PRIMARY AND SECONDARY CUES TO VOICE ASSIMILATION IN FRENCH AND IN SLOVENIAN

Pierre A. HALLÉ,[a,b,d] Juan SEGUI,[b] and Kaja ANDROJNA[c]

[a] LPP (CNRS-Paris 3), [b] LMC (INSERM-Paris 5), [c] ENS Ulm, [d] Haskins Laboratories
pierre.halle@univ-paris3.fr, juan.segui@parsidescartes.fr, kaja.androjna@normalesup.org

ABSTRACT

Several important issues in voice assimilation have been recently revisited. In particular, is it gradient or categorical? Distributional analyses of the voicing ratio in assimilating vs. neutral contexts conducted on French data have yielded somewhat discrepant results, yet suggesting that assimilation is categorical and optional at the same time. In this study, we compare between-word voice assimilation data reported for French with similar data collected in Slovenian: Is assimilation stronger in one language and how categorical is it in each language? We also compare voice assimilation in the two languages for secondary cues to voicing, such as closure duration, and compare the assimilation produced for words vs. nonwords. We find that voice assimilation is more categorical and occurs more often in Slovenian than French, that some secondary cues to voicing resist assimilation in French but not in Slovenian, and that assimilation is similar for words and nonwords in both languages.

Keywords: voice assimilation; French; Slovenian; voicing-ratio; secondary cues.

1. INTRODUCTION

Many aspects of voice assimilation in French are generally agreed on: it is regressive and restricted to contacts between obstruents. However, there is a continued debate on whether voice assimilation may be phonetically complete or not, that is, categorical (all-or-none) or gradient. This debate dates back at least to Grammont [9] who claimed it is more complete within than between words. Rigault [15] did not find such a difference. Later on, Gow published a series of studies bearing on perceptual compensation for assimilation for *place* assimilation [5-7], with acoustic measurements on assimilated vs. non-assimilated materials. Gow found intermediate degrees of assimilation between absence of and complete assimilation (also see [14]). In the same vein, incomplete *voice* assimilation has been found in several studies (French: [16]; Hungarian: [8]). Yet, the notion of incomplete assimilation requires some qualification. The studies mentioned above all reported voicing-ratio data, a measure of the amount of glottal pulsing within a given phonetic segment. Incomplete assimilation for voicing-ratio means that assimilated segments exhibit intermediate values of voicing-ratio between those of voiced and voiceless non-assimilated segments. As discussed in [9-10], voicing-ratio *distributions* must be compared in assimilating vs. non-assimilating contexts to assess the categorical or gradient nature of assimilation. Using this approach, several studies found that voice assimilation in French is categorical with respect to voicing-ratio: it does not occur all the time but when it occurs, it is complete [1-3, 10-11]. ([12] found categorical assimilation in French only in the voicing direction.) Whereas voicing-ratio may be viewed as a primary cue to voicing, as a measure of glottal pulsing, it is not the sole cue to voicing (see [13]). Other cues (usually called secondary) include, among others: duration of the preceding vowel and of the consonant; for stops, closure and release duration and intensity. Abdelli-Beruh found that categorical changes of voicing-ratio are often concurrent with neutralization of the distinctiveness of durational secondary cues [1-2]. More recently, she found that such neutralization was restricted to high speech rates. At moderate speech rates, distinctiveness of duration cues would be maintained [3]. Likewise, [17] showed that residual durational cues are preserved in items that are fully assimilated in terms of voicing-ratio, allowing listeners to recover intended meaning.

In this study, we compare French and Slovenian for voice assimilation, using voicing-ratio as a primary cue to voicing. We also examine Harmonics-to-Noise ratio (HNR) as an alternative primary cue to voicing. One goal of our study is to revisit the issue of categorical vs. gradient voice assimilation: Is the categoricity or gradiency of voice assimilation universal or language-specific? Another goal is to examine whether at least some secondary cues to voicing are preserved in voice-assimilated segments. We therefore examine segmental durations and intensities as secondary cues. Finally, in comparing word and nonword sequences, we ask whether voice assimilation applies only to known word forms, as learned via experience, or also to nonce words, as the result of the application of a phonological rule pertaining to the listeners' competence.

2. TWO PRODUCTION STUDIES

Two parallel production studies have been run with French and Slovenian materials. We focus on the comparison between the French and Slovenian data. Parallel speech materials were constructed for both languages, and the same methodology was used to analyse the French and Slovenian data.

