



Perceptual Correlates of Turkish Word Stress and Their Contribution to Automatic Lexical Access: Evidence from Early ERP Components

Hatice Zora*, Mattias Heldner and Iris-Corinna Schwarz

Department of Linguistics, Stockholm University, Stockholm, Sweden

OPEN ACCESS

Edited by:

Daniela Sammler,
Max Planck Institute for Human
Cognitive and Brain Sciences,
Germany

Reviewed by:

Karen Henrich,
Max Planck Institute for Empirical
Aesthetics, Germany
Cyrille Magne,
Middle Tennessee State University,
USA

*Correspondence:

Hatice Zora
hatice@ling.su.se

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 14 October 2015

Accepted: 08 January 2016

Published: 20 January 2016

Citation:

Zora H, Heldner M and Schwarz I-C
(2016) Perceptual Correlates of
Turkish Word Stress and Their
Contribution to Automatic Lexical
Access: Evidence from Early ERP
Components. *Front. Neurosci.* 10:7.
doi: 10.3389/fnins.2016.00007

Perceptual correlates of Turkish word stress and their contribution to lexical access were studied using the mismatch negativity (MMN) component in event-related potentials (ERPs). The MMN was expected to indicate if segmentally identical Turkish words were distinguished on the sole basis of prosodic features such as fundamental frequency (f_0), spectral emphasis (SE), and duration. The salience of these features in lexical access was expected to be reflected in the amplitude of MMN responses. In a multi-deviant oddball paradigm, neural responses to changes in f_0 , SE, and duration individually, as well as to all three features combined, were recorded for words and pseudowords presented to 14 native speakers of Turkish. The word and pseudoword contrast was used to differentiate language-related effects from acoustic-change effects on the neural responses. First and in line with previous findings, the overall MMN was maximal over frontal and central scalp locations. Second, changes in prosodic features elicited neural responses both in words and pseudowords, confirming the brain's automatic response to any change in auditory input. However, there were processing differences between the prosodic features, most significantly in f_0 : While f_0 manipulation elicited a slightly right-lateralized frontally-maximal MMN in words, it elicited a frontal P3a in pseudowords. Considering that P3a is associated with involuntary allocation of attention to salient changes, the manipulations of f_0 in the absence of lexical processing lead to an intentional evaluation of pitch change. f_0 is therefore claimed to be lexically specified in Turkish. Rather than combined features, individual prosodic features differentiate language-related effects from acoustic-change effects. The present study confirms that segmentally identical words can be distinguished on the basis of prosodic information alone, and establishes the salience of f_0 in lexical access.

Keywords: Turkish, prosody, word stress, lexical access, event-related potential

INTRODUCTION

Turkish stress assignment has so far mostly been investigated from a phonological point of view, with default word stress being reported to be ordinarily on the final syllable (Lees, 1961; Sezer, 1983; Hameed, 1985; Barker, 1989; Inkelas and Orgun, 1998, 2003; Kabak and Vogel, 2001). This default pattern is argued to be neither sensitive to morphological nor to rhythmical considerations; word

stress (henceforth stress) is always final regardless of the suffixes attached to the stem as in *tanı* “know,” *tanı-dık*, “acquaintance,” *tanıdık-lár* “acquaintances” (Sezer, 1983). However, a number of exceptions have been reported, and one of them is associated with irregular roots. In contrast to regular roots that are stressed on the final syllable, irregular roots (e.g., place names) follow a quantity-sensitive rule (Sezer, 1983). According to this rule, stress placement is determined by the syllabic structure: Stress occurs on the antepenult if the penult is light and antepenult is heavy as in *Ánkara*; otherwise it falls on the penult as in *İstánbul*. Regular and irregular roots occasionally create minimal pairs, which are segmentally identical but different in prosodic features such as *bebék* “baby” and *bébek* “a district in İstanbul.” Given that prosodic features play a crucial role in lexical access (Cutler et al., 1997; Friedrich et al., 2004; Zora et al., 2015), even in languages where the number of word pairs distinguished by prosodic features alone is small (Cutler, 2005), this study investigates the impact of prosodic features in the activation of such segmentally identical Turkish pairs without any given context.

The opinions on the properties of Turkish stress have diverged widely in the literature (for an overview of Turkish stress in the first half of 20th century, see Excursus on Stress in Lees, 1961). Although most of the work on this topic has not made a distinction between stress- and pitch-accent to refer to the prominent syllable of a word, some authors have introduced this distinction and claimed Turkish to be a pitch-accent language rather than a stress-accent language¹ (Underhill, 1986; Levi, 2005). Other authors have claimed that there is a difference in the manifestation of final and non-final stress in Turkish; while stress is realized by a pitch-accent on the final syllable, it is realized by a stress-accent on the non-final syllable (Csató and Johanson, 1998; Johanson, 1998).

There are very few studies investigating the phonetic aspects of Turkish stress (Konrot, 1981; Levi, 2005; Pycha, 2006), and these studies have focused on the acoustic correlates rather than on the perceptual ones. These studies have investigated the production of Turkish stress by looking at another exception, which is associated with pre-stressing suffixes. In contrast to stressable suffixes, pre-stressing suffixes require that the stress falls on the preceding adjacent syllable. Konrot (1981), for instance, investigated the role of fundamental frequency (f_0), intensity, duration, and vowel quality in disyllabic minimal pairs created by stressable noun-making suffix—mA² and pre-stressing negative suffix—mA (*kazmá* “pickaxe” vs. *kázma* “do not dig”). Findings indicated only f_0 and intensity as correlates, and furthermore that the correlates were only used in non-final positions. Similarly, Levi (2005) investigated the role of f_0 , intensity and duration in minimal pairs created by stressable locative suffix—dA and pre-stressing instrumental suffix—lA (*metindé* “in the text” vs. *metínle* “with the text”) and by stressable infinitive suffix—mAk and pre-stressing negative suffix—mA (*gezmék* “to travel” vs.

gezmemek “not to travel”). Findings indicated that the stress was realized differently in final and non-final positions. A discriminant analysis confirmed that f_0 was the most robust correlate to mark stress, followed by intensity and then by duration. However, duration and intensity were argued to be less reliable than f_0 ; furthermore, it was claimed that duration is unlikely to be used in the perception of stressed syllables. Pycha (2006) examined acoustic correlates by comparing simplex nouns having final stress and their suffixed versions created by pre-stressing interrogative suffix—mI (*dedé* “grandfather” vs. *dedé mi* “grandfather?”). Here, stress had multiple correlates in both final and non-final positions. However, f_0 was argued to be the primary correlate of stress in both positions despite the non-robustness of f_0 in the final position. The differences in f_0 , intensity, duration, and vowel quality were all shown to be statistically significant. However, only f_0 was argued to function as a perceptual cue because the average differences in intensity and duration between stressed and unstressed positions were below the perceptual threshold of just-noticeable differences.

