

ORIGINAL ARTICLE



WILEY

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

Ferenc Honbolygó^{1,2}  | Andrea Kóbor¹ | Borbála German^{1,3} | Valéria Csépe^{1,4}

¹Brain Imaging Centre, Research Centre for Natural Sciences, Budapest, Hungary

²Institute of Psychology, ELTE Eötvös Loránd University, Budapest, Hungary

³Department of Cognitive Science, Budapest University of Technology and Economics, Budapest, Hungary

⁴Faculty of Modern Philology and Social Sciences, University of Pannonia, Budapest, Hungary

Correspondence

Ferenc Honbolygó, Brain Imaging Centre, Research Centre for Natural Sciences, Magyar tudósok körútja 2., H-1117, Budapest, Hungary.
Email: honbolygo.ferenc@ttk.mta.hu

Funding information

Hungarian Scientific Research Fund, Grant/Award Number: OTKA NK 101087, OTKA K 119365 and OTKA FK 124 412; Hungarian Academy of Sciences

Abstract

Understanding speech at the basic levels entails the simultaneous and independent processing of phonemic and prosodic features. While it is well-established that phoneme perception relies on language-specific long-term traces, it is unclear if the processing of prosodic features similarly involves language-specific representations. In the present study, we investigated the processing of a specific prosodic feature, word stress, using the method of event-related brain potentials (ERPs) employing a cross-linguistic approach. 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 methodology showed that all pseudowords in the deviant position elicited an Early Differentiating Negativity and a Mismatch Negativity 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, but the same stress pattern with a nonnative accent was processed as deviant. This implies that the processing of word stress was based on language-specific long-term memory traces.

KEYWORDS

ERP, MMN, prosody, speech, word stress

1 | 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 & Schlewsky, 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, 2010). 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 noncontrastive). 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 whether stress-related representations

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. Psychophysiology published by Wiley Periodicals, Inc. on behalf of Society for Psychophysiological Research

can be considered as language-specific and whether we find any evidence of language-specific processing at the neuro-cognitive level.

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 frontocentral 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 odd-ball 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 preattentive detection of violation of simple or complex regularities (Winkler, Denham, & Nelken, 2009).

A more complex model for the MMN's emergence is proposed 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 extracts 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 the 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 and 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.

There are a number of studies that applied the MMN paradigm to study the processing of word stress. Weber, Hahne, Friedrich, and Friederici (2004) found that German adults showed an MMN to the word “baba” (“baby” in English) in the deviant position both in the trochaic form (stress on the first syllable, i.e., /'baba/) and in the iambic form (stress on the second syllable, i.e., /ba'ba/), indicating that both stress patterns were equally well discriminated.

Ylinen, Strelnikov, Huotilainen, and Näätänen (2009) studied the effect of prosodic familiarity in the case of Finnish words and pseudowords with familiar (trochaic, i.e., stress on the first syllable) and unfamiliar (iambic, i.e.,

stress on the second syllable) stress patterns. They used the pseudowords with the familiar stress pattern as standard and investigated the MMNs elicited by pseudowords with unfamiliar stress pattern, words with familiar stress pattern, and words with unfamiliar stress pattern. According to the results, the pseudowords with unfamiliar stress pattern elicited two MMNs related to the first and second syllables of utterances and were interpreted as the detection of the unexpected lack of stress on the first syllable and the detection of the unexpected presence of stress on the second syllable, respectively. Words with familiar stress pattern elicited a single MMN in the mid-latency time window, which was interpreted as the detection of the change in the lexical status of the word, and words with unfamiliar stress pattern elicited a single MMN related to the presence of stress on the second syllable. These results demonstrated that the familiarity of the words modulated the processing of stress pattern, and also that the familiarity of stress pattern modulated the processing of words.

Zora, Heldner, and Schwarz (2016) investigated stress processing in Turkish. In the study, the authors presented word and pseudoword stimuli differing in their stress pattern, with stimuli with iambic stress pattern being always standards and stimuli with trochaic pattern being deviants. They manipulated the acoustic features contributing to stress (f_0 , spectral emphasis, duration, all features combined), in order to investigate to what extent they are utilized for lexical access. Results showed that changes in prosodic features elicited ERPs both in words and pseudowords but there were processing differences between the prosodic features. The most prominent feature was the f_0 eliciting an MMN in words but a frontal P3a in pseudowords, and consequently, was claimed to be lexically specified in Turkish by the authors. Results also indicated that the individual prosodic features rather than combined features differentiated language-related effects, since the combined features elicited similar MMN in both words and pseudowords.

The same authors (Zora, Schwarz, & Heldner, 2015) investigated also the contribution of two specific features, f_0 and intensity to stress processing in English. In the study, participants heard stress minimal pair versions (i.e., the same word with trochaic and iambic stress pattern) of the word “upset” and the pseudoword “ukfet”. Deviants—having trochaic stress compared to the standards having iambic stress—differed from the standards in f_0 , intensity, or the combination of the two. According to the results, all deviants elicited the MMN, although intensity change elicited a smaller MMN in pseudowords than f_0 , confirming the importance of both f_0 and intensity to determine words stress and contribute to lexical access.

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 nonspeech-related processing.

