FRENCH /ø-u/ CONTRAST IN JAPANESE LEARNERS WITH AND WITHOUT GESTURE FEEDBACK IN MONOSYLLABLES: A DESCRIPTIVE ACOUSTIC STUDY

Claire Pillot-Loiseau[1], Takeki Kamiyama[2, 1], Ayako Kitamura[3]

[1]Laboratoire de Phonétique et Phonologie (LPP) UMR 7018 CNRS / Sorbonne Nouvelle University, France
[2]UR TransCrit (EA1569), Paris 8 Vincennes - Saint-Denis University, France [3]Sophia University, Tokyo, Japan
claire.pillot@sorbonne-nouvelle.fr, takeki.kamiyama@univ-paris8.fr, ayakokitamura@sophia.ac.jp

ABSTRACT

This study investigated French /ø/-/u/ contrast produced by 2 native speakers (NS) and 12 Japanese-speaking learners; 7 of them received 6 lessons on these vowels with hand and arm gesture feedback (GG), and the others 6 lessons with articulatory explanation (AG). Acoustic measurements (F1, F2, F3, F1/F2 distance, Euclidean distance - ED) of 2 repetitions of /dø/, /ʁø/, /du/ and /ʁu/ are compared to /ø/ and /u/ in isolation, before, after and two months after the lessons. F2 of /u/ decreases more for GG. F3 of /ø/ decreases more for AG. F1/F2 distance is smaller for /ʁø/ and /ʁu/ than for the other monosyllables, while it is smaller and closer to NS for /u/ produced by 4 learners in GG than the others. F2 of /u/ improved due to visual and kinesthetic (gesture) modalities.

Keywords: L2 vowel learning, French, Japanese, gesture as feedback, acoustics.

1. INTRODUCTION AND GOAL

The vowel system of the first language (L1) has a significant impact on the acquisition of the vowels of second languages (L2). For example, front rounded vowels are known to cause difficulty to speakers of languages without such vowels ([1-4], *inter alii*).

In present-day Parisian French, the close-mid front rounded /ø/ and the high back /u/ are phonemically contrasted. Achieving this contrast is thus of great importance for learners of L2 French and often poses a challenge to learners without such vowels in L1, as is the case with Japanese-speaking learners of French (JSLF). Native speakers of Tokyo Japanese tend to produce the French /u/ with a higher F2 than native speakers of French (NSF), resulting in a sound that is perceived mostly as /ø/ by native listeners of French [5]. This stems from the high non-front /u/ in Tokyo Japanese, commonly transcribed [ɯ]. Acoustically, it shows a higher F2 (> 1000 Hz) than the French /u/; articulatorily, the tongue is less retracted and the lips are less rounded than for French /u/ ([6-7]). Besides, it has been observed that JSLF sometimes realize /ø/ as [ɯ] [8], with F2≈1800Hz, F3≈2900Hz for beginners [9], whereas NSF produce this vowel with lower formants (F2≈1600 Hz and F3≈2700Hz for female speakers: [10], *inter alii*). These tendencies lead to the difficulty that JSLF face in realizing the phonemic contrast /ø/-/u/ [9], which is, in PAM-L2 terms [11], a case of single-category or category-goodness assimilation.

Spoken language proficiency is a multimodal task involving auditory, kinesthetic and visual perception; gestures are also part of it. It also involves the production of complex articulatory gestures. The advantage of hand and arm gesture feedback is that it requires no additional equipment, and it can be used in distance learning. This type of embodied feedback was mainly used to improve L2 prosody [12], but little for segmental contrasts [13, 14]. It has been successfully applied to the training of consonants in L2 teaching [13], and to the case of Japanese short and long vowels acquired by Catalan-speaking learners with and without hand gestures [15]. However, to our knowledge, the effect of training to produce the French vowel contrast /ø/-/u/ by JSLF with and without hand and arm gestures has never been studied. Our goal is to explore this effect for monosyllables: does hand gesture feedback improve vowel production over time more than articulatory explanation? What are the differences between learners?

