

PITCH VARIATION FOR CHINESE SYLLABLES OF DIFFERENT INFORMATION LOAD (BASED ON COMMERCIAL AND SOCIAL RADIO ADS)

Svetlana V. Androsova¹⁾, Veronika G. Karavaeva²⁾, Zianwen Zhang³⁾

¹⁾*Amur State University – Blagoveshchensk (Russia)*

²⁾*Moscow City University – Moscow (Russia)*

³⁾*Heihe University – Heihe (China)*

Abstract. F0 is an important cue present in all vocalic segments, a parameter to differentiate between voiced and voiceless consonants, a relative phonological feature of syllable in tonal languages and a crucial feature of intonation. This paper aims to investigate pitch variation patterns in Mandarin Chinese depending on the syllable information load (Factor 1) in one type of discourse – advertising with the subdivision into social and commercial ads (Factor 2) considering gender differences (Factor 3). 1249 syllable tokens were selected for acoustic measurements, each syllable occurring twice – in the informative and uninformative utterance parts. The information load was determined perceptually: the syllable was considered informative when agreed by the minimum of 60% of listeners, other syllables were considered uninformative. Depending on lexical tone (T1 – T4 and T0), the measurements included average pitch values, declination/inclination starting and ending points (mean values). These parameters were used to judge about relevant pitch features – average height and declination/inclination slope. As a result, a consistent parameter increase in commercial ads vs. social ads and on informative syllables vs. uninformative ones was observed for all tones except T0 that showed the opposite trend in expressing information-based partition. Another finding was tone substitution that marked both informative and uninformative parts but was more frequent in the latter. Substitute tone frequency ranks did not depend on any of the three factors except for T0. High predictability of pitch variation patterns make them applicable in teaching Chinese as L2 in terms of speaking and listening for general and specific purposes.

Keywords: advertising discourse; information load; pitch; syllable; acoustic cue

1. Introduction

In any language, fundamental frequency (F0, pitch) is an important acoustic cue that can be affiliated with segmental and suprasegmental phonological units

(i) universally, being present in all vocalic sounds and contrasting voiceless and voiced consonant phonemes in languages like Russian, (ii) universally, being an inherent property of vowel formant structure, (iii) forming lexical tones in languages like Chinese to enlarge syllable repertoire increasing syllable contrastive properties, (iv) universally, being the most crucial intonation parameter (pitch contour). The current paper aims to investigate prosodic patterns variation in Mandarin Chinese depending on the syllable information load in one type of discourse, namely, advertising.

Since the emergence of acoustic phonetics in China in 1920s, beginning with Liu Fu successful attempts to measure F0 values in 1924 using kymograph (as cited in Le and Ge (2021)), most studies of pitch in Chinese as one of the largest dominant languages of the World have revolved around lexical tone patterns and their interaction with intonation, the latter including tone overlap and the use of neutral tone. Thus, Lee (2012) studies F0 variation pattern in spontaneous speech, reading and counting focusing on the position in the utterance and gender differences. Thriving of pitch studies can be easily explained by the role it plays for speech perception in Chinese as a tonal language. According to Li et al. (2021), a Mandarin speaker's mind, unlike a non-tonal speaker's mind is 'equipped' with a 'device' that manages "language-specific distribution of cortical tuning parameters" providing particular sensitivity to tonal parameters. This tuning is so rigid that a speaker of Chinese learning a nontonal language like English would make more mistakes in perceiving lexical stress than a native speaker of that nontonal language. Liu (2019) discovered that erroneous perception triggered particularly by rising intonation led to statistically significant 47% discrimination accuracy decrease in nouns (but not verbs).

