On Improving the Pronunciation of French /r/ in Chinese Learners by Using Real-time Ultrasound Visualization

Yaru Wu, Cédric Gendrot, Pierre Hallé, Martine Adda-Decker

On Improving the Pronunciation of French /r/ in Chinese Learners by Using Real-time Ultrasound Visualization

Yaru Wu[1], Cédric Gendrot[1], Pierre Hallé[2], Martine Adda-Decker[2]

[1] Université Sorbonne-Nouvelle Paris 3 - Laboratoire de Phonétique et Phonologie UMR 7018
[2]Laboratoire de Phonétique et Phonologie UMR 7018
yaru.wu@univ-paris3.fr, cedric.gendrot@univ-paris3.fr, pierre.halle@univ-paris3.fr, martine.adda-decker@univ-paris3.fr

ABSTRACT

Chinese natives are known to perform poorly in /r/ pronunciations when learning French. They tend to produce their native /x/ sound. In this study, we compare two learning methods of French /r/: the traditional method using colloquial explanations of the pronunciation and a newly developed method using ultrasound images to demonstrate tongue movements in real-time. Two distinct groups including 10 Chinese learners participated in the experiments. A preliminary perceptual experiment confirmed their weak discrimination capacity between French /r/ and Mandarin /x/. Acoustic HNR and COG measures on pre- and post-training productions allowed us to quantify their progress in productions. The proposed ultrasound technique is shown to be efficient for foreign language pronunciation learning. By using real-time ultrasound visual feedback, Chinese learners have speeded up their ability to depart from Mandarin /x/ towards French /r/ articulation and to produce a better French /r/ at the end of the training.

Keywords: ultrasound, tongue position, real time, visual feedback, French /r/ learning

1. INTRODUCTION

Ultrasound imaging makes safe and non-invasive real-time visualisation available for language learners [1] [5] [10] [13]. Learners can visualize directly the tongue shape, the tongue position and the tongue movements corresponding to the realisation of a given phoneme. The ultrasound images can be recorded, synchronised with the sound pronounced.

Although ultrasound imaging has its limits, such as the quality of the ultrasound images being speaker-dependant, it delivers visual feedback which may be of great help to language learners in real-time [4] [9] [11] [14].

Chinese learners of French are known to have difficulty pronouncing the French /r/ correctly which may even result in communication problems: French natives may misrecognize words pronounced by Chinese learners.

French /r/ is highly variable depending on phonetic context and speakers. Its main realisation corresponds to a uvular fricative whereas Chinese /x/, which is the closest to the French /r/ sound, is a velar fricative. The realisation difference between the two phonemes is related to the place of constriction of the tongue. Therefore, ultrasound imaging may be successfully envisaged to help Chinese learners to improve their pronunciation, to reduce their foreign accent and to avoid communication errors.

In this study, we analyse the perceptions and the productions of French /r/ in onset position by Chinese learners before and after training sessions.

We attempt to find out:

- whether Chinese learners have perceptual difficulties identifying French /r/
- whether Chinese learners tend to assimilate French /r/ with Chinese /x/ when they pronounce the French /r/ [6]
- whether real-time ultrasound imaging is an efficient tool for improving the pronunciation of French /r/ by Chinese learners

2. EXPERIMENTS

2.1. Subjects

Ten native speakers of Mandarin Chinese were involved in our /r/ study including a perceptual test followed by a production experiment which aimed at comparing a traditional method (TM) to a method including ultrasound (US). Five native speakers of French participated as control subjects in the perception test and one French speaker produced the corresponding reference material (see Figure 1). The Chinese subjects were all born and raised in Mainland China; they all learnt English as a first foreign language and French as a second foreign language; although, they come from different areas of China, they all speak mandarin Chinese fluently from a very early age. The French subjects all speak standard French.

2.2. Equipment

The equipment used for the traditional method (henceforth TM) Group includes a Windows XP system computer, a sound card (UA-25 EX) and a micro phone (Technica ATM 33A) which was fixated at about 20 cm from the mouth. The software used for data recording and processing were Praat 3.9[3] and R 3.0.3[15]. The sampling frequency for the audio recordings was 44100 Hz.

