

Neural correlates of lexical stress: mismatch negativity reflects fundamental frequency and intensity

Hatice Zora, Iris-Corinna Schwarz and Mattias Heldner

Neural correlates of lexical stress were studied using the mismatch negativity (MMN) component in event-related potentials. The MMN responses were expected to reveal the encoding of stress information into long-term memory and the contributions of prosodic features such as fundamental frequency (F0) and intensity toward lexical access. In a passive oddball paradigm, neural responses to changes in F0, intensity, and in both features together were recorded for words and pseudowords. The findings showed significant differences not only between words and pseudowords but also between prosodic features. Early processing of prosodic information in words was indexed by an intensity-related MMN and an F0-related P200. These effects were stable at right-anterior and mid-anterior regions. At a later latency, MMN responses were recorded for both words and pseudowords at the mid-anterior and posterior regions. The P200 effect observed for F0 at the early latency for words developed into an MMN response. Intensity elicited smaller MMN for pseudowords than for words. Moreover, a larger

brain area was recruited for the processing of words than for the processing of pseudowords. These findings suggest earlier and higher sensitivity to prosodic changes in words than in pseudowords, reflecting a language-related process. The present study, therefore, not only establishes neural correlates of lexical stress but also confirms the presence of long-term memory traces for prosodic information in the brain. *NeuroReport* 26:791–796 Copyright © 2015 Wolters Kluwer Health, Inc. All rights reserved.

NeuroReport 2015, 26:791–796

Keywords: event-related potentials, fundamental frequency, intensity, lexical stress, memory trace, mismatch negativity, pitch, prosody

Department of Linguistics, Stockholm University, Stockholm, Sweden

Correspondence to Hatice Zora, MA, Department of Linguistics, Stockholm University, SE-106 91, Stockholm, Sweden
Tel: +46 8 16 29 30; fax: +46 8 15 53 89; e-mail: hatice@ling.su.se

Received 29 May 2015 accepted 17 June 2015

Introduction

To understand the nature of lexical access, it is important to identify the kind of information that is stored in the long-term memory and to study how the brain uses such information. Behavioral studies investigating the contribution of stress information toward lexical access have so far yielded inconclusive results; whereas some of them failed to show the stress effect [1], others indicated listeners' sensitivity to this information in lexical access [2]. Furthermore, the literature so far shows no consensus on the relative contribution of prosodic features such as fundamental frequency (F0) and intensity in stress perception. Whereas F0 is generally claimed to be the most important perceptual correlate of stress, the literature on the role of intensity is inconclusive [3–6].

Neurophysiological and neuroanatomical studies have previously investigated auditory processing and hemispheric status of F0 and intensity perception [7–14]; however, so far no study has directly assessed the interplay between these individual cues and lexical access. F0 and intensity information stored in the long-term memory can be examined by the mismatch negativity (MMN) component in event-related potentials (ERPs). The MMN is a neurophysiological measure that signals the

brain's automatic response not only to any acoustic change in the auditory sensory input such as F0 and intensity but also to higher cognitive processes such as the activation of long-term memory traces for familiar speech sounds and words [11,15,16]. The MMN to familiarity with lexical stress has been investigated previously in a few studies [17–19]. Stress pattern violations in Hungarian words were shown to elicit two distinct MMNs: one to the lack of default stress and one to additional stress, which is argued to show long-term representation of the regular stress pattern in Hungarian [18]. In another study, sensitivity to trochaic and iambic stress patterns in pseudowords was investigated in German infants and adults. Although MMNs were observed in both stress patterns in adults, only the trochaic (frequent) pattern elicited MMN in infants, showing infants' sensitivity to the predominant stress pattern in German [19]. However, these studies could not differentiate language-related effects from the effects of acoustic change on the MMN response. Besides, none of them assessed the interplay between F0 and intensity and lexical access.

The present study aims to investigate the neural correlates of F0 and intensity and their contribution toward lexical access. To separate the effects of these features from the effects of vowel quality, a stress contrastive English verb–noun pair, *upsét-úpsét* was investigated. A pseudoword pair, *ukfét-úkfet*, was used as a control.

This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially.