Lexical tones variation has recently become a popular research issue. In addition to canonical tone sandhi, a number of other issues are examined. The study of emphasis effect on F0 patterns (Chen & Gussenhoven 2008) revealed maximum tones discriminability under emphasis compared to no emphasis condition. Tang et al. (2017) studied tone modifications in terms of hyperarticulation (exaggeration effect) in (i) infant-directed speech, (ii) adult-directed speech in noisy environment as compared to adult-directed speech in normal environment. They found that although the key tone properties remained unchanged, both pitch height and pitch contour underwent modifications in the first type of environment, while in the second type of environment only the first parameter changed (increased in both environment types). Later, Tang and Li (2020) investigated F0 behavior under the contrary conditions of fast speech. They noted a significant effect of speech rate on pitch height (rising), slope (flattening) and curvature (reduced contrasts). Chen and Tseng (2019) who focused on intonation unit and prosodic unit boundary interaction found both anticipatory and carry-over pitch effects expressed in rising and lowering of the values. Wu (2019) who conducted a study based on nonsense

tokens synthesized from randomized disyllabic verbs and nouns found that under masking conditions, final keywords in a multi-tone sentence are better perceived when the intonation is natural (when F0 contour is not decreased or exaggerated). This is a strong motivation to study natural patterns of pitch variation during the interaction of lexical tones and intonation.

Chinese lexical tones vary depending on phrasal intonation whose patterns, besides syntactic rules, are determined by linguistic prominence (see more on linguistic prominence in (Cangemi & Baumann 2020)). One of the crucial factors forming degrees of prominence is type of discourse (Discourse 2020; Zheltukhina et al. 2017; Vikulova et al. 2020) accompanied by the information load that is determined by the information structure of the utterance – the issue extensively covered in the literature (see, e.g. Heusinger (1999) who analyzed the emergence and development of the theory, multiple approaches and terminology variation). However, there are hardly any studies on Chinese phonological units' properties depending on their information load in general or in a particular type of discourse like advertising.

Oral advertising is a powerful means to attract customers encouraging them to buy a particular product, use a particular service etc. Although oral ads phonetic patterns are a powerful tool to achieve that goal, there are almost no studies that would focus on them using Chinese material either in terms of highlighting important information or sounding attractive for a customer. The only work to be mentioned here is the one of Wang Yi and Lu Jia (2018) on intensity parameters that listeners associate with different voice features (e.g. lively, intellectual, exquisite, simple, open, magnetic, full of energy, etc.) when they perceive oral ads.

Our previous perceptual and acoustic study examined the interaction of a number of factors including duration, position in the utterance (beginning, middle, end), information load and gender in shorter and longer oral commercial advertisement (Zhang & Karavaeva 2019). The results showed the tendency of duration increase to highlight more important information. It was also found that the largest amount of syllables evaluated by listeners as informative were located at the beginning (or end) of an advertisement (depending on gender or ad type), while the least number of them were placed in the middle. In this study, the focus is made on pitch variation patterns depending on a syllable information load, ads type and gender. We hypothesize that relative pitch features will expose certain differences depending on these three factors. To test the hypothesis, an acoustic study was performed.

2. This study

2.1. Material and methods

The material was collected from 5 Chinese radio stations (“Autoradio” (交通广播), “Longguang News” (龙广新闻), “City Life” (都市生活), “City Women”

(都市女性), “University Radio” (高校广播)), as well as from advertisement in the streets and trade centers. Transcript of each ad was performed using hieroglyphs (Chinese characters). To avoid the effect of music, the texts of the selected ads were read by 6 volunteer professional Mandarin Chinese speakers (3 males, 3 females, aged 24 – 28). The recordings were conducted in a sound-proof booth at Heihe Radio individually with each speaker.

The total of 59 commercial (30 minutes) and 72 social (30 minutes) reread ads comprised the corpus for the study (5 minutes from each speaker for each advertisement type). Totally, the speakers produced 11290 syllables. Their distribution into the ads types was unequal with 4617 syllables in social ads and 6673 syllables in commercial ads. From the total amount, 1249 syllables were selected for acoustic measurements (611 and 638 from social and commercial ads correspondingly). The reason for selecting a syllable was its occurrence in both parts of the utterance information structure – informative (or more informative) and uninformative (or less informative).