Beyond the equipment used for the TM Group, the ultrasound (henceforth US) Group benefited from a specific ultrasound equipment (Mindray DP6600 engine, a SyncBrightUp synchronization box and a probe fixator).

The audio sampling rate was 44.1 kHz. The software AAA 14 Beta (Articulate Assistant Advanced) and the Mattong tool (based on EdgeTrak and Matlab) [7] were used for recording and processing. US images and the coordinates of 33 equally distributed points of each tongue shape were semi-automatically extracted in AAA and Mattong. The coordinates were then used to draw tongue shapes in R. For both groups, audio data were recorded, segmented and analyzed on Praat. HNR and COG values of the data were computed using Praat scripts [8].

2.3. Materials

2.3.1. Perception tests

The stimuli used for the identification tests were syllables with /r/, /l/, /w/ or Chinese /x/ in onset position followed by /a/, /o/, /ə/, /i/, /u/ or /y/ vowels, since /x/ exists only in onset position in Chinese. The vowels /a/, /o/, / /ə/ /, /i/, /u/ and /y/ were chosen to build syllables in French as well as in Chinese. For example, for the phoneme /r/, the syllables in question were /ra/, /ro/, /rə/, /ri/, /ru/ and /ry/. The syllables were pronounced in the context "Je dis … deux fois" (I say … two times). The phonemes /l/, /w/ and /x/ were used as distractors in the identification tests.

The stimuli were pronounced by a French phonetician who also mastered the pronunciation of Chinese phonemes. The Chinese syllables were pronounced in neutral tone and the pronunciations of these syllables were verified by a native Chinese.

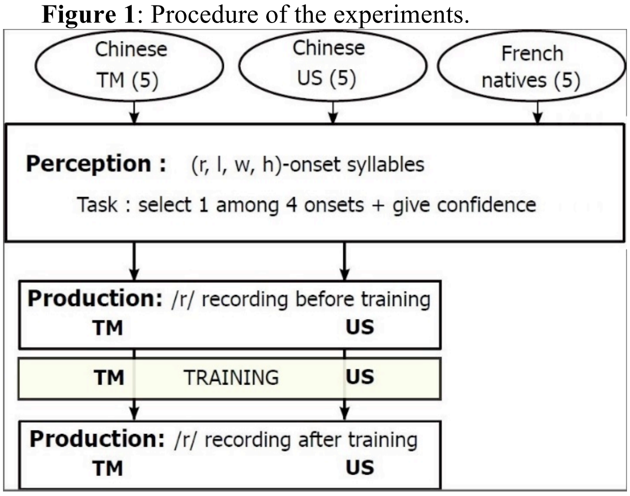
2.3.1. Production

The subjects were asked to pronounce French /r/ and Chinese /x/ in onset position followed by one of the six vowels in Mandarin Chinese. For example, for the phoneme /r/, the syllables in question were /ra/, /ro/, /rə/, /ri/, /ru/ and /ry/. The syllables were pronounced in the context "Je dis … deux fois" (I say … two times).

2.4. Procedure

The training took place in the recording studio of ILPGA, University of Sorbonne Nouvelle. Two training groups of five learners were organized for testing the two methods: one training group without ultrasound imaging but with repeated colloquial explanation (TM) before and after the training session, the other training group with ultrasound imaging (US) before and after the training session (see Figure 1 for details of the experimental flow).

Figure 1: Procedure of the experiments.



2.4.1. Identification tests

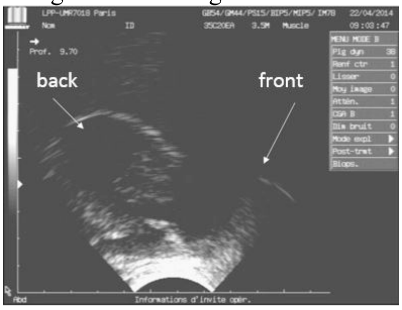
All Chinese subjects were asked to do an identification test before the pronunciation training. The subjects were asked to identify the French /r/, along with /l/, /w/ and Chinese /x/, according to the pre-recorded syllables. Five native French were also asked to participate in the identification tests.