2. METHOD

12 Japanese-speaking learners of French as a Foreign Language FFL (3 male, 9 female) took part in a series of 6 weekly individual lessons of 20 minutes each by videoconference (Zoom) given by an experienced instructor of FFL. An auditory identification test of minimal-pair monosyllables, longer words and phrases with /ø/ and /u/, a reading of similar monosyllables, words and phrases, and a survey (linguistic profile questionnaire at the beginning, satisfaction survey at the end) were administered before (Pre), immediately after (Post), and, for the auditory identification test and the reading, 2 months after the lessons (2M).

Participants were divided into two groups matched for the results of the first perception test, for their age, proficiency level of French and length of stay in

France. 7 learners in the "gesture" group (GG) received the training sessions with the movement of the arms and hands associated with the articulatory features of /ø/ and /u/: G1 (19 y.o., B1 in the Common European Framework of Reference for Languages), G2 (33, C1) G3 (31, A1), G4, (41, B2), G5 (31, B1), G6 (♂, 27, B1) and G7 (65, CEFR level unknown). 5 learners in the "articulatory" group (AG) had training sessions with articulatory explanations on /ø/ and /u/ provided in Japanese but without using gestures: A1 (♂, 62 y.o., B2 in CEFR), A2 (♂, 23, B2), A3 (39, CEFR level unknown), A4 (31, C1) and A5 (38, B1). Seven of them were born in Tokyo (G2, G3, A2, A4) or near Tokyo (G1, G7, A3). They started learning French as adults (learning experience of French: 3 months to 6 years), and had been living in France between 3 months and 21 years at the time of the study. One male (MN, 54 y.o) and one female native speaker (FN, 52 y.o.) of French participated as a reference for comparison.

All training sessions were devoted to vowel(?) production. The training began with the repetition of CV monosyllables (e.g., /ku/, /kø/), followed by the repetition of bisyllabic words containing /ø/ and /u/ in phrase-final (phrase-accented) syllables and non-final syllables (e.g., *couteau* /k**u**to/, *jeudi* /ʒ**ø**di/). Next, learners repeated di- and trisyllabic words and longer phrases containing two tokens of the same target vowel (e.g., *heureux* /**øʁø**/, *kangourou* /kɑ̃g**u**ʁ**u**/) and then two different target vowels (e.g. *coûteux* /k**u**t**ø**/, session 3). They then repeated minimal pairs (e.g., *bourrer* /b**u**ʁe/; *beurrer* /b**ø**ʁe/, session 4). In session 5, they read words and phrases introduced in the first 4 sessions. Finally, they read the lyrics of a song and a short text containing the target vowels (session 6). During the last two sessions, the subjects in GG read the words and phrases first with gestures and then without them. The arm and hand movements of the selected gestures, inspired by [16], are aimed to represent the articulatory characteristics of the target vowel (e.g., tension, tongue position). When pronouncing /u/, the elbows are raised laterally to shoulder height, with the hands and arms placed horizontally (palms down), then the elbows are pulled back vigorously. This gesture is intended to encourage Japanese-speaking learners to place their tongue in a posterior position. For the target vowel /ø/, the arms are stretched forward (hands with palms up and down), as if to hold a long pipe between the two outstretched arms. As articulating /ø/, the arms and hands are moved slightly forward. This gesture is intended to encourage Japanese speakers to move their tongue and lips forward.

The recorded material for each test and each subject included /ø u ɑ̃ ɔ̃/ in isolation and 24 CV monosyllables where V=/ø u ɑ̃ ɔ̃/ and C=/b d g z ʒ ʁ/, presented to the participants in a random order and read aloud by them. The list of these 28 items were repeated twice and a total of 56 monosyllable tokens were obtained. A list of 40 minimal-pair words and phrases of 2 to 5 syllables containing the target vowels were also read aloud twice during this test. Only the monosyllables "eu"/ø/, "ou"/u/ (in isolation), "deu"/dø/, "dou"/du/ (dental context) and "reu"/ʁø/, "rou"/ʁu/ (uvular context) are analysed in this article (12 items per subject).

