

How similar are clusters resulting from schwa deletion in French to identical underlying clusters?

Audrey Bürki¹, Cécile Fougeron², Christophe Veaux³, Ulrich H. Frauenfelder¹

¹ Laboratoire de Psycholinguistique Expérimentale, Université de Genève, Genève, Switzerland

² Lab. de Phonétique et Phonologie, UMR7018, CNRS-Paris3/Sorbonne Nouvelle, Paris, France

³ IRCAM, Analysis-Synthesis Team, Paris, France

{audrey.buerki;ulrich.frauenfelder}@unige.ch, cecile.fougeron@univ-paris3.fr,
veaux@ircam.fr

Abstract

Clusters resulting from the deletion of schwa in French are compared with identical underlying clusters in words and pseudowords. Both manual and automatic acoustical comparisons suggest that clusters resulting from schwa deletion in French are highly similar to identical underlying clusters. Furthermore, cluster duration is not longer for clusters resulting from schwa deletion than for identical underlying clusters. Clusters in pseudowords show a different acoustical and durational pattern from the two other clusters in words.

Index Terms: French schwa, acoustic cues, clusters, duration

1. Introduction

In French, certain words (hereafter “schwa words”) can be pronounced either with schwa (e.g., [ʁəkɛ̃], ‘shark’) or without (e.g., [ʁkɛ̃]). Different proposals have been put forward to account for this alternation between schwa and zero (see [1] for a review). The common view is to consider that schwa is present in the underlying representation of the word and that its absence in the output –at least word internally– is the consequence of a deletion process (but see [2] or [3] for an interpretation in terms of allophony or epenthesis, respectively). The nature of this deletion process is controversial, however. A first dimension along which theories differ concerns its gradual vs. categorical nature. A few authors within the articulatory phonology framework ([4]) have argued in favor of a gradual process, like what has been proposed to account for schwa deletion in English ([5]). Schwa deletion is seen as the end-point of a gradual process of reduction. The absence of schwa is only apparent; schwa is still present but hidden by the overlap of surrounding consonants’ gestures. Most proposals however locate the alternation process in the phonological component, assuming a phonological rule of deletion. The process is thus assumed to be categorical: schwa is either present or absent. Within this framework of a categorical, phonological process, proposals differ according to whether the process is complete (e.g., [6]) or incomplete ([7], [8]).

Detailed acoustic or articulatory analyses of the output forms have been conducted to test these proposals. Most of these studies have compared schwa deleted clusters and identical underlying clusters in quasi-homophonic sequences (e.g., jeu drôle ‘funny game’ vs. jeu d’rôle ‘role-play’). For instance, [8] conducted an acoustic study and observed in a few utterances that the consonant preceding schwa is stronger and longer than coda consonants. The consonant following the schwa is also longer, and stronger if it is [ʀ] or [l]. According

to [8] these results are in favor of an incomplete deletion process: when a (nucleus) schwa is absent in the output, the nucleus of the syllable is nevertheless maintained. [9] compared the duration of the surrounding consonants in schwa deleted clusters and in underlying clusters (in words and pseudowords) in the productions of three speakers. He observed that preceding consonants are significantly longer in schwa deleted clusters than in underlying clusters.

Two articulatory studies have also found differences between schwa deleted and underlying identical clusters. Firstly, [10] found that in schwa deleted clusters, the consonants maintain some articulatory properties of the form with schwa (e.g., amount of linguo-palatal contact, timing between gestures). [11] measured lip rounding and observed maximal rounding for tokens realized with the schwa, medium rounding for schwa deleted clusters and minimum rounding for non-schwa tokens. They conclude that the deletion process is gradual and best explained by a gradual overlap of gestures.

It is important to note that the “homophonic” sequences compared in the studies described above were not selected to control for various factors which are known to influence consonants’ durations and other acoustical properties, such as lexical frequency ([12], [13]), position with regard to syllabic boundaries ([14], [15]), word length ([16]) and position of the syllable in the word ([17]). The question therefore arises whether the differences observed between schwa deleted and underlying clusters in terms of duration or other phonetic characteristics are indeed due to schwa, or attributable to other differences in the words or sequences under consideration.

The main aim of the present study is to shed further light on these questions by a thorough investigation of the durational and acoustical properties of consonant clusters resulting from schwa deletion. We first analyze the clusters’ durations in order to determine whether they are longer for schwa deleted clusters than for identical underlying clusters after the variation explained by other variables known to affect segment duration is taken into account. Secondly, we investigate whether clusters resulting from schwa deletion show acoustic remnants of an underlying vowel that would be incompletely deleted or hidden by gestural overlap.

