Saliency or template? ERP evidence for long-term representation of word stress

Ferenc Honbolygó [1,2], Valéria Csépe [1,2]

[1] Institute of Cognitive Neuroscience and Psychology, Research Centre of Natural Sciences, Hungarian Academy of Sciences
[2] Institute of Psychology, Eötvös Loránd University

Corresponding author: Ferenc Honbolygó
E-mail: honbolygo.ferenc@ttk.mta.hu
Address: Institute of Cognitive Neuroscience and Psychology, Research Centre of Natural Sciences, Hungarian Academy of Sciences, H-1068, Budapest, Szondi u. 83-85.

Manuscript of the article that appeared in:
International Journal of Psychophysiology. 2013 Feb;87(2):165-72.
DOI: 10.1016/j.ijpsycho.2012.12.005.

Abstract

The present study investigated the event-related brain potential (ERP) correlates of word stress processing. Previous results showed that the violation of a legal stress pattern elicited two consecutive Mismatch Negativity (MMN) components synchronized to the changes on the first and second syllable. The aim of the present study was to test whether ERPs reflect only the detection of salient features present on the syllables, or they reflect the activation of long-term stress related representations. We examined ERPs elicited by pseudowords with no lexical representation in two conditions: the standard having a legal stress patterns, and the deviant an illegal one, and the standard having an illegal stress pattern, and the deviant a legal one. We found that the deviant having an illegal stress pattern elicited two consecutive MMN components, whereas the deviant having a legal stress pattern did not elicit MMN. Moreover, pseudowords with a legal stress pattern elicited the same ERP responses irrespective of their role in the oddball sequence, i.e., if they were standards or deviants.

The results suggest that stress pattern changes are processed relying on long-term representation of word stress. To account for these results, we propose that the processing of stress cues is based on language-specific, pre-lexical stress templates.

Keywords: speech perception, word stress, ERP, MMN

Introduction

Stress is a relative emphasis given to certain syllables within words, or to certain words within sentences. It belongs to the suprasegmental or prosodic features of speech, and plays either a culminative or a demarcative role, emphasizing or separating certain parts of the speech stream (for review see Kager, 2007). In the present study we have investigated how the human brain detects changes in the stress pattern of words, and how long-term representations influence this.

Stress plays an important role in the segmentation of continuous speech into words. According to the Metrical Segmentation Strategy hypothesis (Cutler and Norris, 1988) stressed syllables are the starting points of accessing the mental lexicon, because of their prominence and of their concurrence with the beginning of words. Grosjean and Gee (1987) suggested that during speech processing the acoustic input is written into a pre-lexical, intermediary representation that contains the description of phonetical segments as stressed and unstressed, and lexical access is initiated whenever the processing system detects a stressed syllable. The transformation of speech input into abstract stress related representations helps to overcome the problem of the high variability of stress related acoustic cues (for review see Cutler, 2005), as the lexical access can be based on a discrete category, instead of a variable acoustic feature.

Mattys et al. (2005) suggest that besides stress other cues may aid the segmentation process: phonotactical probability, coarticulation, lexical knowledge, sentential context, etc. which are organized in a hierarchical way. The relative dominance of the cues depends on environmental conditions: under optimal circumstances lexical cues are used, but in noisy environments the importance of stress increases. Mattys et al. (2005) also suggest that languages may use different segmentation cues, and in languages with fixed stress patterns where stress is always on the same syllable (e.g., first syllable as in Hungarian or Finnish), stress might play a particularly important role as a segmentation cue.

In order to study the processing and representation of speech information, the Mismatch Negativity (MMN) event-related brain potential (ERP) component have been shown to be particularly valuable (see Näätänen et al., 2007 for review). MMN allows examining how the human brain processes linguistic information almost in an online way without requiring participants to make conscious decisions about the speech stimuli, and thus allowing to avoid the interplay of extra-linguistic processes. The MMN is a negative going auditory component of fronto-central voltage maximum, appearing 100-250 ms after the onset of change. It is elicited in passive oddball paradigms, in which participants are presented with frequently repeated stimuli of the same acoustic features (standard), interspersed by rarely repeated stimuli differing from the standard in some discriminable features (deviant). The MMN is currently interpreted as a brain electrical correlate of the mainly pre-attentive detection of violation of simple or complex regularities (Winkler et al., 2009).

In the case of linguistic information there is evidence that the MMN is elicited relying on long-term representations of regularities and higher-level rules (for review see Näätänen, 2001). Näätänen et al. (1997) found that the MMN elicited by native vs. non-native speech sound contrasts was differently localized in the brain, and this was taken as evidence that the MMN for native speech sounds is based on long-term representations of phonemic information. Pulvermüller et al. (2001) and Shtyrov and Pulvermüller (2002) demonstrated the existence of memory traces for individual spoken words, based on which the MMN is elicited. Pulvermüller and Shtyrov (2006) in a review article suggested that the MMN might reflect the processing of complex linguistic information at the lexical, semantic and syntactic level outside the focus of attention.

There is – at the moment limited in number – experimental evidence that prosodic and more specifically stress related information is also processed pre-attentively by the human brain, and changes in the stress pattern of words may also elicit the MMN component. Weber et al. (2004) found that 5 months old German infants could discriminate words of different stress patterns. Infants showed a positive going Mismatch Response (MMR) to deviant stimuli with stress on the first syllable compared to standard stimuli with stress on the second syllable. Friederici et al. (2007) demonstrated that 4-5 months old German and French infants showed specific

MMR to the native stress patterns: German infants to the deviant with stress on the second syllable and French infants to the deviant with stress on the first syllable. Friedrich et al. (2009) found an early and a late MMR in 4-5 months old German infants to the deviant with a native stress pattern. According to these studies, infants can discriminate changes in the stress patterns of words as early as 4-5 months, and a preference to native stress patterns can be demonstrated as well.

Based on these studies, it can be expected that adults show similar discrimination abilities for stress patterns. Our own study with adult participants (Honbolygó et al., 2004) showed that a word with stress on the second syllable elicited two MMN components when contrasted with a word with stress on the first syllable, which is the legal stress pattern in Hungarian. The consecutive MMN components were consistently synchronized to the changes on the first and second syllable of the deviant word. The interpretation of the results was ambiguous however: the MMNs could signal the detection of acoustic differences on the first and second syllables, or the detection of the stress rule violation.

