Journal of Monolingual and Bilingual Speech

JMBS (PRINT) ISSN 2631-8407
JMBS (ONLINE) ISSN 2631-8415

Article

Schwa Deletion in Word-Initial Syllables of Polysyllabic Words: Investigations Using Large French Speech Corpora

Yaru Wu

Laboratoire de Phonétique et Phonologie, CNRS-Sorbonne Nouvelle, France

Martine Adda-Decker

Laboratoire de Phonétique et Phonologie, CNRS-Sorbonne Nouvelle, France

Lori Lamel

LIMSI, CNRS, Université Paris-Saclay, France

Abstract

This study aims to analyse factors that could influence schwa deletion in word-initial syllables of polysyllabic words in continuous French speech. Both phonological and extralinguistic factors were considered: number of consonants, post-lexical context, speech style, sex and profession. Three large corpora covering different speech styles were explored using forced alignment with optional schwa variants. Formal journalistic ESTER corpus, conversational journalistic ETAPE corpus and casual speech NCCFr corpus were used in this study. We observe that schwa tends to be deleted more for 2C-words than for 3C-words. Words preceded by a consonant or a pause tend to prevent schwa deletion whereas words preceded by a vowel tend to facilitate schwa deletion. The less formal the speech style is, the more schwas are deleted. Males tend to delete schwas more frequently than females. Interestingly, journalists tend to delete more schwas than politicians in our data.

Corresponding author

Yaru Wu
Email: yaru.wu@sorbonne-nouvelle.fr

JMBS VOL 2.2 2020 269–289 https://doi.org/10.1558/jmbs.17311

Keywords: schwa deletion, large corpora, forced alignment, speech style, spontaneous speech, spoken French

Introduction

The realization of schwa (an unstable vowel that alternates with ∅) is considered to be one of the most complicated phenomena in French phonology. According to Grammont (1894), the presence or absence of schwa depends on its consonantal context: schwa becomes mandatory when the surface form resulting from potential schwa deletion has three or more consonants in a row (Grammont's three consonant rule). Since then, the phenomena have been studied and discussed by numerous phoneticians and phonologists (e.g. Anderson, 1982; Tranel, 1987; Morin, 1988; Verluyten, 1988; Charette, 1991; Eychenne, 2006; Durand & Eychenne, 2007; Racine, 2008; Lacheret & Lyche, 2008). However, there are few studies of this phenomenon based on an analysis of large corpora of natural continuous speech.

According to previous studies, there are two completely different points of view on schwa deletion in word initial syllables. Some researchers hold the view that since the initial syllable is the position furthest from the main group accent, the schwa is then in a weak position which favours its deletion (Delattre, 1966; Dell, 1973; Côté, 2000). Other researchers support the idea that the initial syllable position preserves against schwa deletion (Hansen, 1994; Scheer, 2000). A third view is proposed by Lyche and Durand (1996) whose results show that words with a schwa in the initial syllable may be actually produced with or without the schwa in French.

Today, the access to large volumes of annotated fluent speech and the use of automatic speech processing tools allow us to further investigate this question using a corpus linguistic approach (Nguyen & Adda-Decker, 2013). The data may be divided into various subsets according to factors potentially conditioning the realization of schwa, such as intra-word context, cross-word context, speech style, speaker gender and speaker role. We are interested in better understanding schwa deletion in the initial syllables of polysyllabic words and in better understanding phonological factors as well as several extralinguistic factors that could influence the absence/presence of schwa.

The influence of style has been much less systematically explored with the same method on large volumes of data. With respect to style, we suppose that mainstream public speech is more conservative with respect to schwa realization than casual speech: more schwas should be realized in broadcast speech than in casual speech since speech must remain intelligible to a large number of people in the former, whereas in casual speech the complicity among speakers might favour deletions. Broadcast news speech, which is primarily prompt

reading, may be expected to be most conservative (a written 'e' in the prompt tends to be pronounced). We further hypothesize that broadcast speech produced without prompting (answers to questions, comments, discussions) be less conservative and that more schwas may be deleted here as compared to news.

Thanks to progresses in speech technology, we are able to investigate schwa deletion using large corpora including a variety of speech styles. The proposed analyses benefit from a methodology which has been derived from automatic speech processing tools.

Corpora

Three large corpora of continuous speech in Metropolitan French were used for this study. Speakers are considered to speak without any marked regional accent. Each corpus corresponds to a different speech style. The ESTER corpus is a formal journalistic speech corpus containing mainly news (Galliano et al., 2006). The second corpus, that we call ETAPE, is composed of journalistic speech with more casual settings, namely conversations and debates. It contains data from the ETAPE corpus (Gravier et al., 2012) with some additional shows from the REPERE corpus (Giraudel et al., 2012). The Nijmegen Corpus of Casual French (NCCFr) is composed of conversational speech between friends (Torreira et al., 2010). The selected material of the ESTER formal speech style corpus contains 85 hours of radio broadcast news shows., which corresponds to a professional reading style. Only Metropolitan French radio stations were used. The less formal speech style corpus ETAPE composes 80 hours of speech. For the casual speech style, we make use of the full NCCFr corpus, corresponding to 36 hours of spontaneous face-to-face conversations between confederates. All corpora were manually annotated with speaker turns, speaker identities (when available) and a careful orthographic transcription representing what the speakers said, that is including spontaneous speech effects such as hesitations, false starts, repetitions, mispronunciations, etc.