We used perceptual approach to distributing the syllables into the parts. 20 native Mandarin Chinese volunteer listeners participated in the experiment. They listened to the ads and marked each syllable as informative or uninformative, thus performing information-based partition. Syllables marked as informative by 60% of listeners or more were labeled as chosen (Ch), they comprised the core of the ads’ information structure. Others were labeled as not chosen (N-Ch), they were affiliated with the periphery of the ads’ information structure. Only vowel parts of the syllables were taken for measurements. The following measurements were conducted in Praat [Boersma, Weenink, 2016] depending on syllables canonical lexical tones (T1–T4) and T0:

- a. F0 average height for T1 (Level Tone);
- b. F0 on inclination beginning (minimal value = MinV) and on inclination end (maximal value = MaxV) for T2 (Rising Tone);
- c. F0 on declination beginning (maximal value = MaxV1), F0 on declination end which is also inclination beginning (minimal value = MinV), F0 on inclination end (maximal value = MaxV2) for T3 (Falling-Rising Tone);
- d. F0 on declination beginning (maximal value = MaxV) and on declination end (minimal value = MinV) for T4 (Falling Tone).

These parameters were used to judge about relevant pitch features – average height and declination/inclination slope following Torsueva’s method (2009).

Here three comments must be made. First, among 333 realizations of T3, about 2/3 (210 tokens) were full-tone manifestations (T3-full) and a little over 1/3 (123 tokens) were incomplete manifestations (fall beginning from mid-level – T3-incmp). Both variations were measured separately, the latter – using declination measurement technique. T3-incmp was not interpreted as T4 substitute due to higher declination starting point and steeper slope of the latter. Second, for Neutral

Tone (so called toneless syllables), F0 average height was measured. Finally, while processing the data, 3 factors were considered: (i) Factor 1 – information-based partition (Ch vs. N-Ch), (ii) Factor 2 – ads type (commercial vs. social), (iii) Factor 3 – gender (male vs. female).

2.2. Results and discussion

Factor 1 – information-based partition. This factor was considered separately and in combination with Factors 2–3. On the whole, there was a consistent difference between more informative and less informative syllables that were specifically expressed in each tone. Table 1 shows the mean values.

Table 1. Pitch values (Hz) – Ch syllables vs. N-Ch syllables

Tone	T1	T2	T3-full	T3-incmp	T4	T0
Ch	249	168 – 188	185 – 149 – 206	179 – 130	240 – 178	126
N-Ch	227	169 – 187	235 – 197 – 227	172 – 138	231 – 185	141

For syllables realized with T1, Ch-syllables were characterized by an increase of average height compared to N-Ch-syllables. If we zoom in, the tendency is highly visible in female commercials ads (with 55 Hz difference) and quite clearly seen in male social ads (21 Hz difference). In female social ads, there is no difference whatsoever, and male commercial ads show the opposite trend, however the difference is not statistically significant (6 Hz).

T2 Ch- and N-Ch-syllables, on the whole, demonstrated no difference. When zoomed in, in commercial female ads Ch-syllables showed 12 – 13 Hz pitch height increase for both MinV and MaxV compared to N-Ch syllables (210 – 251 vs. 197 – 239). In male counterparts, the opposite trend was seen with 10 – 11 Hz pitch height decrease. No consistent difference for social ads was noticed.

