation expectations. Alternatively, the/da/deviants did not match the [labial] place of articulation expectations set up by/ba/standards, and thus involuntarily created an attention shift. These results would suggest that along with examining MMN responses, underspecification claims could potentially be bolstered by examining P3a responses as well.

#### 4.1.3. Single-subject results

Underspecification is hypothesized to be a language universal phonological phenomenon, which means that it should be observable, if not prevalent, in all participants. However, interpretation of previous underspecification ERP studies is hindered because of a lack of detailed descriptions or analyses of individual participants (Rousselet & Pernet, 2011; Rousselet, Foxe, & Bolam, 2016). Group-level statistics identify effects that are generally consistent across participants, even if those effects are not significant in individuals. By design, group statistics gloss over the single-subject effects that are inconsistent across subjects, especially if the timing or topographies differ. As a result, at least in theory, group-level ERP statistics can be responsible for both false positives and false negatives (Rousselet et al., 2011; Wagenmakers, 2007). This issue is even more important when examining MMN responses, which are generated via a subtraction process that arguably can distort individual results elicited by the standard and deviant stimuli.

The single-subject GLM analyses of the standard and deviant stimuli were used to test the FUL theory of underspecification to see if, in fact, all participants demonstrated [coronal] underspecification. Here, the comparison of the/da/deviants and standards (which created the/da/identity difference wave) was potentially most telling. While most participants demonstrated a significant difference between trial types, which resulted in a visible MMN in the difference wave, three of the 15 did not. The fact that these participants generated no MMN for the/da/stimuli should lend some caution to broad-based universal claims of underspecification.

To be a language-universal theory, an effect should be apparent in all individuals. Given that even with an excellent signal-to-noise ratio, not every participant demonstrates an MMN response (Bishop & Hardiman, 2010; Brandmeyer, Farquhar, McQueen, & Desain, 2013; Diesch & Luce, 1997; Pettigrew et al., 2004), it might be useful to examine evidence of underspecification elsewhere in ERP and EEG data. An alternate potential index of underspecification could instead be examined directly in the responses elicited by the standard syllables. Every participant in the present study demonstrated significant EEG differences between the two standard syllables, with two potential time periods of interest broadly corresponding to the auditory N1 and P2 ERP peaks. Future studies should provide single-subject data of standard and deviant stimuli to provide further evidence for the prevalence of underspecification in all participants.

In addition, the single-subject analyses provided evidence that examining underspecification at just a few electrodes (e.g., Fz) across a brief period of time, might result in null findings. Few participants demonstrated significant effects for any of the analyses at Fz, while other electrode locations across the scalp were often involved. These topography differences could indicate that different neural systems were being recruited across participants for phonological processing. Additional analyses of independent components and/or dipoles would support this idea.

Further, the timeline of the effects was often broadly distributed across participants, which most likely partially represented individual differences in phonological perception, encoding, and processing time. These outcomes provide evidence against the idea of traditional peak-picking measures of mean amplitude and latency values of a pre-specified time window. Permutation or GLM-type analyses involving large time windows and large numbers of electrodes will likely yield more informative data.

#### 4.2. Applying underspecification theory to ERP data

Phonological underspecification is thought to make speech processing easier, as only contrastive or not otherwise predictable phonological information (i.e., distinctive features) is stored for each phoneme. Underspecification suggests that



some phonemes are not stored with as much phonological information as others. Specifically, FUL (Lahiri & Marslen-Wilson, 1991; Lahiri & Reetz, 2002, 2010) predicts that [coronal] exists, but is underspecified as compared to other place of articulation features, such as [labial].

From a neuro-physiological standpoint, an underspecified phoneme could recruit more neuronal populations than a more specified phoneme since the neurons are not coded to not respond. With a larger number of neurons responding, an underspecified phoneme would elicit a larger response than a more specified phoneme. Evidence for a broad, less specified neural response was provided by the standard syllable data: The/da/standards elicited a much larger response than did the/ba/standards.

Interpreting MMN data from a neuro-physiological standpoint can be complex given the nature of how the response was calculated via same-stimulus identify difference waves. To explain, the ERP responses elicited by the standard stimuli, /ba/or/da/, represent the baseline level of neural activation necessary for processing those syllables. When the/ba/deviant was presented within the stream of/da/standards, no place of articulation was established by the/da/standard. This no-mismatch situation meant that the [labial] place of articulation of/ba/met the [no] place of articulation requirements established by the/da/standard, and no heightened neural response was necessary above and beyond what would have been typically elicited by that syllable (i.e., what was elicited by the/ba/standard). Thus, essentially the same level of neuronal activation was elicited by the/ba/standard and/ba/deviant, resulting in the nearly non-existent MMN response.