An additional aim of this study is to determine whether clusters in pseudowords behave like clusters in words. For the majority of schwa words, it is difficult to find in the French lexicon a word with a corresponding underlying cluster matching in terms of syllabic structure. The use of pseudowords for comparison is thus quite convenient for this reason, but one has to be sure that the gestural cohesion pattern in the cluster is not atypical because of their non-word status.

In order to achieve these aims, we conducted three different analyses comparing schwa deleted clusters, identical

underlying clusters in words and identical clusters in pseudowords. The first analysis is concerned with the durations of these different clusters. The second analysis is based on a manual spectrographic comparison and the last analysis involves an automatic classification of all these clusters based on an acoustical distance measure.

2. Method

The present study is part of a larger investigation of the acoustic properties of schwa deleted and underlying clusters. We analyze only a part of the large corpus that we collected.

2.1. Speakers

Thirty-three students from the Psychology Department of the University of Geneva took part in the experiment. They were all monolingual French speakers.

2.2. Corpus constitution

2.2.1. Material selection and preparation

We first selected 34 CC clusters resulting from schwa deletion in French. We found three items for each of these 34 clusters: a schwa word whose schwa deleted variant contained the cluster, an existing word with an identical underlying cluster and a pseudoword with an identical cluster. For instance, the cluster [sɫ] was associated with *cela* ‘this’ (schwa word in which schwa can drop [sla]), *slalom* (existing word with identical underlying cluster) and *slami* (pseudoword). Cluster position in the pseudoword was kept constant between the schwa word and the pseudoword. Furthermore, whenever possible, we chose an existing word with an underlying cluster in the same position with respect to word and syllabic boundaries as in the schwa word and pseudoword. Such a word existed for eight clusters only. For 24 clusters, we could match a word with an identical underlying cluster but in a different position with respect to word and syllabic boundaries. For two clusters, no words could be found containing the cluster, whatever its position.

We then inserted each of these items into a carrier sentence (e.g., *J’aime les chanteurs de rap, Slami est mon préféré* ‘I like rap singers, *Slami* is my favorite’). For the three items containing a given cluster (schwa, underlying and pseudoword), the position of the item in the sentence was identical and the three sentences had a similar number of syllables. Schwa deletion in the schwa words was transcribed with an apostrophe (e.g., *c’la*), as used in many songs or comic books in French.

We created two lists of sentences. The first list contained the items with half of the clusters and the second list the items with the other half.

2.2.2. Recordings

Each participant saw one of the two lists twice. Sentences appeared one by one on the screen of a computer. Participants had to read them as fluently as possible, make them disappear by pressing a button and repeat them. There was a short pause after the first list. Each speaker produced four tokens of each item, that is, one reading and one repetition times two presentations of the list. Items in each presentation were randomized. Vocal responses were recorded using a DAT system in a soundproof cabin.

For the present study, we selected the eight clusters with a word matched for the position of the cluster with respect to

word and syllabic boundaries. It concerned the following clusters: [fr], [lm], [lv], [mn], [pl], [sl], [tn], [vn]. A first automatic alignment of phoneme boundaries was realized using *easyalign* (<http://www.rc.unesp.br/xivbspb/ealign/>). We then undertook a manual correction of these boundaries. Errors and dysfluences were disregarded (113 errors, 7% of tokens).

2.2.3. Analysis of duration

The duration of the whole cluster was computed automatically with Praat. We calculated speech rate for each sentence as its total duration (in ms) divided by the number of syllables. We also considered word length with two measures, number of phonemes and number of syllables. Lexical frequency was taken from the *Lexique* database [18].

2.2.4. Spectrographic analysis

For each cluster, a set of acoustic cues which could be considered as linked to schwa deletion was defined based on previous studies on schwa deletion in French ([8]) and in other languages ([5], [19]). To these predefined acoustic cues other cues were added if observed in the productions of our speakers. These cues could either indicate the remnants of a vowel between the two consonants (e.g., short interval with formant-like structure), or more coarticulation between the two consonants (devoicing of [l] in [pl]). There were between six and fourteen different cues depending on the cluster.

The eight clusters were split into two groups. Two phoneticians each rated the tokens of one group. For each token and for each acoustic cue, they had to decide whether the cue was present or absent.

