Word stress representations are language specific: evidence from event-related brain potentials

Ferenc Honbolygó[1,3], Andrea Kóbor[1], Borbála German[1,2], Valéria Csépe[1,2]

[1] Brain Imaging Centre, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest, Hungary

[2] Department of Cognitive Science, Budapest University of Technology and Economics, Budapest, Hungary

[3] Institute of Psychology, ELTE Eötvös Loránd University, Budapest, Hungary

Corresponding author: Ferenc Honbolygó

Brain Imaging Centre, Research Centre for Natural Sciences, Hungarian Academy of Sciences; Magyar tudósok körútja 2., H–1117, Budapest, Hungary.

E-mail: honbolygo.ferenc@ttk.mta.hu; Phone: + 36 13826615

Abstract

Previous studies investigating the neurocognitive background of word stress processing demonstrated that words differing in their stress pattern are discriminated automatically and unconsciously as shown by the emergence of the Mismatch Negativity (MMN) event-related brain potential (ERP) component. However, it has remained unclear if the processing of word stress relies on language specific representations.

In the present study, Hungarian participants heard disyllabic pseudowords stressed either on the first (legal stress) or on the second (illegal stress) syllable, pronounced either by a Hungarian or a German speaker. Results obtained using a data-driven ERP analysis approach showed that all pseudowords in the deviant position elicited an Early Differentiating Negativity (EDN) and an MMN component, except for the Hungarian pseudowords stressed on the first syllable. This suggests that Hungarian listeners did not process the native legal stress pattern as deviant, implying that the processing was based on language specific long-term memory traces.

Keywords: ERP, MMN, word stress, long-term traces

Introduction

Understanding speech requires the simultaneous processing of segmental and suprasegmental (prosodic) information, suggested to be based on parallel and independent processing mechanisms (Poeppel, 2014), and possibly relying on separate neural pathways (Bornkessel-Schlesewsky & Schlesewsky, 2013). Among the available prosodic information, word stress is a relative emphasis given to certain syllables within words or to certain words in sentences (for review, see Kager, 2007). Word stress can emphasize or separate certain parts of the speech stream; thus, it potentially contributes to the segmentation of continuous speech into words (Cutler & Norris, 1988). However, languages differ in several aspects regarding word stress: in the position of the stressed syllable within multisyllabic words (initial, final, penultimate, etc.); in the variability of the stressed syllable's position (free or fixed); and whether stress can distinguish lexical meaning (contrastive or non-contrastive). Consequently, the language processing system of individuals has to adapt to the specific features of word stress of a given language. In the present study, we investigated how much stress related representations can be considered as language specific, and whether we find any evidence of language specific processing at the neurocognitive level. Word stress provides one of the essential first steps in understanding and learning a native language in infancy by aiding word segmentation (Gervain & Mehler, 2010). Very young infants have been shown to demonstrate language specific discrimination of stress patterns (Friederici, Friedrich, & Christophe, 2007; Ragó, Honbolygó, Róna, Beke, & Csépe, 2014; Weber, Hahne, Friedrich, & Friederici, 2004). In adulthood, when learning a second language (L2), the processing of foreign word stress might impose a difficulty, as demonstrated by studies showing that speakers of a fixed-stress native language have a

reduced ability to exploit free stress for recognizing words in a second language (Dupoux, Peperkamp, & Sebastián-Gallés, 2010; Tremblay, 2008).

Previous studies investigating the neurocognitive basis of word stress processing studied mostly the Mismatch Negativity (MMN) event-related brain potential (ERP) component. The MMN is an auditory ERP component with a negative polarity and a fronto-central voltage maximum, and has been an exceptionally useful tool in studying linguistic processing (for review, see Näätänen, Paavilainen, Rinne, & Alho, 2007). The MMN is usually elicited in passive oddball paradigms where frequently repeated standard stimuli are interspersed by rarely repeated deviant stimuli differing from the standard in some discriminable features. The MMN appears 100-250 ms after the onset of the change and can be elicited in the absence of participants' attention. The MMN is currently interpreted as a brain electrical correlate of the mainly pre-attentive detection of violation of simple or complex regularities (Winkler, Denham, & Nelken, 2009).

Only a few studies have applied the MMN paradigm to study the processing of word stress. Weber et al. (2004) found that German adults showed an MMN to the word with stress on the first syllable as well as to the word with stress on the second syllable. Ylinen et al. (2009) showed that Finnish words and pseudowords with unfamiliar (stress on the second syllable) versus familiar (stress on the first syllable) word stress patterns elicited two MMNs related to the first and second syllables of utterances. At the same time, words with familiar stress pattern elicited a single MMN in the earlier time windows, indicating a stress familiarity effect. Zora, Heldner, & Schwarz (2016) found that in Turkish words and pseudowords with different stress patterns elicited the MMN, which was different according to the prosodic features contributing to stress, f0 being the most salient. The same authors (Zora, Riad, Schwarz, & Heldner, 2016) also demonstrated the appearance of MMN in Swedish words to

changes in stress and a specific enhancement of MMN to phonologically versus lexically stressed words. Finally, in a study comparing the processing of duration-related stress in speech and music in English (Peter, Mcarthur, & Thompson, 2012), the authors found that in the case of speech, only the stress on the first syllable condition elicited an MMN, while in the case of music stimuli, both long-short and short-long patterns (the musical equivalent of stress on the first and stress on the second syllable) elicited an MMN, indicating a fundamental difference between speech and non-speech related processing.

A series of studies investigated word stress processing in Hungarian, a fixed-stress language in which stress falls on the first syllable of disyllabic words without exception. First, it was demonstrated that a word with stress on the second syllable elicited two MMN components when contrasted with a word with stress on the first syllable (Honbolygó, Csépe, & Ragó, 2004). In a subsequent study (Honbolygó and Csépe, 2013), it has also been shown that pseudowords with stress on the second syllable elicited two consecutive MMN components, while pseudowords with a familiar stress pattern in a deviant position did not elicit an MMN, suggesting that stress processing is modulated by top-down processes and supporting the importance of familiarity in word stress processing. This study was also replicated using meaningful words (Garami, Ragó, Honbolygó, & Csépe, 2017), demonstrating that lexicality facilitates the processing of familiar stress patterns but not the unfamiliar ones. These studies provided further evidence that Hungarian speakers rely on long-term representations of the fixed Hungarian stress pattern, which appear to contain a description of the expected word stress pattern that is characteristic for a given language at the phonological level. However, the language specific nature of these representations has not been directly investigated so far; therefore, the present study specifically aimed to clarify this aspect.

The MMN paradigm has also been used to investigate language-specific processing of linguistic representations. It has been demonstrated in several studies (Dehaene-Lambertz, 1997; Kraus & Cheour, 2000; Näätänen et al., 1997; Winkler et al., 1999) that native language phonemes elicit enhanced MMN as compared to non-native phonemes, indicating the modulation of speech processing mechanisms by the exposure to native language. Chandrasekaran, Krishnan, & Gandour (2007) used the MMN paradigm to study the effect of language experience on the processing of pitch contour in Mandarin Chinese in a cross-linguistic study. According to the results, native Chinese speakers showed a larger MMN component to highly dissimilar tones than non-native (English) speakers, demonstrating that the MMN is modulated by the long-term experience with the native language.