The acoustic signals, recorded at 44100 Hz, 16 bits, were subjected to the analysis of the first three formants, semi-automatically measured at 50% of vowel duration, using Praat [17]. This allowed measuring the Euclidean distance (ED) between /ø/ and /u/ in order to quantify the degree of achievement of the phonemic contrast, and calculating the distance between F1/F2 for [u] [10]. French focal vowels [18] in isolation are known to be characterized by a distance of less than 800Hz [10] between the neighbouring formants: F1 and F2 for /u o ɔ (ɑ)/.

The present study is descriptive. ANOVA tests were only applied to the formant frequency measurements of the productions of all same-sex learners in a group (GG or AG, Fig. 1). The small number of measurements per subject and per monosyllable did not allow further acoustic analyses (Fig. 2 to 4).

3. RESULTS

3.1. Formant frequencies

Figure 1 shows the evolution, throughout Pre, Post, and follow-up 2M, of the average frequencies of F1 to F3 of all learners of the same sex combined, broken down into groups ("articulatory" AG vs. "gesture" GG), all monosyllables of the same target vowel combined. The formant values of /ø/ do not change significantly overall and remain close to those of native speakers. Although we observe that from Pre to 2M, F1 of /ø/ and /u/ increases only for the male learner of GG and approaches that of MN, this increase is not statistically significant ($F_{(3,44)}$=0.21; p=0.889 for /ø/; $F_{(3,44)}$=0.21; p=0.929 for /u/). There was a very slight increase in F2 for /ø/ for women and men in GG (respectively $F_{(3,224)}$=0.59; p=0.62, and $F_{(3,44)}$=0.18; p=0.899), while there was a decrease in F3 for /ø/ and /u/, especially for AG. The most salient formant change observed (yellow rectangles, Fig. 1) is a significant decrease in F2 of /u/ for GG, especially between Post and 2M ($F_{(3,224)}$=9.7; p<0.0001 for female and $F_{(3,44)}$=5.71; p=0.0018 for male learners in GG).

3.2. F1-F2 distance for /ø/ and /u/

Figure 2 shows the evolution, throughout Pre, Post and 2M, of the average F1-F2 distance for each speaker for “eu” /ø/, “deu” /dø/ and “reu” /ʁø/. F1-F2 distance of “reu” is smaller than that of “eu”, and “deu” for 10 out of the 12 learners. Moreover, the evolution is variable according to the subjects: there is little difference between the two groups. It does not result in a convergence towards the values of the native subjects (dotted lines). In addition, G1, G2, G4 and G5 had F1-F2 distance close to that of the native subjects already in Pre. Overall, more subjects in GG show F1-F2 distance as small as 800Hz to 1200Hz already in Pre than those in AG (higher values for this latter group).

Figure 3 shows the evolution, throughout Pre, Post and 2M, of the average F1-F2 distance for each speaker for “ou” /u/, “dou” /du/ and “rou” /ʁu/. F1-F2 distance of /u/ in "rou" is smaller than that of "ou" and "dou" for 8 out of the 12 subjects. Moreover, the evolution is variable from one learner to another: the variability is remarkable for subjects G1 and G6 with a decrease of this distance in Post and 2M, except for "dou" for G1. A2 progressed by decreasing his F1-F2 distance by 30% (-400Hz) between Pre and Post, but lost this gain for "ou" and "rou" in 2M. In addition, G2 for "rou", G3, G4 for "dou", G5 and G6 for "rou" showed small F1-F2 distance, close to the native values already in Pre. Overall, more subjects in GG showed F1-F2 distances around 400Hz (value of "rou" for both natives and "ou" for the female native), especially for "rou", than subjects of Group AG.