The pseudoword pair was included to enable a comparison of ERP correlates of F0 and intensity because of lexical processing and nonlexical processing. ERP responses were recorded in a passive oddball paradigm by presenting three distinct deviants interspersed among frequent standard stimuli. Deviants differed from the standards in F0 and intensity alone as well as in both features combined. It is hypothesized that the salience and relevance of stress information in lexical processing will be reflected in the MMN responses, indicating the perceptual discriminability of F0 and intensity and the presence of long-term memory traces for stress information in the brain.

Methods

Participants

The participants were 11 native speakers of American English (seven women; age range 20–32 years). All participants were born and raised in the USA, and reported normal development and hearing. Informed consent was signed before testing, and the study complied with the ethical guidelines.

Materials

The experiment contained one word block and one pseudoword block. The stimuli in the word block consisted of an English disyllabic verb–noun pair in which stress is contrastive: *upsét-úpset*. The stimuli in the pseudoword block were a pseudoword pair, imitating the acoustics of the verb–noun pair: *ukfét,-úkfet*. Each pair was recorded by a female native speaker of American English in an anechoic chamber. The recordings were sampled at a rate of 44.1 kHz with a quantization of 16 bits per sample.

The recordings were manipulated in Praat [20]. The average F0 (in Hz) and the average intensity (in dB) were measured across vowels for each stimulus, and edited to create the deviants. The deviants (with nonfinal stress) were created from the standard (with final stress) by separately manipulating F0 and intensity of the second syllable. As direction and amplitude of the acoustic change might influence the ERP response, deviants were created from the standard by lowering F0 and intensity of the second syllable while the first syllable was held constant in terms of the standard's values. While lowering the values, the second syllables were manipulated to have the same averaged values as derived from the original unstressed syllables.

Procedure

The experiment was run using E-Prime (version 2.0.1.06; Psychology Software Tools, Pittsburgh, Pennsylvania, USA). The stimuli were presented in a multideviant oddball paradigm (one standard, three deviants) through loudspeakers at a comfortable listening level of a 60–65 dB sound pressure level at source. Stimuli were

presented in a random order (standard: $P=8/10$; deviant: $P=2/10$) and the stimulus onset asynchrony was set at 1000 ms. The offset-to-onset interstimulus interval was 575 ms. The experiment had two blocks (word and pseudoword), each block consisting of 720 standard trials and 180 deviant trials, 60 for each deviant. The order of the two blocks was counterbalanced across the participants. A silent documentary was used to take participants' attention off the auditory stimuli.

EEG recordings

The electroencephalography signals were recorded at a sampling rate of 500 Hz using NeuroScan Acquire Software with a SynAmps 2 amplifier (Compumedics Inc., Charlotte, North Carolina, USA). The recordings were made from 30 cap-mounted electrodes (EasyCap; Falk Minow Services, Herrsching-Breitbrunn, Germany) (O2, O1, OZ, PZ, P4, CP4, P8, C4, TP8, T8, P7, P3, CP3, CPZ, CZ, FC4, FT8, TP7, C3, FCZ, FZ, F4, F8, T7, FT7, FC3, F3, FP2, F7, and FP1), horizontal eye electrodes (LO1–LO2), vertical left-eye electrodes (SO1–IO1) and mastoids (M1–M2), and an anterior ground electrode. The impedance was maintained below 5 k Ω at each electrode site. An online band-pass filter of 0.5–70 Hz was applied. The data were referenced to the central, cap-mounted reference electrode during recording.

Event-related potential data analysis

Offline data analysis was carried out in The Math Works Inc., (Natick, Massachusetts, USA) using the EEGLAB toolbox [21]. The EEG data were first filtered with a low-pass filter with a cut-off frequency of 30 Hz and with a high-pass filter with a cut-off frequency of 0.5 Hz. The channels were then re-referenced to both mastoids. To identify and remove eye artifacts, independent component analysis [22] was carried out. Then, the EEG data were segmented into epochs of 800 ms, time-locked to the onset of the second syllable (change onset). A time window of 200 ms before the onset was used for baseline correction. Artifact rejection was set to remove activation exceeding $\pm 100 \mu\text{V}$ at any channel. The grand average was computed per stimulus type and deviant-minus-standard subtractions were calculated for each deviant.

