Is there only one “fenêtre” in the production lexicon?

On-line evidence on the nature of phonological representations of pronunciation variants for French schwa words

Audrey Bürki

Laboratoire de Psycholinguistique Expérimentale

University of Geneva

40, Bd Pont d’Arve

1201 Geneva, Switzerland

Mirjam Ernestus

Radboud University Nijmegen &

Max Planck Institute for Psycholinguistics

P.O. Box 310

6500 AH Nijmegen, The Netherlands

Ulrich H. Frauenfelder

Laboratoire de Psycholinguistique Expérimentale

University of Geneva

40, Bd Pont d’Arve

1201 Geneva, Switzerland

Formation Universitaire A Distance

Centre d'études de Sierre

TechnoArk, 5

Case Postale 218

3960 Sierre, Switzerland

Abstract

This study examines whether the production of words with two phonological variants involves single or multiple lexical phonological representations. Three production experiments investigated the roles of the relative frequencies of the two pronunciation variants of French words with schwa: the schwa variant (e.g., [fənɛtʀ]) and the reduced variant (e.g., [fnɛtʀ]). In two naming tasks and in a symbol-word association learning task, variants with higher relative frequencies were produced faster. This suggests that the production lexicon keeps a frequency count for each variant and hence that schwa words are represented in the production lexicon with two different lexemes. In addition, the advantage for schwa variants over reduced variants in the naming tasks but not in the learning task and the absence of a variant relative frequency effect for schwa variants produced in isolation support the hypothesis that context affects the variants' lexical activation and modulates the effect of variant relative frequency.

Introduction

Many words in connected speech appear in a non-canonical form (e.g., Johnson 2004). Despite this fact, it is only during the last two decades that psycholinguistic studies of speech comprehension have gone beyond studying canonical speech and have begun to examine how listeners recognize non-canonical variants of words. Findings on assimilation (e.g., Gaskell & Marslen-Wilson, 1996, 2001; Snoeren, Segui & Halle, 2008), nasal flap (e.g., Ranbom & Connine, 2007) and schwa deletion (e.g., Connine, Ranbom & Patterson, 2008; Kuijpers, van Donselaar & Cutler, 1996; Racine & Grosjean, 2000, 2005; Spinelli & Gros-Balthazard, 2007) have provided a deeper understanding of the recognition of this everyday form of speech. For production, in contrast, a similar shift in research has not yet taken place. Our current knowledge of how words are represented in the lexicon and encoded during production comes from experiments using canonical word forms only. What we know about phonological variants comes essentially from corpus analyses and acoustic studies. Although recent corpus studies have started to address these questions (see for instance Bell, Brenier, Gregory, Girand & Jurafsky, 2009), so far they have only provided little and circumstantial information about the nature of the lexical representations of words with several variants and about the mechanisms (including their time course) underlying the production of such variants. For more direct information, on-line experimental data are needed. The aim of this work is to provide such data.

Most models of speech production and comprehension can be situated along a continuum with respect to their assumptions about the mental lexicon. Traditional psycholinguistic models are heavily influenced by generative grammar (Chomsky & Halle, 1968), in which words are generally assumed to have only one lexical representation, with

their other pronunciation variants being computed by means of phonological or phonetic rules.

For instance, many authors assume that French schwa is present in the underlying representation of words and may be deleted via a phonological rule (e.g., Dell, 1985; Rialland, 1986). Thus, a word such as *fenêtre* 'window' is stored as /fənɛtʀ/ and a rule of schwa deletion creates [fnɛtʀ]. Within this same framework of a single underlying representation and phonological rules[1], other authors have proposed the inverse, that is, an underlying representation without schwa and a phonological rule of epenthesis (Côté & Morrison, 2007; Tranel, 1981). In both of these single underlying representation accounts the choice of the underlying representation (with or without schwa) is based on linguistic principles. We can imagine a third single representation account, suggested by an anonymous reviewer, in which words differ in whether the schwa or the reduced variant is stored: the mental lexicon contains only a lexical representation for the most frequent variant for a given word.

Further along the continuum, we find the models assuming one abstract lexical representation for each pronunciation variant of a word. We will name such models *abstract variants models*. In these models, there would be, for instance, two phonological representations for the American English word *winter*, one with an underlying /nt/ cluster and one with the nasal flap. Ranbom and Connine (2007) have made such a proposal to account for the comprehension of pronunciation variants of exactly this type.

At the opposite end of the continuum, we find exemplar-based models, which assume that the mental lexicon consists of "clouds" of exemplars, with each exemplar representing one token encountered by the speakers in their own speech or in that of others (e.g., Johnson, 1997). In contrast to the traditional models and the abstract variants models, the lexical

[1] Some models assume that the variants are generated not by phonological but rather by phonetic processes (Barnes & Kavitskaya, 2002; Smorodinsky, 1998, but see the review by Côté & Morrison, 2007). These phonetic accounts are in line with Articulatory Phonology, developed by Browman and Goldstein (1992), which accounts for segment deletion by reduction in the sizes of articulatory gestures and temporal overlap of these gestures.

representations are not abstract but contain all kinds of phonetic information about the token, including speech rate and properties of the speaker (e.g., Goldinger, 1998). Exemplar-based models treat phonological variation in the same way as they treat pronunciation differences between speakers and situations: all pronunciation variants are stored in the mental lexicon. Thus, the lexicon would contain multiple representations for *winter* produced with an [nt] and for *winte*r produced with a nasal flap, each representing different voices and situations.

Our study presents three on-line production studies whose objective is to determine whether the production process is based on only one or on more than one lexical representation for words with several pronunciation variants. Note that our focus is thus not on distinguishing between abstract variants models and exemplar-based models. We examine the role of the relative frequency of each pronunciation variant on its production. Traditional models assume that variant frequency is not stored in the mental lexicon, since storing this frequency would basically amount to including in the lexicon a representation for each pronunciation variant. These models hence predict that variant frequency does not correlate with the production latencies of a variant. Models assuming at least one representation for each variant (abstract variants models and exemplar-based models), in contrast, predict that a higher relative frequency leads to shorter production latencies, since the activation levels of these representations are determined by frequency.

To our knowledge, no study has investigated the role of variant frequency (absolute or relative) in word production. For comprehension, in contrast, Ranbom and Connine (2007) and Pitt, Dilley and Tat (submitted) have presented experimental evidence suggesting that the pronunciation variants with higher frequencies tend to be recognized more easily than the variants of lower frequencies. Both studies are based on the realization of /t/ in American English. It seems that no comparable study has been conducted for another language than

(American) English. This points to the need for cross-linguistic studies that test the generalizability of these findings.

Our study investigates the production of pronunciation variants resulting from so-called schwa deletion in French as it occurs word-internally. Certain words (hereafter called "schwa words") can be pronounced either with schwa (e.g., [ʁəkɛ̃], 'shark') or without (e.g., [ʁkɛ̃]). This phenomenon is not at all marginal since it affects, for example, one out of ten French words in the IlPho Lexicon (Boula de Mareüil et al., 2000). Even though schwa deletion occurs primarily in connected speech, speakers, upon instruction, can easily produce the reduced variant in isolation. Thus, French schwa words provide an excellent medium for the study of on-line production of phonological variants. Research into schwa deletion in French extends psycholinguistic research to a non-Germanic language.

We investigated whether the relative frequency of a variant influences its production latency by means of two different experimental paradigms. In Experiments 1 and 2, we used a picture-naming task in which participants were instructed which variant of the schwa words to produce. In Experiment 1, participants produced words in isolation and in Experiment 2, they produced words both in isolation and preceded by a determiner. In Experiment 3, we used the symbol-word association learning task. Experiments 2 and 3 also investigated issues raised by the results of Experiment 1.

Experiment 1

Method

Participants

Twenty-seven students from the Psychology Department of the University of Geneva took part in the experiment. They were all monolingual French speakers, with no reported hearing or language impairment.

Material

Sixty-four French picturable polysyllabic nouns were selected (see Appendix A). Each word had a schwa in the first syllable and could be produced either with (e.g., [ʀəkɛ̃] 'shark') or without schwa (e.g., [ʀkɛ̃]). The lexical frequencies of these words in *Lexique* (New, Pallier, Ferrand & Matos, 2001), based on a corpus of subtitles for foreign movies, vary between 0.01 and 186.01. The relative frequencies of the pronunciation variants of these words were obtained by means of a rating experiment, as previous investigations have shown that speakers are able to correctly rate the frequencies of phonological variants. For instance, Racine and Grosjean (2002) showed that the frequencies of the two pronunciation variants for schwa words estimated by a group of 18 speakers predicted the variants' frequencies in the productions of 16 different speakers.

The pictures for the words were taken from multiple sources: two from Alario and Ferrand (1999), 50 from the Google picture database (http://www.google.fr), and 12 pictures were drawn specifically for this experiment. Many of the pictures we used were thus not extracted from a psycholinguistic database, and we had no information on their "name agreement" values. Note that since the comparison relevant for our research question is between the two pronunciations of the same word referring to the same picture, the different sources of the pictures will not influence the variables of interest.

Each word was recorded in two variants (with and without schwa) by a female native French speaker on a DAT system. We then created separate sound files for each variant using

Praat (Boersma & Weenink, 2007). These files were used in the familiarization phase described below.

Design and Procedure

The experiment consisted of two naming sessions and a frequency estimation task. The two naming sessions took place in a quiet room and lasted about 30 minutes each. They were separated by at least two weeks. The frequency estimation task took place at least two weeks after the second naming session and lasted for about 15 minutes. It was conducted either at the participant's home or in a quiet room at the University of Geneva.

Variant relative frequency estimation task

Participants were given a booklet with two parts. In the first, the words were presented in isolation; in the second, they were presented in a carrier sentence. The carrier sentences were not identical for all words but all had the same structure (e.g., *j'ai vu un* ***requin*** *hier soir* 'I saw a shark yesterday evening'). Moreover, the position of the target word in the sentence was identical for all words and the number of syllables following and preceding the target word was constant as well, except for one target word.

Each target word was accompanied by a nine point scale. The schwa variant was indicated on the left side of the scale (e.g., *r**e**quin*) and the reduced variant on the right (*r**'**quin*). Participants were instructed to indicate for each word the relative frequencies of the two variants by encircling a value on the scale; a score of 1 meant that the word was always realized with schwa (e.g., [ʀəkɛ̃]), whereas a score of 9 meant that it was always realized without schwa ([ʀkɛ̃]). In order to have a relative frequency value for each variant, we attributed the value on the scale to the reduced variant and ten minus that value to the schwa variant. For example, if a participant circled the value of 8 for the word *requin* (meaning that

the word was almost always produced without its schwa), the reduced variant was given the value 8 and the schwa variant was given the value 2.

