Event-Related Potential investigation of Initial Accent processing in French

Marion Aguilera[1], *Radouane El Yagoubi*[2], *Robert Espesser*[3], *Corine Astésano*[1,3]

[1] U.R.I Octogone-Lordat (E.A. 4156), Université de Toulouse, UTM, Toulouse, France
[2] Laboratoire CLLE-LTC (UMR 5263), Université de Toulouse, UTM, Toulouse, France
[3] Laboratoire Parole & Langage (UMR 7309), Aix-Marseille Université, Aix-en-Provence, France

marion.aguilera@univ-tlse2.fr, yagoubi@univ-tlse2.fr, Robert.Espesser@lpl-aix.fr, astesano@univ-tlse2.fr

Abstract

This study investigates stress processing through the Event-Related brain Potential (ERP) technique. It aims at evaluating whether French listeners can perceive and discriminate the Initial Accent (IA) and whether IA is encoded in the phonological representation. Participants listened to trisyllabic words in two stress-pattern conditions, with (+IA) or without (-IA) initial accenting, in an oddball paradigm. The EEG was recorded in both a passive and an active listening task, and in two different oddball versions: one where standard stimuli were +IA words and deviants -IA words, and the reverse for the other version (-IA standard, +IA deviant). Behavioral results show faster processing and less errors for +IA stimuli. ERP results show larger MisMatch Negativity component for -IA words, pointing out 1) that French listeners are sensitive to *f0* manipulation, and 2) that +IA is the preferred stress template in French. Altogether, our results indicate that French listeners not only discriminate stress patterns but that IA is encoded in long-term memory, hence phonologically relevant.

Index Terms: stress-pattern processing, initial accent, MisMatch Negativity, French

1. Introduction

French prosody presents some peculiarities, which makes its phonological description difficult to fully apprehend and which has consequences on effective propositions for speech processing models. Traditionally, descriptions of the French accentuation system account for a primary Final Accent (FA) co-occurring with intonational boundaries, and a secondary, optional Initial Accent (IA), essentially seen as an extra-metrical phenomenon [1; 2; 5]. Because accentuation (FA) in French is post-lexical and not lexically distinctive, it has been described as a 'language without accent' [1] or a 'boundary language' [2; 3]. In this view, it is also said that French listeners are 'deaf' to accentuation and have no long-term memory representation of stress patterns [4; *inter alia*]. Although FA is now undisputedly seen as a pitch accent [5], IA's phonological status is still unclear. While some models describe it as a pitch accent [6], most models in the frame of the AutoSegmental Metric approach consider IA as a 'loose boundary marker', which peak can be aligned to the first, second or even third syllable of the content word or accentual phrase in some cases [5; 7]. As such, IA is not clearly phonologically implemented in the French prosodic system: its role is that of a secondary, rhythmical balancing device, yielding to FA in case of tonal crowding [5; 8].

However, diverse studies point out systematic use of IA in prosodic structure marking and speech segmentation. Recently for example, IA has been shown to be a more reliable cue to prosodic structuring than FA in the marking of content words and accentual phrases [9]. Namely, IA very commonly marks lexical words inside larger prosodic units, more so than FA, and is a marker of left prosodic boundaries even when in close vicinity to FA and not rhythmically necessary. Other studies indicate that the cohesive prosodic units formed by IA and FA, also described as 'accentual arches' [10] strongly enhance the encoding and segmentation of linguistic units, over units marked by FA alone [11; 12]. Finally, IA ('early rise') is used to segment lexical units in the speech flow [13]. Despite potential 'looseness' of this 'early rise', naïve French listeners systematically perceive IA *on the first syllable* of content words and accentual phrases, even when its peak reaches its maximum further in the unit [14]. This may indicate that IA is strongly linked to the representation of lexical words. Altogether, these production and perception findings point out a more important role of IA than described in prosodic models. Namely, they are in keeping with a possible strong phonological role of IA as a left marker of small units, i.e. accentual phrases and possibly the level of the lexical word.

It is thus interesting to further investigate whether IA is *encoded* at some level of the linguistic representation. The Metrical Segmentation Strategy [15] states that the mental lexicon is accessed through pre-lexical, language-specific stress templates. Finding neural bases for IA processing in French would give further insights as to its phonological role in French. Very few studies have to date investigated the neural bases of accentuation, and mostly do so through the exploration of Event-Related Potential (ERP) components. Stress pattern violations in Dutch in an AX paradigm elicited a N325 component indicating pre-lexical processing of stress [16]. In French, violation of metrical patterns induced by artificial lengthening of the medial syllable in trisyllabic words induced delayed lexical decision. Indeed, lexically *congruent* words in the sentence elicited a larger N400 component when metrically incongruent, indicating that French listeners do have pre-lexical stress template representations [17]. Similar results were found for Chinese [18]. Other results using the oddball paradigm allow investigation of short-term and long-term memory processing of stress patterns, through the MisMatch Negativity ERP component [19; 20; 21].