The above studies established convincingly that the MMN can be reliably used to investigate the neurocognitive background of stress processing and demonstrated the sensitivity of this particular ERP component to acoustic, phonetic, and lexical effects. In a series of studies, we investigated word stress processing in Hungarian. Hungarian is an excellent target language for studying word stress representations because the stress pattern of the disyllabic word is highly regular: stress always falls on the first syllable. Therefore, any changes in the stress pattern are considered as violating the regularity, leading to illegal stress patterns. 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 & Csépe, 2013), it was also 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, supporting the importance of familiarity in word stress processing. This study was 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 that appear to contain a description of the expected word stress pattern characteristic for a given language at the phonological level. However, the language-specific nature of these representations has not been directly investigated in a cross-linguistic experiment so far; therefore, the present study specifically aimed to clarify this aspect.

The MMN paradigm has also been used to investigate the 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 nonnative phonemes, indicating the modulation of speech processing mechanisms by the exposure to the native language. Chandrasekaran, Krishnan, and 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 nonnative (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. In the study, we used naturally produced, meaningless pseudowords having stress either on the first or the second syllable. Moreover, we created a native and nonnative (German) version of both types of pseudowords in order to study the ERP correlates of nonnative stress processing. We assumed that listeners would detect preattentively if a pseudoword was pronounced in a native or nonnative way, and only the native-like stimuli will activate the long-term stress representation. We expected that the activation of the long-term representation would be indicated by not recognizing the native legal stress patterns as deviant (as they do not differ from the expected stress pattern specified in the long-term representation). Based on these assumptions, 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 nonnative version is presented, and no MMN would be elicited by the native legal pseudoword because listeners would not detect this pattern as irregularity.

2 | METHOD

2.1 | 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 Preprocessing). Therefore, 30 participants remained in the final sample (19 females). The participants, aged between 18 and 25 years ($M_{\text{Age}} = 21.7$ 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 (Kóbor et al., 2018). All of them were right-handed according to the Edinburgh Handedness Inventory revised version (Dragovic, 2004) ($M_{\text{LQ}} = 90.4$, $SD = 12.9^1$). 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 a 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 (3,300 HUF, ca. 10 Euros) or course credit for taking part in the experiment.

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

2.2 | Stimuli

Four pseudowords were used as stimuli: two spoken by a Hungarian native speaker and two spoken by a German native speaker (both were female). The same reiterative pseudoword “bébé” (/’be:be:/) was recorded in isolation, using a falling intonation by each speaker with two different stress patterns: one with stress on the first syllable (/’be:be:/ trochaic stress pattern, henceforth referred to as S1), and one with stress on the second syllable (/be:’be:/ 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.² The acoustical properties of stimuli are illustrated in Figure 1 and summarized in Table 1.

2.3 | Procedure

Stimuli were presented in a *passive oddball paradigm* with a stimulus onset asynchrony (SOA) varying randomly between 850 and 1,050 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; (3) German S1 standard with German S2 deviant; (4) 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 1,000 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 different. They had unlimited time to initiate a response for each stimulus pair. The next stimulus pair was presented after the participant gave a 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 the same/different responses was counterbalanced across participants. The entire experimental procedure lasted about 2 hr, including the application of the electrode net. Both the EEG and behavioral experiment were written in and controlled by the Presentation software (v. 17.0, Neurobehavioral Systems).

2.4 | EEG recording and preprocessing

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 a saline electrolyte solution (see Figure 2). Electrode Cz was used as a reference and a sampling rate of 500 Hz was applied.

Data were 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 preprocessing, EEG data were band-pass filtered offline between 0.3 and 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, &

²In order to test if pseudowords are recognizable as trochaic or iambic, we performed an additional behavioral experiment with ten healthy adult native speakers of Hungarian (7 females, $M_{\text{years}} = 26.7$, $SD_{\text{years}} = 6.5$). We recorded three further reiterative pseudowords as fillers uttered by three Hungarian and three German speakers, and presented them together with the original stimulus in a two-alternative forced-choice task. Participants had to decide whether the pseudowords (each repeated 10 times) were stressed on the first or on the second syllable. The results indicated that the number of correct responses were significantly higher than chance level (50%) for all /be:be:/ stimuli (HU S1: $M = 9.90$ (0.316), $t(9) = 49$, $p < .001$; HU S2: $M = 7.30$ (2.497), $t(9) = 2.91$, $p = .017$; GE S1: $M = 8.60$ (1.897), $t(9) = 6$, $p < .001$; GE S2: $M = 8.60$ (1.955), $t(9) = 5.82$, $p < .001$).

