INTERSPEECH 2019
September 15–19, 2019, Graz, Austria

Compensation for French liquid deletion during auditory sentence processing

Sharon Peperkamp, Alvaro Martin Iturralde Zurita

Laboratoire de Sciences Cognitives et Psycholinguistique, ENS - PSL University – EHESS – CNRS, Paris, France

sharon.peperkamp@ens.fr, alvaro.iturraldezurita@mail.mcgill.ca

Abstract

Phonological rules change the surface realization of words. Listeners undo these changes in order to retrieve the canonical word form. We investigate this so-called compensation for a French deletion rule, i.e. liquid deletion. This rule optionally deletes the final consonant of a word-final obstruent-liquid cluster. It can apply both before consonants and before vowels, but its application is about twice as frequent before consonants. Using a word detection task, we find an overall relatively low rate of compensation, which we argue is due to the relatively high perceptual salience of the rule. We also observe a clear effect of context, though: listeners compensate more than twice as often for a deleted liquid before a consonant than before a vowel. This is evidence that compensation involves fine-grained knowledge about the probability of the rule's application in different contexts.

Index Terms: speech perception, word recognition, French, phonological rules, liquid deletion

1. Introduction

All languages have phonological rules that change the canonical sound shape of words when they are pronounced within sentences. During spoken language processing, listeners automatically 'undo' the changes they bring about in order to retrieve the canonical word form [1-9]. Most research on this so-called compensation has focused on assimilation rules, such as English place assimilation. For instance, one of the earliest studies found that an auditory sentence with an assimilated token of the word 'lean' primes the visually presented word LEAN when the context licenses the change (*Sandra will only eat lea[m] bacon*: viable context for dental-to-labial change), but less so when the context does not license the change (*Sandra will only eat lea[m] gammon*: unviable context for dental-to-labial change) [1]. Thus, listeners compensate for the phonological change in a context-sensitive manner. Furthermore, they do so not only when traces of the underlying phoneme are present in the assimilated word, but also – although often to a lesser extent – when assimilation is complete or nearly complete [2, 5-9] (but cf. [4]).

Compensation for phonological rules is not limited to assimilation. For instance, one study showed compensation for a Dutch /t/-reduction rule, which deletes a word-final /t/ (but leaving acoustic traces on the surrounding sounds) in certain phonological contexts [8]. Here, we focus on another reduction rule, i.e. French liquid deletion. Liquid deletion (henceforth: LD) optionally deletes a liquid in word-final obstruent-liquid clusters, e.g., *[tab]* < *[tabl]* 'table' or *[uv]* < *[uvʁ]* 'open$_{IMP}$'. Whether this is a true deletion rule or better described as reduction, with acoustic traces of the deleted liquid being present on the preceding obstruent, is unknown. We will, however, test compensation for complete deletion of the liquid, as explained below.

Our study has two aims. First, LD has been described as occurring before consonants and – to a lesser extent – before pause [10]. Corpus studies, though, have shown LD to apply before vowels as well [11,12], although also not as often as before consonants. We test for compensation for LD before vowels and before consonants, and examine whether listeners likewise compensate for LD in both contexts but more before consonants. Second, word-final /l/ and /r/ are acoustically quite salient, and the difference between word tokens with and without deletion is thus relatively salient too. (This holds especially true before consonants, where LD competes with a rule of schwa epenthesis such that the liquid is either deleted or followed by an epenthetic schwa [10]; see e.g. Figure 1a in section 2.1.2) The relative salience of the change induced by LD might reduce the size of the automatic compensation effect compared to that for less salient changes. Specifically, we will compare compensation for LD to compensation for a French assimilation rule, i.e. voicing assimilation.

Concerning voicing assimilation, previous work has used a word detection task to measure the amount of compensation [7]. In this task, participants are presented auditorily with an isolated target word followed by a sentence, and they are asked to indicate whether the sentence contains the target word or not. The mean detection rates for assimilated words in viable and unviable contexts were 65% and 14%, respectively (in the control condition with unassimilated words the mean detection rate was 96%). In this study, assimilation was complete; in the list of sentences to be read by the speaker who recorded the stimuli, the spelling of the assimilated words reflected the modification; these words hence showed up as pseudo-words (e.g., … *cabe grise* … 'grey cape' for target word *cape* [kap]; cabe [kab] is not a French word).