The oddball paradigm allows presenting listeners with deviant stimuli in a background of standard stimuli, which processing elicits an early negativity in the context. The MisMatch Negativity (MMN) is thus a reconstructed ERP component obtained by subtracting deviants' elicited brainwaves from standards'. The oddball paradigm allows measuring the mnemonic trace that the repetition of standard stimuli leaves in short-term memory [22]. A MMN component emerges on the deviant stimuli when the standard stimuli leave a strong mnemonic trace in short-term memory. This effect is

10.21437/SpeechProsody.2014-63

emphasized by linguistic experience of phonological representations in the native language, i.e. by comparing the auditory stimuli with those representations in the long-term memory. Thus, while the oddball paradigm helps determining whether an auditory stimulus is perceptually discriminated, it can also help reveal whether this discrimination goes beyond lower processing levels [23]. In its pre-attentive version (*passive listening condition*), the MMN allows examining how the brain processes linguistic information online, without requiring listeners to make conscious decisions about the speech stimuli. Hence, results can be interpreted independently from other linguistic levels.

While most studies using MMN manipulate stress patterns' legality in words and pseudo-words [20; 21], the first step of our investigation concerns IA processing at the word level. By manipulating presence or absence of IA on words, our experiment is designed to test 1) whether French listeners *can* perceive and discriminate IA (acoustical low-level processing or salience discrimination), and 2) whether IA is encoded at the word-level (phonological, higher cognitive level or stress template processing). Indeed, although words can sometimes be pronounced without systematic IA in the speech flow, we hypothesize that IA stress patterns are the preferred, expected stress templates in French. Hence, we predict that ± IA stimuli are discriminated (short-term memory) and that a larger MMN will be elicited by -IA deviant patterns in a +IA standards background because they are less expected stress patterns (long-term memory stress pattern representation).

2. Method

2.1. Speech stimuli

The stimuli consisted of two words (*candidat* ('candidate') and *diffusion* ('broadcast')) that were naturally uttered with **IA (+IA)** in a sentence context by a naïve speaker. The words were placed in the sentence so as to appear at the beginning of a major intonation phrase (eg. *Le principe de cette nouvelle émission, dit-elle, et sa* ***diffusion****, pourraient être très mal accueillis par le public* ('the concept of this new program, she said, and its **broadcast,** may not be well accepted by the audience')), hence reinforcing the probability of clear marking of the target word by IA and FA [9].

The words were then extracted and resynthesized *without* **IA (-IA).** FA was kept constant and natural (see Figure 1).

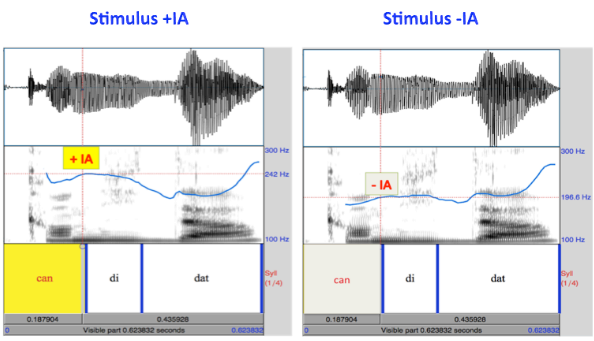

Figure 1: *Example of f0 resynthesis with (+IA) and without (-IA) on the word 'candidat', with quadratic interpolation from the f0 value of the preceding determinant to the f0 value at the beginning of the last stressed syllable for -IA words.*

The *f0* value of the first vowel was lowered near the *f0* value of the preceding (unaccented) determinant ('*le'* or *'sa')*. A quadratic transformation modified the *f0* values to reach progressively the *f0* value at the beginning of the last (accented) vowel, in order to maintain naturalness (micro-prosodic variations and some features of the original *f0* pattern). The +IA stimuli were forward and back transformed to equalize the speech quality between +IA and -IA stimuli.

The duration of the target words was held constant in both stress conditions (+IA; -IA), since only the *f0* parameter was manipulated (*candidat*: 624 ms; *diffusion*: 741 ms).