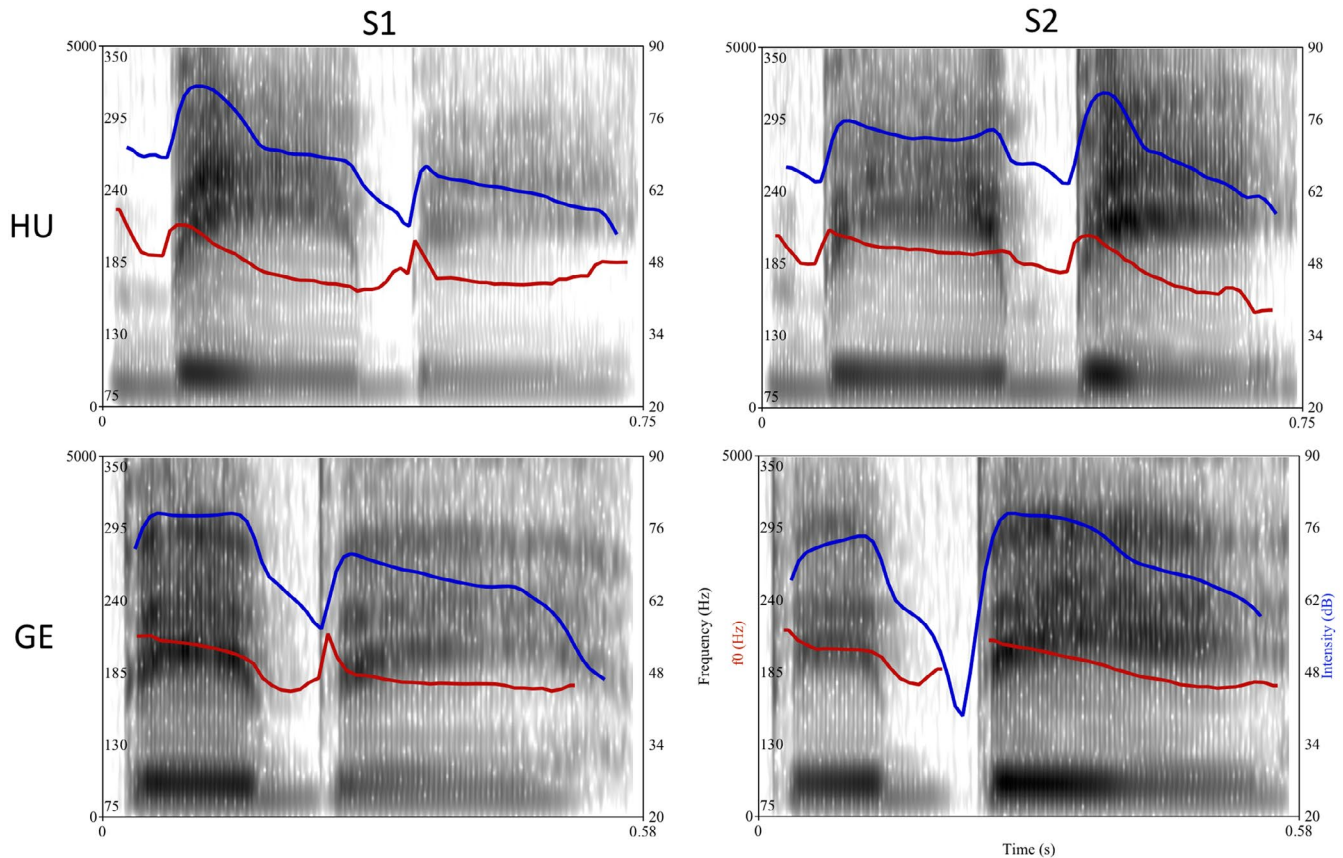


FIGURE 1 Acoustic characteristics of the stimuli. The figures show the spectrogram, f_0 (red line) and intensity (blue line) of the four stimuli. Note the different time axes of Hungarian and German stimuli. S1: stimuli with stress on the first syllable; S2: stimuli with stress on the second syllable; HU: Hungarian stimuli; GE: German stimuli

TABLE 1 Acoustical properties of stimuli

		Hungarian		German	
		1st syllable	2nd syllable	1st syllable	2nd syllable
Intensity (dB) ^a	S1	82.13	66	78.79	70.8
	S2	75.57	80.96	74.35	78.78
f_0 (Hz) ^b	S1	225.32	202.65	213.61	215.44
	S2	211.32	206.36	217.17	209.41
Duration (ms)	S1	360	386	170	403
	S2	350	396	140	433

^aMaximum amplitude measured in the whole length of the syllable.

^bMaximum f_0 measured in the whole length of the syllable.

Makeig, 2007). Using ICA, 1–6 ICA components ($M = 2.93$) were removed, then, the channel-based EEG data were re-composed. Third, EEG data were re-referenced to the average activity of all electrodes. After that, the continuous EEG was segmented into epochs. Epochs extended from -100 to $1,000$ ms relative to the onset of the standard and deviant stimuli, respectively, in each condition. An equal number of epochs ($n = 100$) were selected for both the standard and deviant stimuli in each condition by analyzing only the standards preceding the deviants. 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 the 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 out of the possible 100 ($SD = 5.4$; range: 72–100). After artifact rejection, artifact-free epochs were

