Production of French final stressed syllables in Accentual Phrase by Chinese learners: A pilot study

Lei Xi[1], Sandrine Wachs[2], Rachid Ridouane[1]

[1]Laboratoire de Phonétique et Phonologie (UMR7018, CNRS – Sorbonne Nouvelle), Paris, France
[2]Didactique des langues, des textes et des cultures (EA2288 – Sorbonne Nouvelle), Paris, France
`lei.xi@sorbonne-nouvelle.fr`, `sandrine.wachs@sorbonne-nouvelle.fr`,
`rachid.ridouane@sorbonne-nouvelle.fr`

Abstract

Rhythm plays an essential role in the segmentation of speech, a fundamental process in encoding and decoding messages for speakers and listeners. This study compared productions of French stressed syllables at different rhythmic boundaries, by French native speakers (NS) and Chinese learners (CL).

The study started by retrieving duration of vowels and syllables within the examined utterances, with special focus on the last stressed syllables in Accentual Phrase (AP), for which temporal and non-temporal parameters were examined. Results concerning the temporal patterns showed that both NS and CL produced final lengthening in AP final syllables. However, CL's durations were much longer than the NS'. At the utterance level, CL with low proficiency had a wide temporal variation among segments, while high proficient CL had an isochrony tendency, similar to NS' performance. Results concerning the non-temporal parameters showed that low proficient CL produced a pitch fall between rhythmic units, which was not the case in NS' production. We considered this finding as a negative transfer from resetting of pitch register in Mandarin Chinese. Distribution of pauses was also analyzed, and some inappropriate pauses inserted within sense groups were found in CL. These pauses may weaken listener's decoding process in speech.

Index Terms: L2 rhythm acquisition, stressed syllable, speech segmentation, French, Mandarin Chinese.

1. Introduction and Goal

For a long time, empirical acoustic studies have failed to prove isochrony in speech [1] and it has been widely demonstrated that isochronic occurrence does not exist on the surface timing of speech [2]. Moreover, "*timing is not the only aspect of spoken utterances that might contribute to a listener's sense of their rhythm*" [2]. Nowadays, instead of isochronic occurrence of syllable, stress or mora [3, 4], rhythmic periodicity in speech is more considered as a regular occurrence of certain prosodic events at a particular position. Rhythmic structure serves to combine smaller units into larger units during speech production and division of utterance into smaller prosodic groups in speech perception. Rhythmic chunks work in a hierarchical manner by both temporal and non-temporal cues in order to signal boundary marker, utterance coherence, and process information according to incremental prosodic structure [5].

Despite the fact that rhythm is a crucial tool for segmenting continuous speech, grouping syllables in chunks, and processing information, L2 learners have generally not developed the ability to recognize the durational and melodic cues pertaining to speech rhythm and taught how to use these cues to signal discourse structure. In particular, research and studies examining the problems and difficulties which Chinese learners (CL) face in learning French rhythmic pattern are very rare. In the literature, only a few studies have addressed this question directly [6, 7, 8]. Almost all of the existing studies concerning the question of rhythm acquisition have calculated absolute values of rhythm metrics [9, 10, 11, 12, 13, 14, 15] to compare with the native speaker, which is not sufficient to study rhythm acquisition, at least for the CL of French. The validity of rhythm metrics has been questioned by numerous researchers (among others [1, 16, 17, 18]). Furthermore, studying speech rhythm by rhythm metrics overlooks F0 and prominence, two indispensable factors of the rhythmic system [2, 19].

Stressed syllable, or the last syllable of the last word in a group of words, is a salient feature of French rhythmic pattern, whose role is to constitute a unit of meaning, a sense-group and to mark rhythmic boundary [20, 21, 22]. In French, "*the stress group is effectively the primary unit in sentence phrasing*" [5] and this stress group at the phrasal level assumes a demarcative function in the Accentual Phrase domain (AP). This kind of prominence is acoustically manifested by a lengthening with a fundamental frequency (F0) rise or a F0 peak to expression of continuation in an utterance [23]. Some researchers have classified French as a so-called "boundary language" [20, 23]. In Mandarin Chinese, Cao's research [24, 25, 26] gave a detailed description of rhythmic patterns, whose works have divided the Mandarin rhythmic system into three hierarchic levels: prosodic word (the minor rhythmic unit); prosodic phrase (the intermediate rhythmic chunk) and intonation phrase (the major rhythmic group). Distinguished from French, Mandarin Chinese only has a strong boundary lengthening in the prosodic phrase (the intermediate layer) [25]. In particular, the resetting of pitch register is another important boundary marker between rhythmic units in an utterance. As stated in Cao's study [25], this resetting plays a demarcative role in Chinese Mandarin speech. These differences in prosodic structure between French and Chinese mean that the CL of French may express prominence in areas other than those of French native speakers (NS).

