The effect of schwa duration on pre-schwa mid-vowel lowering in French

Benjamin Storme
Massachusetts Institute of Technology

1 Introduction

In Southern French, mid vowels follow the *loi de position*: close-mid vowels ([e], [ø], [o]) and open-mid vowels ([ɛ], [œ], [ɔ]) are in allophonic distribution, with close-mid vowels occurring in open syllables not followed by schwa, e.g. *hôtel* [**o**tɛl], and open-mid vowels occurring in open syllables followed by schwa and in closed syllables, e.g. *hôtelier* [**ɔ**təlje] and *optique* [**ɔ**ptik]. The fact that open syllables followed by schwa pattern with closed rather than open syllables has long puzzled phonologists (Durand, 1976; Anderson, 1982; Scheer, 2006; Eychenne, 2014).

In this paper, I propose an analysis of the special status of schwa in the *loi de position* following the general hypothesis that mid-vowel lowering in closed syllables in French is perceptually motivated. In Storme (2015), I proposed that mid-vowel lowering is a strategy to enhance the perceptibility of a following consonant which lacks good release cues because it is followed by a segment without formant structure (e.g., by a stop C_2 in a VC_1C_2 sequence). This hypothesis is supported by the results of a perception experiment showing that French hearers are better at identifying stop place contrasts ([p] vs. [t] vs. [k]) in coda position after a low vowel than after a high vowel ([a_na] vs. [u_na] and [i_na]). I also found that the disadvantage of high vowels with respect to low vowels disappears before a liquid ([a_ʁa], [u_ʁa], [i_ʁa]), explaining why vowel lowering does not occur before pre-liquid stops. Close-mid vowels are preferred by default because they are further apart in the acoustic space (Storme, in press), and therefore more distinct from each other than open-mid vowels.

Why does mid-vowel lowering also happen before a pre-schwa consonant in Southern French? French schwa is a short vowel: Fougeron et al. (2007) found that [ə] is 50 ms long in French whereas [ø] and [œ], vowels with formant values similar to [ə], are 67 ms and 65 ms long, respectively. As a short vowel, schwa should provide shorter release transitions to a preceding consonant than a full vowel does. This consonant might then be in need of having its place-features enhanced, explaining why a mid vowel which precedes it is lowered. In this paper, I test one prediction of this approach, namely that the likelihood of lowering pre-schwa mid vowels is correlated to schwa duration: as schwa duration increases, longer release transitions become available to cue the preceding consonant and the need to lower the mid vowel preceding that consonant to improve its perceptibility should decrease.

Section 2 describes the experiment which was run to test this hypothesis. Section 3 presents the results. In section 4, the relationship between pre-schwa lowering and schwa duration is derived using Flemming's (2008) "Realized input" model of phonology, which allows for speaker-specific details of phonetic realization to influence phonotactics.

2 Method

2.1 *Participants* Ten French speakers from different regions in France participated in the experiment on a voluntary basis (7 males and 3 females). Two of them came from Montpellier (Hérault), three from Rodez (Aveyron), three from Clermont-Ferrand (Puy-de-Dôme), and two from Lille (Nord). Montpellier and Rodez are in the South, Clermont-Ferrand in the center, and Lille in the North. Speakers from different regions in

* I would like to thank Edward Flemming in particular for his very helpful feedback on this project. Thanks also to Michael Kenstowicz, Donca Steriade, Alan Yu, the audience at the MIT Phonology circle and at the *Journées FLORaL-PFC 2016* for helpful comments. All errors are mine.

Proceedings of *AMP 2016*

France were chosen because the probability of pre-schwa lowering is expected to vary by region: Southern speakers typically lower mid vowels before schwa whereas Northern speakers typically do not.