The study of Ylinen et al. (2009) provided further MMN data on the processing of stress patterns. The authors investigated the processing of Finnish words and pseudowords with unfamiliar (stress on the second syllable) versus familiar (stress on the first syllable) word stress patterns using multiple versions of CVCV utterances. According to the results, the pseudowords and words with unfamiliar stress pattern elicited two MMNs related to the first and second syllables of utterances. The words with familiar stress pattern however elicited a single MMN in the earlier time windows. Moreover, the MMN was delayed in words with unfamiliar stress. The interpretation of the results was that the MMN signals the lexical status of words, and the unfamiliar stress of meaningful words increases the computational need and delays the processing of words. These results validate our previous results by showing a similar double MMN pattern in the case of pseudowords and words with unfamiliar stress.

Ylinen et al. (2009) thus have demonstrated that the detection of stress pattern change works at the pre-lexical level, since pseudowords can also elicit the double MMN pattern, and argue against an explanation that the double MMN is based on the processing of acoustical differences between stimuli, because they find this ERP pattern when multiple tokens of standard and deviant stimuli are used. However, the origin of the double MMN appearing in the studies of Honbolygó et al. (2004) and Ylinen et al. (2009) is still not clear in terms of underlying representations. The MMNs could indicate the detection of the acoustic / phonetic saliency or the lack of saliency related to stress on the first or second syllables which would require short-term comparison mechanisms. Alternatively, they could indicate the violation of the legality of the stress pattern, which would require a long-term representation of the stress pattern against which the comparison could be made. These two hypotheses can be disentangled in an oddball paradigm where the standard and deviant stimuli are reversed, and the stimuli with legal stress pattern act as deviant. In this situation, the deviant mismatches the short-term trace only, since it has a legal stress pattern that does not violate the supposed long-term stress representation. Therefore if in the reversed situation we obtain the same double MMNs, than we can refute the long-term representation hypothesis, and assume that the double MMN is a result of a short-term comparison process. If however we obtain a different ERP pattern, than we can argue that the perception of stress pattern change is based on long-term representations.

The nature of the long-term stress pattern representations might be assumed to be similar to the phoneme traces suggested by Näätänen (2001). According to this theory, phoneme traces function as recognition templates, and they are activated during the processing of speech sounds by the acoustic characteristics of these sounds. As Näätänen (2001) suggests, the activation of traces corresponds to phoneme categorization that is to the process of matching the speech sounds having variable acoustic characteristics to the abstract phoneme representations. We suggest that there are similar long-term representations in the case of stress information, and we propose to refer to these as stress templates, because the term “trace” does not capture the rule-based nature of these representations. We assume that stress templates are phonological representations relying on rule extraction comprising legal or expected stress patterns of a given language. These representations are probably pre-lexical, as they affect the processing of pseudowords, and since the predominant stress patterns differ between languages, they are probably language-specific as well.

The aim of the present experiment was to investigate if the processing of stress pattern changes are based on short-, or long-term representations, that is whether the comparisons mechanisms leading to the MMN component are based on the processing of the saliency of stimuli, or they reflect the activation of long-term stress templates. In order to investigate this, we created two experimental conditions where we used pseudowords and varied the legality of the deviant stimulus: in the first condition the pseudoword with legal stress pattern (stress on the first syllable) served as standard stimulus and the pseudoword with illegal stress pattern (stress on the second syllable) as deviant stimulus. In the second condition, we reversed the standard and deviant stimuli, and the pseudoword with illegal stress pattern became standard stimulus, and the pseudoword with legal stress pattern became deviant stimulus. We expected to find that the irregular deviant in the first condition would elicit two MMN components similar to previous studies (Honbolygó et al., 2004., Ylinen et al., 2009), but the regular deviant in the second condition would not elicit the same two MMNs. We did not have a specific hypothesis about the exact nature of the ERP pattern in the second condition, but based on the results of Ylinen et al. (2009) who found a single MMN to the word with the familiar stress pattern, the same single MMN could be expected here.

In the present study, unlike Ylinen et al. (2009) we used a single utterance of a pseudoword. As Pulvermüller and Shtyrov (2006) suggest, in order to study the early effects of word processing it is preferable to reduce stimulus variance, and use single words. Since the study of Ylinen et al. (2009) provided evidence that the MMN correlates of stress pattern processing are similar when the acoustical variance is increased by using multiple tokens of words, we were confident that the lack of such variance will not influence our results.

Methods

Participants

Fifteen participants took part in the experiment (6 females). Their age range was 18-26 years (mean age: 20.6 years). All participants were native speakers of Hungarian, were right-handed, and their hearing level was in the normal range according to the audiometry measurement. Participants received payment for their participation in the experiment, and they gave their informed consent according to the Declaration of Helsinki prior to the experiment. The electrophysiological experiment was approved by the Ethical Board of the Research Institute of Psychology, HAS.

Stimuli and procedure

We created two pseudowords differing only in their stress pattern. The pseudoword with the legal stress pattern was bisyllabic, and consisted of the reiteration of the same syllable (“bebe”). This allowed us to create a stimulus with syllables having the same segmental content, but differing in their prosodic properties (the first syllable being stressed, and the second unstressed). The stimulus was created by recording the pseudoword in a sentence context, spoken by a female speaker. In Hungarian, stress in bisyllabic words is always on the first syllable (Siptár and Törkenczy, 2007); therefore this pattern can be considered as legal, while stress on the second syllable is illegal. We use the terms “legal / illegal” instead of “regular / irregular”, since they reflect better that changing the stress pattern in Hungarian does not simply make the word irregular (i.e., less frequent, as in the case of Italian for example, see Colombo, 1992), but also illegal, because it violates the only possible stressing pattern (in the case of two syllable long words).

The pseudoword with the illegal stress pattern was created from the pseudoword with the legal stress pattern by excising and reversing the two syllables by means of sound editor software (Praat, Boersma and Weenink, 2007). This way it was possible to make sure that the standard and deviant stimuli were identical, except for their prosodic structure. The acoustic properties of the stimuli are shown in Figure 1.