T3-full Ch-syllables compared to N-Ch-syllables, on the whole, demonstrated lower level fall-rise contour in terms of MaxV1, MinV and MaxV2, similar declination slope but steeper inclination slope (27 Hz difference). When zoomed in, female commercial ads were characterized by similar MaxV1 and similar declination slope on Ch- and N-Ch-syllables but by much steeper inclination slope on Ch compared to N-Ch with MaxV2 difference reaching 84 Hz. In male counterparts, Ch-syllables had a little steeper inclination, while N-Ch-syllable had much steeper declination due to 12 Hz higher MaxV2 on the former and 40 Hz lower MaxV1 and on the latter. In female social ads, declination slope was a little steeper on N-Ch-syllables, there was no difference in the inclination slope, T3-contour was a little higher on N-Ch-syllables, although the difference was not statistically significant. In male counterparts, canonical T3-full was

substituted with another tone pattern on Ch-syllables, therefore, no comparison of Ch- and N-Ch-syllables can be made so far. Ch- and N-Ch-syllables with T3-incmp showed no differences on the whole. When zoomed in, no consistent tendency was revealed.

T4 Ch-syllables, on the whole, had a little steeper slope compared to N-Ch-syllables. Zooming in, this trend was quite consistent with the only exception of male commercial ads, where the slopes were similar. No clear trend for the contour level was observed.

T0 syllables showed a consistent trend of higher values on N-Ch-syllables on the whole and when zoomed in with the exception of female social ads where no syllables with T0 were found.

Factor 2 – ads type. This factor was considered separately and on the background of Factor 1 (see mean values in table 2).

Table 2. Pitch values (Hz) – Social (Soc.) vs. Commercial (Com.) ads

		T1	T2	T3-full	T3-incmp	T4	T0
Soc.		212	147 – 158	202 – 160 – 203	171 – 131	210 – 158	114
Com.		242	189 – 217	218 – 168 – 230	181 – 136	261 – 205	137
Soc.	Ch	219	147 – 160	212 – 162 – 208	174 – 127	213 – 152	102
	N-Ch	205	146 – 156	191 – 157 – 198	167 – 135	207 – 163	125
Com.	Ch	278	188 – 216	157 – 136 – 203	184 – 132	267 – 204	149
	N-Ch	249	191 – 218	278 – 236 – 256	177 – 140	255 – 206	156

In commercial ads compared to social ads, there was a clear tendency of higher values for T1–T4, T0 on the whole (the first two data lines in Table 2) and for both Ch- and N-Ch-syllables with the exception of T3-Ch-syllables that had lower MaxV1 and MinV. Such data indicate closer correlation of pitch values and ads type than pitch values and information load. T1 on Ch-syllables compared to N-Ch-syllables had higher average height in both ads types, while T0 demonstrated the opposite trend. T2 showed hardly any consistent difference in Ch vs. N-Ch, T3-full contour on Ch-syllables in social ads was located higher than in N-Ch-syllables but in commercial ads it showed the opposite trend. Both T3-incmp and T4 were characterized by steeper slopes on Ch compared to N-Ch.

Factor 3 – gender differences. This factor was considered separately and on the background of Factor 1 (see mean values in table 3).

Table 3. Pitch values (Hz) – male vs. female

		T1	T2	T3-full	T3-incmp	T4	T0
female	Ch	313	197 – 222	269 – 226 – 302	253 – 159	297 – 226	191
	N-Ch	285	192 – 215	265 – 230 – 263	247 – 159	295 – 235	204
male	Ch	178	135 – 150	100 – 78 – 109	126 – 103	187 – 136	118
	N-Ch	170	139 – 154	126 – 87 – 120	134 – 113	183 – 146	130

Besides purely natural biological difference (higher female values, no matter the ads type, no matter the information load, no exceptions), T1 showed higher level on Ch vs. N-Ch with the difference better expressed for female ads. T2 demonstrated no difference for Ch and N-Ch. T3-incmp and T4 had steeper slopes on Ch but the difference was hardly statistically significant. T0 pattern replicated the ones discovered for Factors 1–2 showing higher values on N-Ch-syllables. T3-full for Ch-syllables compared to N-Ch-syllables in female ads demonstrated steeper inclination slope and no difference in the declination phase, while in male ads N-Ch-syllables showed bigger tone contour height compared Ch-syllables. Judging by this, T3-full showed no consistent pattern in expressing information load.