2.2 *Stimuli and recordings* Participants were recorded pronouncing mid vowels in syllable-initial positions in three groups of words: (i) words in which mid vowels occurred in closed syllables (e.g. *optique*), (ii) words in which mid vowels occurred in open syllables followed by schwa (e.g. *hôtelier*), and (iii) words in which mid vowels occurred in open syllables not followed by schwa (e.g. *hôtel*). Group (ii) included both words where schwa is deletable (e.g. *éperon*) and undeletable (e.g. *hôtelier*). In this study, only words with undeletable schwas in group (ii) are analyzed because only those words are consistently realized with schwas by all French speakers. Each word was pronounced in the carrier sentence *Le mot x existe*. Three lists were built, with each list containing all the sentences in pseudo-random order. The target words in the three conditions were:

- closed syllables: *celtique, coltiner, destin, hostile, optique, sceptique*
- open syllables followed by a full vowel: *brevet, échelle, élève, épèle, grommelle, hoquette, hôtel, modèle, tonneau*
- open syllables followed by undeletable schwa: *écrevisse, édredon, hôtelier, saugrenu, sommelier, tonnelier*

Each word was read three times by each speaker, yielding a total of $10 * 3 * (6 + 9 + 6) = 630$ data points.

Speakers were recorded in a quiet room with a head-mounted Shure SM35-XLR connected to a computer via a Shure X2u XLR-to-USB signal adapter. The recordings were made using the Audacity software, with 44 kHz/16 bit sampling. The distance (approx. 5 cm) of the participants to the microphone was held constant across all recording sessions.

2.3 *Phonetic analyses and coding* Two phonetically-trained native French speakers independently judged mid vowel realizations on a 0-100 scale, 0 corresponding to a prototypical close-mid vowel and 100 to a prototypical open-mid vowel. Mid vowel quality was established that way for two main reasons. The close-mid/open-mid distinction does not correspond to a single acoustic variable but to a combination of F1 and F2: open-mid vowels have higher F1 values and more central F2 values than close-mid vowels (Storme, in press). Because the close-mid/open-mid distinction is accompanied by lowering of F2 for front vowels and by raising of F2 for back vowels, it is hard to construct a numerical measure capturing this distinction.

In addition, the consonantal context is expected to affect mid vowel quality quite strongly: for instance, open-mid vowels have similar F1 values before coda [l] as close-mid vowels before onset [ʁ] (Storme, in press). This makes it difficult to use formants only to identify open-mid vowels vs. closed-mid vowels in the absence of control for vowel-consonant co-articulation.

Using judges was a way to avoid these two problems. I found a significant positive correlation of 0.96 (t = 86.545, df = 628, p-value $<$ 2.2e-16) between the two judges' ratings. This led me to use an average of the two judges' ratings as a single measure of mid vowel quality. Because the judges tended to use extreme values of the scale (80 per cent of the ratings were either 0 or 100 across the two judges), averaged numerical judgments were turned into binary judgments, 'close-mid' (if < 50) vs. 'open-mid' (if > 50).

Schwa duration was measured in Praat. Measures of vowel duration included the vocalic segment only, identified through voicing and changes in formant trajectories.

3 Results

Figure 1 shows the posterior probability of lowering mid vowels by context and by speaker.

