Perceptual processing of partially and fully assimilated words in French

Natalie D. Snoeren, Juan Seguí, Pierre Hallé

Perceptual processing of partially and fully assimilated words in French

Natalie D. Snoeren[a], Juan Segui[a], and Pierre A. Hallé[a,b]

[a]Laboratoire de Psychologie et Neurosciences Cognitives , Université René Descartes, CNRS-Paris 5

[b]Laboratoire Phonétique et Phonologie, CNRS-Paris 3

Running head: Perception of voice assimilation in French

Corresponding address:

Natalie D. Snoeren, now at the Department of Psychology,

University of York, Heslington

York, YO10 5DD, United Kingdom

E-mail address: n.snoeren@psych.york.ac.uk

Abstract

Models of speech perception attribute a different role to contextual information in the processing of assimilated speech. The present study examined perceptual processing of regressive voice assimilation in French. This phonological variation is asymmetric in that assimilation is partial for voiced stops and near-complete for voiceless stops. Two auditory-visual cross-modal form priming experiments were used to examine perceptual compensation for assimilation in French words with voiceless versus voiced stop offsets. The results show that, for the former segments, assimilating context enhances underlying form recovery, whereas it does not for the latter. These results suggest that two sources of information -- contextual information, and bottom-up information from the assimilated forms themselves -- are complementary and both come into play during the processing of fully or partially assimilated word forms.

Keywords: voice assimilation, compensation for assimilation, cross-modal form priming

Introduction

A fundamental characteristic of the speech signal is the variability of its phonetic realization. Nonetheless, the human perceptual system copes very well with this variability, and listeners can still access words from their mental lexicon in spite of possible deviations from their canonical pronunciation. This ability raises important challenges for our general understanding of spoken word recognition. The processing of small arbitrary deviations in the speech signal has often been of interest in priming studies (cf. Connine, Blasko, & Titone, 1993; Radeau, Morais, & Segui, 1995; Slowiaczek & Pisoni, 1987). In the present research, we focus on a systematic, regular type of variation, namely regressive voice assimilation in French. In contrast to arbitrary variations, regular variations are present in continuous speech and motivated by language-specific phonological rules. The study of these phenomena might help understand the underlying cognitive processes that allow a listener to recognize a variant form such as [grim] as the underlying form [grin] in the sequence *green beans*.

Over the last decade, a number of studies have addressed the processing implications of regular variations in the speech assimilation, most notably assimilation of place of articulation (cf. Otake, Yoneyama, Cutler & van der Lugt, 1996; Gaskell & Marslen-Wilson, 1996, 1998; Coenen, Zwitserlood & Bölte, 2001; Gow, 2001, 2002, 2003; Gow & Im, 2004; Weber, 2001, 2002; Mitterer & Blomert, 2003; Gumnior, Zwitserlood, & Bölte, 2005). Most of these studies suggested that the following context licensing phonological assimilation plays a major role in the perceptual processing of assimilated segments. Gaskell and Marslen-Wilson (1996) studied the perceptual processing of place assimilation in English, using an auditory-visual cross-modal priming paradigm. Primes were assimilated word forms (e.g., *leam*), canonical forms (e.g., *lean*), or unrelated forms. The magnitude of the priming effects was comparable for assimilated and canonical word forms: *leam* facilitated the processing of LEAN as much as did *lean*, when no right context was presented (Experiment 1) or when the right context licensed labial assimilation (Experiment 2). When the same assimilated word form *leam* was followed

by a contextually inappropriate, "unviable" context such as in *leam gammon* (where the labial place in *leam* is not contextually licensed), priming effects were no longer obtained. This suggests that phonologically lawful variants of word forms do not disrupt lexical access, as long as they occur in phonological contexts that license the change in surface form. The role of phonological context in the perceptual process of assimilated word segments led Gaskell and Marslen-Wilson to interpret these results in terms of a regressive inference mechanism. This mechanism would basically undo the language-specific assimilation rules that apply in production. Listeners would use the context following assimilated segments in order to recover their underlying identity. However, in these form priming experiments, the support for the role of regressive inference in recovering assimilated word forms comes from the negative evidence that phonologically inappropriate contexts are detrimental to lexical activation, not from positive evidence that appropriate contexts help.

More direct support for the benefit of a regressive inference mechanism comes from a phoneme monitoring study reported by the same authors (Gaskell & Marslen-Wilson, 1998). In one experiment, listeners monitored for word-final coronal segments in connected speech. The critical items contained segments that were underlyingly coronal but deliberately pronounced as noncoronals in contextually appropriate versus inappropriate environments. The authors found that listeners hearing *freight* pronounced [freɪp] in the phrase *freight bearer* showed a strong tendency to report hearing a word-final /t/. Gaskell and Marslen-Wilson argued that listeners apply phonological inference *prelexically* to determine that [p] in [freɪp] is an underlying /t/ whose surface form has assimilated to [p] in the context of [b].

Coenen, Zwitserlood, and Bölte (2001) studied both progressive (voice) and regressive (place) assimilation in German, also using cross-modal form priming. Contrary to Gaskell & Marslen-Wilson (1996), they found no priming effect for assimilated words presented in isolation, and graded priming effects for words in context: priming effects were larger for

unassimilated than for assimilated words (e.g., *wort mal* vs. *worp mal*). Gumnior et al. (2005) also reported an advantage of canonical over place-assimilated forms within German compounds. In agreement with Gaskell & Marslen-Wilson (1996; 1998), Coenen et al. did not obtain priming effects in unviable contexts (e.g., *worp kurz*). Their results thus also point to a crucial role of phonological context in the processing of assimilated words. Likewise, Mitterer and Blomert (2003) also showed that right context is used to recover viable but not unviable assimilated word forms (e.g., "tuin" from *tuimbank*, 'garden bench', vs. *tuimstoel* 'garden chair'). ERP data for passive listening revealed that viable but not unviable phonological changes elicited early additional activity (similar to mismatch negativity), presumably related to regressive inference. This would rule out the possibility that recovery from viable assimilation be attributable to attentional and/or decisional processing levels. As Gaskell and Marslen-Wilson (1998) proposed, the underlying process could be rather automatic.

Taken together, the studies mentioned so far suggest that the right context helps listeners to recover words with regressively assimilated speech segments. In these studies, however, assimilation was typically categorical, that is, complete. For example, in Gaskell & Marslen-Wilson's (1996) study, *lean* in "lean bacon" was deliberately pronounced with either [n] or [m]. In natural utterances, place assimilation in languages such as English might not always be complete (Gow & Hussami, 1999; Nolan, 1992). According to Gow (2002), partial assimilation would actually be the rule in natural speech. Our own data (Snoeren, Hallé, & Segui, 2006) suggest that regressive voice assimilation in French is not always complete (also see Kuzla, 2003 [German]; Warner, Jongman, Sereno, & Kemps, 2004 [Dutch]; Wright & Kerswill, 1982 [English]; Jansen & Toft, 2002 [Hungarian]). Partially assimilated segments may be viewed as ambiguous between two phonemic categories. Another approach is to consider that assimilated forms retain acoustic or articulatory cues to both the assimilated and the assimilating segment (Gow, 2002) so that listeners could conceivably exploit two different sources of information: the current information in the assimilated form itself and the upcoming

information in the assimilating context. Logically, then, listeners could in particular use the remaining cues to the underlying form of a partially assimilated segment to recover that form. In this situation, the role of the context information would conceivably be less crucial than when segments are completely assimilated and retain no trace of their underlying value. In other words, context information may be weighted differently according to whether assimilated forms are partially or fully assimilated. Listeners might rely on right context phonemic information when assimilation is complete because bottom-up information does not allow a full recovery of the assimilated segment's underlying identity. When traces of the underlying identity are available, bottom-up information might help to recover this identity and the role of context information could be minimized.