In the present study, we also used the MMN paradigm to investigate the language specific nature of word stress representations in native speakers of Hungarian. Hungarian is an excellent target language for studying word stress representations because the stress pattern of disyllabic word is highly regular: stress always falls on the first syllable. Therefore, any changes in the stress pattern is considered as violating the regularity, leading to illegal stress patterns. In this study, we used naturally produced, meaningless pseudowords having stress either on the first or the second syllable. Moreover, we created a native and a foreign (German) version of both types of pseudowords in order to study the ERP correlates of non-native stress processing. We hypothesized that pseudowords stressed on the second syllable (illegal stress pattern in Hungarian) would elicit the MMN component irrespective of the language. However, we expected that pseudowords stressed on the first syllable (legal stress pattern in Hungarian) would elicit the MMN component only when the non-native version is presented, because listeners would not detect the irregularity in the Hungarian version.

Methods

Participants

Thirty-two healthy young adults took part in the experiment. Two of them were excluded by reason of excessive artifacts (see EEG Recording and Pre-Processing). Therefore, thirty participants remained in the final sample (19 females). The participants, aged between 18 and 25 years (M_{Age} = 21.7 years, *SD* = 1.6 years), were native speakers of Hungarian who have not studied German as L2 according to a questionnaire screening their linguistic background (for details, see Kóbor et al., 2018). All of them were right-handed according to the Edinburgh Handedness Inventory revised version (Dragovic, 2004) (M_{LQ} = 90.4, *SD* = 12.9[1]). Most of the participants were undergraduate students enrolled in different universities in Budapest ($M_{\text{Years of education}}$ = 15.1 years, *SD* = 1.6 years). They had normal or corrected-to-normal vision and normal hearing level according to the audiometry measurement. Prior to their inclusion in the study, participants provided informed consent to the procedure as approved by the Ethical Review Committee for Research in Psychology, Hungary. The study was conducted in accordance with the Declaration of Helsinki. Participants received payment (3300 HUF, ca. 10 Euros) or course credit for taking part in the experiment.

Stimuli

Four pseudowords were used as stimuli: two spoken by a Hungarian native speaker and two spoken by a German native speaker. The same reiterative pseudoword ('bébé'; note that in Hungarian, the acute accent denotes a different phoneme than 'e' and not the stressed syllable) was recorded by each speaker with two different stress patterns: one with stress on

[1] LQ = Laterality Quotient; LQ = -100 means complete left-handedness, LQ = 100 means complete right-handedness.

the first syllable (trochaic stress pattern, henceforth referred to as S1), and one with stress on the second syllable (iambic stress pattern, henceforth referred to as S2). Speakers were asked to produce the pseudowords as naturally as they could, and no further software editing was done on the stimuli, except for equalizing the overall loudness (RMS equalization) and applying a 10 ms long onset and offset ramp. The acoustical properties of stimuli are summarized in Table 1.

INSERT TABLE 1 AROUND HERE

Procedure

Stimuli were presented in a *passive oddball paradigm* with a stimulus onset asynchrony (SOA) varying randomly between 850 and 1050 ms. The duration of the S1 stimulus was 746 ms in Hungarian and 573 ms in German, and the duration of the S2 stimulus was 747 ms in Hungarian and 574 ms in German. In our design, we varied language (Hungarian, German), stress position (S1, S2) and role (standard, deviant) yielding *eight different conditions*: the S1 and S2 as a standard stimulus and the S1 and S2 as a deviant stimulus in Hungarian and German, respectively. Stimuli were presented in *four experimental blocks* according to the following pairings: (1) Hungarian S1 standard with Hungarian S2 deviant; (2) Hungarian S2 standard with Hungarian S1 deviant; (2) German S1 standard with German S2 deviant, (3) German S2 standard with German S1 deviant. Participants were presented 666 stimuli in each block with 566 standard stimuli and 100 deviant stimuli (i.e., the probability of the deviant was 15%). Each block began with at least fifteen standard stimuli, and then deviants were separated by at least two standard stimuli. We created a different pseudorandom sequence of stimuli for each block, and within each block, the same sequence of stimuli was

used for all participants. The order of the four experimental blocks was counterbalanced across participants. During the EEG experiment, participants were seated in a comfortable chair in an acoustically shielded, dimly lit room. They were instructed to ignore the stimulus sequence presented via headphones (Sennheiser PX 200), while they were watching a silent movie.

After the EEG experiment and the removal of the electrode net, a behavioral discrimination task was administered to check whether the variations in stress position were distinguishable within a language. Namely, a *same-different paradigm* with two blocks was used: one block tested the Hungarian stimuli, another tested the German stimuli. Stimuli were presented pairwise with a SOA of 1000 ms. Forty stimulus pairs were presented in each block: The four possible combinations of stimulus pairs (S1-S1, S2-S2, S1-S2, S2-S1) were tested ten times. The order of blocks was counterbalanced across participants. Participants were required to decide as fast and accurate as possible whether the two stimuli presented as a stimulus pair were the same or they were different. They had unlimited time to initiate a response for each stimulus pair. The next stimulus pair was presented after the participant gave behavioral response. We used a Cedrus RB-740 response pad (Cedrus Corporation, San Pedro, CA) to record participants' responses in this task. Assignment of response keys to same/different responses was counterbalanced across participants. The entire experimental procedure lasted about 2 hours, including the application of the electrode net. Both the EEG and the behavioral experiment were written in Presentation (v. 17.0, Neurobehavioral Systems).

EEG Recording and Pre-processing

The continuous EEG activity was recorded using the Electrical Geodesics system (GES 400; Electrical Geodesics, Inc., Eugene, OR) and Net Station 4.5.1 software. We used a 128-channel HydroCel Geodesic Sensor Net with saline electrolyte solution. Electrode Cz was used as a reference and a sampling rate of 500 Hz was applied.

Data was analyzed using BrainVision Analyzer software (Brain Products GmbH, Munich, Germany) and Matlab 8.5 (MathWorks Inc.). Spline interpolation of bad electrodes was performed if necessary. Zero – three (*M* = 0.61) channels per participant were interpolated. As the first step of pre-processing, EEG data was band-pass filtered offline between 0.3 – 30 Hz (48 dB/oct) and notch filtered at 50 Hz to remove additional electrical noise. Second, we corrected horizontal and vertical eye-movement artifacts and heartbeats with independent component analysis (ICA; Delorme, Sejnowski, & Makeig, 2007). Using ICA, 1-6 ICA components (*M* = 2.93) were removed, then, the channel-based EEG data were recomposed. Third, EEG data was re-referenced to the average activity of all electrodes. After that, the continuous EEG was segmented into epochs. Epochs extended from -100 to 1000 ms relative to the onset of the standard and deviant stimuli, respectively, in each condition. To remove artifacts still present in the data after ICA corrections, we used an automatic artifact rejection algorithm implemented in BrainVision Analyzer software. This algorithm rejected epochs where the activity exceeded +/- 100 µV at any of the electrode sites. Only those participant's data were included in further analysis where the percentage of rejected epochs was below 30% in any of the conditions. Accordingly, two participants' data were excluded. The mean number of retained epochs collapsed across all conditions was 95 (*SD* = 5.4; range: 72 – 100). After artifact rejection, artifact-free epochs were baseline corrected based on the mean activity from -100 to 0 ms. Finally, an equal number of epochs were averaged for the

standard and deviant stimuli in each condition by analyzing only the standards preceding the deviants.