2.4.2. Production Training

At the beginning of the training, the subjects of the TM Group were asked to pronounce the French /r/ and Chinese /x/ stimuli in onset position. They were then explained how to pronounce the standard French /r/ and trained to improve their pronunciation of the phoneme. After the training, the subjects were asked to pronounce the stimuli again.

For the US Group, the subjects were asked to pronounce the same stimuli. They were then shown a short ultrasound imaging video and explained the position of the tongue (Figure 2) while a standard French /r/ was pronounced. During the training, the subjects could see in real-time whether they put their tongue in the right position. Subjects hence tried to improve themselves according to the tongue position on the screen. Ultrasound imaging videos and the sound tracks of the stimuli were registered.

Figure 2: Mid-sagittal image of the tongue. The anterior tongue is on the right.



3. RESULTS

Both perception and production results were analysed in this study.

3.1. Perception results

We obtained the results of the identification tests for /r/, /l/, /w/ and /r/ from Praat. Since we are only interested in the results concerning /r/, we show the results concerning /r/ in Table 1. For the results concerning the distractors (/l/, /w/ and /x/), neither the French auditors nor the Chinese auditors had any difficulty identifying these phonemes.

We can see in Table 1 that Chinese learners have difficulty identifying French /r/ and distinguishing French /r/ from Chinese /x/. However, native French don't have difficulty identifying the French /r/ at all.

Table 1: French /r/ identification rates by French and Chinese subjects

	French /r/ identified as			
	r	x	l	w
French auditors	100	0	0	0
Chinese auditors	48	52	0	0

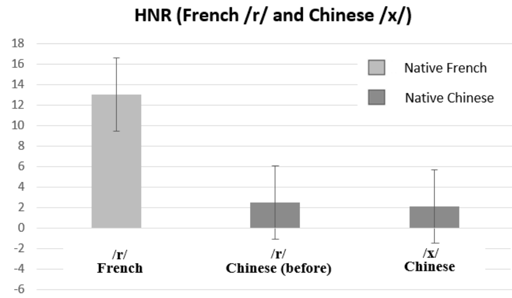
3.2. Production results

HNR (harmonic-to-noise ratio) value and COG (center of gravity) value were analysed for both groups. The higher HNR is, the more voiced the sound is. Therefore, the HNR of French /r/ is higher than that of Chinese /x/. The higher COG is, the higher the mean frequency is. Therefore, the COG of French /r/ is less high than that of Chinese /x/.

HNR values (dB) of French /r/ pronounced by a French phonetician, French /r/ pronounced by Chinese learners before training and Chinese /x/ pronounced by the same Chinese learners of French are shown on Figure 3.

The results of Figure 3 confirm that Chinese learners tend to assimilate French /r/ to Chinese /x/ in terms of HNR values. According to the ANOVA result, the HNR value of /r/ pronounced by the Chinese learners before training is significantly different from that of the standard French /r/($p<0.05$).

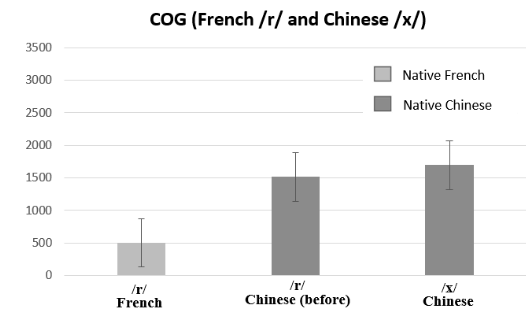
Figure 3: HNR values (dB) on French /r/ pronounced by a native French (left), French /r/ by Chinese learners before the training (middle) and Chinese /x/ by Chinese learners (right).