The studies reviewed above investigated the acoustic manifestation of Turkish stress by comparing final and non-final stress in complex words created by stressable and pre-stressing suffixes. The findings established f_0 as the most salient cue in marking stressed syllables. However, there was disagreement as to where and to what extent prosodic features are employed: While some identified both f_0 and intensity as correlates, but only in non-final position (Konrot, 1981), others identified f_0 , intensity and duration as correlates for both final and non-final positions, but argued that only f_0 can function as perceptual correlate (Pycha, 2006). It should be noted that these studies employed complex words and therefore failed to avoid morphological effects. Moreover, they were not independent from sentence-level prosody; since both pre-stressing negative suffix—mA and pre-stressing interrogative suffix—mI form a full sentence, they may introduce an imperative and a question contour, respectively.

To date, only one study of stress perception in Turkish has employed the ERP technique (Domahs et al., 2013). This study investigated the sensitivity to violations of predictable and unpredictable stress patterns. Monomorphemic trisyllabic words with three different stress patterns such as in *fiasco* /fiyásko/, */fiyasko/, and */fiyaskó/ were examined. In line with the acoustic findings above, the results indicated that final stress was processed differently from non-final stress. Different stress violations elicited different ERP components: While stress violations with final stress elicited an N400 component, violations with non-final stress produced a P300 component. It was argued that the application of the predictable default (final) pattern to the words with non-final stress resulted in higher costs in lexical processing. In contrast, the application of the unpredictable pattern to the words with final-stress resulted in the evaluation of this pattern. It was, accordingly, argued that Turkish speakers were less responsive to stress shifts that lead to final stress patterns than to stress shifts that resulted in non-final stress patterns. The study evaluated this finding with regard to the typology of stress-deafness, which suggests that native speakers of languages with predictable stress are less sensitive to stress

¹According to the phonetic criteria proposed in Beckman (1986), pitch-accent languages employ only fundamental frequency when marking the stressed syllable, while stress-accent languages may also use intensity and duration.

²Capital letters indicate that segment is unspecified and might vary according to vowel harmony.

variations than native speakers of languages with variable stress (Peperkamp and Dupoux, 2002). Thus, stress-deafness in Turkish was claimed to occur only for the final stress pattern. This finding is in agreement with the above-mentioned acoustic studies that found a weakening of acoustic correlates for final stress. However, neither the ERP study nor the acoustic studies were concerned with establishing perceptual correlates of stress in Turkish and their role in lexical access.

So far, no empirical research using the ERP technique has addressed perceptual correlates of stress in Turkish. Perceptual correlates and their role in lexical access can be examined using the mismatch negativity (MMN) component. The MMN is a neurophysiological measure that signals the brain's automatic response not only to any acoustic change in the auditory sensory input (e.g., changes in f_0 , intensity and duration) but also to higher cognitive processes such as the activation of long-term memory traces for lexical information (Dehaene-Lambertz, 1997; Näätänen et al., 1997, 2007; Näätänen and Winkler, 1999; Winkler et al., 1999; Näätänen, 2001; Pulvermüller et al., 2001; Shtyrov and Pulvermüller, 2002; Zora et al., 2015). The MMN is elicited irrespective of the subject's attention to the auditory stimulus and is based on an oddball paradigm. That is, the MMN is elicited when a rare stimulus (deviant) is interspersed among frequent stimuli (standard; Näätänen et al., 1978; Näätänen and Winkler, 1999; for a review, see Näätänen et al., 2007).

In the present study, we studied the neural responses to changes in f_0 , spectral emphasis, and duration in Turkish words that differ in stress but are identical in segmental structure. Spectral emphasis was used rather than overall intensity to better reflect the role of loudness in stress perception. Spectral emphasis is characterized by the relative intensity in the higher frequency bands and has been shown to be a more reliable correlate than overall intensity in both production and perception (Sluijter and van Heuven, 1996; Sluijter et al., 1997; Heldner, 2003). A pseudoword pair with both possible stress patterns was used as control. The pseudoword pair was included to enable a comparison between ERP correlates of prosodic cues on the basis of lexical processing and non-lexical processing. The neural responses were recorded in relation to the second syllable of words and pseudowords in an auditory oddball paradigm by presenting four deviants interspersed among standard stimuli. The deviants differed from the standard in (i) f_0 , (ii) spectral emphasis, and (iii) duration alone as well as in (iv) all features combined. It was hypothesized that the MMN would indicate if segmentally identical Turkish words are distinguished on the basis of these features alone, and that their salience and relevance in lexical access would be reflected in the amplitude of MMN responses.

METHODS

Participants

The participants were 14 native speakers of Turkish (10 males, four females; age range 20–54 years, $M = 30.8$, $SD = 7.8$), currently residing in Stockholm, Sweden. The general inclusion criteria were age of moving to Sweden (≥ 12 years, $M =$

24.7, $SD = 7.2$) and length of residence (≤ 15 years, $M = 6.1$, $SD = 5.1$). Handedness was assessed by the Edinburgh Handedness Inventory (Oldfield, 1971); all participants were right-handed. All of them reported normal development and hearing. The participants were rewarded with movie tickets for their participation. Informed consent was signed prior to testing, and the study was approved by the Stockholm Regional Ethics Committee (2015/63-31).

Materials and Manipulations

The experiment consisted of one word block and one pseudoword block. The material in the word block was a monomorphemic disyllabic Turkish minimal pair in which the location of stress on the first or second syllable led the word to be identified either as *a district in İstanbul* /bébek/ or as *a baby* /bebék/. The material in the pseudoword block was a pseudoword with both possible stress patterns /dedék/ and /dédek/. The stimuli were recorded in a semantically neutral frame sentence produced by a male native speaker of Turkish (from İstanbul, 30 years old) in an anechoic chamber, and were sampled at a rate of 44.1 kHz with 16 bits/per sample.

In order to keep the difference minimal across the blocks, the stimuli in the pseudoword block were created from the word block material by replacing the bilabial segment /b/ with dental segment /d/ in Praat (Boersma and Weenink, 2014)³. All stimuli were matched for duration (497 ms). In both word and pseudoword blocks, stimuli with iambic pattern always served as standards and stimuli with trochaic pattern as deviants. No manipulation was carried out for the standard; that is, the original realization of the iambic pattern was kept. The deviants (i.e., with trochaic pattern) were created out of the standards (i.e., with iambic pattern) by a cross-splicing technique. In order to get an equal ground for the comparison, it was important to ensure that the deviants were identical with the standard up to the onset of the second syllable; the first syllable /be/ in the word block and /de/ in the pseudoword block were therefore kept constant across standards and deviants. Then, the second syllable of the iambic pattern was spliced and acoustic manipulations were carried out. Since direction and amplitude of the acoustic change might influence the neural response, the second syllable was manipulated by lowering the acoustic parameters, which, in turn, lead to having iambic pattern always as standards and trochaic pattern as deviants. Given that word stress is about relations between syllables within words; that is prosodic features are characterized by a comparison of items in sequence (Lehiste, 1970), lowering the acoustic parameters of the second syllable in an iambic pattern should lead us to perceive it as a trochaic pattern.

