PHONOLOGICAL CONTEXT EFFECTS FOR VOICING AND DEVOICING IN FRENCH

Isabelle Darcy & Frank Kügler

Indiana University & Potsdam University
darcy.isa@gmail.com; f.kuegler@gmail.com

ABSTRACT

We examine occurrences of categorical assimilation (neutralizations) in French, the perception of voiced and unvoiced word-final obstruents in different phonological contexts. We first show the categorical nature of the alternation (Exp. 1), supported in Exp. 2 by perceptual categorization data. In Exp. 3, the interpretation of this first percept appears to be corrected in certain contexts, inducing compensation. We argue that context effects are phonological in this case, rather than auditory or phonetic. We conclude that linguistic knowledge of alternations is necessary in compensation for categorical assimilation.

Keywords: voicing and devoicing assimilation, French, context effects, perceptual learning.

1. INTRODUCTION

Speech sounds are not processed individually; the perception or the labeling of a sound can be shifted depending on the properties of the context in which it occurs. This context effect can help listeners recovering from variation in the speech signal, especially from context-induced alternations, like in the cases of coarticulation and assimilation (e.g. voicing assimilation in French: *ro*[b] 'dress'; [b] is devoiced in *ro*[p]#[s]*ombre* 'dark dress', but not in *ro*[b]#[n]*oire* 'black dress'). Different properties of the context have been argued to be responsible for the effect:

- acoustic or phonetic, e.g. [6, 8, 12, 13].
- phonological (related to native language experience) e.g. [1, 2, 5]
- lexical, e.g. [11]
- statistical, e.g. [14].

Studies of compensation for phonological assimilation have observed robust context effects, yet attributing them to different processing levels (acoustic or phonetic [12, 6] vs. phonological [2, 5]). The question remains which processing mechanisms these context effects actually reflect, and whether or not specific linguistic knowledge is involved in compensation.

In auditory processing accounts, context effects arise before a sound is categorized and do not need any specific linguistic knowledge [12]. Compensatory phonetic accounts situate context effects at a post-auditory level, still before categorization; a change is compensated for through reattribution of some of the perceived phonetic cues for one segment to its context, identified as the source of the change [6]. This is linguistic in nature, but does not require specific experience with the assimilation process. The other group of context effects observed in compensation for assimilation [2, 5] relies on higher level phonological knowledge: an assimilatory change presented in a context appropriate for assimilation (e.g. [s] in the above example) will be compensated for more often than when presented in a non-appropriate context (e.g. [n] above). This view has been challenged by context effects obtained with non-humans, in non-speech contexts and at very early latencies [8, 9, 12].

Typically, context effects observed in phonetic categorization tasks require ambiguous stimuli. Both auditory and phonetic context effects in compensation for assimilation arise before sounds are categorized, hence can affect the listener's bias to label the sound as belonging to a particular category; they have been obtained at these levels of processing on the basis of doubly articulated, ambiguous segments [6, 12]. When acoustic information for a sound is unambiguous with respect to category membership (i.e. when changes are categorical), context effects play little role in categorization [7, 4].

If context effects found in compensation for assimilation reflect an early and language independent processing, then they are predicted to be weak in case of categorical assimilation. If on the contrary, context effects emerge despite categorical changes, they must have arisen at a higher processing level (after categorization), and be contingent

on a specific linguistic knowledge. Even though previous research observed context effects with categorical occurrences of assimilation (through cross-splicing or deliberate pronunciation [2, 5]), it has been difficult to clearly rule out auditory or phonetic compensation accounts for these effects due to various factors (different acoustic cues involved, no acoustic analyses, among others).

The present paper seeks to strengthen the claim that specific linguistic knowledge of processes active in a language might be used in compensation for categorical assimilation. To that purpose, we test complete occurrences of French voicing assimilation in perception. We first report the acoustic analysis (Exp. 1) of the stimuli used, confirming the categorical nature of assimilatory changes. Perceptual categorization of the assimilated consonants is performed in Experiment 2. Finally, Experiment 3 assesses the interpretation of assimilatory changes in different contexts.