COG values (Hz) of French /r/ pronounced by a French phonetician, French /r/ pronounced by Chinese learners before training and Chinese /x/ pronounced by the same Chinese learners of French are shown on Figure 4.

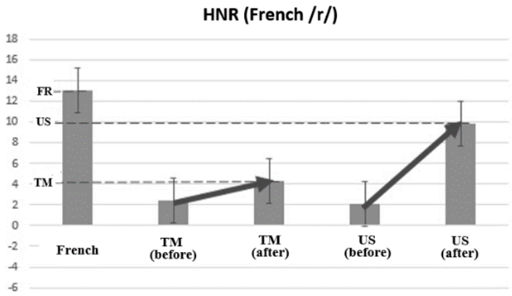
The results of Figure 4 confirmed that Chinese learners tend to assimilate French /r/ to Chinese /x/ in terms of COG values. According to the ANOVA result, the COG value of /r/ pronounced by the Chinese learners before training is significantly different from that of the standard French /r/($p<0.05$).

Figure 4: COG values (Hz) on French /r/ pronounced by a native French (left), French /r/ by Chinese learners before the training (middle) and Chinese /x/ by Chinese learners (right).



On Figure 5, HNR values (dB) of French /r/ pronounced by a French phonetician and French /r/ pronounced by Chinese learners before and after training for both TM Group and US Group can be observed. Figure 5 confirms that subjects of US Group have a better production of French /r/ after the ultrasound visualization training than that of TM Group in terms of HNR values. The ANOVA result of the HNR value of /r/ pronounced by subjects of US Group after the training and that of TM Group after the training is significant ($p<0.05$); the ANOVA result of the HNR value of /r/ pronounced by subjects of TM Group after the training and that of standard French /r/ is significant ($p<0.05$); the ANOVA result of the HNR value of /r/ pronounced by subjects of US Group after the training and that of standard French /r/ is not significant ($p>0.05$).

Figure 5: HNR values of French /r/ pronounced by a native French and by Chinese learners before and after the training for both TM Group and US Group.



On Figure 6, COG values (Hz) of French /r/ pronounced by a French phonetician and French /r/ pronounced by Chinese learners before and after training for both TM Group and US Group can be observed. Figure 6 confirms that subjects of US Group have a better production of French /r/ after the ultrasound visualization training than that of TM Group in terms of COG values. The ANOVA result of the COG value of /r/ pronounced by subjects of US Group after the training and that of TM Group after the training is significant ($p<0.05$); the ANOVA result of the COG value of /r/ pronounced by subjects of TM Group after the training and that of standard French /r/ is significant ($p<0.05$); the ANOVA result of the COG value of /r/ pronounced by subjects of US Group after the training and that of standard French /r/ is not significant ($p>0.05$).

Figure 6: COG values of French /r/ pronounced by a native French and by Chinese learners before and after training for both TM and US Groups.

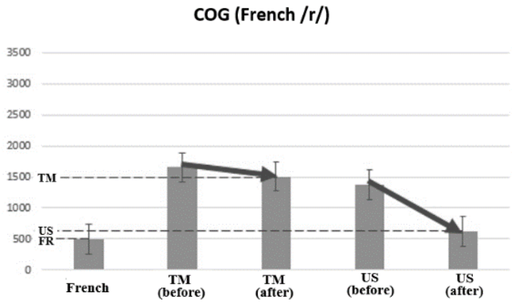
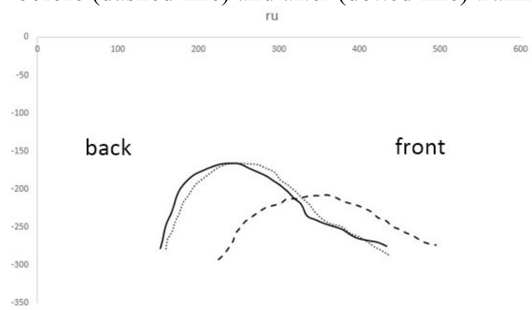


Figure 7 is an example of the tongue shapes of /ru/ pronounced by one of the Chinese learners of US Group. It shows that the subject has a better production of French /ru/ after the ultrasound visualization training. The ANOVA result of the tongue position before and after training is significant ($p<0.05$).

