

# Integrated Preattentive Processing of Vowel and Pitch

## A Mismatch Negativity Study

Pascale Lidji,<sup>a,b,c,d</sup> Pierre Jolicœur,<sup>a</sup> Patricia Moreau,<sup>a</sup>  
Régine Kolinsky,<sup>b,c</sup> and Isabelle Peretz<sup>a</sup>

<sup>a</sup>*BRAMS Laboratory, Université de Montréal, Montréal, Québec, Canada*

<sup>b</sup>*UNESCOG, Université Libre de Bruxelles, Brussels, Belgium*

<sup>c</sup>*Fonds de la Recherche Scientifique-FNRS, Brussels, Belgium*

<sup>d</sup>*McGill University, Montréal, Québec, Canada*

This study examines the additivity of the Mismatch Negativity (MMN) as an index of the early interactions between vowels and pitch when sung. Event-related potentials (ERPs) were recorded while participants were presented with sung vowels. Sixteen percent of stimuli deviated in pitch only, in vowel only, or in both pitch and vowel. All three kinds of deviants elicited an MMN of similar amplitude. The MMNs to vowel and pitch deviants did not show significant additivity. This suggests that vowel and pitch are processed by shared neural substrates at the preattentive level.

**Key words:** phoneme; pitch; ERPs; MMN additivity; vowel; music-language interactions

### Introduction

Phoneme and pitch are the building blocks of language and music, respectively. Therefore, finding evidence of shared or independent processing for these linguistic and musical dimensions will contribute to the debate on music and language specificity.<sup>1</sup> Among phonemes, vowels are the most similar in structure to musical tones by varying both in timbre and pitch. Recent data from Kolinsky *et al.*<sup>2</sup> suggest shared processing between vowel and pitch. Musically naïve participants had to classify bi-syllabic nonwords sung on two-tone intervals, the classification being based on either pitch contour or nonword identity. This speeded classification task<sup>3</sup> revealed that vowel and pitch changes are processed in an interactive, namely *integral*, way.

Irrelevant changes of vowel identity interfered with the classification of pitch intervals and vice versa.

However, Kolinsky *et al.*'s conclusions are limited to speeded classification tasks. These tasks do not specify the processing level at which dimensions interact. For example, Caclin and colleagues<sup>4</sup> recently showed that timbre dimensions leading to an integral (interference) pattern with speeded classification tasks appear to be processed separately at a preattentive level,<sup>5</sup> as indexed by the Mismatch Negativity (MMN) component of event-related potentials (ERPs). The MMN<sup>6</sup> is a negative brain wave occurring between 100 and 250 ms after the onset of a stimulus violating an established acoustic regularity. It reflects preattentive detection of the auditory change. So, the results by Caclin *et al.*<sup>4</sup> suggest that the same dimensions can be found to interact in behavioral studies targeting late attentive processes, while being represented separately in early sensory memory. Similarly,

Address for correspondence: Pascale Lidji, Sequence Production Lab, Department of Psychology, McGill University, 1205 Dr Penfield Avenue, Montréal, Québec H3A 1B1, Canada. pascale.lidji@mcgill.ca

vowel and pitch information could be initially processed independently and integrated at later cognitive stages.

In the present study, we investigate whether vowel and pitch changes are processed as integrated or separated units at an early preattentive level indexed by the MMN in sung stimuli. We designed the experiment to determine whether the MMN responses to a stimulus deviating from the standard along two stimulus attributes could be predicted from the additivity of the MMN responses to each attribute alone.<sup>5,7,8</sup> If the changes of two sound dimensions (here, vowel and pitch) are processed by independent neural generators, then the amplitude of the MMN response to a simultaneous change in both dimensions will be predicted by the sum of the amplitudes of the MMNs to changes in each dimension.

Such an outcome would be consistent with the work by Tervaniemi *et al.*<sup>9</sup> In a magnetoencephalography experiment, these authors found evidence for intra-hemispheric specialization: in each hemisphere, the magnetic MMN (MMNm) source location for a phoneme change was superior to that of a chord change. These results suggest that spatially distinct neural populations are involved in the processing of phonemic and musical stimuli. However, the musical (chord) and the phonemic (vowel) changes were presented in different stimuli and presented in separate blocks. Thus, the independence found by the authors may at least in part be due to acoustical differences between

chords and vowels. Here, we used sung stimuli to avoid this potential confound.