In contrast, the/ba/standards explicitly set up the expectation for a response from only the [labial] neuronal populations. The heightened neural response elicited by the/da/deviants, in comparison to the/da/standards, could represent the fact that a wide variety of place of articulation neuronal populations were trying to match that place of articulation. That is, since no one set of neurons was established as the “place of articulation” neuronal population for/d/, many neuronal populations increased their activation levels, resulting in the heightened MMN response.

While the evidence in the present paper strongly supports the idea of phonological underspecification as proposed by FUL, some aspects of these data could potentially also fit within a memory/usage-based account of language (UBA) (Bybee, 2001, 2002, 2010; Pierrehumbert, 2006).<sup>5</sup> To explain UBA briefly, the [coronal] category contains many more consonants (14: /t d s z n l ʃ j ø ð ʒ ʒ ʃ ʒ ʒ/) in English than does the labial (4: /p b m w/). Given the large number of consonants in the [coronal] category, /d/ can be said to have a “dense neighborhood” while /b/ could be considered to have a “small neighborhood density”. Following UBA, neural responses would be generated based on the neighborhood density of a phoneme. That is, for the MMN response, with/da/as the standard and/ba/as the deviant, it could be difficult for the system to create a strong prediction due to the large neighborhood of/d/, which creates a no mismatch situation. On the other hand, when/ba/is the standard, the system can establish a strong prediction due to the small neighborhood of/b/, which creates a true mismatch situation with/da/.

One way to further test UBA predictions beyond the/ba/and/da/syllable data of the present study is to examine another phoneme, such as [dorsal]/g/. The neighborhood density of/g/is identical to the [labial] phoneme, /b/, as it contains just four phonemes: /k g ŋ w/. Thus, following the above neighborhood density argument for/b/and/d/, it would be predicted that/g/should perform similarly to/b/. The present study incorporated/ga/, but only as a deviant stimulus in both stimulus sets. Since/ga/was never a standard stimulus, identity difference waves cannot be determined for it; thus, neither the MMN response nor the standard response elicited by/ga/can be examined. This usage-based hypothesis will remain a possibility until the predictions of UBA and underspecification are jointly tested. Fully-crossed stimulus sets with a variety of phonemes should be incorporated to elicit a wide range of ERP data, which can then be quantitatively examined to provide support for one theory or the other.

## Acknowledgements

This research was supported by NIH grant numbers R15DC013359 (from the National Institute On Deafness And Other Communication Disorders) and C06RR022088 (from the National Center for Research Resources) awarded to the first author. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We would like to thank Janet Babchishin, Brianna Jallo, Megan Nauman, Meghan Macaulay, Courtney Rowan, and Sheila Cassidy for help with the testing of the participants. The authors do not have any financial or non-financial relationships relevant to the content of this manuscript.

## Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.jneuroling.2017.05.003>.

## References

- Bieniek, M. M., Pernet, C. R., & Rousselet, G. A. (2012). Early ERPs to faces and objects are driven by phase, not amplitude spectrum information: Evidence from parametric, test-retest, single-subject analyses. *Journal of Vision*, 12(13), 12. <https://doi.org/10.1167/12.13.12>.

<sup>5</sup> We owe this interpretation to an anonymous reviewer.