2.2. Participants and experimental tasks

30 healthy right-handed French native speakers, aged 18-35 (mean age 23,8; 27 females), participated in 2 versions of the Oddball paradigm in the same sequential order: a Passive *then* an Active listening task. Task order was not randomized throughout participants because doing the Active task before would have drawn attention to our stress manipulation. During the Passive task, subjects were watching a silent movie while listening to the repetition of the stimulus *'candidat*', in either of the two stress conditions (+IA; -IA). A total of 1092 words were presented with a pseudo-random combination of 106 deviant words and 986 standard words. The Inter Stimuli Interval (ISI) was 576 ms. The total experiment duration was 23 mns, in one single block. After a 10 mns' break, participants performed the Active task, where they were told to press a button as soon as they detected a deviant stimulus. In this task, the stimulus was '*diffusion*' in either of the two stress conditions (+IA; -IA), with the same total number of words and the same proportion of deviant and standard stimuli as in the previous task. The ISI was 459 ms and the experiment lasted 28 mns in total, throughout two blocks of 14 mns each with a 5 mns' break between the two blocks.

The participants were divided into two groups according to the Oddball version presented (see Figure 2): the first group (n= 14) performed the passive and the active task with -IA words as deviants and +IA words as standards (Version1); the reverse was true for the second group (n= 16) who performed both tasks with +IA as deviants and -IA as standards (Version2).

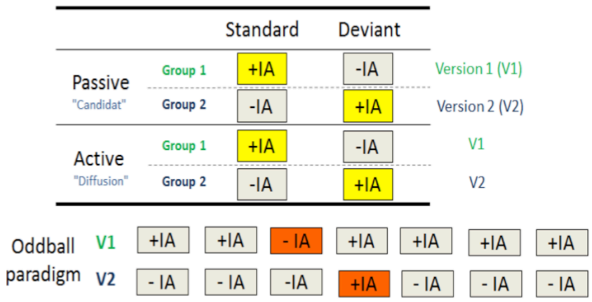

Figure 2: *Two versions of the Oddball Paradigm, where either -IA or +IA is the deviant. Participants are in 2 groups, each group having the same deviant condition in the Passive and the Active condition.*

2.3. EEG recordings

The EEG was recorded in both passive and active tasks, with 32 Ag/AgCl-sintered electrodes mounted on an elastic cap and located at standard left and right hemisphere positions over frontal, central, parietal, occipital and temporal areas (International 10/20 System; Jasper, 1958) at : Fz, Cz, Pz, Oz, Fp1, Fp2, AF3, F3, AF4, F4, C3, C4, P3, P4, PO3, PO4, P5, P6, O1, O2, F7, F8, T3, T4, T5, T6, FC5, FC6, CP1, CP2, CP5 and CP6. The Horizontal ElectroOculoGram (HEOG) was recorded from a bipolar montage with electrodes placed 1 cm to the left and right of the external canthi; the Vertical ElectroOculoGram (VEOG) was recorded from a bipolar montage with electrodes placed beneath and above the left eye, to detect blinks and vertical eye movements. The EEG and EOG were amplified by BioSemi amplifiers (ActiveTwo System) with a band-pass filter of 0.01-30 Hz and was digitized at 512 Hz. Trials containing ocular artefacts, movement artefacts, or amplifier saturation were corrected from averaged ERP waveforms. The data were analysed using Brain Vision Analyser software version 2.0 (Brain Products, Munich, Germany). Each electrode was re-referenced off-line to the algebraic average of the left and right mastoids. Continuous recordings were segmented into 1100 ms duration starting 200 ms before stimulus onset (baseline). Trials were averaged within each of the two stress conditions. Finally, data were averaged across participants to obtain the grand-averages. Subject-averages were aligned to the 200 ms baseline preceding the auditory target. EEG results are presented in the Passive task, while only behavioural data (error rates and RTs) are presented for the Active task.

ERP data were statistically analysed by computing the mean ERP amplitude relative to a 80 ms window centred at the peak latency [100-250 ms] on 9 electrodes (F3, Fz, F4, C3, Cz, C4, P3, Pz and P4). For each comparison (within and between participants) ANOVAs were conducted with Stimulus Type (standard *vs* deviant) and Stress Condition (+IA *vs* -IA). All *p*-values were adjusted with the Greenhouse–Geisser epsilon correction for nonsphericity.