We use the same word detection task and procedure for stimuli recording as [7]. Our design, however, is a modified version of the one in that study. Specifically, as there is no unviable context for LD, we add a distractor condition in which the target word ends in a consonant cluster other than an obstruent-liquid cluster. The accompanying sentence always contains the target word; in the test condition it appears either with or without the final liquid, while in the distractor condition it always appears without the final consonant. Thus, the expected response is 'yes' for all test trials in the condition without deletion and 'no' for all distractor trials (which all have deletion). For the test trials in the crucial condition with deletion, there is no unique expected response. If listeners compensate for LD, they should reply 'yes', and otherwise they should reply 'no'; as LD is more frequent before consonants than before vowels, we expect more 'yes' responses in the consonant than in the vowel condition.

http://dx.doi.org/10.21437/Interspeech.2019-2950

2. Experiment

2.1. Method

2.1.1. Participants

Sixty native speakers of French, 40 women and 20 men aged between 18 and 60 years, participated. Seven additional participants were excluded from the analyses because they were early bilinguals (N=2), they made more than 30% errors on the distractor sentences (N=3), or because of program failure (N=2).

2.1.2. Stimuli

For the test condition we selected 36 monosyllabic nouns ending in an obstruent-liquid cluster. In half of them, the liquid was /l/, in the other half it was /r/. For the distractor condition we selected 24 nouns (22 mono- and 2 disyllabic), all ending in a consonant cluster with no final liquid and not subject to any deletion rule. Finally, for a training phase we selected 10 nouns ending in a single consonant.

Each test item was paired with one vowel-initial and one consonant-initial context; they consisted either of an adjective or of a short prepositional phrases. For instance, *arbre* [aʁbʁ] 'tree' was paired with *exotique* 'exotic', and *japonais* 'Japanese' and *cible* [sibl] 'target' was paired with *en paille* '(made) of straw' and *volante* 'flying'. Consonant-initial contexts were chosen such that: 1/ they did not contain an obstruent-liquid cluster themselves, 2/ their first phoneme was not identical to the obstruent of the obstruent-liquid cluster of the target item, and 3/ the application of LD did not create the environment for any other phonological process (such as voicing assimilation or schwa insertion). Some contexts that produced very frequent, fixed phrases that are typically produced with LD deletion, such as *boucle d'oreille* 'earring', were also avoided.

The two contexts for a given test item had the same number of syllables (between one and three), and the resulting pairs of test item plus context (e.g. *arbre exotique* / *arbre japonais*) were roughly matched in frequency: on google.com the numbers of French pages in which the members of each pair occur do not differ (paired t <1).[1]

For each test item we constructed two sentence frames that could contain both pairs of test item plus context. For half of the test items both these pairs appeared sentence-medially, for the other half they appeared sentence-finally, such that the location of the target word was unpredictable. By way of example, the four sentences for the test item *trouble* [tʁubl] 'disorder' are shown below.

1 a. *Elle a écrit plusieurs livres sur ce* ***trouble affectif****.*
'She's written several books on this emotional disorder.'

b. *Elle a écrit plusieurs livres sur ce* ***trouble de l'humeur****.*
'She's written several books on this mood disorder.'

2 a. *On fait beaucoup d'études sur ce* ***trouble affectif****.*
'There are many studies on this emotional disorder.'

b. *On fait beaucoup d'études sur ce* ***trouble de l'humeur****.*
'There are many studies on this mood disorder.'

For each distractor item we constructed two different sentences, in which the next word was vowel- or consonant-initial, respectively, as illustrated below for *auberge* [obɛʁʒ] 'hostel'.

3 a. *Pour nos vacances on cherche une* ***auberge isolée****.*
'For our vacation we're looking for an isolated hostel.'

b. *Le groupe a décidé d'aller à l'****auberge voisine****.*
'The group decided to go to the nearby hostel.'

Like for the test items, both pairs of distractor item plus context appeared sentence-medially for half of the distractor items and sentence-finally for the other half.

For each training item we constructed a single sentence. In half of them the context was vowel-initial and in the other half it was consonant-initial; and in half of them the pair of training item plus context appeared medially and in the other half it appeared finally.