Manipulation values were determined based on the proportions in the original trochaic pattern. By keeping the proportions between the first and second syllable in the trochaic pattern the same, the second vowel of the iambic pattern was manipulated by lowering (i) f_0 , (ii) spectral emphasis, (iii) duration, and (iv) all features, and four types of deviant stimuli were created. Manipulations did not lead to any lexical cost

³<http://www.praat.org>

TABLE 1 | Mean fundamental frequencies (f_0) in semitones (st) and durations (in ms) of the vowels with final and non-final stress, and deviant-manipulation results.

	bebék (standard)	bebék (original)	bebék (deviant)
/e/− σ 1 mean f_0 in st	85	89	85
/e/− σ 2 mean f_0 in st	87	83	79
Difference		6	6
/e/− σ 1 Duration in ms	71	78	72
/e/− σ 2 Duration in ms	72	63	57
Difference		15	15

in the word block since both the standard and the deviants were real words. As the deviants maintained the same relations (proportions) as the original trochaic words, the manipulations yielded natural sounding words.

The f_0 -values were taken from pitch tracks and were resynthesized in Praat, which uses an acoustic periodicity detection algorithm based on an autocorrelation method (Boersma, 1993). f_0 range settings were 75–600 Hz. Mean f_0 was measured over each vowel. The vowel onset and offset were determined based on the pitch information and were free from the contextual influences of surrounding segments. **Table 1** shows the f_0 measurements in semitones relative to 1 Hz (st) in both stressed and unstressed positions.

Since the first syllable was constant across the standards and deviants, f_0 of the first vowel was 85 st in both stimuli (**Table 1**). The mean f_0 difference between the stressed and unstressed vowels in the original word /bebék/ was 6 st (89 st-minus-83 st). In order to keep this proportion the same in the f_0 deviant, f_0 of the second vowel was set to 79 st (85 st-minus-6 st; **Table 1**). **Figure 1** shows the pitch track of the resynthesized version in Praat.

The duration manipulation was performed in the same manner as the f_0 manipulation. Vowel durations were measured in milliseconds (ms) and manipulated in Praat. Since the standard and deviants shared the same first syllable, duration of the first vowel is 72 ms in both stimuli (**Table 1**). The duration difference between the stressed and unstressed vowels in the original word /bebék/ is 15 ms (78 ms-minus-63 ms). In order to keep this proportion the same in the duration deviant, duration of the second vowel was set to 57 ms (72 ms-minus-15 ms; **Table 1**).

The spectral emphasis manipulation was performed in Adobe Audition CS6. By comparing the second syllables of iambic and trochaic patterns, spectral emphasis differences were measured. By using the fast Fourier transform filter (fft-filter), the spectrum above the f_0 was decreased 4 dB. The spectrum of each syllable is shown in **Figure 2**.

To create the fourth deviant, all features (f_0 , spectral emphasis, and duration) on the second vowel were manipulated at once.

Procedure

The experiment was run using E-Prime (Psychology Software Tools, Pittsburgh, PA, USA). The stimuli were delivered via

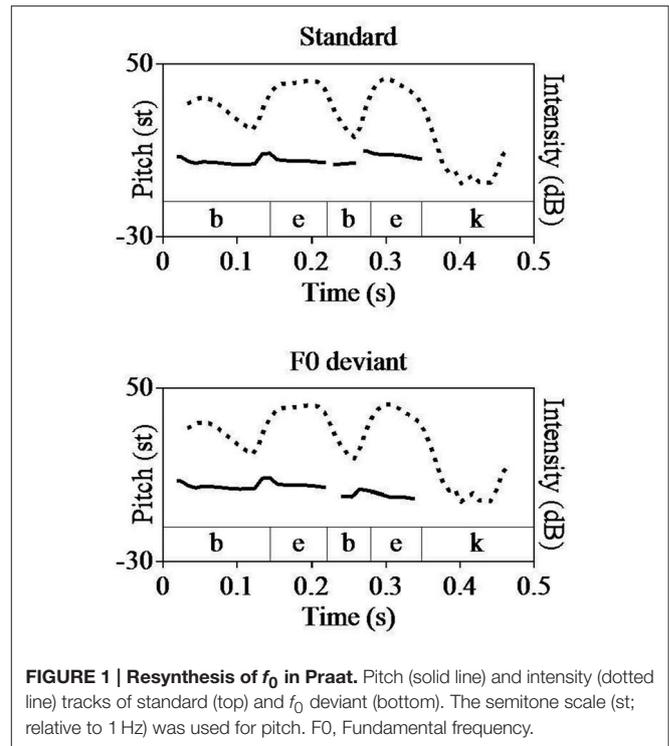


FIGURE 1 | Resynthesis of f_0 in Praat. Pitch (solid line) and intensity (dotted line) tracks of standard (top) and f_0 deviant (bottom). The semitone scale (st; relative to 1 Hz) was used for pitch. F0, Fundamental frequency.

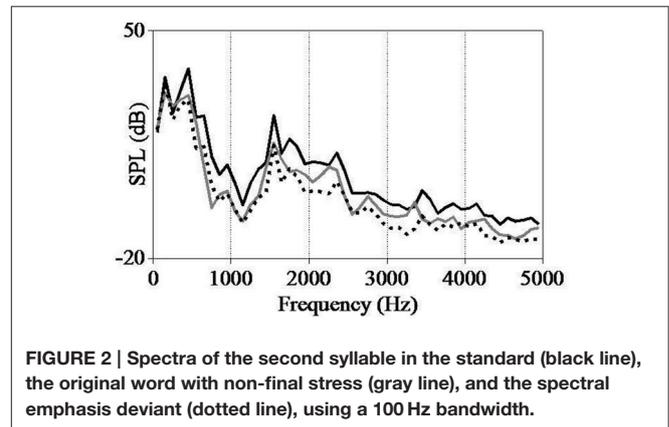


FIGURE 2 | Spectra of the second syllable in the standard (black line), the original word with non-final stress (gray line), and the spectral emphasis deviant (dotted line), using a 100 Hz bandwidth.

loudspeakers at a comfortable listening level of 60–65 dB at source. The stimuli were presented in a 5-stimulus auditory oddball paradigm (1 standard + 4 deviants). The frequently repeated standard stimulus ($p = 8/10$) was randomly replaced by four rare deviant stimuli ($p = 0.5/10$ each), with at least two intervening standards between two consecutive deviants. Of 2000 stimuli, the number of standards was 1600 and the number of deviants was 400 (100 for each). The stimulus onset asynchrony was set at 1000 ms. A silent documentary was used to take the participants' attention off the auditory stimuli. The whole experiment lasted about 1 h 10 min.

Electroencephalography Recordings

The electroencephalography (EEG) signals were recorded at a sampling rate of 250 Hz, using NetStation 4.4 with a Net Amps

300 amplifier (Electrical Geodesic Inc., EGI, Eugene, Oregon, USA). The recordings were made from HydroCel Geodesic Sensor Net of 128 electrodes (EGI, Eugene, Oregon, USA) which employs a non-abrasion high-impedance application method. The impedance was kept below 50 k Ω at each electrode site as specified by EGI for this high impedance system. An online band-pass filter with cut-off frequencies at 0.1 and 70 Hz was applied. The Cz electrode was used as online reference and the ground reference had a centroparietal location.