All participants estimated the relative frequencies of the variants for all words both in isolation and in carrier sentences. When providing their answers in the second part of the experiment, they were instructed not to check the answers they had given in the first part. The order of presentation of the words within each condition (in isolation or in a carrier sentence) was counterbalanced across participants.

We obtained ratings for words in isolation since in the naming experiment the words were produced in isolation as well. We also collected ratings in a sentential context to determine the sensitivity of these ratings to a variable that is known to influence the presence of schwa (isolation versus within sentential context). Since schwa is more often absent in connected speech (Fouché, 1956) than in isolated words, we expected that the variants with schwa would be rated as more frequent when presented in isolation. Moreover, a strong correlation between the two ratings would support the conclusion that the isolated word ratings reflect the variation in real speech. We chose to have the same participants perform the naming and the variant relative frequency estimation tasks in order to investigate the effect of the participants' own ratings on their naming latencies.

Naming task

Each participant took part in two naming sessions, a schwa and a non-schwa session, ran in E-Prime (Schneider, Eschman & Zuccolotto, 2002). At the beginning of the schwa naming session, participants were instructed to produce the variants with schwa only, whereas at the beginning of the non-schwa naming session, they were asked to produce the reduced variants. The session order was counterbalanced across participants.

Each of these two sessions was composed of an initial naming phase, a familiarization phase and a second naming phase. In the initial naming phase, we tested whether the names we had associated with the pictures were also selected spontaneously by our participants. We expected that many pictures would elicit several different names. For instance, the picture for *peloton* 'pack' could elicit words such as *cyclistes* 'cyclists', *course* 'race' or *vélos* 'bikes'.

During the familiarization phase, each picture was displayed once on a computer screen for 1500 ms, while at the same time the word, pronounced as the variant that participants were instructed to produce (with or without schwa), was presented via headphones. A 1500 ms blank screen interval separated trials.

Each trial in the naming phases had the following structure: a fixation cross was shown at the center of the screen for 500 ms, followed by a 300 ms blank screen. The picture was then presented in the middle of the screen and had to be named as quickly as possible. At the end of each response, the experimenter pressed a button, which started the next trial after 1000 ms. No voice key was used to record latencies. Instead, participants' responses were recorded on one track of a DAT while an inaudible 25 ms click, signaling the onset of the picture, was recorded on the second track. Latencies were defined as the time separating the onset of the click and articulation onset.

Each naming phase started with a few training items. During the first naming phase, the 64 target items were presented once in random order. The second naming phase (after the familiarization) consisted of four blocks of the 64 items, with a pause after the first two blocks. In each block, the items were presented in a new random order, which was different for each participant. At the start of the second naming phase, participants were told to use the words they had heard during the familiarization phase. For each participant, the order of presentation of the words in the naming session with schwa was identical to the order in the naming session without schwa.

Results

In the present and subsequent experiments, we analyzed the data by means of mixed effects regression models (see Goldstein, 1987; 1995; Rashbah & Goldstein, 1994; Baayen, Davidson & Bates, 2008, for details on mixed effects applied to psycholinguistic data). Mixed effects models do not only account for the standard fixed predictors considered in simple linear regression modeling but also for the random variation induced by specific words or speakers, by assigning different intercepts to these words or speakers. We only retained those fixed predictors in the models that are statistically significant or figure in statistically significant interactions.

Variant relative frequency estimation task

We first examined the validity and consistency of the frequency estimations given by our participants following the procedure used by Desrochers and Bergeron (2000).

Validity assessment: We compared our ratings with similar ratings in the literature using the Spearman correlation coefficient. To our knowledge, the only estimations available are those of Racine (2007). The correlation coefficient between our estimations for words presented in isolation (values averaged over speakers) and Racine's values from Swiss participants is high ($rho = 0.78$, $S = 9834.1$, $p < 0.0001$), even though the tasks and the contexts in which the words were presented were different. In Racine's study, participants gave a separate rating for each variant and the words were preceded by a determiner.

Consistency assessment: Consistency was assessed by computing the Spearman correlation coefficient between the ratings in isolation and in carrier sentences. The correlation is strong and highly significant ($rho = 0.96$, $S = 1967.6$, $p < 0.0001$), which shows that the ratings for words in isolation are very similar to the ratings for words presented in

carrier sentences. This again suggests that participants are able to rate the variant frequencies of schwa words.

The mean rating for the reduced variants presented in isolation was 4.5 (95% confidence interval: ± 0.41) and the mean rating for the reduced variants presented in carrier sentences was 5.1 (95% confidence interval: ± 0.43). We ran a mixed effects model to test the significance of this difference. We modeled the frequency estimations for reduced variants as a function of context (word in isolation versus in a carrier sentence), with participant and word as crossed random effects. The results show an effect of context: words received higher relative frequency ratings for their reduced variants when presented in carrier sentences than in isolation ($F(1, 3408) = 83.5$, $p < 0.0001$). Apparently, the ratings are sensitive to context, a factor known to influence the presence of schwa. This also suggests that the ratings are reliable estimates for the frequencies of the variants.

Finally, the frequency estimations for the reduced variants averaged over speakers correlate positively with the log of the word frequencies in films (here and in the following analyses, when we refer to word frequency, we always mean the frequencies as given by the database *Lexique* for films, Spearman rho = 0.39, $S = 26851.7$, $p < 0.01$). This finding is in line with previous literature showing that schwa tends to be more often absent in more frequent words (Hansen, 1994; Racine & Grosjean, 2002).

Naming task

Each vocal response was checked for accuracy. Hesitations, productions of non-target words, productions of wrong variants and mispronunciations were considered as errors and removed from the analysis. In the first naming phase (prior to familiarization) there were 2040 errors out of 3366 observations (61%: 76% for the first session and 45% for the second session). Most errors were due either to the production of a non-target word (64%) or to the

production of the wrong variant (27%). In the second naming phase (after familiarization) the error rate dropped to 6% on average (852 errors, 8% in the first session and 5% in the second session). Thus for many items, a familiarization phase was necessary to obtain the intended words. However, the association between words and pictures appeared strong enough for our participants to produce the intended words with few errors after the familiarization phase. An analysis of error type for the second naming phase showed that 108 errors were due to the production of the wrong variant. Overwhelmingly (in 103 cases), participants produced the variant with schwa when they had been instructed to produce the one without. Since the overall error rate was low, no further analyses were conducted on the errors.

Further analyses only concerned the 12972 correct responses from the second naming phase (94%). Unfortunately, we did not have relative frequencies for 21 data points, since the corresponding participants had not provided ratings for the given words. These data points were removed from the data set as well. Finally, visual inspection of the remaining naming latencies showed that the distribution was right skewed. Most of this skewness was removed by taking out the 119 data points above 2000 ms and performing a reciprocal (inverse) transformation (-1/response latencies). The Box-Cox test (Box & Cox, 1964) showed that this inverse transformation was to be preferred over the raw measures or a logarithmic transformation.

The number of observations that were left and on which further analyses were conducted totaled 12832. The mean latency for schwa variants was 765 ms (95% confidence interval: ± 4.9), while the mean latency for reduced variants was 800 ms (95% confidence interval: ± 5.7).

A mixed effects model with the reciprocal of the latencies as the dependent variable and with word and participant as crossed random effects was run on this data set. The most important predictor was the variant relative frequency in isolation for the given variant as

rated by the given participant. Since variant relative frequency was correlated with word frequency (see above), relative frequency was first orthogonalized with word frequency: the raw values for variant relative frequency were replaced by the residuals of a linear model in which variant relative frequency was predicted by log word frequency. Other available variables that could influence word production were entered as fixed predictors as well: variant type (whether the produced variant did or did not contain a schwa), log word frequency and number of repetitions. Following Baayen (2008), residuals larger than 2.5 times the standard deviation (243 data points forming 1.9% of the data) were considered outliers and removed. The final model is summarized in Table 1.[2]

The model shows that latencies decreased with the number of repetitions and with log word frequency. In addition, they were shorter for schwa variants than for reduced variants. These main effects were modulated by two two-way interactions, variant type by number of repetitions and variant type by variant relative frequency. Separate analyses for the variants with schwa and the reduced variants showed that the number of repetitions affected the reduced variants more strongly than the variants with schwa. More importantly, a higher variant relative frequency facilitated the production of reduced variants only. Figure 1 shows the interaction between the residuals of variant relative frequency and variant type, as predicted by the statistical model. Note that the values of the residuals of variant relative frequency range from -4 to 4 (instead of from 1 to 9, which are the original values of variant

[2] Several properties of the words and pictures are known to influence latencies in the picture naming task (see for example Alario et al., 2004). Since two variants of a same word are included in the experiment, these variables cannot act as potential confounding factors. The removal of errors however may lead to a somewhat unbalanced data set. One might argue that the resulting difference in the number of data points between the two variant types may affect the statistical outcomes. We checked this possibility by means of an additional analysis on the basis of only those observations for which latencies were also available for the other variant of the word for the same participant and for the same number of repetitions (12096 observations, 94% of data). The errors' counterparts (i.e., latencies for the same participant, word and repetition but for the other variant) were thus removed. The results we obtain for this more balanced data set closely mirror those for the complete data set. This was expected given that the errors formed only 6% of the data.

relative frequency) because they result from a model predicting variant relative frequency as a function of word frequency.

Table 1.

Summary of the Mixed Effects Model for Experiment 1. The intercept represents a reduced variant produced for the first time.