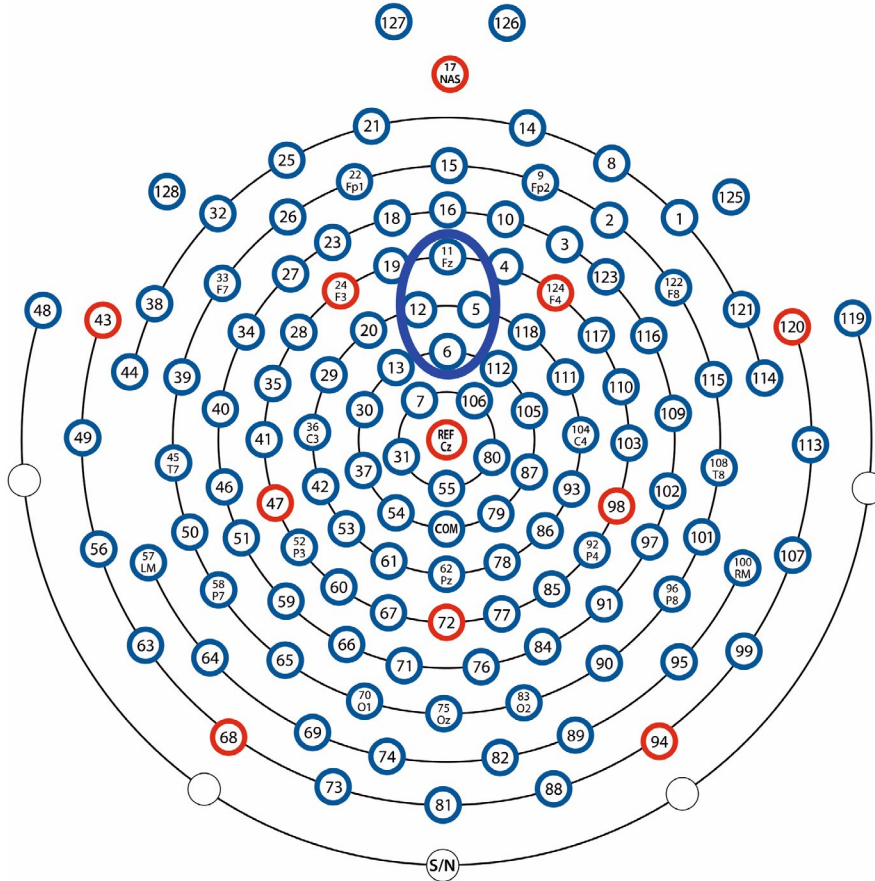


FIGURE 2 The electrode montage with 128 electrodes used in the experiment. The blue ellipse shows electrodes (5, 6, 11, 12) pooled together into the fronto-central electrode pool used for statistical analysis

baseline corrected based on the mean activity from -100 to 0 ms. Finally, the remaining epochs were averaged.

2.5 | Data/ERP analysis

ERP analysis involved a permutation-based nonparametric topographic analysis of variance (TANOVA, Murray, Brunet, & Michel, 2008) conducted in Ragu (Koenig, Kottlow, Stein, & Melie-García, 2011) on the individual ERP waveforms. The aim of the analysis was to determine those latency ranges where there were significant differences between the ERP curves related to the effect of interest (see below). The TANOVA allows the quantification of ERP components without selecting time windows arbitrarily (i.e., where the ERP components' amplitude seems to be maximal or the difference between particular conditions seems to be maximal) or by following previous literature; therefore, it is an unbiased approach (see Brooks, Zoumpoulaki, & Bowman, 2017).

The analysis was conducted separately for the Hungarian and German stimuli. The reason for this was that differences in duration between the stimuli (see Figure 1 and Table 1) made the point-by-point comparison of the ERPs problematic across the languages, as the shorter first syllables in German could potentially elicit ERP components with shorter latencies.

We calculated the TANOVA for topographic dissimilarity (TD) (Koenig & Melie-García, 2009; Wirth et al., 2008), which is a single measure of the distance between two electric field topographies. We assumed that the ERPs to standard and deviant stimuli and to different stress patterns would demonstrate slightly different topographic distributions. The TANOVA was performed with Role (Standard vs. Deviant) and Stress Position (S1 vs. S2) as within-subjects factors. Standard and deviant responses for the same stimuli (i.e., either S1 or S2) taken from different blocks were used in the analysis. This approach enabled us 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).

We performed 1,000 randomization runs and applied a 5% significance threshold on the individual averages during the entire epoch length (no normalization was used). Regarding the main question of this study, we focused on the Stress Position * Role interaction effect, which would show the time range where the difference between deviants and standards is modulated by the stress pattern. As mentioned above, we did not compare the effects directly between languages.

3 | RESULTS

3.1 | 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 than 2,000 ms were excluded from further analysis, thus, accuracy was calculated without these trials. We found the 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.

3.2 | ERP results

Grand average ERP waveforms to the standard and deviant stimuli and the difference curves split by language and stress position are presented in Figure 3. As it can be seen in the figure, stimuli elicited a complex set of ERP waveforms, starting with positive ERP deflections that went in the negative direction from around 200–300 ms. ERPs to the stimuli in the standard and deviant roles started to deviate from each other already from around 100 ms.