2. ACOUSTIC ANALYSIS

2.1. Stimuli and Method

2.1.1. Stimuli

Sixteen French nouns (8 Voicing and 8 Devoicing, C(C)VC structure) were selected as targets. Voicing items ended in a final unvoiced obstruent (e.g. [lak]), devoicing items in a voiced one (e.g. [rob]). Sixteen matched nonwords ([nw]) were constructed by switching the voicing feature of the final obstruents (e.g. robe /rob / 'dress' - rope /rop/ [nw], or lac /lak/ 'lake' - lague /lag/ [nw]). Each target was associated with a triplet of context words (adjectives), each corresponding to one of the experimental conditions (assimilation type)*: viable change, unviable change, and no-change*. For the viable change condition (e.g. *ro*[ps]*al* 'dirty dress'), the adjective's initial consonant was an obstruent agreeing in voicing with the nonword matched to the target item. The adjectives in the unviable change and no-change conditions (e.g. *ro*[pn]*oire*, 'black dress'; *ro*[br]*ouge*, 'red dress') both started with a neutral consonant with respect to the relevant assimilation process. For both sets equally, this neutral consonant was a sonorant.

For each of the 16 target items, 3 sentence frames were constructed (e.g. Elle a mis sa ___ ___ aujourd'hui. 'She put on her ___ ___ today.'). Combining the three conditions with the three sentence frames gave rise to 9 actual sentences associated to each item. This resulted in a total of 144 sentences. All sentences were rotated across three lists, so that each participant heard each item in each condition, but in different carrier sentences. 30 additional filler sentences were constructed, and served as training or distractors. All sentences were recorded by the first author, a female native speaker of French. They were digitized at 16kHz and 16bits on an OROSAU22 sound board, and edited. Changes were deliberate, pronounced naturally and as clear as possible, without release between target and context word.

2.1.2. Acoustic Analyses

In order to validate the full neutralization of the voicing contrast in the sequence Target word + Context word ($C\underline{V_1}\underline{C_1}\#\underline{C_2}\underline{V_2}C$), assimilated word-final consonants must be completely devoiced or voiced to the same extent as their underlyingly voiceless or voiced counterparts. Similarly, vowel duration must not give cues about the underlying voicing of the following consonant [10]. We determined the following critical values: a) Vowel length in V_1 position; b) Closure duration for the $C_1\#C_2$ sequence; c) Duration of voicing into closure for C_1. The duration of (c) indicates the portion of the closure where some glottal activity persists. Voicing into closure for voiced obstruents is comprised between 80% and 100% of the closure duration of C_1, while for unvoiced obstruents it is not necessarily zero; it can be around 30% [15].

2.2. Results

We use a voicing degree measure computed from the duration of voicing into closure as a ratio of the whole C_1 duration. Basis of comparison for each assimilatory change is the non assimilated underlying form (*no D, no V*; see Table 1).

Table 1: Vowel duration and voicing degree according to assimilation type (V=Voicing, D=Devoicing), underlying and surface voicing.

assim. type	duration V1 (ms)	underl. V	surf. V	voicing degree
viable D	72	[+vd]	[p]	15 %
viable V	63	[-vd]	[b]	94 %
unviable D	73	[+vd]	[p]	13 %
unviable V	63	[-vd]	[b]	89 %
no D	79	[+vd]	[b]	87 %
no V	61	[-vd]	[p]	13 %

An underlyingly voiceless consonant (*no V*) following a vowel is voiced for 13% of its closure duration; underlyingly voiced segments (*no D*)

display clear voicing for the most part of C1-closure (87%), comparable to [15]. Comparisons between contexts were performed on the basis of the surface voicing. Neither the analysis of surface voiced conditions ($p>.2$) nor that of surface unvoiced conditions ($p>.5$) revealed any effect of "assimilation type", suggesting that a given surface voicing is similar in all 3 contexts. We conducted an analysis of variance on voicing degree including the variables "assimilation type" (restricted to change conditions: viable and unviable) and "underlying voicing" (voiced, unvoiced). The analysis revealed that "assimilation type" has no effect on the voicing degree ($p>.1$). Not surprisingly, there is a highly significant main effect of "underlying voicing" on the voicing degree ($p<.0001$), but no interaction between both variables ($p>.5$).

