German non-native realizations of French voiced fricatives in final position of a group of words.

Anne Bonneau [1], *Martine Cadot* [2]

[1] LORIA/CNRS Vandoeuvre-lès-Nancy, France
[2] LORIA/UdL Vandoeuvre-lès-Nancy, France

anne.bonneau@loria.fr, martine.cadot@loria.fr

Abstract

We analyzed the realizations of French voiced fricatives /z,ʒ/ by German non-native and French native speakers, in final position of an accentual group, a position where German fricatives are devoiced. Three speaker levels (from beginners to advanced speakers) and different boundary types (depending on whether the fricative is followed by a pause, a schwa, or is directly followed by the first phoneme of the subsequent group), are considered. A set of cues, among which periodicity, fricative duration, and intensity in low frequencies, is used for voicing analysis. Results show that German realizations vary significantly with language, speakers' level and boundary type, and argue in favor of an influence of L1 (German) final devoicing on non-native realizations. We discuss these effects as well as the influence of orthography (the presence of a schwa at the end of words).

Index Terms: Phonetic transfer, fricatives, voicing neutralization, prosodic boundaries, German/French.

1. Introduction

We investigate a typical example of L1-L2 interference concerning the realization of the French voiced fricatives /z,ʒ/ in final position by German learners of French. In German, the opposition between voiced and unvoiced obstruents (fricatives and stops) is neutralized in final position in favour of the realization of unvoiced consonants [1], whereas in French the voicing feature is distinctive in final position. This difference between both systems is known to be a source of error for German speakers, who tend to produce unvoiced obstruents in final position when speaking French instead of the expected voiced consonants.

The voicing feature, which makes a phonological contrast between obstruents sharing identical features but the voicing one (i.e. /f-v/, /s-z/, /ʃ-ʒ/, for fricatives) does not rely only upon the presence or absence of voicing –the articulatory phenomenon corresponding to vocal fold vibration and generating periodicity in the acoustic signals. Other articulatory correlates, such as articulatory strength (associated to the fortis/lenis distinction, and very important in German stop production), play also a role in the categorization of both kind of consonants [2]. From an acoustical point of view, many cues are associated with the voicing contrast [3] [4], such as, for fricatives, the periodicity of the signal during the consonant production, the duration of the consonant, the duration of the preceding vowel, as well as noise intensity [5] .

A number of factors influences the acoustical realization of the voicing contrast. Kuzla *et al* [6] showed prosodic effects on the duration and amount of glottal vibration in German word-initial fricatives /f,v,z/ in assimilatory and non-assimilatory devoicing contexts. Moreover, in languages with voicing neutralization in final position, neutralization appears to be often incomplete [7] and the extent of this phenomenon was found to be dependent on sentence position, phonetic environment, orthography and speaking styles, among other factors (see [8] for a review).

The aim of this study is to examine interferences between L1 and L2 in the case of French voiced fricatives in final positions pronounced by German non-native speakers. The fricatives appeared in final position of an accentual group but not at sentence final position. We pay a special attention to the way speakers realized the prosodic boundary at the vicinity of the fricative and discussed the effects of the boundary type and the speaker level on periodicity and consonant duration. As native and non-native speakers recorded the sentences, German realizations have been systematically compared to French ones. Preliminary results concerning the periodicity of the signal during fricatives by German non-native speakers have been reported in [9], and, for the consonant /ʒ/ in sentence final position, in [10].

2. Experimental Protocol

2.1. Corpus

We extracted a set of French sentences containing fricatives in final position of a group of words (an accentual group, in French) from a bilingual corpus which was recorded by French learners of German and German learners of French in their native and second languages [11]. This corpus has been built under the framework of the IFACSL ANR project (www.ifcasl.org).

The subjects were seated in a quiet room and read the sentences from the screen of a Windows laptop, with a headset microphone (AKG C520) connected to an Audiobox (M Audio Fast track). Fourty German learners of French and fourty French speakers recorded the corpus. The non-native speakers were classified by German teachers of French in three categories: beginners (15 speakers), intermediate (13 speakers) and advanced (12 speakers), corresponding to A, B and C levels, respectively.

