Putting German [ʃ] and [ç] in two different boxes:
native German vs L2 German of French learners

Jane Wottawa[1], Martine Adda-Decker[1], Frédéric Isel[2]

[1]LPP, UMR 7018 CNRS - U. Paris 3 / Sorbonne Nouvelle
[2]MoDyCo, UMR 7114 CNRS - Université Paris Ouest Nanterre La Défense

jane.wottawa@univ-paris3.fr, martine.adda-decker@univ-paris3.fr, fisel@u-paris10.fr

Abstract

French L2 Learners of German (FG) often replace the palatal fricative /ç/ absent in French with the post alveolar fricative /ʃ/. In our study we investigate which cues can be used to distinguish whether FG speakers produce [ʃ] or [ç] in words with the final syllables /ɪʃ/ or /ɪç/. In literature of German as an L2, to our knowledge, this contrast has not yetbeen studied. In this perspective, we first compared native German (GG) productions of [ʃ] and [ç] to the FG speaker productions. Comparisons concerned the F2 of the preceding vowel, the F2 transition between the preceding vowel and the fricative, the center of gravity and intensity of the fricatives in high and low frequencies. To decide which cues are effectively choices to separate [ʃ] and [ç], the Weka interface in R (RWeka) was used. Results show that for German native speech, the F2 of the preceding vowel and the F2 transition are valid cues to distinguish between [ʃ] and [ç]. For FG speakers these cues are not valid. To distinguish between [ʃ] and [ç] in FG speakers, the intensity of high and low frequencies as well as the center of gravity of the fricatives help to decide whether [ʃ] and [ç] was produced. In German native speech, cues furnished only by the fricative itself can as well be used to distinguish between [ʃ] and [ç].

Index Terms: L2 German speech, fricatives, classification, R, Weka

1. Introduction

Learning a second language does not only imply learning grammar and vocabulary but also accurate speech production. The latter is necessary to become an accepted member of a linguistic community. In our research, we are interested in German L2 speech production by French native speakers living in France. German and French do not share the same phonemic inventory. On one hand, there are no nasal vowels in standard German except for French loanwords like *Restaurant* or *Ragout fin*. On the other hand, the consonants /h/, /ŋ/ and /ç/ (*Ich*-Laut) are not phonological in standard French.

Learners of the language from the other have some difficulties producing the speech sounds absent in the phonemic inventory of their L1. With an adapted training, learners are able to learn to produce those speech sounds [1].

In order to provide learners with objective and individual feedback, we aim to use acoustic measurements to characterize the segments produced in German L2 speech. One of our long term aims is to provide teachers with the possibility to rate oral productions of L2 German learners and to quantify errors for each learner individually. We believe that acoustic measures can help teachers to provide quantitative and qualitative pronunciation feedback. In this study, we investigate the acoustic

properties of the German fricatives [ʃ] and [ç] produced by German native speakers (GG) and French learners of German (FG). French learners of German tend to replace the palatal fricative /ç/ by the post alveolar fricative /ʃ/ or to overgeneralize [ç] productions once the phoneme has been well learned.

1.1. The German frcatives /ʃ/ et /ç/

There are several fricatives in German that are not present in the French consonantal inventory. Other than /h/, which is a well-know difficulty among French native speakers [2, 3], there are two voiceless fricatives: one palatal /ç/ (*Ich*-Laut) and one velar /x/ (*Ach*-Laut) that are considered as allophones because they appear in complementary positions in the German lexicon. Anterior vowels (as well as consonants) are followed by [ç], posterior vowels and the vowel /a/ are followed by [x] [4]. Both fricatives appear generally at the end of the word (*Buch* [buːx], book) or at the end of a morpheme (*riech-en* [ʁiːç-ən], smell). In loanwords however, [ç] can also appear word initial (*Chemie* [çemiː], chemistry), *China* [çina], China etc.)

