FRENCH HIGH-MID VOWELS ARE UNDERSPECIFIED FOR HEIGHT

Mirjam J.I. de Jonge & Paul Boersma

University of Amsterdam
m.j.i.dejonge@uva.nl, paul.boersma@uva.nl

ABSTRACT

Asymmetries in Mismatch Negativities (MMNs) have been reported as evidence for phonological underspecification in speech perception. In this study, we investigated whether predictions from Lahiri's Featurally Underspecified Lexicon (FUL) model hold true in a language (French) and a contrast (vowel height) not investigated before, in contrast to predictions from a model with equipollent features or from Element Theory.

The MMNs from French listeners to contrasts among the four vowels [y, u, ø, o] show clear asymmetries that are in line with FUL's predictions for vowel height and place. The change from a back vowel to a front vowel elicits stronger responses than vice versa, which generalises existing findings. The change from a high vowel to a high-mid vowel also elicits stronger responses than the reverse change, which is a new finding that supports the idea that the height contrast in these vowels is expressed by the privative feature [HIGH].

Keywords: mismatch negativity, underspecification, vowel height, vowel place

1. INTRODUCTION

There is a growing body of research demonstrating that the properties of a language's phonological system influence the processing of speech by speakers of that language [5]. One area where this influence is manifest is preattentive auditory processing as reflected by the Mismatch Negativity (MMN) [9]. The MMN is elicited by a change in some auditory regularity, typically the occurrence of a deviant stimulus in a series of repeated 'standard' stimuli. The magnitude of the MMN is thought to reflect the level of perceived contrast between deviant and standard stimuli. From this, it follows that the MMN can be used to assess the perceptual relevance of acoustic differences between stimuli.

Of particular interest from a phonological point of view is the comparison of MMNs for different directions of change within a contrast. Given a contrast between sounds A and B, the MMN that is evoked by the change from one to the other can be similar in both directions or stronger in either direction. For each of these potential experimental outcomes, Lahiri and Reetz [8] proposed that a phonological interpretation can be given regarding the property that distinguishes A and B :

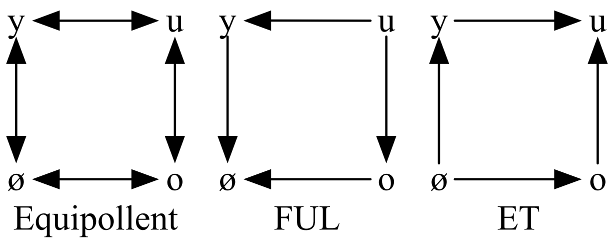
1. symmetrical MMNs indicate equipollent feature specifications of A and B
2. larger MMNs for a change from A to B than vice versa indicate a privative feature specification of A
3. larger MMNs for a change from B to A than vice versa indicate a privative feature specification of B

The assignment of phonological properties to representations is based on the assumption that information that is *detected* in the auditory signal of a deviant stimulus is compared to the information that is *represented* about the standard stimulus. When a standard stimulus is specified for a property, a deviant stimulus that does not match the standard exactly results in a *mismatch* situation and is expected to evoke a strong MMN. When a standard stimulus is not specified or underspecified for a property, a deviant that does not match the standard exactly for this property results in a '*nomismatch'* [8] situation, since there is no information in the representations that the detected information can be compared against.

In this study, we report on the MMN evoked by contrasts between the French vowels [y, u, ø, o] to determine the kind of phonological representations that listeners have for these vowels. The properties that vary among these vowels are their Place, i.e. front or back, reflected acoustically in their F2 values, and their Height, i.e. high or high-mid, reflected acoustically in their F1 values.

For the vowels in this experiment, theoretical accounts of their representations can be formulated aligned with the three options listed above. Using equipollent features one needs [±BACK], [±HIGH] and [±ROUND] to specify all vowel properties exhaustively, and each vowel is assigned the same number of specifications. This predicts symmetric responses as shown in Fig. 1, panel 1. Using the Featurally Underspecified Lexicon model (FUL) [8] one would need the privative features [CORONAL], [DORSAL], [LABIAL] and [HIGH]. In this model the representations of some vowels are sparser than others, partly due to the assumption that [CORONAL] is universally underspecified. This model predicts asymmetric MMN patterns with the stronger responses going leftward and downward as shown in Fig. 1, panel 2. Opposite predictions, as shown in panel 3, arise from phonological representations in terms of the Elements |A|, |I| and |U| [7]. As in FUL, representations in Element Theory (ET) contain privative properties and some vowels contain more specifications than others, but their arrangement is different.