To the authors' knowledge, there is no study examining the problem of the CL in producing final stressed syllables in the AP domain. Thus, the current pilot study is dedicated to this question. In this study, six French sentences were produced by NS and CL with different levels of proficiency. The aim of the present paper is to examine the productions of the final

stressed syllables in French AP by CL. We specifically examine the temporal pattern in utterance and pause distribution, in order to determine the problems of French as a foreign language (FFL) for the CL. By comparing how the CL and the NS express prominence and produce rhythmic pattern in relation to sentence structure, we will be able to provide some evidence that at least some of the CL are applying Mandarin Chinese domains for organizing prominence (notably at the AP level) to French sentences.

2. Hypotheses

Based on the conclusions of previous research and our empirical observations from the production experiment, we have devised several hypotheses for this study: i) when a stressed syllable is located at the end of utterances, we assume that the CL would realize a final lengthening which is similar to NS' production; ii) when a stressed syllable is not located at the utterance level but at the boundary of AP to mark continuation at the phrasal level, the CL would not have a final lengthening and pitch pattern would be problematic; iii) learners with different levels of proficiency would have different rhythmic patterns in their productions.

3. Methods

3.1. Speakers and corpora

Seven female CL of FFL (mean age: 25.43, SD: 4.95) with varying levels of proficiency in French (1 learner with a level B1, 3 with B2, 2 with C1 and 1 with C2 according to the *Common European Framework of Reference for Languages*) took part in the production experiment. All of the CL speak Mandarin Chinese as their mother tongue, started learning French as adults (between 2 to 6 years ago) and had been living in France for at least two years at the time of recording. Six female Parisian French NS (mean age: 24.6, SD: 1.14) participated as references for acoustic comparisons. The recording took place in an anechoic room. Participants were recorded one by one, using Pro Tools program, with an AKG C520L wireless head worn microphone (distance mouth-microphone: 3 cm) and a MOTU ULMK3 sound card. The corpora included 6 French declarative sentences (Table 1). All participants were asked to read the corpora with a natural tone. In total, there were 78 sentences in the corpora.

Table 1: *Corpora recorded in the study*
(stress group in bold with stressed syllable in red).

N°	Sentence with IPA transcription
1	Il est **joli**. /il ɛ ʒɔ li/
2	Tu veux voir **Julie**. /ty vø vwaʁ ʒyli/
3	Il a parlé du **paysage**. /il a paʁle dy peizaʒ/
4	Il a parlé du **paysage**, et de tout autre **chose**. /il a paʁle dy peizaʒ e də tu totʁ ʃoz/
5	Vous aimez **l'anglais** et vous étudiez la **littérature**. /vu zeme lãglɛ e vu zetydje la liteʁatyʁ/
6	La semaine **dernière** à **Paris**, le voyageur égaré vous a **abordé**. /la səmɛn dɛʁnjɛʁ a paʁi lə vwajaʒœʁ egaʁe vu za aboʁde/

3.2. Acoustic study

Our acoustic study started by identifying possible stress groups and syllables (Table 1) in the corpora. Thereafter, all of the sentences that were produced were segmented into syllables and in phones by using an online WebMAUS service [27, 28] for automatic transcription and manual verification and correction were performed. Afterwards, using a script performed on Praat [29], we obtained all of the quantitative values which were necessary for our study: i) duration of each vowel and each syllable was retrieved; ii) values of F0 beginning (point at 10%) and F0 ending (point at 90%) in the stress group (for example, "l'anglais" in "Vous aimez l'anglais et vous étudiez la littérature.") were generated, whose difference (converted into semi-tone (st) with 200 Hz as reference (F0-Diff)) allowed to determine rising/falling pitch. With the values we obtained, we compared differences in several aspects between the CL and the NS productions: i) duration of the last vowel in the last syllable in stress group; ii) duration of the last syllable in stress group; iii) F0-Diff of stress group; iv) durational cue at the utterance level; v) values of ΔS, the standard deviation of syllabic durations [30]; vi) distribution of pauses.