In this paper, we study the voiceless palatal fricative which is often replaced by FG speakers by its closest neighbour the post alveolar fricative /ʃ/. Only few minimal pairs for those two fricatives exist in the German lexicon: (*fischte* [ˈfɪʃtə], to fish (participle) - *Fichte* [ˈfɪçtə], spruce; *misch* [ˈmɪʃ], to mix (vocative) - *mich* [ˈmɪç], me). Nevertheless, German native speakers do not always strictly distinguish between the two fricatives in their speech productions (cf.[5]).

The post alveolar fricative /ʃ/ can appear at the beginning or at the end of a German syllable, e.g. *schnell* [ˈʃnɛl] (fast) and *Fisch* [ˈfɪʃ] (fish) and in the suffix *-isch*. At the end of monosyllabic words, the voiceless palatal fricative /ç/ often appears in a cluster with the plosive /t/: *Licht* [lɪçt] (light), *echt* [ɛçt] (real). Regarding the derivational morphology of German, the voiceless palatal fricative appears in the suffixes *-chen* and *-(l)ich* as well. In this study, we will concentrate on the realization of the suffixes *-isch* [ɪʃ] and *-(l)ich* [(l)ɪç] by German L1 and L2 speakers.

In the following, we will first present our speech material, the acoustic analyses that were carried out and finally the Weka classification carried out on all the parameters.

2. Material

2.1. Speech resource

We performed our research on the *French Learners Audio corpus of German speech* (FLACGS) recorded in 2014/15 in the *Laboratoire de Phonétique et de Phonologie, Paris 3* (LPP) in Paris (France) [6]. This corpus includes both German native

speech and German L2 speech of French native speakers. In this corpus, 40 participants (20 L1 German, 20 L1 French) were recorded. The French native speakers ranked their knowledge of German from A2 up to C2 according to the *Common European Framework of Reference for Languages* (CEFR). The participants performed three tasks of increasing speech complexity: repetition, reading and picture description. In the present study we will only concentrate on the repetition task.

The repetition task counts 55 distinct words in central position in one of the following carrier sentences: *Er sagt ... klar und deutlich* and *Ich sage ... klar und deutlich*. The material were recorded by a German female native speaker. The participants listened through headphones to all the spoken utterances in a randomized order and repeated them.

We decided to focus on the repetition task in this study because our target words appear in the same syntactic environment. Prosodic changes due to syntactic placements do not affect them.

2.2. Material choice

We extracted words with the suffixes -/ɪʃ/ or -/ɪç/ like *solidarisch* [zoliˈdaːrɪʃ] (showing solidarity) or *freundlich* [ˈfrɔ͜ʏntlɪç] (friendly). The word itself appears in a stressed syntactic position whereas the targeted suffix is unstressed.

A German native speaker and trained phonetician perceptively judged the 280 tokens on their realization. This intermediate step was taken in order to ensure the data set and to provide classes that include realizations that are as canonical as possible.

GG canonical	women		men	
	/ʃ/	/ç/	/ʃ/	/ç/
[ʃ]	100.0 %	0.0 %	96.7 %	0.0 %
[ç]	0.0 %	87.5 %	0.0 %	87.5 %
Ambiguous	0.0 %	12.5 %	3.3 %	12.5 %

FG canonical	women		men	
	/ʃ/	/ç/	/ʃ/	/ç/
[ʃ]	73.3 %	40.0 %	53.3 %	22.5 %
[ç]	16.7 %	45.0 %	33.3 %	72.5 %
Ambiguous	10.0 %	15.0 %	13.4 %	5.0 %