We analyzed the fricatives /z/ and /ʒ/ embedded in sentences. There was one sentence per fricative, and each

sentence was uttered once by each speaker, which gave a total amount of 160 sentences. Some data have been removed due to spelling mistakes and the acoustical analysis was based upon 154 sentences. The sentences were made up of three noun phrases and a verb phrase in that order: NP1 VP NP2 NP3, where NP1 was a subject, NP2 a direct object and NP3 an adverbial phrase beginning with a preposition. The fricative was at the end of the last word of NP2, was preceded by the vowel /a/ and the preposition (i.e. the word following the fricative) began with the vowel /a/. The first sentence was "Mon ami a perdu ses bag**ages à** la gare" (My friend lost his luggage at the railway station), and the second one "Les élèves doivent cocher la bonne c**ase a**vec un feutre" (Pupils have to tick the right case with a felt pen). The broad transcription of the standard dialect form for the context surrounding the fricative /ʒ/ is either /aʒa/, /aʒ#a/, or, although less frequent, /aʒə#a/ (same remark for /z/).

2.2. Segmentation and labelling at AG boundaries

We segmented sequences at AG (accentual group) boundaries manually and very carefully with Praat. The labelling of acoustic events was made using precise acoustic criteria; we present below those retained for schwas.

At NP2-NP3 boundaries (the fricatives analysed here are at the end of NP2), we observed the following possible events: a schwa [ə], a silence [#], a glottal stop [ʔ], aspiration [h], as well as the presence of very weak voiced segments at the end of the fricative, noted [hh]. The presence of creaky voice during the vowels was indicated by a diacritic. The distinction between schwas and weak periodic events has been made on the following basis: the duration of schwas must last at least 30 ms, their mean intensity must be no more than 10 dB under that of the vowel preceding the fricative, and they must have a visible formant structure.

We observed a special sequence, produced by non-native speakers but for two cases: a schwa, in general relatively long and intense, directly followed by the vowel /a/. In almost all cases, at least one of both vowels was creaky.

2.3. Boundary types

From observations made during segmentation of the fricative, we considered four types of boundary:

1) the fricative is directly followed by the following vowel, as in /aʒa/,

2) the fricative is followed by a schwa, and a pause, as in /aʒə#a/, the pause being either a silence, either an aspiration, or a glottal stop,

3) the fricative is followed by a pause (silence, glottal stop, or aspiration), as in /aʒ#a/,

4) the fricative is followed by a schwa and then by the vowel /a/, as in /aʒəa/.

The sentences have been split into four groups as a function of their boundary type : G1, G2, G3 and G4; the number assigned to each group matching the rank of the boundary type in the above list.

2.4. Acoustic cues

The periodicity has been estimated by Praat "Voicing Report" function which provides the fraction of locally unvoiced frames, from which we deduced the fraction of *voiced* frames (the percentages of voiced frames during the fricative duration). The estimation of periodicity during fricatives, sounds with often intense noises, being especially difficult [12], we checked the results manually, and corrected them when necessary (there was very few corrections, made for obvious mistakes). The duration of the fricative was provided by segmentation.

3. Results and discussion

We analyse the periodicity of the speech signal and consonant duration for the fricatives /z,ʒ/ as a function of the language, L1 and L2, and, for non-native speakers, the speakers' level, and examine possible effects of boundary types on these factors. Results for /z/ and /ʒ/ have been combined for statistical purposes. Statistical analyses have been made with R.

	G1	G2	G3	G4
A	6	1	16	5
B	7	1	12	5
C	13	1	7	3

	G1	G2	G3	G4
G (levels A, B, C)	26	3	35	13
F	55	1	19	2

Table 1 Number of realizations of G1, G2, G3 and G4 boundaries, as a function of German non-native speakers' levels (A, B, C), and speakers' first language (G for German and F for French)

3.1. Boundary types.

Table 1 presents the frequency of each boundary types, (defined in section 2.3), as a function of languages and, for L2, speaker levels. We used the Pearson's chi-squared test analysis to compare the results. The differences between French and German speakers with respect to boundary types is highly statistically significant ($p < 2\text{x}10^{-5}$).

Almost all realizations of French speakers, and most realizations of German speakers fall into G1 and G3 groups. For most French speakers, the fricative is directly followed by the first vowel of the last accentual group (55 realizations in G1 *vs* 19 in G3) whereas the reverse is true for most German speakers (26 realizations in G1 *vs* 35 in G3). This difference is highly statistically significant ($p < 1.8\text{x}10^{-4}$). Among German speakers, we observe that advanced speakers (C) used more G1 boundary types than G3 ones, as French speakers do, although to a lesser extent, whereas the reverse is true for beginners (A). This difference between A and C speakers is significant ($p < 0.02$). Fluency, due to a better mastering of the language, probably explains why French speakers produced less G3 boundary types (i.e. less pauses) between NP2 and NP3 accentual groups than German non-native speakers. This explanation holds also for the difference observed between advanced speakers and beginners.