The individual test and distractor items were recorded by a male native speaker of French. All the sentences (test, distractor, and training) were recorded by a female native speaker of French at a normal speech rate. She recorded each test sentence twice, once with and once without LD; each distractor sentence once, with final consonant deletion; and each training sentence once, half of them with and half without (N=5) a consonant substitution. For the recordings with a deletion or a substitution, the test or distractor item was written such that it reflected the modification, e.g. (boldface added) "Elle a écrit plusieurs livres sur ce **troub** de l'humeur" (test, cf. *trouble*), "*Le groupe a décidé d'aller à l'****aubère*** *voisine*" (distractor, cf. *auberge*), "La voiture s'est arrêtée en face de la **pielle** géante" (training, cf. *pierre* 'stone').

Figure 1 shows waveforms of *cercle social* 'social circle' and 'cercle intime' from the recordings with and without LD of the sample sentences *Pour les ados, le* ***cercle social/intime*** *est le plus important.* ('For teens, the social/intimate circle is the most important.').

The test sentences were divided into two lists, such that each test item appeared twice in each list, in two different frame sentences, one with a consonant-initial and one with a vowel-initial context. (For instance, for the example above, sentences 1a and 2b would appear in the same list, as would 1b and 2a). These lists were further divided into two, such that in each of them, half of the test items were pronounced with and half without LD (across the two contexts). Finally, all distractor sentences were added to each of the four lists.

2.1.3. Procedure

The participants were randomly divided into four groups, one for each stimulus list, and tested individually in a sound-attenuated booth. On each trial, they were presented auditorily with an individual target word and 500ms later with a sentence. Their task was to indicate as quickly as possible on a button box whether the target word was present in the sentence. They had to press the button on the side of their dominant hand if they thought the word was present and the button on the other side if they thought the word was absent or

[1] We did not use the Google Books Ngram Viewer, as it has a preponderance of academic books, which are characterized by a specialized vocabulary.