Difference curves calculated by subtracting the ERPs to the standard from the ERPs to the deviant elicited by the

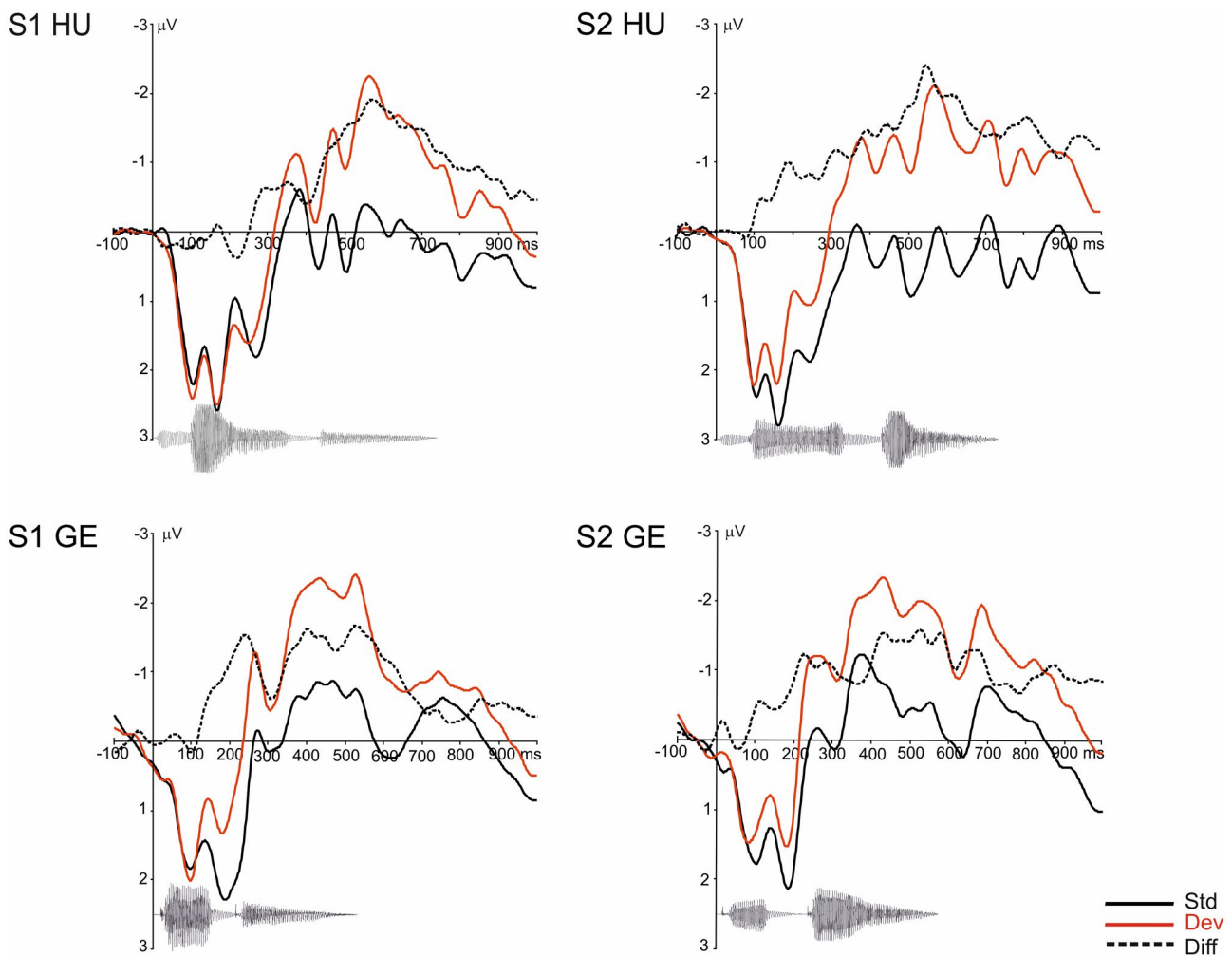


FIGURE 3 Grand average ERP waveforms split by language and stress position to the standard (Std) and deviant (Dev) stimuli together with the difference wave (Diff) calculated by subtracting the ERPs to standard stimuli from that to the corresponding deviant stimuli. Negativity is plotted upwards. ERPs are plotted for the fronto-central electrode pool. The waveform of the respective stimuli below the ERPs illustrates the temporal relation between the stimuli and ERP responses. S1: stimuli with stress on the first syllable; S2: stimuli with stress on the second syllable; HU: Hungarian stimuli; GE: German stimuli

acoustically same stimuli are depicted in Figure 4. Difference waves were calculated as follows: (S1 HU) Hungarian S1 deviant minus Hungarian S1 standard; (S2 HU) Hungarian S2 deviant minus Hungarian S2 standard; (S1 GE) German S1 deviant minus German S1 standard; (S2 GE) German S2 deviant minus German S2 standard. Calculating the difference waveforms as a difference between the standard and deviant versions of the same stimulus recorded in two separate blocks ensured that they reflect only the processing of the deviance between the standard and deviant versions of the stimuli.

The heat maps under the ERPs in Figure 4 show those latency ranges where the TANOVA indicated significant effects (main effect or interaction) for scalp topographies (colors denote the p values). We found a significant Stress position * Role interaction effect only for the Hungarian stimuli between 176–254 ms and 374–418 ms. There was no significant interaction effect for the German stimuli (for the results of the full TANOVA analysis, see Figure S1 in Supplementary material). These two latency windows correspond to two negative ERP deflections, and the significant interaction is due to S1 HU not eliciting or eliciting smaller deflections. As for the German stimuli, these two deflections are also present, but they are not different for the S1 and S2 stimuli.

