(A)symmetry in vowel features in verbs and pseudoverbs: ERP evidence

Sandra Kotzor[1, 2*], Beinan Zhou[1], & Aditi Lahiri[1]

*corresponding author (sandra.kotzor@ling-phil.ox.ac.uk)

[1]Faculty of Linguistics, Philology and Phonetics
University of Oxford
Walton Street
Oxford
OX1 2HG, United Kingdom
sandra.kotzor@ling-phil.ox.ac.uk
aditi.lahiri@ling-phil.ox.ac.uk

[2]School of Education
Oxford Brookes University
Harcourt Hill Campus
Oxford
OX2 9AT, United Kingdom

Neuropsychologia, 143 (2020) 107474 (pages1-8)

Abstract

This paper examines the processing of height and place contrasts in vowels in words and pseudowords, using mismatch negativity (MMN) to determine firstly whether asymmetries resulting from underlying representations found in the processing of vowels in isolation will remain in a word context and secondly whether there is any difference in the way these phonological differences manifest in pseudowords. The stimuli are two sets of English ablaut verbs and corresponding pseudowords (*sit~sat*/**sif~*saf* and *get~got*/**gef~*gof*) contrasting in vowel height ([HIGH] vs. [LOW]) and place of articulation ([CORONAL] vs. [DORSAL]). In line with previous research, the results show a processing asymmetry for place of articulation in both words and nonwords, while different vowel heights result in symmetrical MMN patterns. These findings confirm that an underspecification account provides the best explanation for featural processing and that phonological information is independent of lexical status.

Keywords: speech processing; lexical representation; featural underspecification; mismatch negativity; phonology

1. Introduction

In the languages of the world, we find both symmetries and asymmetries. Although it may seem counterintuitive at first glance, asymmetries are not only common but seem to constitute a key mechanism in language processing. Evidence for asymmetrical patterns in phonological processing has been mounting and studies range from earlier work using gating paradigms (Lahiri & Marslen-Wilson, 1992; Nix et al., 1993; Marslen-Wilson et al., 1995) to more recent psycholinguistic and neurophysiological studies (Eulitz & Lahiri, 2004; Wheeldon & Waksler, 2004; Friedrich et al., 2006, 2008; Cornell et al., 2011, 2013; Roberts et al., 2012, 2014; Kotzor et al., 2016).

The principle underlying these asymmetrical patterns is assumed to be one of economy. Spoken language comprehension is a complex task as the input is highly variable and the perceptual environment may contain a large number of possible options. For example, Turkish listeners need to distinguish four high vowels [i y ɯ u], German listeners three [i y u] and English only two [i u]. In an asymmetric system, certain sounds are easier to process and correctly detect since some features may be fully specified in the mental representation while others may be underspecified. As a consequence, the perceptual system has fewer decisions to make in certain situations, thus reducing the available options and increasing the detection rate. The most crucial prerequisite for such a processing strategy is a system which assumes different amounts of stored information to identify particular speech sounds or phonemes made up of a combination of features. To enable this differentiation, it is necessary to assume that not all features are stored.

One model assuming asymmetric representation is the *Featurally Underspecified Lexicon* (FUL) Model (Lahiri & Reetz, 2002, 2010; Lahiri, 2012), which proposes that not all features present in the surface representation of a sound (speech signal) are necessarily stored in its lexical representation. Unlike other models which propose rich storage of representations (such as exemplar models: Johnson, 1997; Goldinger, 1998; Pierrehumbert, 2002), FUL argues for an abstract representation. There is a select number of features in FUL which have been shown to be *underspecified* – most notably the place feature [CORONAL] (Eulitz & Lahiri, 2004; Friedrich et al., 2006, 2008; Cornell et al., 2011; Roberts et al., 2012, 2014) and the consonantal manner feature

[PLOSIVE] (Cornell et al., 2013). The consequence of asymmetry in representation is asymmetry in processing.

The FUL approach allows for a degree of flexibility in speech processing which is necessary to deal with, for example, predictable place assimilations such as *hand bag* being pronounced as *ha*[mb]*ag*, where the [LABIAL] feature of [b] in *bag* is transferred onto the preceding nasal [n] of *hand* turning it into [m]. The reverse assimilation would be unusual: in *lamb dish*, the final [m] of *lamb* remains unaffected by the initial sound [d] of *dish* and would not change into **la*[nd]*ish.* Thus, phonetic [m] activates /m/ and /n/ but phonetic [n] activates only /n/.

In FUL, the assumption is that although /m/ and /n/ are both specified as [NASAL], /m/ is specified for its place of articulation [LABIAL] while /n/ is unspecified for [CORONAL]. The asymmetry in processing comes about in the following way. The [m] in *ha*[mb]*ag* can activate the unspecified /n/ of *hand* (***no-mismatch***). Thus, both *ha*[m] and *ha*[n] are accepted by the representation *hand.* In contrast, if the [m] in *lamb* is mispronounced as [n], the output **la*[n] would mismatch with the fully specified [LABIAL] /m/. Compared to mismatch cases where the features in question are both fully specified and thus do not tolerate a sound with a different feature, no-mismatch conditions have been shown to result in faster reaction times in behavioural experiments (Bölte & Coenen, 2000; Roberts et al., 2012, 2014) and attenuated MMN responses in neurophysiological studies (cf. Eulitz & Lahiri, 2004; Cornell et al., 2011, 2013; Roberts et al., 2012).

The same logic can be applied to vowels; earlier MMN studies showed that the German vowels [o ø e] resulted in asymmetric processing due to differences in the feature specifications in their representations (Eulitz & Lahiri, 2004). Most of the research on processing asymmetries, however, has focused on isolated vowel and consonantal phonemes or on pseudoword syllables (Cornell et al., 2013). The assumption has been that phonemes have representations independent of lexical entries for words. There are, of course, words consisting of a single vowel in many languages (cf. Bengali [o] 'he/she', French *et* [ɛ] 'and', Italian *o* [ɔ] 'or', but other languages, such as English, for instance, only allow diphthongs as single-phoneme words, e.g. *I* [aɪ] and *Oh?* [oʊ].