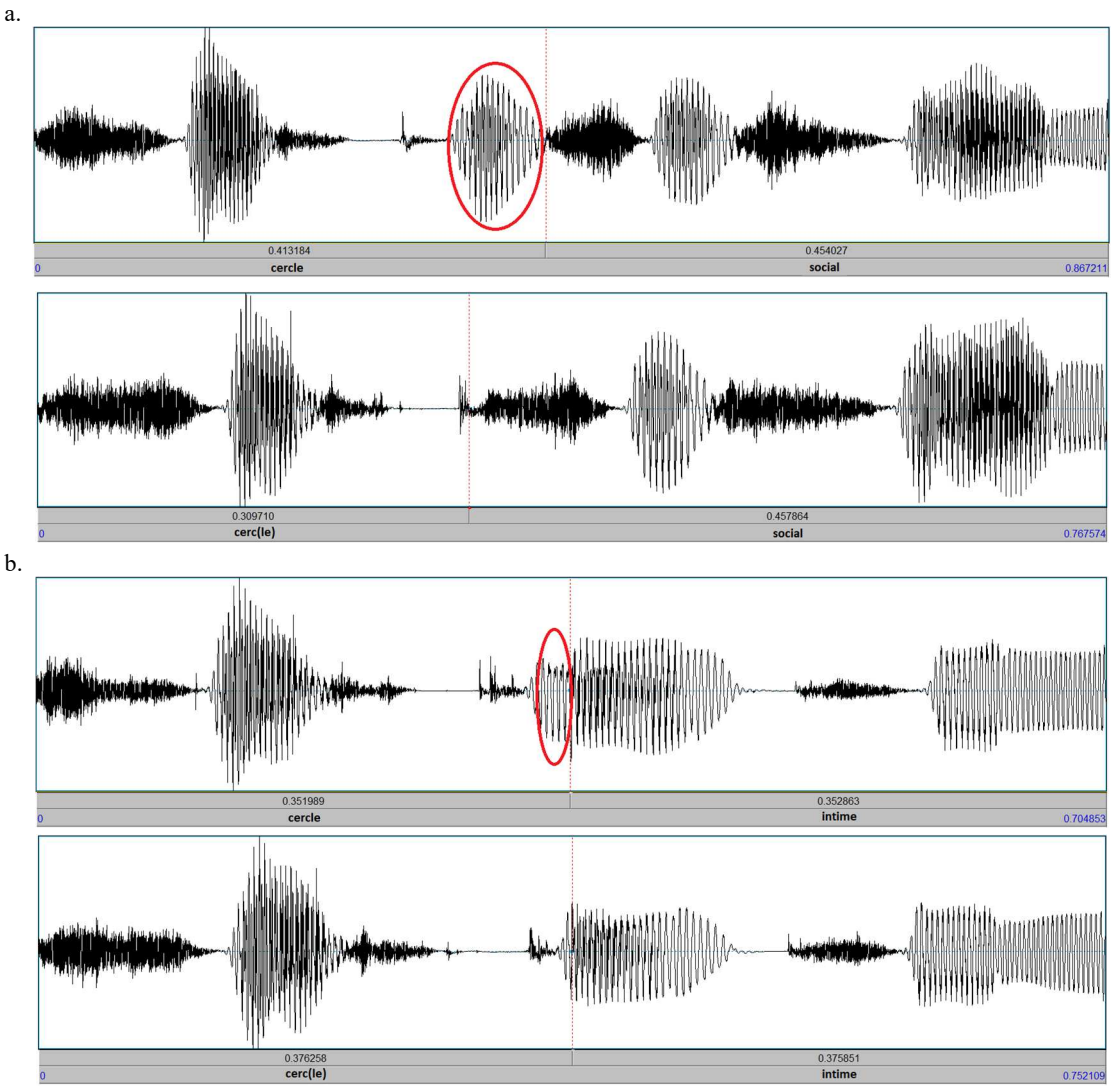

Figure 1: a. *Waveforms of "cercle social" ('social circle') from sentence recordings without LD (top panel; the liquid is followed by an epenthetic schwa) and with LD (bottom panel). b. Waveforms of "cercle intime" ('intimate circle') from sentence recordings without LD (top panel) and with LD (bottom panel). The critical word boundaries are marked by dotted vertical red lines, and the target liquids are shown in red ovals.*

deformed.

The experiment started with the 10 training sentences. Participants received feedback as to their accuracy (*Correct*/*Incorrect*) and response time (*Trop lent!* 'Too slow!' if they responded more than 2s after the end of the sentence; *Trop rapide!* 'Too fast!' if they responded before the onset of the context). For six participants the training session was repeated because they indicated to the experimenter that they were not sure to have understood the task (N=4) or because of a technical error (N=2). The test phase (without feedback) was divided into two blocks, each consisting of 36 test and 24 distractor sentences. Each test or distractor item appeared once per block, with context (vowel or consonant), deletion (yes or no), and position (final or medial) counterbalanced across blocks. Within each block, the sentences were presented pseudo-randomly, such that no more than three trials of the same type (test or distractor) could occur in a row.

Participants could take a short break in between the two blocks. The experiment lasted about 25 minutes.

2.2. Results

One distractor item was identified as an outlier, as its detection rate (46.0%) was higher than the mean plus 3 standard deviations (8.5% + 30.0). This item, *cycliste* 'cyclist', was therefore discarded. The mean detection rates for the remaining distractor and test items are shown in Table 1.

Table 1: *Mean detection rates and standard errors for distractor and test items in consonant and vowel contexts.*

	distractor	**test**	
	deletion	**deletion**	**no deletion**
C	6.74% (3.24)	25.37% (5.62)	97.87% (1.86)
V	2.97% (2.19)	9.54% (3.79)	97.96% (1.82)

The data were analyzed in logistic mixed effects models using the lme4 package [13] in R [14]. All fixed factors were defined using contrast-coding, and significance was assessed through comparison of the full model with models without the

relevant factor or interaction [15]. Slopes for the random variables were included if and only if they improved the model. As the counterbalancing factor Position (medial vs. final) was never significant, we report the results for models without this factor.

First, we focused on all the trials with deletion. We constructed a model with Type (distractor vs test), Context (consonant vs vowel), and their interaction as fixed factors; the random structure included intercepts for participant and item, by-participant slopes for Type and Context, and a by-item slope for Context. There was a main effect of Type, with higher detection rates for deletion in test than in distractor items ($\beta = 0.75$, SE = 0.17, $z = 4.30$, $Chi^2(1) = 15.8$, $p < .0001$), a main effect of Context, with higher detection rates for deletion before consonants than before vowels ($\beta = 0.35$, SE = 0.13, $z = 2.81$, $Chi^2(1) = 4.25$, $p < .04$), and a Type × Context interaction ($\beta = 0.29$, SE = 0.10, $z = 2.96$, $Chi^2(1) = 7.85$, $p < .006$), due to the fact that the difference in detection rates before consonants vs. vowels was larger for the test than for the distractor items.