2.1. Methods

2.1.1. Speakers

Eight speakers (four female) were recorded for each language. Their mean age was 27 (French) and 24 y (Slovenian). The French speakers were from Paris or from the north of France, though not from the north-east part, where final devoicing occurs in some varieties of French. The Slovenian speakers had been all living in Ljubljana for at least 4 years; none came from regions of Slovenia where voicing could be realized slightly differently from the standard dialect of Ljubljana (see [19]).

2.1.2. Materials

For each target language, 48 two-word noun phrases with a C1#C2 stop-stop contact at word boundary were constructed. For half of them, C1 and C2 had opposite underlying voicing (assimilating contacts); C1 and C2 had the same underlying voicing (control contacts) for the other half. C1s and C2s were restricted to labial /b, p/ and dental /d, t/ stops and constrained to be heterorganic to avoid gemination. C1 was preceded by either /a/ or /i/ (controlling for the influence of this vowel) and C2 followed by /a/ where possible (80% of the time) or /o/ otherwise. The first word of the noun phrase was always a monosyllabic noun. The second word was either a two-syllable adjective (French) or a two- or three-syllable adjective where possible, or noun in the genitive case otherwise (Slovenian). This difference between the French and Slovenian materials is due to the fact that, in Slovenian, adjectives generally precede rather than follow the noun they modify. Yet, genitive-case nouns have an adjectival role similar to that of true adjectives, so that both types of modifiers are comparable. Finally, although the second word in the Slovenian materials was often three- rather than two-syllable long, as in the French materials (due to the morphosyntactic constraints of Slovenian), it was always stressed on the second syllable: the French and Slovenian materials thus shared a similar prosody. In addition to 48 noun phrases, 48 nonword sequences were constructed for each language. They were as similar as possible across the two languages, with the same C1#C2 contact conditions and the same context restrictions. Table 1 provides examples of the materials for both languages, for each contact condition, for words and nonwords. All these materials were embedded in a carrier sentence (French: *On parle jamais de__* 'You'd never say__'; Slovenian: *Bolj redko rečeš__* 'It's rare to say__').

Table 1: examples of the speech materials (F/S: French/Slovenian; W/N: word/nonword items).

		C1#C2 contact	
		voiced-voiceless	voiced-voiced
F	W	*guide patient* 'patient guide'	*guide bavard* 'talkative guide'
	N	*chide palcotte*	*chide bafique*
S	W	*zid palače* 'palace wall'	*zid barake* 'cabin wall'
	N	*šid palkota*	*šid bafik*
		voiceless-voiced	voiceless-voiceless
F	W	*pape damné* 'damned pope'	*pape tardif*
	N	*dape davine*	*dape taphile*
S	W	*slap Dolinke* 'Dolinka waterfall'	*slap Tolminke* 'Tolminka waterfall'
	N	*dap davina*	*dap tafil*

2.1.3. Recording procedure

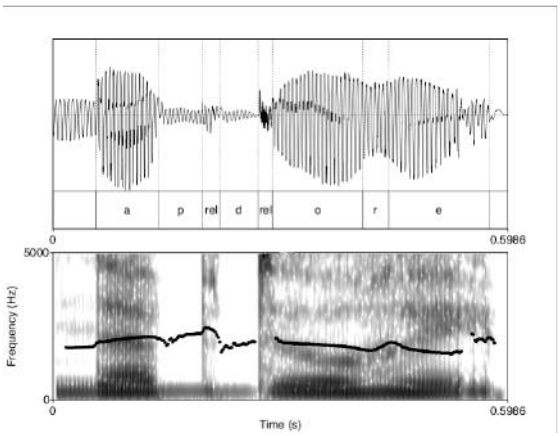
The speakers were asked to read aloud the materials, as fluently as possible, without pauses; whenever they failed to produce a sentence fluently or tripped over a word, they were asked to repeat that sentence immediately. They were given a target speech rate of 6 syllables/s they tried to follow. Nonwords were written in a transparent orthography for both French and Slovenian. For example, *chide bafique* (French) or *šid bafik* (Slovenian) for /ʃid#bafik/. Half of the speakers read the nonword materials first; the other half read the word materials first. Both word and nonword materials were presented in pseudo-random order. The speech was digitized (44.1 kHz, 16 bit) and stored in separate files for each item.