In this paper, we examine the notion of asymmetric feature representation in vowels in both real words and pseudowords to determine whether lexical status affects the representation and processing of the individual phonemes. We examine the vowel height contrast in the ablaut pair *sit~sat* (and the corresponding pseudoword pair **sif~*saf*) where both vowels are fully specified for height, and the place of articulation contrast in *get~got* (and **gef~*gof*) where /ɛ/, as it is [CORONAL], is underspecified for place. All real word stimuli comprise the present and past tense of highly frequent English verbs and the pseudowords are matched in structure as closely as possible. The questions we address are the following: (i) are feature representations for individual phonemes the same as those in existing words and in pseudowords?; (ii) do we see similar asymmetries and symmetries in processing for identical features in words vs. pseudowords? We examine these potential asymmetries using event-related potentials (ERPs), specifically mismatch negativity (MMN) to investigate the featural differences. It has been argued that pseudowords, as they do not have a stored representation and therefore are processed differently (Shtyrov et al., 2010) show different MMN patterns. In what follows, we first provide an overview of the relevant MMN literature and focus on previous evidence for processing asymmetries in vowels.

While vowel height, which is assumed to be fully specified (cf. Lahiri & Reetz, 2002, 2010; Lahiri, 2012), has not yet been examined, other fully specified contrasts, such as manner features [NASAL] /n/ and [STRIDENT] /z/ have resulted in symmetrical mismatch. In a study comparing medial consonants in nonwords in German (e.g. **eni~*ezi *edi~*eni*), Cornell et al., (2013) found symmetrical mismatches for those comparisons which involved [NASAL] (**eni*) and [STRIDENT] (**ezi*) features (both features specified). However, those comparisons involving [PLOSIVE] (underspecified) and [NASAL] (specified) (**edi~*eni*) showed the asymmetry typically resulting from featural underspecification as outlined above, with an attenuated MMN response when the underspecified item served as the standard.

1.1 Mismatch negativity (MMN)

The MMN, an ERP component mainly triggered by automatic pre-attentive change detection in auditory processing (cf. Näätänen & Picton, 1987; Näätänen, 2001; Näätänen et al., 2007 for a review), has previously shown sensitivity to differences on the level of phonological representation and has thus been a key component in establishing processing asymmetries such as those discussed above (cf. Eulitz & Lahiri, 2004; Cornell et al., 2011; 2013; Roberts et al., 2013; de Jonge & Boersma, 2015; Kotzor et al., 2017; Højlund et al. 2019).

Differences in the MMN response have been demonstrated for several linguistic variables such as lexical status (i.e. the presence or absence of a lexical entry) with several studies reporting a larger (and sometimes earlier) MMN response for existing words (cf. among others Pulvermüller et al., 2001, 2004; Shtyrov & Pulvermüller, 2002; Pulvermüller & Shtyrov, 2006; Näätänen et al., 2007; Shtyrov et al., 2010; Bakker et al., 2013), place of articulation (Eulitz & Lahiri, 2004; Cornell et al., 2011) and manner of articulation (Cornell et al., 2013).

The lexical MMN (Shtyrov et al., 2010), in which lexical items elicit a larger MMN response than pseudowords, is based on the activation of a dense neural network in the case of the real word while pseudowords do not activate a stored representation but merely an acoustic neural network in addition to their phonemic information (cf. Pulvermüller et al., 2012, 2014). A 'syntactic MMN' response has also previously been demonstrated for morphologically illegal stimuli (e.g. *flied*) or ungrammatical combinations of items (e.g. incorrect determiner + noun phrase as in German *die* (fem.) *Apfel* (masc.) 'the apple') which has been proposed to be an indicator of morphosyntactic parsing (cf. among others Pulvermüller & Shtyrov, 2003; Hasting et al., 2007; Bakker et al., 2013) with ungrammatical sequences eliciting larger MMN responses. In addition, a priming effect has been observed which accounts for the reduction in the MMN when a deviant is presented after a morphologically related standard (cf. Pulvermüller & Shtyrov, 2003; Bakker et al., 2013).

As crucial acoustic information is already being processed in the early 200ms time window of speech processing (cf. Marslen-Wilson, 1973; Marslen-Wilson & Welsh, 1978; Rastle et al., 2000; Mohr & Pulvermüller, 2002), it is thus clear that the MMN, as an early component, is a useful tool to investigate the nature of the information which affects processing at this early stage. In addition, the MMN is not susceptible to attentional, task-related or strategic biases (Pulvermüller & Shtyrov, 2006: 51) and allows for precise control of stimuli to ensure minimal variation.

1.2. Previous evidence for processing asymmetries in vowels

Evidence for processing asymmetries has been mounting over the last decade and studies employing several different methodologies have provided support for the underspecification of both [PLOSIVE] and [CORONAL]. Since this paper is concerned with the difference between place and height features in vowels, only the evidence for underspecification of place of articulation will be discussed in detail.

In an MMN study, Eulitz and Lahiri (2004) compared the German vowels [e], [o] and [ø] in a standard oddball paradigm. In the underlying representations of these vowels, [e] and [ø] are underspecified for place as they are [CORONAL] while [o] is specified for [DORSAL] (cf. Table 1). As underspecified features are extracted from the speech signal but are not present in the underlying representation, the predicted MMN effect depends on whether a given stimulus is presented as a deviant or a standard. The assumption is that the frequently-occurring standard causes the underlying representation of the expected vowel to be pre-activated, and a mismatch occurs when the appearance of a deviant causes the listener to extract features from the incoming signal that may conflict with this underlying representation. Thus, based on FUL, Eulitz and Lahiri (2004) made the following predictions: both an [e] deviant in a context of an [ø] standard and vice versa should result in a no-mismatch response since both are underspecified for place and thus, even though [CORONAL] is extracted from the signal of the deviant, there is nothing to match it against in the underlying representation activated by the standard. In this case, MMN responses of equivalent amplitude are predicted regardless of which is the standard. In the case of [o] and [ø], however, a difference in MMN amplitude is expected depending on which of the two vowels is presented as the deviant. An [o] deviant in the context of [ø] would still be considered a no-mismatch situation as the standard ([ø]) is underspecified for place and thus the extracted [DORSAL] place feature does not mismatch. In the reverse case, however, the underlying representation of the [DORSAL] standard mismatches with the incoming [CORONAL] place feature extracted from the [ø] deviant.