Figure 1: Predicted MMN patterns given three different types of representations.



2. METHOD

2.1. Participants

The results of 24 native speakers of French are presented here. Participants were between 19 and 33 years old; 6 were male. All participants were right-handed and reported no hearing or neurological problems. Datasets of 7 additional participants were excluded from analysis (see section 2.4).

2.2. Stimuli

Single synthetic tokens of the French vowels [y, u, ø, o] were used as stimuli. Vowels were created using the Klatt synthesiser in Praat [1]. The lower 4 formants were based on the average formant values reported for French males in [2] (see Table 1). Stimuli were 150 ms in duration and had a slight rising-falling F0 contour.

Table 1: Formants 1-4 of the synthesised French vowels [y], [u], [ø], [o].

Vowel	F1 (Hz)	F2 (Hz)	F3 (Hz)	F4 (Hz)
y	308	1750	2166	3184
u	308	764	2166	3184
ø	376	1417	2166	3184
o	376	793	2166	3184

2.3. Experimental paradigm

Participants completed a passive listening task while they watched a silent film with French subtitles. Auditory stimuli were presented in a multi-deviant oddball paradigm with 4 blocks. The order of the blocks was counterbalanced across participants.

In each block, one of the vowels occurred 85% of the time and thus served as the standard. The other 3 vowels served as deviants, each occurring 5% of the time. Every deviant was presented 180 times within a block and separated by at least 4 standards from the next deviant. In between stimuli a silence of 357.1 ms occurred, bringing the stimulus onset asynchrony to 507.1 ms.

The design results in 16 conditions for which event-related potentials (ERPs) can be computed: one standard ERP and 3 deviant ERPs in every block. Because the MMN to a particular contrast is found by subtracting the standard ERP from a deviant ERP, the experiment results in 12 MMNs. Each of the 4 vowels occurs in 3 types of vowel contrast: contrasts in vowel height only, contrasts in vowel place only, and contrasts in vowel height and place simultaneously.

2.4. EEG details

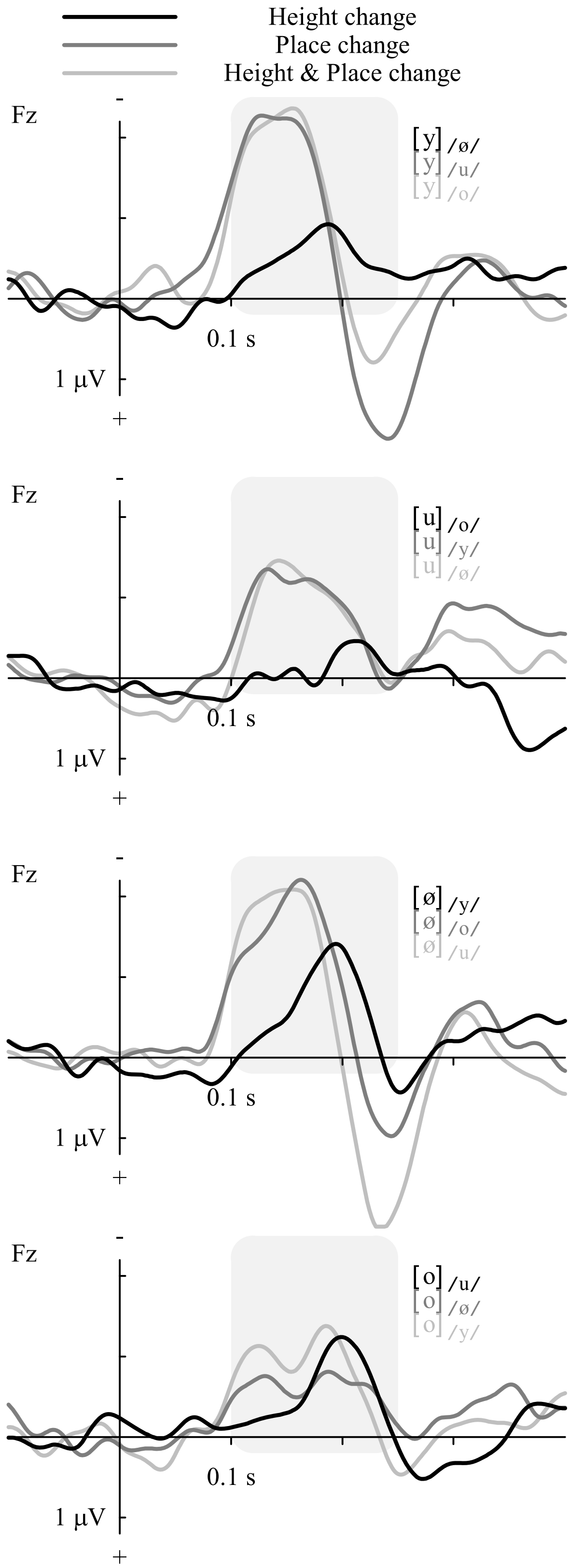
The EEG was recorded using a BioSemi ActiveTwo system with 64 active scalp electrodes and 7 extra electrodes placed on the mastoids, around the eyes and on the tip of the nose. Data were acquired at 8192 Hz and downsampled off-line to 512 Hz, re-referenced to the average mastoids and band-pass filtered from 1 to 30 Hz. Continuous data was segmented into 500-ms epochs with a 100-ms baseline, and all epochs with activity exceeding ±75 μV in any channel were discarded. Participants that had fewer than 120 analysable epochs in any condition were excluded from analysis.