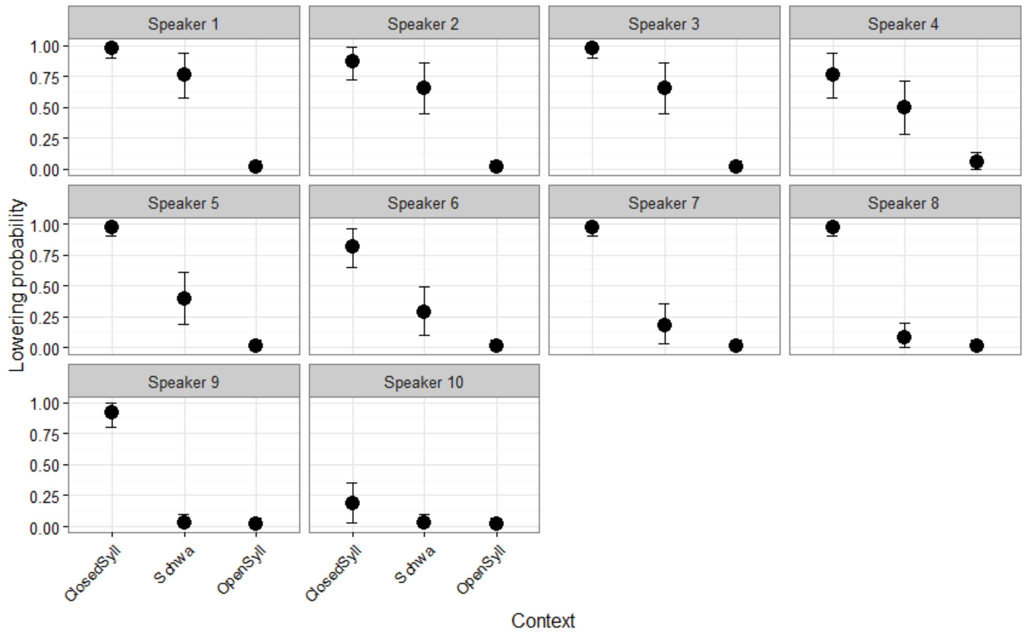


Figure 1: Posterior mean lowering probability by speaker and by context (closed syllables vs. open syllables followed by schwa vs. open syllables not followed by schwa), with 95% highest posterior density interval (HPD). The posterior mean and HPD were computed using the `binom` package (Dorai-Raj, 2014) in R with Jeffrey's prior, which is a Beta(0.5, 0.5) distribution.

The probability of lowering mid vowels in open syllables not followed by schwa is near zero for all speakers. Mid vowels are more likely to be open-mid in closed syllables for all speakers except Speaker 10. For most speakers aside Speaker 10, this probability is close to 1. These results provide further evidence for the hypothesis that French speakers favor close-mid vowels in open syllables not followed by schwa and open-mid vowels in closed syllables. However, they also show that lowering in closed syllables is not universal in French (see Speaker 10).

The probability of pre-schwa lowering varies across speakers. For Speakers 1 and 2, it does not significantly differ from the probability of lowering in closed syllables. For Speakers 8, 9, and 10, it does not significantly differ from the probability of lowering in open syllables not followed by schwa. For the other speakers, the pre-schwa lowering probability has an intermediary value between the two other probabilities.

For all speakers, the probability of lowering before schwa is never greater than the probability of lowering in closed syllables. This is in accordance with the general observation that there seems to be no French variety in which lowering before schwa is attested but lowering in closed syllables is not.

Figure 2 shows the distribution of schwa durations by speaker.

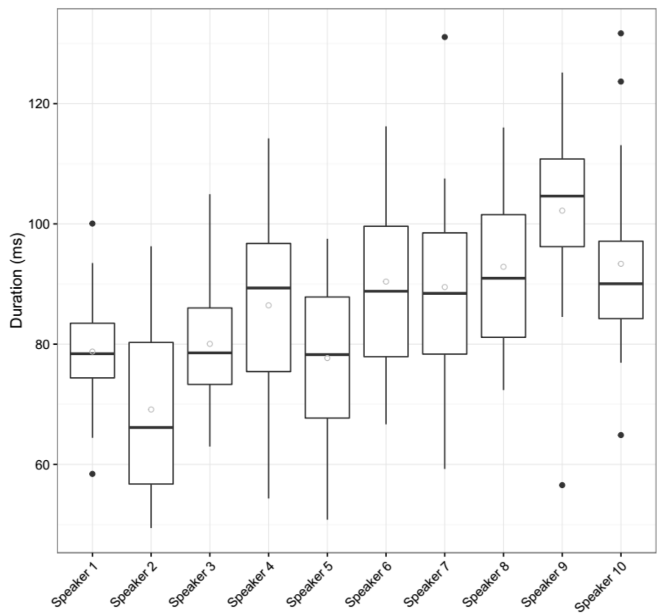


Figure 2: Schwa duration (ms) by speaker. Inside each plot, the boxes indicate the inter-quartile range (IQR), the range between the first and third quartile. The horizontal line indicates the median and the grey circle indicates the mean. The whiskers indicate the range, up to 1.5 times the IQR away from the median. Dot outside the whiskers lie more than 1.5 times the IQR away from the median and are potential outliers.

Overall, schwa duration was found to be longer than in other studies. For instance, Fougeron et al. (2007) and Burki et al. (2011) found that schwa duration is 50 ms on average. In this study, schwas are on average 86 ± 17 ms long. There are several factors which could explain this difference. First, both Fougeron et al. (2007) and Burki et al. (2011) are corpus-based studies of natural speech and therefore speech rate is likely to be faster. Second, schwas in these studies occur in a much wider range of prosodic contexts than in the present study and this could have an effect on mean schwa duration.