Table 1 *Feature specifications and asymmetries for German vowels*

Vowel	Surface representation Place		Underlying representation Place
o	[DORSAL] [LABIAL]	*match*	[DORSAL] [LABIAL]
ø	[CORONAL] [LABIAL]	*no-mismatch*	[—] [LABIAL]
e	[CORONAL]	*no-mismatch*	[—] [—]

Cornell et al. (2011) conducted an MMN study using the same vowels as Eulitz and Lahiri 2004 but embedded in both pseudowords and real words in German. Their results replicated the findings of Eulitz and Lahiri 2004 and showed the same processing asymmetry where an identical acoustic difference between standard and deviant elicited MMNs of different amplitudes depending on which of the stimuli was underspecified. This, again, supports coronal underspecification and highlights the MMN as a reliable detector of phonologically-based processing asymmetries.

Processing asymmetries resulting from underspecified representations have also been found in vowels by de Jonge and Boersma (2015) in an MMN study investigating vowel height as well as place in the French vowels [y u ø o]. Their findings provide additional support for coronal underspecification where the change from a coronal standard to a dorsal or labial deviant elicited smaller MMN responses than that from a dorsal/labial standard to a coronal deviant. In terms of vowel height, de Jonge and Boersma (2015) found that French high-mid vowels are underspecified for vowel height and thus a similar asymmetric pattern is observed, with larger MMN responses for mid-high deviants in the context of high standards than in the reverse condition.

Speech discrimination work by Polka & Bohn (2003, 2011) followed by Masapollo, Polka and Ménard (2017) suggests that vowel perception is asymmetric with respect to the concentration of acoustic energy. They propose that when the adjacent formants (F1 and F2 in particular) of a vowel are closer in frequency, these vowels are considered more 'focal' (Polka & Bohn, 2003, 2011). In their experiments, listeners found it easier to discriminate between two vowels when the change was from a more focal to a more peripheral vowel. Under this description it is easier to perceive a change from [ɛ] to [i] than from [i] to [ɛ]. Such asymmetries are driven purely by the information, and the concentration of energy, in the acoustic space rather than any phonological representational differences. While these studies provide no direct comparisons of the vowel pairs used in the present study, this view will be taken into account in our predictions presented below.

1.3 Predictions

The present study contrasts vowels with different featural specifications in a standard oddball paradigm. The target vowels are presented in two sets of English ablaut (strong) verbs: *sit~sat* and *get~got*. Thus, the same morpheme is used in the competing pairs. In the first pair, *sit~sat*, the FUL model would predict symmetry as both vowels share the same place of articulation ([CORONAL]) and while they differ in height, both [HIGH] and [LOW] are specified in the lexicon (see 2.1 for feature details). As indicated above, the MMN response can be affected by a number of factors and here we lay out our predictions for the most pertinent ones: featural differences, morphological relatedness and lexical status.

In terms of the phonological predictions generated by the differences in specification of the features, there are two sets of predictions: one based on FUL (or any account assuming featural underspecification) and one based on models which assume full specification. If we assume a full-specification account (cf. for example Bybee, 2001; Gaskell and Marslen-Wilson, 1996, 1998; Johnson, 1997; Gaskell, 2003) with equipollent features for both place and height, the data should show MMN responses of similar amplitude and latency for all combinations of standards and deviants in our stimulus set as, if all features are fully specified, standards and deviants will always mismatch to the same degree.

An underspecification account, however, while predicting the same symmetrical pattern for *sit~sat* which differ in vowel height, would make asymmetric predictions for *get~got*. If *get* is heard as the standard and *got* as the deviant, we would expect an attenuated MMN response since, while the incoming /ɔ/ is not a match for the underspecified /ɛ/, it does not result in a mismatch. In the opposite case, with *got* as standard and *get* as deviant, [CORONAL] is extracted from the acoustic signal and will mismatch with the [LABIAL] feature of the standard and we would expect an MMN of similar magnitude as that elicited by *sit~sat*.

Although it is difficult to translate the discrimination evidence into predictions here, employing the Natural Referent Vowel framework (NRV; Polka & Bohn, 2011), we could render our standard/deviant distribution in terms of the referent vowels in their discrimination tasks. Thus, given the F1/F2 acoustic measures of our vowels, an NRV account would predict that [æ] to [ɪ] would be easier to discriminate than the reverse, while both an underspecification account and a full specification account would expect a symmetric pattern for [æ] vs. [ɪ]. The main acoustic difference between [ɛ] and [ɔ] is in F2, with [ɛ] displaying a larger difference between F1 and F2. Consequently, we would expect a change from [ɔ] to [ɛ] to be easier to discriminate than one from [ɛ] to [ɔ], which would fit with an underspecification account. We have summarised the predictions based on these three accounts in Table 2 below, which shows that the three accounts predict different patterns of MMN responses and thus our data will be able to lend support to only one of these theories.

Table 2 *Phonological predictions for full specification & underspecification accounts and NRV framework*

	/ɪ/[æ]*	**/æ/[ɪ]**	**/ɛ/[ɔ]**	**/ɔ/[ɛ]**
full specification account	high MMN	high MMN	high MMN	high MMN
	symmetry		*symmetry*	
underspecification account	high MMN	high MMN	lower/no MMN	high MMN
	symmetry		**asymmetry**	
NRV framework	higher MMN	lower MMN	lower MMN	higher MMN
	asymmetry		**asymmetry**	

***/STD/[DEV]**

In terms of their morphological status, both real word pairs in this study are irregular but differ by the same magnitude as both are ablaut verbs with a change in the vowel. Thus, in terms of the predictions for the present study, any morphologically driven effects should be similar for both pairs as they display similar patterns of (ir)regularity. Both are highly frequent in the language and the only difference between the patterns is the magnitude of the phonological change in terms of the number of features which differ between the simple and complex forms.

The last factor to be addressed here is that of lexical status. As discussed above, word stimuli which activate a stored representation have been shown to elicit earlier and larger MMN responses (Shtyrov et al., 2010, Bakker et al., 2013) and we would expect to see this effect in the form of a difference between the word and pseudoword stimuli. However, we are not primarily interested in word/pseudoword differences as such, but in the question of whether the predicted phonological processing (a)symmetry in vowels will be reflected in both word and pseudoword stimuli or whether it is contingent on or affected by lexical status (and therefore the presence of an underlying lexical representation of the whole word rather than the individual sound). Therefore, our main aim is not to compare word and pseudoword stimuli directly in our analyses. There may, however, be a difference in the N400 component since we may see a priming-induced reduction in N400 in the related real-word pairs which we do not expect to see in the nonword pairs.