In incomplete assimilation situations, the assimilated segment also contains acoustic cues to its assimilating context. This allows at least partial anticipation of the following context. Indeed, Gow (2001, 2003), using partially place-assimilated forms such as *tem* in *ten buns*, demonstrated that the labial cues in *tem* facilitate the detection of the following /b/. Similar findings have been reported in Japanese for the assimilated moraic /N/ (Otake, et al., 1996; also see Lahiri & Marslen-Wilson, 1991; Quené, van Rossum & van Wijck, 1998). In contrast, fully assimilated forms such as [freɪp] in freight bearer do not enhance the detection of /b/ in *bearer* (Gaskell & Marslen-Wilson, 1998). Progressive contextual effects, just like regressive contextual effects, thus also seem to depend on the complete versus incomplete nature of the assimilation process. To sum up, according to the nature of assimilation, complete with deliberate full-feature change as in Gaskell and Marslen-Wilson (1996, 1998) and other studies, or incomplete –and perhaps more representative of natural speech assimilations– as in the work of Gow (2001, 2002, 2003; Gow & Im, 2004), the relative weights of the two sources of information –current form and upcoming context– called on to either recover the underlying form of assimilated segments or anticipate the upcoming segment, may be tuned differently.

Alternatively, regardless of the complete versus incomplete nature of assimilation, the processing system may "blindly" rely on a fixed combination of the sources of information to recover underlying forms.

The present study asks whether different proportions of the two sources of information considered here are involved according to the nature of assimilation. On one extreme, bottom-up information from the current word form could be sufficient to recover its underlying form in the case of partial assimilation, whereas, on the opposite extreme, information from the upcoming context only could be used to the same effect in the case of complete assimilation. The latter scenario may be termed "regressive inference". We propose that the two sources of information are complementary and both come into play during the processing of assimilated forms. In the absence of acoustic traces of the underlying segment in completely assimilated speech segments, listeners can only rely on the following context to derive their underlying identity, whereas in the presence of acoustic traces in partially or weakly assimilated segments, listeners can rely on this information to access their underlying forms with a lesser role of context. To test for this prediction, we compared two situations of natural regressive voice assimilation in French. One is devoicing of underlyingly voiced segments, as in *coude plié* ('bent elbow'). The other is voicing of underlyingly voiceless segments, as in *note grave* ('low tone'). These two situations are not symmetrical as one could expect. Our previous study (Snoeren et al., 2006) indeed established that voice assimilation is generally incomplete in the former situation and almost complete in the latter one. This finding was substantiated by both perceptual and acoustic data. In naturally produced voice assimilations, as in *coude plié* and *note grave*, the word-final consonant was perceived mainly /d/ in *note*, whereas it was perceived slightly less often /t/ than /d/ in *coude*. Acoustically, the word-final consonant was assimilated to a "lesser degree" in *coude* than in *note*. (We proposed a straightforward measure of assimilation degree based on the proportion of voicing within stop closure.) Importantly, the correlation between perceived and measured voicing was quite high, which makes the observed

asymmetry all the more reliable. Therefore, regressive voice assimilation in French *naturally* provides a nearly ideal contrast to test for the prediction stated above. Underlyingly voiceless segments are prone to complete voicing while underlyingly voiced segments only lead to partial devoicing. We therefore predict that context will be used to a larger extent in the former than in the latter situation.

To test for this prediction, we used the auditory-visual form priming paradigm, as in the previous studies of Gaskell & Marslen-Wilson (1996) and Gow (2001, 2002, 2003), to measure the priming effect of assimilated speech forms on visual targets. The cross-modal priming paradigm is sensitive to lexical rather than prelexical speech properties (Marslen-Wilson, Tyler, Waksler, & Older, 1994; Marslen-Wilson, Moss, & van Halen, , 1996; also see Spinelli & Gros-Balthazard, in press). Intra-modal priming (e.g., auditory-auditory) rather reveals prelexical relationships such as rhyming relationships (cf. Norris, McQueen, & Cutler, 2002; Radeau, Morais, & Segui, 1995; also see Utman, Blumstein, & Burton, 2000). Priming effects in auditory-visual cross-modal form priming rather are symptomatic of lexical pre-activation by the primes –not just phonetic or phonemic similarities between primes and targets– and are thus potentially sensitive to lexical activation mechanisms other than bottom-up, such as regressive inference mechanisms. This is an important motivation for using cross-modal priming in our study, whose goal is to assess the relative roles of bottom-up and regressive inference information in recovering underlying word forms according to degree of assimilation.

Throughout the present study, the auditory primes consisted of short noun phrases (article + noun + adjective) such as *une note grave*, in which the adjective's initial consonant licensed voice assimilation of the noun's final consonant. The visual target (NOTE in this example) was presented at the offset of the noun. In Experiment 1, the primes were presented without the adjective, that is, without the assimilating context (e.g., *une note* in the example above). In Experiment 2, the entire primes were presented (e.g., *une note grave*). This design,

similar to that used in Gaskell and Marslen-Wilson (1996), allowed us to examine the role of assimilating context in the processing of assimilated word forms. We begin by describing the materials which were the same in the two experiments.

Selection of speech materials

Initial stimulus set. Thirty-six monosyllabic noun words ending with a voiced stop consonant and 36 with a voiceless consonant were first selected. The two sets of words were matched at best in terms of frequency of occurrence and lexical competition.[1] There were 12 words for each of the six stops /p, t, k, b, d, g/. For all words, swapping word-final voicing did not produce another word (e.g., changing /p/ into /b/ in *note*, 'note, tone', produced [nɔd], which is not a French word). Hence, none of these words was potentially ambiguous under a change of voicing of the final consonant. Each noun word was inserted in two article+noun+adjective noun phrases: one in which the right context licensed voice assimilation, and the other not (e.g., *note* was inserted in *une note grave* and in *une note salée*). Three native speakers of French judged that all the constructed noun phrases were semantically plausible. The adjective's initial consonant always had a place of articulation different from that of the preceding noun's final consonant so as to avoid possible gemination (as could occur in *note tenue* [nɔtːəny] or *note douce* [nɔdːus]). These 144 noun phrases (72 nouns x 2 contexts) are listed in the Appendix. They were recorded, together with a pool of filler speech materials (also noun phrases) to be used in the main experiments, by a male native speaker of French from the Paris region and directly stored to computer files (20 kHz sampling rate, 16 bit precision). The speaker was instructed to produce fluent speech without pauses between words. Each noun phrase was recorded three times, and the best token with respect to fluency and naturalness, chosen by the first author, was retained.

Selected set. From the initial set, we proceeded to select a set of items showing the asymmetric pattern of assimilation (stronger degree of assimilation for voiceless than voiced stops), which we planned to exploit to test for the possibly differential role of assimilatory context according to degree of assimilation: ideally, full versus partial assimilation. A perception pretest was run on the 72 noun phrases with an assimilatory context[2] to determine how much assimilated each noun was perceived by French listeners, based on their categorization of the noun's final stop as voiced or voiceless. We expected that most of the speech items fit in the asymmetric pattern of assimilation found by Snoeren et al. (2006). The 72 phrases were presented auditorily without the assimilatory context (e.g., *une note grave* up to *note*) to avoid biasing participants' judgments. For this purpose, the adjective was excised from each noun phrase; the cut-off point in the speech wave was always the end of the release burst of the noun-final stop (at the nearest zero crossing to avoid audible click); the release burst was located from visual inspection of the spectrogram; finally, care was taken to equalize peak acoustic intensity across the stimuli. Twenty undergraduate students at Paris 5 René Descartes University participated in the pretest. All of them were native speakers of French and none of them reported any hearing problem. The pretest consisted of a test phase preceded by a training phase. In the test phase, participants received the 72 truncated phrases in a randomized order and were asked to categorize each utterance-final consonant by choosing one of two alternative responses (e.g., 'd' or 't' for *une note*), then to indicate how well they thought their choice matched the presented item, using a 1-5 scale in which 1 = "poor match" and 5 = "excellent match." Participants were warned that they would be presented with either words or nonwords and had to ignore the lexical status of what they heard: they just had to focus on the final consonant of each item, and choose the more appropriate phonemic label proposed to them. In the training phase, participants received 12 nonwords ending with a stop. This was intended to discourage participants to use lexical knowledge to categorize utterance-final consonants. Underlyingly voiceless stops (as in *note*) produced an average 85% of "voiced" responses, whereas underlyingly voiced stops (as in

coude) produced an average 59% of "voiceless" responses. The mean ratings were 3.8 and 3.6 for voiceless and voiced stops, respectively, indicating that participants were fairly confident in their responses. The results thus suggest that, overall, voiceless stops were perceived as voice-assimilated to a larger extent than voiced stops, replicating the asymmetric assimilation pattern reported in Snoeren et al. (2006). However, three words with a voiceless final stop (*coupe*, *jupe*, *lampe*) and three with a voiced stop (*fougue*, *stade*, *robe*) ran opposite to the dominant assimilation profile: the former ones only received an average 27% of "voiced" responses and the latter almost 100% of "voiceless" responses. These six items were thus excluded from the final set. After this exclusion, the 33 remaining items with an underlyingly voiceless stop can be considered as completely or near-completely voice-assimilated (they received an average 90% of "voiced" judgments), whereas the 33 items with an underlyingly voiced stop can be considered as incompletely voice-assimilated (they received an average 45% of "voiced" judgments).[3] The high rate of "voiced" judgments for the items with an underlying voiceless stop suggests that participants responses showed little lexical bias. Moreover, the percentage of voiced occlusion measured in the assimilated stops (see Snoeren et al., 2006) paralleled the perceptual measures: 96% and 58% in average for voiceless and voiced stops, respectively.