Variable	**β**	**F**	**p**
Number of repetitions	$-1.012\ 10^{-4}$	2778.82	<0.0001
Log word frequency	$-1.62\ 10^{-5}$	5.38	<0.05
Variant type (with schwa)	$-1.33\ 10^{-4}$	167.11	<0.0001
Variant relative frequency	$-4.90\ 10^{-6}$	2.25	>0.1
Variant type by Number of repetitions	$2.44\ 10^{-5}$	51.89	<0.0001
Variant type by Variant relative frequency	$7.44\ 10^{-6}$	11.18	<0.001

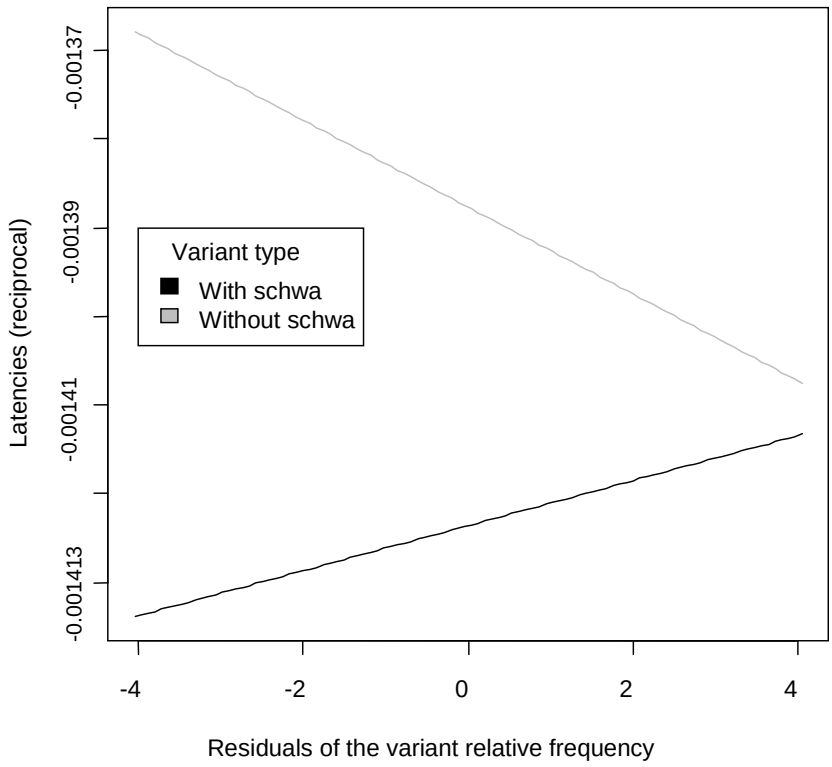

Figure 1. The effects of variant type and the residuals of the variant relative frequency on the naming latencies in the statistical model of Experiment 1.

Discussion

Experiment 1 showed an effect of variant relative frequency for the reduced variants in naming. This effect suggests that information about the use of both variants of schwa words is stored in the lexicon, since relative frequency is computed on the basis of the frequencies of both variants. In order to obtain this information, the lexicon needs to keep a frequency count of the productions of each variant. Hence, on the basis of Experiment 1, we can conclude that both schwa variants and reduced variants are stored.

If indeed both variants are represented in the lexicon, the question arises why only the production of the reduced variants is affected by variant relative frequency. An effect for the

schwa variants would be expected as well. The answer to this question is probably found in other variables influencing the production of schwa words. According to Fouché (1956), words are more often produced with schwa in isolation than in sentences, as also appears from the results of our variant relative frequency estimation task. We hypothesize that context plays a role in the effects of variant type and variant relative frequency and their interaction in the production of schwa words. Context may favor one of the variants, which is consequently produced more quickly (effect of variant type). Moreover, effects of variant relative frequency might be overruled by the contextual bias in the production of that favored variant. The picture-naming task used in Experiment 1 is an example of a context where only one variant type is appropriate. This task is similar to any speech act in which one names a concept, object or picture in isolation. In such acts, the schwa variant is almost exclusively used. In contrast, in contexts where both variants are equally probable, variant relative frequency effects may be observed for both variants.

In order to investigate whether context affects the roles of variant relative frequency and variant type, we conducted a second naming experiment. In this experiment, participants did not only produce the nouns in isolation, but also preceded by a determiner. Reduced variants are more common in such small phrases than in isolation. For these phrases, we therefore predicted a smaller effect of variant type and an effect of variant relative frequency for both variant types.

Experiment 2

Method

Participants

Twenty-four students from the Psychology Department of the University of Geneva took part in the experiment. They were all monolingual French speakers, with no reported hearing or language impairment, and they had not participated in Experiment 1.

Material

Eight items were removed from the target words used in Experiment 1. Some words were disregarded because they could not be used in the singular form (e.g., *cheveux* 'hair'), while others were removed in order to obtain a balanced set according to the words' grammatical gender. In addition, four pictures were changed so that the picture represented the singular form of the word instead of its plural form and one word was changed in order to better correspond to its associated picture (*renard* 'fox' became *renardeau* 'fox cub'). In summary, 56 words were retained for this experiment (see Appendix A), 55 of which also occurred in Experiment 1.

A native speaker of French recorded each word in the two variants (with and without schwa) on a DAT system. Both variants were produced in two contexts: in isolation and preceded by the possessive determiner *mon* or *ma* 'my' (the selected form depending on the word's grammatical gender). We then created separate sound files for each variant in each context using Praat. These files were used in the familiarization phase of the naming task.

Design and Procedure

The experiment again consisted of two naming sessions and a frequency estimation task, conducted as in Experiment 1.

Variant relative frequency estimation task

The variant relative estimation task was identical to the one in Experiment 1.

Naming task

The naming task differed in a number of respects from the one in Experiment 1. First, the experimenter was not present during the experiment. Second, there was no naming phase prior to the familiarization phase, since Experiment 1 had shown that for most items participants need a familiarization phase in order to produce the intended words. Third, the familiarization phase consisted of the variants produced in isolation for the first naming session and of the variants produced with the possessive determiner for the second naming session. Fourth, in each naming session the participants produced the variants twice in isolation (one token in each of two blocks) and twice preceded by the possessive determiner (again one token in each of two blocks). The order of these two contexts was counterbalanced across participants. The order of the target words in the two blocks of one context condition was identical to the order in the other condition. Fifth, a voice key was used to record the naming latencies. If no response was given, the next item started after 2500 ms. Sixth,, in order to have a similar pace in the two context conditions, the intertrial interval was reduced to 750 ms for the words in isolation, but was still 1000 ms for the context 'determiner + noun'. Finally, the experiment was run with DMDX (Forster & Forster, 2003).

Results

Variant relative frequency estimation task

The ratings provided by the participants of Experiment 2 are highly similar to those obtained in Experiment 1 (Spearman rho = 0.94, S = 1691.9, $p < 0.0001$). Consequently, as in Experiment 1, the correlation coefficient between the estimations for words presented in isolation (values averaged over speakers) and Racine's values was high and statistically significant (Spearman rho = 0.83, S = 50816.1, $p < 0.0001$), and so was the correlation between the ratings for the reduced variants in isolation and the words' frequencies in films (as given by the database *Lexique*, Spearman rho = 0.38, S = 17274.6, $p < 0.01$). The

correlation between the ratings in isolation and in carrier sentences was also again high and significant (Spearman rho = 0.95, S = 1539.4, $p < 0.0001$), and a mixed effects model showed again that context affected the ratings. Words obtained a higher relative frequency for their reduced variants when presented in carrier sentences (mean: 4.6; 95% confidence interval: ± 0.45) than in isolation (mean: 4.6; 95% confidence interval: ± 0.43; $\beta = 0.26$, $F(1, 2205) = 10.6$, $p < 0.01$).

In summary, the results for the variant relative frequency estimation task in Experiment 2 were very similar to the results obtained for the same task in Experiment 1. They show that the ratings are valid and consistent and that they can be considered as reliable estimates for the relative frequencies of the variants.

Naming task

Each vocal response was checked for accuracy. Hesitations, dysfluencies, productions of non-target words, productions of the wrong variant, and mispronunciations were considered as errors and removed from the analysis. For the naming of words in isolation, the number of errors totaled 579 (11%). For the naming of words preceded by a determiner, there were 719 errors (13%).

An analysis of error type showed that in isolation, 56 errors were due to the production of the wrong variant. Overwhelmingly (in 50 cases), participants produced the variant with schwa when they had been instructed to produce the one without. For the words produced with the determiner, 61 errors were due to the production of the wrong variant. In 37 cases, participants produced the schwa variant when instructed to produce the reduced variant and in 24 cases, they produced the reduced variant instead of the schwa variant. Thus, whereas the isolation context showed a preference for the schwa-variant, this preference was weaker for the 'determiner + noun' context.

Further analyses were restricted to the 9454 correct responses. We removed from this data set another 14 data points for which we did not have relative frequencies (the given participants had not provided ratings for these words). The remaining latencies were adjusted whenever necessary using the software CheckVocal (Protopapas, 2007) in order to eliminate differences in voice key accuracy caused by the nature of the two first phonemes (Kessler, Treiman & Mullennix, 2002). Visual inspection of the resulting latencies showed that the distribution was right skewed. Most of this skewness was removed by taking out the 87 data points above 1800 ms and performing a reciprocal transformation, again following the Box-Cox test.

The number of observations that were left and on which further analyses were conducted totaled 9353 (87% of the data). Table 2 gives the mean latencies and 95% confidence intervals as a function of production context and variant type.

Table 2.

Mean naming latencies (in milliseconds) and 95% confidence intervals as a function of production context and variant type in Experiment 2.

	Isolation (n = 4737)		**With determiner** (n = 4616)	
	Mean	**95% confidence interval**	***Mean***	**95% confidence interval**
Schwa variants	770	± 8.5	740	± 8.0

	Isolation (n = 4737)		**With determiner** (n = 4616)	
	Mean	**95% confidence interval**	***Mean***	**95% confidence interval**
Reduced variants	809	± 9.6	786	± 8.5

We analyzed the data again by means of a mixed effects model with the reciprocal latencies as the dependent variable and with word and participant as crossed random effects. As in Experiment 1, the most important predictor was the variant relative frequency for the given variant as rated by the given participant. We used only the ratings obtained for the words in isolation. We did so also for the words produced after the possessive determiner, since we felt that the determiner context was closer to the isolation context than to the sentence context of the relative frequency estimation task. Also as in Experiment 1, we included as fixed predictors variant type, log word frequency and number of repetitions. In addition, we included production context (word in isolation versus preceded by the determiner), and familiarization context (whether the items in the familiarization phase of the given session had been presented in the same context or in the different context).

Residuals larger than 2.5 times the standard deviation (159 data points forming 1.7% of the data) were considered outliers and removed. The final model is summarized in Table 3. This model includes the raw variant relative frequencies, instead of variant relative frequencies orthogonalized with word frequency as in Experiment 1, since word frequency appeared not to be a significant predictor for the naming latencies.

Table 3.

Summary of the Mixed Effects Model for Experiment 2. The intercept represents a schwa variant produced for the first time in isolation after a familiarization in which the word was preceded by a determiner.