Data/ERP Analysis

ERP analysis involved two phases. First, in order to identify the putative time windows where the effect of stress position on the ERPs varied as a function of language, a permutation based non-parametric topographic analysis of variance (TANOVA, Murray, Brunet, & Michel, 2008) was conducted in Ragu (Koenig, Kottlow, Stein, & Melie-García, 2011) on the individual differences waves. For this purpose, four difference waves were created by subtracting the waveform for the standard stimuli from the waveform for the deviant stimuli for each participant as follows: (1) Hungarian S1 deviant minus Hungarian S1 standard; (2) Hungarian S2 deviant minus Hungarian S2 standard; (3) German S1 deviant minus German S1 standard; (4) German S2 deviant minus German S2 standard. This approach enabled to measure genuine MMN effects as we compared physically identical stimuli within each language differing only in their role (standard vs. deviant) in the experimental blocks (see also Honbolygó & Csépe, 2013; Jacobsen & Schröger, 2003).

INSERT FIGURE 1 AROUND HERE

To assess topographic dissimilarity (TD) (Koenig & Melie-Garcia, 2009; Wirth et al., 2008), a single measure of the distance between two electric field topographies, we used TANOVA

with Language (Hungarian vs. German) and Stress Position (S1 vs. S2) as within-subjects factors. We performed 1000 randomization runs and applied a 5% significance threshold on the individual averages (difference waves) during the entire epoch length (no normalization was used). Regarding the main question of this study, we focused on the Language * Stress Position interaction effect. A significant interaction would show that the ERPs elicited by stress on the first or the second syllable varies as a function of language. The interaction was significant within the following time intervals: (1) 116-286 ms; (2) 378-444 ms; and (3) 530-818 ms (Fig. 2).

INSERT FIGURE 2 AROUND HERE

Results of the TANOVA were used in the second step of ERP analysis. We first averaged the activity of electrodes 5, 6, 11, 12 (see Fig. 3) into a fronto-central electrode pool. Second, grand average ERP waveforms were calculated for each of the eight conditions (Hungarian S1 standard and deviant, Hungarian S2 standard and deviant, German S1 standard and deviant, German S2 standard and deviant, see Fig. 1). Third, we selected those three time windows where the Language * Stress Position interaction was significant in the TANOVA. This way, following a data-driven approach (see e.g., Brooks, Zoumpoulaki, & Bowman, 2017), we minimized the chance of selecting time windows arbitrarily (i.e., where the ERP components' amplitude seemed to be maximal or the difference between particular conditions seemed to be maximal) or by merely following previous literature. As a final step, therefore, we quantified the ERP components as the mean amplitude within the three time windows specified above in each of the eight conditions. To maintain the objective

approach, we refer to the activity measured in the time windows simply as ERPs and identify the components only in the discussion part.

INSERT FIGURE 3 AROUND HERE

The mean amplitude of the ERPs was then entered into three-way repeated measures ANOVAs with Language (Hungarian vs. German), Stress Position (S1 vs. S2), and Role (standard vs. deviant) as within-subjects factors. In all ANOVAs, partial eta-squared (η_p^2) is reported as the measure of effect size. To control for Type I error, we used Tukey HSD tests for pair-wise comparisons.

Results

Behavioral results

One participant's data in the same-different paradigm are missing because of technical reasons. Thus, behavioral data were analyzed in 29 participants. As Hungarian and German stimuli differed in duration, we did not compare the RT of responses across the different contrasts. Instead, accuracy, defined as the percentage of correct responses, was used as a measure of performance. Responses slower that 2000 ms were excluded from further analysis, thus, accuracy was calculated without these trials. We found ceiling effect on this task as the mean accuracy for the same pairs was 99.48% (*SD* = 1.55%) in the Hungarian block and 98.79% (*SD* = 2.18%) in the German block. For the different pairs, this was 98.45% (*SD* = 3.02%) in the Hungarian block and 99.12% (*SD* = 2.38%) in the German block. Therefore, participants were able to discriminate pseudowords with different stress pattern both in Hungarian and German.

ERP results

Grand average ERP waveforms to the standard and deviant stimuli split by language and stress position are presented in Fig. 1. Results of the following ANOVAs are summarized in Table 2.

INSERT TABLE 2 AROUND HERE

In the first latency window (116-286 ms), the main effects of Language, $F(1, 29) = 65.83$, $p < .001$, $\eta_p^2 = .694$, Stress Position, $F(1, 29) = 7.93$, $p = .009$, $\eta_p^2 = .215$, and Role, $F(1, 29) = 57.54$, $p < .001$, $\eta_p^2 = .665$, were significant on the mean amplitude of the ERPs. The two-way interactions of Language * Role, $F(1, 29) = 10.82$, $p = .003$, $\eta_p^2 = .272$, and Stress Position * Role, $F(1, 29) = 5.20$, $p = .030$, $\eta_p^2 = .152$, were also significant. These effects were qualified by the three-way interaction of Language * Stress Position * Role, $F(1, 29) = 9.28$, $p = .005$, $\eta_p^2 = .242$. Pair-wise tests showed that the mean amplitude of the ERPs was more negative for the deviant role than for the standard for Hungarian S2 ($p = .002$), German S1 ($p < .001$), and German S2 stimuli ($p = .001$) but not for the Hungarian S1 stimuli ($p = .999$) (see Fig4A). This indicates that no Role effect was elicited for the Hungarian S1 stimuli in this early latency window.

INSERT FIGURE 4 AROUND HERE

In the second latency window (378-444 ms), the main effects of Language, $F(1, 29) = 108.73$, $p < .001$, $\eta_p^2 = .789$, and Role, $F(1, 29) = 62.85$, $p < .001$, $\eta_p^2 = .684$, were significant. These

effects were qualified by the three-way interaction of Language * Stress Position * Role, $F(1, 29) = 6.57$, $p = .016$, $\eta_p^2 = .185$. Again, pair-wise tests showed that the mean amplitude of the ERPs was more negative for the deviant role than for the standard for Hungarian S2 ($p < .002$), German S1 ($p < .001$), and German S2 stimuli ($p < .001$) but not for the Hungarian S1 stimuli ($p = .267$) (see Fig4B). This indicates that no Role effect was elicited for the Hungarian S1 stimuli in this middle latency window.

In the third latency window (530-818 ms), the main effect of Role was significant, $F(1, 29) = 78.40$, $p < .001$, $\eta_p^2 = .730$, as well as the interaction of Language * Role, $F(1, 29) = 6.31$, $p = .018$, $\eta_p^2 = .179$. The mean amplitude of the ERPs was more negative for the deviant role than for the standard; however, this difference was larger for Hungarian stimuli (1.57 µV) than for German stimuli (0.93 µV) (see Fig4C).

Discussion

This study investigated for the first time the language-specific nature of word stress representations in a fixed-stress language. We studied the ERP correlates of the processing of disyllabic pseudowords having stress on the first or the second syllable produced by either a native Hungarian or a native German speaker. Using a data-driven analysis approach, we found three time windows that demonstrated stimulus-related effects.