Experiment 1

We first examined the priming effect of the nouns of the selected set, in their assimilated and non-assimilated versions, presented in the original noun phrases in which they were produced but with the right context removed. For example, *une note* from *une note grave* (assimilated version) and *une note* from *une note salée* (canonical version), were presented as auditory primes to the visual target NOTE, thus following the basic design of Gaskell and Marslen-Wilson's (1996) Experiment 1. The issues addressed are of whether assimilated and canonical forms produce comparable priming effects, and whether degree of assimilation modulates the size of priming effects.

Method

Participants. Sixty-one undergraduates students at the Psychology Department of Paris 5 René Descartes University, native speakers of French, participated in the experiment (4 male and 57 female students, mean age = 23 years, range 18-47 years), took part in Experiment 1. None of them reported hearing or vision problems. None of them had participated in the pretest. Each participant filled in a language background questionnaire before the experiment was run.

Design and Materials. The printed forms (in uppercase) of the 66 words in the selected set were used as visual targets. The primes were either unrelated to the target (e.g., *un acte*—NOTE) or form-related (e.g., *une note*—NOTE), with the critical noun in its non-assimilated, "canonical" form or in its assimilated form. There were thus three types of priming, which we label "canonical," "assimilated", and "unrelated", hence a total of 198 (66 x 3) test trials. Three lists of 66 test trials were constructed in counterbalancing the three types of priming so that the subjects assigned to a given list saw all 66 test targets only once and received all three trial types. Another 150 filler trials were constructed, 108 of which had a nonword target and the remaining 42 a word target. Each subject hence received an equal number of trials with a word and a nonword target. The primes in the filler trials were always noun phrases. Amongst the 108 trials with a nonword target, 72 had a noun prime phonologically related to the target (e.g., *bière* /bjɛr/ 'beer' for BIEVE, a nonword whose plausible pronunciation is /bjɛv/) and 36 had a phonologically unrelated noun prime (e.g., *nymphe* /nɛ̃f/ 'nymph' for REUX, /rø/). The form-related filler trials with a nonword target were included to discourage participants from associating phonological relatedness, present in two thirds of the test trials, with a "word" response (see Lukatela, Eaton, Sabadini, & Turvey, 2004). In addition to the test and fillers trials, 10 similar practice trials and two warm-up trials were constructed.

Procedure. We followed the standard auditory-visual cross-modal priming lexical decision procedure (cf. Grosjean & Frauenfelder, 1996): visual targets were presented on a computer screen at the acoustic offset of the prime in the auditory stimulus and remained displayed until the subject's response with a three seconds time-out. (Responses entered outside this time window were counted as misses.) The time location of each prime offset was determined by visual inspection of its spectrogram as the end of the release burst of the final stop of the noun. Participants were instructed to respond on the visual target in each trial as quickly and accurately as possible, by pressing a "yes" button or a "no" button for positive or negative lexical decision, respectively. The "yes" button was assigned to the participants' better skilled hand. Participants were informed that they were to receive a recall test after they completed the main test. The recall test was intended to incite participants to attend to the auditory stimuli. Participants were tested individually in a dimly lightened, quiet room. The auditory stimuli containing the prime were presented via headphones at a comfortable listening level. Targets were displayed using 14 point Arial font in black on a white background, centered on the computer screen. The buttons of a Logitech Wingman gamepad were used to enter responses, ensuring a 1 ms precision for response times. The experiment was run on a PC-compatible micro-computer using the DMDX software (Forster & Forster, 2003). The experiment begun with a 10-trials training phase; participants did not receive feedback on their responses during training phase but were welcome to ask for clarification explanations after they had completed training. This was followed by the test phase, which begun with two warming-up trials for which responses were not recorded. Participants were allowed to pause midway during the test phase. The second half of the test again begun with two warming-up trials. After the test phase was completed, participants received, as announced, a "recall test." They received a recognition sheet containing 30 words, 15 of which occurred as visual targets in the previous

test phase. Participants were instructed to circle in the list the words that seemed familiar to them. The total duration of the experiment was about 30 minutes.

Results

The data files for three participants were not retained, due to high error rates (> 20%) and long mean response times (> 850 ms). For the 58 retained participants, RTs longer than 1200 ms (1.4 %) were not included in the RT analyses. After these exclusions, the mean response times were 502 ms for the canonical condition, 555 ms for the assimilated condition, and 595 ms for the unrelated condition. The RT and error data (%) are shown in Table 1.

Table 1 about here

Response times. Two-way analyses of variance were conducted, by subject (F_1) and by item (F_2), with Priming type (canonical, assimilated, and unrelated), and target Voicing (voiceless vs. voiced word offset) as main factors.[4] The effect of Priming type was highly significant, $F_1(2, 114) = 102.97$, $p < .0001$; $F_2(2, 128) = 98.63$, $p < .0001$. The effect of target Voicing was significant, $F_1(1,57) = 40.42$, $p < .0001$; $F_2(1,64) = 4.67$, $p < .05$: voiceless targets (e.g., NOTE) were responded faster overall than voiced ones (e.g., COUDE). The interaction between Priming and Voicing was not significant, $F_1(2,114) = 1.94$; $F_2(2,128) = 1.18$, both ps >.15.

Paired comparisons showed that RTs were faster for canonical than assimilated primes and for assimilated than unrelated primes, for either voiced or voiceless targets (e.g., COUDE or NOTE), at least at the p<.0005 level.

Error rates. The error data largely reflected the RT data. The effect of Priming was significant, $F_1(2, 114) = 6.83$, $p < .001$; $F_2(2, 128) = 10.04$, $p < .001$. The effect of Voicing was significant by subject, $F_1(1,57) = 15.28$, $p < .001$, and marginally significant by item, $F_2(1, 64) = 4.38$, $p =$

0.056: there were less errors for voiceless than voiced targets. Again, the interaction between these two factors was not significant, $F_1(2, 114) = 1.54$; $F_2(2, 128) = 1.35$, both *p*s> .2.

Discussion

Experiment 1 indicated that unassimilated and assimilated primes give rise to different priming patterns. Priming effects were larger for unassimilated (canonical) than assimilated forms, and were equivalent for underlyingly voiceless and voiced words, suggesting that, in the absence of context, fully and partially assimilated forms activate underlying forms to the same extent. These results differ from those obtained by Gaskell and Marslen-Wilson (1996), who did not find any difference in priming effects between canonical and (fully) assimilated conditions. However, in their study, each sentence containing the critical auditory prime was preceded by a semantically biasing sentence. For instance, the sentence *We have a house full of fussy eaters* preceded the critical sentence *Sandra will only eat* ***lean*** bacon. In this situation, the predictability of the prime may very well have increased participants' tolerance for mismatch. In contrast, we exclusively used simple noun phrases in which the noun was never predictable. The clear advantage we found for canonical over assimilated forms in terms of priming efficacy may be due to the absence of predictability for the critical stimuli. Our results also differ from those reported by Coenen et al. (2001), who found no priming *at all* for (fully) assimilated prime forms, although they used materials similar to Gaskell and Marslen-Wilson's (1996), consisting of an introductory sentence followed by a critical sentence, in which the prime word was embedded. Thus, our results are intermediate between the dramatically opposed patterns in Coenen et al. (2001) and Gaskell and Marslen-Wilson (1996) studies.