ERP Data Analysis

The EEG data was first filtered with a low-pass filter with a cutoff frequency of 30 Hz and with a high-pass filter with a cutoff frequency of 1 Hz. The channels were then re-referenced to both mastoids. The EEG data was segmented into epochs of 800 ms, time-locked to the onset of second syllable (200 ms before onset to 600 ms after onset). The onset of second syllable was used as the zero point in the data analysis because, given that the standards and deviants were same up to the second syllable, information about any difference between the standards and deviants would be present only after this point

(see Pulvermüller et al., 2001; Pulvermüller and Shtyrov, 2003). A time window of 200 ms prior to the onset was used for baseline correction. Artifact rejection was set to remove activity exceeding $\pm 100 \mu\text{V}$ at any channel. 60% of deviant trials had to be artifact-free in order for the retained material to form a valid base for the conclusions. A grand average was computed for each stimulus type for all participants and deviant-minus-standard subtraction signals were calculated for each deviant type.

Statistical Analysis

Statistical analysis was performed in SPSS (International Business Machines Corp., Armonk, NY, USA). The electrodes were grouped together in five regions of interest (ROI): left, right, frontal, central, and parietal. The **Figure 3** illustrates ROIs. The measurement window was determined by visual inspection of grand average difference waveforms. For MMN quantification, amplitudes were computed as a mean voltage within a 50-ms-window centered at the peak latency in the grand-average waveforms. Amplitude data extracted from deviant-standard subtraction curves was used for statistical analysis (see Näätänen

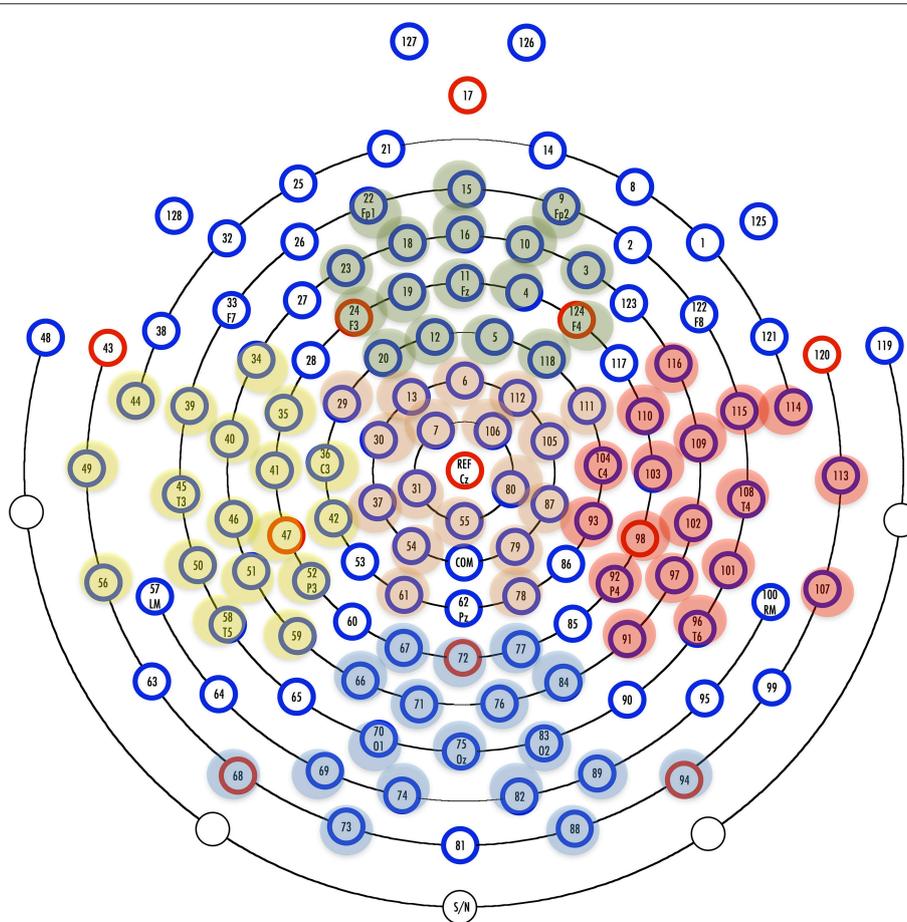


FIGURE 3 | 128-channel HydroCel GSN and Regions of Interest (ROIs). Yellow marked electrodes, Left hemisphere; red marked electrodes, Right hemisphere; green marked electrodes, Frontal sites; orange marked electrodes, Central sites; Blue marked electrodes, Parietal sites.

and Winkler, 1999; Näätänen et al., 2007; Winkler, 2007; Kappenman and Luck, 2012).

A Three-way repeated-measures ANOVA with factors of Lexicality (two levels: word and pseudoword), Prosody (four levels: f_0 , spectral emphasis, duration, and all combined), and ROI (five levels: left, right, frontal, central, and parietal) was performed. If significant interactions occurred, follow-up ANOVAs were performed and the levels were then compared in *post-hoc* pairwise comparisons. SPSS Bonferroni adjusted *p*-values are reported for *post-hoc* comparisons. Additional two-tailed *t*-tests were used to look at topographical distributions closer. *P*-values are given with Greenhouse-Geisser correction in case of sphericity violations. Effect sizes are reported with η^2 (partial η^2).

RESULTS

Event-Related Potential Data

The grand average difference waves and scalp topographies for deviants are shown for word and pseudoword blocks in Figures 4, 5, respectively.

Changes in prosodic features elicited neural responses at around 230 ms after change onset (i.e., onset of second syllable) in both words (Figure 4, top) and pseudowords (Figure 4, bottom), confirming the brain's automatic response to any change in auditory sensory input. The processing differences between words and pseudowords were, on the other hand, reflected in the amplitude, polarity, and topography of the neural responses.