To summarise, based on the phonological differences proposed by an underspecification account, we predict an asymmetry in the difference in amplitude for *get~got* and the corresponding nonword pair **gef~*gof* depending on which item is the standard/deviant, while *sit~sat* and its matched nonword pair **sif~*saf* should result in a symmetrical pattern regardless of which member of the pair is the standard/deviant.

2. Methods and analysis

2.1 Stimuli

The stimulus set for this study consisted of four words and four pseudowords which only differed from the real words in the final consonant: real words *sit~sat* and *get~got* and pseudowords **sif~*saf* and **gef~*gof.*

Table 3 *Feature specifications for the four vowels used in the experiment*

		Surface representation		Underlying representation	
	Vowel	**Place**	**Height**	**Place**	**Height**
Experiment 1	**ɪ**	[CORONAL]	[HIGH]	[ ]	[HIGH]
	æ	[CORONAL]	[LOW]	[ ]	[LOW]
Experiment 2	**ɛ**	[CORONAL]	*variable*	[ ]	[ ]
	ɔ	[DORSAL]	*variable*	[DORSAL]	[ ]

The mid vowels [ɛ ɔ] are not specified for height and the phonetic context dictates which features the perceptual system extracts. See Figure 1 for average F1 and F2 values across the four tokens for each of the stimuli.

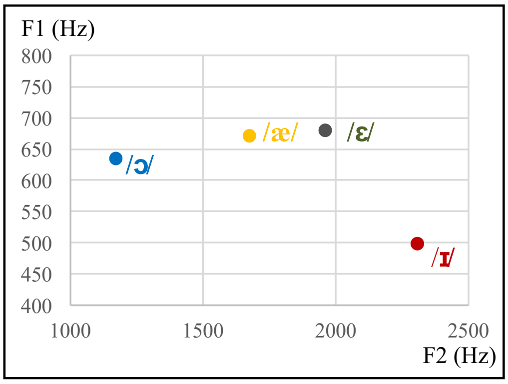

Figure 1 *Average formant frequencies (F1/F2) of all four tokens for /ɪ/, /æ/, /ɛ/ and /ɔ/*

2.1.1 Stimulus recording

The stimuli were recorded by a female native speaker of Southern British English in a sound-attenuated room using a professional quality USB microphone (Røde NT-USB) at a sampling rate of 44.1 kHz. The words were presented in a list of unrelated words in a randomised order. The speaker provided eight tokens of each word.

2.1.2 Stimulus construction

We extracted and manipulated the stimuli using the speech analysis software PRAAT (Boersma & Weenink, 2012). As the aim was to keep natural variation in the vowels, we cut four tokens of each vowel from the original recordings. For the initial and final consonants, we chose one token for each of the four consonants /g, s, t, f/ to ensure they were identical across all stimuli. All individual sounds were taken from recordings of real words rather than pseudowords. The initial /g/ for *get~got* and **gef~gof* was recorded in the context of a schwa (in *gorilla*) since cross-splicing any of the other tokens of /g/ from the test stimuli resulted in anomalous auditory effects due to the difference in place of articulation depending on the following vowel. All individual sounds were normalised before cross-splicing to ensure consistent volume.

The individual sounds were cross-spliced using PRAAT to create four tokens of each of the eight stimuli. We then adjusted the duration of the vowel by cutting individual, non-consecutive glottal pulses to ensure all tokens occurring within one block of recording were of the same duration (see Table 4 for duration values). Finally, to ensure the naturalness of the stimuli, tokens were rated by five native speakers of Southern British English who were asked to listen to the stimuli and comment on any they thought sounded unusual. The final set of tokens consisted only of those which were deemed to sound natural.

Table 4 *Durations (ms) for all stimuli*[1]

	Block	Stimulus	Vowel	Token 1	Token 2	Token 3	Token 4
Experiment 1	1 & 2	*sit*	/ɪ/	550	551	549	550
		sat	/æ/	550	550	550	550
	3 & 4	**sif*	/ɪ/	624	625	623	624
		**saf*	/æ/	624	624	625	622
Experiment 2	1 & 2	*get*	/ɛ/	379	378	378	379
		got	/ɔ/	379	379	378	379
	3 & 4	**gef*	/ɛ/	453	453	453	453
		**gof*	/ɔ/	453	454	453	454

2.2 Experimental design

The study was constructed using a standard oddball paradigm with a standard-to-deviant ratio of 85%/15% in blocks of 700 trials. The ISI was 600ms and stimuli were pseudo-randomised. Each of the eight stimuli was used once as a standard and once as a deviant, resulting in eight blocks with an average duration of fourteen minutes.

2.3 Participants

We recorded data from 24 native speakers of Southern British English who were all students at the University of Oxford (average age: 24.2). They did not report any hearing deficits or neurological conditions and all were right-handed (assessed using the Edinburgh Handedness questionnaire; Oldfield 1971). Participants were compensated for their participation.

2.4 Procedure and recording

Both experiments were conducted with the same participants. Due to the number and duration of blocks, data was acquired in two separate recording sessions. The order of blocks over the two sessions was randomised across both experiments with participants hearing both /g/ and /s/ blocks in each session.

Participants were seated in an electrically shielded, sound-attenuated EEG booth at a comfortable distance from the screen. They first performed an electro-occulogram

1 Stimulus duration was closely matched within blocks but not across blocks as these stimuli would never be compared directly.

(EOG) calibration task before the start of the experimental task. During the main task, participants watched a nature documentary without sound while the stimuli were presented through headphones (Sennheiser PX200). The volume of the auditory stimuli was kept constant across all participants. Participants were instructed to pay no particular attention to the auditory stimuli. No recording session exceeded 90 minutes (including set-up) and breaks were scheduled every 20 minutes during the recording (after every two blocks).

2.5 Data acquisition

A BiosemiActiveTwo amplifier was used to record EEG from 64 sintered Ag/AgCl pin electrodes arranged on the scalp in a 10-10 montage. The recording was online-referenced to the mastoids. Four facial electrodes (IO1, IO2, LO1, LO2) were used to measure EOG activity. All electrode offsets (comparable to impedance) were kept below 20mV and signals were digitised at a sampling rate of 2048Hz.

