Brain and Language 176 (2018) 26–35

Contents lists available at ScienceDirect

Brain and Language

journal homepage: www.elsevier.com/locate/b&l

Short communication

Using rotated speech to approximate the acoustic mismatch negativity response to speech

Ellen Marklund*, Francisco Lacerda, Iris-Corinna Schwarz

Stockholm University, Department of Linguistics, SE-106 91 Stockholm, Sweden

A B S T R A C T

The mismatch negativity (MMN) response is influenced by the magnitude of the acoustic difference between standard and deviant, and the response is typically larger to linguistically relevant changes than to linguistically irrelevant changes. Linguistically relevant changes between standard and deviant typically co-occur with differences between the two acoustic signals. It is therefore not straightforward to determine the contribution of each of those two factors to the MMN response. This study investigated whether spectrally rotated speech can be used to determine the impact of the acoustic difference on the MMN response to a combined linguistic and acoustic change between standard and deviant. Changes between rotated vowels elicited an MMN of comparable amplitude to the one elicited by a within-category vowel change, whereas the between-category vowel change resulted in an MMN amplitude of greater magnitude. A change between rotated vowels resulted in an MMN ampltude more similar to that of a within-vowel change than a complex tone change did. This suggests that the MMN amplitude reflecting the acoustic difference between two speech sounds can be well approximated by the MMN amplitude elicited in response to their rotated counterparts, in turn making it possible to estimate the part of the response specific to the linguistic difference.

1. Introduction

Mismatch negativity (MMN) is an event-related component that reflects pre-attentive change detection (e.g., Näätänen, Paavilainen, Rinne, & Ahlo, 2007). Its amplitude reflects the magnitude of acoustic difference between the standard and deviant sounds (e.g., Tiitinen, May, Reinikainen, & Näätänen, 1994). The MMN amplitude also reflects the presence of linguistic relevance in the change between standard and deviant (e.g., Sharma & Dorman, 1999). Since a linguistically relevant change between standard and deviant in most cases co-occur with a difference between the two acoustic signals, it is not straightforward to determine the relative contribution of each of those two factors to the MMN response.

The terms and definitions for making the distinction between linguistically relevant and linguistically irrelevant information in the speech signal vary. In this paper, the terms *linguistic* and *acoustic* will be used. *Linguistic* refers to the abstract content conveyed by the speech signal, whereas *acoustic* refers to the properties of the speech signal itself. The term acoustic thus includes both acoustic information that may function as a cue to linguistic content (e.g., spectral differences that results in two sounds being perceived as two different vowels) and acoustic information that does not (e.g. spectral differences that only signal allophonic variation). Importantly, in the case of a linguistic difference between two sounds, there is in most cases also an acoustic difference between them. In some rare cases, a linguistic difference can be perceived without altering the acoustic signal (by modifying the acoustic context or the listener's expectations instead), but as a rule, a linguistic difference always co-occurs with (and is conveyed by) an acoustic difference.

1.1. MMN responses to linguistic and acoustic differences between standard and deviant

A growing body of evidence shows that the amplitude of the MMN is influenced by the type of difference that is present between standard and deviant. Specifically, a combination of a linguistic and an acoustic difference results in a greater amplitude than a difference that is acoustic only, provided that the magnitude of the acoustic differences are matched (see Section 1.2).

When contrasting the MMN responses to speech and non-speech stimuli, it is likely that the speech condition includes linguistic differences, whereas the non-speech condition by definition includes only acoustic differences (e.g., Christmann, Berti, Steinbrink, & Lachmann, 2014; Sorokin, Alku, & Kujala, 2010). In a study by Christmann et al.

* Corresponding author.
E-mail addresses: ellen.marklund@ling.su.se (E. Marklund), francisco.lacerda@ling.su.se (F. Lacerda), iris-corinna.schwarz@ling.su.se (I.-C. Schwarz).

https://doi.org/10.1016/j.bandl.2017.10.006
Received 2 April 2015; Received in revised form 31 August 2017; Accepted 21 October 2017
Available online 22 November 2017

(2014), the speech condition consisted of changes between vowels, making the difference between standard and deviant both acoustic and linguistic in nature. In their non-speech conditions on the other hand, differences between standard and deviants were of course acoustic-only. The authors found larger MMNs in their speech condition than in their two non-speech conditions (Christmann et al., 2014). In the same vein, Sorokin et al. (2010) found that their speech condition overall had a larger MMN than their non-speech condition. Further, differentiating between changes of linguistic relevance and other changes *within* the speech condition, they found that for the most part, changes between standard and deviant that were comprised of both an acoustic and a linguistic change had enhanced MMN responses compared to those that were comprised of an acoustic-only change (Sorokin et al., 2010). On the whole, making the distinction between linguistic and acoustic changes is more informative than merely differentiating between speech stimuli and non-speech sound stimuli, since both types of stimuli invariably entail acoustic changes.

Zeroing in on speech stimuli, between-category changes between standard and deviant comprise a combination of a linguistic and acoustic change, whereas within-category changes by design are comprised of an acoustic-only difference. As expected in studies making this distinction between conditions, the amplitude of the MMN is greater in response to between-category changes than to within-category changes (e.g., Dehaene-Lambertz, 1997; Sharma & Dorman, 1999).