3.3. Euclidean Distance (ED)

Figure 3 shows the evolution, throughout Pre, Post, and follow-up 2M, of the average Euclidean Distance (ED) for each speaker for “eu-ou” /ø-u/, “deu-dou” /dø-du/ and “reu-rou” /ʁø-ʁu/ contrasts. A larger ED suggests that /ø/ and /u/ were produced more distinctly. There is an increase of ED for G1 for “eu-ou”, and from Pre to Post for “deu-dou” and “reu-rou”, G04 (for “reu-rou” and “eu-ou” at 2M), G6 (for “eu-ou” and “deu-dou” at 2M), G7 for “deu-dou” at 2M), but also for “reu-rou” for A2. However, there is a decrease in ED for G3 (except for “deu-dou”), G4 for “deu-dou”, G5 (except for “eu-ou”) and G7 for “reu-rou”. Except for subjects G1 for “eu-ou”, G3 for “eu-ou” and “reu-rou”, G6 for “eu-ou” and “deu-dou” at 2M, ED is not greater for participants in GG than for those in AG. In addition, subjects A1 and A2 increase their ED for vowels in isolation from Pre to Post, and then to 2M.

4. DISCUSSION

The results for ED show that the use of arm and hand gestures moderately helps Japanese learners to better distinguish /ø/ and /u/, compared to learners in AG. However, 4 subjects out of 7 in GG decreased F2 and F1-F2 distance for /u/ in isolation and in dental and uvular contexts. The preceding consonant influences F1-F2 distance of /u/ more for AG than for GG, but differently depending on the subject. In GG, F1-F2 distance of "reu" and "rou" is globally smaller due to the uvular context that retracts the vowels /ø/ and /u/.

The gestures used for L2 Spanish acquisition by native Dutch speakers improved only /u/, not /θ/ [14]. Prosodic (not segmental) gestures improved French /y, ø, œ/ for 57 Catalan learners [19]. However, in both of these studies, learners only *looked* at the gesture. In our study, learners were additionally asked to *produce* this gesture at the same time as they pronounced the target vowels.

The moderate effect of the gesture action is probably due to the fact that the training was mainly based on repetition rather than reading. In addition, each training session was short (20mn). However, subjects were more motivated by gesture instruction, even when following the lessons through Zoom.

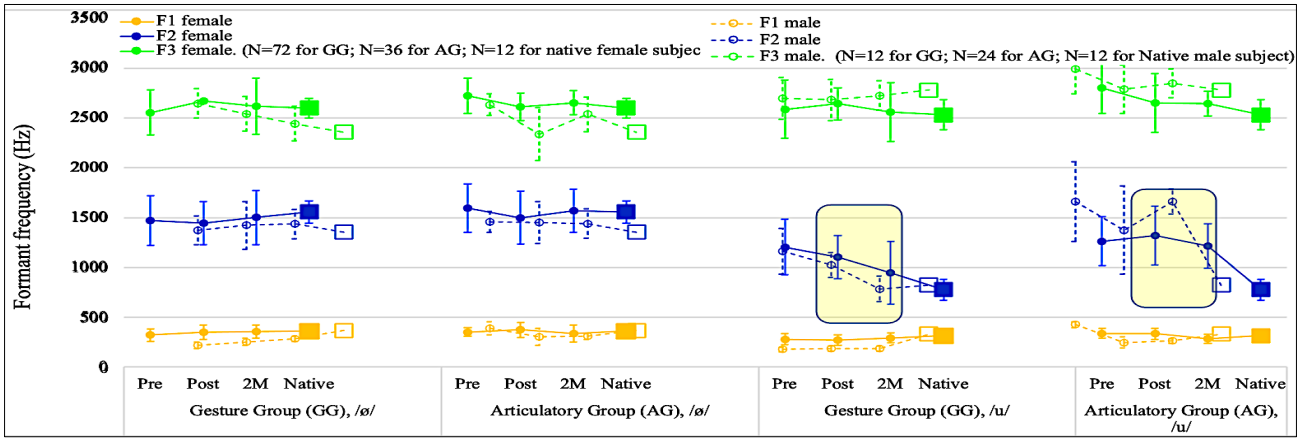

Figure 1: F1 to F3 frequencies before (Pre), just after (Post) and 2 months after (2M) the 6 lessons (all monosyllables included, MN and FN, and all subjects of Groups GG and AG). Light yellow rectangles: the most relevant changes between GG and AG (F2 lowering for /u/).