3. Results

Because of an abnormal error rates (>40%), data from five participants were excluded from the analyses. Moreover, one more participant was excluded due to a large number of artifacts on the ERP data. Thus, a total of 24 participants (version -IA: n=12; version +IA: n=12) were included on the behavioral data analyses and the ERP grand averages.

3.1. **Behavioural results from the Active Task: Reaction Times and Errors Rates by deviant type**

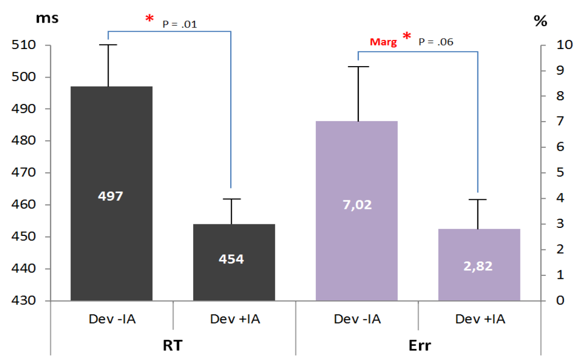

Figure 3: *Reaction Times (left; in ms) and Error Rates (right; in %) for -IA and +IA deviants. * indicates statistic significance.*

Paired two-tailed t tests revealed that RTs were significantly faster for detecting +IA deviants than for -IA deviants (454 ms vs. 497 ms; $t(23) = 3.27$, $p < .01$). Moreover, error rates differences between the two stress conditions were marginally significant ($t(23) = 1.1$, $p = .06$), -IA deviants being slightly more prone to errors than +IA deviants (7.02% *vs* 2.82%; see Figure 3).

3.2. **Event-Related Potential results from the Passive Task**

3.2.1. *Within participants' MMN comparison*

Results show a significantly larger MMN amplitude for -IA deviants than for +IA deviants ($F(1, 23) = 6.52$, $p < .01$), indicating that participants not only process the acoustic cues to stress, but that -IA deviants in a +IA standards environment are less expected than +IA deviants in a -IA standards environment.

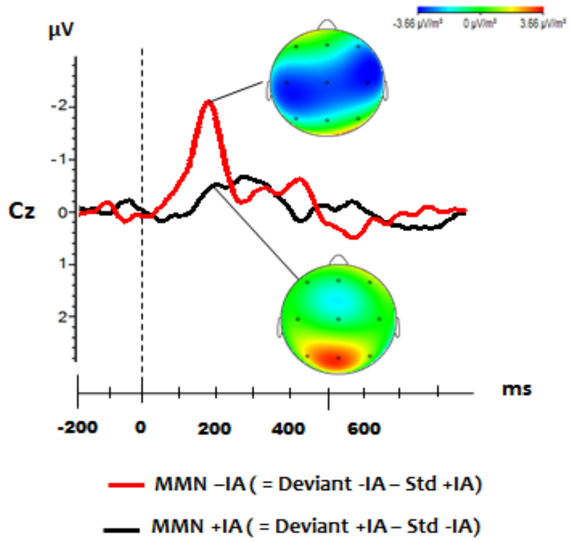

Figure 4: *MMN component for -IA and +IA deviants, for each participants' group. Grand average ERPs recorded at the Cz (central) electrode. Amplitude (µV) is represented on the ordinate, with negative voltage up, and time (msec) on the abscissa. Topographical maps are computed at the MMN peak latency.*

3.2.2. *Between participants' MMN comparison*

The previous within participants' results indicate that +IA deviants do not elicit an MMN. In order to unravel whether it is merely due to low-level, acoustic processing or whether it reflects higher-level processes (stress pattern rarity processing), we calculated MMN effects between identical deviant and standard stimuli (-IA deviants subtracted from -IA standards, and same procedure for +IA stimuli) across participant groups.

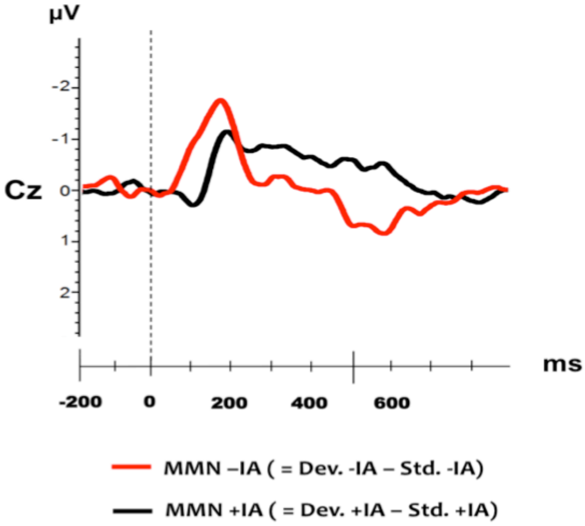

Figure 5: *Difference waves for (deviant – standard) -IA stimuli, and +IA stimuli respectively, across participants. Identical ERPs' graph and statistical settings than for Figure 4.*

Results for deviants and standards ± IA identical stimuli show a significantly larger MMN amplitude for -IA stimuli than for +IA stimuli ($F(1, 23) = 2.94$, $p < .05$).