2.2.5. Automatic acoustical analysis

All the CC clusters were extracted from the target items and a Mel-cepstrum analysis was performed to characterize them acoustically. Given \mathbf{x} and \mathbf{y} , the vector sequences of mel-cepstrum coefficients (mfcc) associated with two clusters, we defined an acoustic distance between these clusters as the Mahalanobis difference between \mathbf{x} and \mathbf{y} . Since the clusters had unequal length, we used dynamic time warping ([20]) to establish a non linear mapping that minimizes the cost of matching \mathbf{x} to \mathbf{y} . This acoustic distance was used to define a classification tree (CART) for each group of clusters as described in section 3.3.

3. Results and Discussion

3.1. Duration

We calculated the mean cluster duration for each condition (schwa cluster, underlying cluster, pseudoword cluster). Recall that for the ‘schwa cluster’ condition, we considered only tokens produced in their non-schwa variants. The cluster is shortest in the schwa condition (127 ms, $sd = 34$), longest in the pseudoword condition (152 ms, $sd = 36$) and intermediate for the underlying condition (133 ms, $sd = 34$).

We ran a mixed effect model on the 1422 observations with speaker and cluster as random terms, cluster duration as the dependent variable and condition as the only predictor. The model shows a main effect of condition $F(2,1419) = 98.44$, $p < 0.0001$. Each contrast is significant after the Bonferroni correction has been applied (pseudoword longer than schwa or underlying: $p < 0.0001$, schwa shorter than

underlying: $p < 0.01$). Figure 1 shows the boxplots for cluster duration in each condition, for each cluster. It suggests much variation in the duration pattern for the different clusters.

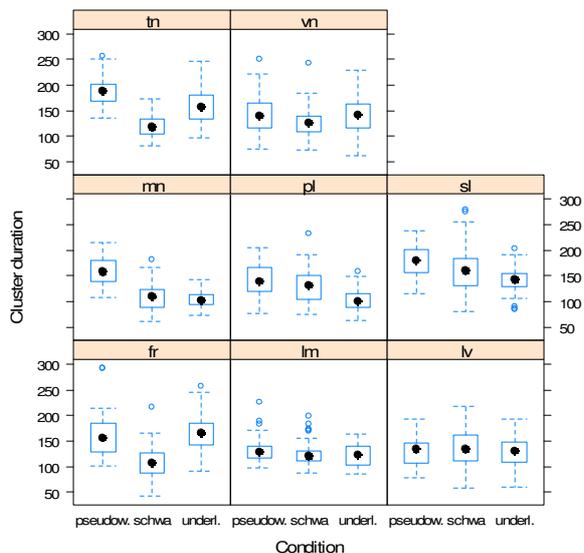


Figure 1: Boxplots showing cluster duration in each condition, for each cluster.

In the next analysis we examined whether these differences in duration were best explained by condition (schwa, underlying and pseudoword) or by other predictors known to affect segment duration. We again ran a mixed effect model with cluster and speaker as random terms and cluster duration as the dependant variable. We added several predictors: number of phonemes, number of syllables, speech rate and lexical frequency. As the number of syllables and the number of phonemes were highly correlated (0.86), they were orthogonalized by replacing the variable “number of syllables” by the residuals of a linear model in which the number of phonemes predicted the number of syllables. We then used a stepwise procedure; predictors were entered sequentially in the model, and retained only if significant (at $p < 0.05$). Thus the contribution of each predictor is determined once the contribution of the previous predictors is taken into account. This new model shows four main effects (see Table 1 for statistical values). Cluster duration decreases with lexical frequency, word length (number of phonemes) and speech rate and is influenced by condition. Pseudowords have again a longer cluster duration than underlying and schwa clusters. There is no difference between schwa and underlying clusters. We can thus conclude that the difference in duration we observed between the schwa and underlying conditions is not due to schwa per se but to other variables known to affect segment duration.

Table 1: Duration model: regression coefficients, F and p values.

Predictors	β	F	p
Frequency	-0.06	7.33	<0.01
Nb of phonemes	-10	196.1	<0.0001
Speech rate	-18180	471.3	<0.0001
Condition:		20.1	<0.0001
Pseudow. vs. schwa	12.36		<0.0001
Pseudow. vs. underl.	9.69		<0.0001
Schwa vs. underl.	-2.7		=0.2

3.2. Spectrographic analysis

As the set of acoustic cues differed for each cluster examined, we analyzed each cluster separately. For the cues which showed some variation in the responses (i.e., present and absent responses) between the conditions, a generalized mixed effect model with presence/absence of the cue as the response and condition as the fixed effect was run.