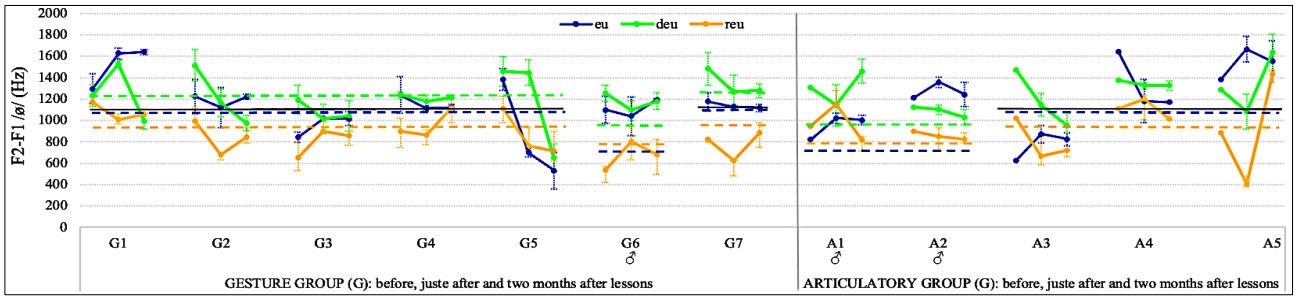

Figure 2: Mean (N = 2 for each monosyllable) F1-F2 distance and standard deviation (SD) for each speaker for "eu" /ø/, "deu" /dø/ and "reu" /ʁø/. Within each speaker, the lines connect Pre, Post and 2M. Horizontal dotted lines: values of MN and FN for the same monosyllables. Black horizontal solid lines: value of /ø/ by 40 native French-speaking women [10].

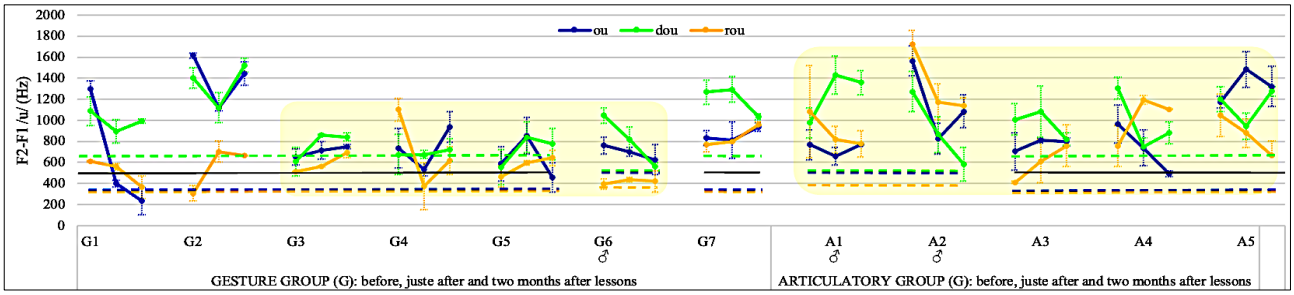

Figure 3: Mean (N = 2 for each monosyllable) F1-F2 distance and SD for each speaker for "ou" /u/, "dou" /du/ and "rou" /ʁu/. Within each speaker, the lines connect Pre, Post and 2M. Horizontal dotted lines: values of MN and FN for the same monosyllables. Black horizontal solid lines: value of /u/ by 40 native French-speaking women according to [10]. Light yellow rectangles: most relevant changes between GG and AG (smaller F1-F2 distance for most subjects in GG).