While eliciting a small negativity in words (Figure 5, top, left), the spectral emphasis deviant elicited a large negativity in

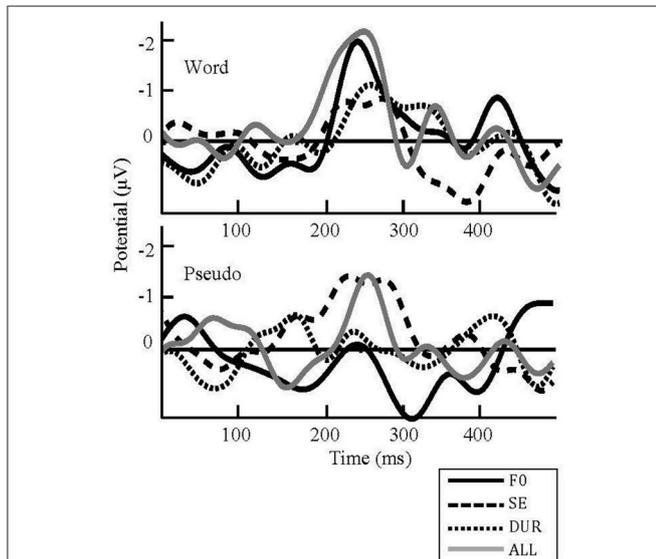


FIGURE 4 | Grand average difference waveforms for all deviants on FZ in word (top) and pseudoword (bottom) blocks. F0, Fundamental frequency; SE, Spectral emphasis; DUR, Duration; ALL, Fundamental frequency, spectral emphasis and duration combined. Negativity is plotted upward.

pseudowords (Figure 5, bottom, left). These negativities were considered to be N200 responses. There are two negative components in the time range of N200: N2a or MMN and N2b (Näätänen and Gaillard, 1983; Näätänen, 1992). In contrast to the MMN, which reflects automatic processing and is elicited in unattended conditions, N2b is usually elicited in attended conditions (for a review, see Näätänen and Gaillard, 1983). However, when a stimulus deviation is wide, the N2b may also occur in unattended conditions (Näätänen et al., 1982). The presence of N2b has been indicated with a broad scalp distribution similar to that of auditory N100 (Näätänen, 1992). The spectral emphasis-related negativity in words was considered to be an MMN response whereas it was considered to be an N2b response in pseudowords due to its distribution in a wider area and due to the nature of the stimuli. This will be further discussed in Section Conclusion.

The most noticeable processing difference between words and pseudowords seems to be the f_0 -related activation; while eliciting a right-lateralized frontally-maximal negativity in words (Figure 5, top, second from left), the f_0 manipulations elicited a frontal positivity in pseudowords. This f_0 -related positivity in pseudowords was considered to be a P300 response. If categorized according to their functional correlates and scalp distributions, P300 is typically divided into two subcomponents: P3a and P3b (Squires et al., 1975; Linden, 2005; Patel and Azzam, 2005; Polich, 2007). The P3a shows a frontally maximum scalp distribution and indexes the orienting of attention to unexpected events. The P3b shows a parietally maximum scalp distribution and indexes the updating of working memory. While the P3b is task relevant, the P3a is elicited without a task (Squires et al., 1975; Linden, 2005; Patel and Azzam, 2005; Polich, 2007). The positivity in pseudowords was argued to be a P3a response (Figure 5, bottom, second from left) due to task-independent elicitation and a frontal scalp distribution.

Although not so prominent, a duration-related frontally distributed negativity was present in both words (Figure 5, top, third from left) and pseudowords (Figure 5, bottom, third from left). The all-combined deviant elicited a fronto-centrally

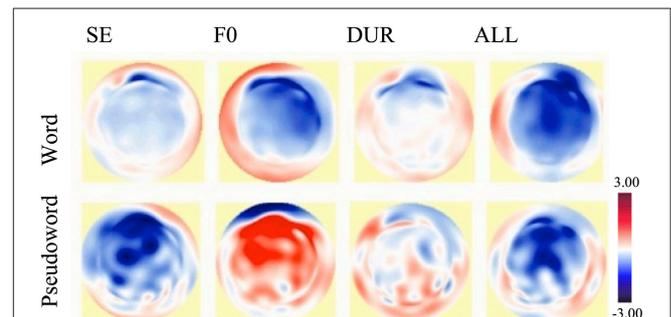
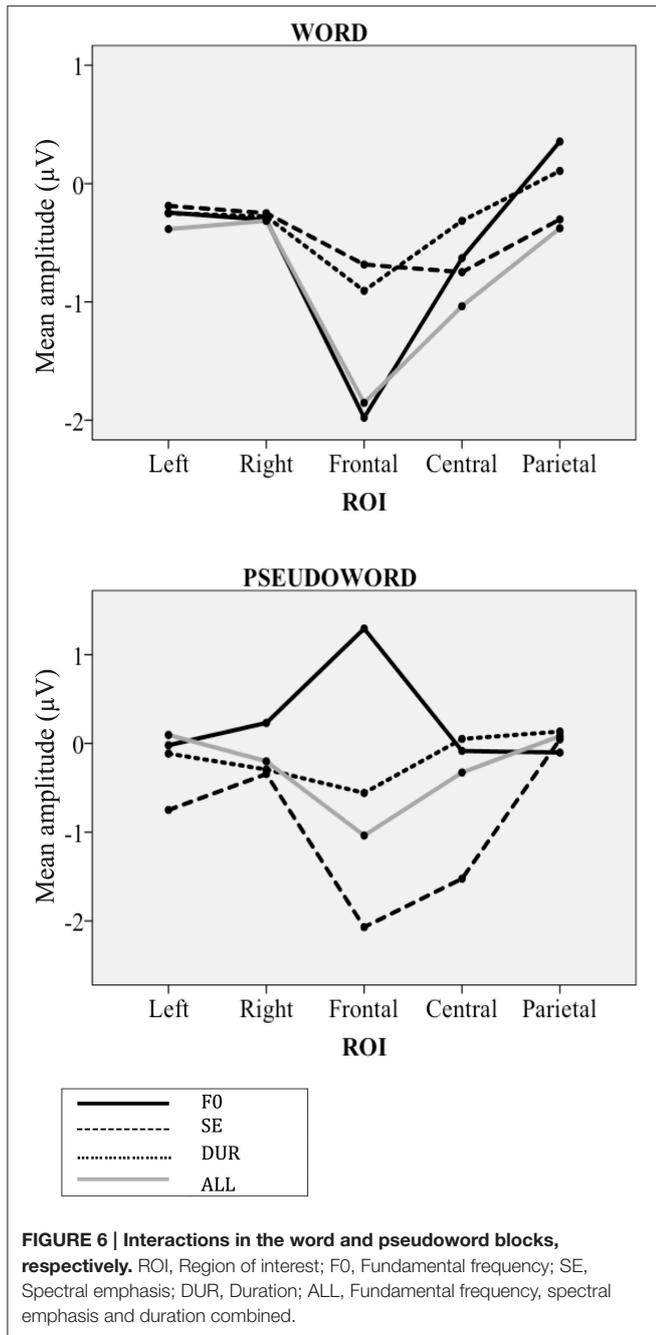


FIGURE 5 | Topographic difference maps at 240 ms after change onset for the word block (top) and for the pseudoword block (bottom): MMN for SE, for F0, for DUR, and for ALL, respectively, (from left to right). MMN, Mismatch negativity; F0, Fundamental frequency; SE, Spectral emphasis; DUR, Duration; ALL, Fundamental frequency, spectral emphasis and duration combined.



that only f_0 elicited positivity at frontal sites; that duration did not elicit any negativity at central sites, and that the negativity for spectral emphasis was still prominent at central sites.