Method

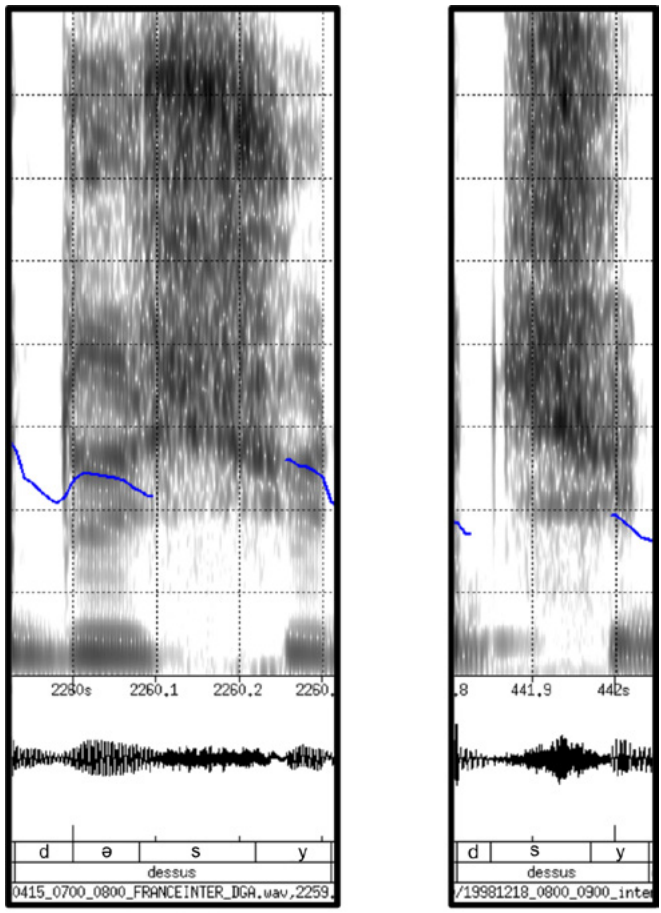
The manually transcribed corpora are processed with the help of automatic forced alignment, to locate the transcribed words in the speech signal. 'Forced' indicates that the system has no choice to modify the word sequence (as an automatic recognition system normally does). However, the system remains free to choose the best matching variant, provided the system is informed about such pronunciation variants.

Speech alignment using variants

The aim of the forced alignment mode is to locate the optimal match between speech segments and sequences of acoustic HMM (statistical hidden Markov model) phone models according to the phonemic pronunciation lexicon. The speech data were automatically segmented and labelled using the LIMSI speech transcription system (Gauvain et al., 2002, 2005) in forced alignment mode. Pronunciation variants were added to the pronunciation lexicon of the system in order to better account for the various productions (with or without schwa) of speakers. As a result, the forced alignment system with variants remains constrained by the word sequence, but is able to choose the most suitable pronunciation variant, given the speaker's realization. In this study, schwa variants were introduced into the pronunciation lexicon of transcription system by adding a variant without schwa for each word containing a schwa in its canonical pronunciation (e.g. *dessus* 'above', /dəsy/ versus /dsy/). Thus, the absence/presence of schwa could be automatically decided via forced alignment, which outputs word and phone labels and boundaries for the speech stretches. For the interspersed non-speech stretches, the system produces silence, respiration and noise labels and boundaries. In what follows (e.g. Table 4), word-token counts do not include labels describing these non-speech stretches. The minimum duration of a segment is 30 ms, corresponding to 3 acoustic frames (Adda-Decker & Lamel, 2000).

Forced alignment with variants may serve to test different hypotheses in phonetics and phonology using big speech data (Adda-Decker et al., 1999). By locating words and finding their best matching variants according to the pronunciation dictionary in large corpora, we are able to quantify the realization of the variant in question. In this study, the deletion or not of schwa was automatically decided using this approach of forced alignment with specific variants. This approach is very different from that used in the PFC project (Durand et al., 2009), on French regional variation, where the absence/presence of schwa was decided through auditory perception.

Figure 1 shows spectrograms of the word dessus (/dəsy/, 'above') with (Figure 1 (a), [dəsy]) and without (Figure 1 (b), [dsy]) schwa in the automatic alignment. A total of two hours of speech were randomly selected from the three corpora and the automatic labelling of presence or absence of schwa in word-initial syllables was verified manually, using both acoustic and perceptual cues. The correct performance of the automatic alignment with respect to the absence/presence of schwa is estimated to be 95% (ESTER 95% versus ETAPE 96% versus NCCFr 93%).



(a) **(b)**

Figure 1. Examples of the output of forced alignment with automatic schwa variant selection and phone segmentation of the word dessus (/dəsy/, above) produced by the same speaker.

Data preparation

Large speech corpora, especially sources corresponding to news and political debates tend to include a large number of proper nouns from many different countries across the world. In order to keep our focus on typical French words, we decided to remove all proper nouns. To this end, we made use of the Lexique380 database (New et al., 2004, 2007), which is a large lexical database developed over the last decades to include the most frequent French word forms (New et al., 2001). Lexique380 was thus used as a reference to filter our data to establish the list of polysyllabic words with schwa in word-initial

syllable. The Lexique380 reference pronunciation was used to quantify the absence and presence of schwa in the alignment pronunciation form (i.e. production of speakers aligned by the LIMSI transcription system; see the examples given in Table 1). Word pronunciations were thus categorized into 'schwa present' (i.e. schwa not deleted) and 'schwa absent' (i.e. schwa deleted).

Table 1. Examples of canonical and aligned pronunciations.

Word Pattern	Orthographic word form	Reference pronunciation (Lexique380)	Aligned pronunciation (LIMSI system)
#CəCV	*semaine* ('week')	səmɛn	səmɛn (ə present) smɛn (ə absent)
#CəCCV	*retraité* ('retired')	rətrɛte	rətrɛte (ə present) rtrɛte (ə absent)

In order to study the influence of the number of surrounding left and right consonants on the absence/presence of schwa, we looked at different word patterns (Table 2). In this study, #CəCV words (e.g., *semaine* /səmɛn/ 'week') and the more complex #CəCCV (e.g., *retraité* /rətrɛte/ 'retired') words were selected since they are the most frequently used word-tokens (≈ 99%) and word-types (≈ 98%) in our corpora. Words like *prenons* /pʁənɔ̃/ (1st person plural form of 'take', #CCəCV) or *restructuration* ('restructuring', #CəCCCV) were discarded, as there were only few word-tokens (< 1%) and word-types in our corpora (< 2%). Only speakers with more than 50 occurrences of word-tokens starting in #CəCV or #CəCCV were included (19305 word-tokens in total). Corresponding information (i.e. aligned pronunciation, contexts of the

Table 2. Word patterns of schwa in word-initial syllables of polysyllabic words.

Nb. of Cs before the second vowel of the word	Nb. of Cs in the 1st syllable		
	1C	2Cs	3Cs
2Cs	#CəCV e.g. semaine /səmɛn/ 'week'	/	/
3Cs	#CəCCV e.g. secrétaire /səkʁetɛʁ/ 'secretary'	#CCəCV e.g. prenons /pʁənɔ̃/ '*1st person* plural form of 'take'	/
4Cs	#CəCCCV e.g.restructuration /ʁəstʁyktyʁasjɔ̃/ 'restructuring'	/	/

words in question, corpus (indicating speech style), speaker identity, sex and socio-professional status) was produced for each selected word-token.

