

A Voice for the Voiceless: Voice Assimilation in French

Natalie D. Snoeren and Juan Segui

Laboratoire de Psychologie Expérimentale, CNRS-Paris V, Paris, France

E-mail: snoeren@psycho.univ-paris5.fr, segui@psycho.univ-paris5.fr

ABSTRACT

Past research has shown that the speech processing system is, under certain circumstances, relatively tolerant to surface variations in connected speech. The goal of this study is to explore regressive voice assimilation in French, a surface variation in which the voicing of a consonant is modified by the voicing of the following consonant.

French subjects pronounced sentences containing words with plosive codas inserted in either assimilatory or non-assimilatory contexts. We carried out acoustic analyses and defined an index for voice assimilation. We observed that regressive voice assimilation in French is a graded rather than an all-or-none phonological phenomenon. The strength of assimilation did not depend on the following context (stops or fricatives). There was a significant effect of the underlying voicing: voiceless stops assimilated more strongly than voiced stops. Such acoustic trends should be taken into account when studying the perceptual consequences of voice assimilation.

1. INTRODUCTION

Continuous speech is susceptible to various surface variations. Words can deviate from their underlying form, either by accident or because of regular, phonologically motivated surface variations. Yet, surface variations do not necessarily hamper speech perception and therefore have received a lot of attention from psycholinguists [2,3,4,11].

A known example of regular surface variation in running speech is place assimilation [3,4,8]. Assimilation may be defined as the spreading of some phonetic features from one sound to an adjacent sound [1,5]. For instance, the velar feature of /g/ in *sweet girl* can spread to the preceding alveolar stop /t/, as a consequence *sweet girl* can be realised phonetically as *sweek girl*.

The perceptual consequences of assimilation have been extensively studied, especially in English [3,4], but also in German [2], and in Japanese [11]. In the literature several hypotheses have been put forward in order to explain how listeners cope with place assimilation: the under-specification theory, or more recently the regressive inference account [3]. The processing system would include regressive inference mechanisms that allow for a re-analysis of the assimilated segment according to its context, and eventually recover its underlying form. In a phoneme monitoring experiment [3], subjects detected /t/ in *freight bearer* although /t/ had assimilated to /p/ and *freight* was pronounced [freip]. This suggests that subjects

can infer the labial feature of the stop from the right context *bearer*, and recover the underlying coronal nature of /t/ in *freight*. Recently, Gow [4] found that listeners do not necessarily use contextual information to “recover” the underlying consonant. Gow used potentially ambiguous phrases such as *right berries*, that can be confused with *ripe berries* after assimilation. Listeners heard the intended *right*, even in case of strong assimilation, and they didn’t make use of regressive inference mechanisms that should have made them hear *ripe* and not *right*. Gow concluded that place assimilated consonants contain residual acoustic cues that signal *two* places of articulation, one related to the underlying consonant and the other to the assimilatory consonant. Listeners would be able to exploit simultaneously these acoustic cues and identify the underlying nature of the assimilated consonant. Gow argued that the acoustic properties of the stimuli explain the apparently discrepant results found by different investigators. In their experiences, Gaskell and Marslen-Wilson [3] used *deliberately* assimilated stimuli, whereas Gow used stimuli that were naturally produced and were less assimilated. Thus, different degrees of assimilation might yield different processing strategies. Near-complete assimilation requires the use of the following context to recover the underlying identity of the segment, hence the “regressive inference” effects found by Gaskell and Marslen-Wilson and others [2,3]. Intermediate and, arguably, natural assimilation allows anticipation of the upcoming segments because they should be consistent with the assimilation: hence the “progressive” effects [4, 11].

This brief review suggests that an appropriate phonetic description of phonologically motivated variations, such as assimilation, is essential in order to account for differences in perceptual consequences. In this study, we therefore focused on the phonetic-acoustic properties of voice-assimilated consonants in continuous speech in French. Many aspects of voice assimilation in natural running speech have not been extensively studied, and in particular its graded *versus* all-or-none nature. We strived to fill that gap.

Regressive voice assimilation can occur when two consonants C_1 and C_2 with different voicing are in direct contact. C_1 typically “inherits” the voicing feature [+voice] or [-voice] of C_2 [1,5,9]. For instance, the final stop /b/ in *robe* (‘dress’) which is normally voiced, as in *robe brune* (‘brown dress’), can become a voiceless /p/ in the sequence *robe sale* (‘dirty dress’), due to regressive

voice assimilation induced by the following voiceless fricative /s/ in *sale*.