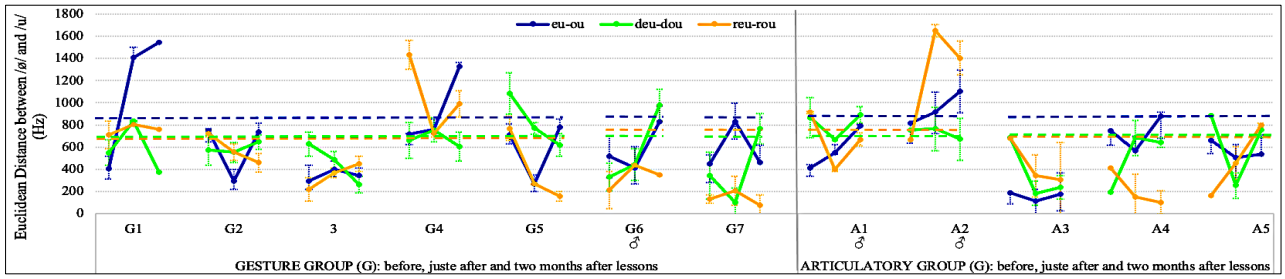

Figure 4: Mean (N = 2 for each monosyllable) ED and SD for each speaker between "eu-ou" /ø u/, "deu-dou" /dø du/ and "reu-rou" /ʁø ʁu/. Within each speaker, the lines connect Pre, Post and 2M. Horizontal dotted lines: values of MN and FN for the same monosyllables.

Furthermore, although the learners in GG were more satisfied with the training than those in AG (satisfaction survey), motivation and the ability to reproduce the gesture faithfully to the model varied among learners, according to the instructor who led their training (the third author). The nature of the gesture can also influence the production of the vowel phoneme: the gesture of /u/ insists more on the posterior position of the tongue than on the rounding and protrusion of the lips: this may be the reason why we do not observe a drop in F3 of this vowel for the learners of GG, contrary to those of AG: only the latter received an articulatory explanation on the lip rounding of /u/ and /ø/.

The learner's individual profile is another effect to be considered, as reported by [20]: for example, learner A2 is fluent in English, Spanish and French (which is his 4th language). Subject G7 is 65 y.o. and expressed great difficulty in pronouncing /ø/ and /u/. She started learning French after the age of 40.

5. CONCLUSION AND PERSPECTIVES

Such a multimodal approach can train the ear to perceive unfamiliar speech features by presenting them through visual and kinesthetic modalities. In addition, controlling pronunciation through hand and arm gestures bypasses ingrained patterns in the natural voice that may be difficult to correct [21].

This preliminary research needs to be complemented by examining more data from these monosyllables and other elements of the corpus already recorded (other monosyllables and phrases); studying vowel stability taking into account formant measures at 25% and 75% of each vowel; conducting a perceptual evaluation of these learners' productions by native listeners of French.

6. ACKNOWLEDGMENTS

This work is part of the program "Investissements d'Avenir" overseen by the French National Research Agency, ANR-10-LABX-0083 (Labex EFL).

7. REFERENCES

[1] Flege, J. E. 1987. The production of new and similar phones in a foreign language: evidence for the effect of equivalence classification. *J. of Phonetics* 15, 47-65.

[2] Levy, E. S., Law II, F. F. 2010. Production of French vowels by American-English learners of French: Language experience, consonantal context, and the perception-production relationship. *J. Acoust. Soc. Am.* 128(3), 1290-1305.

[3] Rochet, B. L. 1995. Perception and production of second-language speech sounds by adults. In: Strange, W. (ed), *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore: York Press, 379-410.

[4] Lauret, B. 2007. *Enseigner la prononciation du français : questions et outils*. Paris: Hachette.