The ERP activity in the first (116-286 ms) and second (378-444 ms) time windows showed a difference between the standard and the deviant stimuli in all conditions except for the Hungarian stress on the first syllable (S1) stimulus. This means that Hungarian listeners detected the difference between the native stress patterns when the stimulus had stress on the second syllable (S2) but not when the stress was on the first syllable. At the same time, they detected the difference between the German stress patterns in both S1 and S2. We

interpret these results as evidence for language-specific processing of stress patterns. The lack of Hungarian S1 effect means that the legal stress patterns was processed as not being deviant. Based on our earlier results (Garami et al., 2017; Honbolygó & Csépe, 2013; Honbolygó, Kolozsvári, & Csépe, 2017), we assume that the legal stress pattern is processed in relation to long-term stress representations, so called stress-templates.

The results can be interpreted in the AERS (Auditory Event Representation System) predictive coding model suggested by Winkler and Schröger (2015). According to the model, acoustic information reaching the auditory system undergoes an initial sound analysis which extract the proto-features of sounds (for speech, these are speech-related acoustic-phonetic information, such as voicing, formant transitions, fundamental frequency, intensity, etc.). This is followed by binding these separate features to unitary sensory memory representations (for speech, these are phonetic representations like phonemes, prosodic features, etc.). These representations are then compared against the predictions derived from the model of the acoustical environment. The predictive model stores representations of the current regularities of the acoustical environment but it is also affected by long-term experience, which is available for native linguistic representations (speech sounds, prosodic patterns, words, etc.). The predictive model is the basis of the mechanisms comparing incoming information with the regularity representations. If predictions from the model fail, the model needs to be corrected via an updating process, which is reflected by the MMN. In our case, the lack of ERP difference in the first and second time windows for the Hungarian S1 can be explained by assuming that the native stress pattern in the deviant position did not start the updating process because it was fully predicted by the model based on long-term stress-related representations. Thus, in the model, the long-term representations overwrote or overweighed the short-term representations based on the stimulus sequence,

in which the deviant stimuli still differed from the standard. In all other cases, the incoming deviant stimuli did start the updating process, because they violated both the short-term (standard vs. deviant) and the long-term (irregular native stress, regular and irregular non-native stress) regularities.

We suggest that the ERP activity in the first time window is not an MMN component, we rather consider it to be an Early Differentiating Negativity (EDN). The EDN sensitive to feature changes was investigated by Kuuluvainen, Alku, Makkonen, Lipsanen, and Kujala (2016) who studied the discrimination of five syllable features (consonant, vowel, vowel duration, fundamental frequency, and intensity), including both segmental and prosodic phonetic changes. They found an expressed EDN elicited by all feature changes around 100-200 ms after stimulus onset in pre-school children. Because the EDN was larger for speech than for non-speech stimuli, the authors suggested that this component was related to speech-specific processing mechanisms. They left open the possibility that the EDN could reflect a difference in the neural refractoriness resulting from an enhanced N2 component for the novel deviant features. In another study investigating stress discrimination (Cunillera, Toro, Sebastián-Gallés, & Rodríguez-Fornells, 2006), the authors found an increased N2 component for the stressed stimulus compared with the unstressed one. However, in our study, this component was larger for the unstressed syllable compared with the stressed one, which suggests that it is in fact a different kind of component, reflecting different processing mechanisms. Since data on the EDN component is still scarce, we have to be cautious about the functional interpretation. However, as our data suggest, the EDN is sensitive to both the acoustical information and the regularity of the stimuli.

We considered the ERP deflection in the second time window (378-444 ms) as a genuine MMN component. This component has been previously found in relation to stress

processing both in our studies investigating Hungarian language (Garami et al., 2017; Honbolygó & Csépe, 2013; Honbolygó, Kolozsvári, et al., 2017) and in other studies investigating Finnish (Ylinen et al., 2009), German (Weber et al., 2004) and English (Peter et al., 2012) languages. As described above, the MMN component was elicited in the Hungarian S2 and German S1 and S2 conditions but not in the Hungarian S1 condition, indicating a language-specific processing of stress related linguistic information. The functional difference between the EDN and MMN is currently unclear and requires further investigation.

Finally, the ERP activity in the third time window was regarded as a Late Discriminative Negativity (LDN) component. The LDN demonstrated a Language by Role interaction, showing a larger amplitude for the Hungarian than for the German deviant stimuli. This difference again indicates a language-related effect, but in this case, it was not modulated by the specific stress pattern of the pseudowords. The LDN is an ERP component found in oddball paradigms appearing around 300-550 ms after stimulus onset in both adults and children (Bishop, Hardiman, & Barry, 2011; Cheour, Korpilahti, Martynova, & Lang, 2001; Korpilahti, Krause, Holopainen, & Lang, 2001; Korpilahti, Lang, & Aaltonen, 1995), and it is suggested to be associated with more complex cognitive processes, such as attention (Shestakova, Huotilainen, Ceponiene, & Cheour, 2003) and long-term memory access (Zachau et al., 2005). Shestakova et al. (2003) found that children attending bilingual schools showed an increased LDN to the second language after only two months. Jakoby, Goldstein and Faust (2011) suggested that the increased LDN reflects an increasingly larger relevance of the foreign language learnt, leading to better cognitive evaluation and stimulus encoding. In our case, however, it was the native stimulus eliciting a larger LDN, which may indicate that it was the native and not the foreign stress pattern that underwent an enhanced

processing, probably because the native long-term stress representations were more elaborate than the foreign stress representations. Further studies are needed to investigate the development of foreign stress representations, and the possible effects of native language on the formation of such representations (cf. stress deafness effect, Dupoux et al., 2010; Honbolygó, Kóbor, & Csépe, 2017).

Our results are compatible with the assumptions of the neurobiological language model proposed by Bornkessel-Schlesewsky and Schlesewsky (2013). This model suggests that language is processed along a dual auditory stream network consisting of an antero-ventral or "what" stream responsible for the recognition of linguistic elements in an order-insensitive way, and a postero-dorsal or "where" stream performing predictive sequential processing of linguistic information in successively larger temporal windows (e.g., sounds, words, sentences, discourse). The postero-dorsal stream includes the prosodic segmentation of the input, and it can be argued that word stress plays a major role in the segmentation of continuous speech input into words (Cutler & Norris, 1988; Mattys, White, & Melhorn, 2005; Norris, McQueen, & Cutler, 1995; van Donselaar, Koster, & Cutler, 2005). We assume that the processing of words stress might be based on the postero-dorsal auditory stream, and the MMN effects found in the present study suggest that word stress information is processed according to predictive coding. Further brain imaging studies should be performed in order to unveil the neural networks involved in word stress processing.

Summary and conclusion

In summary, we found evidence for the language specific processing of word stress patterns indicated by the lack of early ERP (EDN and MMN) activities to the native stress pattern in the deviant position. Our results are compatible with a predictive coding model of auditory

event representations and suggest the word stress is processed predictively, based on long-term memory traces activated only by native language input. Further studies are required to extend these findings to languages with variable stress, and to show the limitations imposed by the native memory traces when learning a second language with different stress pattern.

Acknowledgment

The authors are thankful for Gabriella Baliga, Renáta Szűcs, Lívia Elek for their help in EEG recording and data collection, and Dorottya Gyarmathy and Mária Gósy for their help in recording the stimuli. The study was supported by the Hungarian Scientific Research Fund (OTKA K 119354, PI: V.Cs.), the János Bolyai Research Fellowship of the Hungarian Academy of Sciences (F.H.), and the Postdoctoral Fellowship of the Hungarian Academy of Sciences (to A. K.)