In order to provide a more conventional analysis of the ERP effects, we also performed repeated-measures ANOVAs. For this analysis, we first averaged the activity of electrodes 5, 6, 11, 12 (see Figure 2) into a frontocentral 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. Third, for the ANOVA conducted on the Hungarian stimuli, we selected those two time windows where the Language * Stress Position interaction was significant based on the TANOVA (176–254 ms, 374–418 ms). For the ANOVA conducted on the German stimuli, because the TANOVA did not reveal a significant interaction effect, we selected two time windows that encompassed the peaks of the EDN and MMN components visible in Figure 4 (i.e., 150–250 ms and 350–450 ms; for the topoplots showing the amplitude distribution of EDN and MMN components, see Figures S2 and S3 in the Supplementary materials). As a final step, we quantified the ERP components as the mean amplitude within the time windows specified above in each of the eight conditions. The mean amplitudes of the ERPs were then entered into two-way repeated-measures ANOVAs, separately for Hungarian and German stimuli, with 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 of the ANOVAs of the Hungarian stimuli in the two time windows are summarized in Table 2. For the Hungarian stimuli, in the first latency window

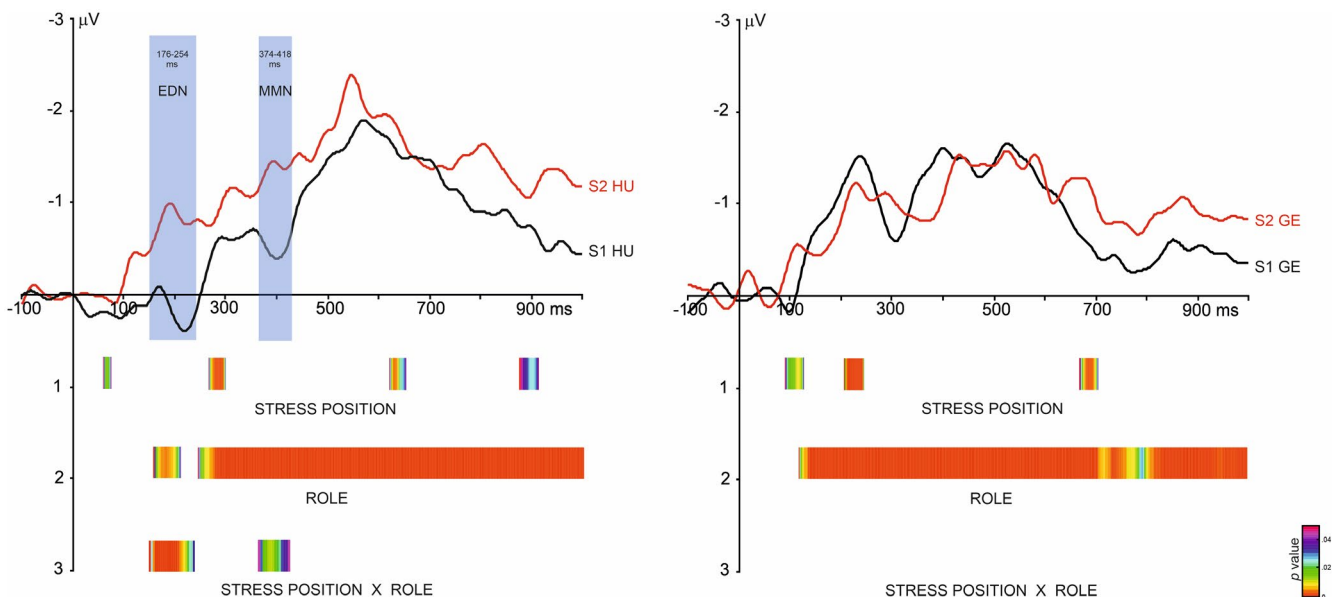


FIGURE 4 Difference waves in the four different conditions split by language (Hungarian: left side; German: right side) and the TANOVA results. ERPs are plotted for the fronto-central electrode pool. The results of the TANOVA are depicted below the ERPs. The colored intervals show those time windows where the TANOVA indicated significant main effects of Stress Position, Role, and Stress Position * Role interaction. Only those time windows are shown where the effect was significant for at least 10 consecutive time points. The p value of the significance is shown as a colored scale in the range of 0–0.05. Shaded areas show the time windows on the ERPs selected for later analysis. 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

TABLE 2 Summary of results from ANOVAs of the Hungarian stimuli performed on the ERPs in the two time windows

Time window	Stress position		Role		Stress position * Role	
	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>
176–254 ms	0.26	.610	9.52	.004	21.17	<.001
374–418 ms	0.59	.448	31.79	<.001	13.04	.001

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

TABLE 3 Summary of results from ANOVAs of the German stimuli performed on the ERPs in the two time windows

Time window	Stress position		Role		Stress position * Role	
	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>
150–250 ms	14.15	<.001	43.17	<.001	2.55	.121
350–450 ms	1.67	.207	36.69	<.001	1.43	.242

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

(176–254 ms), the main effect of Role, $F(1, 29) = 9.52$, $p = .004$, $\eta_p^2 = .247$, and the Stress Position * Role interaction, $F(1, 29) = 21.17$, $p < .001$, $\eta_p^2 = .422$, were significant. Pair-wise tests showed that the mean amplitude of the ERPs was more negative for the deviant than for the standard for Hungarian S2 stimuli ($p < .001$) but not for the Hungarian S1 stimuli ($p = .560$). This indicates that the Hungarian S1 stimuli did not elicit a significantly different ERP response for the deviant versus standard role in this early latency window.