Table 1: Perception results of /ɪʃ/ and /ɪç/ in GG and FG

Table 1 summarizes the perception results. The table shows three perceptive categories: [ʃ], [ç], and a third category that is not clearly identified as [ʃ] or [ç]. In both groups, the acoustic properties of the third category are more like the properties of the [ç] than of the [ʃ]. That is why in the following analysis, the third group will be annotated as [ç]. In the FG group, we observe confusions between /ʃ/ and /ç/ that are absent in the GG group regarding the repetition task. Substitutions of /ç/ with [ʃ] have been reported by other research teams that are interested in German as an L2 for French learners [7]. But that /ʃ/ can also be replaced by [ç] is not mentioned by Jouvet *et al.* (2015). Speakers who did the latter, against all expectations, did not often substitute /ç/ with [ʃ]. This behaviour could indicate that the replacement of /ʃ/ by [ç] is rather linked to hypercorrection than to articulatory difficulties.

3. Acoustic Analyses

The acoustic measures aim to define which one of the two fricatives FG speakers (and GG speakers alike) produce in the different words. Acoustic measures were carried out with Praat [8]. Statistical analyses have been carried out with R [9], privileging two-way ANOVAs with unequal sample sizes as statistical tests.

The following measures have been carried out to distinguish between [ʃ] and [ç]:

- F2 transition
- Center of gravity
- Intensity in low and high frequency bands of the fricative

3.1. Formant analyses

The first three formants of the suffix /ɪʃ/ and /ɪç/ are plotted in Figure 1 for both the GG and the FG group. The figure shows that in GG speakers, there is a clear difference for the F2 transition between [ɪʃ] and [ɪç]. In the center of the vowel, the group mean of the F2 for [ɪʃ] is 1622 Hz (sd=276), for [ɪç], 2149.861 Hz (sd=242). This difference is significant ($p \leq 0.001$). At the vowel/fricative transition point, the group mean of the F2 for [ɪʃ] is 2236 Hz (sd=443) and 2199 Hz (sd=232) for [ɪç], which is not significantly different ($p \geq 0.05$). No normalization of the formant values was undertaken. In the FG group, the mean at the center of the vowel is 2112 Hz (sd=420) for [ɪʃ] and 2155 Hz (sd=284) for [ɪç]. At the vowel/fricative transition point, the F2 mean for [ɪʃ] is 2185 Hz (sd=376) and for [ɪç] the F2 mean is 2242 Hz (sd=270). Neither point shows any significant difference in the FG group's F2 productions ($p \geq 0.05$).
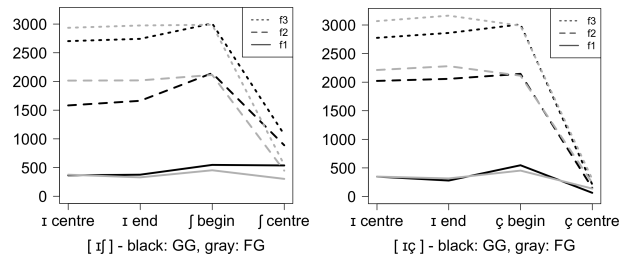

Figure 1: Transition of the first three formants for [ɪʃ] and [ɪç] in GG (black) and FG (gray), **abscissa:** syllable point, **ordinate:** frequency in Hz

GG speakers show a fairly large variability of the F2 which is conditioned by the following fricative. The GG group's F2 transition indicates that both fricatives have different articulatory places. In the FG group however, the less variable F2 transition suggests that the articulatory places for [ʃ] and [ç] are globally less well separated.

In studies carried out on L1 speech, F2 transitions were found useful to distinguish between fricatives. It was found that F2 transition works as a valid predictor in CV contexts. In VC contexts however, F2 transitions are less solid [10, 11, 12].

As proposed by Żygis and Padgett (2010), we calculated the F2 slope by using the following equation:

$$slope\ F2 = \frac{F2_{VC\ boundary} - F2_{V\ midpoint}}{duration\ between\ these\ two\ points} \quad (1)$$

Results of the formant transition are plotted in Figure 2. In the GG group, the slope difference between the syllables -/ɪʃ/

and -/ɪç/ is significant. In the FG group, we find no significant difference. Overall, the F2 slopes in both the FG and GG groups are similar for -/ɪç/. As we saw in Figure 1, F2 in the FG group for both syllables are as high as in the GG group for -/ɪç/ and vary little while transitioning from the vowel to the fricative.