2.6. ERP analysis

The acoustic difference between the initial consonants in the two pairs (/g/ vs. /s/) may result in latency differences in ERPs elicited by the two sets of stimuli and this was addressed by time-locking the analysis epoch to the beginning of the vowel. There was also no direct comparison between the two stimulus sets but a comparison within paired blocks with identical stimuli (both as standards and deviants).

The continuous EEG data was filtered with a .03Hz High-pass and a 30Hz Low-pass filter. Pre-experimental EOG data was used to capture characteristic scalp topographies of eye artifacts, which were used in an EOG correction algorithm (Ille, Berg & Scherg, 2002) applied across the experimental data. In order to remove other sources of non-EEG noise, two procedures were applied. All data was visually inspected and any epochs containing noisy data or badly corrected eye artifacts were rejected by hand. In addition to this, trials at which at least one electrode exceeded an amplitude of 100 mV or a gradient of 75 mV/division were also rejected in a semi-automatic procedure. Any participant whose data showed too many artifacts (>20% of trials rejected in a single condition) was excluded from further analysis. EEG epochs were averaged time-locked to the onset of the vowel (deviance point) with a pre-stimulus baseline period of 100ms and a window of 600ms from vowel onset.

The time window for the MMN response was determined by averaging the latency of the maximum amplitude (averaged across all participants) at Fz in each condition per experiment and then selecting a window of 60ms centred around this latency. The window for the extraction of mean amplitude for the MMN in Experiment 1 (/s/-stimuli) was 155-215ms while that for Experiment 2 (/g/-stimuli) was slightly earlier from 115-175ms. The ERP data was analysed using a linear mixed effects (LME) model (for a discussion, see for example Newman et al. 2012) with *Subjects* and *Electrode* as random effects (intercepts) and *Condition* as a fixed effect.

3. Results

3.1 Experiment 1

The linear mixed model shows a significant main effect of *Condition* overall (F(11, 827) = 107.04, p < .001) and individual planned comparisons within each block show significant MMN responses for all deviant stimuli compared to the standards (see Table 5 for a summary of results). There is no difference in the MMN responses based on the role of each item (i.e. whether items were presented as standard or deviant). In both word and nonword pairs, [æ] deviants in the context of /ɪ/ standards (*sit~sat* (t(827) = 2.56, p = .010; Est: 0.509; SE: 0.199) and **sif~*saf* (t(827) = 4.67, p < .001; Est: 0.929; SE: 0.199)) resulted in significant MMN responses, as did [ɪ] deviants in the context of /æ/ standards (*sat~sit* (t(827) = 6.20, p < .001; Est: 1.231; SE: 0.199) and **saf~*sif* (t(827) = 7.44, p < .001; Est: 1.478; SE: 0.199)). This pattern indicates symmetrical mismatches between the featural information of these items, in this case the features [LOW] and [HIGH] which are both fully specified.

Table 5 *T-test results for all conditions in Experiments 1 & 2*

	Block	Std	Amp	Dev	Amp	t-test	
Exp 1	1 & 2	*sit* /ɪ/	-3.11μV	*sat* [æ]	-3.62μV	t(827) = 2.56, p=.010	*
		sat /æ/	-3.02μV	sit [ɪ]	-4.26μV	t(827) = 6.20, p<.001	**
	3 & 4	**sif* /ɪ/	-3.59μV	**saf* [æ]	-4.51μV	t(827) = 4.67, p<.001	**
		**saf* /æ/	-3.28μV	**sif* [ɪ]	-4.76μV	t(827) = 7.44, p<.001	**
Exp 2	1 & 2	*get* /ɛ/	0.05μV	*got* [ɔ]	0.35μV	t(827) = 1.67, p=.096	ns
		got /ɔ/	0.94μV	*get* [ɛ]	0.13μV	t(827) = -4.55, p<.001	**
	3 & 4	**gef* /ɛ/	-0.11μV	**gof* [ɔ]	-0.05μV	t(827) = 0.35, p=.727	ns
		**gof* /ɔ/	0.31μV	**gef* [ɛ]	-0.38μV	t(827) = -3.85, p<.001	**

3.2 Experiment 2

In Experiment 2, there is a significant difference between the role of the two stimuli (standard or deviant) in both word and nonword pairs. There is a main effect of *Condition* ($F(11, 827) = 10.40$, $p < .001$) and planned comparisons show a difference between those blocks where the standard contains the vowel /ɔ/ and those which have items containing /ɛ/ as the standard.

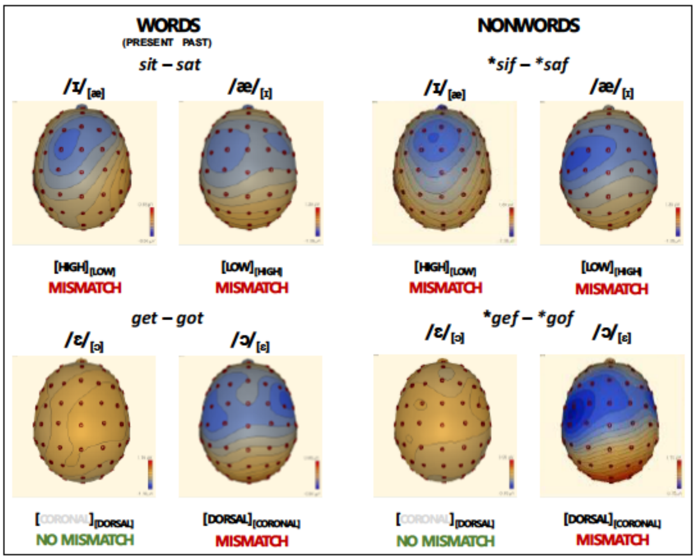

Figure 2 *Topographic maps for all conditions in Experiment 1 and 2 at Fz at maximum MMN amplitude*

In the pairs where *got* or **gof* is presented as the standard, the pattern is the same as that shown in Experiment 1. Both the word and nonword pair, i.e. *got~get* ($t(827) = -4.55$, $p < .001$; Est: -0.817; SE = 0.1796) and **gof~*gef* ($t(827) = -3.85$, $p < .001$; Est -0.691; SE = 0.1796) show significant MMN responses for the deviant compared to the standard. However, the results of the two blocks where the standard is underspecified for place, i.e. those with an /ɛ/ as standard and an [ɔ] as deviant, do not show a significant MMN response for the deviant compared to the standard. This is the case for both the word pair *get~got* ($t(827) = 1.67$, $p = .096$; Est 0.2997; SE = 0.1796) and the nonword pair **gef~*gof* ($t(827) = 0.35$, $p =.727$; Est 0.063; SE = 0.1796). This indicates that the deviant [ɔ] does not mismatch with the stored representation of the standard /ɛ/ since /ɛ/ is underspecified for place and thus the [DORSAL] place feature in /ɛ/ is accepted. The differences in the topographic maps at Fz at maximum MMN amplitude for all experimental blocks are shown in Figure 2 and the individual waveforms for the three electrodes included in the analysis (Fz, FCz, Cz) per stimulus pair can be seen in Figures 3 and 4).