Factors related to schwa deletion

In this section, we describe the different factors that we investigate which are potentially related to the absence/presence of schwa (Grammont, 1894; Durand et al., 2009). We consider not only phonological factors, but also extralinguistic factors. Table 3 gives an overview of the factors studied.

Table 3. Overview of the factors analysed in this study. '#' stands for word boundary; 'V' stands for full vowels (schwa excluded); 'C' represents consonants; 'P' stands for pause.

Category	Factor	Description (example)
Phonological factor	Number of consonants	2C: #CəCV (e.g. *semaine* /səmɛn/, 'week') 3C: #CəCCV (e.g. *retraité* /ʁətʁɛte/, 'retired')
	Post-lexical context (left context)	V# (e.g. *la semaine* /la#səmɛn/, 'the week') C# (e.g. *cette semaine* /sɛt#səmɛn /, 'this week') P# (*Regarde*! /ʁəgaʁd/, 'Look!')
Extralinguistic factor	Speech style	ESTER (formal journalistic speech), broad audience ETAPE (informal journalistic speech), broad audience NCCFr (casual conversational speech), private audience
	Sex	male female
	Profession	journalist politician

With regard to phonological factors, we are particularly interested in the number of consonants resulting from potential schwa deletion: '2C' for #CəCV and '3C' for #CəCCV. According to Grammont's three consonant rule, we suppose that schwa tends to be deleted more in 2C-words (#CəCV) than in 3C-words (#CəCCV).

The post-lexical context (i.e. cross-word left contexts: pause 'P#' versus consonant 'C#' versus vowel 'V#') of our target words was also considered. As a matter of fact, when the post-lexical context is C#, words with #CəCV and #CəCCV patterns get an additional consonant in the immediate neighbourhood (3 or 4 consonants in a row in case of schwa deletion). In line with Grammont's three consonant rule, we further hypothesize that schwa tends to be deleted more when the word in question is preceded by a vowel (i.e. post-lexical context 'V#') than when it is preceded by a consonant (i.e. post-lexical context 'C#'), given that the latter would add a consonant to the consonant cluster resulting from schwa deletion. Moreover, when the left

post-lexical context is a pause ('P#'), words in question (i.e. polysyllabic words containing a schwa in initial syllable) are at the start of the intonational phrase (IP initial). In this case, the post-lexical context may reveal the influence of prosodic position on schwa deletion. Côté (2000) observed that the schwa in the initial syllable of a polysyllabic word preceded by a pause (e.g. *demandez-la* 'request it' /dəmɑ̃de#la/) is optional. Our assumption is that schwa might be deleted less when preceded by a pause than in other positions, as it may lead to non-admissible onset consonant clusters.

Speech style, sex and socio-professional status were considered as extralinguistic factors in this study (Labov, 1976; Encrevé, 1977). The speech styles are formal journalistic speech (ESTER), informal journalistic speech (ETAPE), and casual speech between friends (NCCFr). We suppose that schwa deletion occurs more frequently in casual speech than in journalistic speech. As for journalistic speech, we expect more schwas in the formal journalistic speech (ESTER) than in the informal journalistic speech (ETAPE), as ETAPE contains more interactive speech with conversations and debates. Casual speech (NCCfr) is expected to have the highest schwa deletion rate among the three speech styles. Sex differences were also considered in this study. Simpson (2000) and Adda-Decker and Lamel (2005), among others, showed that male speakers, overall, tend to articulate with less precision than female speakers. In this study, we question whether the effect of sex persists when it comes to schwa deletion and we looked into whether male speakers tend to produce fewer schwas than female speakers.

Journalists are known to have a particular speech style when speaking in public as news presenters (see Boula de Mareüil et al., 2012 for example); as public speakers, politicians are also known for their marked speech style (e.g. Duez, 1997). We are therefore also interested in comparing the production of speakers of the two professions, given that both groups of speakers are public speakers with characteristic speech styles. More precisely, we aim to have a better understanding of whether or not journalists and politicians behave similarly as far as schwa deletion is concerned. We may suppose that journalists tend to delete more schwas than politicians since the former are more conscious about allocated time slots. The effect of socio-professional status/profession (i.e. journalist versus politician) was investigated only using the less formal journalistic corpus (ETAPE), given the high participation of both journalists and politicians in this corpus.

Results

In this section, we first present quantitative data (word-tokens and word-types) concerning words containing at least one schwa in the canonical form in

our corpora. This data quantification is followed by analyses of inter- and intra-speaker variation with respect to schwa deletion. Finally, schwa deletion rates are reported as a function of the different factors of variation.

Frequencies of schwas by speech style

In this section, we present the frequencies of words (with or without schwa) in different speech styles in order to have a general idea about polysyllabic words containing a schwa in initial syllable with respect to other words used in the corpora.

Table 4 gives an overview of the number of word-tokens and word-types in each speech style, excluding proper nouns. While there are 12–16% of the word-tokens that have at least one schwa (16% for the ESTER corpus, 15% for the ETAPE corpus and 12% for the NCCFr corpus), only 2–5% of the words in the corpora are polysyllabic words containing at least one schwa (5% for the ESTER corpus, 4% for the ETAPE corpus and 2% for the NCCFr corpus) and about half of these are polysyllabic words with a schwa in the initial syllable.

Figure 2 illustrates the percentage of words containing at least one schwa in their canonical forms (according to Lexique380) in our pooled corpora. Among the word-tokens containing at least one schwa, 71% are monosyllabic function words, even though there are only 9 word-types containing schwa (i.e. *de* ('of'), *le* ('the'), *me* ('me'), *te* ('you'), *se* ('himself'), *ce* ('this'), *ne* ('not'), *que* ('that'), *je* ('I')). Among the 29% of the polysyllabic word-tokens containing schwa(s), about half of these are polysyllabic words with schwa in the word-initial syllable.

Table 4. Number of word-tokens and word-types in each speech style (i.e. formal/ESTER, less formal/ETAPE, casual speech/NCCFr): (i) total; (ii) words with schwa; (iii) polysyllabic words with schwa; (iv) polysyllabic words with schwa in word-initial syllables.