Our data and those of Coenen et al. (2001) agree in that they do not seem to support the "underspecified representations" account of tolerance for assimilated forms, proposed first in Lahiri and Marslen-Wilson (1991), and later elaborated in the "featurally underspecified lexicon" (FUL) model (Lahiri & Reetz, 2000). FUL assumes that coronality of the offset

consonant is not specified in English words such as *lean* or in German words such as *Wort*, hence that place assimilated and unassimilated forms equally match a lexical representation in which coronal place is not specified. Likewise, FUL could assume that voicing is unspecified in the offset stop of French words such as *note* (or, alternatively such as *coude*, were the unmarked case voiced instead of voiceless), hence predict that the voiced and voiceless surface forms [nɔd] and [nɔt] equally match the lexical representation of *note*. This prediction is not borne out by either the German data in Coenen at al. (2001) or our French data, whereas it is congruent with the English data in Gaskell and Marslen-Wilson (1996). For the French [voice] feature, however, "unviable" context assimilations such as [nɔd#sale] for *note salée* ('long bill') or [kut#blese] for *coude blessé* ('wounded elbow') have not been tested yet, but context viability should not play a major role in FUL, other than disambiguate ambiguous forms (e.g., between *right* and *ripe*) with the help of higher level constraints.

If the assimilating context helps to recover the underlying form of assimilated words, we should find that its presence enhances the priming efficacy of assimilated primes, especially, perhaps, completely assimilated forms. We address this issue in Experiment 2, in which the entire noun phrases are presented. The comparison between the results obtained in the absence of context (Expt. 1) and those obtained in the presence of context (Expt. 2) may allow us to evaluate the role of context for fully and partially assimilated forms.

Experiment 2

Experiment 2 was identical to Experiment 1 in all respects except that the noun phrases were presented entirely instead of truncated after the noun prime (e.g., *une note grave* instead of *une note* for the target NOTE).

Method

Participants. Sixty- two undergraduate students at the Psychology Department of Paris 5 René Descartes University, native speakers of French, participated in the experiment (mean age 23 years, range 19-52 years, 11 male and 51 female). None of them reported hearing of vision problem. None of them had participated in the pretest or in Experiment 1.

Design, Materials, and Procedure. The only difference with Experiment 1 was that the auditory noun phrase were not truncated, that is, included the right context of the noun, assimilatory or not. As in Experiment 1, visual targets were presented at the acoustic offset of the noun for each trial.

Results

The data for four participants were not retained, due to long mean response times (> 800 ms). For the 58 participants retained, response times longer than 1200 ms (0.48%) were excluded from the RT analyses. After these exclusions, the mean response times were 508 ms for the canonical condition, 539 ms for the assimilated condition, and 591 ms for the unrelated condition. The RT and error data are shown in Table 2.

Table 2 about here

Response times. As in Experiment 1, two-way analyses of variance were conducted by subject and by item, with Priming type (canonical, assimilated, and unrelated) and target Voicing (voiceless vs. voiced word offset) as main factors.

The effect of Priming was highly significant, $F_1(2,114) = 74.16$, $p < .0001$; $F_2(2,128) = 92.92$, $p < .0001$. Voicing was significant too, $F_1(1,57) = 25,42$, $p < .0001$, $F_2(1,64) = 4.47$, $p < .05$. The interaction between these two factors was significant by subject $F_1(2,114) = 5,37$, p, $< .01$, but not by item, $F_2(2,128) = 1.59$, $p = 0.21$. The interaction reflects the fact that the

magnitude of the priming effect differs as a function of Voicing. Indeed, as can be seen from Table 2, fully assimilated voiceless primes gave rise to a priming effect of 67 ms, whereas the priming effect was only 36 ms for partially assimilated voiced items. These results contrast with those observed in Experiment 1, in which both types of assimilated primes gave rise to comparable priming effects and no interaction was observed between Voicing and Priming.

Paired comparisons showed, that RTs were faster for canonical than assimilated primes and for assimilated than unrelated primes, as in Experiment 1. All the comparisons were significant at least at the $p<.0005$ level, except for the canonical versus assimilated comparison for voiceless targets ($t_1(57) = 3.21$, $p = .0022$; $t_2(32) = 2.93$, $p = .0062$).

Error rates. The error data largely reflected the RT data. The effect of Priming was significant, both *p*s <.001. That of Voicing was significant as well, $F_1(1, 57) = 29.28$, $p <.0001$; $F_2(1, 64) = 5.99$, $p <.05$: there were less errors for voiceless than voiced targets.[5] As in the RT data, the interaction between Priming and Voicing was significant by subject, $F_1(2, 114) = 3.77$, $p <.05$, but did not reach significance by item, $F_2(2, 128) = 1.93$, $p = .15$.

Combined analyses Experiments 1 and 2. A combined analysis of Experiments 1 and 2 was performed to assess more precisely the role of context in the perceptual processing of voiceless and voiced items. To this end, the results of Experiments 1 and 2 corresponding to assimilated word primes, were combined. A two-way ANOVA was conducted, with Context (absence in Expt. 1 vs. presence in Expt. 2) and target Voicing as the main factors. This analysis revealed a significant effect of Voicing, $F_1(1, 114) = 26.56$, $p <.0001$; $F_2(1, 64) = 4.34$, $p <.05$: RTs to voiceless targets were faster than to voiced ones. The effect of Context was not significant in the subjects analysis, $F_1(1, 114) = 1.22$, $p = .27$, whereas it was significant in the items analysis, $F_2(1, 64) = 10.07$, $p <.01$. Importantly, the interaction between these two factors was

significant by subject, $F_1(1, 114) = 10.25$, $p < .01$, and marginally significant by item, $F_2(1, 64) = 3.37$, $p = 0.07$.

For voiceless targets with fully assimilated primes, as in *note* [nɔd], Context did significantly affect RTs, $F_1(1, 114) = 4.37$; $p < .05$; $F_2(1, 64) = 9.89$, $p < .01$. For these items, RTs were faster in the presence than in the absence of context. For voiced targets with partially assimilated, primes, as in *coude* [ku$^{t}_{d}$], context did not affect RTs, both Fs < 1.

A similar analysis was conducted for the results obtained in Experiments 1 and 2 with canonical primes. This analysis indicated a main effect of Voicing, $F_1(1,114) = 25.83$, $p < .0001$; $F_2(1,64) = 5.28$, $p < .05$. No effect was obtained for the Context factor, both Fs <1. The interaction between Context and Voicing was not significant, $F_1(1,114) = 1.53$; $F_2(1,64) = 0.26$, both ps >.2.

Figure 1 illustrates the priming effects for voiced and voiceless targets in the assimilated and canonical conditions, according to the presence or absence of the right context. As can be seen from this figure, priming effects for fully assimilated (underlyingly voiceless) primes increased dramatically with the presence of context, whereas priming effects for partially assimilated (underlyingly voiced) primes were virtually not affected by the presence of context. Not surprisingly, priming effects for targets that follow canonical primes were unaffected by the presence of the context.

Figure 1 about here

Discussion

Experiment 2 indicates that, in the presence of assimilating context, priming effects are greater for voiceless than for voiced offset assimilated primes, that is, for fully than for partially assimilated primes, whereas no difference was found in Experiment 1, in which

context was not presented. In other words, assimilating context helps to recover assimilated words such as note pronounced [nɔd] but not words such as coude pronounced [ku$^{t}_{d}$].

A possible explanation for this difference between note and coude nouns could be that, in the case of assimilatory context, noun phrases such as *note grave* are more likely than noun phrases such as *coude plié*. However, co-occurrence counts of the involved noun-adjective pairs rather indicate the opposite trend.[6] Hence, the difference between *note* pronounced [nɔd] and *coude* pronounced [ku$^{t}_{d}$] cannot be due to differential lexical co-occurrence frequencies.

We might therefore conclude that the presence of assimilating context benefits to completely but not partially assimilated speech. This facilitatory effect could be explained in terms of an on-line phonological inference mechanism, which is called for when physical word forms markedly differ from canonical forms, that is, in the case of complete or near-complete assimilation, but not when physical word forms retain some cues of the canonical forms.

Across the two experiments, the priming effects obtained –less priming for assimilated than "canonical," unassimilated forms– show that assimilated speech has a processing cost compared to canonical, unassimilated speech. This is in line with Gumnior et al.'s (2005) finding that priming effects are greater for canonical than for assimilated forms in the presence of assimilating context, using German compounds such as *Bahngleis* /ba:nglais/ with unassimilated or assimilated /n/ ([n] or [ŋ]).