Figure 7: Tongue shapes of /ru/ of a native French (full line) and a Chinese learner of US Group before (dashed line) and after (dotted line) training.



4. CONCLUSION AND DISCUSSION

The perception study confirms that Chinese learners have difficulty identifying French /r/ and they tend to assimilate it to Chinese /x/. Our analyses confirm that Chinese learners tend to assimilate French /r/ to Chinese /x/ when they pronounce the French /r/. The improvement of the pronunciation of French /r/ produced by Chinese learners who benefited from an ultrasound imaging related training confirms that real-time ultrasound imaging is an efficient tool for language learning in terms of improving the pronunciation of language learners.

We have not asked native French to identify the /r/ pronounced by Chinese learners yet, but the experience is underway. We will ask native French to identify and to rate the /r/ pronounced both before and after the training experiences for both groups.

A current study of pronunciation training of French /r/ involving real-time ultrasound imaging in the contexts other than the onset position for Chinese learners is taking place. We hope to determine a training system by using ultrasound imaging and help language learners improve their pronunciation efficiently.

5. ACKNOWLEDGEMENTS

This work was supported by the French Investissements d'Avenir - Labex EFL program (ANR-10-LABX-0083).

It was also partially funded by the ANR young researchers program (ANR-13-JSH2-0005) and the European Commission via the i-Treasures (FP7-ICT-2011-9-600676) project.

6. REFERENCES

[1] Archangeli, D., Kennedy, R., Baker, A., Racy, S. 2004. Ultrasound techniques for phonological research. Department of Linguistics Colloquium: University of Arizona.

[2] Bernhardt, B., Gick, B., Bacsfalvi, P., Adler-Bock, M. 2005. Ultrasound in speech therapy with adolescents and adults. Clinical Linguistics and Phonetics, 19, 605-617.

[3] Boersma, P. and Weenink, D. "Praat: doing phonetics by computer" (Version 5.3.22) [Computer program]. From http://www.praat.org/, Retrieved September 15, 2012.

[4] Chi-Fishman, G. 2005. Quantitative lingual, pharyngeal and laryngeal ultrasonography in swallowing research: A technical review. Clinical Linguistics and Phonetics, 19, 589-604.

[5] Davidson, L. 2006. Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. J. Acoust. Soc. Am. 120, 407-415.

[6] Flege, J. E. 1995. "Second language speech learning: Theory, findings, and problems." in W. Strange, Speech perception and linguistic experience: Issues in cross-language research. Baltimore, York Press, 233-277.

[7] Fux, T. « Mattong », http://thibautfux.com/mattong/.

[8] Gendrot Cédric. « Script analyse tout » http://gendrot.ilpga.fr/scripts.htm/

[9] Gick, B., Bird, S., Wilson, I. 2005. Techniques for field application of lingual ultrasound imaging. Clinical Linguistics and Phonetics, 19, 503-514.

[10] Iskarous, K. 2005. Patterns of tongue movement. *Journal of Phonetics* 33(4), 363-381.

[11] Li, M., Kambhamettu, C., Stone, M. 2005. Automatic contour tracking in ultrasound images. Clinical Linguistics and Phonetics, 19, 545–554.

[12] Mielke, J., Baker, A., Archangeli, D., Racy, S. 2005. Palatron: A technique for aligning ultrasound images of the tongue and palate. *Coyote Papers* 14, 97-108.

[13] Parthasarathy, V., Stone, M., Prince, J. 2005. Spatiotemporal visualization of the tongue surface using ultrasound and kriging (surfaces). Clinical Linguistics and Phonetics, 19, 529–544.

[14] Stone, M. 2005. A guide to analysing tongue motion from ultrasound images. Clinical Linguistics and Phonetics, 19, 455–501.

[15] The R Project for Statistical Computing http://www.r-project.org/.