Corpus	Tokens			Types			Examples
	ESTER	ETAPE	NCCFr	ESTER	ETAPE	NCCFr	
(i) Total number of words in corpus	774691	729312	427635	23446	21605	11532	
Including:							
(ii) words with schwa (Cə)	122365	110865	53089	2769	2577	1229	*ce* ('this') *demain* ('tomorrow') *simplement* ('simply')
(iii) polysyllabic words with schwa (CəC)	35229	28406	10159	2759	2567	1220	*demain* ('tomorrow') *simplement* ('simply')
(iv) polysyllabic words with schwa in initial syllables (#CəC)	17391	13969	5526	1022	983	544	*demain* ('tomorrow') *secret* ('secret')

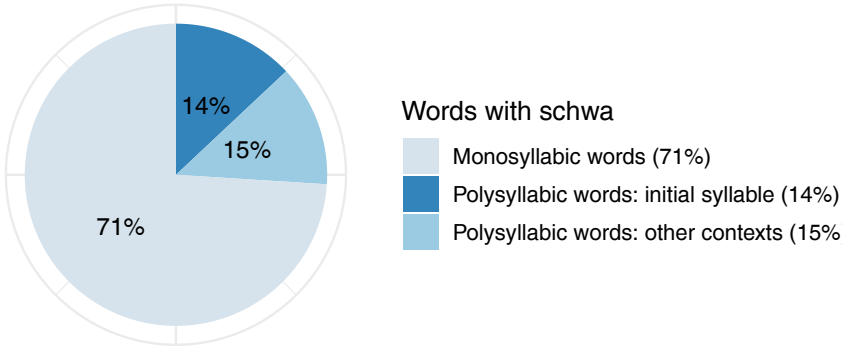


Figure 2. Schwa distribution in word-tokens containing at least one schwa in the canonical form (all three corpora pooled). Words are categorized into monosyllabic words, polysyllabic words with schwa in word-initial syllable and polysyllabic words with schwa in other positions.

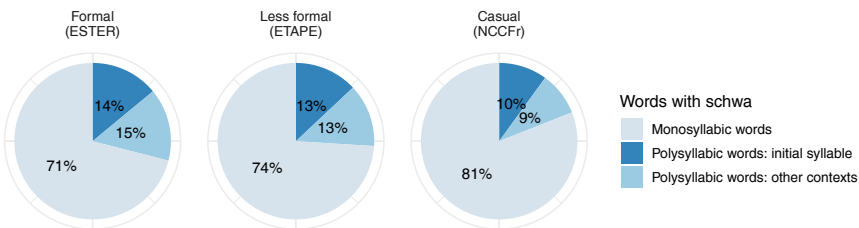


Figure 3. Schwa distribution of word-tokens containing at least one schwa in the canonical form as a function of the three speech styles: Formal journalistic (ESTER), Less formal journalistic (ETAPE), Casual (NCCFr).

The percentage of word-tokens containing at least one schwa in its canonical form as a function of speech style is presented in Figure 3. As for the pooled corpora, monosyllabic function words are by far the most frequent in each individual corpus (71%, 74% and 81% for ESTER, ETAPE and NCCFr respectively). Interestingly, more monosyllabic function words are observed in casual speech than in journalistic speech. With regard to content words, we tend to observe slightly more polysyllabic words in formal speech than in casual speech (29%, 26% and 19% for ESTER, ETAPE and NCCFr respectively).

Table 5 shows the number of word-tokens and word-types for polysyllabic words with schwa in word-initial syllables: words starting with #CəCV and #CəCCV patterns and 'others'. As can be expected, the table shows that there are more word-tokens and word-types for the simpler #CəCV pattern than for the more complex #CəCCV pattern in all corpora. With respect to speech style, it can also be seen that 3C-words (word-tokens) appear more often in journalistic speech (28.2% and 26.3% for ESTER and ETAPE respectively)

Table 5. Number and proportion of word-tokens and word-types for polysyllabic words with schwa in word-initial syllables as a function of speech style (formal/ESTER, less formal/ETAPE, casual speech/NCCFr). Words are separated into three classes according to word-initial syllable structure (#CəCV #CəCCV, others).

	Tokens				Types			
	ESTER	**ETAPE**	**NCCFr**	**Total**	**ESTER**	**ETAPE**	**NCCFr**	**Total**
#CəCV	12385 (71.2%)	10178 (72.9%)	4643 (84.0%)	27206 (73.8%)	761 (74.5%)	721 (73.3%)	413 (75.9%)	1064 (75.3%)
#CəCCV	4898 (28.2%)	3678 (26.3%)	860 (15.6%)	9436 (25.6%)	237 (23.2%)	243 (24.7%)	120 (22.1%)	325 (23.0%)
Others	108 (0.6%)	113 (0.8%)	23 (0.4%)	244 (0.7%)	24 (2.3%)	19 (2.0%)	11 (2.0%)	24 (1.7%)
Total	17391 (100%)	13969 (100%)	5526 (100%)	36886 (100%)	1022 (100%)	983 (100%)	544 (100%)	1413 (100%)

than in casual speech between friends (15.6% for NCCFr corpus). That is, speakers tend to use simpler words in casual communication situations in contrast to formal speech.

Inter- and intra-speaker variability

Before discussing the influence of different factors on whether or not schwa is deleted, it is interesting to present some examples of inter- and intra-speaker variability.

First of all, there are probably no two individuals who have pronunciations which are identical in every respect. In order to control for the variability linked to the number of consonants, speech style, sex and profession, we decided to focus on 2C-words produced by four male journalists from the ETAPE corpus. These four speakers were chosen since they had the most data of interest in the ETAPE corpus. Figure 4 shows the schwa absent/present rates (x-axis) for each of the four speakers (AA, LN, OT, CR). The observed schwa deletion rate is much higher for speaker AA (≈64%) than for speaker CR (≈25%). Although both are male journalists, speaker AA deletes more schwas in 2C-words than speaker CR in the corpus of conversational journalistic speech. The schwa deletion rates in 2C-words of the two other male journalists (speakers LN and OT) lie in between the two extremes.

These examples illustrate that people of the same gender and the same profession do not necessarily produce the same word patterns in the same way, even in a given communication situation.