References

Bishop, D. V. M., Hardiman, M. J., & Barry, J. G. (2011). Is auditory discrimination mature by middle childhood? A study using time-frequency analysis of mismatch responses from 7 years to adulthood. *Developmental Science*, *14*(2), 402–416. https://doi.org/10.1111/j.1467-7687.2010.00990.x

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013). Reconciling time, space and function: A new dorsal-ventral stream model of sentence comprehension. *Brain and Language*, *125*(1), 60–76. https://doi.org/10.1016/j.bandl.2013.01.010

Brooks, J. L., Zoumpoulaki, A., & Bowman, H. (2017). Data-driven region-of-interest selection without inflating Type I error rate. *Psychophysiology*, *54*(1), 100–113. https://doi.org/10.1111/psyp.12682

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2007). Mismatch negativity to pitch contours is influenced by language experience. *Brain Research*, *1128*(1), 148–156. https://doi.org/10.1016/j.brainres.2006.10.064

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*, *6*(1), 2–11.

Cunillera, T., Toro, J. M., Sebastián-Gallés, N., & Rodríguez-Fornells, A. (2006). The effects of stress and statistical cues on continuous speech segmentation: An event-related brain potential study. *Brain Research*, *1123*(1), 168–178.

Cutler, A., & Norris, D. (1988). The Role of Strong Syllables in Segmentation for Lexical Access. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(1), 113–121.

Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport*, *8*(4), 919–924.

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.

Dragovic, M. (2004). Towards an improved measure of the Edinburgh Handedness Inventory: A one-factor congeneric measurement model using confirmatory factor analysis. *Laterality: Asymmetries of Body, Brain and Cognition*, *9*(4), 411–419. https://doi.org/10.1080/13576500342000248

Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2010). Limits on bilingualism revisited: stress “deafness” in simultaneous French-Spanish bilinguals. *Cognition*, *114*(2), 266–275. https://doi.org/10.1016/j.cognition.2009.10.001

Friederici, A. D., Friedrich, M., & Christophe, A. (2007). Brain Responses in 4-Month-Old Infants Are Already Language Specific. *Current Biology*, *17*(14), 1208–1211. https://doi.org/10.1016/j.cub.2007.06.011

Garami, L., Ragó, A., Honbolygó, F., & Csépe, V. (2017). Lexical influence on stress processing in a fixed-stress language. *International Journal of Psychophysiology*, *117*, 10–16. https://doi.org/10.1016/j.ijpsycho.2017.03.006

Gervain, J., & Mehler, J. (2010). Speech perception and language acquisition in the first year of life. *Annual Review of Psychology*, *61*, 191–218. https://doi.org/10.1146/annurev.psych.093008.100408

Honbolygó, F., & Csépe, V. (2013). Saliency or template? ERP evidence for long-term representation of word stress. *International Journal of Psychophysiology*, *87*(2), 165–172. https://doi.org/10.1016/j.ijpsycho.2012.12.005

Honbolygó, F., Csépe, V., & Ragó, A. (2004). Suprasegmental speech cues are automatically processed by the human brain: a mismatch negativity study. *Neuroscience Letters*, *363*(1), 84–88. https://doi.org/10.1016/j.neulet.2004.03.057

Honbolygó, F., Kóbor, A., & Csépe, V. (2017). Cognitive components of foreign word stress processing difficulty in speakers of a native language with non-contrastive stress. *International Journal of Bilingualism*. https://doi.org/10.1177/1367006917728393

Honbolygó, F., Kolozsvári, O., & Csépe, V. (2017). Processing of word stress related acoustic information: A multi-feature MMN study. *International Journal of Psychophysiology*, *118*, 9–17. https://doi.org/https://doi.org/10.1016/j.ijpsycho.2017.05.009

Jacobsen, T., & Schröger, E. (2003). Measuring duration mismatch negativity. *Clinical Neurophysiology*, *114*(6), 1133–1143. https://doi.org/10.1016/S1388-2457(03)00043-9

Jakoby, H., Goldstein, A., & Faust, M. (2011). Electrophysiological correlates of speech

perception mechanisms and individual differences in second language attainment. *Psychophysiology*, *48*(11), 1517–1531. https://doi.org/10.1111/j.1469-8986.2011.01227.x

Kager, R. (2007). Feet and metrical stress. In P. de Lacy (Ed.), *The Cambridge Handbook of Phonology* (pp. 195–228). Cambridge University Press, Cambridge.

Kóbor, A., Honbolygó, F., Becker, A. B. C., Schild, U., Csépe, V., & Friedrich, C. K. (2018). ERP evidence for implicit L2 word stress knowledge in listeners of a fixed-stress language. *International Journal of Psychophysiology*, *128*, 100–110. https://doi.org/10.1016/j.ijpsycho.2018.04.006

Koenig, T., Kottlow, M., Stein, M., & Melie-García, L. (2011). Ragu: A free tool for the analysis of EEG and MEG event-related scalp field data using global randomization statistics. *Computational Intelligence and Neuroscience*, *2011*, 1–15. https://doi.org/10.1155/2011/938925

Koenig, T., & Melie-Garcia, L. (2009). Statistical analysis of multichannel scalp field data. In *Electrical Neuroimaging* (pp. 169–189). Cambridge University Press New York.

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 and Language*, *76*(3), 332–339.

Korpilahti, P., Lang, H., & Aaltonen, O. (1995). Is there a late-latency mismatch negativity (MMN) component? In *Electroencephalography and Clinical Neurophysiology* (Vol. 95, p. 96P). https://doi.org/10.1016/0013-4694(95)90016-G

Kraus, N., & Cheour, M. (2000). Speech sound representation in the brain. *Audiology & Neuro-Otology*, *5*(3–4), 140–150. https://doi.org/13876

Kuuluvainen, S., Alku, P., Makkonen, T., Lipsanen, J., & Kujala, T. (2016). Cortical speech and

non-speech discrimination in relation to cognitive measures in preschool children. *European Journal of Neuroscience*, *43*(6), 738–750. https://doi.org/10.1111/ejn.13141

Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of Multiple Speech Segmentation Cues: A Hierarchical Framework. *Journal of Experimental Psychology: General*, *134*(4), 477–500.

Murray, M. M., Brunet, D., & Michel, C. M. (2008). Topographic ERP analyses: A step-by-step tutorial review. *Brain Topography*, *20*(4), 249–264. https://doi.org/10.1007/s10548-008-0054-5

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., … 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*(12), 2544–2590.

Norris, D., McQueen, J. M., & Cutler, A. (1995). Competition and segmentation in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(5), 1209–1228.

Peter, V., Mcarthur, G., & Thompson, W. F. (2012). Discrimination of stress in speech and music: A mismatch negativity (MMN) study. *Psychophysiology*, *49*(12), 1590–1600. https://doi.org/10.1111/j.1469-8986.2012.01472.x

Poeppel, D. (2014). The neuroanatomic and neurophysiological infrastructure for speech and language. *Current Opinion in Neurobiology*, *28*, 142–149. https://doi.org/10.1016/j.conb.2014.07.005

Ragó, A., Honbolygó, F., Róna, Z., Beke, A., & Csépe, V. (2014). Effect of maturation on

suprasegmental speech processing in full- and preterm infants: A mismatch negativity study. *Research in Developmental Disabilities*, *35*, 192–202. https://doi.org/10.1016/j.ridd.2013.10.006

Shestakova, A., Huotilainen, M., Ceponiene, R., & Cheour, M. (2003). Event-related potentials associated with second language learning in children. *Clin Neurophysiol*, *114*(8), 1507–1512.