Variable	**β**	**F**	**p**
Familiarization context (same)	$3.46\ 10^{-5}$	10.14	<0.01
Number of repetitions (second repetition)	$-1.14\ 10^{-4}$	587.57	<0.0001
Production context (determiner)	$7.54\ 10^{-5}$	42.25	<0.0001
Variant type (reduced variant)	$1.07\ 10^{-4}$	241.15	<0.0001
Variant relative frequency	$4.69\ 10^{-6}$	0.023	>0.1
Variant Type by Variant relative frequency	$-8.83\ 10^{-6}$	0.81	>0.1
Production context by Variant relative frequency	$-1.19\ 10^{-5}$	9.83	<0.01
Production context by Variant type	$-4.96\ 10^{-5}$	3.32	>0.1
Production context by Familiarization context	$-1.04\ 10^{-4}$	116.25	<0.0001
Production context by Variant type by Variant relative frequency	$1.33\ 10^{-5}$	13.42	<0.001

The model showed main effects for all predictors, except variant relative frequency and word frequency. In addition, there were several interactions with production context and a three way interaction between production context, variant type and variant relative frequency. In order to better understand these interactions, we ran separate models for each production context (isolation and "with determiner"). The models are summarized in Table 4. Importantly, both contexts showed that schwa variants were produced faster than reduced variants. The two contexts, in contrast, showed different effects for the variant relative frequency. The isolation context replicated the results from Experiment 1. The interaction between variant type and variant relative frequency shows that there is only a variant relative frequency effect for the reduced variants. The 'determiner+noun' context shows a main effect of variant relative frequency, but no interaction with variant type: variant relative frequency affected both variants equally. Figure 2 illustrates the three-way interaction between production context, variant type and variant relative frequency in the complete data set, as predicted by the statistical model for all latencies (in isolation and with the determiner).

Table 4.

Summary of the Mixed Effects Models for words produced in isolation (left) and with a determiner (right) in Experiment 2. The intercept represents a schwa variant produced for the first time after a familiarization in the other context.

	Isolation			**With determiner**		
Variable	**β**	**F**	**p**	**β**	**F**	**p**

Variant relative frequency	$4.89\ 10^{-6}$	0.017	>0.1	$-4.83\ 10^{-6}$	9.85	<0.01
Familiarization (same)	$3.50\ 10^{-5}$	27.51	<0.0001	$-7.05\ 10^{-5}$	114.52	<0.0001
Number of repetitions (second repetition)	$-1.14\ 10^{-4}$	290.18	<0.0001	$-1.16\ 10^{-4}$	332.12	<0.0001
Variant type (reduced variant)	$1.09\ 10^{-4}$	86.46	<0.0001	$7.74\ 10^{-5}$	143.96	<0.0001
Variant Type by Variant relative frequency	$-9.25\ 10^{-6}$	6.19	<0.05	-	-	-

The models also show additional effects. Participants were quicker in producing a variant in a given context the second time. Furthermore, participants reacted more quickly if they had to produce a variant in the same context as in the familiarization phase, but only in the ‘determiner + noun’ context.

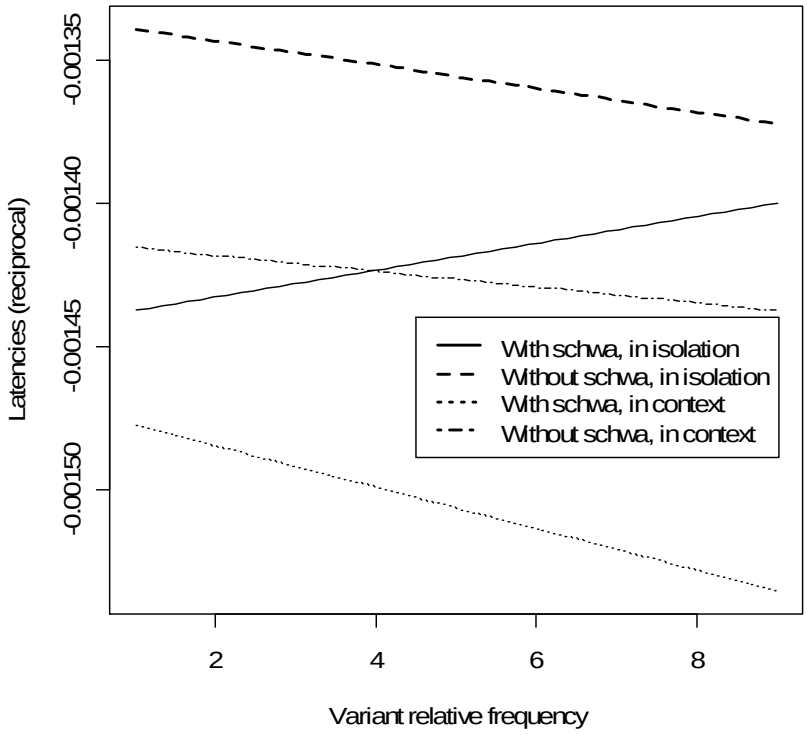

Figure 2. The effects of production context, variant type and variant relative frequency on the naming latencies in the statistical model of Experiment 2.

Discussion

Experiment 2 investigated the influence of contextual factors on the effect of variant type and on the effect of variant relative frequency. Most importantly, the results are in line with our hypothesis that context plays a role in the determination of an effect of variant relative frequency. Just as in Experiment 1, only reduced variants were influenced by the variants' relative frequencies when the words were produced in isolation. In contrast, when preceded by the possessive determiner, a context in which both variants are appropriate, the variants' relative frequencies facilitated the naming of both variant types. These results support our hypothesis that context modulates the effect of variant relative frequency.

In addition, we had predicted that if the advantage for schwa variants was due to contextual factors, this advantage would be smaller in the context 'determiner + noun' than in the isolation context. The errors of Experiment 2 confirm this prediction as they show a less pronounced preference for the schwa variants when the words had to be produced after the possessive determiner (37 erroneous production of schwa variants versus 24 erroneous productions of reduced variants) than in isolation (50 versus 6). The response latencies, however, do not show an effect of the manipulation of context on the schwa variant advantage in our experiment. In the two production contexts (isolation and with determiner), the advantage for schwa variants over reduced variants was comparable. It is likely that in picture naming, the context to be produced before the schwa word has to be longer than just a determiner in order for the reduced variant to be substantially more likely than it is in isolation context.

In fact, it is probably impossible to test such sufficiently long contexts in naming tasks. The naming task is only sensitive to the planning of a single phonological unit, and this puts severe restrictions on the sizes of the sequences we can use. Costa, Navarrete and Alario (2006), for instance, showed that in sequences such as *the dog and the car*, only the effect of the frequency of the first noun could be detected.

A final result worth mentioning is the absence of an effect of word frequency. We believe that it is just due to lack of statistical power. In Experiment 2, we collected only half of the number of data points for the different conditions that we did in Experiment 1, since the participants produced all variants in all contexts only twice, instead of four times.

In order to further investigate the effects of variant relative frequency and variant type and their modulation by context, we conducted a third experiment with a different experimental paradigm. Participants performed a symbol-word association learning task, similar to the one reported in Levelt and Wheeldon (1994). Participants learned associations

between abstract sequences of symbols and auditorily presented words or pseudowords. They then produced these (pseudo)words whenever the corresponding symbols appeared on the computer screen. In this paradigm, participants are thus required to utter a phonological form in response to a sequence of abstract symbols, instead of an object or a concept as in the naming task. This paradigm has the advantage that the participant's task is not part of the linguistic acts speakers use on a regular basis. As a consequence, it should favor neither the variant with schwa nor the reduced variant. We thus expected an influence of variant relative frequency for both variants and no advantage for schwa variants.

In addition, Experiment 3 was designed to test simultaneously two alternative explanations for the advantage for schwa variants observed in Experiments 1 and 2. First, the schwa variant may have a privileged status because it is supported by orthography. The systematic advantage found for non-reduced forms over reduced forms in the word recognition literature is often explained by an appeal to the existence of orthographic representations for these non-reduced forms (Racine & Grosjean, 2002; Ranbom & Connine, 2007). The same could hold for word production.

Second, the slower production of reduced variants compared to schwa variants could be due to structural differences between the two types of variant. These differences may affect ease of production, which includes ease of lexical retrieval, phonological encoding, and articulation. Several studies have shown an advantage for words with simple (as opposed to complex) onsets (Santiago, MacKay, Palma & Rho, 2000, but see Roelofs, 2002), for words with initial syllables of higher frequencies (Cholin, Levelt & Schiller, 2006; Laganaro & Alario, 2006), for words with higher phonotactic probabilities (Vitevich, Armbrüster & Chu, 2004), for words with higher neighborhood densities (Vitevich, 2002) and neighborhood frequencies (Vitevich & Sommers, 2003), and for words that have fewer phonemes (Roelofs, 2002) or syllables (Santiago et al., 2000, but see Bachoud-Lévi, Dupoux, Cohen & Mehler

1998; Alario et al., 2004; Meyer, Roelofs & Levelt, 2003). Most of these effects would lead to faster latencies for variants with schwa. Firstly, all schwa variants have simple onsets and all reduced variants have complex onsets. Furthermore, reduced variants often have rare or unattested initial syllables or consonant clusters, leading to lower syllable frequencies and lower phonotactic probabilities. As for neighborhood properties, we know that neighborhood density correlates positively with phonotactic probability (Vitevich & Luce, 1999) and as a consequence reduced variants probably also have fewer neighbors and lower neighborhood frequencies. Only word length should favor reduced variants, since they all have one fewer syllable and one fewer phoneme than their schwa variant counterparts. The relative impact upon production latencies of these different phonological properties has not yet been determined. If the multiple properties favoring faster naming of the schwa variants outweigh the word length effect, this would produce an advantage for the schwa variants.

If the advantage for the schwa variant observed in Experiments 1 and 2 is due only to the experimental context, we would expect no effect of variant type in the symbol-word association learning task used in Experiment 3, since this task favors neither the schwa variant nor the reduced variant. By contrast, if the advantage for schwa variants is due to intrinsic rather than contextual factors, such as the presence of *e* in the orthographic representations of the words, or differences in structural properties between the two variant types, schwa variants should also show an advantage over reduced variants in the symbol-word association learning task.