The hypothesis that the probability of pre-schwa lowering is correlated with mean schwa duration across speakers was tested as follows. Schwa duration was modeled as a linear function of speaker. The binary variable 'close-mid' vs. 'open-mid' was modeled as a function of speaker in a logistic regression. The parameters of the two regressions were estimated in a Bayesian way, using the R package `MCMCpack` (Martin et al., 2011). A Bayesian analysis was adopted because the parameters of the logistic regression fit in a frequentist framework had very large standard deviations, in particular for Speaker 9 who never lowers any pre-schwa mid vowel. Putting a prior on the standard deviation for the parameters helped solving that issue. Uninformative priors were used in both regressions, with a mean equal to zero for the intercepts and the speaker parameters, a precision equal to 0 for the linear regression and equal to 0.01 for the logistic regression. For each regression, I used three MCMC chains with 100,000 samples, and a thinning interval of 10. The first 10,000 samples of each chain were used for burn-in. Convergence of the chains on the posterior distribution was assessed using the Gelman-Rubin statistic: it was equal to 1 in both regressions, indicating that the samples are representative of the posterior distribution.

Figure 3 shows the relationship between the coefficients of the linear regression predicting schwa duration and the logistic regression predicting lowering probability. Speaker 10 was excluded from this analysis: as he hardly lowers mid vowels in any context, the absence of pre-schwa lowering could be due to a general dispreference for lowering rather than to his schwas being longer overall.

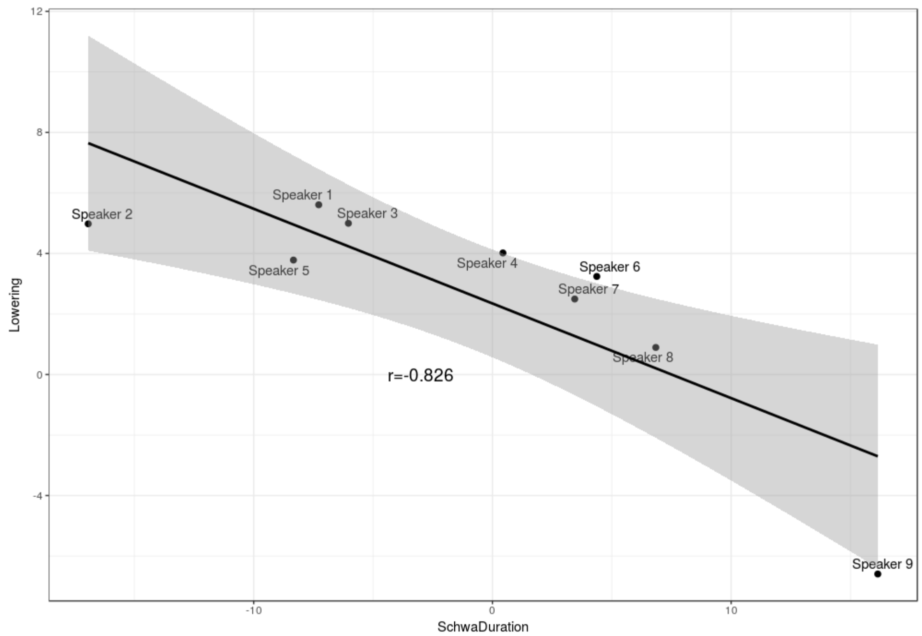


Figure 3: Pre-schwa lowering as a function of schwa duration. On the x-axis: the parameters indicate individual deviations from the mean schwa duration across speakers. On the y-axis: the parameters indicate individual deviations from the mean log-transformed odds of lowering pre-schwa mid vowels across speakers. Larger y-values correspond to higher probabilities of lowering pre-schwa mid vowels.

There is a significant negative correlation between schwa duration and lowering, r=-0.822 (t = -3.8132, df = 7, p-value = 0.0066): as the mean schwa duration of a speaker decreases, the likelihood that that speaker will lower pre-schwa mid vowels increases.[1] This is in accordance with the hypothesis presented in the introduction.