Additional two-tailed t -tests were used to compare the amplitudes obtained for deviants in words against those in pseudowords in each ROI (Table 3). In the right hemisphere, f_0 elicited negativity in words ($M = -0.306 \mu\text{V}$, $SD = 0.150$) while eliciting positivity in pseudowords ($M = 0.231 \mu\text{V}$, $SD = 0.225$, $p = 0.011$). In frontal sites, f_0 elicited negativity in words ($M = -1.978 \mu\text{V}$, $SD = 0.813$) while eliciting

TABLE 3 | Two-tailed t -tests with the amplitudes obtained for deviants in words and pseudo words in each ROI.

Prosody	Lexicality level	p
LEFT		
f_0	Word-pseudo	0.244
SE	Word-pseudo	0.055
DUR	Word-pseudo	0.592
ALL	Word-pseudo	0.052
RIGHT		
f_0	Word-pseudo	0.011*
SE	Word-pseudo	0.717
DUR	Word-pseudo	0.916
ALL	Word-pseudo	0.591
FRONTAL		
f_0	Word-pseudo	0.001*
SE	Word-pseudo	0.098
DUR	Word-pseudo	0.739
ALL	Word-pseudo	0.512
CENTRAL		
f_0	Word-pseudo	0.074
SE	Word-pseudo	0.044*
DUR	Word-pseudo	0.351
ALL	Word-pseudo	0.085
PARIETAL		
f_0	Word-pseudo	0.056
SE	Word-pseudo	0.207
DUR	Word-pseudo	0.928
ALL	Word-pseudo	0.226

f_0 , Fundamental frequency; SE, Spectral emphasis; DUR, Duration; ALL, Fundamental frequency, spectral emphasis and duration combined. * $p < 0.05$.

positivity in pseudowords ($M = 1.294 \mu\text{V}$, $SD = 0.504$, $p = 0.001$). In central sites, spectral emphasis elicited larger negativity in pseudowords ($M = -1.524 \mu\text{V}$, $SD = 0.309$) than in words ($M = -0.748 \mu\text{V}$, $SD = 0.231$, $p = 0.044$). f_0 -related MMN activation was right-lateralized in words. f_0 -related difference between words and pseudowords was maximal over frontal sites. The spectral emphasis-related activation was more pronounced over central sites in pseudowords in comparison to words.

DISCUSSION

By recording neural responses to changes in fundamental frequency, spectral emphasis and duration, the present study investigated for the first time to what extent each perceptual correlate of stress is utilized for lexical access in Turkish. It was predicted that the neural responses would indicate differences in the contributions of fundamental frequency, spectral emphasis and duration in stress perception and lexical access. Given that studies on Turkish stress have so far investigated only the phonetic correlates in acoustic studies (Konrot, 1981; Levi, 2005; Pycha, 2006), this study makes an important contribution to the literature by investigating the perceptual correlates with an electrophysiological approach. In contrast to

the previous acoustic studies, which employed complex words and therefore failed to avoid morphological effects, the present study investigated the perceptual correlates of Turkish stress using simplex (monomorphemic) words. It is also worth noting that the present study used a spectral emphasis measure to better assess the importance of loudness considering that the contribution of higher frequency bands to the perceived intensity is much greater (Sluijter and van Heuven, 1996; Sluijter et al., 1997; Heldner, 2003). It is therefore the first study to introduce spectral emphasis measures to Turkish word stress and to an ERP study.

The findings indicated that the MMN responses were maximal over frontal and central scalp locations in line with the previous findings (Näätänen and Winkler, 1999; Näätänen et al., 2007). The findings further indicated that changes in prosodic features elicited neural responses at around 230 ms after change onset in both words and pseudowords, confirming the brain's automatic response to any change in auditory sensory input which typically peaks at 150–250 ms from change onset (Näätänen et al., 2007).

However, there were processing differences of prosodic information between words and pseudowords. First, while eliciting a frontal MMN in words, spectral emphasis manipulations elicited a fronto-centrally maximal N2b in pseudowords. The presence of N2b in pseudowords was indicated by its topography, which usually shows a broad scalp distribution somewhat similar to that of an auditory N100 (Näätänen, 1992). Apart from the topographical factor, another factor favoring the N2b-effect interpretation was that the negativity had a larger amplitude in pseudowords: If it were an MMN, reflecting an acoustic processing, then the amplitude of negativity would be the same in words and pseudowords since they are identical in their acoustic features; if it were an MMN, reflecting lexical processing, then the amplitude of negativity would be larger in words than in pseudowords. However, negativity was larger in the amplitude of pseudowords. Given that the presence of N2b has been suggested to be a result of stimulus-directed attention (Näätänen et al., 2007), negativity to spectral emphasis manipulations in pseudowords might be simply due to a larger attentional load, indicating that the spectral emphasis change was rather unexpected in the absence of lexical processing. The spectral emphasis might, therefore, play a crucial role in lexical processing. Given that this study is the first to introduce spectral emphasis measures into ERP, previous studies cannot explain the current results; the role of spectral emphasis in lexical processing is therefore a subject for further ERP research.

Second, while eliciting a frontally maximal MMN in words, fundamental frequency manipulations elicited a frontal positivity in pseudowords. This fundamental frequency-related positivity in pseudowords was considered to be a P3a response since P3a shows a task-independent, frontally maximum activation (Squires et al., 1975; Linden, 2005; Patel and Azzam, 2005; Polich, 2007). Alternatively, one could argue positivity to be a P200 response as a reflection of pitch sensitivity. The peak latency of P200 is at about 200 ms and the scalp distribution is typically noted to be maximal over central regions (Näätänen, 1990). However, the positivity here is later in latency and shows a frontal

distribution. Moreover, the positivity elicited here is believed to be not just a reflection of pitch change but rather an attention orientation. It is therefore argued to be a P3a response. Given that P3a reflects involuntary allocation of attention to salient changes (Squires et al., 1975; Näätänen, 1990; Linden, 2005; Polich, 2007), this fundamental frequency-related P3a in pseudowords could be due to attention orientation to pitch change. In contrast to the MMN that indexes lexical as well as acoustic processing, P3a activation is limited to physical parameters and reflects only the acoustic processing. The fact that fundamental frequency changes elicited an MMN in words whereas eliciting a P3a in pseudowords suggests therefore that fundamental frequency is lexically specified in Turkish.

Although not as prominent, duration-related frontally distributed MMN was present in both words and pseudowords. Duration is therefore claimed to be present as an acoustic cue for word stress perception although it does not make a significant contribution to automatic lexical access. This is not surprising given that the vowel length is not phonemically distinctive in Turkish (Kornfilt, 1997; Nimz, 2015). Previous acoustic studies (Konrot, 1981; Levi, 2005; Pycha, 2006) ruled out duration as an acoustic correlate in Turkish word stress. In this regard, this study makes an important contribution to the literature by indicating duration as a perceptual correlate. However, although being a potential perceptual correlate, the role of duration in automatic lexical access could not be established.