Canonical lexical tone substitutions. In this section, only the substitutes that run contrary to canonical tonal sandhi are viewed. Table 4 shows the number of substitutions depending on the three factors.

Table 4. Tone substitutions

Factor	Information load	Tone Change (%)
Factor 1		
	Ch	38
	N-Ch	62
Factor 2		
Soc.	Ch	40
	N-Ch	60
Com.	Ch	36
	N-Ch	64

Factor 3		
Female	Ch	40
	N-Ch	60
Male	Ch	37
	N-Ch	63

As can be seen from Table 4, there is a consistent difference of the number of tone substitutions depending on the information load alone (Factor 1), on the background of ads types (Factor 2) and on the background of gender (Factor 3). There is a strong tendency of less substitutions on informative syllables compared to uninformative ones. If we zoom in, on the whole, T3 (both varieties) and T2 were most sensitive to tone change (70.7% and 50.4% correspondingly out of all T3 and T2 occurrences), while T1 and T4 were considerably less sensitive (29.4% and 39.6% correspondingly). The tendency was sustained when separated into Ch and N-Ch with maximum differences.

As far as the substitutes, their ranks did not seem to depend much on any of the three factor for T1–T4. Thus, T1 was most frequently substituted with T2 – Rank 1 substitute (74% on Ch and 85% on N-Ch) and much less frequently with T4 – Rank 3 substitute (26% on Ch and 15% on N-Ch). T4 was almost always substituted with T1 – Rank 1 (100% on Ch and 91% on N-Ch) and very rarely with T2 – Rank 2 (9% on N-Ch). T2's most frequent substitute was T1 – Rank 1 (73% on Ch and 61% on N-Ch) and T4 substituted it not so often – Rank 2 (27% on Ch and 39% on N-Ch). Only in 4–5% T3-full was changed into T3-incmp which is a sandhi change. In 54% on Ch and 60% on N-Ch, T4 (Rank 1) was used instead, in 34% on Ch and 31% on N-Ch, T2 (Rank 2) was used. As for T3-incmp, unlike T3-full, it was most often replaced with T1 – Rank 1 (56% on Ch and 59% on N-Ch). T2 was Rank 2 substitute (31% on Ch and 29% on N-Ch) for T3-incmp. In 13% on Ch and 12% on N-Ch, T3-full was used instead of T3-incmp.

Contrary to T1 – T4, T0 showed a considerable difference in substituting tone ranks depending on all the three factors. According to Factor 1, T1 was Rank 1 substitute on Ch-syllables, while T4 was Rank 2 substitute (43% vs. 38%). On N-Ch syllables it was vice versa (T1 – 35%, T4 – 49%). T2 and T3-full as substitute had similar ranks (Rank 3 and Rank 4 correspondingly). According to Factor 2, in commercial ads, there was a wider spectrum of substitute: T1 – Rank 1 (51%), T4 – Rank 2 (28%), T2 – Rank 3 (16%), T3-full – Rank 4 (5%). In social ads, T4 was a major substitute accounting for 99% of tone changes (the remaining 1% was T2 substitute). Factor 3 differences were quite striking: in female ads, T1 was Rank 1 substitute, while in male ads, it was T4, both accounting for the vast majority of tone changes (85% and 78% correspondingly). Similarly, T2 was Rank 2 substitute for both female and male ads. In male ads, T1 was not used as a substitute, T3-full was Rank 3 substitute (6%), while in female ads, T3-full was not used as a substitute. In our material, T3-incmp was never used as a T0

substitute. This variation is consistent with Sun and Shih's results (2020) on anticipatory tonal coarticulation affecting Neutral tone.