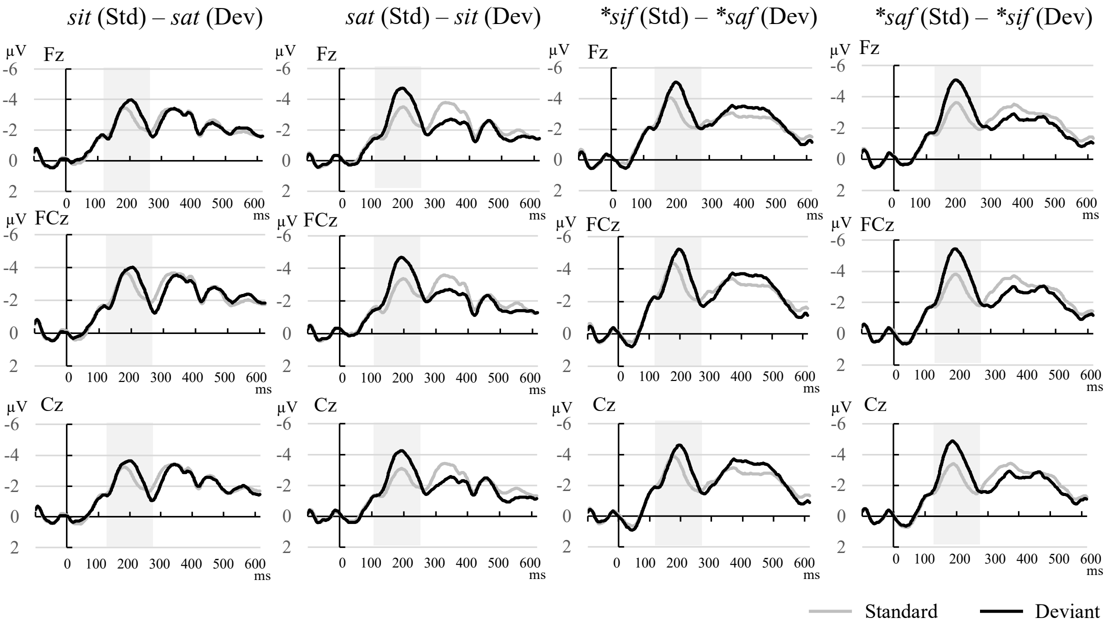

Figure 3 *Experiment 1 grand average waveforms at Fz, FCz and Cz for all blocks; negativity is plotted upwards and shaded area indicates the expected time-window for the MMN component.*

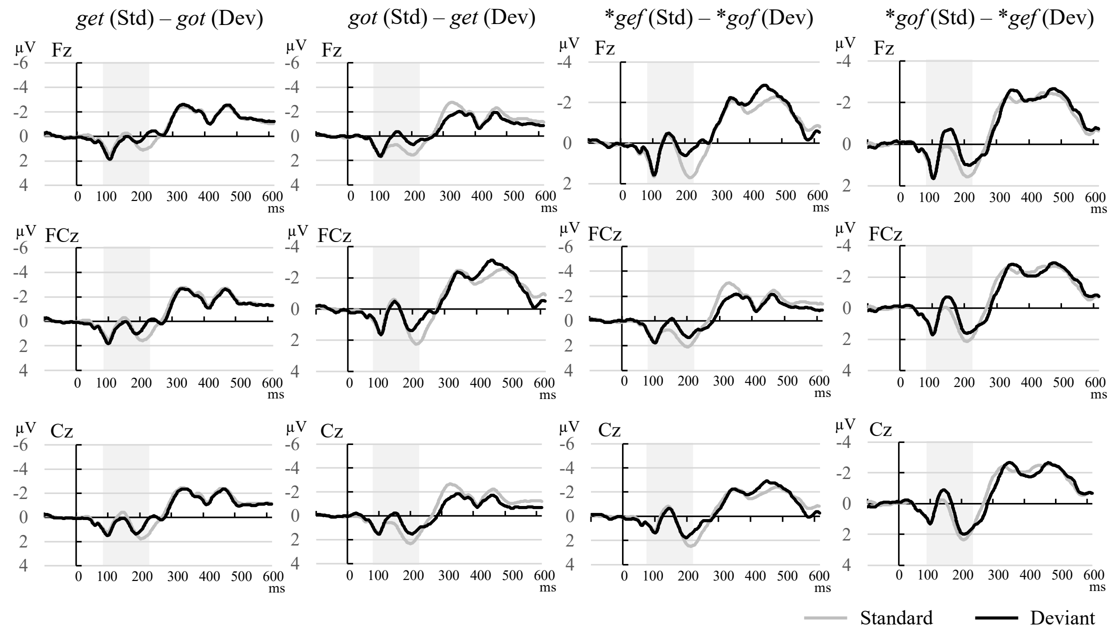

Figure 4 *Experiment 2 grand average waveforms at Fz, FCz and Cz for all blocks; negativity is plotted upwards and shaded area indicates the expected time-window for the MMN component.*

4. Concluding Discussion

In this paper, we present the results of two MMN experiments investigating asymmetric processing comparing vowel height and place of articulation differences in English vowels, within real words and pseudowords. We compared MMN responses to two sets of ablaut (strong) verbs (*sit~sat* and *get~got*) and their corresponding pseudowords (**sif~*saf* and **gef ~*gof*). The vowels /ɪ/ and /æ/ mismatch in height, [HIGH] vs. [LOW], while the vowels [ɛ] and [ɔ] differ in place of articulation, [CORONAL] vs. [DORSAL]. The height features are mutually exclusive and conflict with each other. In contrast, the place features are asymmetric in representation and therefore also in processing. The vowel [ɛ] is unspecified for place in its representation while [ɔ] is specified for [DORSAL]. Thus /ɔ/[ɛ] is a mismatch condition while /ɛ/[ɔ] is an instance of no-mismatch. The former is predicted to elicit a larger MMN while the latter may not result in an MMN at all.