In addition, speakers do not always speak in the same way in the similar communication situation on the same word-type. Figure 5 gives the absence

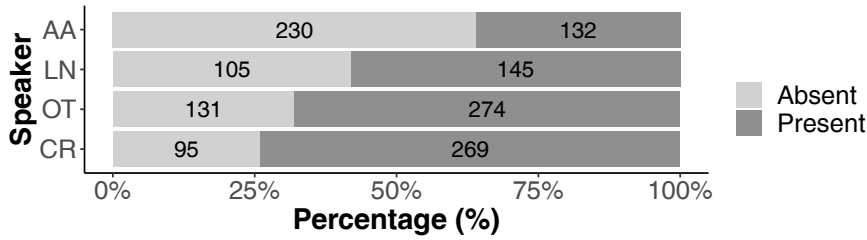


Figure 4. Schwa absent/present rates in 2C-words for four male journalists in the ETAPE corpus. Each bar also indicates the number of absent versus realized schwas.

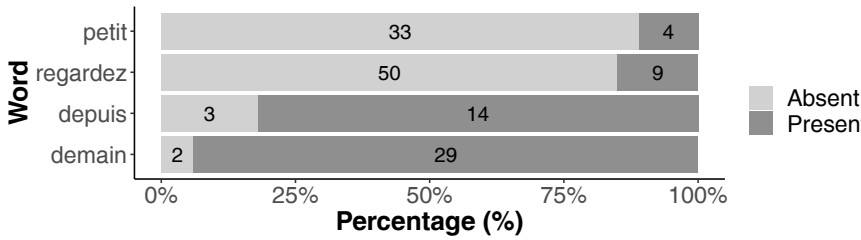


Figure 5. Schwa absent/present rates for four 2C-words (*demain* 'tomorrow', *depuis* 'since', *regardez* 'look', and *petit* 'small') spoken by journalist AA in the ETAPE corpus. Each bar also indicates the number of absent versus realized schwas.

rate of schwa for four 2C-words (#CəCV: *demain* 'tomorrow', *depuis* 'since', *regardez* 'look', and *petit* 'small') produced by one of the journalists (speaker AA, ETAPE corpus). Schwa absent/present rates are seen to vary substantially across the lexical items, even though the speaker and word pattern are the same. This may simply be due to the fact that four different word-types are involved. Different factors could explain these differences: the influence of the post-lexical context (here not controlled to keep enough tokens per speaker) and the possible influence of phonetic convergence (Aubanel, 2011), for instance.

Phonological and extralinguistic factors of schwa deletion

In this section, we propose an analysis of absence/presence of schwa considering both phonological and extralinguistic factors. The investigated phonological factors concern the number of word-initial consonants in a row in case of schwa deletion (2C, 3C) and the post-lexical context, namely whether the preceding word ends in a consonant or a vowel, or whether the left context is a pause. The investigated extralinguistic factors are speech style (formal journalistic speech (ESTER); less formal journalistic speech (ETAPE); casual speech (NCCFr)), sex (male, female) and socio-professional status. The latter factor is limited to the ETAPE corpus of public speech, for which two well represented

professions (journalists and politicians) were selected. The results with respect to profession are shown in a separate figure.

To validate the proposed analyses, generalized linear mixed models (GLMM) were employed (McCulloch & Neuhaus, 2001) using the package lme4 (Bates et al., 2015) in R (R Development Core Team, 2013). The effects of number of consonants (in the surface form resulting from potential schwa deletion), post-lexical context, speech style and sex are considered within the same model. To avoid collinearity related problems, a separate model was used to account for the effect of profession.

The 'number of consonants' (2C for #CəCV and 3C for #CəCCV, reference: 2C), 'post-lexical context' (P#, C# or V#, reference: V#), 'speech style' (ESTER, ETAPE or NCCFr, reference: ESTER), and 'sex' (male or female, reference: female) were used as fixed effects. The following random terms were considered in the model: a random intercept per speaker and one per word, a by-speaker slope for the effects of number of consonants and post-lexical context and a by-word slope for the effects of post-lexical context, speech style and sex. Model-based post-hoc tests were performed to obtain level information for each fixed effect with more than two levels (i.e. post-lexical context and speech style).

A separate model is conducted to analyse the effect of profession. Profession (journalist versus politician, reference: journalist) was included as a fixed effect. Post-lexical context and sex were included as covariates. Random terms were included in the model: a random intercept per speaker and one per word, a by-speaker slope for the effect of post-lexical context and a by-word slope for the effects of profession and sex.

Figure 6 illustrates the influence of phonological factors on schwa deletion in several contrasting conditions: the number of consonants (2C versus 3C) and post-lexical/left-boundary contexts. From top to bottom, 'the three left post-lexical contexts (V#, C# and P#) are shown for the '2C' and '3C' word forms. Schwas tend to be realized more for 3C-words than for 2C-words, which is in line with Grammont's three consonant rule. With respect to the post-lexical context, when the word in question is preceded by a vowel (V#), schwa tends to be absent more often than in the C# and P# contexts. Pause ('P#') and consonant (C#) contexts both seem to favour maintaining schwa. The deletion rate is consistently slightly lower for P# than C#.

Results from GLMM suggest that the realization or not of schwa is influenced by number of consonants, post-lexical context, speech style and sex. '3C' (#CəCCV) has a significant positive effect on the retention of schwa with respect to that observed in '2C' (#CəCV) (log odds ratio = 0.6902, $|Z| = 4.425$, $p < 0.001$). Therefore, there is a tendency to observe more schwas for '3C' words than '2C' ones. Both 'C#' and 'P#' contexts have a significant positive

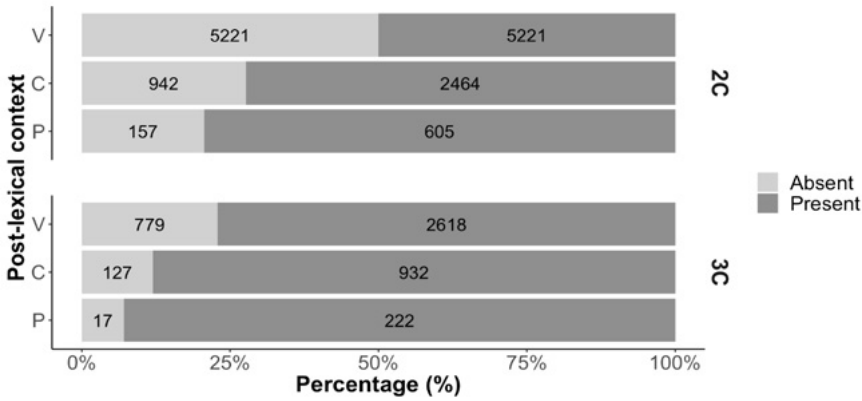


Figure 6. Combined view of schwas absent/present rates for two different factors: post-lexical contexts (V# versus C# versus P#, see left y-axis) and number of consonants (2C versus 3C, see right y-axis). Each bar also indicates the number of absent versus realized schwas.