2.1.4. Segmentation and acoustic analyses

For each sequence around a C1#C2 contact, acoustic measurements were made on the C1 stop and on the preceding vowel V. Measurements on C2 and the next vowel are underway and are not reported here. We distinguish between primary and secondary cues to voicing and, within the latter, between local and non-local cues, following in that [4]. We compared two primary cues: voicing-ratio (henceforth, *v-ratio*) and HNR (see [12] for details). We examined the duration and intensity of the closure and release of C1, as local secondary cues, and the duration of V as

a non-local secondary cues to voicing. As a first step for performing the measurements, the sequences were segmented into V, C1 closure, and C1 release from visual inspection of spectrograms and, when needed, with the help of a spectral derivative function (see [12] for details). Fig. 1 illustrates the labelling of one French item.

Figure 1: French item *nappe dorée* /nap#dore/.



2.2. Results

2.2.1. v-ratio and HNR

Tables 2A-B summarize the v-ratio ([0,1]) and HNR (dB) data, respectively. Similar patterns are observed: v-ratio and HNR are larger for voiced than voiceless C2, as well as for voiced than voiceless C1. The Δ variation shown in Tables 2A-B between assimilation and control contacts measures the overall amplitude of assimilation. We ran subjects analyses of variance on these data with *Language* (French vs. Slovenian) as a between factor, *Lexical status* (word vs. nonword items), C1-voicing (*VC1*: voiced vs. voiceless), and C2-voicing (*VC2*: idem) as within factors. The significance of the latter factor directly indicates that of assimilation effects. Due to space limitations, we do not discuss the effects of C1 place (labial vs. dental) and V context (/a/ vs. /i/).

Table 2A-B: (a) v-ratio, (b) HNR data: control and assimilation contact, assimilation–control variation (Δ). (V/NV: voicing; F/S and W/N: see Table 1).

(A)		NV-NV	NV-V	Δ	V-V	V-NV	Δ
F	W	0.19	0.65	+0.46	0.99	0.50	- 0.49
	N	0.19	0.51	+0.32	0.92	0.48	- 0.44
S	W	0.19	0.82	+0.63	0.93	0.21	-0.72
	N	0.19	0.71	+0.52	0.86	0.20	-0.66

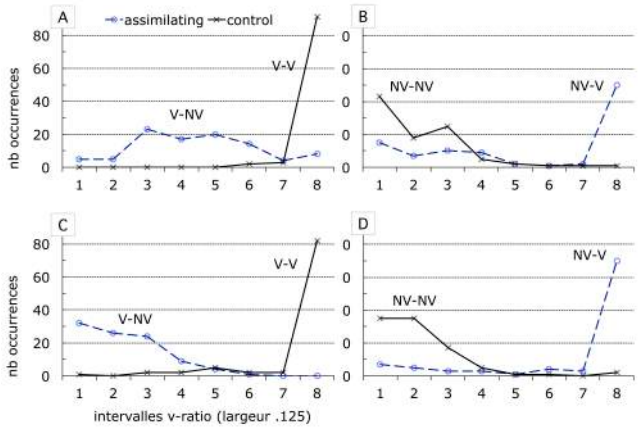
(B)		NV-NV	NV-V	Δ	V-V	V-NV	Δ
F	W	- 95	- 26	+69	11	- 58	- 69
	N	- 92	- 49	+43	2	- 58	- 60
S	W	- 91	-4	87	4	-87	-91
	N	- 98	-18	80	-2	-90	-88

VC2 was highly significant, *p*s<.00001, for both v-ratio and HNR: assimilation strongly affects these primary cues. For v-ratios, the *VC2* x *Language* , *VC2* x *VC1*, and *VC2* x *Lexical status* interactions were significant at least at the *p*<.01 level, indicating stronger assimilation overall for Slovenian than French (|Δ|=.63>.43), in the devoicing than voicing direction (|Δ|=.58>.48), and for word than nonwords (|Δ|=.58>.49). The same differences were found for HNR but with marginal significance except with *Lexical status* (*Language*: |Δ|=87>60 dB; *VC1*: |Δ|=77>70 dB; *Lexical status*: |Δ|=79>68 dB).

In more detail, the difference between devoicing and voicing assimilations was significant for both HNR and v-ratio only for nonwords; those between words and nonwords, and between Slovenian and French were significant across the board.