- 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. <https://doi.org/10.1111/j.1469-8986.2009.00970.x>.
- Brandmeyer, A., Farquhar, J. D. R., McQueen, J. M., & Desain, P. W. M. (2013). Decoding speech perception by native and non-native speakers using single-trial electrophysiological data. *PLoS One*, 8(7), e68261. <https://doi.org/10.1371/journal.pone.0068261>.
- Bullmore, E. T., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E., & Brammer, M. J. (1999). Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. *IEEE Transactions on Medical Imaging*, 18(1), 32–42.
- Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.
- Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, 14, 261–290.
- Bybee, J. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press. Retrieved from [https://www.amazon.com/Language-Usage-Cognition-Joan-Bybee/dp/0521616832/ref=sr\\_1\\_1?ie=UTF8&qid=1467564600&sr=8-1&keywords=Bybee%2C+Joan.+2010.+Language%2C+usage+and+cognition.+Cambridge%3A+Cambridge+University+Press](https://www.amazon.com/Language-Usage-Cognition-Joan-Bybee/dp/0521616832/ref=sr_1_1?ie=UTF8&qid=1467564600&sr=8-1&keywords=Bybee%2C+Joan.+2010.+Language%2C+usage+and+cognition.+Cambridge%3A+Cambridge+University+Press).
- Čeponiene, R., Yaguchi, K., Shestakova, A., Alku, P., Suominen, K., & Näätänen, R. (2002). Sound complexity and “speechness” effects on pre-attentive auditory discrimination in children. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, 43(3), 199–211.
- Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.
- Cornell, S. A., Lahiri, A., & Eulitz, C. (2011). “What you encode is not necessarily what you store”: Evidence for sparse feature representations from mismatch negativity. *Brain Research*, 1394, 79–89. <https://doi.org/10.1016/j.brainres.2011.04.001>.
- Cornell, S. A., Lahiri, A., & Eulitz, C. (2013). Inequality across consonantal contrasts in speech perception: Evidence from mismatch negativity. *Journal of Experimental Psychology. Human Perception and Performance*, 39(3), 757–772. <https://doi.org/10.1037/a0030862>.
- Csépe, V. (1995). On the origin and development of the mismatch negativity. *Ear and Hearing*, 16(1), 91–104.
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>.
- Dien, J., Spencer, K. M., & Donchin, E. (2004). Parsing the late positive complex: Mental chronometry and the ERP components that inhabit the neighborhood of the P300. *Psychophysiology*, 41(5), 665–678. <https://doi.org/10.1111/j.1469-8986.2004.00193.x>.
- Diesch, E., & Luce, T. (1997). Magnetic mismatch fields elicited by vowels and consonants. *Experimental Brain Research*, 116(1), 139–152.
- Eulitz, C., & Lahiri, A. (2004). Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition. *Journal of Cognitive Neuroscience*, 16(4), 577–583.
- Friedrich, C. K., Eulitz, C., & Lahiri, A. (2006). Not every pseudoword disrupts word recognition: An ERP study. *Behavioral and Brain Functions: BBF*, 2(36). <https://doi.org/10.1186/1744-9081-2-36>.
- Gaskell, M. G., & Marslen-Wilson, W. D. (1996). Phonological variation and inference in lexical access. *Journal of Experimental Psychology. Human Perception and Performance*, 22(1), 144–158.
- Gaskell, M. G., & Marslen-Wilson, W. D. (1998). Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology. Human Perception and Performance*, 24(2), 380–396.
- Gimson, A. C. (1989). *An introduction to the pronunciation of English* (Fourth ed.). London: Edward Arnold.
- Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, 48(12), 1711–1725. <https://doi.org/10.1111/j.1469-8986.2011.01273.x>.