In the second latency window (374–418 ms), the main effect of Role, $F(1, 29) = 31.79$, $p < .001$, $\eta_p^2 = .523$, was significant. This effect was qualified by the Stress Position * Role interaction, $F(1, 29) = 13.04$, $p = .001$, $\eta_p^2 = .310$. Again, pair-wise tests showed that the mean amplitude of the ERPs was more negative for the deviant than for the standard for Hungarian S2 stimuli ($p < .001$) but not for the Hungarian S1 stimuli ($p = .107$). This indicates that no Role effect (deviant < standard) was elicited for the Hungarian S1 stimuli in this second latency window.

Results of the ANOVAs of the German stimuli in the two time windows are summarized in Table 3. For the German stimuli, in the first latency window (150–250 ms), the main effects of Stress Position, $F(1, 29) = 14.15$, $p < .001$, $\eta_p^2 = .328$, and Role, $F(1, 29) = 43.17$, $p < .001$, $\eta_p^2 = .598$, were significant. The mean amplitude of the ERPs was more negative for the deviant (0.59 μV) than for the standard (1.58 μV) and for the S2 stimuli (0.86 μV) than for S1 (1.31 μV).

In the second latency window (350–450 ms), the main effect of Role, $F(1, 29) = 36.69$, $p < .001$, $\eta_p^2 = .559$, was significant, and the mean amplitude of the ERPs was more negative for the deviant (−2.14 μV) than for the standard (−0.84 μV).

4 | DISCUSSION

This study investigated 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 two time windows that demonstrated stimulus-related effects.

The ERP activity in the first (176–254 ms for the Hungarian, 150–250 ms for the German stimuli) and second (374–418 ms for the Hungarian, 350–450 ms for the German stimuli) 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 Hungarian 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. We suggest that 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. As we suggested previously (Honbolygó & Csépe, 2013), the lack of MMN in S1 cannot be explained simply with reference to acoustic differences between the stimuli, as in the case of the trochaic stress pattern, the deviant actually contains more salient acoustic features than the standard. One could argue that the MMN in the German conditions could reflect the difference in both the

speech sounds and prosodic patterns between the two languages. We believe that we could disentangle the ERP responses to the stress patterns from the ERP responses to the speech sounds because, in a given experimental block, either the Hungarian or the German stimuli were presented, and the stimuli differed only in their stress pattern (i.e., Hungarian S1 and S2; German S1 and S2). Therefore, we assume that the ERP differences obtained reflect solely the processing of the stress pattern differences.

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 sensitivity of the EDN 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 in preschool children. They found that the EDN was elicited around 100–200 ms after stimulus onset by all feature changes except for the vowel duration, and nonspeech consonant and f_0 changes. Because the EDN was larger for speech than for nonspeech 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 enhanced N2s 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 initially stressed stimuli compared with the initially unstressed ones. However, in our study, the EDN component was larger for the initially unstressed syllable compared with the initially stressed one, which suggests that the negative deflection appearing in our study is different from the N2 components. Since data on the EDN component are 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 (374–418 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 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 that stress-related linguistic information was processed in a language-specific way. The functional difference between the EDN and MMN is currently unclear and requires further investigation. However, based on the temporal relations between the syllables and the ERP deflections (see Figure 3) we

can speculate that the EDN and MMN reflect the processing of different information: because its latency is very short, the EDN is probably related to processing the onset of the first syllable, while the MMN may be related to the processing of acoustic-phonetic changes inside the first syllable. Because the second syllable starts around 360 ms, it is unlikely that the MMN would be related to the processing of that.

These results can be interpreted in the AERS model described in the Introduction (Winkler & Schröger, 2015). In our case, the lack of ERP difference in the two 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 nonnative stress) regularities.

5 | LIMITATIONS

One of the most important limitations of our study is that the experimental stimuli were not acoustically controlled. We wanted to create the stimuli as natural as possible; therefore, we did not modify their acoustical properties and did not match Hungarian and German stimuli (apart from the overall loudness levels that were equalized across the stimuli). The advantage of this approach, in our opinion, is that it is unbiased, as it does not take into account any expectations regarding the acoustical characteristics of stress. The disadvantage is that the acoustical properties of stimuli are uncontrolled, which makes the cross-linguistic comparison difficult. In our case, both Hungarian and German speakers used intensity as the most prominent cue, while f_0 followed the same falling pattern in all stimuli, due to the instruction used during the recording.

Furthermore, the trochaic and iambic versions of the stimuli were created by asking the speakers to utter the pseudowords this way. This might have been a more challenging task for the Hungarian speaker because the iambic stress pattern does not exist in Hungarian. Nevertheless, as the acoustic measurements demonstrated, there were at least some cues (e.g., intensity) that differed considerably between the stressed and unstressed syllables of the two versions of the pseudowords. We believe that despite the uncontrolled nature of the stimuli, we could still create trochaic and iambic versions of the pseudowords in both languages that were highly discriminable, as shown by the high accuracy values in the same-different task.