3. RESULTS

The MMN for all 12 contrast conditions was computed by subtracting the standard ERP from the 3 deviant ERPs of the same vowel, to make sure the resulting difference wave reflected the phonological contrast rather than the acoustic difference between stimuli within a block. Fig. 2 shows all Grand Average difference waves at electrode Fz grouped by quality of the stimulus vowel. To analyse the directionality of the responses with respect to the predictions listed in section 1, only conditions in which a single dimension was changed are included in the further discussion of these results, i.e. only the Place and Height contrasts.

The MMN magnitude was quantified separately in each condition by finding the grand average negative peak within the 100-250 ms window following the onset of change (0 ms) and computing the mean amplitude in a 40-ms window around this peak [3]. The magnitudes of the MMN in the relevant contrast conditions depicted in Fig. 3.

Figure 2: Grand Average difference waves (deviant minus standard) at Fz relative to average mastoids for all 12 contrasts. Negativity is plotted upwards. Shaded areas indicate the window in which the MMN is expected to occur.



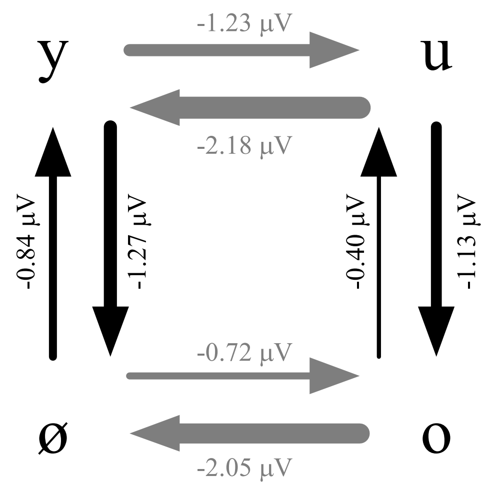
In order to assess which of the theoretical accounts best predicted the pattern of evoked responses, directionality was analysed for the Height and Place contrast separately.

We found that a change in the Height of a vowel, regardless of direction, resulted in an average MMN of -0.91 µV, compared to -1.55 µV for a change in vowel Place. This difference of 0.64 µV is highly significant ($t(23) = 4.584$, $p = .0001$) in a paired samples *t*-test, indicating that our listeners responded more strongly overall to changes between front and back vowels than to changes between high and mid-high vowels.

On average, a change from a high vowel towards a high-mid vowel evoked an MMN that was 0.58 µV larger (i.e. more negative) than vice versa, meaning a multiplication factor of 1.9. This difference in MMN magnitude between the directions of change approaches significance in a two-tailed paired-samples *t*-test ($t(23) = 2.028$, $p = .054$).

A change from a back vowel to a front vowel evoked an MMN that was on average 1.14 µV larger than vice versa, meaning a multiplication factor of 2.2, and this directional difference in MMN magnitude is highly significant in a two-tailed paired-samples *t*-test ($t(23) = 3.319$, $p = .003$).

Figure 3: MMN magnitudes for contrasts in vowel Height and vowel Place in each direction.



4. DISCUSSION

The current study set out to determine the properties of phonological representations of French vowels by measuring listeners' neural responses to vowel contrasts. The potential outcomes could support three different theoretical accounts of phonological representations, employing differently valued properties. The crucial parameter that differentiated the accounts was symmetry versus asymmetry, and in case of asymmetry its direction. The pattern of obtained results in Fig. 3 most closely resembles the predictions made by the FUL model shown in Fig. 1, panel 2, both for the vowel Place and the vowel Height dimension.