4. Discussion

The aim of this study was to determine whether French listeners discriminate and process stress patterns in French. Since initial accents (IA) seem to play an important role in prosodic structuring and speech segmentation, the focus of this study was on IA processing at the word level. Despite still inconsistent accounts in prosodic models as to its phonological role, and its description as a 'loose boundary marker', our results indicate that French listeners readily perceive IA and that +IA stress patterns on lexical words are the expected, preferred patterns.

We chose to investigate IA processing through the MisMatch Negativity ERP component because it allows accounting for stress pattern processing at a pre-attentive level and independently from other levels of linguistic processing. In our study, participants were also submitted to an active version of the task in order to enrich the electrophysiological investigation with behavioral data.

Behavioral results show significantly faster Reaction Times to detect +IA deviants in a -IA standard background than to detect -IA deviants in a +IA standard background, which indicates facilitating processing for +IA stimuli. Error rates are also informative insofar as +IA deviants are significantly less numerous than -IA deviants. Altogether, behavioral results revealed that +IA words are discriminated and processed more easily and automatically than -IA words.

ERP data analysis revealed that French listeners process -IA and +IA word stimuli differently. If French listeners did not process stress and were 'deaf' to accentuation, both oddball versions would have yielded similar results, i.e. no MMN component would have emerged when -IA were deviants in +IA stimuli background and vice-versa. The presence or absence of f0 acoustical saliency would not have been discriminated. On the contrary, our results not only indicate that listeners process saliency differences but that they have a preferred stress pattern. Indeed, a large MMN emerges for -IA deviants, indicating that -IA deviants are surprising in a +IA context and showing that the absence of f0 variation on the word was processed (Figure 4). But if listeners were sensitive to low-level acoustic features only, a similar MMN should emerge when +IA deviants occur in a -IA context, indicating a mere processing of acoustic difference. However, Figure 4 shows no such MMN for +IA deviants. Because the MMN is a difference wave between deviant and standard stimuli, this latter result calls for further investigation. The MMN is a measure of the rarity of a stimulus in a context. This short-term memory rarity processing may be emphasized by pattern comparison in long-term memory. It may be that -IA patterns are not expected stress patterns and elicit a negativity wave even when presented as standards. Or it could be that -IA stimuli do not leave a strong enough trace in the short-term memory so as to elicit an ample MMN for +IA deviants, because they do not call upon long-term memory representations. Hence, the rarity processing of +IA deviant stimuli is reduced in the difference wave and no MMN emerges. To explore the rarity processing effect independently from acoustic and prosodic phonological effects, we calculated MMN for identical deviant and standard stimuli *across* participants. Figure 5 reveals a MMN for +IA rare stimuli but this MMN is significantly less ample than for -IA stimuli. Rare -IA stimuli thus remain more surprising than +IA deviants in matched contexts. Thus, a mere acoustic interpretation of ±IA processing has to be ruled out, because a MMN is still present between similar acoustical stimuli. On the contrary, our results show that ±IA stimuli both rely on short-term memory processing (the *absence* of saliency is surprising) and on long-term memory representation (-IA stress patterns are more surprising than +IA stress patterns and mismatch the long-term stress pattern representation).

Altogether, these results point out a preference for +IA stress patterns in French.

5. Conclusions and Perspectives

This study is, to our knowledge, the first investigation of stress pattern processing in French using the oddball paradigm. Results indicate that French listeners not only discriminate stress patterns but show preference for +IA stress patterns. This is a first step towards showing that IA is encoded at the lexical word level and it reinforces its role in speech encoding and decoding processes. These results echo similar findings in other languages [19; 20; 21]. Further MMN investigations will be conducted with more ±IA lexical items in order to generalize our interpretation of a phonological representation of +IA stress patterns. Following Honbolygo [21], we will also extend our investigation to pseudo-words, to emphasize the interpretation of +IA stress templates in long-term memory. Finally, this paradigm will also be extended to FA processing to address the potential similar role of IA and FA in French prosodic phonology.