Statistical analysis

Statistical analysis was carried out using SPSS (International Business Machines Corp., Armonk, New York, USA). The measurement window was determined by visual inspection of grand average difference waveforms. Amplitudes were computed as a mean voltage within a 20-ms-window centered at the peak latency in the grand-average waveforms. Initially, four-way repeated-measures ANOVAs were performed in two time windows. The factors were 'Lexicality' with the two levels, Word and Pseudo-word; 'Prosody' with the three levels, F0, intensity (Int) and F0 and intensity combined (F0Int); 'Hemisphere' with the three levels, left (LH),

mid (MH), and right (RH); and ‘Anteriority’ with the two levels, anterior and posterior. If significant interactions occurred, follow-up ANOVAs were performed and the levels were then compared with post-hoc pairwise comparisons using Bonferroni correction. *P*-values are given with Greenhouse–Geisser correction in case of sphericity violations. Effect sizes are reported with η^2 (partial η^2).

Results

Event-related potential data

The grand average difference waves and scalp topographies to deviants are shown for word and pseudoword blocks, respectively, in Figs 1 and 2.

In the word block (Fig. 1a), an early MMN occurred for an intensity between 100 and 150 ms. This slightly right-lateralized response, as shown in the topographical map (Fig. 2, top row, right), is in line with neuroanatomical evidence indicating that the discrimination of sound intensity is reflected in the right frontoparietal network [7]. In contrast, F0 gave rise to a P200, which showed a slightly right-lateralized distribution (Fig. 2, top row, left). This F0-related finding is in agreement with previous results, which indicated that the amplitude of P200

is modulated by the presence of a pitch accent and pitch discrimination is indexed by an increased activity in the right prefrontal cortex [8,9]. These early neural responses were absent in the pseudoword block.

The difference between the two blocks seems to disappear around 200–300 ms after change onset when both words and pseudowords elicited MMNs. The positivity observed for the F0 in words in the first time window developed into an MMN response. In addition, the intensity-related MMN seems to be smaller in pseudowords compared with the MMN in words. The MMNs in words are also more centrally distributed as opposed to a frontal distribution in pseudowords as shown in the topographical maps for all three deviants in both the word (Fig. 2, mid row) and the pseudoword blocks (Fig. 2, bottom row).