The deviant combining fundamental frequency, spectral emphasis, and duration elicited negativity in both words and pseudowords. Combined with previous findings, this suggests that it is individual prosodic features that differentiate lexical processing from non-lexical processing rather than combined features. This, further, provides answer to the question whether prosodic features are processed separately or holistically in the auditory system. A number of MMN studies indicated that acoustic features are represented separately, indicating independent storage of these features (for a review see Caclin et al., 2006). Separate processing of acoustic features has been favored by studies of MMN generation location (Giard et al., 1995; Rosburg, 2003). This argument was further supported by studies that showed additivity of responses to single deviants; that is, the MMN elicited by multiple deviants can be predicted by the sum of the MMNs of the corresponding single deviants (Wolff and Schröger, 2001). However, deviants with three deviating features might not follow the MMN additivity argument and might elicit a smaller MMN than predicted by the sum of the single deviants, indicating complex processing of simultaneously deviating features (Paavilainen et al., 2001). MMN amplitude is argued to index the probability of a specific feature rather than the probability of feature combinations (Deacon et al., 1998). In line with these findings, present results support a model where acoustic features are processed separately.

The brain does not only detect prosodic changes but also uses them in lexicality decisions; there is a difference in how the brain treats prosodic cues depending on lexicality. Prosodic changes consistently elicit MMN in words while eliciting different components in pseudowords depending on the deviating cue. If the MMN responses were to indicate pre-lexical processing

only, then the neural responses would be the same in words and pseudowords since they are identical in their acoustic features. Different components in pseudowords reflect acoustic-change effects on the neural responses rather than language-related effects. Prosodic feature manipulations, in the absence of lexical processing, trigger an increase in attentional load and cause reevaluation of what was heard.

In summary, the present study is the first to demonstrate the neural correlates of prosodic features in Turkish word stress and their contribution to lexical access. The findings indicate that there are memory traces for prosodic information in the brain and that they play a significant role in lexical access without any given context. The findings show that segmentally identical Turkish words can indeed be distinguished on the sole basis of prosodic features. In line with the previous acoustic studies, which found fundamental frequency as being the most salient cue in marking stressed syllables (Konrot, 1981; Levi, 2005; Pycha, 2006), fundamental frequency has been found as the most prominent perceptual correlate. This finding is potentially in agreement with the phonetic criterion, which claims Turkish to be a pitch-accent language rather than a stress-accent language (Underhill, 1986; Levi, 2005). Apart from confirming the perceptual salience of fundamental frequency, the present study also indicates the contribution of fundamental frequency in lexical access and therefore argues that fundamental frequency is lexically specified in Turkish. It should again be noted that the role of spectral emphasis in lexical access is subject to further research. In short, fundamental frequency is argued to be the most prominent perceptual correlate and lexically specified due to the fact that it elicited the largest negativity in words, and that the most remarkable processing difference

between words and pseudowords was related to fundamental frequency.

CONCLUSION

Perceptual correlates of Turkish word stress and their contribution to lexical access were studied. Neural responses to changes in fundamental frequency, spectral emphasis and duration were recorded in Turkish words and pseudowords. The findings indicate that memory traces for Turkish words are indeed activated on the sole basis of prosodic information. The presence of prosodic representations in lexical representations has a differential effect on the processing of prosodic changes. The manipulations of prosodic features, in the absence of lexical processing, increase the attentional load and cause reevaluation of the auditory stimuli. Fundamental frequency as the most salient perceptual correlate contributes to lexical access and is therefore lexically specified in Turkish.

AUTHOR CONTRIBUTIONS

HZ, MH, IS: The conception and design of the work; the analysis of the data; drafting the work and revisiting it critically; final approval of the version to be published.

ACKNOWLEDGMENTS

The authors thank Anders Eriksson and Hartmut Traunmüller for their feedback and the reviewers for their valuable comments. This work was supported by Forskarskolan i Språkvetenskap (FoSpråk), Stockholm University.

REFERENCES

- Barker, C. (1989). "Extrametricity, the cycle, and Turkish word stress," in *Phonology at Santa Cruz 1*, eds J. Itó and J. Runner (Santa Cruz, CA: Syntax Research Center), 1–33.
- Beckman, M. (1986). *Stress and Non-Stress Accent*. Dordrecht: Foris.
- Boersma, P. (1993). "Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound," in *Proceedings of the Institute of Phonetic Sciences*, Vol. 17 (Amsterdam: University of Amsterdam), 97–110.
- Boersma, P., and Weenink, D. (2014). *Praat: Doing Phonetics by Computer (Version 5.3.24)*. Amsterdam. (Accessed September 1, 2014).
- Caclin, A., Brattico, E., Tervaniemi, M., Näätänen, R., Morlet, D., Giard, M. H., et al. (2006). Separate neural processing of timbre dimensions in auditory sensory memory. *J. Cogn. Neurosci.* 18, 1959–1972. doi: 10.1162/jocn.2006.18.12.1959
- Csató, É. Á., and Johanson, L. (1998). "Turkish," in *The Turkic languages*, eds L. Johanson and É. Á. Csató (London; New York, NY: Routledge), 203–235.
- Cutler, A. (2005). "Lexical stress," in *The Handbook of Speech Perception*, eds D. B. Pisoni and R. E. Remez (Oxford: Blackwell), 264–289.
- Cutler, A., Dahan, D., and van Donselaar, W. (1997). Prosody in the comprehension of spoken language: a literature review. *Lang. Speech* 40, 141–201.
- Deacon, D., Nousak, J. M., Pilotti, M., Ritter, W., and Yang, C. M. (1998). Automatic change detection: does the auditory system use representations of individual stimulus features or gestalts? *Psychophysiology* 35, 413–419. doi: 10.1111/1469-8986.3540413
- Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *Neuroreport* 8, 919–924. doi: 10.1097/00001756-199703030-00021
- Domahs, U., Genc, S., Knaus, J., Wiese, R., and Kabak, B. (2013). Processing (un-) predictable word stress: ERP evidence from Turkish. *Lang. Cogn. Process.* 28, 335–354. doi: 10.1080/01690965.2011.634590
- Friedrich, C. K., Kotz, S. A., Friederici, A. D., and Alter, K. (2004). Pitch modulates lexical identification in spoken word recognition: ERP and behavioral evidence. *Cogn. Brain Res.* 20, 300–308. doi: 10.1016/j.cogbrainres.2004.03.007
- Giard, M. H., Lavikainen, J., Reinikainen, K., Perrin, F., Bertrand, O., Pernier, J., et al. (1995). Separate representation of stimulus frequency, intensity, and duration in auditory sensory memory: an event-related potential and dipole-model analysis. *J. Cogn. Neurosci.* 7, 133–143. doi: 10.1162/jocn.1995.7.2.133
- Hameed, J. (1985). Lexical phonology and morphology of Modern Standard Turkish. *Cah. Linguist. Ottawa* 14, 71–95.
- Heldner, M. (2003). On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish. *J. Phon.* 31, 39–62. doi: 10.1016/S0095-4470(02)00071-2
- Inkelas, S., and Orgun, C. O. (1998). "Level (non)ordering in recursive morphology: evidence from Turkish," in *Morphology and its Relation to Phonology and Syntax*, eds S. G. Lapointe, D. K. Brentari, and P. M. Farrell (Stanford, CA: CSLI), 360–410.
- Inkelas, S., and Orgun, C. O. (2003). Turkish stress: a review. *Phonology* 20, 139–161. doi: 10.1017/S0952675703004482
- Johanson, L. (1998). "The structure of Turkic," in *The Turkic Languages*, eds L. Johanson and É. Á. Csató (London; New York, NY: Routledge), 30–66.