Our aim was to establish (i) whether the evidence found for processing asymmetries resulting from coronal underspecification in vowels in earlier studies (cf. Eulitz & Lahiri 2004), where vowels were presented in isolation, would also hold when the same featural combinations are examined within a syllable; and (ii) whether the lexical status of the item in which the vowel is embedded has an effect on the processing asymmetry previously found for individual vowel pairs where one item is underspecified for place of articulation. The real word pairs *get~got* and *sit~sat* are all verbs, present and past tense of the same root. Thus, there is no difference in meaning and both verbs are very frequent and deeply entrenched in the language, having been attested for centuries. The pseudowords are minimally different, only replacing the final consonant: **gef~*gof* and **sif~*saf.*

Our results show that the asymmetry in place of articulation found in isolated vowels (Eulitz & Lahiri, 2004) is also found when these vowels are embedded in real words. Thus, the MMN for the deviant [ɔ] with respect to the standard /ɛ/ was smaller than that for the deviant [ɛ] when the standard was /ɔ/. The assumption was that the place feature [CORONAL] is unspecified for /ɛ/ and thus the phonetic feature [DORSAL] extracted from the deviant [ɔ] is tolerated (see Table 3). In contrast, the [DORSAL] feature of the standard /ɔ/ is specified and thus the extracted [CORONAL] feature information from the deviant [ɛ] mismatches and results in a larger MMN. This also holds for the pseudoword pair **gef~*gof* which shows an identical pattern.

The real word pair which differed in vowel height (*sit~sat*) showed, as predicted, symmetrical large MMNs indicating full specification of height and thus conflict. This is in line with findings for other fully specified features such as [NASAL] /n/ and [STRIDENT] /z/, which also showed symmetrical mismatch (Cornell et al. 2013). Both the features [LOW] of /æ/ and [HIGH] of /ɪ/ are fully specified. Thus, the corresponding features mismatch with the representation in either combination of standard and deviant. Equally large MMNs are observed for /ɪ/[æ] and /æ/[ɪ]. As with the results for /ɛ/ and /ɔ/ above, this also holds for the pseudoword pair **sif~*saf.*

As discussed earlier (cf. 1.3), differing proposals of phonological representation (e.g. full-specification vs. underspecfication accounts) and phonetic-acoustic processing preferences (e.g. NRV framework) led to clear predictions for the MMN pattern in our data (cf. Table 2). Models which propose full specification of equipollent features (e.g. Bybee, 2001; Gaskell and Marslen-Wilson, 1996, 1998; Johnson, 1997; Gaskell, 2003) would predict symmetrically high MMN responses regardless of the distribution of standards and deviants. However, this is not borne out by our data. The NRV framework (cf. Polka & Bohn, 2011), which is based on the differences in the distribution of acoustic energy in vowels, would predict asymmetries in both cases since one vowel in each pair displays a greater concentration of acoustic energy in the F1/F2 space, which would result in a change from one vowel to another being easier to discriminate than the reverse direction. Again, our data only shows an asymmetric pattern in the vowel pair which is distinguished by place of articulation (i.e. *get~got/*gef~*gof)* but not for the pair with differences in vowel height *(sit~sat/*sif~*saf).*

Thus, only an underspecification account (cf. Table 2) can explain our data, as the underspecification of [CORONAL] leads to an asymmetric pattern depending on which stimulus is presented as standard and which as deviant. This is not the case for the vowel height pair where both features are fully specified. In addition, it appears that the representation of features for vowels in the mental lexicon is independent of word status as the same pattern was observed for word and pseudoword pairs. This seems logical since any single vowel has the potential to be an independent word (e.g. Bengali [o] 'he/she', French *et* [ɛ] 'and', Italian *o* [ɔ] 'or'). This data thus supports accounts such as FUL (Lahiri & Reetz, 2002, 2010) which propose featural underspecifications and shows feature representations of sounds to be independent of the lexical status of the items they are embedded in.

Acknowledgements

This research was funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement number: 695481 awarded to A. Lahiri). We would also like to thank Holly Kennard, Swetlana Schuster and Kai Alter for their help in assessing stimuli, recording participants and preparing the data for analysis. In addition, we owe thanks to Henning Reetz for his advice on stimulus generation and Emily Darley, Steven Kaye and Adam Roberts for providing comments on earlier drafts of the manuscript. Finally, we thank all our participants, without whom this study would not have been possible.

References

Bakker, I., MacGregor, L. J., Pulvermüller, F., & Shtyrov, Y. (2013). Past tense in the brain's time: Neurophysiological evidence for dual-route processing of past-tense verbs. *NeuroImage*, 71, 187–195.

Boersma, P., & Weenink, D. (2012). *PRAAT: Doing Phonetics by Computer* (ver.5.2.26). Amsterdam: Institute for Phonetic Sciences.

Bölte, J., & Coenen, E. (2000). Domato primes paprika: Mismatching words activate semantic and phonological representations. *Proceedings of the workshop on Spoken Word Access Processes (SWAP)*. MPI for Psycholinguistics, Nijmegen, The Netherlands.

Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.

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.

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.

de Jonge, M. J. I., & Boersma, P. (2015). French high-mid vowels are underspecified for height. In *Proceedings of the 18th International Congress of Phonetic Sciences,* Glasgow: The University of Glasgow.

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, 2*, 1–36

Friedrich, C. K., Lahiri, A., & Eulitz, C. (2008). Neurophysiological Evidence for Underspecified Lexical Representations: Asymmetries with Word Initial Variations. *Journal of Experimental Psychology: Human Perception and Performance*, 34(6), 1545–1559.

Gaskell, M. G. (2003). Modelling regressive and progressive effects of assimilation in speech perception. *Journal of Phonetics*, 31(3/4), 447–463.

Gaskell, M. G. & Marslen-Wilson, W. D. (1996). Phonological variation and interference 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.

Goldinger, S. D. (1998). Echoes of Echoes? An Episodic Theory of Lexical Access. *Psychological review*, 105(2), 251–279.

Hasting, A., Kotz, S.A., & Friederici, A., 2007. Setting the stage for automatic syntax processing: the mismatch negativity as an indicator of syntactic priming. *Journal of Cognitive Neuroscience*, 19(3), 386–400.