The symbol-word association learning task allows for the testing of pseudowords (Cholin et al., 2006). Importantly, pseudowords allow us to also test whether the effect of variant relative frequency is driven by structural properties of the words. The two variants of a word differ in their structural properties and this difference varies across words. Since these structural properties most likely influence ease of production, the difference in ease of

production between the schwa and the reduced variant may vary across words as well. We cannot exclude a priori that variant relative frequency is highly correlated with these differences in ease of production. Consequently, the effect of variant relative frequency that we observed in Experiments 1 and 2 may simply be a by-product of these structural differences. If this were the case, response latencies for pseudowords should be correlated with the relative frequencies of the variants to which these pseudowords are highly phonologically similar. If, in contrast, as we have assumed so far, the variant relative frequency is an idiosyncratic property of the word, and therefore stored with that word, these frequencies should not correlate with the production latencies of phonologically similar pseudowords.

Additionally, the pseudowords can show more directly whether structural differences are responsible for the effect of variant type. We constructed two types of pseudowords: pseudowords that are very similar to the schwa variants of the words and pseudowords that are very similar to the reduced variants. If Experiment 3 also shows a main effect of variant type, the question is whether this effect is equally large for the pseudowords as for the words. A similar effect would indicate that structural differences between the schwa and reduced variants are likely responsible for the advantage for schwa variants. In contrast, a greater effect for the words than for the pseudowords would indicate that the advantage for schwa variants results at least partly from context or the existence of orthographic representations for these variants.

Experiment 3

Method

Participants

Twenty-eight students from the Psychology Department of the University of Geneva took part in the experiment. They were all monolingual French speakers, with no reported hearing or language impairment, and they had not participated in Experiments 1 or 2.

Material

Thirty French polysyllabic nouns were chosen (see Appendix B). Each word had a schwa in the first syllable and could be produced either with (e.g., [ʀəkɛ̃]) or without the schwa (e.g., [ʀkɛ̃]). Out of these 30 words, 23 were also part of the stimuli of Experiment 1 and 18 of them were part of the stimuli of Experiment 2.

For each of the 30 target words, two pseudowords were constructed: one with structural properties similar to the variant with schwa and the other with properties similar to the reduced variant. We constructed the pseudowords associated with the schwa variants by changing at least the schwa and one more vowel but keeping the consonants constant (e.g., [ʀəkɛ̃] => [ʀikø]). We constructed the pseudowords associated with the reduced variants by deleting the first vowels of the pseudowords associated with the schwa variants (e.g., [ʀikø] => [ʀkø]). The syllabic structures and lengths in phonemes or syllables of the pseudowords thus equaled those of the associated variants of the words.

Four different lists of 30 items were created: word variants with schwa, word reduced variants, pseudowords structurally similar to the word variants with schwa and pseudowords structurally similar to the word reduced variants. In each list, items were paired to form 15 different blocks. The pairing was done pseudo-randomly, since we avoided close semantic or phonological relationships between the two items of a pair. The pairing of items differed for each participant but was identical for the four items corresponding to a target word (i.e., the word variant with schwa, the word reduced variant, the pseudoword similar to the word

variant with schwa and the pseudoword similar to the word reduced variant) for a given participant.

All items corresponding to a target word were associated with the same arbitrary sequence of identical symbols (e.g., *$$$$$$*). Since some sequences of symbols might be easier to process than others, symbols-word associations were varied across participants.

An acoustic version of each item was recorded by a female native speaker of French on a DAT system.

Design and Procedure

The experiment was run in two sessions, with DMDX. In the first session, participants performed the production task with the two lists of pseudowords (the order of the two lists counterbalanced across participants), and then a variant frequency estimation task of the words. During the second session, the participants performed the production task with the two lists of words (again the order being counterbalanced across participants). Sessions one and two were separated by at least two weeks.

Variant relative frequency estimation task

The procedure for the variant relative frequency estimation task was identical to the one used in Experiment 1, except that the task was now performed in a quiet room at the University of Geneva by all participants.

Production tasks

Participants were tested individually, in a quiet environment. Both production sessions contained 31 blocks (1 training block followed by two lists of 15 blocks each) and lasted about 45 minutes. Every block consisted of three phases: a learning phase, a practice phase, and a test phase.

During the learning phase, participants were asked to memorize the associations between the sequences of symbols and the two items of the block. They were instructed to listen carefully to the segmental content of the items in order to be able to reproduce them as faithfully as possible. Each sequence of symbols was presented on the screen for 2000 ms while the corresponding item was presented auditorily over headphones. Every association was presented twice, leading to a total of four trials per learning phase for each block.

During the practice phase, each trial had the following structure. A number (either *3*, *12*, *13*, or *15*) first appeared on the screen for 800 ms. Participants had to name this number as quickly as possible. After a 500 ms blank screen a sequence of symbols was displayed on the screen for 1500 ms, and participants had to produce the item associated with it as quickly as possible. The sequence of symbols disappeared and after a 500 ms delay, the associated stimulus was presented auditorily so that participants could judge the adequacy of their response. The next trial started after a 800 ms intertrial interval. Each of the two symbol sequences was presented three times, leading to a total of six trials per practice phase for each block. The inclusion of a number to be named before the naming of each symbol sequence was to prevent facilitation effects due to the articulation of two identical words in a row. Such sequences of identical words occurred since the two symbol sequences were presented more than once and in random order. There was no phonological overlap between these numbers and the words to be produced (see Cholin et al., 2006, for a similar use of fillers).

Test phases were identical to the practice phases except for the following four aspects. The voice key was activated at the onset of the presentation of the sequence of symbols and for 2000 ms. Symbol sequences disappeared with the response and no feedback was given. Finally, each of the two symbol sequences was presented four times, leading to a total of eight trials per test phase for each block.

The presentation order of the 15 blocks in a list differed for each participant but was identical in the four lists for a given participant (i.e., all four items associated with a given target word appeared in the same positions in the lists).

Results

Variant relative frequency estimation task

The ratings provided by the participants, who only rated the words, again appeared highly reliable. The correlation coefficient between the ratings obtained for the words presented in isolation (averaged over speakers) and Racine's values was high and statistically significant (rho = 0.88, S = 554.2, $p < 0.0001$) and so was the correlation between the ratings in the isolation condition and in the context condition (rho = 0.92, S = 358.8, $p < 0.0001$). A mixed effects model showed again that, as expected, words obtained higher relative frequencies for their reduced variants when they were presented in carrier sentences (mean: 4.1; 95% confidence interval: ± 0.67) than in isolation (mean: 3.6; 95% confidence interval: ± 0.73; $F(1, 1587) = 32.4$, $p < 0.0001$). In contrast to Experiments 1 and 2, however, the frequency estimations for the reduced variants averaged over speakers did not correlate with log word frequency (Spearman rho = -0.01, S = 4544.5, p = 0.96, see below for a possible explanation).

Symbol-word association learning task

Each vocal response was checked for accuracy. One participant did not take the experiment seriously and often gave inadequate responses. His responses were therefore excluded from the analysis. For the remaining participants, hesitations, productions of non-target words, productions of the wrong variant, and mispronunciations were considered as errors and removed from the data set. For pseudowords, these errors formed 16% of the data. For the words, participants produced errors in 5% of the trials, and only six of these errors

resulted from the production of the wrong variant: three items were incorrectly produced without schwa and three items with schwa.

Latencies for correct responses were adjusted whenever necessary using the software CheckVocal. The latency distributions were skewed both for the words and the pseudowords. To reduce this skewness we applied a (natural) log transformation and took out the one data point below 100 ms, the three data points above 1500 ms for pseudowords and the 20 data points above 1200 ms for words. According to the Box-Cox test, the logarithm transformation was to be preferred over the raw measures or the reciprocal transformation. In addition, we removed the 30 data points (15 words and the 15 corresponding pseudowords) for which we did not have variant relative frequencies since the given participant had not provided a rating. The number of observations that were left and on which further analyses were conducted totaled 11539 (90% of the data). Table 5 gives the mean latencies as a function of lexicality and variant type.

Table 5.

Mean naming latencies (in milliseconds) and 95% confidence intervals as a function of lexicality and variant type in Experiment 3.

	Words (n = 6134)		**Pseudowords** (n = 5405)	
	Mean	**95% confidence interval**	***Mean***	**95% confidence interval**
Schwa variants	466	±5.0	486	±5.6

	Words (n = 6134)		**Pseudowords** (n = 5405)	
	Mean	**95% confidence interval**	***Mean***	**95% confidence interval**
Reduced variants	472	±5.3	481	±5.8

We ran a mixed effects model with word and participant as crossed random effects and (natural) log transformed latencies for the correct responses as the dependent variable. Lexicality, variant type, variant relative frequency, and number of repetitions were entered as fixed predictors. Again, residuals larger than 2.5 times the standard deviation were considered outliers and these 177 data points (1.5 % of the data) were removed. The model refitted without these observations only showed a main effect of lexicality ($F = 58.99$, $p < 0.0001$). Words were produced with shorter latencies (average naming latency: 469 ms) than pseudowords (484 ms). Importantly, unlike Experiments 1 and 2, this experiment did not show an effect of variant type. Furthermore, there was an interaction between lexicality and variant relative frequency ($\beta = -0.0037$, $F = 5.87$, $p < 0.05$). This interaction is illustrated in Figure 3.

In order to investigate in detail the effect of variant relative frequency, we analyzed the words (n = 6134) and pseudowords (n = 5405) separately. We first ran a mixed effects model for the words. Along with variant relative frequency and variant type we included log word frequency as a fixed predictor. The 99 data points (1.6 % of the data) with residuals larger than 2.5 times the standard deviation were considered as outliers and removed from the data

set. The model refitted without these observations showed that variant relative frequency was the only significant predictor of the latencies (β = -0.0034, F = 11.05, p < 0.001). Higher variant relative frequencies implied shorter latencies (see Figure 3). Contrary to Experiments 1 and 2, there was no effect of variant type and no interaction between variant type and other factors.

We finally analyzed the response latencies for the pseudowords. They showed no effect of variant type or variant relative frequency (both ps > 0.1).