effect on the presence of schwa with respect to that observed in 'V#' (C#: log odds ratio = 1.1520, $|Z| = 8.590$, $p < 0.001$; P#: log odds ratio = 1.4416, $|Z| = 5.845$, $p < 0.001$). This means that there are fewer schwas when the post-lexical context is 'V#' than when the context is 'C#' or 'P#'. Post-hoc tests based on the model show that the influence of the context 'V#' is significantly different from that of the contexts C# ($p < 0.001$) and 'P#' ($p < 0.001$). Results on P# context also suggests a potential influence of prosodic position on schwa deletion, given that the P# context includes words at the beginning of an intonation phrase (IP), in contrast to the C#/V# contexts which cover words that are not located in such a position. That is, schwa in word-initial syllables of polysyllabic words at the beginning of an IP is less likely to be delated than words that are inside of an IP.

Figure 7 shows the influence of speech style and speakers' sex differences on the absence and presence of schwa. The bar plot indicates the relative percentages of deleted versus realized schwas for the different factors. As in the previous figures, the number of occurrences is also reported. Important differences can be seen with regard to speech style. Whereas in the most formal speech (ESTER), the schwa deletion rate in word-initial syllable of polysyllabic words is rather low (about 20%), this rate is above 70% for casual interactions between friends (NCCFr). Schwas are seen to be deleted more often in conversational broadcast data (ETAPE) than in the formal journalistic news corpus (ESTER). Males are seen to delete more schwas than females for all speech styles. Statistical analyses based on GLMM confirm that the probability of observing a schwa is significantly lower in both the ETAPE corpus (log odds

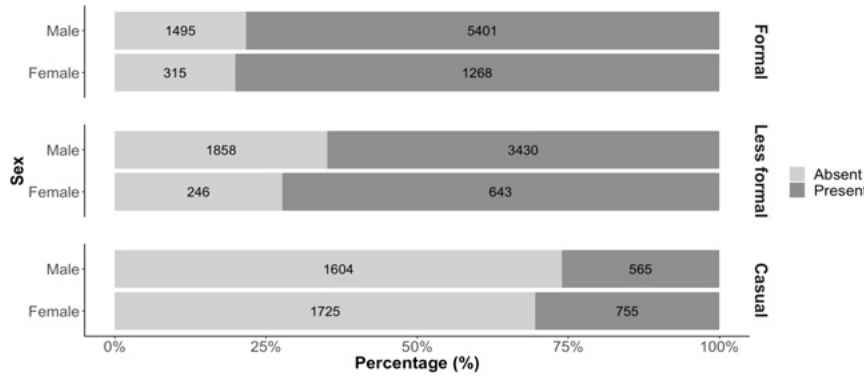


Figure 7. Combined view of schwas absent/present rates for two factors: speech style (from top to bottom: formal/ESTER, less formal/ETAPE, casual speech/NCCFr) and sex (male versus female) on y-axis. Each bar also indicates the number of absent versus realized schwas.

ratio = –0.6373, $|Z| = 4.924$, $p < 0.001$) and in the NCCFr corpus (log odds ratio = –2.9299, $|Z| = 19.679$, $p<0.001$) with respect to that observed in the ESTER corpus. Post-hoc tests show that the three corpora differ significantly from each other as far as schwa deletion is concerned (ESTER versus ETAPE: $p < 0.001$; ETAPE versus NCCFr: $p < 0.001$; ESTER versus NCCFr: $p < 0.001$). Finally, the probability of observing a schwa shows a slight, nonetheless significant decline for male speakers with respect to that observed for female speakers (log odds ratio = –0.2864, $|Z| = 1.966$, $p < 0.05$). This indicates that male speakers tend to produce fewer schwas than female speakers.

The influence of socio-professional status is shown in Figure 8. Interestingly, the plot shows that journalists tend to produce fewer schwas than politicians. Statistical analyses based on GLMM confirms that the probability of observing a schwa increases significantly for politicians (log odds ratio = 0.8118, $|Z| = 2.837$, $p < 0.01$) with respect to that observed for journalists.

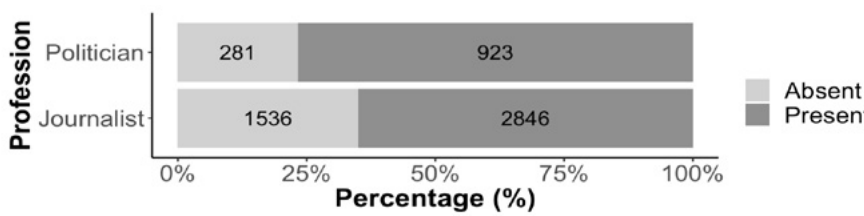


Figure 8. Combined view of Schwas absent/present rates for the profession of the speakers (journalist versus politician). Each bar also indicates the number of absent versus realized schwas.

Discussion and conclusions

In this study, schwa deletion in French was analysed using large corpora with various speech styles and tools from automatic speech recognition. A quantitative analysis of a total of 200 hours of transcribed speech data with respect to potential schwa sites reveals that a large amount (about 70%) of such sites are located in monosyllabic function words, such as the articles *le* ('the') and *de* ('of'). The remaining 30% of schwa sites is almost equally distributed in word-initial versus word-internal position of polysyllabic words.

This study focused on schwa in word-initial syllables, namely #CəCV (2C) and #CəCCV (3C) word patterns which account for ≈ 99% of all word tokens with schwa in the word-initial syllable, and explored different factors that may influence absence/presence of schwa in fluent speech. Both phonological and extralinguistic factors, five in total, were investigated: number of neighbouring consonants, post-lexical context, speech style, sex and profession.

Our results on the number of consonants in the surface form resulting from potential schwa deletion ('2C' for #CəCV and '3C' for #CəCCV) are consistent with Grammont's three consonant rule (1894), which states that sequences of three consonants in a row are avoided by maintaining (or inserting) a schwa. Schwas are more easily deleted in 2C-words than in 3C-words: in our corpora 43% versus 20% of schwas deleted respectively. However, given the measured schwa deletion rate of 20% in the '3C' case, our results suggest that the three consonant rule is a principle or trend rather than a law.