The finding that Place differences evoked larger responses overall compared to Height differences is not predicted by a particular theory of phonological representations, but could arise from the simple fact that the Place contrasts in the current stimulus set span larger distances in the vowel space than the Height contrasts.

The finding that changes in vowel place and in vowel height evoke responses of which the magnitude depends on the direction of change is more relevant from a phonological perspective, although the acoustic-phonetic dimension should not be ignored. For both contrast dimensions, the strongest MMNs were evoked when the change constituted a formant increase rather than a decrease. The compatibility of the current results with phonological accounts could then be a consequence of the alignment of phonological dimensions with acoustic-phonetic dimensions.

Asymmetric MMNs for back versus front vowels have been reported before [4,6] using natural German stimuli with multiple tokens per vowel type in an attempt to "force the processing system to map the incoming signals onto more abstract representations" [6] (p. 581). Our results suggest that it is not necessary to use natural and varied speech stimuli to get asymmetric MMN responses, indicating that the mechanisms responsible for this influence on auditory processing operate without the need to be 'tempted' into an abstract listening mode. Neither do these asymmetries rely on peculiarities of German, since French lacks the morphologically conditioned Umlaut process that turns some German back vowels into front vowels [10], potentially prompting German listeners to be more attentive to this type of change.

The asymmetry in the MMN pattern for contrasts in vowel height is statistically less robust than for place contrasts but proportionally very comparable. This might be a matter of statistical power, as MMNs were generally smaller for changes in height, leading to a worse signal-to-noise ratio than for place contrasts. Either way, the clear trend towards an asymmetry for vowel height is evidence in favour of using a privative feature [HIGH] in the representation of French high vowels while leaving high-mid underspecified. To our knowledge, this is the first piece of MMN evidence for the underspecification of vowel height.

Given that the French vowel system is generally described as having 4 height levels and that FUL only has two features ([HIGH] and [LOW]) to deal with vowel height, it is an open and prudent question how the non-terminal height levels are represented and distinguished from one another.

5. SUMMARY AND CONCLUSION

In this study, we used a multi-deviant oddball paradigm to elicit MMN patterns in French listeners in response to contrasts among the vowels [y, u, ø, o]. We found asymmetric response patterns for both Height and Place differences that are in line with predictions from FUL, while contradicting predictions based on phonological representations containing Elements or equipollent features.

Considering that our results provide the first evidence for a previously untested corollary of the FUL model regarding vowel height and extend previously reported findings regarding vowel place with a new type of stimuli in a new population, we believe the FUL model is a fruitful model of phonological representations that can guide research endeavours to bridge the gap between theoretical models of linguistic representations on the one hand and on the other hand the neural mechanisms that are involved in parsing and interpreting speech signals.

6. REFERENCES

[1] Boersma, P., Weenink, D. 2014. *Praat: doing phonetics by computer* [Computer program]. Version 5.3.71, retrieved 9 April 2014 from http://www.praat.org/

[2] Calliope. 1989. *La parole et son traitement automatique*. Paris: Masson.

[3] Chládková, K., Escudero, P., Lipski, S.C. 2013. Pre-attentive sensitivity to vowel duration reveals native phonology and predicts learning of second-language sounds. *Brain and Language* 126, 243–252.

[4] 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.

[5] Cutler, A. 2012. *Native Listening: Language Experience and the Recognition of Spoken Words*. Cambridge: MIT Press.

[6] Eulitz, C., Lahiri, A. 2004. Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition. *Journal of Cognitive Neuroscience* 16, 577-583.

[7] Harris, J., Lindsey, G. 1995. The elements of phonological representation. In: Durand, J., Katamba, F. (eds), *Frontiers of phonology: atoms, structures, derivations*. Harlow, Essex: Longman, 34-79.

[8] Lahiri, A., Reetz, H. 2010. Distinctive Features: Phonological underspecification in representation and processing. *Journal of Phonetics* 38, 44-59.

[9] Näätänen, R. 2001. The perception of speech sounds by the human brain as reflected by the mismatch

negativity (MMN) and its magnetic equivalent MMNm. *Psychophysiology* 38, 1–21.

[10] Scharinger, M., Lahiri, A., Eulitz, C. 2010. Mismatch negativity effects of alternating vowels in morphologically complex word forms. *Journal of Neurolinguistics* 23, 383–399.