4 Discussion

In this section, I propose to model the relationship between schwa duration and pre-schwa lowering using Flemming's (2008) "Realized input" model of phonology. This model posits an intermediary level, called the realized input, between underlying phonological representations and surface phonological representations, allowing for speaker-specific details of phonetic realization (e.g., schwa duration) to influence phonotactics (e.g., the distribution of mid vowels before schwa).

In section 4.1, I show how different schwa durations can be derived for different speakers in a phonetic grammar with continuous constraints. In section 4.2, I show how the relationship between schwa duration and pre-schwa lowering can be derived in a phonological grammar taking the output of the phonetic grammar in section 4.1 as input and including categorical constraints om mid vowel quality referring to segment duration.

4.1 *Phonetic grammar* The observation that schwa duration may vary across speakers requires to derive schwa duration in a grammar. I adopt Flemming's (2001) Harmonic grammar framework to evaluate candidates in a continuous space and implement it in a Maximum Entropy framework (Hayes & Wilson, 2008) in order to derive within-speaker variation in the realization of schwas.

The realization of schwa duration is determined by two conflicting constraints. SHORTSCHWA, a faithfulness constraint, requires schwa to be realized as a short vowel. A vowel qualifies as short if its duration is equal to $T_{V_{short}}$, the durational target for short vowels. LONGV, a markedness constraint, requires

[1] To make sure that this is an effect of schwa duration rather than speech rate, I normalized schwa duration by dividing it by word duration and tested the correlation between that normalized schwa duration and pre-schwa lowering. I also found a significant negative correlation, indicating that the effect cannot be attributed to speech rate. The effect is not driven by Speaker 9 alone, as pre-schwa lowering and schwa duration remain negatively correlated when Speaker 9 is excluded.

vowels to be realized as long. A vowel qualifies as long if its duration is equal to $T_{V_{long}}$, the durational target for long vowels, with $T_{V_{long}} > T_{V_{short}}$.

The two constraints and the cost incurred by candidates violating them are shown in Table 1. The cost incurred by a candidate D(ə) violating one of the constraints on vowel duration is obtained by multiplying the weight for that specific constraint by the square of the deviation from the target (Flemming, 2001). This ensures that the more D(ə) deviates from the target specified by the relevant constraint on vowel duration, the more penalized it is by that constraint.

	Constraint	Cost of violation
SHORTSCHWA	$D(ə) = T_{V_{short}}$	$w_s(D(ə) - T_{V_{short}})^2$
LONGV	$D(ə) = T_{V_{long}}$	$w_l(D(ə) - T_{V_{long}})^2$

Table 1: Constraints and cost of violations. Weights are positive real numbers.

The overall cost incurred by a candidate D(ə), noted h(D(ə)), is equal to the sum of the violation costs.

$$h(D(ə)) \quad = w_s(D(ə) - T_{V_{short}})^2 + w_l(D(ə) - T_{V_{long}})^2$$

Flemming (2001) shows that the optimal realization, i.e. the one which minimizes h(D(ə)), is equal to the weighted sum of the target for long vowels and the target for short vowels, i.e. $(w_s * T_{V_{short}} + w_l * T_{V_{long}})/(w_s + w_l)$. Assuming fixed targets for long and short vowels, different speakers will have different schwa durations if they have different weights w_s and w_l.

There is a single constraint regulating the duration of non-schwa vowels, LONGV.[2] The analysis predicts that the optimal duration for non-schwa vowels is equal to the target for long vowels, $T_{V_{long}}$.

$$h(D(V_{\neq ə})) \quad = w_l(D(V) - T_{V_{long}})^2$$

This framework only derives a single vowel duration by speaker and vowel. However, speakers produce different vowel durations for a given vowel category and the distribution of durations typically follows a normal distribution (see Figure 2). To make the phonetic grammar stochastic, I adopt the Maximum Entropy framework for Harmonic Grammars described in Hayes & Wilson (2008). A Maximum Entropy Harmonic Grammar defines a probability distribution over a space of candidates. The probability of a given candidate is computed in two steps. First, the maxent value $P^*(x)$ of each candidate x is computed based on the harmony score of that candidate, $h(x)$.