When comparing the MMN responses to native versus non-native contrasts, studies are often designed in such a way that the two sounds in the non-native contrast will be perceived as different exemplars of the same speech sound in the native language (e.g., Sharma & Dorman, 2000; Winkler et al., 1999). In essence, the distinction between native versus non-native contrasts is the same distinction as that between between-category changes and within-category changes, but set in the context of the language background of the listeners. This means that the difference between native speech sounds is both linguistic and acoustic, whereas the difference between non-native speech sounds is acoustic-only. The study by Winkler et al. (1999) illustrates this point beautifully: a pair of vowels that to Hungarian speakers belongs to different categories is considered a variation of a single vowel by speakers of Finnish, and vice versa. The difference between the same two vowels is thus either both linguistic and acoustic or acoustic-only, depending on the native language of the listener. In the study, the combination of linguistic and acoustic difference elicited larger MMN responses than the acoustic-only difference (Winkler et al., 1999). Similarly, Sharma and Dorman's (2000) difference between the sounds in a Hindi contrast was linguistic and acoustic to Hindi speaking participants. To the English speaking participants, the difference was acoustic-only, since the two stimuli were different exemplars of the same speech sound. Consequently, the Hindi speakers showed larger MMN responses than the English speakers (Sharma & Dorman, 2000).

In sum, three patterns have been observed regarding the MMN response to acoustic and linguistic changes:

(1) Speech stimuli result in larger MMNs than non-speech stimuli (Christmann et al., 2014; Sorokin et al., 2010).
(2) Native contrasts result in larger MMNs than non-native contrasts (Sharma & Dorman, 2000; Winkler et al., 1999).
(3) Between-category changes result in larger MMN responses than within-category changes (Dehaene-Lambertz, 1997; Sharma & Dorman, 1999).

These patterns can be collapsed into one single distinction: combined linguistic and acoustic differences between standard and deviant result in larger MMN responses than acoustic-only differences (of corresponding acoustic magnitude, see Section 1.2 below) between standard and deviant.

1.2. Effect of difference magnitude between standard and deviant on MMN amplitude

The magnitude of the acoustic difference between standard and deviant influences the amplitude of the MMN response, with a greater acoustic difference resulting in a higher amplitude. This pattern is found for:

(1) Non-speech stimuli, when the difference between standard and deviant is consequently acoustic-only (e.g., Nikjeh, Lister, & Frisch, 2009; Novitski, Tervaniemi, Huotilainen, & Näätänen, 2004; Tiitinen et al., 1994).
(2) Speech stimuli, when the difference between standard and deviant is acoustic-only (Aaltonen, Eerola, Hellström, Uusipaikka, & Lang, 1997; Marklund, Schwarz, & Lacerda, 2014; Pakarinen et al., 2013).
(3) Speech stimuli, when the difference between standard and deviant is both acoustic and linguistic (Deguchi et al., 2010; Näätänen et al., 1997; Pakarinen et al., 2013; Sittiprapaporn, Tervaniemi, Chindaduangratn, & Kotchabhakdi, 2005).

The pattern of greater MMN amplitude in response to greater difference between standard and deviant is found as long as the difference between standard and deviant is of the same kind (acoustic-and-linguistic or acoustic-only), and stimuli are of the same type (non-speech or speech). The pattern is *not* found when comparing conditions with different types of differences; the combination of a linguistic and an acoustic difference between standard and deviant results in an overall higher MMN amplitude than an acoustic-only difference (see Section 1.1).

Considering the overall difference between standard and deviant instead of only the acoustic, this is however not surprising, assuming the two acoustic differences are equal. In the first condition, a linguistic difference of some magnitude greater than zero is added, whereas in the second condition, there is a linguistic difference of zero (no linguistic difference). There is then a greater overall difference in the first condition than in the second. Following the pattern established above, the first condition has a greater difference magnitude than the second and will thus yield a larger MMN response.

1.3. Variations in MMN amplitude in response to linguistic differences

In general, not much is known about the ways in which a linguistic difference influences the amplitude of the MMN response, except for the fact that its mere presence enhances the response (see Section 1.2 above). However, asymmetries in speech sound discrimination have been demonstrated in the MMN amplitude for both vowels (e.g., Eulitz & Lahiri, 2004) and consonants (Scharinger, Merickel, Riley, & Idsardi, 2011). This type of asymmetries entails greater MMN amplitude when the change between two sounds occurs in one direction than when it occurs in the other direction. For example, the MMN is more negative when /y/ is presented as a deviant among /u/ standards than when /u/ is presented as deviant among /y/ standards (de Jonge & Boersma, 2015). The relative size of these asymmetries has been proposed to reflect different degrees of specifications of phonemic characteristics of speech sounds (de Jonge & Boersma, 2015; Schluter, Politzer-Ahles, Al Kaabi, & Almeida, 2017). However, studies of this kind assume that the influence of the acoustic difference between standard and deviant can be eliminated by measuring the MMN in a difference wave created by subtracting the ERP responses to the same physical stimulus presented both as deviant and standard (in different blocks). This is a problematic assumption, since the exact morphology of the ERP response to the deviant sound depends not only on its own characteristics, but also on the degree to which the same neuronal populations are activated by the deviant and the standard sound (May & Tiitinen, 2010). That is, the deviant-among-standards ERP waveform is inevitably influenced by the characteristics of the standard sound, and directional asymmetries in