Interestingly, the v-ratio and HNR patterns were closely similar. This was expected since both are measuring voicing degree, though in different ways. We now focus on v-ratios to address the issue of whether voice assimilation is differently categorical in French and Slovenian. Fig. 2 shows the v-ratio distributions for words, in the assimilating and control C1#C2 contacts, as a function of C1's underlying voicing. For control contacts in French, v-ratios are concentrated in the last interval].875, 1] (C1=/b, d/), and in the first three intervals [0, .375] (C1=/p, t/), defining the voiced and voiceless categories without assimilation; v-ratio distribution for NV-V contacts (Fig. 2B) suggests a switch of the two categories. In the case of V-NV contacts (Fig. 2A), v-ratios fall in between the two categories. Assimilation in French thus appears categorical in the voicing, not the devoicing direction. The pattern is different for Slovenian: assimilation appears to be categorical in both directions. The distributions for nonword items follow the same patterns.

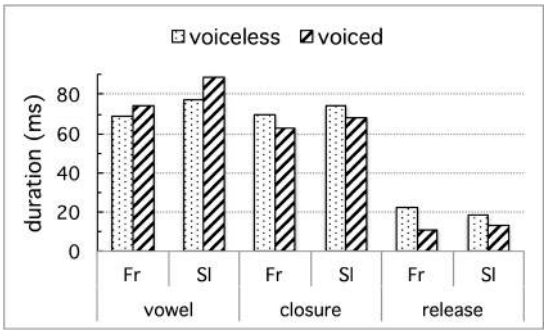
Figure 2: v-ratio distributions for voiced (A,C) and voiceless (B,D) C1, for the French (A,B) and Slovenian (C,D) data (word items).



2.2.2. Durations

Typically, vowels are longer before voiced than voiceless obstruents, whereas voiced obstruents are shorter than voiceless obstruents. In stops, both closure and release are shorter for voiced stops. We found this pattern for both the French and Slovenian data and for both words and nonwords in the control C1#C2 contacts, as summarized in Fig. 3.

Figure 3: Duration measurements (words and nonwords pooled) in control contacts.

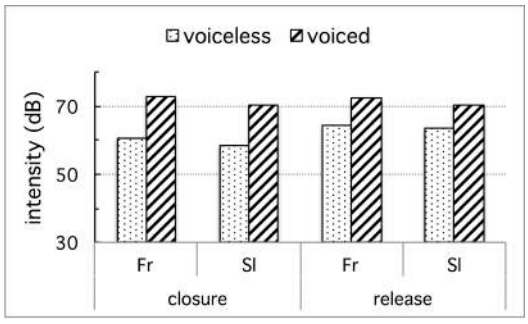


Were these patterns affected by assimilation, the durations of voiced C1s should increase, and that of preceding Vs decrease in assimilating contexts. Opposite changes should obtain for voiceless C1s. Such changes occurred significantly for C1 release duration in both languages but asymmetrically (45% increase vs. 15% decrease), in line with stronger devoicing than voicing assimilation (§2.2.1). V and C1 closure durations also changed significantly in the expected direction for Slovenian (~5%); for French, they changed in the expected direction but not significantly, and for devoicing only (-3%). Note that in all cases, voice assimilation only tended to neutralize the duration pattern and did not revert it.

2.2.4. Intensities

Closure and release intensities are larger for voiced than voiceless C1s in control contacts for both French and Slovenian, as shown in Fig. 4.

Figure 3: Intensity measurements (words and nonwords pooled) in control contacts.



Assimilation contacts induce significant intensity changes relative to control contacts in the expected direction for both closures and releases: ~13% decrease (devoicing) vs. ~9% increase (voicing). The changes are symmetrical in the Slovenian data (~12.5% in either direction) but not in the French data (~13% decrease vs. 6.5% increase), in line with a stronger devoicing than voicing assimilation.

3. DISCUSSION

This study showed four main points about voice assimilation: (i) it is "stronger" in Slovenian than French, in the sense it induces a larger change in the primary cues to voicing (voicing-ratio, HNR); (ii) it is categorical in Slovenian for both the voicing and devoicing directions; in French it is categorical only for the voicing direction; (iii) it is similar for words and nonwords; (iv) some secondary cues to voicing (preceding vowel and stop closure durations) resist assimilation in French but not in Slovenian. Finally, there is a trend, showing across primary and secondary cues, for stronger assimilation in the direction of devoicing than voicing.