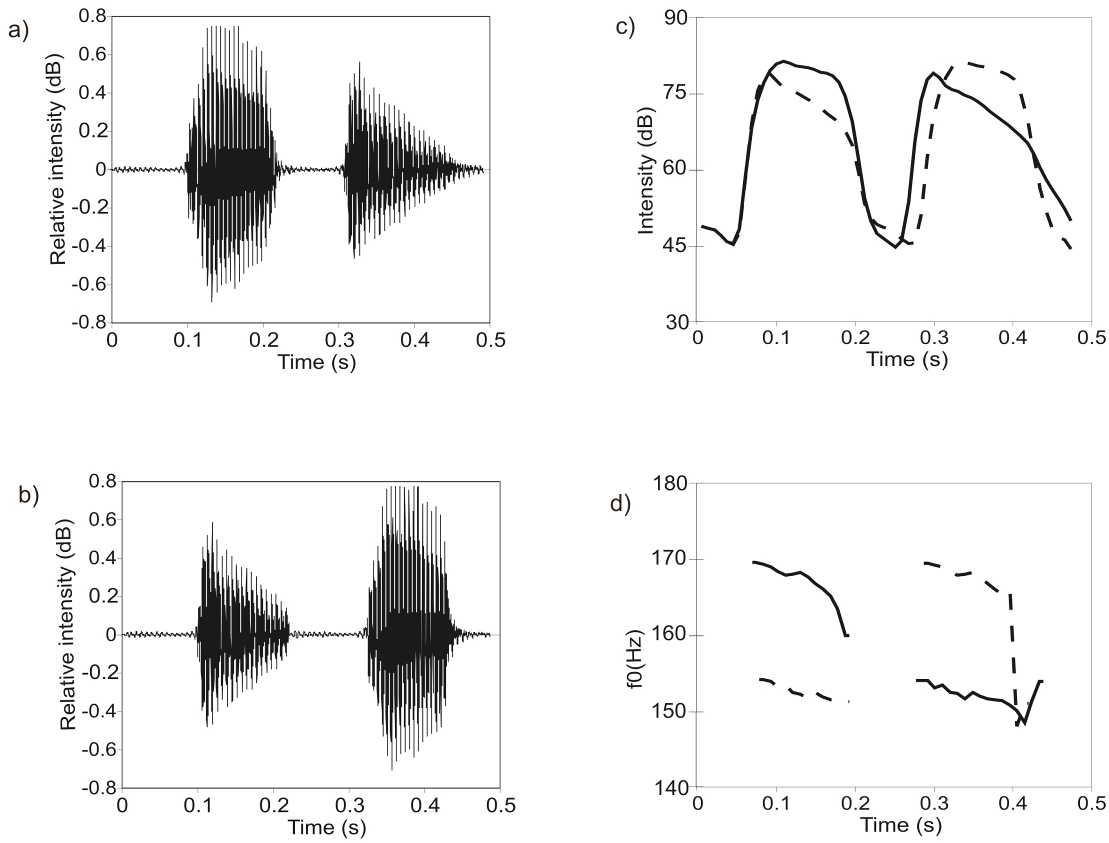

Figure 1. Acoustic characteristics of the stimuli used in Experiment 1. a) and b) show the oscillogram of the legal (stress on the first syllable) and illegal (stress on the second syllable) pseudowords respectively. c) shows the intensity envelope curve, and d) the f0 contour of the stimuli. Continuous line: legal stimulus; broken line: illegal stimulus.

According to the acoustic measurements, the stressed and unstressed syllables differed in their maximal intensities (84.5 vs. 82 dB) and maximal f0s (170 vs. 155 dB). This is in accordance with the assumption that in Hungarian word stress is realized by changes of both intensity and pitch (Varga, 2002).

Stimuli were presented in a passive oddball paradigm with a stimulus onset asynchrony (SOA) varying randomly between 730-830 ms. The duration of stimuli were 539 ms. Participants were seated in a comfortable chair in a sound proof and electrically shielded room, and they had to ignore the stimuli presented via headphones (AKG Varimotion System, K401), while they were watching a silent movie of their choice. We used two conditions. In the first condition (legal condition), the pseudoword with the legal stress pattern was the standard stimulus, and the pseudoword with the illegal stress was the deviant stimulus. In the second condition (illegal condition), the stimuli were reversed. In each condition, stimuli were presented in four blocks, with a pre-specified quasi-random order. Each block began with at least ten standard stimuli, and deviants were separated by at least one standard stimulus. The probability of the deviant stimuli was 20% (n = 200). The order of presentation of the two conditions was counterbalanced between participants. Stimuli were presented using Presentation software (v. 12.1). The experiment lasted for about 2 hours, including the application and removal of electrodes.

EEG recording and analysis

EEG activity was measured using a 32 channel recording system (BrainAmp amplifier and BrainVision Recorder software, BrainProducts GmbH). The Ag/AgCl sintered ring electrodes were mounted in an electrode cap (EasyCap) on the scalp according to the 10% equidistant system at the following positions: Fp1, Fp2, F9, F7, F3, Fz, F4, F8, F10, FC5, FC1, FC2, FC6, T9, T7, C3, Cz, C4, T8, T10, CP5, CP1, CP2, CP6, P7, P3, P4, P8, O1, O2, P9, P10. We used Pz as a reference, and the electrode position between Fz and Fpz as ground. Electrode contact impedances were kept below 10 kΩ. EEG data was recorded with a sampling frequency of 500 Hz, using a band-pass online filter between 0.1-70 Hz.

The EEG data was analyzed offline by using BrainVision Analyzer software. Eye-movement artifacts were corrected with the help of independent component analysis (ICA). ICA is a method capable of blindly separating signals having different sources, and therefore it is suggested to be efficient for separating the EEG signal from non-cephalic artifacts (Delorme et al., 2007). In order to correct eye-movement artifacts, the raw EEG was first decomposed into ICA components using the Infomax algorithm, and then 2-5 components related to eye-movements were selected by visual inspection by an expert, relying on both the time course and the spatial maps of the components. This was followed by the reconstruction of EEG from the remaining ICA components, thus leaving out the eye-movement related activity without loosing data. The data was then band-pass filtered between 0.3-30 Hz (12 dB/oct), and notch filtered at 50 Hz. After that, the continuous EEG was segmented into epochs synchronized to the onset of stimuli from -100 to 800 ms, separately for the standards and deviants, and baseline corrected using the pre-stimulus segment. We applied an automatic artifact rejection algorithm to reject those segments where the activity exceeded +/-80 μV. This was necessary in order to remove artifacts still remaining in the data after the ICA correction. Finally, remaining epochs were averaged, and the data was re-referenced to the average activity of all electrodes. For the standard and deviant stimuli an equal number of epochs were averaged together, by analyzing only the standards preceding the deviants.