The F2 transition appears to be a solid cue to distinguish between [ʃ] and [ç] in German native speakers but not for German L2 speakers who have French L1.

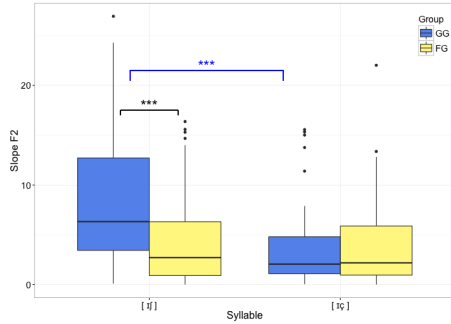Figure 2: Comparison of the F2 transition in GG and FG
* $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$

3.2. Center of gravity

The center of gravity (CoG) is the measure of the frequency mean of an frequency interval which is weighted to its amplitude. Several research teams presented this measure to distinguish between fricatives in native speech [11, 12, 13, 14].

In order to know how the center of gravity progresses over the whole fricative, we extracted the center of gravity in the beginning, the center and the end of the fricative in our target words using Praat.

Results of two way ANOVAs with unequal sample sizes for GG groups's [ʃ] and [ç] show first no significant difference for the beginning of the fricatives, then significant differences for the other CoG values that were extracted (mean CoG: $p \leq 0.05$, CoG in the center and the end of the fricatives: $p \leq 0.01$). In the FG group, statistical analysis show significant CoG values all over the fricative (mean CoG, CoG of the center of the fricatives: $p \leq 0.01$, Cog of the beginning and the end of the fricatives: $p \leq 0.001$). The significant differences in both groups are rather due to the frequency span the CoG occupies than to different frequency positions. For instance in German native speakers, [ʃ] occupies a frequency span between 2898 and 4038 Hz (regarding the values of the second and third quartile) whereas [ç] occupies a frequency span which is located between 2379 and 5283 Hz.

The CoG values of [ʃ] between both speaker groups, GG and FG, do not show any significant difference for the CoG that was extracted of the fricative. The same comparison carried out on [ç] show significant differences in CoG in all three places of the fricative and the mean CoG ($p \leq 0.001$). These results suggest on one hand that FG speakers produce two different fricatives. On the other hand, FG speaker's [ç] is not the same voiceless palatal fricative we can find in GG speakers. For [ç], both groups show a wide frequency span in their CoG but in the GG group the CoG is situated in higher frequencies than in FG. In the FG group, maximal values of the CoG for [ç] are equal to those found in their [ʃ] whereas the GG group tend to produce [ç] with CoGs situated in higher frequencies than in their [ʃ].

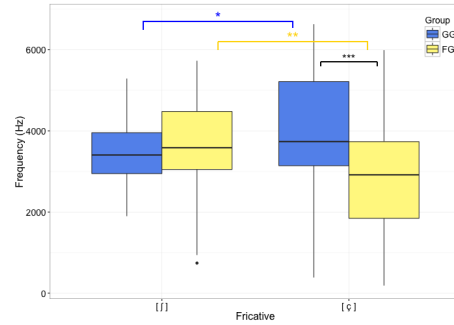Figure 3 illustrates the differences of the mean CoG for the fricatives [ʃ] and [ç] in both speaker groups.

Figure 3: mean CoG for the fricatives [ʃ] and [ç]
* $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$

3.3. Intensity in low and high frequencies

We extracted the intensity of low (1kHz-3kHz, 1kHz-4kHz) and high (3kHz-6kHz, 4kHz-7kHz) frequency bands because both fricatives, /ʃ/ and /ç/ spread their energy in an unequal way over the spectrogram. For /ʃ/, low frequency bands are more charged with intensity than high frequencies. The contrary is true for /ç/. That is why we want to use the intensity of low and high frequency bands as a local measure to distinguish between [ʃ] and [ç] in German native and non-native speech.