Next, we focused on all test trials. We constructed a model with Deletion (yes vs no), Context (consonant vs vowel), and their interaction as fixed factors; the random structure included intercepts for participant and item as well as by-item and by-participant slopes for Deletion. There was a main effect of Deletion, with higher detection rates for items without than with deletion ($\beta = 3.57$, SE = 0.23, $z = 15.7$, $Chi^2(1) = 110.0$, $p < .0001$), a main effect of Context, with higher detection rates before consonants than before vowels ($\beta = 0.44$, SE = 0.09, $z = 5.08$, $Chi^2(1) = 24.6$, $p < .0001$), and a Deletion × Context interaction ($\beta = 0.47$, SE = 0.09, $z = 5.35$, $Chi^2(1) = 27.5$, $p < .0001$), due to the fact that the higher detection rates before consonants was present only for the items with deletion.

Finally, we used linear regressions to examine the role of item frequency on the detection of test items with deletion. Neither the log frequency of the test items (adjusted $R^2 = -0.01$, $p > .1$) nor the log frequency of the sequence of test item plus context (adjusted $R^2 = 0.03$, $p > .1$) correlated with the mean detection rate.

3. Discussion

Using a word detection task, we examined compensation for French liquid deletion (LD) in spoken sentences. The results show that listeners indeed compensate for this rule, since for trials with deletion the detection rate is higher for test than for distractor items. Yet, the size of the effect is small, with a mean detection rate of 17.5% for test items with deletion compared to 97.9% for test items without deletion.

The results also show that compensation is context-sensitive, as shown both by the Type × Context interaction in the analysis for the test and distractor items with deletion, and by the Deletion × Context interaction in the analysis for the test items with and without deletion. In particular, the mean detection rates before consonants and vowels were 25.4% and 9.5%, respectively, a ratio of 2.7. This context effect reflects the production patterns reported in two sociolinguistic studies: First, for middle-class speakers 65% LD was found before consonants and 23% before vowels in mostly conversational speech in [11]. Second, an analysis of the subset of middle-class speakers in France (as opposed to Belgium and Switzerland) in [12] reveals very similar rates for conversational speech, i.e. 67.1% before consonants and 29.4 before vowels. The pattern for read speech in this second study is similar but shows overall lower deletion rates, i.e. 16.5% before consonants and 11.8% before vowels.

Even in the consonantal context, the mean detection rate in our experiment is considerably smaller than the one reported in a similar experiment on compensation for French voicing assimilation, i.e. 65% [7]. Like LD, voicing assimilation is optional: a study on a large corpus of journalistic speech reported a mean rate of 22% [16], and a production experiment found a mean rate of 48% [17]. Moreover, both our experiment and the one on voicing assimilation used a word detection task with identical instructions, and both tested on sentences in which the items had undergone a complete change. The comparison between experiments is thus especially interesting, despite differences in their exact design and a few other aspects (specifically, the sentences were recorded by different speakers, and the experimenter was a native French speaker for the voicing assimilation experiment but an L2 French speaker for the current experiment). We hypothesize that the relatively small compensation effect for LD compared to that for voicing assimilation is due to its relative perceptual salience. That is, the deletion of a word-final liquid is arguably more perceptible than a voicing change on a word-final obstruent. Participants might thus have been aware of some of the deletions. This, in turn, might have deflated the detection rate of test items with deletion, as participants had been asked to give a 'no' response whenever they thought the target word was either absent or deformed.