Statistical analysis

For the statistical analysis, peak amplitudes and latencies were measured automatically on the individual ERPs separately for the standard and deviant stimuli. Peaks were searched in latency windows corresponding to the ERP components identified on the grand average and difference wave ERPs (see Figure 2 and 4). In both legal and illegal conditions two latency windows between 100-200 ms and 200-300 ms were used to assess early ERP effects (P2-N2), and three latency windows between 320-420 ms, 420-520 ms and 520-620 ms were used to evaluate MMN effects. In order to test the MMN effects, i.e., if the standard and deviant stimuli elicited different ERPs, a 6x2 repeated measures ANOVA with Electrodes (F3, Fz, F4, C3, Cz, C4) and Stimulus (standard vs. deviant) as within-subjects factors was performed in the two conditions separately. We also analyzed if the pseudowords with legal and illegal stress patterns elicited different ERPs in the standard and deviant positions in the legal and illegal conditions by running a 6x2x2 repeated measures ANOVA with Electrodes (F3, Fz, F4, C3, Cz, C4), Stimulus (standard vs. deviant) and Legality (legal vs. illegal) as within-subjects factors. We used the same analysis in order to test the early ERP effects. Where necessary, the Greenhouse-Geisser method (Greenhouse and Geisser, 1959) was used to correct the violation of sphericity assumption. We used Tukey HSD as a post-hoc test. Peak latencies were compared by running paired sample t-tests.

Results

ERP results

Figure 2 shows grand average ERP curves to the standard and deviant stimuli and the difference wave obtained by subtracting the ERPs to the standard stimulus from that of the deviant in the legal condition. As it is visible on the difference waves, the deviant elicited two large negative components at 370 and 570 ms, and a smaller in between at 450 ms. The two larger negativities were considered as two separate

MMNs. Figure 2 illustrates that both components were maximal at fronto-central electrode sites, and showed a polarity reversal at the occipital sites, further supporting that both components were genuine MMNs.

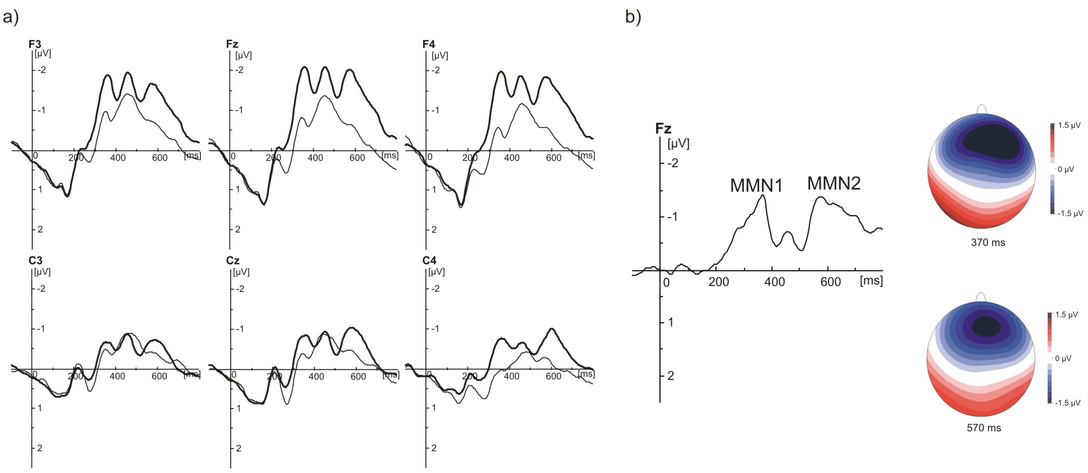

Figure 2. ERP results in the legal condition. a) Grand average ERPs to the standard with the **legal** stress (thin line) and deviant with the **illegal** stress (thick line) on six fronto-central electrodes. b) Difference wave obtained by subtracting the ERPs to the standard from the ERPs to the deviant stimuli (Fz electrode). The amplitude map shows the amplitude distribution of the MMN components on the scalp. ERPs are low-pass filtered by 15 Hz for the purpose of presentation only here and on the following ERP figures, and negativity is plotted upwards.

Figure 3 depicts grand average ERPs and the difference wave in the illegal condition. The difference wave shows three negativities: two smaller ones at 350 and 510 ms, and a larger one at 410 ms. These three negative components were in the same latency windows as the negative components in the legal condition, but their amplitudes were rather different: the first and third components were relatively small, while the second one was larger than these two, which was the opposite pattern compared to components in the legal condition.

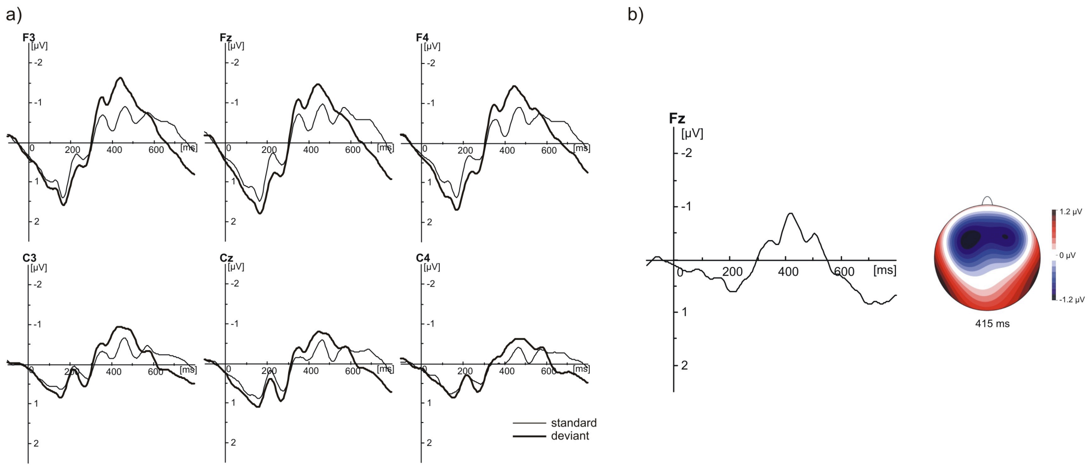

Figure 3. ERP results in the illegal condition. a) Grand average ERPs to the standard with **illegal** stress (thin line) and deviant with **legal** stress (thick line) on six fronto-central electrodes. b) Difference wave obtained by subtracting the ERPs to the standard from the ERPs to the deviant stimuli (Fz electrode). The map shows the amplitude distribution of the MMN component on the scalp.

In order to account for ERP effects resulting from comparing physically different stimuli, we plotted ERPs to the same legally and illegally stressed pseudoword in the standard and deviant positions (Figure 4). We also calculated difference waves by subtracting ERPs to physically identical stimuli in the two conditions (deviant minus standard in the *other* condition). According to the difference waves on Figure 4, the pseudoword with legal stress pattern elicited a broad, low-amplitude negativity in the deviant conditions compared to the standard conditions, while the illegal pseudoword elicited three consecutive negativities. Moreover, the legally stressed pseudoword elicited early ERP components (an early positivity, P2 and an early negativity, N2) that differed between the standard and deviant stimuli, whereas there was no such difference in the case of the illegally stressed pseudoword.