For most clusters, there is no difference between the schwa and underlying conditions. The clusters [mn], [fr] and [lm] show no difference between the three conditions, whatever the acoustic cue considered. The clusters [pl] and [lv] only show differences between pseudowords and the other two categories. For the cluster [pl], pseudowords show less devoicing of the second consonant than schwa and underlying clusters. For the cluster [lv], pseudowords have fewer tokens with a formant structure between the two consonants and more tokens with an increase in the amplitude of the signal between the two consonants.

The three remaining clusters show differences between the schwa and the underlying conditions. For the [sl] cluster, schwa clusters differ from underlying clusters but are similar to pseudowords. There are more tokens produced with an increase in amplitude of the signal between the two consonants in underlying clusters than in schwa and pseudoword clusters. This increase in amplitude between the two consonants could be attributed to a schwa remnant. However, the fact that this increase is observed more frequently in underlying clusters rather than in schwa clusters refutes this proposal. Interestingly, the [vn] cluster shows the inverse pattern; there are more tokens with an increase in signal amplitude between the two consonants for schwa than for underlying or pseudoword clusters, thus at the position where an underlying schwa could be assumed. Finally, for the tokens involving the cluster [tn], there are more tokens realized with either a burst, a double burst, a noisy/fricative release portion or a silent portion between the two consonants for underlying clusters than for schwa clusters. This difference could be interpreted in terms of the timing between the two consonants: in the schwa-less pronunciation of schwa clusters, the two consonants would be more coarticulated/overlapped compared to underlying clusters, as was found in [10].

3.3. Automatic acoustical analysis

For each group of clusters [fr], [lm], [lv], [mn], [pl],[sl], [tn] and [vn], a CART method ([21]) was used to build a decision tree whose questions best minimize the average acoustical distance between all the members of each leaf of the tree. The criteria used for CART questions were speaker identity, cluster type (schwa / underlying / pseudoword), cluster duration and a random value.

The tree building algorithm selected only the questions that were the most significant among that list. We reported in Table 2 the first question used by the CART for each cluster. The random value was used as the first criterion for all groups except [lm], [lv] and [tn]. For these three clusters, the first classification criterion is the pseudoword condition and the second one is the random value. Thus, for all clusters except [lm], [lv] and [tn], cluster type (schwa / underlying / pseudowords) has no significant impact on the acoustical distance between clusters. For these three clusters, the pseudoword condition differs from the schwa and underlying conditions, but belonging to the schwa or underlying condition

does not affect the acoustical distance between clusters. Overall, these results suggest that clusters from the schwa and underlying conditions do not differ acoustically. Furthermore, some clusters differ in pseudowords compared to the two other conditions.

Table 2: First classification criterion for each cluster.

Cluster	First classification criterion
fr, mn, pl, sl, vn	random
lm, lv, tn	pseudoword

4. Discussion and conclusion

We conducted three acoustic analyses of clusters resulting from schwa deletion in French and compared them to identical underlying clusters in words and identical clusters in pseudowords. We characterized the similarities and differences between schwa and underlying clusters as well as between schwa and pseudoword clusters.

Our comparisons between schwa and underlying clusters in existing words lead to several interesting findings. First, contra to [8] and [9], we showed that cluster duration does not differ between these two cluster types when the variation accounted for by word length, lexical frequency and speech rate is taken into account. According to autosegmental phonology [22], segments and time are assumed to be represented on different tiers, each segment being associated with a temporal unit. In many languages, segment deletion leads to compensatory lengthening: the temporal unit left empty is taken by the preceding consonant, and this consonant is lengthened. Our results suggest that schwa deletion does not lead to compensatory lengthening.

Secondly, we found that acoustical differences between schwa and underlying clusters are rare and not systematic. The automatic analysis suggests that the acoustic distance between the clusters is not affected by condition. In the manual acoustic comparison, most clusters do not differ at all between these two conditions. When they do differ, the differences only concern very few acoustic cues. Furthermore, for very few clusters these differences may be interpreted as signaling remnants of a vowel in the schwa condition whereas in other clusters, they suggest more coarticulation between the two consonants. A possible explanation for this rather complex picture might be that, as shown for cluster duration, other variables affect the acoustic realization of these consonants, such as lexical frequency, word length or speech rate.

Concerning the comparison between underlying clusters and the same clusters in pseudowords, the three analyses give convergent results. They show that the two types of clusters are dissimilar. This finding suggests that it might not be appropriate to use pseudowords instead of existing words in the comparison with schwa deleted clusters.