General Discussion

The purpose of the present research was to study the perceptual consequences of regressive voice assimilation in French. We examined in particular whether clear-cut differences in degree of assimilation entail differences in the role of contextual information. Voice assimilation in French allowed us to examine the impact of such differences, because it naturally provides two clearly contrasted cases of voice assimilation: voiceless stops are

strongly assimilated in a voiced environment, whereas voiced stops are incompletely assimilated in a voiceless environment.

In Experiment 1, using an auditory-visual cross-modal form priming paradigm, we found that the unassimilated –"canonical"- forms of word primes such as *note* or *coude* presented *without context*, strongly primed their printed counterpart by about 93 ms, whereas the assimilated forms had a significantly lesser priming effect of about 40 ms. Although the voiced final stops as in *coude* were only half devoiced in assimilated forms and the voiceless stops as in *note* almost completely voiced, both types of assimilated forms produced analogous, significant priming effects. In Experiment 2, right context was made available to listeners. The overall advantage in priming effect for unassimilated over assimilated forms still obtained. However, whereas the priming effect for assimilated voiceless-stop words such as *note* was significantly increased by the presence of assimilating context, that for voiced-stop words was not. This clear-cut difference was assessed by a combined statistical analysis of Experiments 1 and 2. To sum up, the presence of the assimilating context seems to help to process strongly voice-assimilated word forms, such as *note* pronounced [nɔd], whereas it does not help for partially voice-assimilated forms, such as *coude* approximating [kut] but retaining traces of voicedness.

The robust priming differences obtained in Experiment 1 between canonical and assimilated items presented without context contrasts with the absence of difference observed by Gaskell and Marslen-Wilson (1996). In their study, however, the carrier sentence with the critical prime item was somewhat predictable in that it was preceded by a semantically biasing sentence. This feature may very well have increased participants' tolerance for mismatch. In our Experiment 1, we exclusively used article+noun noun phrases, in which the nouns were in no way predictable. Another possible explanation of these divergent results pertains to the fact that voice assimilation is different in its acoustic implementation from place assimilation. Gow

and Im (2004) remark that "voicing cues inherently play out over a longer interval than place cues" (Gow & Im, 2004: 286). This difference may have important perceptual consequences so that a comparison between voice assimilation in French and place assimilation in Germanic languages such as English is unwarranted, although both types of regressive assimilation belong to the same class of phonological alternation processes.

The results of Experiment 1 showed an analogous priming pattern for assimilated forms of voiceless-stop words such as *note* and voiced-stop words such as *coude*. If the magnitude of the priming effect was to reflect form-closeness to canonical forms, assimilated voiced-stop word forms (e.g., *coude*) should induce greater priming than voiceless-stop word forms (e.g., *note*) because the latter are more strongly assimilated, hence depart more markedly from canonical form. However, we did not find a significant difference between the priming effects produced by the two types of primes. Priming efficacy thus is not determined by prime form-similarity to canonical form.

By comparison with the results obtained for the assimilated primes in Experiment 1, the presence of the right context in Experiment 2 clearly enhanced the priming effect of voiceless-stop items, but not that of the voiced-stop ones. This suggests that the role of the right context in the perception of assimilated speech depends on the extent to which segments are assimilated. In earlier studies, the role of context has been assessed by comparing contextually viable with unviable assimilation (e.g., Gaskell & Marslen-Wilson, 1996; Coenen, Zwitserlood, & Bölte, 2001; Mitterer & Blomert, 2003). These studies only reported negative evidence for regressive contextual effects, showing that, for example, an inappropriate combination of labial assimilation and velar context blocked the recovery of underlying coronal place, as in *leam gammon*. In the present study, we focused on the positive evidence for the role of postassimilation context in viable assimilations. Our results suggest that postassimilation context enhances the priming efficacy of near-completely assimilated word forms (in line with the findings of Coenen et al., 2001), but not that of partially assimilated

word forms. The data thus support our initial prediction of quantitative differences in the role of assimilating context according to degree of assimilation. In the case of strongly assimilated forms, we tentatively interpret the substantial role of context as attributable to a phonological inference mechanism. In the case of partially assimilated forms, in which no regressive contextual effect is observed, we assume that cues to underlying voicing, still present in the acoustic signal, are sufficient to restore the intended word. How does this pattern fit with a "regressive inference" account? On the "activation" metaphor, which is widely used in the context of priming effects, the greater priming efficacy obtained for fully than partially assimilated primes in Experiment 2, where the assimilating context is present, suggests that an intended word is more strongly activated by a fully than partially assimilated auditory word form. Such differential level of activation clearly does not parallel closeness to canonical word form. It can only be explained if we assume that activation is solely determined by bottom-up evidence in the case of partially assimilated word forms, but results from a (full) "restoration" mechanism in the case of fully assimilated word forms. Restoration in the latter case simply means that when bottom-up evidence is insufficient for "immediate" integration at the lexical level, lexical resolution is achieved with the additional integration of the upcoming acoustic information. This type of mechanism is called "delayed commitment" in the general context of word recognition (see Mattys, 1997, for a review). In the present case, we call it "regressive inference," in the sense of a restoration mechanism that compensates for assimilation and eventually produces a stronger activation than the direct, bottom-up integration of partially assimilated word forms.

We stated that context helps to recover from strong assimilation, not from partial assimilation. Yet, in our design, the "assimilation strength" factor was intentionally confounded with underlying voicing because we wished to capitalize on a natural asymmetry in French voice assimilation. A complete demonstration of the "assimilation strength" account could be provided by the opposite situation of fully assimilated voiced stops compared to

partially assimilated voiceless stops (e.g., *coude* [kut] vs. *note* [nɔ$^{t}_{d}$]), however unnatural these assimilations may be. We therefore cannot already conclude that the presence of right context helps to recover completely but not incompletely assimilated forms. The important point we make, however, is that two sources of information in speech utterances that undergo assimilation are exploited in combination. One is strictly bottom-up and independent from context. It seems to apply to weakly assimilated forms (or for some reason, to voiced-stop words such as *coude*), presumably drawing on the traces of original voicedness that remain after incomplete assimilation. The other one is contextual and seems to apply to strongly assimilated forms (or for some reason, to voiceless-stop words such as *note*). We proposed that the active role of context information be attributable to a regressive inference mechanism such as the one posited by Gaskell and Marslen-Wilson (1996). But is there an alternative account of the role of assimilating context?

Gow's recent research (2001, 2002, 2003; Gow & Im, 2004) suggests that both regressive and progressive contextual effects observed in assimilation situations can be explained by a universal mechanism of feature cue parsing, whereby not only assimilating context helps to disambiguate partially assimilated segments, but partially assimilated segments also facilitate processing upcoming context. In essence, the *feature parsing model* elegantly accounts for how the temporally dispersed acoustic features that are present in the speech signal are optimally assigned to speech segments. If *right* in *right berries* is partially assimilated, it contains acoustic cues to both coronal and labial place: in standard phonological description, the privative (single-valued) features [coronal] and [labial] both are present. In *right berries*, the strong evidence for labial place in *berries* would attract away the weaker evidence for labial place in *right*, "leaving only evidence for coronal place" (Gow & Im, 2004: 282). In the absence of the labial context *berries*, the assimilated form of *right* would remain ambiguous between [raɪt] and [raɪp], *ripe* (cf. Gow, 2002, Experiment 4). In the phrase *ripe*

berries, [raɪp] contains no cues to coronal place and is not discernable from a fully labial-assimilated form or *right*. Here the feature parsing mechanism cannot restore a putatively intended *right*: context does not help. Thus, the feature parsing account predicts that partially assimilated forms are more likely to be restored than fully assimilated ones. This prediction does not seem to apply to our results, which showed the opposite pattern. However, it should be noted that we used word forms that could not be lexically ambiguous (e.g., /nɔd/ is not a French word). The role of context may be limited in that case. Also, as Gow and Im (2004: 293) note, listeners "also engage in top-down schema-driven grouping processes." The schema can refer to stored lexical representations and word form recovery in our data could be simply lexically driven. If such was the case, however, *note* [nɔd] in *note grave* should not induce stronger priming effects than *coude* [ku$^{t}_{d}$] in *coude plié*.