Tremblay, A. (2008). Is second language lexical access prosodically constrained? Processing of word stress by French Canadian second language learners of English. *Applied Psycholinguistics*, *29*(4), 553–584. https://doi.org/10.1017/S0142716408080247

van Donselaar, W., Koster, M., & Cutler, A. (2005). Exploring the role of lexical stress in lexical recognition. *The Quarterly Journal of Experimental Psychology A, Human Experimental Psychology*, *58*(2), 251–273.

Weber, C., Hahne, A., Friedrich, M., & Friederici, A. D. (2004). Discrimination of word stress in early infant perception: Electrophysiological evidence. *Cognitive Brain Research*, *18*(2), 149–161.

Winkler, I., Denham, S. L., & Nelken, I. (2009). Modeling the auditory scene: predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences*, *13*(12), 532–540. https://doi.org/10.1016/j.tics.2009.09.003

Winkler, I., Lehtokoski, A., Alku, P., Vainio, M., Czigler, I., Csépe, V., … Näätänen, R. (1999). Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations. *Cognitive Brain Research*, *7*(3), 357–369.

Winkler, I., & Schröger, E. (2015). Auditory perceptual objects as generative models: Setting the stage for communication by sound. *Brain and Language*, *148*, 1–22. https://doi.org/10.1016/j.bandl.2015.05.003

Wirth, M., Horn, H., Koenig, T., Razafimandimby, A., Stein, M., Mueller, T., … Strik, W. (2008). The early context effect reflects activity in the temporo-prefrontal semantic system: Evidence from electrical neuroimaging of abstract and concrete word reading. *NeuroImage*, *42*(1), 423–436. https://doi.org/10.1016/j.neuroimage.2008.03.045

Ylinen, S., Strelnikov, K., Huotilainen, M., & Näätänen, R. (2009). Effects of prosodic familiarity on the automatic processing of words in the human brain. *International Journal of Psychophysiology*, *73*(3), 362–368.

Zachau, S., Rinker, T., Körner, B., Kohls, G., Maas, V., Hennighausen, K., & Schecker, M. (2005). Extracting rules: early and late mismatch negativity to tone patterns. *Neuroreport*, *16*(18), 2015–2019.

Zora, H., Heldner, M., & Schwarz, I. C. (2016). Perceptual correlates of turkish word stress and their contribution to automatic lexical access: Evidence from early ERP components. *Frontiers in Neuroscience*, *10*(JAN), 1–11. https://doi.org/10.3389/fnins.2016.00007

Zora, H., Riad, T., Schwarz, I. C., & Heldner, M. (2016). Lexical specification of prosodic information in Swedish: Evidence from mismatch negativity. *Frontiers in Neuroscience*, *10*(NOV), 1–12. https://doi.org/10.3389/fnins.2016.00533

Tables

Table 1. Acoustical properties of stimuli.

		Hungarian		**German**	
		1st syllable	2nd syllable	1st syllable	2nd syllable
Intensity (dB)[1]	S1	81.88	65.48	78.73	70.29
	S2	75.34	80.52	73.95	78.57
f0 (Hz) [2]	S1	236.53	201.57	222.66	218.79
	S2	221.52	213.66	242.02	218.53

Note. [1]:maximum amplitude measured in the whole length of syllable; [2]: maximum f0 measured in the whole length of syllable

Table 2. Summary of results from ANOVAs performed on the ERPs in the three time windows.

	Language		Stress Position		Role		Language * Stress Position		Language * Role		Stress Position * Role		Language * Stress Position * Role	
Time window	*F*	*p*	*F*	*p*	*F*	*p*	*F*	*p*	*F*	*p*	*F*	*p*	*F*	*p*
116-286 ms	65.83	< **.001**	7.93	**.009**	57.54	< **.001**	2.59	.118	10.82	**.003**	5.20	**.030**	9.28	**.005**
378-444 ms	108.73	< **.001**	0.93	.343	62.85	< **.001**	< 0.01	.955	2.52	.123	3.53	.070	6.57	**.016**
530-818 ms	3.86	**.059**	0.36	.553	78.40	< **.001**	2.06	.162	6.31	**.018**	3.86	.059	< 0.01	.959

Note. p-values below .050 are **boldfaced**.