4. Results

4.1. Temporal cues

4.1.1. Final lengthening in stressed syllable

After having compared vocalic and syllabic durations of stressed syllables between the CL and the NS, results showed that all CL produced a final lengthening, whether it is at the utterance level or the phrasal level. This result is contrary to our second hypothesis, which assumed that the CL would not lengthen final syllables at the phrasal level. Nevertheless, in the CL's realizations, duration of final lengthening is much longer than the NS', especially in low proficient learners. Here we show an example of sentence N° 5 produced by a low proficient learner, compared with a NS' realization (Figure 1). From the figure, we can see that at the non-stressed syllable positions, the CL and the NS have a similar duration, whereas at the phrasal and utterance boundaries, the CL's lengthening was much longer than that of the NS. Comparing all of the values of final lengthening, we noticed a general tendency: the lower the level of proficiency, the longer the final lengthening produced by our CL in this study.
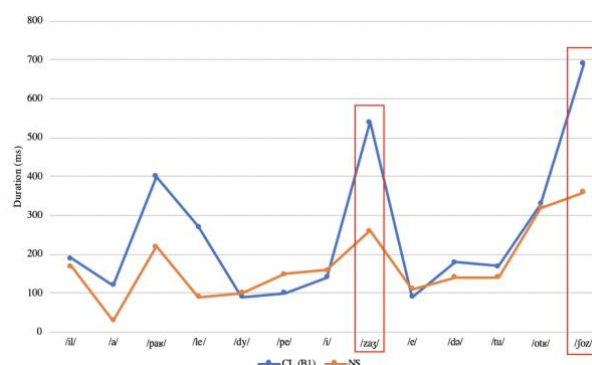

Figure 1: *Comparison of the duration of stressed syllables in sentence N°4 produced by a CL (B1) and a NS.*

4.1.2. Temporal pattern over full utterances

The temporal variation at the utterance level was also compared and we noticed that NS' performance had an isochronic tendency, with the exception of a moderate lengthening at boundary position. However, by observing durational cues, we found a large temporal variation between segments in some CL, which is far from an isochronic tendency. It was always the case in low proficient learners producing sentences (N°4, 5 and 6) with the AP boundaries at the phrasal level. As level of proficiency increases, the temporal pattern remains stable in the entire utterance. Three temporal patterns of sentence N°6 produced by two CL (B1 and C2 levels of proficiency, respectively) and one NS are shown as examples in Figure 2. From this figure, we can see that the NS and the CL with C2 had a similar temporal variation: a duration between a range of 100-200 ms for the segments at the non-stressed positions and a moderate lengthening in the stressed syllables (/njɛʁ/ in "dernière"; /ri/ in "Paris"; /ʒœʁ/ in "voyageur"; /boʁ/ in "abordé"). Contrary to the advanced C2 learner, the beginner had a large durational difference between segments in her temporal pattern.
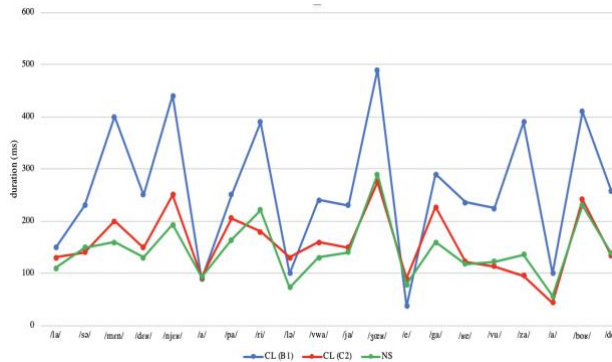

Figure 2: *Comparison of syllabic durations in sentence N°6 produced by two CL (B1: blue; C2 red) and a NS (green).*

In order to have a clearer view about how temporal pattern changes with the proficiency level of the CL, the values of ΔS were calculated for each sentence produced by each subject and ΔS values of our NS subjects were averaged for each sentence. ΔS values in each CL were compared with the averaged ΔS in NS by carrying out a dependent sample t-test on R [31]. Results of the test are reported in Table 2:

Table 2: *Results of dependent t-test for comparison of ΔS values.*

CL N°	P values
1	**2.048e-05 ***
2	**0.01586 ***
3	**0.005952 ***
4	**0.009132 ***
5	0.07216
6	0.505
7	0.8614

From the statistical test, we found a significant difference (p< 0.05) of syllabic duration patterns between NS and learners at the intermediate level (N°1-4). Nevertheless, p values indicated that there is no significant difference between the NS and advanced learners (N°5-7), which means that realizations of duration for these learners seemed similar to NS'

production. The statistical test results were in keeping with our expectation: as acquisition progresses, timing pattern becomes more stable and more target like in the CL.