The influence of the post-lexical context on whether or not a segment is realized has rarely been studied before at a large scale. Since we were able to conduct our work on large amounts of data, we were able to look at the post-lexical context to investigate its potential impact of the neighbouring word. In particular, we defined the post-lexical context as the preceding word (or pause), and more narrowly the last segment of the preceding word. We were thus able to test the influence of the number of consonants within polysyllabic words including the left post-lexical context. We analysed the left post-lexical contexts according to the nature of its last segment: V# (vowels), C# (consonants) and P# (pauses). Schwa in the initial syllable was found to be less frequently deleted when the polysyllabic word is preceded by a consonant (C#CəCV or C#CəCCV, which would result in 3 or even 4 consonants in a row in the surface form in case of schwa deletion) than when it is preceded by a vowel (V#). These results highlight that schwa deletion depends not only on the word pattern itself, but also on its preceding post-lexical context, which has been shown to be of major influence (43%, 24% and 17% for V#, C# and P# respectively). Schwa deletion is therefore favoured if the preceding word ends with a vowel. A possible explanation for this result is that in this case

(post-lexical context V#), some of the consonants may be redistributed in the coda (empty) of the previous syllable. In addition, post-lexical P# context has a special status: it provides prosodic information about the syllable in question. Our results on the influence of the P# context could therefore be interpreted as follows: when the initial syllable schwa of polysyllabic words is at the beginning of a prosodic unit, it tends to be maintained. Results on post-lexical context show that it is essential to take it into consideration while analysing continuous speech.

Cases of schwa deletion in sequences involving three or four consonants are not negligible in our data; in particular in the casual conversational speech corpus (NCCFr). This may be explained by a process of resyllabification in speech production. Indeed, these consonantal sequences resulting from schwa deletion may be considered as two consonantal sequences distributed over the coda position of the preceding syllable and the onset position of the following syllable. Post-analyses verification show that these pairs of consonantal sequences are almost all (> 99.6%) allowed by the resyllabification of the French language and by the phonotactic constraints of French, according to the syllabification rules of spoken French (Wioland, 1985; Adda-Decker et al., 2005) and the Sonority Sequencing Principle (Clements, 1990; Côté, 2004).

The presented results also demonstrate the importance of extralinguistic factors. Among the investigated factors, speech style has proven to have a crucial impact, whereas the impact remains relatively low for speaker sex and profession. With respect to style, schwa tends to be deleted much more often in casual speech (NCCFr, ≈ 70%) than in journalistic speech (≈ 20% for ESTER and ≈ 35% for ETAPE). Furthermore, schwa tends to be deleted more in conversational journalistic speech (ETAPE) than in the more formal journalistic speech (ESTER). Our results on the influence of speech style show that the less formal the corpus is, the more the schwa is deleted; which is consistent with what has been proposed by Léon (1971), Adda-Decker and Lamel (1999) and Brognaux and Drugman (2014) for other linguistic phenomena. As for the influence of sex difference, male speakers tend to delete slightly more schwas than female speakers, which is in line with previous findings on differences between male and female speakers' articulation (Simpson, 2000; Adda-Decker & Lamel, 2005).

Our analyses on public speaking professionals (journalists versus politicians; a previously little studied contrast), suggest different strategies related to professional particularities within a particular speaking style (i.e. informal journalistic style) as far as schwa deletion is concerned. In our data, journalists tend to produce fewer schwas than politicians in general, which might be related to the fact that journalists are very conscious of tightly allocated time slots.

Our results on extra-linguistic factors suggest that public speakers in journalistic speech corpora (ESTER and ETAPE) articulate more accurately than young adults in a casual communication situation (NCCFr) as far as schwa deletion is concerned. It would be interesting to extend our study to other reduction phenomena and compare professional speakers speaking in a formal setting with speakers conversing with their friends in a casual one. Indeed, public speakers such as journalists and presenters are trained to speak at a high articulation rate and to articulate with high precision within limited time as their speech is primarily addressed to a wide (and generally hidden) public in a somewhat disconnected mode. Friends who talk to each other may also speak at a high rate; but as they share a common conversational context, they may articulate less carefully without loss of intelligibility.

In conclusion, this study provides a fine-tuned quantitative description of schwa deletion in word-initial syllables of polysyllabic words in French. The analysis of large volumes of speech from various communication settings was enabled thanks to speech technology tools from automatic speech recognition. In particular, the applied method for phonetic and phonological analyses using forced alignment with schwa-specific pronunciation variants was found to be efficient and reliable. Various linguistic hypotheses could thus be explored and quantified in contrastive communication settings as represented by the different large speech corpora. Overall, the results reveal a complex interplay of both linguistic and sociolinguistic factors. The produced results contribute to a more in-depth understanding of these different variation factors on schwa deletion.

References

Adda-Decker, M., Boula de Mareüil, P., Adda, G. & Lamel, L. (2005). Investigating syllabic structures and their variation in spontaneous French. *Speech communication*, *46*(2), 119–39. https://doi.org/10.1016/j.specom.2005.03.006

Adda-Decker, M., Boula de Mareüil, P. & Lamel, L. (1999). Pronunciation variants in French: schwa & liaison. In *Proceedings of the XIVth International Congress of Phonetic Sciences* (pp. 2239–42). Berkeley, CA: University of California, Department of Linguistics.

Adda-Decker, M. & Lamel, L. (1999). Pronunciation variants across system configuration, language and speaking style. *Speech Communication*, *29*(2–4), 83–98. https://doi.org/10.1016/s0167-6393(99)00032-1

Adda-Decker, M. & Lamel, L. (2000). The use of lexica in automatic speech recognition. In F. Van Eynde & D. Gibbon (eds), *Lexicon development for speech and language processing* (pp. 235–66). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-010-9458-0_8

Adda-Decker, M., & Lamel, L. (2005). Do speech recognizers prefer female speakers? In *Ninth European conference on speech communication and technology*. Lisbon.

Anderson, S. (1982). The Analysis of French Schwa. *Language*, *58*, 534–73.

Aubanel, V. (2011). Variation phonologique régionale en interaction conversationnelle Doctoral dissertation, Univ. Aix-Marseille 1.

Bates, D., Mächler, M., Bolker, B. M. & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Boula de Mareüil, P., Rilliard, A. & Allauzen, A. (2012). A diachronic study of initial stress and other prosodic features in the French news announcer style: corpus-based measurements and perceptual experiments. *Language and Speech*, *55*(2), 263–93. https://doi.org/10.1177/0023830911417799

Brognaux, S. & Drugman, T. (2014). Variations phonétiques: Impact de la situation de communication. *Nouveaux Cahiers de Linguistique Francaise*, *31*, 335.