Rigault [9] has studied in a rather detailed way voice assimilation of the /d/ in the context of /s/. He found that, /d/ indeed assimilated to /t/ and, moreover, became undistinguishable (acoustically *and* perceptually) from /t/ in /t/ + /s/, while it radically differed from /d/ in /d/+z/. Rigault concluded that the “fortis”-“lenis” distinction previously proposed by French linguists (e.g. Grammont, [5]) does not apply to voice assimilation in French. That is, there is no acoustic or perceptual difference between [t] resulting from an underlying /d/ devoiced by assimilation, as *medecin* and [t] as the plain realisation of /t/ in *Hauteserve*. To summarise, Rigault defended the view of a radical, complete voice assimilation in French, both within words as in *medecin* and between words as in *robe sale*. Other authors offer a less extreme view proposing, in particular, that assimilation is more optional between than within words [1].

The goal of the present study is to explore voice assimilation from a phonetic-acoustic point of view. We focused on assimilation between rather than within words, because between words assimilation potentially leaves more room to variation in French [1]. Our main concern was whether voice assimilation is complete when it occurs, as Rigault claimed [9], or whether it can be graded (as Nolan [8], proposed for place assimilation in English). We also wanted to explore assimilation of voiced *versus* voiceless word-final stops, and the potential role of the manner of articulation of the following consonants that cause assimilation.

2. PRELIMINARY EXPERIMENT

2.1 Stimuli

We inserted French words in assimilatory and non-assimilatory sentential contexts. Assimilatory contexts stand for those that can trigger voice assimilation. For instance, the final voiceless stop /t/ of the French word *frites* /frit/ (“fried potatoes”) may be pronounced as a /d/ when followed by a voiced consonant in *les frites de Bruxelles* (“French fries from Brussels”). A non-assimilatory context cannot give rise to voice assimilation. For example, when *frites* is followed by a voiceless consonant /k/ such as in *les frites croustillantes* (“crunchy French fries”), *frites* cannot be pronounced as /frid/ (with a /d/). In a pre-test, we presented a list of sentences to 12 native speakers of French, who had to rate the plausibility of the sentences, that is their syntactic and semantic well-formedness, on a 1-5 scale. Twenty sentences that were judged highly plausible remained and contained the same 10 plosive-coda monosyllabic words: half of the words were inserted in assimilatory contexts, the other half in non-assimilatory contexts. Seven out of 10 critical words had a voiceless coda and three a voiced coda. Seventeen out of 20 sentences had the critical word followed by a stop and three had the critical word followed by a

fricative.

2.2 Subjects and procedure

We presented the experimental set of sentences in print to four native speakers of French (two men and two women). They were researchers at the Experimental Psychology Laboratory at the University René Descartes (Paris V). We asked them to produce the sentences fluently at a normal speech rate. The speakers were not informed of the goal of the experiment. All the sentences were recorded on DAT and transferred to computer files for acoustic analyses.

2.3 Acoustic analyses: degree of voice assimilation

For every critical word, such as *frites* inserted in an assimilatory context (*les frites de Bruxelles*), we measured several acoustic cues to voicing using spectrograms. We measured (1) the duration of the vowel preceding the occlusion, (2) the duration of the occlusion, and (3) the duration of the *voiced* portion of the occlusion. Vowel duration, relative or not to occlusion duration, did not prove to be a reliable index for voicing. On the other hand, the relative duration of the voiced part of the occlusion, that is the (3)-to-(2) ratio seemed to provide a more reliable index, a degree of voicing. To ascertain this point, we extracted the critical items from their sentential contexts and presented them auditorily to 22 French participants whose task was to categorise the word-final consonant (forced-choice identification with rating, among the choice-set {p, t, k, b, d, g, s, z}). There was a highly significant correlation between the degree of voicing and the rating weighted participants’ judgment of voicing that indicated the perceptual value of this acoustic cue. In the following, we used the ratio defined above, -we call it *V-ratio*- as an index of voicing. The *V-ratio* allows for the definition of the “degree of voice assimilation”: it stand for the amount of voicing found in originally voiceless stops in an assimilatory context, or for the amount of devoicing in originally voiced stops. Expressed in percentage, the degree of assimilation is thus $100 \times v\text{-ratio}$ for originally voiceless stops, and $100 \times (1-V\text{-ratio})$ for voiced stops.

2.4 Results

Table 1 shows the degrees of voice assimilation for voiced and voiceless word-final stops in assimilatory contexts calculated for four speakers.

Speaker	Voiceless	Voiced
1	97	31
2	67	22
3	100	59
4	100	34
Mean	91	37

Table 1: Degrees of voice assimilation (%) per speaker for originally voiceless and voiced word-final stops.

A closer look at the data showed that voice assimilation was incomplete (degree of assimilation in the [20%-80%] range) about 33% of the data. The data in Table 1 also suggests that word-final voiceless stops assimilate more

strongly (91%) relative to word-final voiced stops (37%). However, we are not in the position to make any strong claims as to the apparent asymmetry since our stimulus-set was not balanced (there were 7 voiceless stop codas and only 3 voiced stop codas).