3. Conclusion

This study was aimed to examine pitch pattern variation depending on 3 factors: information load (Factor 1), ads type (Factor 2) and gender (Factor 3). Each factor was viewed separately and in combination with other factors: Factor 1 + Factor 2 + Factor 3, Factor 2 + Factor 1, Factor 3 + Factor 1. The patterns were considered for each lexical tone separately: T1 – T4 and T0. The results prove the hypothesis in a number of ways. First, excluding purely biological male-female difference in pitch, the closest correlation was obtained for pitch values in the combination of Factor 2 and Factor 1 with a consistent increase in commercial ads vs. social ads and on informative syllables vs. uninformative ones with the exception of T0 that consistently increased on uninformative syllables. Second, depending on the tone, either the average height was increased or the declination/inclination slope was steeper.

Third, unlike in Tang and Li (2020), in this study, tonal substitutions were found. A clear correlation was revealed between the number of substitutes and the information structure of the ads that was consistent taking into account the ads type and gender. T3 and T2 were most frequently substituted with other tones. The substitutes themselves, however, did not demonstrate any dependence on any of the three factors. Most frequent substitutes were: T2 for T1; T1 for T2, T3-incmp and T4; T4 for T3-full. Unlike T1 – T4, T0 substitutes were strongly dependent on all the three factors. The differences mainly concerned T1 and T4 ranks as substitutes and the number of tones used as substitutes.

The obtained data show that lexical tones are characterized by quite consistent variation pattern that is predictable depending on information structure, speech genre/sub-genre and gender. The results provide certain implications for Chinese as L2 teaching methods in the field of developing speaking skills and identifying more and less important information in an oral text. General recommendations for Chinese as L2 learners that arise from the obtained results, depending on the particular tone are: (i) to increase pitch height and make inclination and/or declination slopes steeper to highlight more important words, (ii) to work on the interaction patterns of lexical tones and intonation and focus on most frequent tone substitutes that do not violate tones usage by Mandarin Chinese speakers and are common among them. This will contribute into speech naturalness that, according to Wu (2019), is crucial for tones discrimination.

The obtained results provide a tool to skillfully manipulate pitch together with other prosodic parameters like speech rate based on segments duration and intensity to function as effective speakers and listeners of Chinese as L2 for both general and specific purposes.

This study is limited within one speech genre and six speakers (even though they are balanced according to sex, age, education and profession). In advertising that has strict time limits, prosodic manifestations of information-based partition seems not so contrastive as it might be, e.g. in spontaneous casual speech, reports,

lectures, etc. Therefore, involving more genres with more speakers would enable to draw more general conclusions as well as conclusions specific for each genre.