6 | 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 that 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 patterns.

ACKNOWLEDGMENT

The authors are thankful for Gabriella Baliga, Renáta Szűcs, Livia Elek for their help in EEG recording and data collection, and Dorottya Gyarmathy and Mária Gósy for their help in recording the Hungarian stimuli. We would like to thank the Baby- & KinderLab Uni Tübingen for the help in recording the German stimuli. The study was supported by the Hungarian Scientific Research Fund (project number: OTKA NK 101087, OTKA K 119365, PI: V.C.; OTKA FK 124412, PI: A.K.), the János Bolyai Research Fellowship of the Hungarian Academy of Sciences (F.H., A.K.), and the Postdoctoral Fellowship of the Hungarian Academy of Sciences (A.K.).

CONFLICT OF INTEREST

None.

ORCID

Ferenc Honbolygó  <https://orcid.org/0000-0002-5266-0188>

REFERENCES

- 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>
- 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. <https://doi.org/10.1016/j.brainres.2006.09.046>
- 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. <https://doi.org/10.1037/0096-1523.14.1.113>
- Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport*, *8*(4), 919–924. <https://doi.org/10.1097/00001756-199703030-00021>
- Delorme, A., Sejnowski, T., & Makeig, S. (2007). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *NeuroImage*, *34*(4), 1443–1449. <https://doi.org/10.1016/j.neuroimage.2006.11.004>
- 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>
- 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>
- 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., Koložsvári, 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/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](https://doi.org/10.1016/S1388-2457(03)00043-9)
- Kager, R. (2010). Feet and metrical stress. In P. de Lacy (Ed.), *The Cambridge handbook of phonology* (pp. 195–228). Cambridge, UK: Cambridge University Press. <https://doi.org/10.1017/cbo9780511486371.010>
- 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-García, L. (2009). Statistical analysis of multichannel scalp field data. In C. M. Michel, T. Koenig, D. Brandeis, L. R. Gianotti, & J. Wackermann (Eds.), *Electrical neuroimaging* (pp. 169–190). Cambridge, UK: Cambridge University Press. <https://doi.org/10.1017/CBO9780511596889.009>
- Kraus, N., & Cheour, M. (2000). Speech sound representation in the brain. *Audiology & Neuro-Otology*, *5*(3–4), 140–150. <https://doi.org/10.1159/000013876>
- 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>
- 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(6615), 432–434. <https://doi.org/10.1038/385432a0>
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, 118(12), 2544–2590. <https://doi.org/10.1016/j.clinph.2007.04.026>
- 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>
- Poeppl, 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>
- 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. <https://doi.org/10.1016/j.cogbrainres.2003.10.001>
- 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. [https://doi.org/10.1016/S0926-6410\(98\)00039-1](https://doi.org/10.1016/S0926-6410(98)00039-1)
- 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. <https://doi.org/10.1016/j.ijpsycho.2009.05.013>
- 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(January), 1–11. <https://doi.org/10.3389/fnins.2016.00007>
- Zora, H., Schwarz, I.-C., & Heldner, M. (2015). Neural correlates of lexical stress. *NeuroReport*, 26, 791–796. <https://doi.org/10.1097/WNR.0000000000000426>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

Figure S1. Results of the TANOVAs conducted separately on the Hungarian and German stimuli. The graphs depict the p -values of the different effects (main effects of Stress Position and Role, interaction between Stress Position and Role) separately for the Hungarian and German stimuli during the whole ERP segments (from -100 ms to 1000 ms). White areas show those latency ranges where the effects were significant ($p < .050$).

Figure S2. Topoplots of the EDN and MMN components for the Hungarian stimuli. Plots depict the amplitude distribution of the average activity of the standard, deviant, and difference (deviant minus standard) ERPs in latency windows determined in the TANOVA, separately for the S1 and S2 stimuli. The distribution shows a frontal negativity accompanied by a posterior positivity in the S2 difference waves in both latency windows, which is missing in the S1 difference waves. Note: HU1: Hungarian stimuli with stress on the first syllable; HU2: Hungarian stimuli with stress on the second syllable; STD: stimuli in the standard role; DEV: stimuli in the deviant role; DIFF: difference of the deviant minus standard.

Figure S3. Topoplots of the EDN and MMN components for the German stimuli. Plots depict the amplitude distribution of the average activity of the standard, deviant, and difference (deviant minus standard) ERPs in latency windows determined by visual inspection, separately for the S1 and S2 stimuli. The distribution shows a frontal negativity accompanied by a posterior positivity for both S1 and S2 stimuli. Note: GE1: German stimuli with stress on the first syllable; GE2: German stimuli with stress on the second syllable; STD: stimuli in the standard role; DEV: stimuli in the deviant role; DIFF: difference of the deviant minus standard.

How to cite this article: Honbolygó F, Kóbor A, German B, Csépe V. Word stress representations are language-specific: Evidence from event-related brain potentials. *Psychophysiology*. 2020;57:e13541. <https://doi.org/10.1111/psyp.13541>