This exploratory experiment obviously called for more rigorously planned comparisons. We therefore carried out a second production experiment with a more appropriate design. First, we used the same number of voiced and voiceless items. Second, we controlled for the kind of right context determining voice assimilation using the same number of fricative and stop consonants. Indeed, the degree of assimilation might depend on the manner of articulation of the assimilatory context, maybe with fricatives giving more assimilation [5]. A further improvement was having speakers from a homogenous population with respect to age, social background and regional “dialect”. To sum up, the second experiment was designed to investigate more thoroughly the graded nature of voice assimilation in French, the possible asymmetry of assimilation for voiced and voiceless word-final stops, and the possible role of the manner of articulation of the right context on assimilation.

3. EXPERIMENT 2

3.1 Stimuli

Two sets of 20 monosyllabic noun words with a stop-coda were selected : the final stop was voiced in one set, voiceless in the other set. The two sets were balanced in terms of frequency (38.71 occurrences per million for voiced items and 39.23 occurrences per million for voiceless items, most words that we used in Experiment 1 were “recycled”). Compared to Experiment 1, the critical words were inserted in a simpler embedding context: usually a noun phrase (e.g. *une jupe grise*) instead of a complete sentence; the context was always assimilatory, that is the critical word-final stop and the next consonant had a different voicing. In each set of 20 test utterances, half of them had a fricative context (voiced: /v, Z/; voiceless: /f, s/), the other half a plosive context (voiced: /b, d, g/; voiceless /p, t, k/). Forty fillers were added to the materials, so that participants had to read aloud eighty utterances in all.

3.2 Participants and Procedure

Five male and five female students at the University René Descartes, Paris, participated in the experiment. All of them were native speakers of French and came from the Paris region. They were presented with the list of 80 utterances and asked to read them at a normal speech rate, as fluently as possible. All recordings were made on DAT and transferred to computer files for acoustic analyses.

3.3 Results

Figures 1 and 2 show the distribution of the degrees of voice assimilation observed for word-final consonants followed by a stop (Figure 1) or a fricative (Figure 2). These distributions illustrate that between-word voice assimilation in French is not an all-or-none but rather a

graded phenomenon. Second, no significant difference was found between the assimilation caused by stops and that caused by fricatives. Finally, the tendency found in Experiment 1 of more assimilation for underlying voiceless than voiced stops was confirmed. The degree of voice assimilation was significantly greater for voiceless stops (mean 75%) than for voiced stops (mean 29%), $t(19)=9.34, p<.0001$.

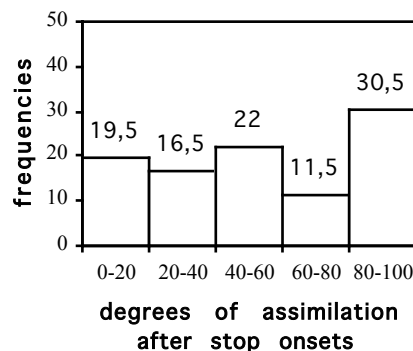


Figure 1: Distribution of the degree of voice assimilation for stop consonants followed by (assimilatory) stops.

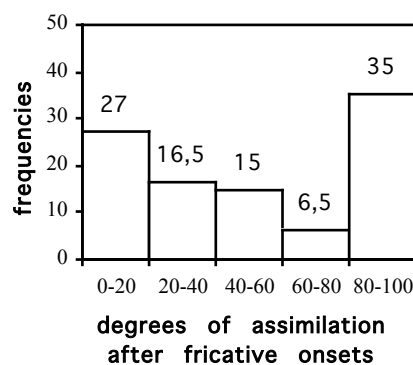


Figure 2: Distribution of the degree of voice assimilation for stop consonants followed by (assimilatory) fricatives.

4. CONCLUSIONS

In this study, we explored regressive voice assimilation in French. The study focused on the phonetic-acoustic characterisation of voice assimilation between words. An accurate description of assimilation in this situation is a prerequisite to understand how French listeners recognise words in spite of the form variation presumably caused by voice assimilation. We had French participants to pronounce sentences containing words with a plosive coda, embedded in non-assimilatory and assimilatory contexts. In a preliminary experiment (Experiment 1), we carried out acoustic analyses that served to define an index of voicing that was used to define a “degree of voice assimilation”. We observed that in about one third of the cases, voice assimilation was not complete: voice assimilation in French might be a graded rather than an all-or-none process (this is reminiscent of place assimilation in English, [8]). We also observed a tendency for word-final voiceless stops to be more strongly affected

by assimilation than voiced stops. The trends observed in Experiment 1 were confirmed in Experiment 2, in which a rigorously designed set materials was used. In particular, the graded nature of between-word voice assimilation was confirmed as well as the asymmetric treatment of voiced and voiceless stop codas. The design of Experiment 2 also allowed for exploring the possible differences in assimilation caused by two kinds of assimilatory contexts: stop consonants (/b, d, g/ and /p, t, k/) and fricative consonants (/v, Z/ and /f, s/). No significant difference in degree of voice assimilation was found between these two kinds of context.