$$P^*(x) = exp(-h(x))$$

The probability of x, $P(x)$, is calculated by determining its share in the total maxent values of all possible forms in the space of candidates Ω, a quantity designated as Z. Candidates with higher harmony scores have lower maxent values and therefore lower probabilities.

$$P(x) = \frac{P^*(x)}{Z}$$

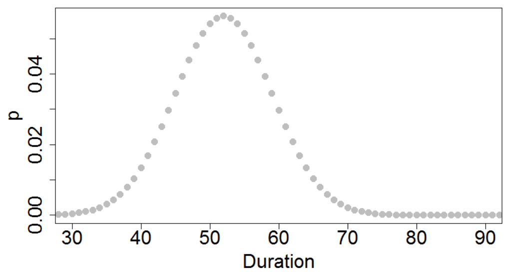
$$\text{where } Z = \sum_{y \in \Omega} P^*(y)$$

As duration is continuous, the candidate space for schwa durations is infinite and the probability of all candidates cannot be calculated as a sum. To overcome this problem, I use a discrete variable as a proxy for duration: the continuum between 0 and 100 ms is represented as a vector of integers between 0 and 100.

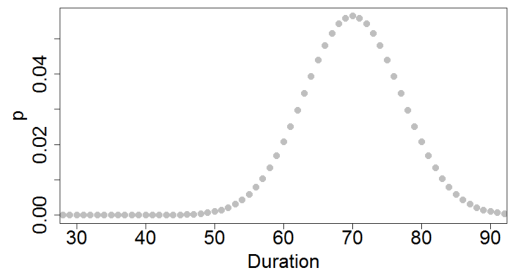
Figure 4 shows the probability distribution over (schwa and non-schwa) vowel durations predicted by the grammar for two speakers with different weights w_s and w_l. I call these two speakers Speaker A and Speaker B. They illustrate the kind of variation observed for real French speakers in the experiment. The model derives normal distributions for vowel duration. Speaker A gives a higher priority to realizing schwas as

[2] To simplify, I ignore the fact that vowel duration varies with height, with higher vowels being shorter than lower vowels (Lehiste, 1970).

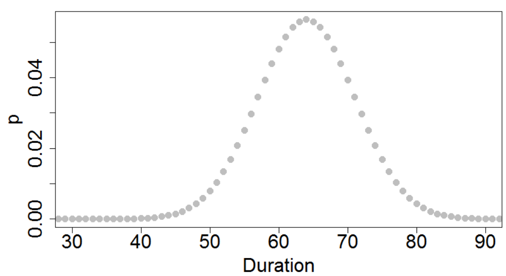
short vowels than Speaker B (the weight w_s is larger for Speaker A), resulting in the mean of the distribution being smaller. For both speakers A and B, schwas are shorter that non-schwa vowels. Keeping the sum of the weights constant ensures that the standard deviation remains equal across speakers and across vowels.



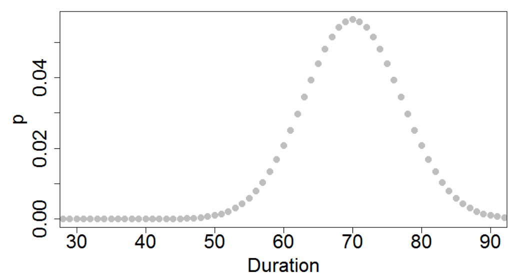
(a) Speaker A, schwa
($w_s = 0.009$, $w_l = 0.001$)



(b) Speaker A, non-schwa vowel
($w_l = 0.009 + 0.001$)



(c) Speaker B, schwa
($w_s = 0.003$, $w_l = 0.007$)



(d) Speaker B, non-schwa vowel
($w_l = 0.003 + 0.007$)

Figure 4: Schwa and non-schwa durations for speakers with different weights w_l and w_s. Targets for long and short vowels are fixed across speakers: $T_{V_{long}}$=70 ms, $T_{V_{short}}$=50 ms