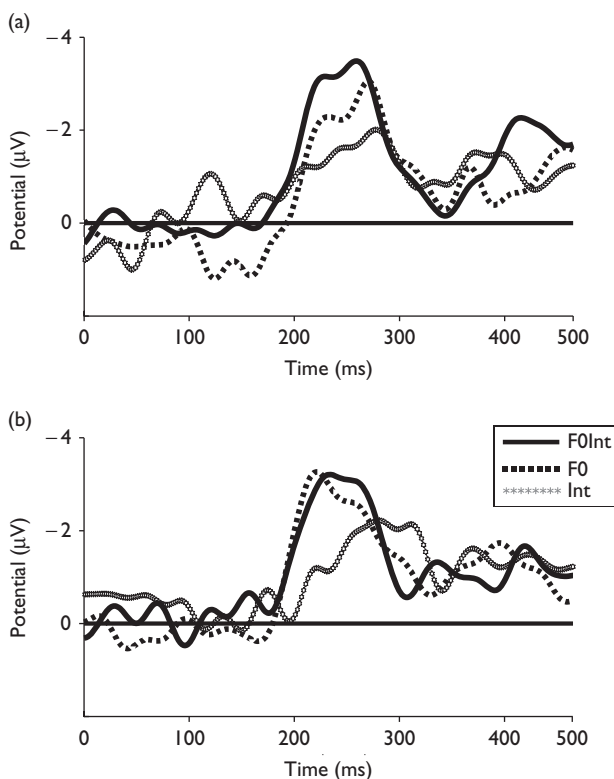
Statistical data

The analysis in the time window 130–150 ms showed a significant three-way interaction of lexicality with anteriority and prosody [$F_{(2,20)} = 3.573$, $P = 0.047$, $\eta^2 = 0.263$]. Follow-up analyses confirmed the interaction between lexicality and prosody at the anterior sites ($F_{(2,20)} = 4.258$, $P = 0.029$, $\eta^2 = 0.299$), but not at the posterior sites [$F_{(2,20)} = 0.854$, $P = 0.441$, $\eta^2 = 0.079$]. A main effect of prosody was then observed in the word block [$F_{(2,20)} = 4.529$, $P = 0.024$, $\eta^2 = 0.312$], but not in the pseudoword block [$F_{(2,20)} = 0.372$, $P = 0.694$, $\eta^2 = 0.036$]. Post-hoc comparisons in the word block indicated that Int (mean = -0.515 μV , $\text{SD} = 0.323$) elicited significantly larger negativity than F0 (mean = 1.181 μV , $\text{SD} = 0.352$, $P = 0.006$), but not than F0Int (mean = 0.469 μV , $\text{SD} = 0.597$, $P = 0.457$). In the word block, only the intensity deviant elicited an MMN and a positive deflection (P200) occurred for the F0 deviant.

The analysis further indicated a trend toward a three-way interaction of hemisphere with anteriority and prosody [$F_{(4,40)} = 2.567$, $P = 0.053$, $\eta^2 = 0.204$]. Post-hoc comparisons at the anterior sites indicated that F0 elicited larger P200 in the RH (mean = 1.459 μV , $\text{SD} = 0.246$) than in the LH (mean = 0.761 μV , $\text{SD} = 0.339$, $P = 0.024$), and that Int elicited larger MMN in the MH (mean = -0.643 μV , $\text{SD} = 0.278$) than in the LH (mean = -0.004 μV , $\text{SD} = 0.263$, $P = 0.012$). Both the P200 and the MMN effects were stable over the right-anterior and mid-anterior regions.

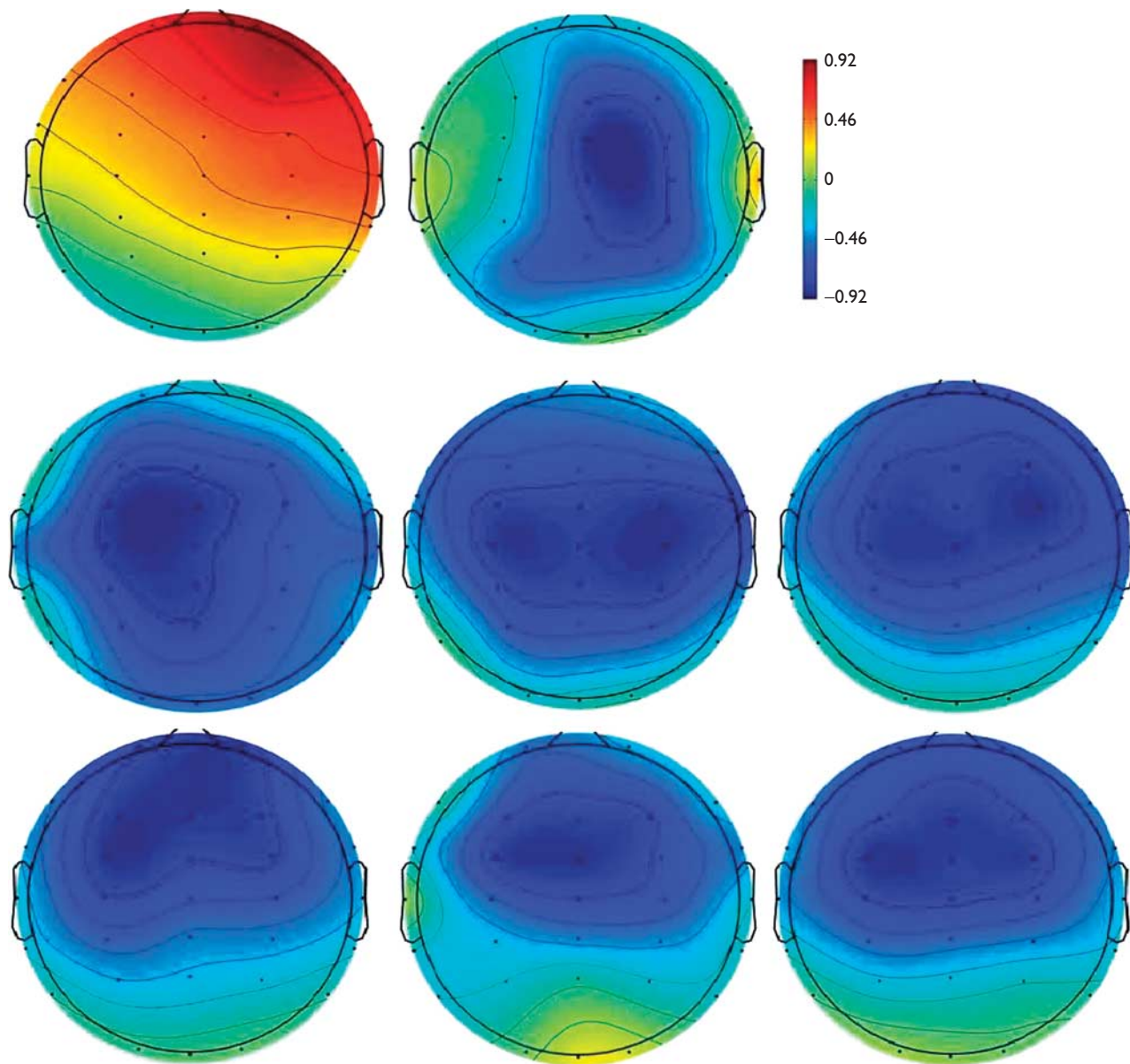
The analysis in the time window 230–250 ms showed a significant main effect of hemisphere [$F_{(2,20)} = 12.515$, $P < 0.001$, $\eta^2 = 0.556$]. Pairwise comparisons indicated that MMN over MH (mean = -1.384) was larger compared with the MMN over LH (mean = -0.504 , $P = 0.012$) and over RH (mean = -0.552 , $P = 0.004$). The analysis further showed a significant two-way interaction of anteriority with lexicality [$F_{(1,10)} = 27.019$, $P < 0.001$, $\eta^2 = 0.730$]; a significant two-way interaction of lexicality