6. Acknowledgments

This research is funded by a French National Research Agency (ANR) three-year award to the Principal Investigator Corine Astésano (n° ANR-12-BSH2-0001-01; '*PhonIACog*' project: 01/01/13 - 31/12/15).

7. References

[1] Rossi, M. (1980). Le français, langue sans accent ? In I. Fónagy & P. Léon (Eds.). *L'accent* en *français contemporain (Studia Phonetica), 15*. : 13-51.

[2] Vaissière, J. (1991). 'Perceiving rhythm in French?' *ICPhS'91*, 4, Aix-en-Provence: 258-261.

[3] Beckman, M.E. (1992). 'Evidence for Speech Rhythms across Languages'. In *Speech Perception, Production and Linguistic Structure*. Tohkura; Vatikiotis-Bateson; Sagisaka (eds.), Tokyo: 457-463.

[4] Dupoux, E.; Pallier, C.; Sebastian, N.; Mehler, J. (1997). A destressing "deafness" in French? *Journal of Memory & Language, 36,* 406-421.

[5] Jun, S.-A.; Fougeron, C. (2000). A phonological model of French intonation. In Botinis, A. (Ed.). *Intonation: Analysis, Modelling and Technology*. Kluwer, Boston: 209–242.

[6] Post, B. (2000). *Tonal and Phrasal Structures in French Intonation*. Thesus, The Hague.

[7] Welby, P. (2003). *The slaying of Lady Mondegreen, being a study of French tonal association and alignment and their role in speech segmentation*. PhD dissertation, Ohio State University.

[8] Di Cristo, A. (1999, 2000). Vers une modélisation de l'accentuation du français. *French Language Studies, 9*, 143-179 & *10*, 27-45.

[9] Astésano, C.; Bard, E.; Turk, A. (2007). Structural influences on Initial Accent placement in French. *Language and Speech, 50 (3)*: 423-446.

[10] Fonagy I. (1980). L'accent en français : accent probabilitaire. In Ivan Fonagy et Pierre Léon (eds.), *l'accent en français contemporain*. Studia Phonetica (pp. 123.233). Paris: Didier.

[11] Rolland, G.; Lœvenbruck, H. (2002). Characteristics of the Accentual Phrase in French: An acoustic, articulatory and perceptual study. In *Speech Prosody 2002, International Conference*.

[12] Bagou, O.; Frauenfelder, U. H. (2006). Stratégie de segmentation prosodique: rôle des proéminences initiales et finales dans l'acquisition d'une langue artificielle. *Proceedings of the XXVIèmes Journées d'Etude sur la Parole*, 571-574.

[13] Welby, P. (2007). The role of early fundamental frequency rises and elbows in French word segmentation. *Speech Communication*, *49*(1), 28-48.

[14] Astésano, C.; Bertrand, R.; Espesser, R.; Nguyen, N. (2012). Perception des frontières et des proéminences en français. In *Actes de la conférence conjointe JEP-TALN-RECITAL 2012, volume 1: JEP,* 353–360.

[15] Cutler, A.; Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 113-121.

[16] Böcker, K. B.; Bastiaansen, M.; Vroomen, J.; Brunia, C. H.; Gelder, B. (1999). An ERP correlate of metrical stress in spoken word recognition.*Psychophysiology*, *36*(6), 706-720.

[17] Magne, C.; Astésano, C.; Aramaki, M.; Ystad, S.; Kronland-Martinet, R.; Besson, M. (2007). Influence of syllabic lengthening on semantic processing in spoken French: behavioral and electrophysiological evidence. *Cerebral cortex*,*17*(11), 2659-2668.

[18] Li, X.-q.; Ren, G.-q. (2012). How and when accentuation influences temporally selective attention and subsequent semantic processing during on-line spoken language comprehension: An ERP study. *Neuropsychologia, 50*, 1882–1894.

[19] Honbolygó, F.; Csépe, V.; Ragó, A. (2004). Suprasegmental speech cues are automatically processed by the human brain: a mismatch negativity study. *Neurosci. Lett. 363*, 84–88.

[20] 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 733*, 362–368.

[21] Honbolygó, F.; Csépe, V. (2012). Saliency or template? ERP evidence for long-term representation of word stress. *International Journal of Psychophysiology 87,* 165-172.

[22] Winkler, I. (2007). Interpreting the mismatch negativity. *Journal of Psychophysiology 21 (3/4)*, 147–163.

[23] 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:* 2544–2590.