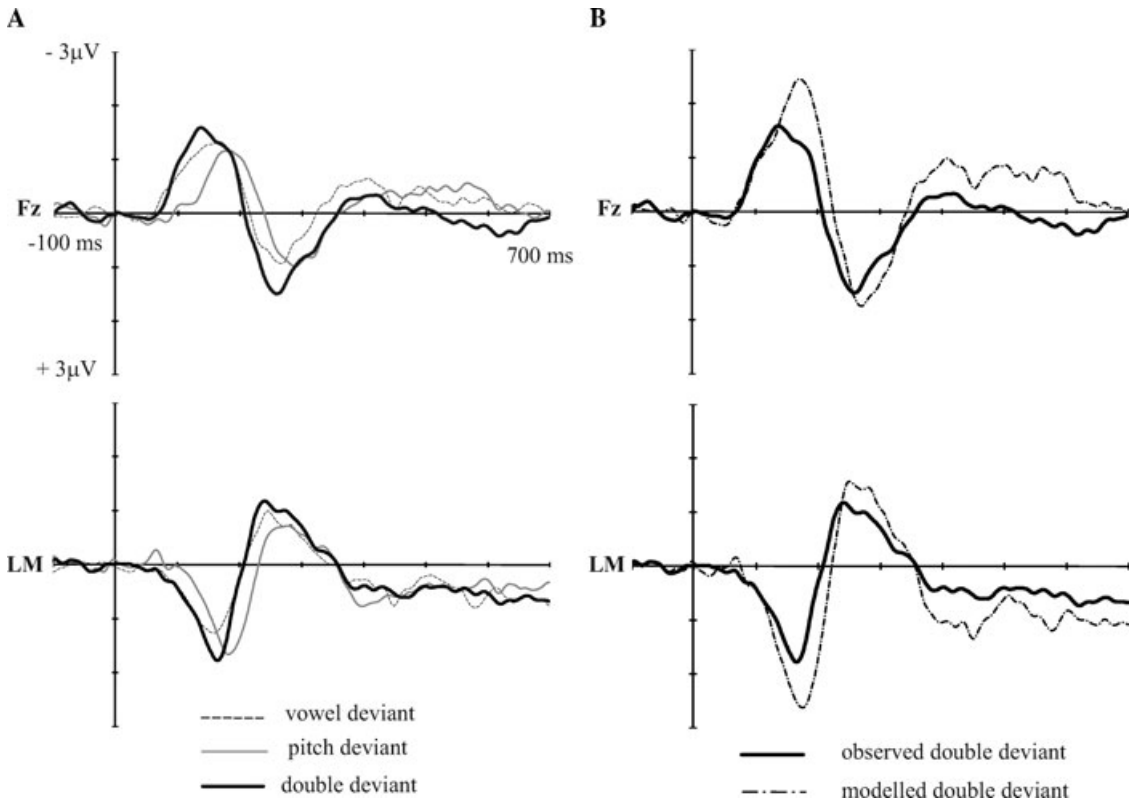
## Methods

ERPs were recorded from 12 right-handed nonmusicians aged 18–38 (7 women and 5 men). Subjects had to watch a silent, subtitled movie, while sequences of auditory stimuli including repetitive standard sounds and rare deviant sounds were presented through headphones at 70 dB SPL. The stimuli were 300-ms synthesized vowels / $\epsilon$ / and / $\upsilon$ / sung on C3 (130 Hz) and C<sup>#</sup>3 (138 Hz), with a fixed inter-stimulus interval of 400 ms. The deviants could differ from the standards in three different ways: pitch height only, vowel identity only, and both pitch height and vowel identity. As illustrated in Table 1, each stimulus served as the standard, and the other three as deviants, across different blocks of stimulus presentations.

The EEG was recorded (bandpass 0.1–70 Hz, sampling rate 256 Hz, impedance < 10 K $\Omega$ ) from 64 Ag-AgCl electrodes at standard 10–10 scalp sites. The tip of the nose served as the reference. The EEG data were corrected for eye movements and filtered with a 0.1–30 Hz bandpass filter (24 dB/octave). Artifacts exceeding  $\pm 100 \mu\text{V}$  were rejected. Epochs of 800 ms, including a 100-ms pre-stimulus interval, were averaged separately for each oddball block (depending on the identity

**TABLE 1.** Illustration of the Experimental Design: Example of Sequence for One Participant

Stimulus category	Block 1	Block 2	Block 3	Block 4
Standard ( $P = 0.82$ )	/ $\epsilon$ / C3 $N = 1050$	/ $\upsilon$ / C3	/ $\epsilon$ / C3#	/ $\upsilon$ / C3#
Vowel deviant ( $P = 0.06$ )	/ $\upsilon$ / C3 $N = 75$	/ $\epsilon$ / C3	/ $\upsilon$ / C3#	/ $\epsilon$ / C3#
Pitch deviant ( $P = 0.06$ )	/ $\epsilon$ / C3# $N = 75$	/ $\upsilon$ / C3#	/ $\epsilon$ / C3	/ $\upsilon$ / C3
Double deviant ( $P = 0.06$ )	/ $\upsilon$ / C3# $N = 75$	/ $\epsilon$ / C3#	/ $\upsilon$ / C3	/ $\epsilon$ / C3



**Figure 1.** (A) MMN difference waves (deviant minus standard) for each category of deviant (vowel, pitch, double) at Fz (top) and LM (bottom). (B) MMN to the observed and the modeled double deviant at Fz and LM.

of the standard) and type of sound (standard, pitch, vowel, and double deviants). To control for stimulus-specific ERP variations, epochs for the same type of deviant were pooled.