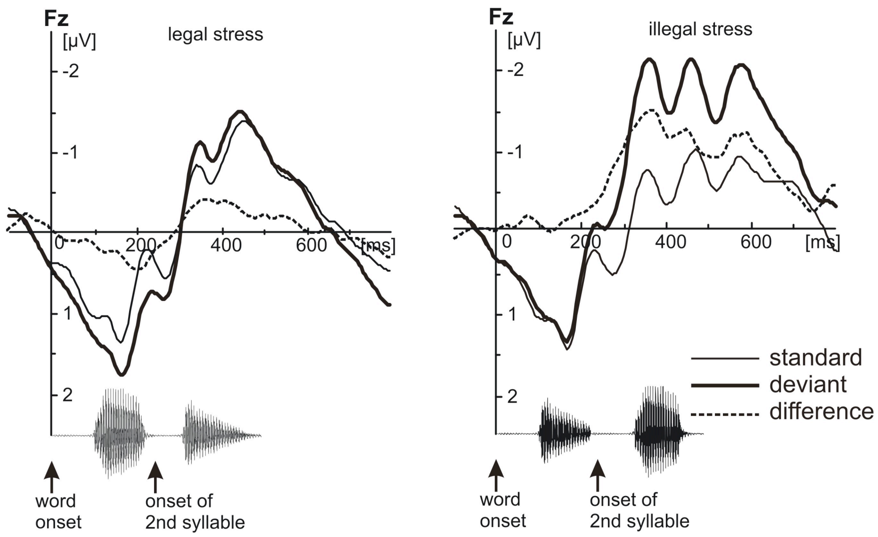

Figure 4. ERPs elicited by the pseudoword with legal and illegal stress pattern in the standard (thin line) and deviant (thick line) positions. The figures depict ERPs to the same stimulus in two different positions, and the difference wave (dashed line) obtained by subtracting the ERPs to the same pseudoword in the deviant and standard positions. The schematic depiction of stimuli illustrates the timing relation between the stimuli and the ERPs.

Statistical results

Legal condition

In the first latency window (320-420 ms) the repeated measures ANOVA resulted in significant Electrode ($F(5,70) = 39.34$, $p < 0.01$) and Stimulus ($F(1,14) = 30.73$, $p < 0.01$) main effects, and a significant Electrode x Stimulus interaction ($F(5,70) = 9.77$, G-G corrected $p < 0.01$). The post-hoc Tukey HSD test indicated that except for the C3 electrode, the standard and deviant stimuli elicited significantly different ERPs on all electrodes (all ps < 0.05). There was no significant difference between the latencies of the standard and deviant peaks. These results support that a significant MMN appeared at 370 ms.

In the second latency window (420-520 ms) we found a significant Electrode ($F(5,70) = 31.98$, $p < 0.01$) and Stimulus ($F(1,14) = 7.6$, $p < 0.02$) main effect, and a tendency for Electrode x Stimulus interaction ($F(5,70) = 2.38$, G-G corrected $p = 0.06$). The post-hoc Tukey HSD test showed that the difference between the standard and deviant was significant on the Fz and F4 electrodes ($p < 0.05$). There was no significant difference between the latencies of the standard and deviant peaks. These results show that the negative component between the two MMNs was different for the standard and deviant stimuli, but this difference was significant on certain electrodes only.

In the third latency window (520-620 ms), the Electrode ($F(5,70) = 19.98$, $p < 0.01$) and Stimulus ($F(1,14) = 25.76$, $p < 0.01$) main effects and the Electrode x Stimulus interaction ($F(5,70) = 8.3$, G-G corrected $p < 0.01$) were all significant. The Tukey HSD test demonstrated that the standard and deviant significantly differed on all electrodes (all *ps* < 0.01). There was no significant difference between the latencies of the standard and deviant peaks. These results support that a significant MMN appeared at 570 ms.

Illegal condition

In the first latency window (320-420 ms), the repeated measures ANOVA showed a significant main effect of Electrode ($F(5,70) = 18.89$, $p < 0.01$) and Stimulus ($F(1,14) = 5.67$, $p < 0.05$). The Tukey HSD test showed a significant difference between the standard and deviant on the Fz and F4 electrodes ($p < 0.01$). There was no significant difference between the latencies of the standard and deviant peaks. According to these results, the deviant elicited a slightly, but significantly larger negativity than the standard in the first latency window.

In the second latency window (420-520 ms), the Electrode ($F(5,70) = 14.07$, $p < 0.01$) and Stimulus ($F(1,14) = 15.21$, $p < 0.01$) main effects were significant, but there was no significant interaction between the two factors. According to the Tukey HSD test, the standard and deviant differed on the F3 and F4 electrode ($p < 0.05$). We found a significant difference in the peak latencies ($t(14) = 2.44$, $p < 0.05$), the deviant having an earlier peak (about 18 ms) than the standard. Although the difference between the ERPs elicited by the standard and deviant stimuli were larger than in the previous latency window, it was significant only on certain electrodes.

In the third latency window (520-620 ms), there was only a significant main effect of Electrode ($F(5,70) = 12.23$, $p < 0.01$). The peak latencies were not significantly different. The third negativity visible on the difference waves (Figure 3) could not be statistically confirmed.

Interaction between legality and stimulus type

To compare the ERPs obtained in the two different conditions for the two different stimulus types, we conducted a 6x2x2 repeated measures ANOVA with factors of Electrodes, Stimulus and Legality on the amplitude and latency data obtained in the two early ERP latency windows and the three MMN latency windows.

In the first early ERP latency window (100-200 ms) we did not find any significant effects. In the second window (200-300 ms) however, we found a significant main effect of Legality ($F(1,14) = 6.44$, $p < 0.05$), a significant Stimulus x Legality interaction ($F(1,14) = 13.23$, $p < 0.01$), and a significant Electrode x Stimulus x Legality interaction ($F(5,70) = 3.1$, G-G corrected $p < 0.05$). The Tukey HSD post-hoc test calculated on the Electrode x Stimulus x Legality interaction showed a significant difference between the ERPs elicited by the standard and the deviant in case of the legally but not the illegally stressed pseudoword, and only on the F3 and Fz electrodes ($p < 0.01$). Moreover, the deviant elicited a larger negativity in the case of the illegally as compared to the legally stressed pseudoword, on electrodes F3, Fz, F4 ($p < 0.05$).