The present study's data seem, at least superficially, in agreement with the predictions of Gaskell's (2003) recurrent network model, an extension of a previous model by Gaskell, Hare, and Marslen-Wilson (1995), which only treated complete place assimilation. Gaskell's (2003) model integrates the possibility of partial place assimilation in languages such as English (i.e., place assimilation is restricted to underlying coronal place). The network uses three sets of output nodes, representing the current input segment, and the previous and upcoming segments. This architecture allows for evaluating progressive and regressive contextual effects. In Gaskell's (2003) model, intermediate degrees of assimilation are implemented by assigning complementary weights to, for example, coronal and labial features in the case of labial assimilation (e.g., 40% coronal and 60% labial). After (statistical) training, both regressive and progressive context effects obtain depending on assimilation strength. The model produces progressive, i.e. anticipatory effects for moderately assimilated segments (between 20% and 80% non-coronal). Stronger assimilation (80%-100% non coronal) does not produce progressive, anticipatory contextual effects but produces regressive effects that can

readily be interpreted as corresponding to regressive inference. Our results, which only address regressive context effects, exhibited the general pattern of a regressive context effect restricted to fully assimilated forms.

The present research provides, to our knowledge, the first empirical data supporting the hypothesis that the role of context is modulated by assimilation strength in the perceptual processing of assimilated speech. (Coenen et al., 2001, showed that fully assimilated forms require assimilating context to be recovered.) We have tried to show that in the processing of assimilated speech, two sources of information are exploited. They loosely correspond to two distinct mechanisms proposed in the literature. The perceptual consequences on the processing of assimilated speech are elegantly captured in Gaskell's (2003) model. It remains to be seen whether Gaskell's model can accommodate assimilation phenomena other than the English-specific place assimilation it was initially designed to model. Future cross-linguistic comparisons are crucial because they will allow us to dissociate language-specific from universal perceptual mechanisms, contributing to the current debate on the role of language-specific vs. universal processes of "compensation for assimilation" (Darcy, 2003; Gow & Im, 2004, Mitterer, Csépe, Honbolygo, & Blomert, 2006). It is hoped that future cross-linguistic modeling work as well as empirical work can shed some more light on these complex issues.

References

Coenen, E., Zwitserlood, P., & Bölte, J. (2001). Variation and assimilation in German: Consequences for lexical access and representation. *Language and Cognitive Processes*, *16*, 535-564.

Connine, C. M., Blasko, D. G., & Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language*, *32*, 193 210.

Darcy, I. (2003). *Assimilation phonologique et reconnaissance des mots* [Phonological assimilation and word recognition]. Unpublished Doctoral Thesis, EHESS, Paris.

Dufour, S., Peereman, R., Pallier, C., & Radeau, M. (2002). Vocolex: une base de données lexicales sur les similarités phonologiques entre les mots français. *L'Année Psychologique*, *102*, 725-746.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments and Computers*, *35*, 116-124.

Gaskell, M. G., Hare, M., & Marslen-Wilson, W. D. (1995). A connectionist model of phonological representation in speech perception. *Cognitive Science*, *19*, 407-439.

Gaskell, M. G., & Marslen-Wilson, W. D. (1996). Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 144-158.

Gaskell, M. G., & Marslen-Wilson, W. D. (1998). Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 380-396.

Gaskell, M. G., & Marslen-Wilson, W. D. (2001). Lexical ambiguity resolution and spoke word recognition: bridging the gap. *Journal of Memory and Language*, *44*, 325-349.

Gaskell, M. G., & Marslen-Wilson, W. D. (2002). Representation and competition in the perception of spoken words. *Cognitive Psychology*, *45*, 220-266.

Gaskell, M. G. (2003). Modelling regressive and progressive effects of assimilation in speech perception. *Journal of Phonetics*, *31*, 447-463.

Gow, D. W. & Hussami, P. (1999, November). *Acoustic modification in English place assimilation*. Paper presented at the meeting of the Acoustic Society of America, Columbus, OH.

Gow, D. W. (2001). Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language*, *45*, 133-159.

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

Gow, D. W. (2003). Feature parsing: Feature cue mapping in spoken word recognition. *Perception and Psychophysics*, *65*, 575-590.

Gow, D. W., & Im, A. M. (2004). A cross-linguistic examination of assimilation context effects. *Journal of Memory and Language*, *51*, 279-296.

Gumnior, H., Zwitserlood, P., & Bölte, J. (2005). Assimilation in existing and novel compounds. *Language and Cognitive Processes*, *20*, 465-488.

Grosjean, F., & Frauenfelder, U. (1996). A Guide to Spoken Word Recognition Paradigms: Introduction. *Language and Cognitive Processes*, *11*, 553-558.

Jansen, W. & Toft, Z. (2002). On sounds that like to be paired (after all): An acoustic investigation of Hungarian voicing assimilation. *SOAS Working Papers in Linguistics, 12*, 19-52.

Kuzla, C. (2003). Prosodically-conditioned variation in the realization of domain-final stops voicing assimilation of domain-initial fricatives in German. In M.J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 2829-2832). Barcelona, Spain.

Lahiri, A. & Marslen-Wilson, W.D. (1991). The mental representation of lexical form; A phonological approach to the recognition lexicon. *Cognition*, *38*, 245-294.

Lahiri, A., & Reetz, H. (2002). Underspecified recognition. In C. Gussenhoven & N. Warner (Eds.), *Papers in Laboratory Phonology 7* (pp. 637-675). Berlin: Mouton de Gruyter.

Lukatela, G., Eaton, T., Sabadini, L., & Turvey, M. T. (2004). Vowel duration affects visual word identification: Evidence that the mediating phonology is phonetically informed. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 151 162.

Marslen-Wilson, W. , Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, *101*, 3-33.

Marslen-Wilson, W., Moss, H., & van Halen, S. (1996). Perceptual distance and competition in lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1376-1392.

Mattys, S.L. (1997). The use of time during lexical processing and segmentation: A review. *Psychonomic Bulletin and Review, 4*, 310-329.

Mitterer, H., & Blomert, L. (2003). Coping with phonological assimilation in speech perception: Evidence for early compensation. *Perception and Psychophysics*, *65*, 956-969.

Mitterer, H., Csépe, V., Honbolygo, F., & Blomert, L. (2006). The recognition of phonologically assimilated words does not depend on specific language experience. *Cognitive Science*, *30*, 451-479.

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004) Lexique 2 : A New French Lexical Database. *Behavior Research Methods, Instruments, & Computers*, *36*, 516-524.

Nolan, F. (1992). The descriptive role of segments: Evidence from assimilation. In G. J.Docherty & D. R. Ladd (Eds.), *Laboratory phonology II: Gesture, segment, prosody* (pp. 261-280). Cambridge: Cambridge University Press.

Norris, D., McQueen, J. M., & Cutler, A. (2002). Bias effects in facilitatory phonological priming. *Memory and Cognition*, *30*, 399-411.

Otake, T., Yoneyama, K., Cutler, A., & van der Lugt, A. (1996). The representation of Japanese moraic nasals. *Journal of the Acoustical Society of America*, *100*, 3831-3842.

Quené, H., van Rossum, M., & van Wijck, M. (1998, December). *Assimilation and anticipation in word perception.* Paper presented at the Proceedings of the Fifth International Conference on Spoken Language Processing, Sydney, Australia.

Radeau, M., Morais, J., & Segui, J. (1995). Phonological priming between monosyllabic spoken words. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1297-1311.

Slowiaczek, L.M., & Pisoni, D.B. (1986). Effects of phonological similarity on priming in auditory lexical decision. *Memory and Cognition*, *14*, 230-237.

Snoeren, N. D., Hallé, P. A., & Segui, J. (2006). A voice for the voiceless: Production and perception of assimilated stops in French. *Journal of Phonetics*, *34*, 241-268.

Spinelli, E. & Gros-Balthazard, F. (in press). Phonotactic constraints help to overcome effects of schwa deletion in French. *Cognition*.

Utman, J. A., Blumstein, S. E., & Burton, M. W. (2000). Effects of subphonetic and syllable structure variation on word recognition. *Perception and Psychophysics*, *62*, 1297-1311.

Warner, N., Jongman, A., Sereno, J., & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics*, *32*, 251-276.

Weber, A. (2001). Help or hindrance: How violation of different assimilation rules affects spoken-language processing. *Language and Speech*, *44*, 95-118.

Weber, A. (2002). Assimilation violation and spoken-language processing: A supplementary report. *Language and Speech*, *45*, 37-46.