The low number of items in the second group (G2), for French and German speakers, shows that before a pause, there are very few sufficiently clear evidences of a schwa, whatever the speaker nationality. However, results for G4

shows that German speakers pronounced a (relatively long and intense, in general) schwa directly followed by the first vowel of the next accentual group in 17% of the cases (*vs* only 3% for French speakers).

Before analyzing acoustic cues, let us comment results for G4. In French, fricatives in (oral) final position are followed by a "e" in writing, as in the word "neige" ("snow"). This schwa is either not pronounced (/nɛʒ/) or very weak (/nɛʒə/) in standard French. In similar contexts (fricatives followed by a "e" in writing at the end of a word), the schwa is pronounced in German and tend to be more intense than the French schwa, although always in unstressed position, and this difference in both languages might be the source of mistakes. If German speakers, influenced by German spelling rules, pronounced the word containing the fricative the German way, as it might be the case in G4 when the schwa is intense, then this fricative is not in final position for these speakers. Interestingly, we observed that the voicing distribution for consonants in this group is particular: all percentages are above 40%, and 62% of the data are fully voiced (against 15% on the averaged for fricatives in other groups).

Since we have pointed out that fricatives in G4 might not be in devoicing position for German speakers, and since G2 contains very few items, we will analyze and comment results with and without G4 and G2.

3.2. Periodicity

The fractions of locally voiced frames observed during fricative productions have been split into 10 or 5 intervals in figures 1 and 2. For the statistical analysis used for periodicity analysis (Chi2), it was necessary to reduce the number of intervals to have a sufficient number of data in each case, and we considered two intervals [0,50] and]50, 100].

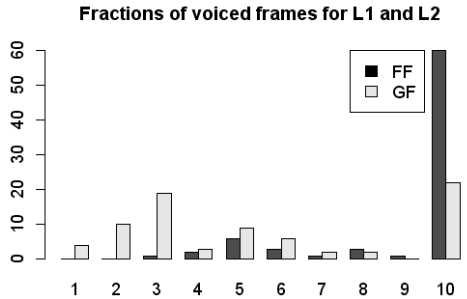

Figure 1 Frequencies of fractions of voiced frames for French speakers (L1, in black) and German non-native speakers (L2, in grey). On the x-axis, "1"stands for the interval [0-10%], "2" for]10-20%] ...

There are strong and very significant differences between French and German realizations of French voiced fricatives in word final position. As shown by Figure 1, which represents the frequency of voiced fractions for the fricatives /z,ʒ/, German realizations are on average less "voiced" than French ones: there are more slightly voiced and less fully voiced realizations for German than for French speakers ($p < 2x10^{-9}$). Such results demonstrate very clearly the influence of L1 (German) on L2 (French) realizations. If we remove G4, so as to keep only fricatives which are without doubt in word final position for German speakers, the distribution is very similar to the one represented in Figure 1, but with a reduced amount of fully voiced realizations for German speakers, which increases the level of significance ($p < 5x10^{-11}$).

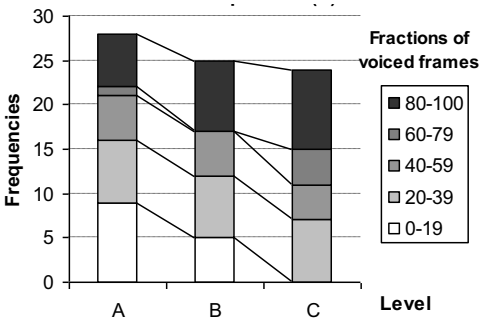

Figure 2. Frequencies of fractions (in percentages) of voiced frames for beginners (A), intermediate (B) and advanced German speakers (C).

The percentages of voicing during fricatives vary also in a statistically significant manner as a function of speakers' level (the *p_value* is equal to 0.037, when all levels are taken into account, and 0.014, when only advanced speakers and beginners are considered). The differences between realizations due to speakers' levels are slighly more significant when data from the fourth group are removed. Thus, as might be expected, more advanced speakers realized more "voiced" realizations than beginners (see Figure 2).