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1–2), 67–99. <https://doi.org/10.1016/j.cognition.2003.10.011>.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews. Neuroscience*, 8(5), 393–402. <https://doi.org/10.1038/nrn2113>.
- Horváth, J., Winkler, I., & Bendixen, A. (2008). Do N1/MMN, P3a, and RON form a strongly coupled chain reflecting the three stages of auditory distraction? *Biological Psychology*, 79(2), 139–147. <https://doi.org/10.1016/j.biopsycho.2008.04.001>.
- Jakobson, R., Fant, G., & Halle, M. (1952). *Preliminaries to speech analysis. the distinctive features and their correlates*. Cambridge, MA: MIT, Acoustics Laboratory. Technical Report No. 13.
- Johnson, K. (1997). Speech perception without speaker normalization. In K. Johnson, & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 145–166). New York: Academic Press.
- Johnson, K. (2005). Speaker normalization in speech perception. In D. B. Pisoni, & R. E. Remez (Eds.), *The handbook of speech perception* (pp. 363–389). Malden, MA: Blackwell Publishing.
- Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics*, 34(4), 485–499. <https://doi.org/10.1016/j.wocn.2005.08.004>.
- Jung, T. P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2000). Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 111(10), 1745–1758.
- Kabak, B. (2007). Hiatus resolution in Turkish: An underspecification account. *Lingua*, 117(8), 1378–1411.
- Kiparsky, P. (1985). Some consequences of lexical phonology. *Phonology Yearbook*, 2, 85–138.
- Korpilahti, P., Krause, C. M., Holopainen, I., & Lang, A. H. (2001). Early and late mismatch negativity elicited by words and speech-like stimuli in children. *Brain and Language*, 76(3), 332–339. <https://doi.org/10.1006/brln.2000.2426>.
- Ladefoged, P. (2007). Articulatory features for describing lexical distinctions. *Language*, 83(1), 161–180.
- Lahiri, A. (2007, August). Non-equivalence between phonology and phonetics. In *Presented at the 16th international congress of phonetic sciences* (Saarbrücken, Germany).
- Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition*, 38(3), 245–294.
- Lahiri, A., & Reetz, H. (2002). Underspecified recognition. In C. Gussenhoven, & N. Warner (Eds.), *Labphon* (Vol. 7, pp. 637–676). Berlin: Mouton.
- Lahiri, A., & Reetz, H. (2010). Distinctive features: Phonological underspecification in representation and processing. *Journal of Phonetics*, 38, 44–59.
- Luck, S. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press. Retrieved from <https://mitpress.mit.edu/books/introduction-event-related-potential-technique>.
- Luck, S., & Lopez-Calderon, J. (2012). *ERPLAB toolbox (Version 3.0.2.1)*. Davis: University of California.
- Maddieson, I. (1984). *Patterns of sounds*. Cambridge University Press.
- Manly, B. F. J. (2006). *Randomization, bootstrap and Monte Carlo methods in biology* (Third Ed.). CRC Press.
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164(1), 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>.
- Morr, M. L., Shafer, V. L., Kreuzer, J. A., & Kurtzberg, D. (2002). Maturation of mismatch negativity in typically developing infants and preschool children. *Ear and Hearing*, 23(2), 118–136.
- Näätänen, R. (1995). The mismatch negativity: A powerful tool for cognitive neuroscience. *Ear and Hearing*, 16(1), 6–18.
- Näätänen, R., Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42(4), 313–329.
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., ... Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, 385(6615), 432–434. <https://doi.org/10.1038/385432a0>.