4.2 *Phonological grammar* The output of the phonetic grammar serves as input to the phonological grammar. I propose to analyze the distribution of mid vowels with two markedness constraints, *OPENMIDV and *CLOSEMIDV/_C$_{\text{PRT}}$, where C$_{\text{PRT}}$ stands for "consonant with poor release transitions". *OPENMIDV is a general constraint against open-mid vowels. *CLOSEMIDV/_C$_{\text{PRT}}$ penalizes close-mid vowels before consonants which are poorly cued by the subsequent context, either because they lack long enough release transitions or because they lack release transitions. Only vowels with a duration greater than a durational threshold θ ($D(V) > \theta$) provide long enough release transitions. *OPENMIDV and *CLOSEMIDV/_C$_{\text{PRT}}$ are formulated as markedness constraints, but can be conceptually interpreted as constraints on contrasts (Flemming, 2002).

The two constraints and the cost incurred by candidates violating them are shown in Table 2. The cost incurred by a candidate violating one of the constraints on mid vowel quality is equal to the weight for that specific constraint.

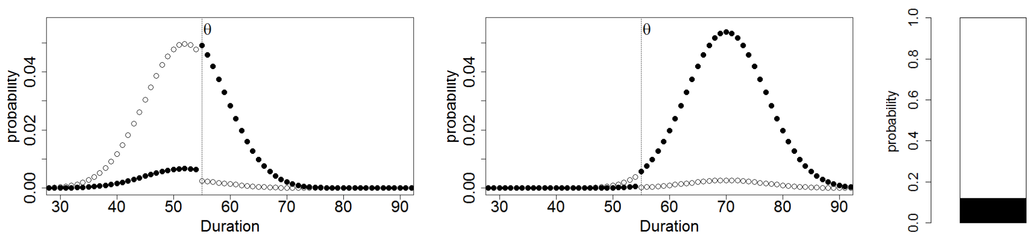
	Constraint	Cost of violation
*OPENMIDV	mid vowels are close	w_o
*CLOSEMIDV/_C$_{\text{PRT}}$	mid vowels are open before C_1C_2 and $CV_{D(V)<\theta}$	w_c

Table 2: Constraints and cost of violations. Weights are positive real numbers.

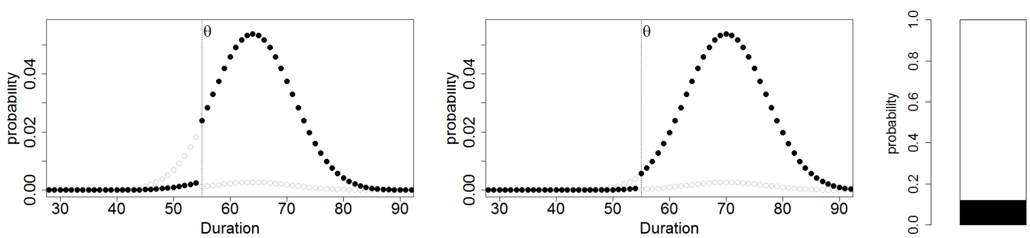
Close-mid vowels in open syllables and followed by a vowel with duration D(V) will violate *CLOSEMIDV/_C$_{\text{PRT}}$ depending on whether D(V) is longer or shorter than the threshold θ. Open-mid vowels will always violate *OPENMIDV regardless of the following context. The harmony score for a sequence containing a mid vowel varies as a function of the quality of that mid vowel (close-mid vs. open-mid), its segmental context, and the duration of the vowel in the following syllable if the mid vowel occurs in an open syllable:

$$\begin{cases} h(V_{close-mid}CV) & = w_c \quad \text{if } D(V) < \theta \\ & = 0 \quad \text{if } D(V) > \theta \\ h(V_{close-mid}CC) & = w_c \\ h(V_{open-mid}) & = w_o \end{cases}$$

The phonological grammar is also implemented in a Maximum Entropy framework to allow for variation (see Figure 1 for evidence for variation). Figure 5 shows the probabilities of close-mid vs. open-mid realizations predicted by the model in the three relevant contexts for Speaker A and Speaker B, assuming $w_c > w_o$. The phonological grammar operates on segments whose durations have already been fixed by the phonetic grammar and therefore the distributions of vowel durations are the same as in Figure 4.