Analyses of V1 vowel durations according to surface voicing did not reveal any difference ($p>.8$) between voiced (68 ms) and unvoiced segments (69 ms), nor any difference due to assimilation type ($p>.8$). In conclusion, acoustic analyses confirm that stimuli are categorical with respect to voicing degree; they do present occurrences of "complete assimilation" (neutralization) where assimilation was intended.

3. WORD FINAL CATEGORIZATION

3.1. Procedure and participants

All target words were excised out of the carrier sentences and presented in isolation in a forced-choice categorization task. Participants had to tick the consonant they heard on a response sheet: They were always given a choice between the original consonant and the assimilated one. For the word *robe* 'dress' for example, the choice was between [b] (unchanged) and [p] (underwent voice assimilation). A free cell allowed them to report any better matching sound, if needed. Responses are measured as "congruent response", i.e. consistent with underlying voicing. Eighteen French native speakers who did not participate in Experiment 3 were recruited for this experiment.

3.2. Results

One item in the Devoicing set containing an affricate (*badge*) turned out to induce high error rates (mean error: 30%), and was excluded from analyses and from the subsequent Experiment 3. No participant was excluded.

Table 2: Mean percentages of "responses congruent with underlying voicing" according to assimilation type, underlying and surface voicing (N = 18)

assim. type	underl. V	% congruent	SE	Surf. V	% as [+vd]
viable D	[+ vd]	7,1	1,3	[p]	7,1
viable V	[- vd]	1,8	0,6	**[b]**	98,1
unviable D	[+ vd]	0,8	0,4	[p]	0,8
unviable V	[- vd]	1,3	0,6	**[b]**	98,6
no D	[+ vd]	97,8	0,5	**[b]**	97,8
no V	[- vd]	98,3	0,4	[p]	1,6

Analyses of variance restricted to both change conditions revealed that devoicing produced significantly ($p=.012$) more "congruent" judgments (3.9%) than voicing items (1.6%); this is due to the presence of voiceless geminate closures, which are most difficult to identify. Globally, stimuli were perceived mostly categorically as voiced or voiceless, reflecting very closely the intended surface voicing and the acoustic properties of our stimuli. This suggests that no traces of the underlying voiced or voiceless counterpart, for example minimal cues following from coarticulation, is retained in the assimilated conditions, inducing no unexpected perceptual biases.

4. WORD RECOGNITION IN CONTEXT

4.1. Procedure and participants

To assess context effects, we use a word detection task: first, a prime (i.e. a target word in citation form, recorded by a male voice in order to avoid pure acoustic/phonetic matching during detection) is presented auditorily, the sentence follows after 500 ms. Participants are instructed to press "yes" as soon as possible when they think that the prime presented was the same (i.e. had the same *form*) as the target in the sentence, "no" otherwise. Our measure is detection rate in each condition. Eighteen French native speakers who did not participate in Exp 2. (categorization) participated in this experiment.

4.2. Results

The item *badge* is removed from the analyses (see 3.2). Results (Fig. 1) show clearly that compensation is higher for both voicing and devoicing when the context is appropriate for assimilation. For both, detection is low in the unviable context, and highest when no change occurred (main effect of "assimilation type" significant by subjects and

items, p<.0001). There is a main effect of underlying voicing on detection rate (p<.0001 by subjects, p<.001 by items). Considering each condition separately, we observe that detection rates are higher for Voicing as compared to Devoicing (p<.001, participants & items) in the viable condition only, being equal (p>. 1, participants & items) in both other conditions.