Interestingly, though, perceptual salience cannot account for the context effect. To see why, recall that before consonants, either the liquid is deleted or a different rule of final schwa insertion applies obligatorily. The difference between test items with and without deletion is therefore larger before consonants than before vowels (see Figure 1). In other words, deletion is perceptually more salient when it occurs before a consonant than when it occurs before a vowel. Yet, compensation for LD occurs more often before consonants than before vowels. This context effect, then, can only be explained by the fact that in production, LD applies more frequently before consonants than before vowels. Thus, performance in our task involves probability matching. In this respect, the absence of an effect of frequency – especially that of the sequences of test item plus context – is puzzling. Indeed, under the assumption that these sequences are also more likely to undergo LD in production, one would expect higher detection rates for more frequent sequences. The absence of this effect could be due either to our frequency measure being too noisy or to the above-mentioned assumption being false.

To conclude, we have shown that French listeners compensate for the optional rule of LD. The size of the effect is larger before consonants than before vowels, despite the fact that LD is arguably more perceptible in the former context. This is evidence that listener are sensitive to the differential frequency of occurrence of LD, with the rule indeed applying more often before consonants than before vowels.

4. Acknowledgements

This study was supported by ANR-17-EURE-0017 and ANR-17-CE28-0007-01. We would like to thank Anne Christophe for recording the stimuli and Mathieu Avanzi for sharing the raw data of [12] with us, and three reviewers for comments.

5. References

[1] M.G. Gaskell and W.D Marslen-Wilson, "Phonological variation and inference in lexical access." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 22, pp. 144-158, 1996.

[2] M.G. Gaskell & W.D. Marlsen-Wilson, "Lexical ambiguity resolution and spoken word recognition." *Journal of Memory and Language,* vol. 44, pp. 325–349, 2001.

[3] E. Coenen, P. Zwitserlood, and J. Bölte, "Variation and assimilation in German: Consequences for lexical access and representation." *Language and Cognitive Processes*, vol. 16, no. 5-6, pp. 535-564, 2001.

[4] D. Gow, "Does English coronal place assimilation create lexical ambiguity?" *Journal of Experimental Psychology: Human Perception and Performance*, vol. 28, no. 1, pp. 163-179, 2002.

[5] M.G. Gaskell and N.D. Snoeren, "The impact of strong assimilation on the perception of connected speech." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 34, no. 6, pp. 1632-1647, 2008.

[6] N.D. Snoeren, J. Segui, and P. Hallé, "Perceptual processing of partially and fully assimilated words in French." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 34, no. 1, pp. 193-204, 2008.

[7] I. Darcy, F. Ramus, A. Christophe, K. Kinzler, and E. Dupoux, "Phonological knowledge in compensation for native and non-native assimilation." In F. Kügler, C. Féry & R. van de Vijver (Eds.), *Variation and Gradience in Phonetics and Phonology*. Berlin: Mouton De Gruyter, pp. 265-309, 2009.

[8] H. Mitterer and J. McQueen, "Processing reduced word-forms in speech perception using probabilistic knowledge about speech production." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 35, no. 1, pp. 244–263, 2009.

[9] H. Mitterer, S. Kim, and T. Cho, "Compensation for complete assimilation in speech perception: The case of Korean labial-to-velar assimilation." *Journal of Memory and Language,* vol. 69, no. 1, pp. 59-83, 2013.

[10] F. Dell, *Les règles et les sons*. Paris: Hermann, 1985.

[11] Z. Boughton, "Social class, cluster simplification and following context. Sociolinguistic variation in word-final post-obstruent liquid deletion in French." *Journal of French Language Studies*, vol. 25, pp 1-21, 2015.

[12] S. Brognaux and M. Avanzi, "Sociophonetics of phonotactic phenomena in French." in *ICPhS 2015 - 18th International Congress of Phonetic Sciences, August 10-14, Glasgow, UK, Proceedings*, 2015, paper number 0744.

[13] D. Bates, M. Maechler, B. Bolker & S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, 1-48, 2015.

[14] R Core Team, "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/, 2014.

[15] R. Levy, "Using R formulae to test for main effects in the presence of higher-order interactions," arXiv preprint arXiv:1405.2094, 2014.

[16] M. Adda-Decker and P. Hallé, "Bayesian framework for voicing alternation & assimilation studies on large corpora in French," *Proceedings of the 16th Congress of Phonetic Sciences*, Saarbrücken, ID1562, 2007.

[17] N.D. Snoeren and P. Hallé, "A voice for the voiceless: Production and perception of assimilated stops in French," *Journal of Phonetics*, vol 34, pp. 241-268, 2006.