Wright, S., & Kerswill, P. (1989). Electropalatography in the analysis of connected speech. *Clinical linguistics and Phonetics*, *3*, 49-57.

Appendix: Stimulus Sentences

Canonical Prime	Assimilated Prime	Unrelated Prime	Target
Word-final stop = /p/			
une *coupe* transversale	une *coupe* droite	un *acte* final	COUPE
une *grippe* contagieuse	une *grippe* durable	un *orgue* portatif	GRIPPE
une *jupe* serrée	une *jupe* grise	un *blanc* laiteux	JUPE
une *nappe* tâchée	une *nappe* déchirée	une *phase* critique	NAPPE
Le *pape* triste	le *pape* débonnaire	une *aire* protégée	PAPE
une *soupe* corse	la *soupe* délicieuse	la *cendre* volcanique	SOUPE
une *troupe* comique	une *troupe* gaie	une *anse* métallique	TROUPE
un *type* sensé	un *type* galant	une *boîte* noire	TYPE
une *lampe* cassée	Une *lampe* de cheveux	une *cible* monumentale	LAMPE
une *pompe* tordue	une *pompe* grinçante	une *firme* japonaise	POMPE
une *rampe* tympanique	une *rampe* glissante	un *jour* férié	RAMPE
un *groupe* solidaire	un *groupe* difficile	un *moine* réfugié	GROUPE
Word-final stop = /t/			
des *bottes* confortables	des *bottes* brillantes	un *nœud* double	BOTTES
une *brute* sanguinaire	une *brute* violente	une *gamme* complète	BRUTE
une *chute* chaotique	une *chute* brutale	un *œuf* cuit	CHUTE
un *doute* persistant	un *doute* grandissant	une *pierre* cassée	DOUTE
une *faute* prévisible	une *faute* grossière	une *pluie* diluvienne	FAUTE
une *note* salée	une *note* grave	le *ventre* plein	NOTE
la *route* perdue	la *route* goudronnée	un *angle* différent	ROUTE
un *vote* secret	un *vote* blanc	une *bague* empruntée	VOTE
la *datte* séchée	la *datte* garnie	une *base* militaire	DATTE
une *grotte* préhistorique	une *grotte* blanche	le *beurre* naturel	GROTTE
la *lutte* continue	la *lutte* brutale	une *blouse* blanche	LUTTE
des *gouttes* scintillantes	des *gouttes* brûlantes	le *but* principal	GOUTTES
Word-final stop = /k/			
la *banque* populaire	la *banque* d'Algérie	un *couple* heureux	BANQUE
un *bloc* plastifié	un *bloc* défectueux	un *drap* humide	BLOC
des *briques* posées	Des *briques* déstabilisées	un *casque* protecteur	BRIQUES
un *choc* terrifiant	un *choc* brutal	une *gloire* éphémère	CHOC
un *grec* patriotique	un *grec* drôle	un *cuir* souple	GREC
un *lac* pollué	un *lac* desséché	la *clef* verte	LAC
la *nuque* tendue	la *nuque* dégagée	un *cadre* familial	NUQUE
une *plaque* tordue	une *plaque* découpée	un *centre* culturel	PLAQUE
un *sac* troué	un *sac* démesuré	une *canne* jaune	SAC
un *truc* particulier	un *truc* débile	un *dieu* omniscient	TRUC
un *flic* pointilleux	un *flic* décidé	une *dose* forte	FLIC
des *claques* sonores	des *claques* violentes	un *masque* facial	CLAQUES

Canonical Prime	Assimilated Prime	Unrelated Prime	Target
Word-final stop = /b/			
une *bombe* destructrice	une *bombe* terrifiante	un *loup* domestiqué	BOMBE
un *club* gastronomique	un *club* touristique	un *vase* simple	CLUB
un *globe* doré	un *globe* terrestre	la *dette* nationale	GLOBE
une *jambe* galbée	une *jambe* cassée	un *titre* national	JAMBE
une *robe* droite	une *robe* serrée	une *branche* professionnelle	ROBE
une *tombe* grandiose	une *tombe* somptueuse	une *corse* sociale	TOMBE
le *tube* digestif	le *tube* cathodique	la *ferme* conservatoire	TUBE
l'*aube* glacée	l'*aube* colorée	une *fiche* personnelle	AUBE
des *bribes* disséminées	des *bribes* signifiantes	un *poste* permanent	BRIBES
un *crabe* délicieux	un *crabe* farci	une *poche* pleine	CRABE
un *cube* dense	un *cube* saillant	une *plaie* profonde	CUBE
un *snob* agréable	un *snob* silencieux	le *linge* sale	SNOB
Word-final stop = /d/			
une *aide* bancaire	une *aide* sociale	une *feuille* quadrillée	AIDE
la *bande* magnétique	la *bande* passante	le *prince* charmant	BANDE
la *blonde* belge	la *blonde* suédoise	un *arc* traditionnel	BLONDE
le *coude* blessé	le *coude* plié	un *gosse* gâté	COUDE
le *guide* breton	le *guide* prévoyant	la *cloche* royale	GUIDE
la *mode* britannique	la *mode* parisienne	une *source* disparue	MODE
le *stade* bruyant	le *stade* complet	une *panne* majeure	STADE
une *viande* braisée	une *viande* saignante	une *poudre* suspecte	VIANDE
une *bride* mauve	une *bride* soudée	une *gare* sympathique	BRIDE
la *dinde* gratinée	la *dinde* savoureuse	une *zone* industrielle	DINDE
les *soldes* budgétaires	les *soldes* précédentes	un *pacte* secret	SOLDES
la *sonde* gastrique	la *sonde* perdue	une *marge* supérieure	SONDE
Word-final stop = /g/			
la *langue* basque	la *langue* pendue	la *face* cachée	LANGUE
les *seringues* doseuses	les *seringues* trouées	les *armes* chimiques	SERINGUES*
une *figue* délicieuse	une *figue* sucrée	la *reine* norvégienne	FIGUE
un *gang* dangereux	un *gang* terrifiant	un *culte* privé	GANG
la *ligue* dissociée	la *ligue* portugaise	le *sable* rouge	LIGUE
les *digues* basses	les *digues* submersibles	les *cerises* mûres	DIGUES
un *dingue* bruyant	un *dingue* paumé	une *bosse* douloureuse	DINGUE
la *drogue* brute	la *drogue* parfaite	le *verbe* conjugué	DROGUE
la *vogue* branchée	la *vogue* française	le *peuple* migrateur	VOGUE
les *fringues* bizarres	les *fringues* sportives	une *marche* funèbre	FRINGUES
les *fugues* d'adolescents	les *fugues* proposées	le *miel* contaminé	FUGUES
la *fougue* disciplinée	la *fougue* passionnée	la *housse* moulante	FOUGUE

* The bisyllabic word *seringues* was inserted in the materials by mistake. Yet, in phrases such as *les seringues*, schwa deletion may occur, so that *seringues* is actually pronounced [sʁɛ̃g]. Because this target did not yield divergent RT patterns compared to the other targets, it was not excluded from the experimental materials.

Author Note

This research was supported by a “MENRT” doctoral fellowship to the first author. We are indebted to Gareth Gaskell, Harald Baayen, and Mirjam Ernestus for their comments and constructive discussion during the preparation of this manuscript. We also thanks two anonymous reviewers who greatly helped improving an earlier version of this paper.

Footnotes

[1] Frequencies of occurrence were drawn from the "film" subpart of the LEXIQUE database (New, Pallier, Brysbaert, & Ferrand, 2004), which contains 16.6 million words and is fairly representative of spoken French. Nouns with a voiceless final stop tended to be more frequent than those with a voiced final stop but not significantly so (o.p.m.: 62.0 vs. 31.8, $p = .095$; log frequencies [log10(o.p.m.)]: 1.39 vs. 1.13, $p = .075$). For all the items but two, the uniqueness point was not reached within the word-form (in "grec" /grɛk/ and "bribe" /brib/, the uniqueness point was the last phoneme). Two indices of lexical competition were tabulated, using the "Vocolex" database (Dufour, Peereman, Pallier, & Radeau, 2002): Cohort size at word offset (this was relevant because virtually all the items were embedded monosyllabic words), and density of "dangerous" (i.e., more frequent) phonological neighbors in number of types or tokens. For voiceless vs. voiced offset items, Cohort size was 23.5 vs. 16.6 (n.s.), number of dangerous neighbors was 2.83 vs. 2.89 (types) or 1715 vs. 1005 (tokens) (n.s. for both). In summary, voiceless offset nouns such as "note" tended to be slightly more frequent than voiced offset nouns such as "coude" but, on the other hand, tended (numerically, not statistically) to be challenged by slightly more lexical competition.