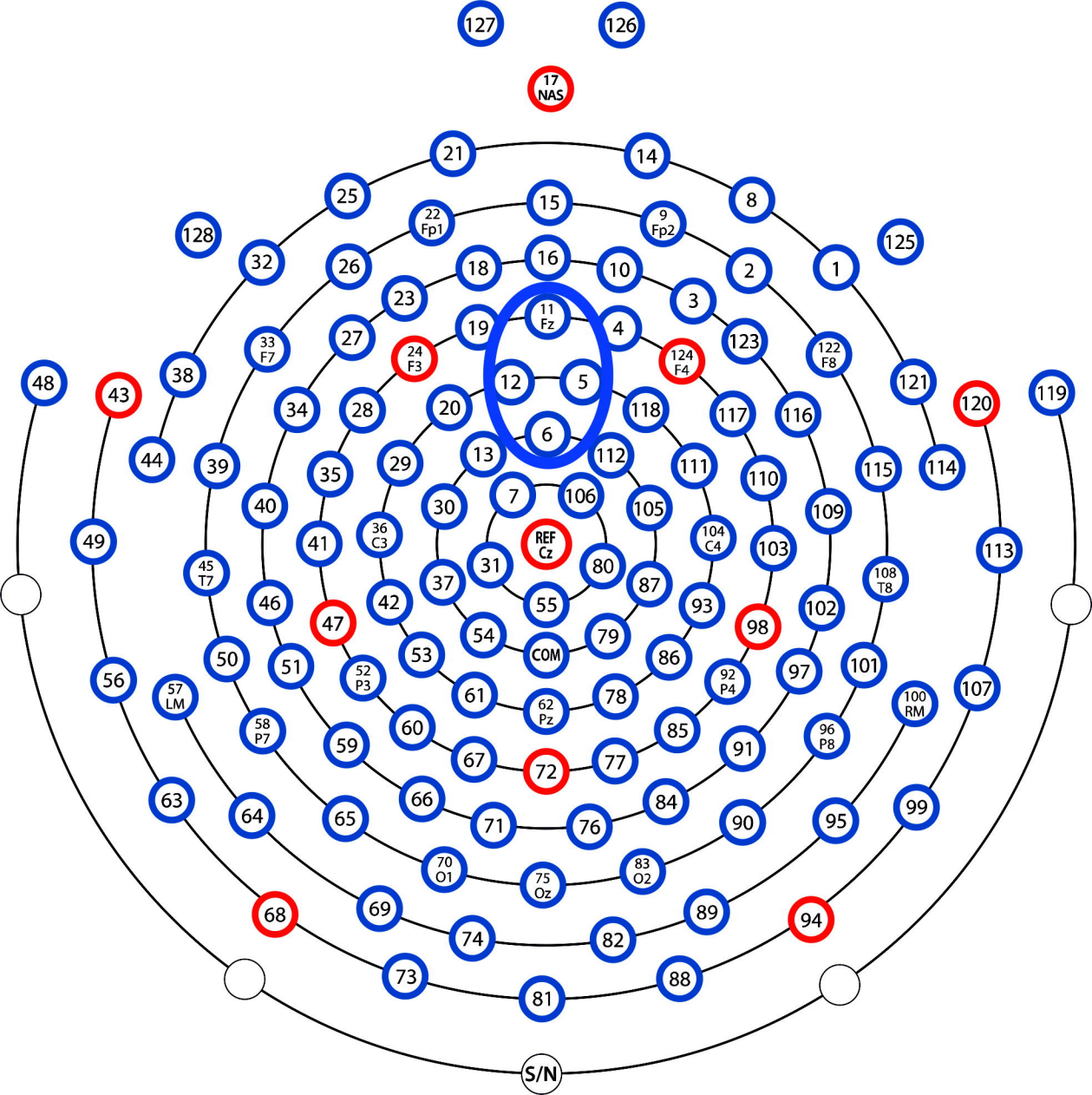
Figure captions

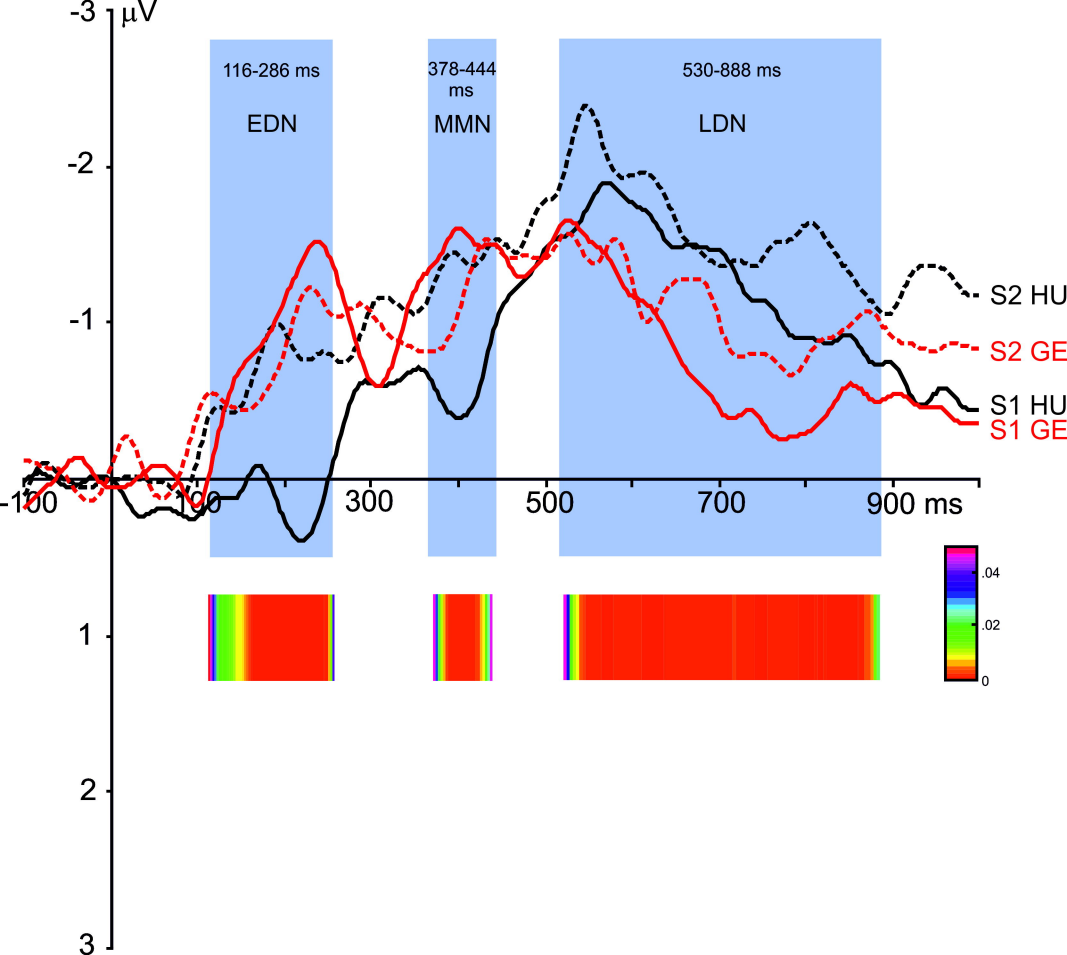
Figure 1. Grand average ERP waveforms split by language and stress position to the standard (Std) and deviant (Dev) stimuli together with the difference wave (Diff) calculated by subtracting the ERPs to standard stimuli from that to the corresponding deviant stimuli. Negativity is plotted upwards. Topoplots depict the average activity of the difference wave in three time windows determined in the TANOVA analysis. The waveform of the respective stimuli below the ERPs illustrates the temporal relation between the stimuli and ERP responses. *Note*: S1: stimuli with stress on the first syllable; S2: stimuli with stress on the second syllable; HU: Hungarian stimuli; GE: German stimuli.

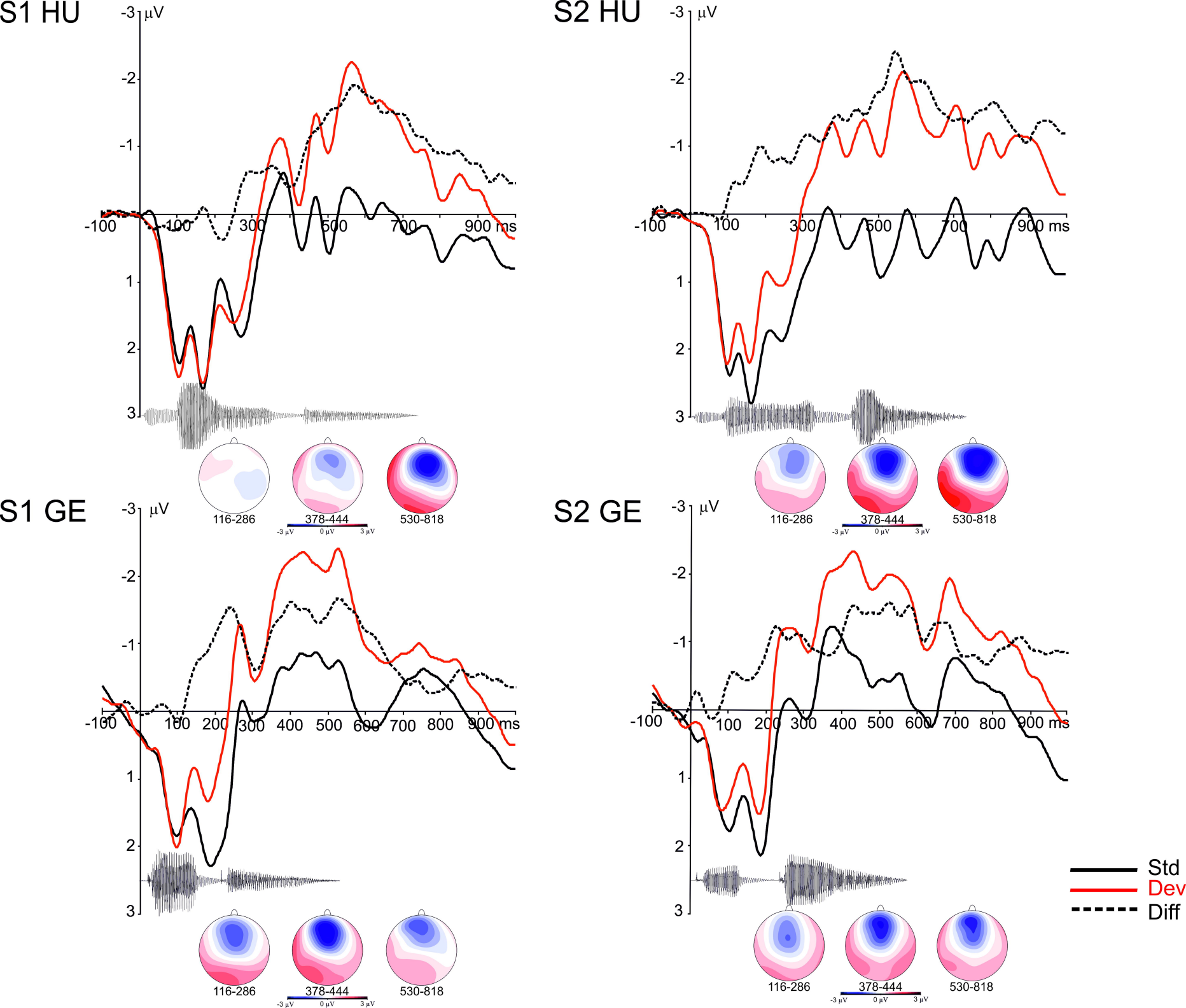
Figure 2. Difference waves in the four different conditions and the TANOVA results. The results of the TANOVA analysis are depicted below the ERPs. In the time windows where the colored intervals are shown, the Language * Stress position interaction was significant. The *p*-value of the significance is shown as a colored scale in the range 0-0.05. Shaded areas show the time windows on the ERPs selected for later analysis.
Note: S1: stimuli with stress on the first syllable; S2: stimuli with stress on the second syllable; HU: Hungarian stimuli; GE: German stimuli. EDN: Early Differentiating Negativity; LDN: Late Discriminative Negativity.