The question arises as to whether boundary types have an effect on voicing. For French speakers (Figure 3, Left), we have to split the data into fully voiced (percentage of voiced frames equal to 100) and non fully voiced consonants, since there is few non fully voiced consonants. The ratio of fully to non-fully voiced fricatives is above 10 for consonants in G1, and less than 1 for consonants in G3, which tends to show that consonants, even at word boundary, are more voiced between two vowels, than before a pause. However, the number of data for non-fully voiced consonants in G3 just falls short to allow us to use the result from Chi2 test ($p < 0.001$) without violating prerequisites for this analysis. We are presently working on another set of data for French speakers, so as to obtain an exploitable result. For German speakers (Figure 3, Right), with data split into [0,50[and [50,100] intervals, results are significant ($p < 0.014$, three levels taken into account), which might be due to different causes: first, the position of the fricatives (at the end of a word but nevertheless between two vowels in G1 *vs* before a pause in G3); secondly, the higher number of advanced speakers in G1.

If we consider results as a function of the speakers' levels for G1 and G3 separately, we found statistically significant differences among advanced speakers and beginners in G1, but not in G3. Indeed, based upon the observation of voicing frame distributions, we split the data into [0-25] and]25-100] intervals, and found that the number of "slightly" voiced consonants with respect to "more voiced" consonants was less important for advanced speakers than for

beginners ($p < 3x10^{-3}$). We will comment this result in the following paragraph.

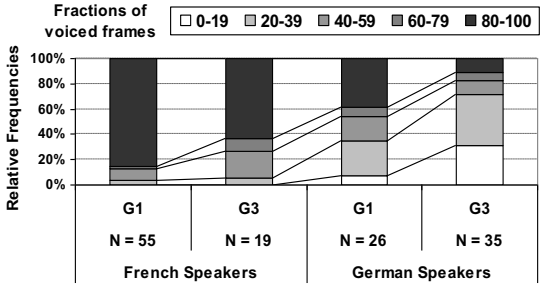

Figure 3. Fractions of voiced frames for French speakers and German speakers as a function of G1 and G3 boundary types.

3.3. Consonant duration

Table 2 provides the median and the mean for the duration of the consonants, as a function of languages (L1 and L2), speakers' level (A, B and C, for L2) and boundary types (G1 and G3). Figure 4 shows the boxplots for consonant duration as a function of the same factors. We used pairwise comparisons using Wilcoxon rank sum test for the statistical analysis, with Benjamini, Hochberg [13] correction for multiple comparisons (Table 3). On the left of table 3, we can see that for a same group (G1 or G3), consonants pronounced by native speakers are always significantly shorter than those pronounced by non-native ones (see French G1 *vs* German G1 and French G3 *vs* German G3). This result can be attributed either to categorization problems, due to L1L2 interference -voiceless fricatives tend to be longer than voiced ones-, or to lack of fluency. We believe both factors play a role here. For a given language, French or German, duration in G1 is significantly shorter than duration in G3 (see French G1 *vs* French G3 and German G1 *vs* German G3), which can be explained by the fact that the last syllable of an accentual group is longer before a pause. There is no statistical difference between duration for French G3 and German G1, which underlines the importance of the type of boundary on native and non-native realizations.

		Medians				**Means**			
		G1	G3	G4	Tot.	G1	G3	G4	Tot.
GERMAN	A	131	134	106	125	138	104	99.2	119
	B	83	135	71	86	82.7	85	70.8	78.3
	C	78	87	70	79	77.1	109	75	78.6
	Tot.	81	126	77	92	92.7	99.3	82.7	90
French		63	80	53	66	64.9	81	53	64.8
Tot. G+F		68	94	73	78	73.8	94.8	78.7	75.4

Table 2 Summaries for consonant duration (expressed in ms).

Let's now talk about speakers' level and groups. On the right of table 3, results show, not surprizingly, that there is a very significant difference between realizations of advanced speakers in G1 and realizations of beginners in G3. More interestingly, advanced speakers realized shorter consonants than beginners when there is no pause after the consonant (G1 group). Once again, we believe categorization problems and fluency might explain this difference between advanced speakers and beginners in G1. The same kind of relation has also been observed for periodicity: in G1, advanced speakers realized more voiced fricatives than beginners. Thus, advanced speakers perform better (are closer to French speakers) than beginners in G1, whatever the acoustic cue considered (consonant duration or periodicity). Such a difference, statistically significant, between speakers, is not found in G3. Since there are few and unbalanced data in G3, we prefer to be cautious before claiming that the boundary type G3, contrary to G1, do not induce statistically significant differences between advanced speakers and beginners. More data would be necessary to confirm this hypothesis.