[2] In the Snoeren et al.'s (2006) study, word-final voiceless and voiced stops in non-assimilatory contexts had voicing ratios of 32% and 100%, respectively and were judged as voiced or voiceless 16% or 76% of the time, respectively. It is thus plausible that voiceless stops extracted from running speech objectively and subjectively sound as somewhat voiced, whereas voiced stops would always be, and sound as fully voiced. This asymmetric pattern is in part due to the voicing trail from a preceding vowel into the occlusion portion of a stop: the proportion of voiced occlusion is rarely zero or even close to zero (it was about 0.3 in the Snoeren et al.'s data), whereas the entire occlusion portion may be voiced in unassimilated

voiced stops. Relevant for the present study, however, is that underlyingly voiceless stops in assimilatory context would usually reach nearly full assimilation whereas underlyingly voiced stops would not, and remain halfway between voiced and voiceless.

[3] The lexical characteristics (frequency and competition) for the 66 retained items hardly differed from those tabulated for the initial set of 72 items (see footnote 1). For voiceless vs. voiced offset items, lexical frequency (from the LEXIQUE database, New et al., 2004) was, in average, 65.7 vs. 31.9 o.p.m. (p = .086); Cohort size (from the "Vocolex" database, Dufour et al., 2002) was, in average, 24.0 vs. 17.2 (n.s.), number of "dangerous" neighbors (from "Vocolex") was, in average, 2.91 vs. 2.88 (types) or 1828 vs. 1079 (tokens) (n.s. for both).

[4] We also ran analyses including final stop Place of the final stop (labial, dental, velar) as a factor. Place was far from significance and did not interact with the other factors. For each level of Place, the Voicing x Priming interaction was far from significant (Fs < 1).

[5] Both the error and the RT data of Experiments 1 and 2 show that voiced targets are more difficult overall than voiceless targets (there is no sign of a speed-accuracy trade-off), which runs contrary to the numerical difference in log frequency between the two types of words.

[6] The frequency of co-occurrence for all the noun-adjective pairs we used were tabulated using the LEXIQUE's movie subtitle database (16.7 million word occurrences). Noun-adjective pairs such as *coude plié* are more frequent than pairs such as *note grave*: 3.2 vs. 0.8 occurrences in average; the difference, however, is not significant, $t(70)=1.44$, $p=0.15$.

Table 1. Mean RTs (ms; SD between parentheses) and error rates (%) for lexical decisions in Experiment 1.

	Prime Type		
	Canonical	Assimilated	Unrelated
Target Type			
voiceless final stop (e.g., NOTE)	[nɔt]	[nɔd]	[vɑ̃tʁ]
RT	487 (83)	550 (91)	582 (83)
error rate	1.84	3.51	4.83
voiced final stop (e.g., COUDE)	[kud]	[ku$^{t}_{d}$]	[gɔs]
RT	516 (80)	560 (95)	608 (83)
error rate	3.45	6.73	9.73

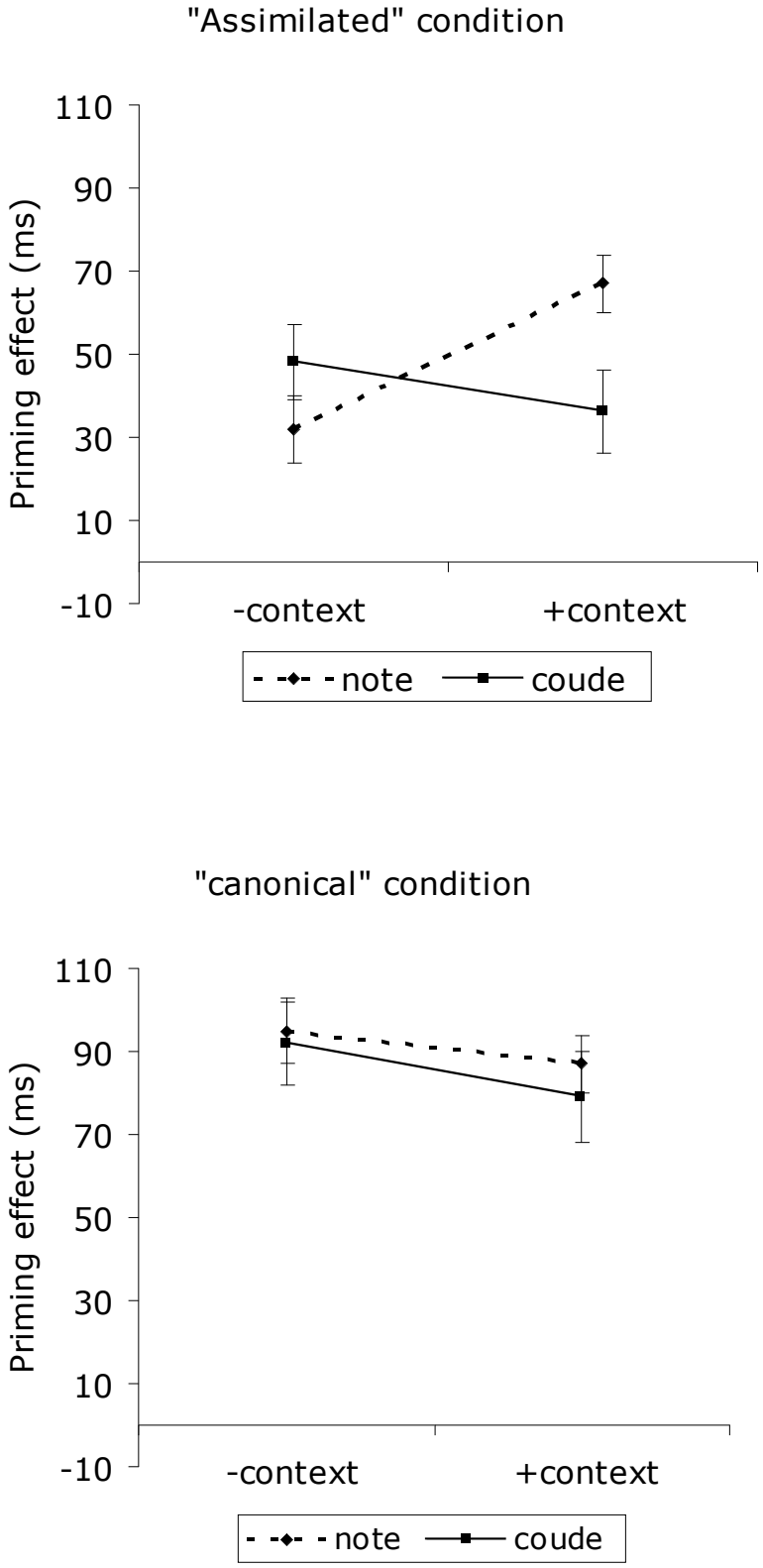
Table 2. Mean RTs (ms; SD between parentheses) and error rates (%) for lexical decisions in Experiment 2.

	Prime Type		
	Canonical	Assimilated	Unrelated
Target Type			
voiceless final stop (e.g., NOTE)	[nɔt]	[nɔd]	[vɑ̃tʁ]
RT	499 (71)	519 (74)	586 (78)
error rate	1.88	1.72	3.93
voiced final stop (e.g., COUDE)	[kud]	[ku$^{t}_{d}$]	[gɔs]
RT	516 (77)	559 (84)	595 (87)
error rate	4.23	4.24	9.53

Figure caption

Figure 1. Priming effects for visual target words with a voiced vs. voiceless offset stop (e.g., COUDE vs. NOTE) primed by assimilated word forms (upper panel) or canonical (unassimilated) forms (lower panel); the absence vs. presence of context in the prime is noted "-context" vs. "+context" (Experiments 1 vs. 2).

Figure 1


"Assimilated" condition
Priming effect (ms)
110
90
70
50
30
10
-10
-context
+context
note
coude
"canonical" condition
Priming effect (ms)
110
90
70
50
30
10
-10
-context
+context
note
coude