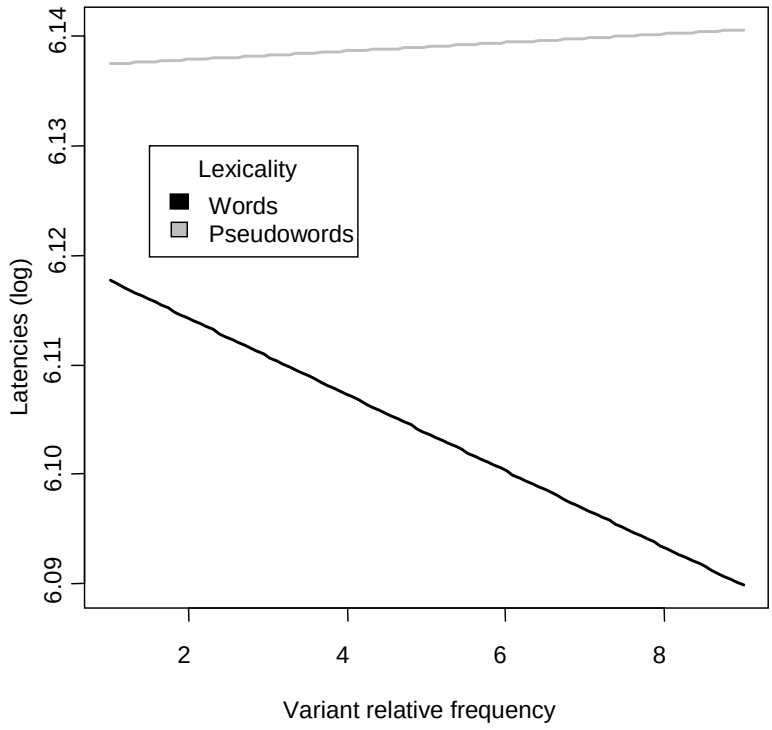

Figure 3. The effects of lexicality and variant relative frequency on the naming latencies in the statistical model of Experiment 3.

The words and the corresponding pseudowords were identical in their word onsets and word lengths (both in numbers of segments and syllables). Hence, our finding that the naming latencies for the pseudowords are not correlated with variant relative frequency shows that the

observed effect of variant relative frequency for the words is not a by-product of differences in these structural properties between the variants among the words. However, since the words and the pseudowords differed in the quality of their first vowels (schwa versus full vowel), they were not matched perfectly on the frequency of this first syllable, their phonotactic probability, their neighborhood density, and their neighborhood frequency, which are all variables known to affect ease of production (e.g., Laganaro & Alario, 2006; Vitevich, Armbrüster & Chu, 2004; Vitevich, 2002; Vitevich & Sommers, 2003). Hence, the results for the pseudowords do not eliminate completely the possibility that the variant relative frequency effect observed for the words is driven by differences in these four properties between the variants among the words. We investigated this possibility in further analyses.

First, in order to assess the relevance of these four structural properties for our type of data (French schwa variants and reduced variants in a symbol-word association learning task), we investigated their ability to predict the naming latencies for the words and pseudowords above and beyond lexicality. Initial syllable frequency and phonotactic probability, operationalized by positional diphone frequency, predicted the naming latencies (both $ps < 0.06$). Neighborhood density showed a trend in the expected direction ($p = 0.1$), while neighborhood frequency showed a significant effect for the pseudowords (interaction with lexicality, $p < 0.01$). Given that all four structural properties have some predictive value for our naming latencies, we then conducted two series of analyses examining their relationship to variant relative frequency.

In the first series, we examined whether the variant relative frequencies are correlated with the differences in these structural properties between the schwa variants and the corresponding reduced variants among the words. In order to characterize the differences in these structural properties between the variants, we computed a difference score for each structural property by subtracting the value for this property for one variant of a given word

from the value for the corresponding variant and divided the outcome by the value for the latter. In this way we computed a difference score for the frequency of the first syllable, for the phonotactic probability of the words' diphones, for neighborhood density, and for neighborhood frequency. None of these difference scores was correlated with variant relative frequency. These results are supported by the following examples. The reduced variants of the word *demeure* 'residence' and *demoiselle* 'young lady' both start with the syllable /dəm/. Nevertheless, whereas *demeure* is seldom produced without schwa according to our participants' estimations (mean relative frequency for reduced variant: 1.5), the reduced variant for the word *demoiselle* is rather frequent (4.7). Additionally, some reduced variants with very easy to produce onset clusters received very low frequencies (e.g., *belette* 'weasel': 1.8; *querelle* 'quarrel': 1.1).

In a second series of analysis, we checked whether the production latencies still showed an effect of variant relative frequency when we added these structural variables (both the difference scores and the values for the variants themselves) as predictors in the statistical models. We reran the main statistical model for Experiment 3 (for words and pseudowords) several times, each time adding one of these variables. Results for all models are identical: the variable capturing the difference in structural property between the schwa variant and the reduced variant is not significant and does not affect the significance of the other predictors. Hence, we conclude that the effect of variant relative frequency on production latencies is real and at best marginally driven by differences in structural properties between the variants among the words.

Discussion

Experiment 3 was conducted in order to replicate the variant relative frequency effect with another experimental task, to investigate the roles of context and structural properties of

the words in this variant relative frequency effect, and to determine the roles of context, orthography, and structural properties of the words in the advantage for schwa variants over reduced variants in Experiment 1 and 2.

The results for Experiment 3 replicate the effect of variant relative frequency for words: the production latencies for both the schwa variants and the reduced variants were negatively correlated with variant relative frequency. This supports the hypothesis that the production lexicon contains at least one lexical representation for each pronunciation variant.

Importantly, the results of this experiment also provide additional support for our hypothesis that experimental context modulates the effect of variant relative frequency. In the symbol-word association learning task both the schwa and the reduced variants are equally appropriate. This is also true, but to a lesser extent, for the naming of nouns preceded by determiners, a task we used in Experiment 2. In both cases, we found that variant relative frequency affected the production of both variants. In contrast, in situations where only one variant is appropriate (e.g., the naming of nouns in isolation in Experiments 1 and 2), variant relative frequency only facilitates the production of the non-favored variant.

Experiment 3 also aimed to determine whether the observed variant relative frequency effect was just a by-product of the structural properties of the words' pronunciation variants that we showed to affect naming latencies. This might have been the case if the relative frequencies of the variants were highly correlated with differences in structural properties between the variants among the words. We argued that this is not the case. First, the production latencies for the pseudowords derived from the words show no variant relative frequency effect. Second, neither the variant relative frequencies nor the naming latencies for the words were highly correlated with the differences in the structural properties between the schwa variants and the reduced variants for the words. This suggests that the variant relative frequency effect is at best marginally due to phonological structure. It confirms that speakers

have stored the variant relative frequencies for words and that their lexicons contain a representation for each pronunciation variant.

This experiment was also designed to investigate the nature of the naming advantage for schwa variants over reduced variants in Experiments 1 and 2. We hypothesized that if contextual characteristics were the main source of this advantage, it would be absent in Experiment 3 since neither variant is a priori preferred or privileged in the symbol-word association learning task. If, by contrast, it were structural differences between the two variants of a word that are responsible for the schwa variant advantage, an effect of variant type would be expected for both the words and pseudowords. Finally, if the advantage for the schwa variants was driven by the existence of an orthographic representation for these variants, we would expect this advantage to surface in Experiment 3 for the words only. The experimental results showed no effect of variant type at all. We thus conclude that the advantage for the schwa variants in Experiments 1 and 2 were due above all to contextual factors.

The presence of an effect of variant relative frequency for words strongly suggests that the symbol-word association learning task we used in this experiment involved lexical processing. One may wonder why the results showed no effect of word frequency, especially since word frequency effects have been found in very similar studies involving a symbol-word association task (Levelt & Wheeldon, 1994; Cholin et al., 2006). There are two possible reasons. Firstly, the words were not selected to test for a word frequency effect. The words differed little in their frequencies (word frequency ranged from 0.01 to 21 occurrences per million) and word frequency showed a strongly right skewed distribution (median = 3.01, mean = 4.62). This would also explain why we did not find a correlation between word frequency and the relative frequency of the reduced variants of the words. Secondly, participants had a practice session with three repetitions of each word before their latencies

were recorded. As a consequence, any frequency effect could well have been attenuated or lost before the real experimental trials began.

To summarize, the results obtained in this third experiment replicate the effect of variant relative frequency and therefore support the hypothesis that schwa words have at least two phonological representations in the mental lexicon. Furthermore, the results strongly suggest that the effect of variant type found in Experiment 1 and 2 was likely due to contextual factors. Finally, the structural properties of the variants cannot explain the effects of variant relative frequency or variant type.

General discussion

The aim of this study was to better understand the nature of the lexical representations involved in the production of words with more than one pronunciation variant. The key question is whether these words are represented by one or by more lexical representations. We conducted three experiments testing whether the production of the two pronunciation variants of schwa words in French is affected by their relative frequencies. If so, speakers have stored these relative frequencies, and their lexicons therefore contain both pronunciation variants. Two naming tasks and a symbol-word association learning task showed correlations between the production latencies of the pronunciation variants and their relative frequencies. We therefore conclude that schwa words are lexically represented by at least one phonological representation for each pronunciation variant.

This conclusion is based on two assumptions. The first assumption is that variant relative frequency is largely an idiosyncratic property of the word. Hence, it cannot entirely be predicted on the basis of the structural properties of its variants, which are related to the ease of production of the variants. In order to explain our variant relative frequency effects, these frequencies thus have to be stored in the lexicon. We tested this in several ways. First

we showed that pseudowords that were structurally matched with the words in our experiment do not show an effect of the variant relative frequency of these corresponding words. Secondly, the variant relative frequencies we obtained in the rating experiments did not correlate with various difference scores expressing the differences in ease of production between the variants of the words. Thirdly, the production latencies show an effect of variant relative frequency even when the effects of these differences in ease of production are partialled out. In fact, the difference scores had no explanatory value at all. These results all support the assumption that the observed effect of variant relative frequency effect is real and not a by-product of structural differences between the words.

The second assumption is that variant relative frequency has a gradient rather than a categorical effect on the production latencies. We tested this assumption for each experiment. We compared our statistical models with similar models in which variant relative frequency was a two-level factor (high versus low variant frequency). Results shows that this categorical measure of variant relative frequency does not predict latencies of Experiment 1. It does explain the latencies of Experiments 2 and 3, but comparisons between the models with the categorical measure and the models with the gradient measure, based on likelihood ratio tests, show that the models with the gradient variable outperform the ones with the categorical variable (Experiment 2: gradient measure: AIC = -110602, BIC = -110504; categorical measure: AIC = -110596, BIC = -110498; Experiment 3: gradient measure: AIC = 1646.91, BIC = 1704.78; categorical measure: AIC = 1647.41, BIC = 1705.27). Our experimental results thus support a gradient effect of variant relative frequency.