Fig. 1



Grand average difference waveforms for all three deviants on FCZ for word (a) and pseudoword (b) blocks. F0, fundamental frequency; Int, intensity; F0Int, fundamental frequency and intensity combined. Negativity is plotted upward.

Fig. 2



Topographic difference maps for the word block in the first time window (top row): P200 for F0 (left) and MMN for Int (right). Topographic difference maps for the word block in the second time window (mid row): MMN for F0 (left), for Int (mid), and for F0Int (right). Topographic difference maps for the pseudoword block in the second time window (bottom row): MMN for F0 (left), for Int (mid), and for F0Int (right). F0, fundamental frequency; Int, intensity; F0Int, fundamental frequency and intensity combined; MMN, mismatch negativity.

with prosody [$F_{(2,20)}=4.918$, $P=0.018$, $\eta^2=0.330$]; and a significant three-way interaction of lexicality with anteriority and prosody [$F_{(2,20)}=5.805$, $P=0.010$, $\eta^2=0.367$]. Follow-up analyses showed that the two-way interaction of lexicality with prosody was significant at the anterior sites [$F_{(2,20)}=6.319$, $P=0.007$, $\eta^2=0.387$], but not at the posterior sites [$F_{(2,20)}=1.708$, $P=0.207$, $\eta^2=0.146$]. At the anterior sites, prosody had a main effect in pseudowords [$F_{(2,20)}=4.434$, $P=0.025$, $\eta^2=0.307$], but not in words [$F_{(2,20)}=2.661$, $P=0.094$, $\eta^2=0.210$]. In the word

block, all deviants elicited MMNs, but there were no significant differences (mean = -2.445 μV , SD = 0.691 , for F0; mean = -1.778 μV , SD = 0.564 , for Int; mean = -3.233 μV , SD = 0.748 , for F0Int, $P>0.05$). Post-hoc comparisons in the pseudoword block indicated that MMN for F0 (mean = -3.051 μV , SD = 0.334) was larger than MMN for Int (mean = -0.939 μV , SD = 0.594 , $P=0.016$), but not larger than MMN for F0Int (mean = -2.437 μV , SD = 0.512 , $P=0.448$). All deviants elicited MMNs in both word and pseudoword blocks. In the word block, there

were no significant differences between deviants, whereas in the pseudoword block, there was a significant difference between the F0 and the intensity deviant.