Concerning the GG group, results show significant intensity differences in both low frequency bands as well as in the high frequency band 3kHz-6kHz ($p \leq 0.001$) whereas there is no significant difference of the intensity values on the second high frequency band (4kHz-7kHz). In the FG group, in both high and low frequency bands, we find significant differences ($p \leq 0.001$). Regarding group differences, GG speakers and FG speakers do not show any significant differences for the intensity of high frequencies in [ʃ] and the low frequencies in [ç]. However, for low frequency bands in [ʃ] (1kHz-3kHz, 1kHz-4kHz: $p \leq 0.001$) and high frequency bands in [ç] (3kHz-6kHz: $p \leq 0.05$, 4kHz-7kHz: $p \leq 0.01$), the intensity values in GG and FG speakers are significant different.

To summarize, the intensity of frequency bands is more distinctive for the fricative production in FG speakers than in GG speakers. Nevertheless, this local measure seems to have the potential of becoming a valid cue to distinguish between [ʃ] and [ç] in both GG speakers and FG speakers.

In the following, we are going to present our classification results for both speaker groups on [ʃ] and [ç] using Weka.

4. Results and Discussion

In order to distinguish between [ʃ] and [ç], the measures presented in Section 3 have been analyzed with the Weka implementation in R [15]. Figure 4 presents the results of this analysis for female speakers of both speaker groups (GG and FG). Separating the data set according to speaker group (GG or FG) and sex helped to increase the accuracy of the decision tree.

Weka was developped by the Waikato University (New Zealand) and is often used as a classification tool which also allows machine learning [16]. Weka in R creates a J48 tree. This decision tree uses a labelled data set and performs a relevance analysis of the features in order to decide which ones help best predict the labelled class (highest normalized information gain)
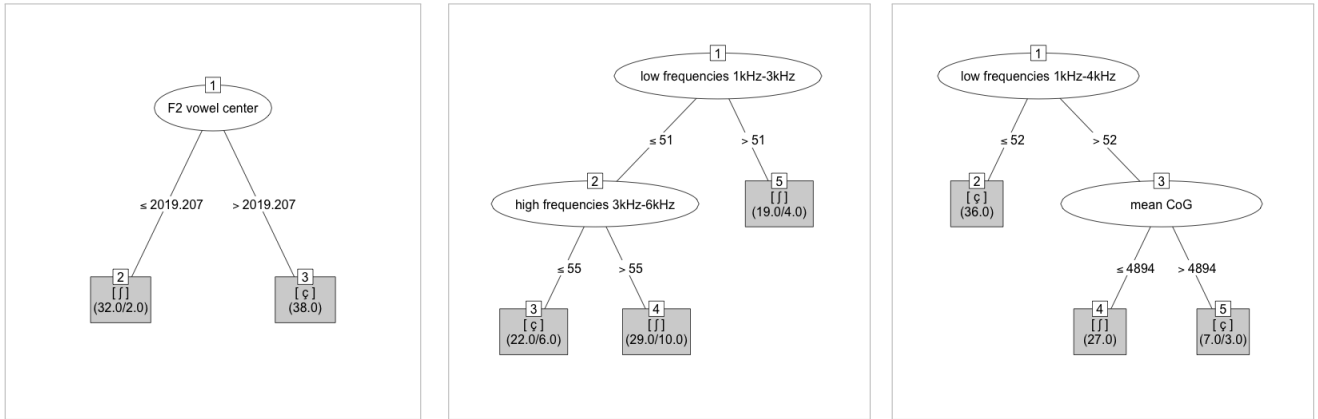
Figure 4: Classification of [ʃ] and [ç] with Weka, **left:** GG female (+ contextual measures), **center:** FG female (+ contextual measures), **right:** GG female (- contextual measures)

[17]. In our case, the labelled class was the perception result of /ʃ/ and /ç/. The data set contained contextual measures (F2 of the preceding vowel, vowel duration, frequency band intensities of the vowel, and F2 transition) and local measures (duration of the fricative, CoG values, frequency band intensities, and ratios of high and low frequency band intensities).