4.2. Melodic cue

Pitch movement in stress group was also examined, both at the utterance boundary and at the phrasal boundary. Quantitatively, we calculated F0-Diff in all the stress groups produced by the CL and the NS; qualitatively, we generated melodic contour on Praat [29] in these stress groups to compare pitch movements.

Firstly, we found that all CL had no problem in producing a falling pitch at the utterance level to mark the end of sentence. Their performance was comparable with NS' realization. However, the visual inspection of productions revealed a problematic production in some CL at the phrasal boundary: some intermediate learners realized a falling pitch at the phrasal level boundary in sentences N°4, 5 and 6, which was contrary to the cases in the NS. Taking sentence N°4 as a typical example, the quantitative and qualitative analyses are shown below:

- quantitatively, values of F0-Diff in the stress group "paysage" were compared:

Table 3: *Values of F0-Diff in the stress group "paysage" in sentence N°4 produced by the CL and the NS.*

CL N°	F0-diff (st)	NS N°	F0-Diff (st)
1 (B1)	-10.4	1	+5.7
2 (B2)	-12.7	2	+4.3
3 (B2)	-7.6	3	+5.4
4 (B2)	+6.5	4	+3.7
5 (C1)	+6.2	5	+4
6 (C1)	-1.3	6	+5.7
7 (C2)	+0.5		

The F0-Diff showed that all NS had a rising pitch movement at the phrasal level, but in learners with low proficiency (N°1-3), they performed a falling pitch.

- qualitatively, pitch contours in a CL and a NS are compared in Figure 3 from which a clear fall can be seen in the CL's production (right).
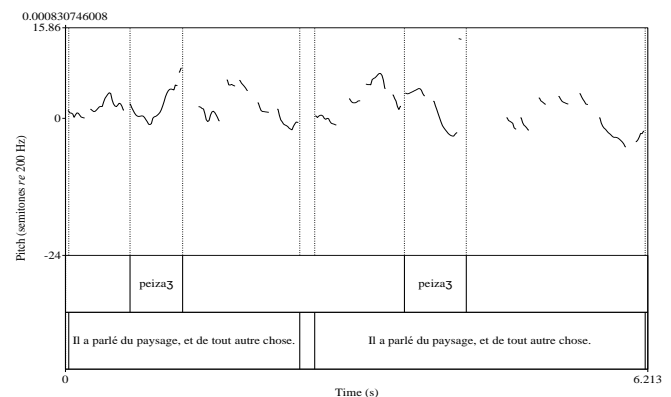

Figure 3: *Pitch realization in the stress group "paysage" of sentence N°4 in a NS (left) and a CL (right).*

Falling pitch contour at the phrasal level was found in most of the realizations of CL but not in all cases. We noticed that the same learner can sometimes produce a rising pitch, just like a NS did, and sometimes a falling pitch, both at the phrasal level. It is commonly the case in intermediate learners. We presumed that intermediate CL did not seem to be able to master French pitch patterns and to produce an appropriate melodic contour in rhythm production.

4.3. Distribution of pauses

Pause is another important element in rhythmic pattern, whose function is to mark the boundary of chunks. In this study, distribution of pauses was analyzed, and we identified all the pauses inserted in a sense-group by the CL. This is shown in Table 4.

Table 4: *Pauses inserted in a sense group by the CL.*

CL N°	Inserted pause (/)	Duration (ms)
1	… et vous étudiez la / littérature.	70
1	… le voyageur / égaré …	260
2	… le voyageur / égaré …	771
3	… le voyageur / égaré …	182
	vous / a …	120
4	… parlé / du paysage …	230
5	… le voyageur / égaré …	74
6	… le voyageur / égaré …	210

However, in the NS productions, no pause was inserted in a sense group. The chunks in NS productions were organized by coherent sense-groups and separated by pauses.

5. Discussion and Conclusions

Modern cognitive linguistics suggests that speech understanding responds to discourse chunks. It is also true that in human language production, people neither utter a sentence without any break nor word by word separately, especially for a longer sentence or syntactic phrase. Rhythmic organization is a chunking strategy concerned mainly with boundary marking by prosodic events.