The conclusion that schwa words are stored with two phonological representations is in line with data from an off-line task reported in the literature. Racine and Grosjean (2002) showed that French speakers are quite good at estimating how often they produce a particular French schwa word with and without schwa. The correlation between participants' estimates

of the relative frequencies and the frequencies of the variants observed in a speech corpus was 0.47. This off-line experiment thus shows that speakers store frequency information about the two pronunciation variants of schwa words.

While our data only allow us to draw conclusions about the production lexicon, two comprehension studies suggest the same conclusions for the comprehension lexicon. Ranbom & Connine (2007) showed that listeners recognize the realization of /nt/ as a nasal flap in American English more quickly for words that are more often produced with this nasal flap (instead of [nt]). Pitt and colleagues (submitted) investigated the recognition of American English words with word-medial /t/ and documented that response latencies to the different pronunciations of these words correlate with the absolute frequencies of these pronunciation variants for the words.

The similarity between the results for these comprehension studies and our results for production is striking. It is in line with the hypothesis that there is just one lexicon for both production and comprehension (e.g., Alport, 1984; MacKay, 1987). Alternatively, if there are separate phonological systems for production and comprehension (e.g., Dell, Schwartz, Martin, Saffran & Gagnon, 1997; Levelt, Roelofs & Meyer, 1999), these systems are very similar in that they use at least one lexical representation for each pronunciation variant of a word.

Our three production experiments have also begun to document the status of the phonological representations of the variants of a word as a function of context. We have clearly shown that contextual variables affect the status of the variants. In the picture naming tasks (Experiment 1 and 2), participants more often produced erroneously the schwa variant than the reduced variant for words in isolation, and they produced the schwa variants more quickly than the reduced variants. These experiments are relatively close to natural situations in which speakers name objects in isolation or in very short phrases. In these situations they

almost exclusively use schwa variants. In the symbol-word association learning task (Experiment 3), in contrast, there was no advantage for schwa variants whatsoever. This task is rather artificial, not comparable to any situation in our everyday life, and, as a consequence, it does not favor one variant type over the other. Hence, context appears to crucially affect the status of pronunciation variants in the production process.

Interestingly, a privileged status for canonical forms has also been reported in comprehension studies (e.g., Racine & Grosjean, 2005; Ernestus & Baayen, 2007; Ranbom & Connine, 2007). For comprehension, however, no attempt has been made so far to investigate the reasons for this privileged status.

Our finding that the mental lexicon may contain different representations for different pronunciation variants may be accommodated in models assuming abstract lexical representations. Traditional models (e.g., Levelt et al., 1999; Dell, 1986, 1990) assume that each word has a lexical representation consisting of two components, a lemma (i.e., a semantically and syntactically specified representation) and one lexeme (i.e., a phonological representation). If a word has more than one pronunciation variant, only one pronunciation is stored and the others are derived by means of phonological or phonetic processes. Our results suggest that these traditional models need to be extended and to store one lexeme for each phonological variant. This extension results in what we have called in the introduction of this paper an abstract variants model.

This extension is partially anticipated in Levelt (1989)'s account of highly frequent reduced forms. He suggested that a word may have two phonological representations (a reduced and an unreduced one) provided that the reduced pronunciation is highly frequent. Our data suggest that as far as French schwa is concerned, a variant need not to be highly frequent in order to be lexically stored. Moreover, the data on French schwa show that the

number of words with more than one phonological representation is far from negligible since schwa words are abundant in French.

The abstract variants models can also account for the role of variant relative frequency, if one assumes that the two variants are in competition. This assumption would be in line with recent results showing competition in production between phonologically similar words (Vitevich & Stamer, 2006). In addition, competition can account for the role of context. Variants that fit better with the context receive higher activation. Higher activation will not only lead to faster production latencies, but will also decrease the competition from the other variant. If the activation for the non-favored variant is much lower than that for the favored variant, the non-favored variant will not compete with the favored variant, and the latter in turn will thus not show a variant relative frequency effect.

Our results can also be accommodated within exemplar-based models. These models assume that every token of a word produced or perceived by the language user is stored in the mental lexicon. In order to account for the variant relative frequency effect reported in this study, these models have to assume that exemplars representing the same pronunciation variant are stored together. In addition to the word nodes proposed by Johnson (1997), exemplar-based models should thus incorporate variant nodes dominating all exemplars of a given variant. The way these models account for the observed variant relative frequency effect and the effect of context is probably very similar to the account that we sketched above for the abstract variants models; it would involve competition between the variant nodes and context increasing the activation of the favored variant node.

In conclusion, this research represents to our knowledge the first attempt to capture the psycholinguistic processes and representations underlying the production of pronunciation variants with on-line experimental methods. Our study clearly shows that on-line psycholinguistic investigations of the production of non-canonical pronunciation variants are

crucial for our understanding of the mental lexicon and speech encoding. Such studies provide information that cannot be obtained by the study of canonical forms alone. This study provides evidence that words with regular pronunciation variation, such as schwa words in French, are represented in the (production) lexicon with at least two lexemes, which requires modifications of current abstractionist and exemplar-based models. Furthermore, this research shows that the canonical form does not necessarily have a privileged status in the production process: context appears to be more important than the canonicity of the word form for production. Further research is necessary to determine the extent to which data on other phonological processes of variation, in French and other languages, lead to similar conclusions. Our results underline the need to go beyond the study of canonical speech to improve our models of the mental lexicon and speech encoding.

Acknowledgments

We would like to thank Emmanuel Cuche for his drawing of pictures for Experiments 1 and 2. This study was partly funded by a European Young Investigator Award to the second author and by the Marie Curie Research Training Network Sound to Sense.

References

Alario, F.-X. & Ferrand, L. (1999). A set of 400 pictures standardized for French: Norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. *Behavior Research Methods, Instruments, & Computers,* 31, 531-552.

Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H. & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers,* 36, 140-155.

Alport, D. A. (1984). Speech production and comprehension: One lexicon or two? In W. Prinz & A. F. Sanders (Eds.), *Cognition and motor processes* (pp. 209-228), Berlin: Springer-Verlag.

Baayen, R. H. (2008). *Analysing linguistic data. A practical introduction to statistics using R*. New York: Cambridge University Press.

Baayen, R. H., Davidson, D. J. & Bates, D. M. (2008). Mixed effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390-412.

Bachoud-Lévi, A.-C., Dupoux, E., Cohen, L. & Mehler, J. (1998). Where is the length effect? A cross-linguistic study of speech production. *Journal of Memory and Language*, 39, 331-346.

Barnes, J. & Kavitskaya, D. (2002). Phonetic analogy and schwa deletion in French. *Berkeley Linguistics Society*, 28, 39-50.

Bell, A., Brenier, J., Gregory, M., Girand, C. & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, 60(1), 92-111.

Boersma, P. & Weenink, D. Praat: Doing phonetics by computer [Computer program] (Version 4.6.15) Retrieved August 23, 2007, from http://www.praat.org/.

Boula de Mareüil, P., Yvon, F., d'Alessandro, C., Aubergé, V., Vaissière, J. & Amelot, A. A French phonetic lexicon with variants for speech and language processing. *Proceedings of LREC*, Athens, Greece (2000), pp. 273-276.

Box, G.E. & Cox, D.R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society*, Series B 26, 211-252.

Browman, C. P. & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49, 155-180.

Cholin, J., Levelt, W. J. M. & Schiller, N. O. (2006). Effects of syllable frequency in speech production. *Cognition*, 99, 205-235.

Chomsky, N. & Halle, M. (1968). *The sound pattern of English*. New York: Harper and Row.

Connine, C. M., Ranbom, L. J. & Patterson, D. J. (2008). Processing variant forms in spoken word recognition: The role of variant frequency. *Perception & Psychophysics*, 70, 403-411.

Costa, A., Navarrete, E. & Alario, F.-X. (2006). Accessing object names when producing complex noun phrases: Implications for models of lexical access. *Cognitiva*, 18, 3-23.

Côté, M.-H. & Morrison, G. S. (2007). The nature of the schwa/zero alternation in French clitics: Experimental and non-experimental evidence. *Journal of French language studies*,17, 159-186.

Dell, F. (1985). *Les règles et les sons*. Paris : Hermann.

Dell. G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93, 283-321.

Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes*, 5, 313-349.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M. & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, 104, 801-838.

Desrochers, A. & Bergeron, M. (2000). Valeurs de fréquence subjective et d'imagerie pour un échantillon de 1916 substantifs de la langue française. *Revue Canadienne de Psychologie Expérimentale*, 54, 274-325.

Ernestus, M. & Baayen, R. H. The comprehension of acoustically reduced morphologically complex words: The roles of deletion, duration and frequency of occurrence. *Proceedings of ICPHS*, Saarbrücken, Germany (2007), pp. 773-776.

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

Fouché, P. (1956). *Traité de Prononciation française*. Paris: Klincksieck.

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

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

Goldinger, S. D. (1998). Echoes of Echoes? An episodic theory of lexical access. *Psychological Review*, 105, 251-279.

Goldstein, H. (1987). *Multilevel models in educational and social research*. London: Griffin.
Goldstein, H. (1995). Hierarchical data modeling in the social sciences. *Journal of Educational and Behavioral Statistics*, 20, 201-204.

Hansen, A. B. (1994). Etude du E caduc - stabilisation en cours et variations lexicales. *Journal of French Language Studies*, 4, 25-54.

Johnson, K. (1997). The auditory/perceptual basis for speech segmentation. *Ohio State University Working Papers in Linguistics*, 50, 101–113.

Johnson, K. (2004). Massive reduction in conversational American English. In K. Yoneyama & K. Maekawa (Eds.), *Proceedings of the 10th International Symposium: Spontaneous Speech: Data and Analysis* (pp. 29-54), Tokyo, Japan.

Kessler, B., Treiman, R. & Mullennix, J. (2002). Phonetic biases in voice key response time measurements. *Journal of Memory and Language,* 47, 145-171.

Kuijpers, C., van Donselaar, W. & Cutler, A. Phonological variation: Epenthesis and deletion of schwa in Dutch. *Proceedings of ICSLP*, Philadelphia, USA (1996), pp. 149-152.

Laganaro, M. & Alario, F.-X. (2006). On the locus of the syllable frequency effect in language production. *Journal of Memory and Language*, 55, 178-196.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. M., Roelofs, A. & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1-75.

Levelt, W. J. M. & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition*, 50, 239-269.

MacKay, D. G. (1987). *The organization of perception and action: A theory for language and other cognitive skills*. New York: Springer-Verlag.