Finally this study illustrates the importance of respecting several methodological constraints in the investigation of phonetic phenomena like schwa deleted clusters. First, it has demonstrated the need to include various linguistic factors as covariants in the statistical analyses and the importance of careful matching of the linguistic material. Furthermore, it also provides a warning against simply using pseudowords instead of words because the former can behave quite differently than words.

5. References

- [1] Côté, M. H. and Morrison, G., "The nature of the schwa/zero alternation in French clitics: Experimental and non-experimental evidence", *Journal of French Language Studies*, 17: 159-186, 2007.
- [2] Martinet, A., "La nature phonologique d'e caduc", in A. Valdman [Ed], *Papers in linguistics and phonetics to the memory of Pierre Delattre*, The Hague: Mouton, 1972.
- [3] Tranel, B., "Concreteness in generative phonology: Evidence from French", Berkeley: University of California Press, 1981.
- [4] Browman, C. and Goldstein, L., "Articulatory phonology: An overview", *Phonetica*, 49: 155-180, 1992.
- [5] Davidson, L., "Schwa elision in fast speech: Segmental deletion or gestural overlap?", *Phonetica*, 63: 79-112, 2006.
- [6] Dell, F., "Les règles et les sons", Paris : Hermann, 1973.
- [7] Charrette, M., "Conditions on phonological government", Cambridge: Cambridge University Press, 1991.
- [8] Rialland, A., "Schwa et syllabes en français", in L. Wetzels & E. Sezer [Eds], *Studies in Compensatory Lengthening*, Foris Publications : Dordrecht, 1986.
- [9] Lebel, J. G., "Allongement compensatoire de quelques consonnes par suite de la chute du <e> caduc", *Revue de Phonétique Appliquée*, 7 : 53-79, 1968.
- [10] Fougeron C. & Steriade D., "Does deletion of French schwa lead to neutralization of lexical distinctions?", *Proceedings of Eurospeech*, Rhodes, Greece, 943-946, 1997.
- [11] Barnes, J. and Kavitskaya, D., "Phonetic analogy and schwa deletion in French", *Berkeley Linguistics Society*, 28: 39-50, 2002.
- [12] Pluymaekers, M., Ernestus, M. and Baayen, H., "Lexical frequency and acoustical reduction in spoken Dutch", *Journal of the Acoustical Society of America*, 118: 2561-2569, 2005.
- [13] Kuperman, V., Pluymaekers, M., Ernestus, M. and Baayen, H., "Morphological predictability and acoustic duration of interfixes in Dutch compounds", *Journal of the Acoustical Society of America*, 121: 2261-2271, 2007.
- [14] Browman, C. and Goldstein, L., "Gestural syllable position effects in American English", in F. Bell-Bertin & L. Rafael [Eds], *Producing speech: Contemporary issues for Katherine Safford Harris*, New York: American Institute of physics, 1995.
- [15] Kochetov, A., "Syllable position effect and gestural organization: Articulatory evidence from Russian", in L. Goldstein, D. Whalen & C. Best [Eds], *Papers in Laboratory Phonology 8*, Berlin: Mouton de Gruyter, 2006.
- [16] Crystal, H. and House, A., "Articulation rate and the duration of syllables and stress groups in connected speech", *Journal of the Acoustical Society of America*, 88: 101-112, 1990.
- [17] O'Shaughnessy, D., "A study of French vowels and consonants duration", *Journal of Phonetics*, 9: 390-404, 1981.
- [18] New, B., Pallier, C., Ferrand, L. and Matos, R., "Une base de données lexicales du français contemporain sur internet: LEXIQUE", *L'Année Psychologique*, 101 : 447-462, 2001.
- [19] Manuel, S. Y., Shattuck-Hufnagel, S., Huffman, M., Stevens, K. N., Carlson, R. and Hunnicutt, S., "Studies of vowel and consonant reduction", *Proceedings of the 2nd Conference on Spoken Language Processing*, Banff, Canada, 943-946, 1992.
- [20] Black, A. and Taylor, P., "Automatically clustering similar units for unit selection in speech synthesis", *Proceedings of Eurospeech*, Rhodes, Greece, 601-604, 1997.
- [21] Breiman, L., Friedman, J., Olshen R. and Stone C., "Classification and regression trees", Wadsworth & Brooks, Pacific Grove, CA., 1984.
- [22] Goldsmith, J., "Autosegmental phonology", Indiana. MIT PhD Dissertation, 1976.