In this study, we have examined the production of French prominence in final AP in the NS and the CL. In our acoustic analysis, final prominence was firstly analyzed from temporal and non-temporal aspects. Afterwards, we examined temporal patterns in entire utterances by durational cues and ΔS values. Finally, we looked into the distribution of pauses.

By our analysis, we concluded that the CL realized a lengthening, whether at the phrasal level or the utterance level. However, lengthening in stressed syllables produced by the CL was much longer than in the NS. Maybe this is due to a relatively slower speech rate in the CL. What is more problematic resides in non-temporal aspects: we found that the CL at an intermediate level always produced a falling pitch instead of a rising one. We consider this to be a negative transfer from Mandarin Chinese. As mentioned in the introduction, Cao's study [25] revealed that declination inside a prosodic phrase and resetting of pitch register at the beginning of the next phrase are particularities of the Chinese Mandarin rhythmic pattern, which play a demarcative role. Yet in French, the NS use a F0 rise or a F0 peak to express continuation [20, 23] and a F0 fall for termination. The CL in our study may be interfered by their L1 and realized a falling pitch.

The comparison of durational cues between the CL and the NS informed us that timing patterns in L2 speech vary as proficiency increases. We looked at durational cues in entire utterances produced respectively by the NS (with an isochronic tendency), advanced learners (with a small variation in duration, more target-like), and low proficient learners (with a large variation in duration). Values of ΔS were calculated and investigated by a dependent t-test. Results of the test indicated that productions in low proficient learners had a significant difference compared with the NS' production, but it is not the case for the advanced CL. Our study confirmed conclusions resulting from other studies, which have suggested that "*timing control in L2 speech production improves as acquisition progresses, and rhythm becomes more stable*" [32]. Timing control ability is not only concerned with durational cue, but also with how to distribute pauses in speech. Most of the CL in this study put pauses inside sense group, which lead to an incorrect demarcation. According to various studies [33, 34], rhythmic boundaries are used to organize words into larger syntactic phrases. The factors investigated in our study showed that low proficient CL were unable to use appropriate prosodic events to make such boundaries in their speech.

However, there are several limitations with the present study: i) we have only examined six French sentences in the study. We need to have a larger corpus covering more sentences and more speech styles (spontaneous speech in particular) in order to see whether our conclusions from this pilot study are still relevant; ii) perception test may also be necessary to determine whether CL can distinguish prominence and rhythmic pattern between two languages. For future research, these factors should be taken into consideration.

6. Acknowledgements

7. References

[1] A. Arvaniti, "Rhythm, Timing and the Timing of Rhythm," *Phonetica*, 66 (1-2), pp. 46-63, 2009.

[2] A. Turk, S. Shattuck-Hufnagel, "What is speech rhythm? A commentary on Arvaniti and Rodriquez, Krivokapić, and Goswami and Leong," in *Laboratory Phonology 4(1)*. Berlin: Mouton de Gruyter, pp. 93-118, 2013.

[3] K. Pike, *The intonation of American English*. Ann-Arbor: University of Michigan Press, 1945.

[4] D. Abercrobie, *Elements of general phonetics*. Edinburgh: Edinburgh University Press, 1967.

[5] P. Martin, *The Structure of Spoken Language. Intonation in Romance*. Cambridge: Cambridge University Press, 2015.

[6] C. Gabriel, J. Stahnke, J. Thulke, "On the acquisition of French speech rhythm in a multilingual classroom: Evidence from linguistic and extra-linguistic data," in *CMLF 2014 - Congrès Mondial de Linguistique Française*, 2014.

[7] S. Landron, J. Gao, Y. Chang, Y. Tian, "Les sinophones", in *La prononciation du français dans le monde : du natif à l'apprenant*. Paris : CLE-International, pp. 193-199, 2017.

[8] C. Pillot-Loiseau, H. Xie, "Transfert rythmique du chinois mandarin au français dans l'apprentissage du Français Langue Étrangère," in *CMLF 2018 - Congrès Mondial de Linguistique Française*, 2018.

[9] F. Ramus, M. Nespor, J. Mehler, "Correlates of linguistic rhythm in the speech signal," *Cognition*, 73, pp. 265-292, 1999.

[10] E. L. Low, E. Grabe, F. Nolan, "Quantitative characterizations of speech rhythm: 'Syllable-timing' in Singapore English," *Language and Speech*, 43(4), pp. 377-401, 2000.