Meyer, A. S., Roelofs, A. & Levelt, W. J. M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language,* 48, 131-147.

New, B., Pallier, C., Ferrand, L. & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE. *L'Année Psychologique*, 101, 447-462. http://www.lexique.org

Pitt, M. A., Dilley, L. & Tat, M. (submitted). Exploring the role of exposure frequency in recognizing pronunciation variants.

Protopapas, A. (2007). Checkvocal, a program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavioral Research Methods*, 39, 859-862.

Racine, I. Effacement du schwa dans des mots lexicaux : constitution d'une base de données et analyse comparative. *Proceedings of the Journées d'études linguistiques*, Nantes, France (2007), pp. 119-124.

Racine, I. & Grosjean, F. (2000). Influence de l'effacement du schwa sur la reconnaissance des mots en parole continue. *L'Année Psychologique*, 100, 393-417.

Racine, I. & Grosjean, F. (2002). La production du E caduc facultatif est-elle prévisible? Un début de réponse. *Journal of French Language Studies*, 12, 307-326.

Racine, I. & Grosjean, F. (2005). Le coût de l'effacement du schwa lors de la reconnaissance des mots en français. *Canadian Journal of Experimental Psychology*, 59, 240-254.

Ranbom, L. J. & Connine, C. M. (2007). Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language*, 57, 273-298.

Rasbash, J. & Goldstein, H. (1994). Efficient analysis of mixed hierarchical and cross-classified random structures using a multilevel model. *Journal of Educational and Behavioural Statistics*, 19, 337-350.

Rialland, A. (1986). Schwa et Syllabes en Français. In L. Wetzels & E. Sezer (Eds.), *Studies in compensatory lenghtening* (pp. 187-226), Dordrecht, Holland: Foris Publications.

Roelofs, A. (2002). Syllable structure effects turn out to be word length effects: Comment on Santiagio et al. (2000). *Language and Cognitive Processes*, 17, 1-13.

Santiago, J., MacKay, D .G., Palma, A. & Rho, C. (2000). Sequential activation processes in producing words and syllables: Evidence from picture naming. *Language and Cognitive Processes,*15, 1-44.

Schneider, W., Eschman, A. & Zuccolotto, A. (2002). E-prime user's guide. Pitsburg, PA: Psychology Software Tools, Inc.

Smorodinsky, I. (1998). The phonology and phonetics of schwas in Parisian French: An articulatory analysis. *Texas Linguistic Forum*, 41, 161-171.

Snoeren, N. D., Segui, J. & Hallé, P. A. (2008). Perceptual processing of partially and fully assimilated words in French. *Journal of Experimental Psychology: Human Perception and Performance,* 34, 193-204.

Spinelli, E. & Gros-Balthazard, F. (2007). Phonotactic constraints help to overcome effects of schwa deletion in French. *Cognition*, 104, 397-406.

Tranel, B. (1981). *Concreteness in generative phonology: Evidence from French*. Berkeley, CA: University of California Press.

Vitevich, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28, 735-747.

Vitevitch, M. S., Armbrüster, J. & Chu, S. (2004). Sublexical and lexical representations in speech production: Effects of phonotactic probability and onset density. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 30, 514-529.

Vitevich, M. S. & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, 40, 374-408.

Vitevich, M. S. & Sommers, M. S. (2003). The facilitative influence of phonological similarity and neighborhood frequency in speech production in younger and older adults. *Memory and Cognition*, 31, 491-504.

Vitevich, M. S. & Stamer, M. K. (2006). The curious case of competition in Spanish speech production. *Language & Cognitive Processes*, 21, 760-770.

Appendix A. Word stimuli (Experiment 1 and 2)

Word	Schwa variant	Word	Schwa variant
belette 'weasel'	bəlɛt	**pelote** 'ball of wood'	pəlɔt
besace 'pouch'	bəzas	**peloton** 'pack'	pəlɔtɔ̃
cerises 'cherries'	səʀiz	**pelouse** 'lawn'	pəluz
cerisier 'cherry tree'	səʀizje	**peluches** 'plush'	pəlyʃ
chemin 'path'	ʃəmɛ̃	**pelure** 'peel'	pəlyʀ
cheminée 'chimney'	ʃəmine	**petit** 'small boy'	pəti
cheminot 'railroader'	ʃəminɔ	**petite** 'small girl'	pətit

Word	Schwa variant	Word	Schwa variant
chemise ‘shirt’	ʃəmiz	**recette** ‘recipe’	ʀəsɛt
chenil ‘kennel’	ʃənil	**recyclage*** ‘recycling’	ʀəsiklaʒ
chenille ‘caterpillar’	ʃənij	**redingote*** ‘frock coat’	ʀədɛ̃gɔt
cheval ‘horse’	ʃəval	**reflet** ‘reflection’	ʀəflɛ
chevalet ‘easel’	ʃəvalɛ	**regard** ‘look’	ʀəgaʀ
chevalier ‘knight’	ʃəvalje	**registre** ‘register’	ʀəʒistʀ
cheveux* ‘hair’	ʃəvø	**remède** ‘medicine’	ʀəmɛd
cheville ‘ankle’	ʃəvij	**remorque** ‘trailer’	ʀəmɔʀk
chevreuil ‘deer’	ʃəvʀœj	**renard**** ‘fox’	ʀənaʀ
demeure ‘residence’	dəmœʀ	**renoncule*** ‘buttercup’	ʀənɔ̃kyl
demoiselle ‘young lady’	dəmwazɛl	**repas** ‘meal’	ʀəpa
fenêtre ‘window’	fənɛtʀ	**repassage** ‘ironing’	ʀəpasaʒ
fenouil ‘fennel’	fənuj	**reporter*** ‘reporter’	ʀəpɔʀtɛʀ
gelée ‘jelly’	ʒəle	**requin** ‘shark’	ʀəkɛ̃
genou ‘knee’	ʒənu	**ressort** ‘spring’	ʀəsɔʀ
genouillère ‘knee pad’	ʒənujɛʀ	**revues** ‘magazines’	ʀəvy
guenon* ‘female	gənɔ̃	**secouriste** ‘rescuer’	səkuʀist

Word	Schwa variant	Word	Schwa variant
monkey’			
jetons ‘tokens’	ʒətɔ̃	**secrétaire** ‘secretary’	səkʀɛtɛʀ
levier ‘lever’	ləvje	**semaine** ‘week’	səmɛn
levure ‘yeast’	ləvyʀ	**semelle** ‘sole’	səmɛl
melon ‘melon’	məlɔ̃	**semoule** ‘semolina’	səmul
menottes* ‘handcuffs’	mənɔt	**seringue** ‘syringe’	səʀɛ̃g
menu* ‘menu’	məny	**tenaille** ‘pincers’	tənaj
menuisier ‘carpenter’	mənɥizje	**velours** ‘velvet’	vəluʀ
meringue ‘meringue’	məʀɛ̃g	**venin** ‘venom’	vənɛ̃

* Words not in Experiment 2

** Word changed into ‘renardeau’ in Experiment 2

Appendix B. Words and pseudowords stimuli (Experiment 3)

	Word		Pseudoword	
	Schwa variant	Reduced variant	As schwa variant	As reduced variant
belette ‘weasel’	bəlɛt	blɛt	balyt	blyt
chenapan ‘scallywag’	ʃənapã	ʃnapã	ʃunipa	ʃnipa
cheveu ‘hair’	ʃəvø	ʃvø	ʃɛvɔ̃	ʃvɔ̃

	Word		Pseudoword	
	Schwa variant	**Reduced variant**	**As schwa variant**	**As reduced variant**
demeure ‘residence’	dəmœʀ	dmœʀ	dimuʀ	dmuʀ
demoiselle ‘young lady’	dəmwazɛl	dmwazɛl	dimɥizal	dmɥizal
genou ‘knee’	ʒənu	ʒnu	ʒina	ʒna
genouillère ‘knee pad’	ʒənujɛʀ	ʒnujɛʀ	ʒanajɛʀ	ʒnajɛʀ
guenon ‘female monkey’	gənɔ̃	gnɔ̃	gani	gni
jeton ‘token’	ʒətɔ̃	ʒtɔ̃	ʒatɔ	ʒtɔ
melon ‘melon’	məlɔ̃	mlɔ̃	məly	mly
menace ‘threat’	mənas	mnas	munis	mnis
menotte ‘handcuff’	mənɔt	mnɔt	minat	mnat
menuisier ‘carpenter’	mənɥizje	mnɥizje	minwazjo	mnwazjo
meringue ‘meringue’	məʀɛ̃g	mʀɛ̃g	miʀag	mʀag
peloton ‘pack’	pəlɔtɔ̃	plɔtɔ̃	pɔlata	plata
peluche ‘plush’	pəlyʃ	plyʃ	paliʃ	pliʃ
pelure ‘peel’	pəlyʀ	plyʀ	puliʀ	pliʀ
pesanteur ‘gravity’	pəzɑ̃tœʀ	pzɑ̃tœʀ	pazityʀ	pzityʀ

	Word		Pseudoword	
	Schwa variant	**Reduced variant**	**As schwa variant**	**As reduced variant**
querelle ‘quarrel’	kəʀɛl	kʀɛl	kaʀɛ̃l	kʀɛ̃l
rebelle ‘rebel’	ʀəbɛl	ʀbɛl	ʀibɔl	ʀbɔl
redingote ‘frock coat’	ʀədɛ̃gɔt	ʀdɛ̃gɔt	ʀɔ̃digat	ʀdigat
relique ‘relic’	ʀəlik	ʀlik	ʀalɛk	ʀlɛk
remorque ‘trailer’	ʀəmɔʀk	ʀmɔʀk	ʀymiʀk	ʀmiʀk
reporter ‘reporter’	ʀəpɔʀtɛʀ	ʀpɔʀtɛʀ	ʀapiʀtɛʀ	ʀpiʀtɛʀ
requête ‘request’	ʀəkɛt	ʀkɛt	ʀykat	ʀkat
requin ‘shark’	ʀəkɛ̃	ʀkɛ̃	ʀikø	ʀkø
revue ‘magazine’	ʀəvy	ʀvy	ʀavø	ʀvø
tenaille ‘pincers’	tənaj	tnaj	tanij	tnij
velours ‘velvet’	vəluʀ	vluʀ	valyʀ	vlyʀ
venin ‘venom’	vənɛ̃	vnɛ̃	vynu	vnu