Figure 3. The electrode montage with 128 electrodes used in the experiment. The blue ellipsis shows those electrodes (5, 6, 11, 12) that were pooled together in the fronto-central electrode pool used for statistical analysis.

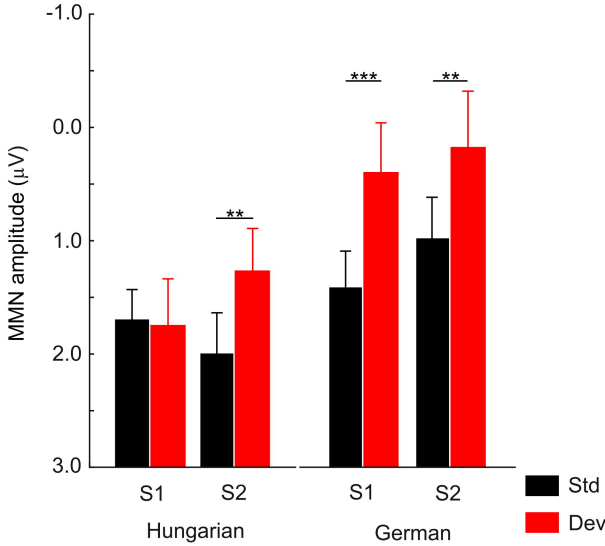
Figure 4. Results of the statistical analysis in the three time windows: A) first time window; B) second time window; C) third time window. Bar plots indicate the amplitude of ERPs in the four conditions in the standard (Std) and deviant (Dev) positions.
Note: ** $p < .01$; *** $p < .001$.


127
126
17 NAS
21
14
25
8
15
22 Fp1
9 Fp2
128
125
32
26
18
16
10
2
1
23
3
11 Fz
19
4
33 F7
27
24 F3
124 F4
123
122 F8
48
43
38
12
5
121
120
119
34
28
20
118
117
116
6
44
39
35
29
13
112
111
110
115
114
40
30
7
106
105
109
49
41
36 C3
REF Cz
104 C4
103
113
45 T7
31
80
108 T8
37
87
46
42
55
93
102
47
98
54
79
50
51
53
COM
86
56
52 P3
92 P4
97
101
57 LM
61
78
107
62 Pz
100 RM
58 P7
60
85
96 P8
59
91
67
72
77
63
66
84
64
65
71
76
90
95
99
70 O1
75 Oz
83 O2
68
69
89
94
74
82
73
88
81
S/N

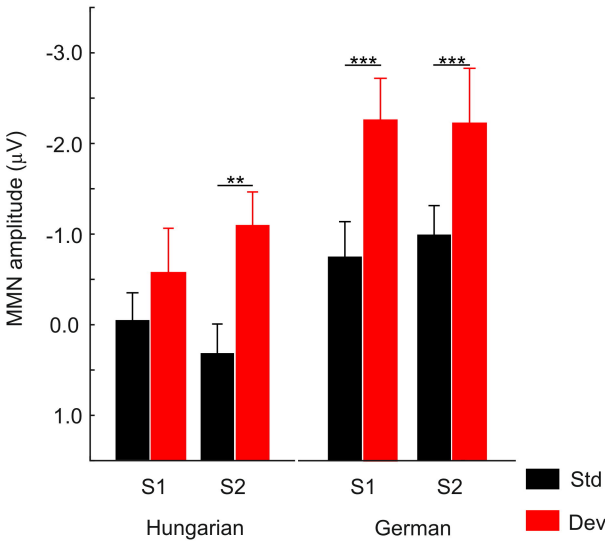

-3
μV
-2
-1
1
2
3
116-286 ms
EDN
378-444 ms
MMN
530-888 ms
LDN
S2 HU
S2 GE
S1 HU
S1 GE
-100
100
300
500
700
900 ms
.04
.02
0


S1 HU
S2 HU
S1 GE
S2 GE
µV
-3
-2
-1
1
2
3
-100
100
300
500
700
900 ms
200
400
600
800
116-286
378-444
530-818
-3 µV
0 µV
3 µV
Std
Dev
Diff

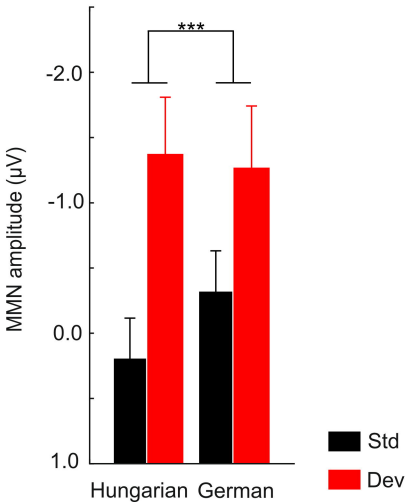
A


-1.0
0.0
1.0
2.0
3.0
MMN amplitude (µV)
***
**
**
S1
S2
S1
S2
Hungarian
German
Std
Dev

B


-3.0
-2.0
-1.0
0.0
1.0
MMN amplitude (µV)
***
***
**
S1
S2
S1
S2
Hungarian
German
Std
Dev

C


***
-2.0
-1.0
0.0
1.0
MMN amplitude (µV)
Hungarian
German
Std
Dev