- Näätänen, R., Paavilainen, P., & Reinikainen, K. (1989). Do event-related potentials to infrequent decrements in duration of auditory stimuli demonstrate a memory trace in man? *Neuroscience Letters*, 107(1–3), 347–352.
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, 118(12), 2544–2590. <https://doi.org/10.1016/j.clinph.2007.04.026>.
- Näätänen, R., Teder, W., Alho, K., & Lavikainen, J. (1992). Auditory attention and selective input modulation: A topographical ERP study. *Neuroreport*, 3(6), 493–496.
- Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*, 125(6), 826–859.
- Paradis, C., & Prunet, J.-F. (1991). *The special status of Coronals: Internal and external Evidence: Phonetics and phonology*. San Diego, CA: Academic Press.
- Pernet, C. R., Chauveau, N., Gaspar, C., & Rousselet, G. A. (2011). LIMO EEG: A toolbox for hierarchical Linear MOdeling of ElectroEncephaloGraphic data. *Computational Intelligence and Neuroscience*, 2011(831409). <https://doi.org/10.1155/2011/831409>.
- Pettigrew, C. M., Murdoch, B. E., Ponton, C. W., Finnigan, S., Alku, P., Kei, J., ... Chenery, H. J. (2004). Automatic auditory processing of english words as indexed by the mismatch negativity, using a multiple deviant paradigm. *Ear and Hearing*, 25(3), 284–301.
- Picton, T. W., Alain, C., Otten, L., Ritter, W., & Achim, A. (2000). Mismatch negativity: Different water in the same river. *Audiology & Neuro-Otology*, 5(3–4), 111–139. <https://doi.org/10.1016/j.jam.2007.04.001>.
- Pierrehumbert, J. (2006). The next toolkit. *Journal of Phonetics*, 34, 516–530.
- Pisoni, D. B. (1993). Long-term memory in speech perception: Some new findings on talker variability, speaking rate and perceptual learning. *Speech Communication*, 13(1–2), 109–125.
- Ranbom, L. J., & Connine, C. M. (2007). Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language*, 57(2), 273–298. <https://doi.org/10.1016/j.jml.2007.04.001>.
- Rinne, T., Särkkä, A., Degerman, A., Schröger, E., & Alho, K. (2006). Two separate mechanisms underlie auditory change detection and involuntary control of attention. *Brain Research*, 1077(1), 135–143. <https://doi.org/10.1016/j.brainres.2006.01.043>.
- Rousselet, G. A., Foxe, J. J., & Bolam, J. P. (2016). A few simple steps to improve the description of group results in neuroscience. *The European Journal of Neuroscience*, 44(9), 2647–2651. <https://doi.org/10.1111/ejn.13400>.
- Rousselet, G. A., Gaspar, C. M., Wiczorek, K. P., & Pernet, C. R. (2011). Modeling single-trial ERP reveals modulation of bottom-up face visual processing by top-down task constraints (in some subjects). *Frontiers in Psychology*, 2(137). <https://doi.org/10.3389/fpsyg.2011.00137>.
- Rousselet, G. A., & Pernet, C. R. (2011). Quantifying the time course of visual object processing using ERPs: It's time to up the game. *Frontiers in Psychology*, 2(107). <https://doi.org/10.3389/fpsyg.2011.00107>.
- Salvia, E., Bestelmeyer, P. E. G., Kotz, S. A., Rousselet, G. A., Pernet, C. R., Gross, J., et al. (2014). Single-subject analyses of magnetoencephalographic evoked responses to the acoustic properties of affective non-verbal vocalizations. *Frontiers in Neuroscience*, 8(422). <https://doi.org/10.3389/fnins.2014.00422>.
- Scharinger, M., Bendixen, A., Trujillo-Barreto, N. J., & Obleser, J. (2012). A sparse neural code for some speech sounds but not for others. *PLoS One*, 7(7), e40953. <https://doi.org/10.1371/journal.pone.0040953>.
- Scharinger, M., Monahan, P. J., & Idsardi, W. J. (2012). Asymmetries in the processing of vowel height. *Journal of Speech, Language, and Hearing Research: JSLHR*, 55(3), 903–918. [https://doi.org/10.1044/1092-4388\(2011/11-0065\)](https://doi.org/10.1044/1092-4388(2011/11-0065)).
- Scharinger, M., Merickel, J., Riley, J., & Idsardi, W. J. (2011). Neuromagnetic evidence for a featural distinction of english consonants: Sensor- and source-space data. *Brain and Language*, 116(2), 71–82. <https://doi.org/10.1016/j.bandl.2010.11.002>.
- Shafer, V. L., Morr, M. L., Kreuzer, J. A., & Kurtzberg, D. (2000). Maturation of mismatch negativity in school-age children. *Ear and Hearing*, 21(3), 242–251.
- Snoeren, N. D., Gaskell, M. G., & Di Betta, A. M. (2009). The perception of assimilation in newly learned novel words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(2), 542–549. <https://doi.org/10.1037/a0014509>.
- Stemberger, J. P., & Stoel-Gammon, C. (1991). The underspecification of coronals: Evidence from language acquisition and performance errors. *Phonetics and Phonology*, 2, 181–199.
- Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement Science Review*, 2(2), 1–12.
- Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, 14(5), 779–804.
- Walter, M. A., & Hacquard, V. (2004). MEG evidence for phonological underspecification. In *Proceedings of the 14th biennial BIOMAG conference* (Boston, MA).
- Wheeldon, L., & Waksler, R. (2004). Phonological underspecification and mapping mechanisms in the speech recognition lexicon. *Brain and Language*, 90(1–3), 401–412. [https://doi.org/10.1016/S0093-934X\(03\)00451-6](https://doi.org/10.1016/S0093-934X(03)00451-6).
- Wilcox, R. R. (2005). *Introduction to robust estimation and hypothesis testing* (Second ed.). New York, NY: Elsevier Academic Press.
- Winkler, I., Lehtokoski, A., Alku, P., Vainio, M., Czigler, I., Csépe, V., ... Näätänen, R. (1999). Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations. *Brain Research. Cognitive Brain Research*, 7(3), 357–369.
- Woodman, G. F. (2010). A brief introduction to the use of event-related potentials in studies of perception and attention. *Attention, Perception & Psychophysics*, 72(8), 2031–2046. <https://doi.org/10.3758/APP.72.8.2031>.
- Yip, M. (1991). Coronals, consonant clusters, and the coda condition. In C. Paradis, & J. F. Prunet (Eds.), *The special status of coronals* (pp. 61–78). San Diego, CA: Academic Press.
- Zimmerer, F., Reetz, H., & Lahiri, A. (2009). Place assimilation across words in running speech: Corpus analysis and perception. *The Journal of the Acoustical Society of America*, 125(4), 2307–2322. <https://doi.org/10.1121/1.3021438>.