		French		German	
		G1	G3	G1	G3
F	G1	-	-	-	-
	G3	1E-2	-	-	-
G	G1	2E-05	0.512	-	-
	G3	2E-12	1E-04	4E-04	-

		Level A		Level C	
		G1	G3	G1	G3
A	G1	-	-	-	-
	G3	0.747	-	-	-
C	G1	0.047	4E-04	-	-
	G3	0.354	0.09	0.23	-

Table 3 P-values for pairwise comparisons. On the left side, groups G1 and G2 crossed with first languages French and German. On the right side, groups crossed levels for German speakers

We are presently studying the duration of the preceding vowel, which is also, as expected, very interesting with respect to this study.

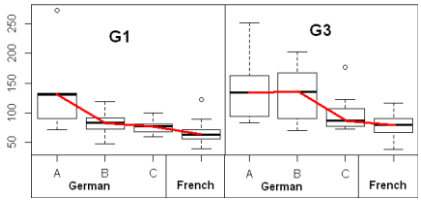

Figure 4 Boxplots of consonant durations (in ms) as a function of language, Levels and Groups. In red, the medians of table 2.

4. Conclusions

We analyzed the realizations of French voiced fricatives /z,ʒ/ embedded in sentences and appearing in final position of an accentual group, as a function of German speakers' levels and boundary types. Results show that German speakers produced consonants which are less voiced –with respect to signal periodicity- and longer than French speakers, which demonstrate the influence of L1 (fricatives in final position are devoiced in German) on L2. Results depend upon speakers' level (advanced speakers performing better than beginners) but also upon boundary types: the presence or absence of a pause after the fricative affect the realizations. We also found an effect of L1 spelling rules on L2 realizations, since German speakers pronounce 17% of fricatives followed by a (relatively intense and long) schwa, and in this context, fricatives tend to be fully voiced.

5. References

[1] R. Wiese, The Phonology of German, Oxford: Clarendon Press, 1996.

[2] K. Kohler, «Phonetic explanations in phonology: The feature fortis/lenis,» *Phonetica,* vol. 31, pp. 150-174, 1984.

[3] M. Jessen, "Phonetic implementation of the distinctive auditory features [voice] and [tense] in Stop consonants," in *Recent developments in distinctive feature theory*, T. Hall and U. Kleinhenz, Eds., Berlin, Mouton de Gruyter, 2001.

[4] L. Lisker and A. Abramson, "A cross-language study of voicing in initial stops," *Word,* pp. 384-422, 1964.

[5] A. Jongman, R. Wayland and S. Wong, "Acoustic characteristics of English fricatives," *J. Acoust. Soc. Am., Vol. 108,* vol. 108, no. 3, 2000.

[6] C. Kuzla, T. Cho and M. Ernestus, "Prosodic strengthening of German fricatives in duration and assimilatory devoicing," *Journal of Phonetics,* vol. 35, pp. 301-320, 2007.

[7] R. Port and P. Crawford, "Incomplete neutralization and pragmatics in German," *Journal of Phonetics,* vol. 17, pp. 257-282, 1989.

[8] F. Kleber, T. John and J. Harrington, "The implications for speech perception of incomplete neutralization of final devoicing in German," *Journal of Phonetics,* vol. 38, pp. 185-196, 2010.

[9] A. Bonneau, "Realizations of French voiced fricatives by German learners," in *ICPhS*, Glasgow, 2015.

[10] C. Fauth and A. Bonneau, "L1-L2 interference: the case of devoicing of French voiced obstruents in final position by German learners - Pilot study," in *Workshop on Multilinguality in Speech Research: Data, Methods and Models*, Dagstuhl – Germany, 2014.

[11] C. Fauth, A. Bonneau, O. Mella, B. Andreeva, V. Colotte, D. Fohr, D. Jouvet, B. Moebius and J. Trouvain, "Designing a bilingual speech corpus for french and German language learners," in *LREC*, Reykjavic, 2014.

[12] S. Fuchs, J. Brunner and A. Busler, "Temporal and spatial aspects concerning the voicing contrast in alveolar and postalveolar fricatives," *Advances in Speech-Language Pathology,* vol. 9, no. 1, pp. 90-100, 2007.

[13] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *Journal of the Royal Statistical Society,* vol. 57, p. 289–300, 1995.