REFERENCES

- BOERSMA, P., WEENINK, D., 2016. *Praat: Doing phonetics by computer* (Version 5.4.15) [Computer Program], <https://www.fon.hum.uva.nl/praat>, last accessed 2016/04/07.
- CANGEMI, F., BAUMANN, S., 2020. Integrating phonetics and phonology in the study of linguistic prominence. *Journal of Phonetics*. 81, 1 – 6.
- CHEN, Y., GUSSENHOVEN, C. (2008). Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics*, 36(4), 724 – 746.
- CHEN, A. C-H. & TSENG, S. C., 2019. Prosodic encoding in Mandarin spontaneous speech: Evidence for clause-based advanced planning in language production. *Journal of Phonetics*. 76, 1 – 22. <https://doi.org/10.1016/j.wocn.2019.100912>.
- DISCOURSE, 2018. *Diskurs kak universal'naya matrica verbal'nogo vzaimodejstviya* / D. D. Kholodova, G. N. Manaenko, S. N. Plotnikova [et al.]. In: O. A. Suleymanova (ed.). Moscow, URSS: Lenand. ISBN 978-5-9710-5080-3.
- HEUSINGER, K. VON., 1999. *Intonation and Information Structure*. Habilitationsschrift, accepted by the Faculty of Philosophy, University of Konstanz.
- LE, A., GE, C., 2021. Studies in Chinese phonetics. In Z. YE (ed.), *The Palgrave Handbook of Chinese Language Studies*. Singapore: Palgrave Macmillan. https://doi.org/10.1007/978-981-13-6844-8_21-1
- LEE, M.-K. (이무경 (대구보건대학교)), 2012. Variance characteristics of speaking fundamental frequency and vocal intensity depending on utterance conditions (발화조건에 따른 기본주파수 및 음성강도 변동의 특징). *Phonetics and Speech Sciences. The Korean Society of Speech Sciences (말소리와 음성과학)*. <https://doi.org/10.13064/ksss.2012.4.1.111>
- LI, Y., TANG, C., LU, J. WU, J. & CHUNG, E. F., 2021. Human cortical encoding of pitch in tonal and non-tonal languages. *Nature Communications*. 12(1161). <https://doi.org/10.1038/s41467-021-21430-x>
- LIU, Y., 2019. The influence of pitch contour on Mandarin speakers' perception of English stress. *University of Pennsylvania Working Papers in Linguistics*. 25(1), Article 20. <https://repository.upenn.edu/pwpl/vol25/iss1/20>
- ZHELTUKHINA M. R., Biryukova E. V., Gerasimova S. A. et al., 2017. Modern media advertising: effective directions of influence in business and political communication. *Man in India*. 97(14), 61 – 71.

- SUN Y., SHIH C., 2021. Boundary-conditioned anticipatory tonal coarticulation in Standard Mandarin. *Journal of Phonetics*. **84**, 1 – 27. <https://doi.org/10.1016/j.wocn.2020.101018>
- TANG, P., LI, S., 2020. The acoustic realization of Mandarin tones in fast speech. *Proc. Interspeech*. pp. 1938 – 1941. <https://doi.org/10.21437/Interspeech.2020-1274>.
- TANG, P., RATTANASONE, N. X., DEMUTH, K., 2017. Acoustic realization of Mandarin neutral tone and tone sandhi in infant-directed speech and Lombard speech. *JASA*, **142**, 2823. <https://doi.org/10.1121/1.5008372>
- TORSUEVA, I. G., 2009. *Intonatsiya i smysl vyskazyvaniya*. Moscow: LIBROKOM Press.
- VIKULOVA L. G., ZHELTUKHINA M. R., GERASIMOVA S. A., MAKAROVA I. V., 2020. *Communication. Theory and practice : textbook*. Moscow: VKN. ISBN 978-5-7873-1738-1.
- WEISS, B., TROUVAIN, J., BARKAT-DEFRADAS, M., OHALA, J., 2020. *Voice Attractiveness: Studies on Sexy, Likable, and Charismatic Speakers*. https://doi.org/10.1007/978-981-15-6627-1_hal-02965919
- WU, M., 2019. Effect of F0 contour on perception of Mandarin Chinese speech against masking. *PLoS ONE*, **14**(1): e0209976. <https://doi.org/10.1371/journal.pone.0209976>
- ZHANG, J., KARAVAEVA, V. G., 2019. Temporalnye kharakteristiki kitayskoy kommercheskoy radioreklamy [Temporal characteristics of Chinese commercial radio advertisement]. *Teoreticheskaya i prikladnaya lingvistika [Theoretical and Applied Linguistics]*, **5**(3), 248 – 272. https://doi.org/10.22250/2410-7190_2019_5_3_248_272

✉ **Svetlana V. Androsova**
ORCID iD: 0000-0001-6714-9684
Amur State University
Blagoveshchensk, Russia

✉ **Veronika G. Karavaeva**
ORCID iD: 0000-0003-2608-7362
Moscow City University
Moscow, Russia
E-mail: karavaeva.vg@yandex.ru

✉ **Zianwen Zhang**
ORCID iD: 0000-0001-9867-0245
Heihe University
Heihe, China