For the creation of the decision trees, a threshold was fixed: at least 10% of all tokens are sharing the same property. Fixing this threshold presents several advantages:

1. The decision for each branch is based on productions of several speakers. Therefor, the selected property can be considered as a repetitive pattern.

2. Error rates are still relatively low.

3. Tree complexity is reduced.

For German native speakers, the contextual measures as cues for the classification of [ʃ] and [ç] reduce the error rate. In the left tree of Figure 4, we see that in female GG speakers the second formant value extracted of the center of the preceding vowel is solid enough to separate [ʃ] and [ç] with a low error rate (2.9% of incorrectly classified instances). All of the incorrectly classified instances were labelled as [ʃ]. A similar result was obtained for male GG speakers.

The tree in the center of Figure 4 shows female FG speakers. In the data set which was used for the classification, contextual measures were also present. Only the intensity of low and high frequencies was used to build up the decision tree. In male FG speakers, low frequencies and CoG are the strongest cues to separate [ʃ] from [ç]. To decide whether a FG speaker produces [ʃ] or [ç], contextual measures are not of great use which is due to their vowel quality (compare Section 3.1). Local measures are more solid cues to distinguish between [ʃ] and [ç] in FG speakers. Nevertheless, in our example, error rates for female FG speakers are high (28.6% incorrectly classified instances). These values could point out that the production of the unfamiliar fricative [ç] and its separation from [ʃ] show high variability. This variability does not only depend on the background of the speaker (learning process, regional phonetic influences) [18, 19]. L2 speech within the same speaker can show high variability as well [20].

As contextual measures are less relevant for FG speakers with a threshold fixed at 10% of the tokens, we created a Weka decision tree for female GG speakers leaving contextual measures out of the data set (see the right tree in Figure 4). Com-

pared to the first tree (on the left), the new one has increased in complexity by one level. Without the contextual measures, the error rate slightly increases (4.2% of incorrectly classified instances). All incorrectly classified instances are labelled as [ç] productions. The overall error rate increased from 2.9% to 4.2% involving a different fricative. Leaving out the contextual measures in GG speakers still leads to satisfying classification results. The acoustic differences between [ʃ] and [ç] are strong in GG speakers. Combined measures of the low frequency band intensity and the CoG help to separate [ʃ] from [ç]. In male GG speakers, only the intensity of different low frequency bands was used to build up the tree which leads to a lower error rate than using contextual measures in this speaker subset.

5. Discussion and Perspectives

We carried out different acoustic measures for [ʃ] and [ç] in a VC context: formant analysis of the preceding vowel, CoG of the fricative and intensity in low and high frequency bands. Our results show that in order to distinguish between [ʃ] and [ç], contextual measures are only solid cues in German native speech. Vowel quality in German L2 speech of French L1 speakers does not allow a solid distinction in a VC context. Leaving out contextual measures in German native speech does not lead to an increased error rate regarding classification.

Using only local measures extracted on the fricative may present another advantage: acoustic measures and decision thresholds do not have to be revised with respect to the preceding vowel or position of the fricative (VC context *versus* CV context). In further studies, we will investigate this hypothesis to distinguish [ʃ] and [ç] in native and non-native productions in reading and picture description. The analysis presented in this study are going to help rate speech productions of adult learners of German in a more accurate way in order to identify both accurate and erroneous pronunciation and to quantify L2 pronunciation progress over time.