The analysis further indicated a trend toward a three-way interaction of hemisphere with anteriority and lexicality [$F_{(2,20)} = 2.890$, $P = 0.079$, $\eta^2 = 0.224$]. Follow-up analyses indicated that the anteriority and lexicality interaction was restricted to MH [$F_{(1,10)} = 8.285$, $P = 0.016$, $\eta^2 = 0.453$]. Subsequent analyses showed that anteriority had a main effect only in pseudowords [$F_{(1,10)} = 37.205$, $P < 0.001$, $\eta^2 = 0.788$]. Pairwise comparisons in the pseudoword block indicated that anterior sites elicited larger negativity (mean = $-2.005 \mu\text{V}$, $\text{SD} = 0.453$) than the posterior sites (mean = $-1.502 \mu\text{V}$, $\text{SD} = 0.425$, $P < 0.001$). In the word block, no significant differences occurred between the anterior (mean = $-2.270 \mu\text{V}$, $\text{SD} = 0.531$) and the posterior sites (mean = $-2.158 \mu\text{V}$, $\text{SD} = 0.424$, $P = 0.528$). In the word block, the posterior sites are also involved in lexical access as opposed to the pseudoword block with only anterior activation.

Discussion

The present study studies neural correlates of lexical stress, specifically the effect of F0 and intensity on lexical access. The ERP responses are expected to reveal the perceptual discriminability of F0 and intensity and the presence of long-term memory traces for stress information in the brain. In contrast to previous ERP studies on word stress [18,19], the present study controlled the direction of stress change by creating deviants by lowering F0 and intensity of the standard, and therefore avoided an intrinsic increase in the ERPs because of an increase in auditory features. Word and pseudoword contrasts were used to differentiate the language-related effects from the possible acoustic-change effects on the neural responses. Thus, language-relevant ERP effects were either absent or smaller for pseudowords compared with words. The present study also broke down stress patterns into several parameters, hence establishing the relevance of F0 and intensity in stress perception and lexical access.

There were significant differences not only between words and pseudowords but also between prosodic features. Early processing of prosodic information in words was indexed by an intensity-related MMN deflection and an F0-related P200. Considering their absence in pseudowords, these early responses in words reflect a language-related process, implying that there are well-developed prosodic representations in the brain to support an automatic discrimination of prosodic information and accelerate lexical access. In addition, the right-anterior region seems to be more sensitive to changes in prosodic features in early latencies, which is in line with neuroanatomical evidence [7,9].

Another enhancement of the MMN was observed around 220 ms after change onset for both words and pseudowords. All deviants elicited MMN responses, confirming the brain's automatic response to changes in the auditory sensory input [10–12]. The positivity observed for the F0 deviant earlier for words developed into an MMN response that implies a language-related process. It is worth noting that this study is the first to show this transition of F0-related ERPs in the processing of lexical stress. In addition, the contribution of intensity toward the measured MMNs was smaller in pseudowords, confirming the importance of intensity in lexical access. This finding makes a significant contribution toward the inconclusive literature on the role of intensity in stress [3–6]; even though not as salient as F0, intensity is an important prosodic cue to determine word stress and contributes to lexical access. Moreover, in the word block, the posterior sites are also involved in lexical access as opposed to the pseudoword block with only anterior activation. This indicates that a larger brain area is recruited to the processing of words than to the processing of pseudowords.

Conclusion

ERPs to prosodic changes were obtained in two different time windows for words while restricted to one time window for pseudowords. An intensity-related MMN and F0-related P200 reflected early processing of prosodic information in words in the first time window. Another MMN was recorded for both words and pseudowords in the second time window. The positivity observed earlier for F0 for words developed into an MMN response in this later latency, and intensity elicited larger MMN for words than for pseudowords. Considering their absence for pseudowords, these early and larger ERP responses for words cannot be attributed only to simple acoustic changes; they are probably a result of pre-existing memory traces for prosodic information. In conclusion, apart from the acoustic change-detection process, a language-related process contributed toward the ERPs; the brain not only detected prosodic changes but also used them in lexical access.