A robust finding across the two experiments was that the degree of assimilation was clearly greater for voiceless than for voiced stops. That is, the /p/ in *une jupe droite* changed more clearly into a [b] than the /b/ in *une robe claire* changed into a [p]. At first sight, this seems a rather counter-intuitive result considering the phonological principle of coda-obstruent devoicing (*[+VOICE/_] in its usual form), which has frequently been posited as accounting for a universal tendency [10]. Coda-obstruent devoicing is partly motivated by articulatory-phonetic considerations: a voiced stop coda, for example, presumably requires a quite perfect timing of laryngeal gestures in order to maintain voicing during occlusion, then inhibit voicing at occlusion release –if the next segment is voiceless. Producing a voiceless occlusion right away is presumably an easier manoeuvre. These are precisely the situations encountered in word-final stop voice assimilation. The coda-devoicing principle would let one expect “more” assimilation for underlying voiced than voiceless stops. Yet we observed the reverse pattern in French across two production experiments. A possible explanation of this apparent paradox is that the phonological principle of coda-obstruent devoicing is based on an all-or-none view of voicing. Yet, as we observed in this study, there are intermediate degrees of voicing. If the coda stops that were not fully voiced were to be considered as voiceless, a very different picture would have emerged in terms of how many voiceless versus voiced coda stops did or did not “assimilate.” Assimilation would, perhaps, be found *more frequent* for voiced than for voiceless stops. Another kind of explanation could stem from differential lexical frequencies. There are more words with a voiceless than with a voiced coda stop in French (according to “Lexique” [7]): 2556 against 1091). One explanation of our results is in terms of “compensation,” or, equivalently, in terms of “preservation” of the marked rather than the unmarked form: voiceless coda stops may be considered as the default, unmarked case, and are more easily altered. This view implies that speakers implicitly use lexical frequencies in language use (see [6]). If this is correct, we should find different voice assimilation patterns in languages with a different distribution of voiced and voiceless stop codas. At any rate, we believe that an accurate phonetic-acoustic description of voice assimilation should provide a good basis for further investigations.

REFERENCES

- [1] Carton, F., *Introduction à la phonétique du français*, Paris: Bordas, 1974.
- [2] Coenen, E., Zwitserlood, P., & Bölte, J., “Variation and assimilation in German: consequences for lexical access and representation”, *Language and Cognitive Processes*, vol. 16, pp. 535-564.
- [3] Gaskell, M.G. & Marslen-Wilson, W.D., “Mechanisms of phonological inference in speech perception”, *J. Exp. Psychology: Human perception and Performance*, vol. 24, pp. 380-396, 1998.
- [4] Gow, David W., “Does English coronal place assimilation create lexical ambiguity?”, *J. Exp. Psychology: Human perception and Performance*, vol. 28, pp. 163-179, 2002.
- [5] Grammont, M., *Traité de Phonétique*, Paris : Delagrave, 1933.
- [6] Hay, J., Pierrehumbert, J., & Beckman, M., (in press), “Speech perception, wellformedness and the statistics of the lexicon”, *Proceedings of 6th Conference on Laboratory Phonology*, York, Great Britain: University of York.
- [7] New, B., Pallier, C., Ferrand, L., & Matos, R., “Une base de données lexicales du français contemporain sur internet : LEXIQUE”, *L'Année Psychologique*, vol. 101, pp. 447-462, 2001.
- [8] Nolan, F., “The descriptive role of segments: Evidence from assimilation”, in D.R. Ladd & G. Docherty (Eds.), *laboratory Phonology II: Gesture, Segment, Prosody*, Cambridge: Cambridge University Press, pp. 261-280, 1992.
- [9] Rigault, A., “L’assimilation consonantique de sonorité en français : Etude acoustique et perceptuelle”, in B. Hala, M. Romportel, & P. Janota (Eds.), *Proceedings of the 6th International Congress of Phonetic Sciences*, pp. 763-766, 1967.
- [10] Steriade, D., “The phonology of perceptibility effects: the P-map and its consequences for constraint organization”, unpublished manuscript.
- [11] Otake, T., Yoneyama, K., Cutler, & A., van der Lugt, “The representation of Japanese moraic nasals”. *J. of the Acoustic Society of America*, vol. 100, pp. 3831-3842, 1996.