In the first MMN latency window (320-420 ms), the repeated measures ANOVA indicated a significant main effect of Electrode ($F(5,70) = 31.09$, $p < 0.01$), Stimulus ($F(1,14) = 39.02$, $p < 0.01$) and Legality ($F(1,14) = 5.95$, $p < 0.05$), a significant Stimulus x Legality interaction ($F(1,14) = 8.26$, $p < 0.05$), and a tendency for Electrode x Stimulus x Legality interaction ($F(5,70) = 2.75$, G-G corrected $p = 0.051$). The Tukey HSD test performed on the Electrode x Stimulus x Legality interaction showed that the illegal stimulus elicited a significantly larger negativity in the deviant position than in the standard position (on all electrodes except for C3, $p < 0.01$), but there was no such difference for the legal stimulus (see Figure 5 for the illustration of this effect on the Fz electrode). Moreover, the illegal pseudoword in the deviant position elicited a significantly larger negativity than the legal one (on all electrodes except for C3 and Cz, $p < 0.01$), but there was no such difference between the pseudowords in the standard position.

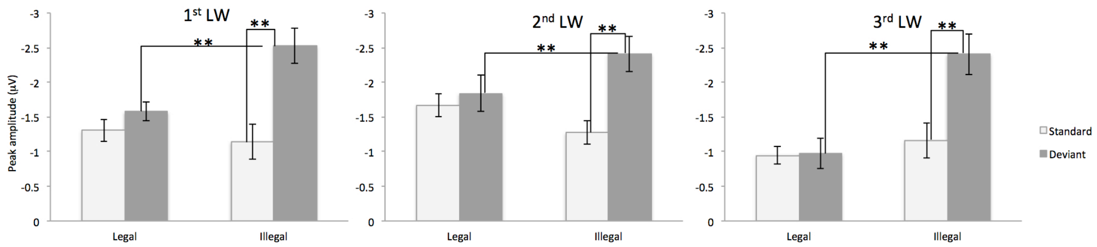

Figure 5. Illustration of the Stimulus x Legality interaction on the Fz electrode. LW: latency window; **: $p < 0.01$

In the second latency window (420-520 ms), we found a significant main effect of Electrode ($F(5,70) = 27.02$, $p < 0.01$) and Stimulus ($F(1,14) = 17.72$, $p < 0.01$), a significant Stimulus x Legality ($F(1,14) = 8.43$, $p < 0.05$), and Electrode x Stimulus x Legality interaction ($F(5,70) = 5.16$, G-G corrected $p < 0.05$). The Tukey HSD test performed on the Electrode x Stimulus x Legality interaction demonstrated that the illegal stimulus elicited a larger negativity in the deviant position compared to the standard position (only on the frontal electrodes, $p < 0.01$) but the legal stimulus did not. At the same time there were only sporadic differences between the illegal and legal stimuli in the deviant (Fz only $p < 0.01$), and standard (F3 only, $p < 0.01$) positions.

In the third latency window the repeated measures ANOVA revealed a significant main effect of Electrode ($F(5,70) = 21.09$, $p < 0.01$), Stimulus ($F(1,14) = 13.08$, $p < 0.01$) and Legality ($F(1,14) = 9.39$, $p < 0.01$), a significant Stimulus x Legality ($F(1,14) = 14.17$, $p < 0.01$), Electrode x Stimulus x Legality interaction ($F(5,70) = 4.31$, G-G corrected $p <0.01$). The Tukey HSD test calculated on the Electrode x Stimulus x Legality interaction showed similar results as in the first latency window: the illegal, but not the legal stimulus elicited a significantly larger negativity in the deviant than in the standard position (on all electrodes except for C3, $p < 0.01$), and the illegal stimulus elicited a significantly larger negativity than the legal stimulus when it was in the deviant, but not in the standard position (on all electrodes except for C3, $p < 0.01$)

To summarize the above results, the pseudoword with the illegal stress pattern elicited a significantly larger negativity in all three MMN latency windows when it was in the deviant compared to the standard position. At the same time the legal pseudoword did not elicit significantly different ERP amplitudes in the standard and deviant positions in the MMN latency windows, but ERPs in the early latency windows did differ, which was not the case for the illegal pseudoword. Furthermore, the ERPs elicited by illegal pseudoword in the deviant position were larger in all three latency windows compared to the ERPs elicited by the legal pseudoword in the deviant position. Based on that, we can conclude that only the illegal pseudoword elicited significant MMN components in the deviant position, but the legal pseudoword did not.

Discussion

In the present study we investigated the processing of stress pattern change in the case of pseudowords with no lexical representation, having either a legal (stress on the first syllable), or illegal (stress on the second syllable) stress pattern. We manipulated which stress pattern served as standard and deviant. We found that the pseudoword with the illegal stress pattern elicited three consecutive negative components. This was confirmed in the analysis comparing ERPs elicited by physically identical stimulus in the standard and deviant situations (for a similar approach see Jacobsen et al., 2003). Of the three negativities, the first and third were considered as MMNs reflecting the processing of stress changes related to the first and second syllables of the pseudowords: the detection of the lack of stress on the first syllable of the deviant, and the detection of the additional stress on the second syllable. These results are similar to those found in previous experiments (Honbolygó et al., 2004; Ylinen et al., 2009). We also found a negative component between the two MMNs, which was smaller than the MMNs, and was significant only on two frontal electrodes. The interpretation of this component is not clear: it may reflect the processing of the onset of the second syllable, but at the same time it may reflect a deviance detection process because it was larger for the deviant stimulus. Further research is needed to clarify the nature of this negativity.

It could be debated whether the second MMN appearing in the legal condition of the present study and in previous studies (Honbolygó et al., 2004; Ylinen et al., 2009) is an MMN or a late discriminative negativity (LDN). LDN has been shown to appear for complex speech sounds and reflect higher-order, cognitive processing of stimuli (Ceponiene et al., 2004), but has been found mostly in the case of children (Korpilahti et al., 2001; Cheour et al., 2001). At the same time, Horváth et al. (2009) obtained LDN in adults, and attributed its appearance to the complexity of the paradigm, and to the richness of the stimuli. However, since the functional interpretation of LDN is not yet clear, the results are better explained by assuming two consecutive MMNs related to the processing of stress changes on the two syllables of pseudowords.