Acknowledgements

The authors thank Mikael Roll, Merle Horne, and Yury Shtyrov for their constructive comments and feedback.

This work was supported by Forskarskolans i Språkvetenskap (FoSpråk).

Conflicts of interest

There are no conflicts of interest.

References

- 1 Cutler A. Forbear is a homophone: lexical prosody does not constrain lexical access. *Lang Speech* 1986; **29**:201–220.

- 2 Cooper N, Cutler A, Wales R. Constraints of lexical stress on lexical access in English: evidence from native and nonnative listeners. *Lang Speech* 2002; **45**:207–228.
- 3 Fry DB. Experiments in the perception of stress. *Lang Speech* 1958; **1**:126–152.
- 4 Lieberman P. Some acoustic correlates of word stress in American English. *J Acoust Soc Am* 1960; **32**:451–454.
- 5 Beckman ME. *Stress and non-stress accent*. Dordrecht: Foris; 1986.
- 6 Sluijter AM, van Heuven VJ. Spectral balance as an acoustic correlate of linguistic stress. *J Acoust Soc Am* 1996; **100** (Pt 1):2471–2485.
- 7 Belin P, McAdams S, Smith B, Savel S, Thivard L, Samson S, Samson Y. The functional anatomy of sound intensity discrimination. *J Neurosci* 1998; **18**:6388–6394.
- 8 Heim S, Alter K. Prosodic pitch accents in language comprehension and production: ERP data and acoustic analysis. *Acta Neurobiol Exp (Wars)* 2006; **66**:55–68.
- 9 Zatorre RJ, Evans AC, Meyer E. Neural mechanisms underlying melodic perception and memory for pitch. *J Neurosci* 1994; **14**:1908–1919.
- 10 Näätänen R, Gaillard AW, Mäntysalo S. Early selective-attention effect on evoked potential reinterpreted. *Acta Psychol (Amst)* 1978; **42**:313–329.
- 11 Näätänen R, Paavilainen P, Rinne T, Alho K. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin Neurophysiol* 2007; **118**:2544–2590.
- 12 Näätänen R, Winkler I. The concept of auditory stimulus representation in cognitive neuroscience. *Psychol Bull* 1999; **125**:826–859.
- 13 Pardo PJ, Mäkelä JP, Sams M. Hemispheric differences in processing tone frequency and amplitude modulations. *Neuroreport* 1999; **10**:3081–3086.
- 14 Koso A, Hagiwara H. Event-related potential evidence of processing lexical pitch-accent in auditory Japanese sentences. *Neuroreport* 2009; **20**:1270–1274.
- 15 Dehaene-Lambertz G. Electrophysiological correlates of categorical phoneme perception in adults. *Neuroreport* 1997; **8**:919–924.
- 16 Shtyrov Y, Pulvermüller F. Neurophysiological evidence of memory traces for words in the human brain. *Neuroreport* 2002; **13**:521–525.
- 17 Ylinen S, Strelnikov K, Huotilainen M, Näätänen R. Effects of prosodic familiarity on the automatic processing of words in the human brain. *Int J Psychophysiol* 2009; **73**:362–368.
- 18 Honbolygó F, Csépe V, Ragó A. Suprasegmental speech cues are automatically processed by the human brain: a mismatch negativity study. *Neurosci Lett* 2004; **363**:84–88.
- 19 Weber C, Hahne A, Friedrich M, Friederici AD. Discrimination of word stress in early infant perception: electrophysiological evidence. *Brain Res Cogn Brain Res* 2004; **18**:149–161.
- 20 Boersma P, Weenink D, Praat. Doing phonetics by computer [Computer program], version 5.3.24. Available at: <http://www.praat.org>. [Accessed 2 June 2012].
- 21 Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 2004; **134**:9–21.
- 22 Jung TP, Makeig S, Humphries C, Lee TW, McKeown MJ, Iragui V, Sejnowski TJ. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* 2000; **37**:163–178.