Højlund, A., Gebauer, L., Mc Gregor, W. B., & Wallentin, M. (2019). Context and perceptual asymmetry effects on the mismatch negativity (MMNm) to speech sounds: an MEG study. *Language and Cognitive Neuroscience*, 34(5), 545-560.

Ille, N., Berg, P., & Scherg, M. (2002). Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. *Journal of Clinical Neurophysiology*, 19, 113–24.

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson & J.W. Mullennix (eds), *Talker variability in speech processing.* San Diego, CA: Academic, 145–165.

Kotzor, S., Wetterlin, A., Roberts, A. C. & Lahiri, A. (2016). Processing of consonant length: behavioural evidence from Bengali. *Language and Speech*, 1–30.

Kotzor, S., Wetterlin, A., & Lahiri, A. (2017). Symmetry or asymmetry: Evidence for underspecification in the mental lexicon. In: A. Lahiri & S. Kotzor, *The Speech Processing Lexicon: Neurocognitive and behavioural approaches*. Berlin: Mouton de Gruyter, 85–106.

Lahiri, A., Marslen-Wilson, W. D. (1992). Lexical processing and phonological representation. In R. D. Ladd & G. Docherty (eds), *Papers in Laboratory Phonology II: gesture, segment, prosody*. Cambridge: Cambridge University Press, 229–254.

Lahiri, A., & Reetz, H. (2002). Underspecified recognition. In C. Gussenhoven & N. Warner (eds), *Laboratory Phonology VII*, Berlin: Mouton de Gruyter, 637–677.

Lahiri, A., & Reetz, H. (2010). Distinctive Features: Phonological underspecification in representation and processing. *Journal of Phonetics*, 38, 44–59.

Lahiri, A. (2012). Asymmetric phonological representations of *words* in the mental lexicon. In A. C. Cohn, C. Fougeron, & M. K. Huffman (eds), *The Oxford Handbook of Laboratory Phonology*, Oxford: Oxford University Press, 146–161.

Marslen-Wilson, W. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature*, 244, 522–523.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology,* 10, 29–63.

Marlsen-Wilson, W., Nix, A., & Gaskell, G. (1995). Phonological variation in lexical access: abstractness, interference and English place assimilation. *Language and Cognitive Processes*, 10 (3/4), 285–308.

Masapollo, M., Polka, L., Molnar, M., & Ménard, L. (2017). Directional asymmetries reveal a universal bias in adult vowel perception. *Journal of the Acoustical Society of America*, 141(4).

Mohr, B., & Pulvermüller, F. (2002). Redundancy gains and costs in cognitive processing: effects of short stimulus onset asynchronies. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28(6), 1200–1223.

Näätänen, R. (2001). The perception of speech sounds in the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology,* 38, 1–21.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*, 24(4), 375–425.

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, 2544–2590.

Newman, A. J., Tremblay, A., Nichols, E. S., Neville, H. J., & Ullman, M. T. (2012). The influence of language proficiency on lexical semantic processing in native and late learners of English. *Journal of Cognitive Neuroscience*, 24(5), 1205–1223.

Nix, A., Gaskell, G., & Marslen-Wilson, W. (1993). Phonological variation and mismatch in lexical access. *Proceedings of Eurospeech 93*, Berlin, Germany, 1–4.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia,* 9, 97–113.

Pierrehumbert, J. B. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner (eds), *Laboratory Phonology* 7, Berlin: de Gruyter, 101–139.

Polka, L., & Bohn, O.-S. (2003). Asymmetries in vowel perception. *Speech Communication*, 41, 221–231.

Polka, L., & Bohn, O.-S. (2011). Natural Referent Vowel (NRV) framework: an emerging view of early phonetic development. *Journal of Phonetics*, 39, 467–478.

Pulvermüller, F., Kujala, T., Shtyrov, Y., Simola, J., Tiitinen, H., Alku, P., Alho, K. et al. (2001). Memory traces for words as revealed by the mismatch negativity. *NeuroImage,* 14*,* 607–616.

Pulvermüller, F., & Shtyrov, Y. (2003). Automatic processing of grammar in the human brain as revealed by the mismatch negativity. *NeuroImage*, 20(1), 159–172.

Pulvermüller, F., & Shtyrov, Y. (2006). Language outside the focus of attention: The mismatch negativity as a tool for studying higher cognitive processes. *Progress in Neurobiology,* 79, 49–71.

Pulvermüller, F., Shtyrov, Y., Kujala, T., & Näätänen, R. (2004). Word-specific cortical activity as revealed by the mismatch negativity (MMN). *Psychophysiology,* 41, 106–112.

Pulvermüller, F., Garagnani, M., & Wennekers, T. (2014). Thinking in circuits: toward neurobiological explanation in cognitive neuroscience. *Biological Cybernetics*, 108(5), 573–593.

Pulvermüller, F., Cappelle, B., & Shtyrov, Y. (2012). Brain basis of meaning, words, constructions and grammar. In T. Hoffmann & G. Trousdale (eds), *The Oxford Handbook of Construction Grammar*, Oxford: Oxford University Press, 396–416.

Rastle, K., Davis, M. H., Marslen-Wilson, W. D., & Tyler, L. K. (2000). Morphological and semantic effects in visual word recognition: a time-course study. *Language and Cognitive Processes*, 15(4/5), 507–537.

Roberts, A. C., Wetterlin, A., & Lahiri, A. (2012). Aligning mispronounced words to meaning: Evidence from ERP and reaction time studies. *The Mental Lexicon,* 8(2), 140–163.

Roberts, A. C., Kotzor, S., Wetterlin, A., & Lahiri, A. (2014). Asymmetric processing of durational differences – electrophysiological investigations in Bengali. *Neuropsychologia,* 58, 88–98.

Shtyrov, Y. & Pulvermüller, F. (2002). Neurophysiological evidence of memory traces for words in the human brain. *Neuroreport*, 13(4), 521–525.

Shtyrov, Y., Kujala, T., & Pulverm ller, F. (2010). Interactions between language and attention systems: early automatic lexical processing? *Journal of Cognitive Neuroscience*, 22 (7), 1465–1478.

Wheeldon, L. R., & Waksler, R. (2004). Phonological underspecification and mapping mechanisms in the speech recognition lexicon. *Brain and Language*, 90, 401–412.