In the reversed condition, when the deviant had a legal stress pattern, and the standard had an illegal one, three negative ERP components were elicited by the deviant, but only the first and the second were significantly different from the ERPs elicited by the standard. However, the analysis comparing ERPs elicited by physically identical stimuli in the standard and deviant situations showed that none of these components were significant, that is the pseudoword with the regular stress pattern did not elicit any significant MMN when compared to its physically identical counterpart in the standard condition, instead of comparing it to the illegal standard in the same condition. Thus the negativities visible on the difference wave in Figure 3 are probably due to the processing of the physical differences between the stimuli. This result contradicts our expectations about the ERP results in the reversed condition, which was based on the results of Ylinen et al. (2009) who found a single MMN to the word with the familiar stress pattern. Although a broad negativity can be discerned on Figure 4 to the legal deviant pseudoword, it did not reach statistical significance. The reason for that could be that Ylinen et al. (2009) did not calculate the difference between physically identical stimuli, as we did, and in our case the more careful comparison helped to eliminate the effect of the processing of physical differences between standard and deviant.

Besides of MMN responses, we have found that early ERPs (P2-N2), especially the early negativity showed an exquisite sensitivity to legality and stimulus effects: in the case of standards, the N2 elicited by the legal and illegal words was always the same, but a larger N2 was elicited by the illegal deviant than the legal deviant (see Figure 4). The results found by Cunillera et al. (2006) in a stress discrimination paradigm would provide a possible explanation of the N2 effect found, if the direction of the effect was not the opposite of that found in our study. Cunillera et al. (2006) found a larger N2 component to the stressed non-words than to the unstressed ones, whereas in our study the pseudoword with unstressed first syllable was the one that elicited a larger N2. These results indicate that the N2 component is sensitive both to the frequency of the stimulus, as shown by the effects occurring only on the deviant elicited responses, and to the legality of the stimulus. This effect cannot be associated simply with the acoustic differences, as it was larger for the illegal deviant of less salient acoustic features (i.e., features related to the lack of stress).

Our results thus show that different ERP patterns were elicited in the legal and illegal conditions, confirming the hypothesis that the perception of stress pattern change is based on long-term representations. The MMNs appearing in the two conditions were in fact determined by the legality of the stress pattern and not just by the role of the stimulus in the oddball sequence, as only the illegal pseudoword in the deviant position elicited the MMNs, but the legal pseudoword in the deviant position did not. We suggest that the illegal pseudoword mismatched both the short-term and long-term traces, and therefore elicited the MMNs. However, the legal pseudoword did not elicit the MMN because it did not mismatch the long-term stress representation, only the short-term trace built up in the oddball sequence. At the same time the short-term trace established by the illegal standard stimulus was insufficient to elicit a significant MMN, probably because it itself mismatched the long-term representation. This result is further strengthened by the early N2 having larger amplitude for the illegal deviant pseudoword, which could be interpreted as a first detection of the deviation of the stimulus from the long-term representation. Therefore we assume that two different processes give rise to the ERP changes found; the first one shown by the N2 effect and a second one shown by the first MMN, both associated with the template violation and a later one elicited by the deviation from the trace built up by the standards. However, further studies are needed to clarify the role of the two processes assumed to contribute to the early ERP effects found in the automatic detection of legal and illegal stress patterns.

Our results can be interpreted with reference to lexical processing theories put forward by Pulvermüller et al. (2001) and Jacobsen et al. (2004). Pulvermüller et al. (2001) suggest the lexical trace hypothesis that assumes that the MMN depends on the familiarity of the deviant (whether it is a word or a pseudoword), but the lexical status of the standard does not influence that. Words elicit larger MMN irrespective of the lexical status of the standard, because they activate lexical representations in addition to acoustic and phonetic representations. Jacobsen et al. (2004) refer to the familiar context hypothesis, and suggest that the representation of familiar sounds of any kind is more elaborate, sharp, and rich than the representation of unfamiliar sounds. The familiar context hypothesis assumes that when the standard sounds of an oddball sequence are familiar, then this creates a regularity representation or context that makes the processing of deviant features more elaborate. In fact it is the familiar context that is important, because it makes it possible to better encode the deviant. The authors found evidence for a larger MMN when the standards were familiar words as compared to when standards were unfamiliar words. Our results are in line with the familiar context hypothesis, since the MMNs obtained were determined by the status of the standard stimuli: MMNs appeared when the deviant was presented in a familiar context, i.e., the standards had a legal stress pattern, and no MMN appeared when the deviant was presented in an unfamiliar context, i.e., standards had an illegal stress pattern.

According to the familiar context hypothesis, the familiar stimuli, in our case the pseudowords with a legal stress pattern, have a more elaborate and rich representation. The origin of the familiarity is supposed to be learning: it can be either long-term, as in the case of speech sounds (Näätänen et al., 1997; Winkler et al., 1999), or short term (Atienza and Cantero, 2001; Näätänen et al., 1993). To account for the familiarity of the legal stress patterns we proposed in the Introduction that stress might be represented in the form of long-term, pre-lexical, language-specific representation, called stress templates. Stress templates might play an important role in speech perception by helping the categorization of syllables as stressed or unstressed. As we have mentioned in the Introduction, the acoustic information related to stress in a syllable is often unreliable. However, the speech processing system needs to detect stressed and unstressed syllables even under suboptimal (or especially in suboptimal, see Mattys et al., 2005) conditions. According to the metrical segmentation strategy, the importance of the detection of stressed syllables is that these help the segmentation of speech stream into words, serving as starting points in the process of lexical access (Cutler and Norris, 1988; Mattys and Samuel, 1997; Mattys et al., 2005). Grosjean and Gee (1987) explicitly suggest the existence of pre-lexical, intermediary representations containing the description of phonetical segments as stressed or unstressed. Our proposal is that stress templates might play a role in the construction of these intermediary representations, or that stress templates might be the actual intermediary representations.

Our ERP data might be explained with reference to the suggested stress templates. We propose that the processing of stress information relies on a complex matching process, where the perceptual system compares the input both with the previous short-term acoustic traces built up by the standard stimulus, and with the long-term stress template. In the case of the deviant illegal stress pattern, the system detects already at the first syllable that the stimulus do not match the features specified in the template, and this gives rise to a neural process reflected by the MMN component, and probably by the modulation of the early N2 component. At the same time, no MMN appears if the standard mismatches the template, as in the case of the illegal condition.

To conclude, our ERP results demonstrate that during the processing of word stress, pseudowords with an illegal stress pattern elicited two consecutive MMN components, while pseudowords with a legal stress pattern in a deviant position did not elicit MMN. We suggest that these results can be explained with reference to long-term, pre-lexical, language-specific representations of stress information that we called stress templates. The role of stress templates is to help the categorization of speech information into stressed and unstressed syllables, which is an important process in accessing the mental lexicon.