Analyses were run at Fz and on the left mastoid (LM) MMN amplitude was maximal. The MMN was computed by subtracting the waveform to the standard from the waveform to each type of deviant, yielding three MMN waves: pitch deviant, vowel deviant, and double deviant. The MMN amplitude was quantified by the mean voltage within a 40-ms window centered at the each individual's peak between 100 and 250 ms after sound onset. To test MMN additivity, a *modeled double deviant* MMN was computed for each subject as the sum of the pitch deviant and the vowel deviant difference waves.

## Results

Each type of deviant elicited a significant MMN at Fz (double deviant,  $-1.6 \mu\text{V}$ ,  $t(11) = -5.43$ ,  $P < 0.0001$ ; vowel deviant,  $-1.3 \mu\text{V}$ ,  $t(11) = -8.29$ ,  $P < 0.0001$ ; and pitch deviant,  $-1.1 \mu\text{V}$ ,  $t(11) = -5.22$ ,  $P < 0.0001$ ). There was a significant polarity inversion at LM for the double [ $1.4 \mu\text{V}$ ,  $t(11) = 6.59$ ,  $P < 0.0001$ ], the vowel [ $1 \mu\text{V}$ ,  $t(11) = 5.61$ ,  $P < 0.0001$ ], and the pitch [ $1.4 \mu\text{V}$ ,  $t(11) = 5.49$ ,  $P < 0.0001$ ] deviants. As can be seen in Figure 1A, the experimental condition did not significantly modulate the amplitude of the MMN at Fz or LM,  $F(2, 22) = 1.34$ , and  $F(2, 22) = 2.093$ , respectively, both  $P$ s  $> 0.1$ .

The mean amplitude of the observed double deviant MMN was compared to that of

the modeled double deviant MMN by paired-samples *t*-tests (Fig. 1B). The observed double deviant MMN was significantly smaller than the modeled MMN at both Fz,  $t(11) = 2.89$ ,  $P < 0.02$ , and LM,  $t(11) = -3.58$ ,  $P < 0.005$ . In other words, the MMNs to vowel changes and to pitch changes are underadditive.

## Conclusions

The MMNs elicited by vowel and pitch changes in sung stimuli do not produce additive effects when manipulated simultaneously in a single stimulus (double deviant). This suggests that the early, preattentive processing of these dimensions is not independent. Thus, contrary to timbre dimensions,<sup>4,5</sup> vowel and pitch would not be processed separately at initial stages of processing and integrated later on.<sup>2</sup> Our results confirm the idea that at least some of the building blocks of music (pitch) and language (vowels) can be processed conjointly.

This finding seems to contradict the claim of Tervaniemi *et al.*<sup>9</sup> that vowel and chord changes are processed by anatomically different sources. However, Tervaniemi and colleagues<sup>9</sup> used chords and vowels, in other words, stimuli differing in both physical structure and cognitive domain (music for chords, speech for vowels). Because MMN is sensitive to the acoustical features of the auditory stimuli,<sup>6</sup> the generator difference observed by Tervaniemi *et al.*<sup>9</sup> might be related to the physical difference between vowels and chords as well as to their abstract linguistic and musical nature. By contrast, our use of sung stimuli introduced vowel and pitch differences in the same acoustical stream.

However, the existence of shared neural generators for the detection of vowel and pitch

changes in the same acoustical stream requires further research. For example, by using MEG, we might be able to demonstrate shared MMNm sources for the detection of the vowel and pitch deviances with sufficient spatial resolution to rule out distinct, but proximal, sources.

## Conflicts of Interest

The authors declare no conflicts of interest.

## References

1. Patel, A.D. 2008. *Music, Language, and the Brain*. Oxford University Press. New York. p. 528.
2. Kolinsky, R. *et al.* 2009. Processing interactions between phonology and melody: vowels sing but consonants speak. *Cognition* In press.
3. Garner, W.R. 1974. *The Processing of Information and Structure*. Erlbaum. Potomac, MD.
4. Caclin, A. *et al.* 2007. Interactive processing of timbre dimensions: a Garner interference study. *Brain Res.* **1138**: 159–170.
5. Caclin, A. *et al.* 2006. Separate neural processing of timbre dimensions in auditory sensory memory. *J. Cogn. Neurosci.* **18**: 1959–1972.
6. Näätänen, R. *et al.* 2007. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* **118**: 2544–2590.
7. Paavilainen, P., S. Valppu & R. Näätänen. 2001. The additivity of the auditory feature analysis in the human brain as indexed by the mismatch negativity:  $1 + 1$  approximate to  $2$  but  $1 + 1 + 1 < 3$ . *Neurosci. Lett.* **301**: 179–182.
8. Levänen, S. *et al.* 1993. Responses of the human auditory cortex to changes in one vs. two stimulus features. *Exp. Brain Res.* **97**: 177–183.
9. Tervaniemi, M. *et al.* 1999. Functional specialization of the human auditory cortex in processing phonetic and musical sounds: a magnetoencephalographic (MEG) study. *NeuroImage* **9**: 330–336.