- Kabak, B., and Vogel, I. (2001). The phonological word and stress assignment in Turkish. *Phonology* 18, 315–360. doi: 10.1017/S0952675701004201
- Kappenman, E. S., and Luck, S. J. (2012). “ERP components: the ups and downs of brainwave recordings,” in *The Oxford Handbook of ERP Components*, eds S. J. Luck and E. S. Kappenman (New York, NY: Oxford University Press), 3–30.
- Konrot, A. (1981). Physical correlates of linguistic stress in Turkish. *Univ. Essex Lang. Cent. Occas. Papers* 24, 26–53.
- Kornfilt, J. (1997). *Turkish*. London: Routledge.
- Lees, R. (1961). *The Phonology of Modern Standard Turkish*. Bloomington, IL: Indiana University Press.
- Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.
- Levi, S. V. (2005). Acoustic correlates of lexical accent in Turkish. *J. Int. Phon. Assoc.* 35, 73–97. doi: 10.1017/S0025100305001921
- Linden, D. E. (2005). The p300: where in the brain is it produced and what does it tell us? *Neuroscientist* 11, 563–576. doi: 10.1177/1073858405280524
- Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event related potentials and other brain measures of cognitive function. *Behav. Brain Sci.* 13, 201–288. doi: 10.1017/S0140525X00078407
- Näätänen, R. (1992). *Attention and Brain Function*. Hillsdale, MI: Lawrence Erlbaum Associates.
- Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology* 38, 1–21. doi: 10.1111/1469-8986.3810001
- Näätänen, R., and Gaillard, A. W. K. (1983). “The orienting reflex and the N2 deflection of the event-related potential (ERP),” in *Tutorials in Event Related Potential Research: Endogenous Components*, eds A. W. K. Gaillard and W. Ritter (Amsterdam: North-Holland Publishing Company), 119–141.
- Näätänen, R., Gaillard, A. W. K., and Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychol. (Amst)*. 42, 313–329. doi: 10.1016/0001-6918(78)90006-9
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huottilainen, M., Iivonen, A., et al. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385, 432–434. doi: 10.1038/385432a0
- Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118, 2544–2590. doi: 10.1016/j.clinph.2007.04.026
- Näätänen, R., Simpson, M., and Loveless, N. E. (1982). Stimulus deviance and evoked potentials. *Biol. Psychol.* 14, 53–98. doi: 10.1016/0301-0511(82)90017-5
- Näätänen, R., and Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychol. Bull.* 125, 826–859. doi: 10.1037/0033-2909.125.6.826
- Nimz, K. (2015). “The Perception of vowel quality and quantity by Turkish learners of German as a foreign language,” in *The Segment in Phonetics and Phonology*, eds E. Raimy and C. E. Cairns (Hoboken, NJ: Wiley-Blackwell), 253–266.
- Oldfield, R. C. (1971). The assessment and analysis of handedness. *Neuropsychologia* 9, 91–113. doi: 10.1016/0028-3932(71)90067-4
- Paavilainen, P., Valppu, S., and Näätänen, R. (2001). The additivity of the auditory feature analysis in the human brain as indexed by the mismatch negativity: 1+1~2 but 1+1+1<3. *Neurosci. Lett.* 301, 179–182. doi: 10.1016/S0304-3940(01)01635-4
- Patel, S. H., and Azzam, P. N. (2005). Characterization of N200 and P300: selected studies of the event-related potential. *Int. J. Med. Sci.* 2, 147–154. doi: 10.7150/ijms.2.147
- Peperkamp, S., and Dupoux, E. (2002). “A typological study of stress deafness,” in *Laboratory Phonology*, eds C. Gussenhoven and N. Warner (Berlin: Mouton de Gruyter), 203–240.
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019
- Pulvermüller, F., Kujala, T., Shtyrov, Y., Simola, J., Tiitinen, H., Alku, P., et al. (2001). Memory traces for words as revealed by the mismatch negativity. *Neuroimage* 14, 607–616. doi: 10.1006/nimg.2001.0864
- Pulvermüller, F., and Shtyrov, Y. (2003). Automatic processing of grammar in the human brain as revealed by the mismatch negativity. *Neuroimage* 20, 159–172. doi: 10.1016/S1053-8119(03)00261-1
- Pycha, A. (2006). *A Duration-Based Solution to the Problem of Stress Realization in Turkish*. UC Berkeley Phonology Lab Annual Report. 141–151.
- Rosburg, T. (2003). Left hemispheric dipole locations of the neuromagnetic mismatch negativity to frequency, intensity and duration deviants. *Cogn. Brain Res.* 16, 83–90. doi: 10.1016/S0926-6410(02)00222-7
- Sezer, E. (1983). On non-final stress in Turkish. *J. Turk. Stud.* 5, 61–69.
- Shtyrov, Y., and Pulvermüller, F. (2002). Neurophysiological evidence of memory traces for words in the human brain. *Neuroreport* 13, 521–525. doi: 10.1097/00001756-200203250-00033
- Sluijter, A. M. C., and van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *J. Acoust. Soc. Am.* 100, 2471–2485. doi: 10.1121/1.417955
- Sluijter, A. M. C., van Heuven, V. J., and Pacilly, J. J. A. (1997). Spectral balance as a cue in the perception of linguistic stress. *J. Acoust. Soc. Am.* 101, 503–513. doi: 10.1121/1.417994
- Squires, N. K., Squires, K. C., and Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalogr. Clin. Neurophysiol.* 38, 387–401. doi: 10.1016/0013-4694(75)90263-1
- Underhill, R. (1986). “Turkish,” in *Studies in Turkish Linguistics*, eds D. Slobin and K. Zimmer (Amsterdam; Philadelphia, PA: John Benjamins Publishing Company), 7–21.
- Winkler, I. (2007). Interpreting the mismatch negativity. *J. Psychophysiol.* 21, 147–163. doi: 10.1027/0269-8803.21.34.147
- Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., et al. (1999). Brain responses reveal the learning of foreign language phonemes. *Psychophysiology* 36, 638–642. doi: 10.1111/1469-8986.3650638
- Wolff, C., and Schröger, E. (2001). Human pre-attentive auditory change-detection with single, double, and triple deviations as revealed by mismatch negativity additivity. *Neurosci. Lett.* 311, 37–40. doi: 10.1016/S0304-3940(01)02135-8
- Zora, H., Schwarz, I.-C., and Heldner, M. (2015). Neural correlates of lexical stress: MMN reflects fundamental frequency and intensity. *Neuroreport* 26, 791–796. doi: 10.1097/wnr.0000000000000426

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Zora, Heldner and Schwarz. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.