Interestingly enough, our results confirm the suggestion of Iván Fónagy, who proposed decades ago that "stress is a psychological category, the experience of which is the result of the quick and unconscious analysis of the sounds" (Fónagy, 1958, p. 23.): indeed, we have found that the processing of stress is based on abstract phonological templates, and that during the processing of speech, stress is analyzed quickly and without conscious attention by means of a neural mechanism generating the MMN component.

Acknowledgements

The authors wish to thank Gabriella Baliga for her help in the data acquisition, and Orsolya Kolozsvári for her help in preparing the figures. We thank the anonymous reviewers for their helpful comments on the earlier version of this paper. This study was supported by the Hungarian Scientific Research Fund (OTKA 47381 and OTKA-NK 72465, Principal Investigator: Valéria Csépe).

References

Atienza, M., Cantero, J. L., 2001. Complex sound processing during human REM sleep by recovering information from long-term memory as revealed by the mismatch negativity MMN. Brain Res 901, 151-60.

Boersma, P., Weenink, D., 2007. Praat: doing phonetics by computer. Retrieved from http://www.praat.org/.

Cheour, M., Korpilahti, P., Martynova, O., Lang, A. H., 2001. Mismatch negativity and late discriminative negativity in investigating speech perception and learning in children and infants. Audiol Neurootol 61, 2-11.

Ceponiene, R., Lepistö, T., Soininen, M., Aronen, E., Alku, P., Näätänen, R., 2004. Event-related potentials associated with sound discrimination versus novelty detection in children. Psychophysiology 41, 130-141.

Colombo, L.,1992. Lexical stress effect and its interaction with frequency in word pronunciation. J Exp Psychol Human, 18, 987-1003.

Cunillera, T., Toro, J., Sebastián-Gallés, N., Rodrìguez-Fornells, A., 2006. The effects of stress and statistical cues on continuous speech segmentation: An event-related brain potential study. Brain Res 1123, 168-178.

Cutler, A., Norris, D., 1988. The Role of Strong Syllables in Segmentation for Lexical Access. J Exp Psychol Human 141, 113-121.

Delorme, A., Sejnowski, T., Makeig, S., 2007. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. NeuroImage 34, 1443-1449.

Friederici, A. D., Friedrich, M., Christophe, A., 2007. Brain responses in 4-month-old infants are already language specific. Curr Biol 17, 1208-1211.

Friedrich, M., Herold, B., Friederici, A. D., 2009. ERP correlates of processing native and non-native language word stress in infants with different language outcomes. Cortex 455, 662-76.
Fónagy, I., 1958. A hangsúlyról [On stress]. Nyelvtudományi Értekezések, 18.
Greenhouse, S. W., Geisser, S., 1959. On methods in the analysis of profile data. Psychometrika 242, 95-112.
Grosjean, F., Gee, J. P., 1987. Prosodic structure and spoken-word recognition. Cognition 25, 135-155.
Honbolygó, F., Csépe, V., Ragó, A., 2004. Suprasegmental speech cues are automatically processed by the human brain: a Mismatch Negativity study. Neurosci Lett 3631, 84-88.
Horváth, J., Roeber, U., Schröger, E., 2009. The utility of brief, spectrally rich, dynamic sounds in the passive oddball paradigm. Neurosci Lett 461, 262–265.
Jacobsen T, Schröger E. 2003. Measuring duration mismatch negativity. Clin Neurophysiol. 114, 1133-1143.
Jacobsen, T., Horváth, J., Schröger, E., Lattner, S., Widmann, A., Winkler, I., 2004. Pre-attentive auditory processing of lexicality. Brain Lang 881, 54-67.
Kager, R., 2007. Feet and metrical stress. In: de Lacy, P. Ed, The Cambridge Handbook of Phonology, Cambridge University Press, Cambridge, pp. 195-228.
Korpilahti, P., Krause, C. M., Holopainen, I., Lang, A. H., 2001. Early and late mismatch negativity elicited by words and speech-like stimuli in children. Brain Lang 76, 332–339.
Mattys, S. L., Samuel, A. G., 1997. How Lexical Stress Affects Speech Segmentation and Interactivity: Evidence from the Migration Paradigm. J Mem Lang 361, 87-116.
Mattys, S. L., White, L., Melhorn, J. F., 2005. Integration of Multiple Speech Segmentation Cues: A Hierarchical Framework. J Exp Psychol Gen 1344, 477-500.
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 3801, 1-21.
Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., Allik, J., Sinkkonen, J., Alho, K., 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. Nature 385, 432-434.

Näätänen, R., Paavilainen, P., Rinne, T., Alho, K., 2007. The mismatch negativity MMN in basic research of central auditory processing: a review. Clin Neurophysiol 118, 2544-2590.
Näätänen, R., Schröger, E., Karakas, S., Tervaniemi, M., Paavilainen, P., 1993. Development of a memory trace for a complex sound in the human brain. Neuroreport 45, 503-506.
Pulvermüller, F., Kujala, T., Shtyrov, Y., Simola, J., Tiitinen, H., Alku, P., Alho, K., Martinkauppi, S., Ilmoniemi, R., Näätänen, R., 2001. Memory traces for words as revealed by the mismatch negativity. NeuroImage 143, 607-616.
Pulvermüller, F., Shtyrov, Y., 2006. Language outside the focus of attention: the mismatch negativity as a tool for studying higher cognitive processes. Prog Neurobiol 791, 49-71.
Shtyrov, Y., Pulvermüller, F., 2002. Neurophysiological evidence of memory traces for words in the human brain. Neuroreport 134, 521-5235.
Siptár, P., Törkenczy, M., 2007. The phonology of Hungarian. Oxford University Press.
Varga, L., 2002. Intonation and stress: evidence from Hungarian. Palgrave Macmillan, New York.
Weber, C., Hahne, A., Friedrich, M., Friederici, A., 2004. Discrimination of word stress in early infant perception: Electrophysiological evidence. Cognitive Brain Res 182, 149-161.
Winkler, I., Denham, S. L., Nelken, I., 2009. Modeling the auditory scene: predictive regularity representations and perceptual objects. Trends Cogn Sci 1312, 532-540.
Winkler, I., Lehtokoski, A., Alku, P., Vainio, M., Czigler, I., Csépe, V., Aaltonen, O., Raimo, I., Alho, K., Lang, H., Iivonen, A., Näätänen, R., 1999. Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations. Cognitive Brain Res 73, 357-69.
Ylinen, S., Strelnikov, K., Huotilainen, M., Näätänen, R., 2009. Effects of prosodic familiarity on the automatic processing of words in the human brain. Int J Psychophysiol 733, 362-368.