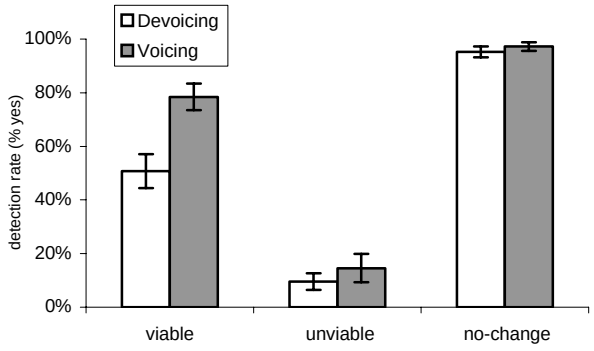

Figure 1: Compensation results (% yes) in each assimilation type according to underlying voicing, N=18

This unexpected asymmetry cannot be explained by the higher "congruent" judgment rate for devoicing items in Exp. 2, which would lead here to higher detection responses. Interestingly, this asymmetry parallels the one observed in naturally produced assimilation: devoicing is systematically less categorical (only 6% are complete) than voicing assimilation (52% are categorical) [15].

5. DISCUSSION AND CONCLUSION

On the basis of robust context effects obtained with categorical instances of voicing assimilation in French, the present paper provides strong evidence in favor of the claim that specific linguistic knowledge of processes active in a language is used in compensation for assimilation.

Experiments 1 to 3 have shown that a sound, categorized clearly without context as voiced or unvoiced, is given the other value in appropriate contexts, inducing compensation for voicing assimilation. We argue that this context effect is phonological, because it applies to categorical instances. The correct reinterpretation depends on prior linguistic knowledge of the voicing alternation in French, rather than on auditory or phonetic cues present in the stimuli. The asymmetry observed in the viable change condition suggests that an input-driven, detailed knowledge of the processes at work in a language can shape our perception: because listeners are used to incomplete occurrences of devoicing assimilation in French [15], their word recognition system is not optimally prepared to compensate for categorical changes. By contrast, listeners' word recognition system is very effective in compensating for categorical voicing assimilation [3].

The question remains as to where during processing this phonological knowledge applies, and whether it interacts with auditory or post-auditory processing stages. Further research will need to delineate their respective contribution in categorization, compensation and word recognition.

6. REFERENCES

[1] Beddor, P. S., Krakow, R. A. 1999. Perception of coarticulatory nasalization by speakers of English and Thai: Evidence for partial compensation. *JASA* 106,2868-2887.

[2] Darcy, I., Peperkamp, S., Dupoux, E. in press. Bilinguals play by the rules: perceptual compensation for assimilation in late L2-learners. In: Cole, J., Hualde, J. (eds), *Papers in Laboratory Phonology 9*.

[3] Darcy, I., Kügler, F. in preparation. Voicing and devoicing in French: a perceptual asymmetry.

[4] Gaskell, G. 2003. Modelling regressive and progressive effects of assimilation in speech perception. *J. Phon.* 31 447-463.

[5] Gaskell, G., Marslen-Wilson, W. D. 1996. Phonological variation and inference in lexical access. *J. Exp. Psych., HPP* 22, 144-158.

[6] Gow, D. W. 2002. Does English coronal place assimilation create lexical ambiguity? *J. Exp. Psych., HPP* 28, 163-179.

[7] Holt, L. L. 2006. Speech categorization in context: Joint effects of nonspeech and speech precursors. *JASA* 119, 4016-4026.

[8] Lotto. A. J., Kluender, K. R. 1998. General contrast effects in speech perception: Effect of preceding liquid on stop consonant identification. *Percept. Psychophys.* 60, 602-619.

[9] Lotto A. J., Kluender, K. R., Holt, L. L. 1997. Perceptual compensation for coarticulation by Japanese Quail. *JASA* 102, 1134-1140.

[10] Mack, M. 1982. Voicing-dependent vowel duration in English and French: Monolingual and bilingual production. *JASA* 71, 173-178.

[11] Magnuson, J. S., McMurray, B., Tanenhaus, M. K., Aslin, R. N. 2003. Lexical effects on compensation for coarticulation. The ghost of Christmash past. *Cogn. Science* 27, 285-298.

[12] Mitterer, H., Blomert, L. 2003. Coping with phonological assimilation in speech perception: Evidence for early compensation. *Percept. Psychophys.* 65, 956-969.

[13] Repp, B. H. 1982. Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychol. Bull.* 92, 81-110.

[14] Pitt, M. A., McQueen, J. M. 1998. Is compensation for coarticulation mediated by the lexicon? *J. Mem. Lang.* 39, 347-370.

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