Our data are consistent with the scarce literature on Slovenian: [18] found complete assimilation in the devoicing direction but not in the voicing direction. For French, previous data in [10, 11, 2] showed either no such asymmetry or stronger voicing than devoicing assimilation [16]. (But this latter study did not use control contexts to quantify degrees of assimilation.) Our findings for secondary cues in French are partly in line with those of [3] for moderate speech rates, those of [17], and those of [4]: vowel and closure durations tend to be preserved in assimilations or tongue twister slips of the tongue. Note that the distributional data shown in Fig. 2, showing a more categorical assimilation in the voicing direction is in line with the assumption made in [17] that French *soute* can fully assimilate to [sud] in terms of v-ratio. Yet, the residual secondary cues to the voicelessness of *soute* were sufficient for listeners to recover the speakers's intended meaning. Our present study confirms that these residual cues are indeed substantial.

Finally, the close similarity between the word and nonword data, suggests that, at least for read speech, assimilation is governed by phonological rules rather than learned alternations presumably stored in memory as "exemplars," which could be more strongly memorized for frequent words and induce much stronger assimilation than nonwords.

4. ACKNOWLEDGEMENTS

This study was supported by LabEx EFL (first two authors) and an ENS doctoral fellowship to the third.

4. REFERENCES

[1] Abdelli-Beruh, N. 2004. The stop voicing contrast in French sentences: Contextual sensitivity of vowel duration, closure duration, aspiration duration and closure voicing. *Phonetica*, 61, 201-219

[2] Abdelli-Beruh, N. 2011. Voicing and devoicing assimilation of French /s/ and /z/. *Journal of Psycholinguistic Research*. 1-16.

[3] Abdelli-Beruh, N. 2012. Voicing assimilation of French /t/. *J. Phonetics* 40, 521-534.

[4] Goldrick, M., Blumstein, S. 2006. Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes* 21, 649-683.

[5] Gow, D. 2001. Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language* 45, 133-159.

[6] Gow, D. 2002. Does English coronal place assimilation create lexical ambiguity? *Journal of Experimental Psychology: Human Perception and Performance* 28, 163-179.

[7] Gow, D. 2003. Feature parsing: Feature cue mapping in spoken word recognition. *Perception and Psychophysics* 65, 575-590.

[8] Gow, D., Im, A. 2004. A cross-linguistic examination of assimilation context effects. *Journal of Memory and Language* 51, 279-296.

[9] Grammont, M. 1939. *Traité de Phonétique*. Paris: Delagrave.

[10] Hallé, P., Adda-Decker, M. 2007. Voicing assimilation in journalistic speech. *Proc. 16th ICPhS Saarbrucken*, 2325–2328.

[11] Hallé, P., Adda-Decker, M. 2011. Voice assimilation in French obstruents: Categorical or gradient? In: Goldsmith, J., Hume, E., Wetzels, L. (eds.), *Tones and features*. De Gruyter Mouton, 149–175.

[12] Hallé, P., Androjna, K., Segui, J. 2012. L'assimilation de voisement en français : Elle vaut pour les non-mots autant que les mots. *Proc. 29th JEP* Grenoble.

[13] Lisker, L. 1986. "Voicing" in English: a catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech* 29. 3-11.

[14] Nolan, F. 1992. The descriptive role of segments: Evidence from assimilation. In: Docherty, G., Ladd, D. (eds.), *Laboratory phonology II: Gesture, segment, prosody*. Cambridge: Cambridge University Press, 261-280.

[15] Rigault, A. 1967. L'assimilation consonantique de sonorité en français: étude acoustique et perceptuelle. *Proc. 6th ICPhS* Prague, 763-766.

[16] Snoeren, N., Hallé, P., Segui, J. 2006. A voice for the voiceless: Production and perception of assimilated stops in French. *J. Phonetics* 34, 241-268.

[17] Snoeren, N., Segui, J., Hallé, P. 2008. Perceptual processing of partially and fully assimilated words in French. *Journal of Experimental Psychology: Human Perception and Performance* 34, 193-204.

[18] Srebot-Rejec, T. 1990. Zveze dveh zapornikov v slovenščni in angleščini. [Two-stop consonant clusters in Slovene and English]. *Slavističina revija* 38, 265-28438, 265-284.

[19] Toporišič, J. 1976/2004. *Slovenska slovnica* [*Slovene grammar*]. Založba Obzorja: Maribor.