Charette, M. (1991), *Conditions on phonological government*. Cambridge: Cambridge University Press.

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. *Papers in Laboratory Phonology*, *1*, 283–333. https://doi.org/10.1017/cbo9780511627736.017

Côté, M. H. (2000). Consonant cluster phonotactics: A perceptual approach. Doctoral dissertation, Massachusetts Institute of Technology.

Côté, M. H. (2004). Consonant cluster simplification in Québec French. *Probus*, *16*(2), 151–201. https://doi.org/10.1515/prbs.2004.16.2.151

Delattre, P. (1966). *Studies in French and comparative phonetics*. The Hague: Mouton.

Dell, F. (1973). *Les règles et les sons* (Vol. 25). Paris: Hermann.

Duez, D. (1997). Acoustic markers of political power. *Journal of Psycholinguistic Research*, *26*(6), 641–54.

Durand, J. & Eychenne, J. (2007). Remarks on schwa deletion and epenthesis in French. *Proceedings of the 5th Journées d'Etudes Linguistiques* (pp. 89–94). Nantes: Laboratoire de Linguistique de Nantes.

Durand, J., Laks, B. & Lyche, C. (2009). *Phonologie, variation et accents du Français*. Paris: Hermès.

Encrevé, P. (1977). Présentation: linguistique et sociolinguistique. *Langue française*, *34*(1), 3–16. https://doi.org/10.3406/lfr.1977.4814

Eychenne, J. (2006). Aspects de la phonologie du schwa dans le français contemporain. Doctoral dissertation, Université de Toulouse-Le Mirail, France.

Galliano, S., Geoffrois, E., Gravier, G., Bonastre, J. F., Mostefa, D. & Choukri, K. (2006). Corpus description of the ESTER evaluation campaign for the rich transcription of French broadcast news. *Proceedings of LREC*, *6*, 315–20.

Gauvain, J. L., Adda, G., Adda-Decker, M., Allauzen, A., Gendner, V., Lamel, L. & Schwenk, H. (2005). Where are we in transcribing French broadcast news? In *Ninth European conference on speech communication and technology* (pp. 1665–8). Lisbon. https://doi.org/10.1109/icassp.1997.596002

Gauvain, J. L., Lamel, L. & Adda, G. (2002). The LIMSI broadcast news transcription system. *Speech Communication*, *37*(1), 89–108. https://doi.org/10.1016/s0167-6393(01)00061-9

Giraudel, A., Carré, M., Mapelli, V., Kahn, J., Galibert, O. & Quintard, L. (2012). The REPERE Corpus: a multimodal corpus for person recognition. *Proceedings of LREC*, *11*, 1102–7.

Grammont, M. (1894). Le patois de la Franche-Montagne et en particulier de Damprichard (Franche-Comté). IV: La loi des trois consonnes. *Mémoires de la Société de linguistique de Paris*, *8*, 53–90. https://doi.org/10.2143/bsl.90.1.2002532

Gravier, G., Adda, G., Paulson, N., Carré, M., Giraudel, A. & Galibert, O. (2012). The ETAPE corpus for the evaluation of speech-based TV content processing in the French language. *Proceedings of LREC*, *12*, 114–18.

Hansen, A. B. (1994). Étude du e caduc: stabilisation en cours et variations lexicales. *Journal of French Language Studies*, *4*(1), 25–54. https://doi.org/10.1017/s0959269500001964

Labov, W. (1976). *Sociolinguistique*. Paris: Editions de Minuit.

Lacheret, A & Lyche, C. (2008). Looking at French schwa in initial position through the glasses of prosody. Retrieved from www.isca-speech.org/archive/sp2008/sp08_523.html.

Léon, P. R. (1971). *Essais de phonostylistique*, vol. 4. Montréal: Didier.

Lyche, C. & Durand, J. (1996). Testing government phonology ou pourquoi le choix du schwa. *Current Trends in Phonology: Models and Methods*, *2*, 443–71.

McCulloch, C. E. & Neuhaus, J. M. (2001). *Generalized linear mixed models*. Chichester: John Wiley & Sons.

Morin, Y. C. (1988). *Explaining schwa in French*. Occasional Papers of the University of Essex: French Sound Patterns Changing Perspectives. Colchester: University of Essex.

New, B., Pallier, C., Ferrand, L. & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE™//A lexical database for contemporary French: Lexique. *L'année psychologique*, *101*(3), 447–62. https://doi.org/10.3406/psy.2001.1341

New, B., Pallier, C., Brysbaert, M. & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 516–24. https://doi.org/10.3758/bf03195598

New, B., Brysbaert, M., Veronis, J. & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*, *28*(4), 661–77. https://doi.org/10.1017/s014271640707035x

Nguyen, N. & Adda-Decker, M. (2013). *Méthodes et outils pour l'analyse phonétique des grands corpus oraux*. Paris: Hermès-Lavoisier.

Racine, I. (2008). Les effets de l'effacement du Schwa sur la production et la perception de la parole en français. Doctoral dissertation, University of Geneva, Switzerland.

R Development Core Team (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Scheer, T. (2000). L'immunité de schwa en début de mot. *Langue française*, *126*, 113–26. https://doi.org/10.3406/lfr.2000.992

Simpson, A. P. (2000). Gender-specific differences in the articulatory and acoustic realization of interword vowel sequences in American English. In *5th Seminar on Speech Production: Models and Data* (pp. 209–12). Kloster Seeon: Elsevier Science.

Torreira, F., Adda-Decker, M. & Ernestus, M. (2010). The Nijmegen corpus of casual French. *Speech Communication*, *52*(3), 201–12. https://doi.org/10.1016/j.specom.2009.10.004

Tranel, B. (1987). French schwa and nonlinear phonology. *Linguistics*, 25(5), 845–66. https://doi.org/10.1515/ling.1987.25.5.845

Verluyten, S. P (ed.). (1988). *La phonologie du schwa français*, vol. 16. Amsterdam: John Benjamins Publishing.

Wioland, F. (1985). *Les structures syllabiques du français: fréquence et distribution des phonèmes consonantiques, contraintes idiomatiques dans les séquences consonantiques*, vol. 31. Paris: Editions Slatkine.