[11] E. Grabe, E. L. Low, "Durational variability in speech and the rhythm class hypothesis," in *Laboratory Phonology 7*. Berlin: Mouton de Gruyter, pp. 515-546, 2002.

[12] V. Dellwo, "Rhythm and speech rate: A variation coefficient for deltaC," *Language and Language-Processing: Proceedings of the 38th Linguistics Colloquium*, pp. 231-241, 2006.

[13] L. White, S. L. Mattys, "Calibrating rhythm: First language and second language studies," *Journal of Phonetics*, 35, pp. 501-522, 2007a.

[14] L. White, S. L. Mattys, "Rhythmic typology and variation in first and second languages," in *Segmental and Prosodic Issues in Romance Phonology*. Amsterdam: John Benjamins, pp. 237-257, 2007b.

[15] J. M. Liss, L. White, S. L. Mattys, K. Lansford, A. J. Lotto, S. M. Spitzer, J. N. Caviness, "Quantifying speech rhythm abnormalities in the dysarthrias," *Journal of speech, language, and hearing*, 52(5), pp. 1334-1352, 2009.

[16] A. Arvaniti, "The usefulness of metrics in the quantification of speech rhythm," *Journal of Phonetics*, 40, pp. 351-373, 2012.

[17] L. Wiget, L. White, B. Schuppler, I. Grenon, O. Rauch, S. Mattys, "How stable are acoustic metrics of contrastive speech rhythm?," *Journal Acoustical Society of America*, 127(3), pp. 1559-1569, 2010.

[18] U. Gut, "Rhythm in L2 speech," *Speech and Language Technology*, v.14/15, pp. 83-94, 2012.

[19] A. Patel, *Music, language and the brain*. New York: Oxford University Press, 2010.

[20] J. Vaissière, "Le français, langue à frontière par excellence," in *Frontière, du linguistique au sémiotique*. Limoges : Lambert Lucas, 2010.

[21] P. Delattre, "Comparing the prosodic features of English, French, German and Spanish," *International Review of Applied Linguistics*, 1(3), pp. 193-210, 1962.

[22] P. Delattre, "Accent de mot et accent de groupe," *The French Review*, 13(2), pp. 141-146, 1939.

[23] J. Vaissière, "Rhythm, accentuation and final lengthening in French," in *Music, Language, Speech and Brain*. London: Macmillan Press, pp. 108-120, 1991.

[24] J. Cao, "Rhythm of spoken Chinese - linguistic and paralinguistic evidences," in *ICSLP 2000 - 6th International Conference on Spoken Language Processing, October 16-20, Beijing, China*, 2000.

[25] J. Cao, "The rhythm of Mandarin Chinese," *Journal of Chinese Linguistics*, 17, pp. 53-80, 2001.

[26] J. Cao, "Restudy of segmental lengthening in Mandarin Chinese,", in *Speech Prosody 2004, March 23-26, Nara, Japan*, 2004.

[27] F. Schiel, "Automatic Phonetic Transcription of Non-prompted Speech," in *ICPhS 1999 - 17th International Congress of Phonetic Sciences, August 17-21, Hong Kong, China*, 2011.

[28] T. Kisler, U. Reichel, F. Schiel, "Multilingual processing of speech via web services," *Computer Speech & Language*, 45, pp. 326-347, 2017.

[29] P. Boersma, D. Weenink, "Praat: doing phonetics by computer [Computer program]," *version 6.0.54*, 2019.

[30] P. Mok, "On the syllable-timing of Cantonese and Beijing Chinese," *Chinese Journal of Phonetics*, 2, pp. 148-154, 2009.

[31] S. Urbanek, H. Bibiko, S. M. lacus, "R [Computer program]," *version 3.6.0*, 2019.

[32] M. Ordin, L. Polyanskaya, "Perception of speech rhythm in second language: the case of rhythmically similar L1 and L2," *Frontiers in Psychology*, 6, 2015.

[33] A. Kennedy, W. S. Murray, F. Jennings, C. Reid, "Parsing complements: Comments on the generality of the principle of minimal attachment," *Language and Cognitive Processes*, 4, pp. 51-76, 1989.

[34] J. Pynte, B. Prieur, "Prosodic breaks and attachment decisions in sentence parsing," *Language and Cognitive Processes*, 11, pp. 165-192, 1996.