[5] Kamiyama, T., Vaissière, J. 2009. Perception and production of French close and close-mid rounded vowels by Japanese-speaking learners. *Acquisition et interaction en langue étrangère Aile... Lia* 2, 9-41.

[6] Bothorel, A., Simon, P., Wioland, F., Zerling, J. P. 1986. *Cinéradiographie des voyelles et consonnes du français*. Strasbourg : Publications de l'Institut de Phonétique de Strasbourg.

[7] Uemura, Y., Takada, M. (Kokuritsu Kokugo Kenkyûjo [National Institute for the Japanese Language]) 1990. *Nihongo no boin, shiin, onsetu: chouon undou no jikken-onseigakuteki kenkyû* [Vowels, consonants and syllables in Japanese: an experimental phonetic study on articulatory gestures]. Tokyo: Shûei shuppan.

[8] Kamiyama, T., Detey, S., Kawaguchi, Y. 2017. Les japonophones. In: S. Detey, I. Racine, Y. Kawaguchi, J. Eychenne (eds), *La prononciation du français dans le monde : du natif à l'apprenant*. Hachette, 155-161.

[9] Marushima, N., Detey, S., Kawaguchi, Y. 2010. Caractéristiques phonétiques des voyelles orales arrondies du français chez des apprenants japonophones. *Flambeau* 36, 53-72.

[10] Georgeton, L., Paillereau, N., Landron, S., Gao, J., Kamiyama, T. 2012. Analyse formantique des voyelles orales du français en contexte isolé : à la recherche d'une référence pour les apprenants de FLE. *Proc. Journées d'Étude sur la Parole-TALN-RECITAL* Grenoble, 145–152.

[11] Best, C. T., Tyler, M. D. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In: M. J. Munro & O.-S. Bohn (eds), *Language experience in second language speech learning: In honor of James Emil Flege*. John Benjamins, 13-34.

[12] Baills, F., Alazard-Guiu, C., Prieto, P. 2022. Embodied Prosodic Training Helps Improve Accentedness and Suprasegmental Accuracy. *Applied Linguistics* 43(4), 776-804.

[13] Xi, X., Li, P., Baills, F., Prieto, P. 2020. Hand gestures facilitate the acquisition of novel phonemic contrasts when they appropriately mimic target phonetic features. *Journal of Speech, Language, and Hearing Research* 63(11), 3571-3585.

[14] Hoetjes, M., Van Maastricht, L. 2020. Using gesture to facilitate L2 phoneme acquisition: The importance of gesture and phoneme complexity. *Frontiers in Psychology*, *11*, 575032.

[15] Li P, Baills F, Prieto P. 2020. Observing and producing durational hand gestures facilitates the pronunciation of novel vowel-length contrasts. *Studies in Second Language Acquisition* 42(5),1015-1039.

[16] Gladić, V. A. 1982. *Le graphisme phonétique : du geste phonétique au geste graphique*. Nathan.

[17] Boersma, P., Weenink, D. 2010. *Praat: doing phonetics by computer* [Computer program]. Version 6.2.23, http://www.praat.org/

[18] Schwartz, J.-L., Boë, L.-J., Vallée, N., Abry, C. 1997. The Dispersion-Focalization Theory of vowel systems. *Journal of Phonetics* 25(3), 255-286.

[19] Li, P., Baills, F., Baqué, L., Prieto, P. 2022. The effectiveness of embodied prosodic training in L2 accentedness and vowel accuracy. *Second Language Research*, 02676583221124075.

[20] Li, J. J., Ayala, S., Harel, D., Shiller, D. M., McAllister, T. 2019. Individual predictors of response to biofeedback training for second-language production. *J. Acoust. Soc. Am.* 146(6), 4625-4643.

[21] D'alessandro, C., Rillard, A., Lebeux, S. 2011. Chironomic stylization of intonation. *J. Acous. Soc. Am.* 3129, 1594–1604.