(a) Speaker A, VCə **(b)** Speaker A, VCV **(c)** Speaker A, VCC



(d) Speaker B, VCə **(e)** Speaker B, VCV **(f)** Speaker B, VCC

Figure 5: Probability of mid vowels being close-mid (in black) vs. open-mid (in white) in open syllables before schwa (VCə), in open syllables followed by a non-schwa vowel (VCV), and in closed syllables (VCC). Speaker A and B have the same phonological grammar (w_c=5, w_o=3, θ=55 ms).

Speaker A and Speaker B have the same phonological grammar (w_c=5, w_o=3, θ=55 ms). However, because they have different phonetic grammars and, as a result, different schwa durations, they have different pre-schwa lowering probabilities: Speaker A is more likely to lower pre-schwa mid vowels than Speaker B because more of the distribution of schwa durations is below the threshold θ for Speaker A. Non-schwa vowels are long enough so that most of their distribution is above the threshold θ. As a consequence, mid vowels are not lowered in open syllables before non-schwa vowels for either speaker. In closed syllables, mid-vowels are much more likely to be realized as open-mid than close-mid because $w_c > w_o$ for both speakers.

The model correctly predicts that the likelihood of lowering pre-schwa mid vowels in open syllables is inversely correlated to mean schwa vowel duration across speakers: as the mean schwa duration for a given speaker becomes smaller, the likelihood that that speaker produces schwas with a duration smaller than θ increases and, as a consequence, the likelihood that she will lower pre-schwa mid vowels increases. The model further predicts that the correlation should be found within each speaker. The data are compatible with this hypothesis, but more data should be collected to confirm it.

The model also predicts that pre-schwa lowering entails lowering in closed syllables in a statistical sense, i.e. the probability of pre-schwa lowering is never larger than the probability of lowering in closed syllables. This is because close-mid vowels in closed syllables always violate *CLOSEMIDV/_C_{PRT} whereas close-mid vowels in open syllables followed by schwa violate *CLOSEMIDV/_C_{PRT} only for schwa durations below θ. This prediction is welcome, as the probability of pre-schwa lowering was never found to be larger than the probability of lowering in closed syllables in the experiment (see Figure 1) and pre-schwa mid-vowel lowering is only reported in a variety of French (i.e. Southern French) which also has mid-vowel lowering in closed syllables.

References

Anderson, Stephen R. (1982). The Analysis of French Schwa, or How to Get Something for Nothing. *Language* 58, 534–573.

Burki, Audrey, Cécile Fougeron, Cedric Gendrot & Ulrich H. Frauenfelder (2011). Phonetic reduction versus phonological deletion of French schwa: Some methodological issues. *Journal of Phonetics* 39, 279–288.

Dorai-Raj, Sundar (2014). *binom: Binomial Confidence Intervals For Several Parameterizations*. URL `https://CRAN.R-project.org/package=binom`. R package version 1.1-1.

Durand, Jacques (1976). Generative phonology, dependency phonology and southern french. *Lingua e Stile* 11, 3–23.

Eychenne, Julien (2014). Schwa and the *loi de position* in southern french. *French Language Studies* 24, 223–253.

Flemming, Edward (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18, 7–44.

Flemming, Edward (2002). *Auditory Representations in Phonology*. Routledge.

Flemming, Edward (2008). The Realized Input.

Fougeron, Cécile, Cédric Gendrot & Audrey Burki (2007). On the phonetic identity of french schwa compared to /ø/ and /œ/.

Hayes, Bruce & Colin Wilson (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39, 379–440.

Lehiste, Ilse (1970). *Suprasegmentals*. The MIT Press, Cambridge.

Martin, Andrew D., Kevin M. Quinn & Jong Hee Park (2011). MCMCpack: Markov chain monte carlo in R. *Journal of Statistical Software* 42:9, p. 22, URL `http://www.jstatsoft.org/v42/i09/`.

Scheer, Tobias (2006). How yers made Lightner, Gussmann, Rubach, Spencer and others invent CVCV. *Studies in Constraint-based Phonology*.

Storme, Benjamin (2015). The *loi de position* in French. Generals paper.

Storme, Benjamin (in press). The *loi de position* and the acoustics of French mid vowels. *Glossa: a journal of general linguistics* .