6. Acknowledgements

7. References

[1] J. E. Flege, N. Takagi, and V. Mann, "Japanese adults can learn to produce English /ɹ/ and /l/ accurately," *Language and Speech*, vol. 38, no. 1, pp. 25–55, 1995.

[2] F. Zimmerer and J. Trouvain, "Productions of /h/ in German: French vs. German speakers," in *Sixteenth Annual Conference of the International Speech Communication Association*, 2015.

[3] J. Wottawa, M. Adda-Decker, and F. Isel, "Segmental difficulties in French learners of German," in *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2015*, 2015.

[4] K. Kohler, "German," *Journal of the International Phonetic Association*, vol. 20, no. 01, pp. 48–50, 1990.

[5] S. Jannedy, M. Weirich, and L. Helmeke, "Acoustic analyses of differences in [ç] and [ʃ] productions in Hood German," in *Proceedings of the 18th International Congress of Phonetic Sciences*, 2015.

[6] J. Wottawa and M. Adda-Decker, "French learners audio corpus of German speech (FLACGS)," in *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC'16)*, Portorož (Slovenia), May 2016.

[7] D. Jouvet, A. Bonneau, J. Trouvain, F. Zimmerer, Y. Laprie, and B. Möbius, "Analysis of phone confusion matrices in a manually annotated French-German learner corpus," in *Workshop on Speech and Language Technology in Education*, 2015.

[8] P. Boersma and D. Weenink, *Praat: doing phonetics by computer [Computer program]*, 2016, version 6.0.15, retrieved 23 March 2016. [Online]. Available: http://www.praat.org/

[9] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2008, ISBN 3-900051-07-0. [Online]. Available: http://www.R-project.org

[10] P. C. Delattre, A. Berman, and F. S. Cooper, "Formant transitions and loci as acoustic correlates of place of articulation in american fricatives," *Studia Linguistica*, vol. 16, no. 1-2, pp. 104–122, 1962.

[11] M. Żygis and J. Padgett, "A perceptual study of Polish fricatives, and its implications for historical sound change," *Journal of Phonetics*, vol. 38, no. 2, pp. 207–226, 2010.

[12] A. Benamrane, "Acoustic study of fricatives in standard Arabic (Algerian speakers)," Theses, Université de Strasbourg, Dec. 2013. [Online]. Available: https://tel.archives-ouvertes.fr/tel-01089132

[13] R. L. Kemp, "The perception of German dorsal fricatives by native speakers of English," Ph.D. dissertation, Master thesis, University of Georgia. Cerca con Google, 2011.

[14] S. Schmid and S. Negrinelli, "Palatal obstruents in two Rhaeto-romance varieties: acoustic analysis of a sound change in progress," in *18th International Congress of Phonetic Sciences*, 2015.

[15] K. Hornik, C. Buchta, and A. Zeileis, "Open-source machine learning: R meets Weka," *Computational Statistics*, vol. 24, no. 2, pp. 225–232, 2009.

[16] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The weka data mining software: an update," *ACM SIGKDD explorations newsletter*, vol. 11, no. 1, pp. 10–18, 2009.

[17] T. C. Sharma and M. Jain, "WEKA approach for comparative study of classification algorithm," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 2, no. 4, pp. 1925–1931, 2013.

[18] N. Kartushina, A. Hervais-Adelman, U. H. Frauenfelder, and N. Golestani, "The effect of phonetic production training with visual feedback on the perception and production of foreign speech sounds)," *The Journal of the Acoustical Society of America*, vol. 138, no. 2, pp. 817–832, 2015.

[19] E. Simon, M. Debaene, and M. Van Herreweghe, "The effect of L1 regional variation on the perception and production of standard L1 and L2 vowels," *Folia Linguistica*, vol. 49, no. 2, pp. 521–553, 2015.

[20] T. K. Perrachione, J. Lee, L. Y. Ha, and P. C. Wong, "Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design," *The Journal of the Acoustical Society of America*, vol. 130, no. 1, pp. 461–472, 2011.