

# How much does orthography influence the processing of reduced word forms? Evidence from novel-word learning about French schwa deletion

Malte C Viebahn<sup>1,2,3</sup>, James M McQueen<sup>1,4</sup>, Mirjam Ernestus<sup>1,2</sup>, Ulrich H Frauenfelder<sup>3</sup> and Audrey Bürki<sup>3,5</sup>

Quarterly Journal of Experimental Psychology  
2018, Vol. 71(11) 2378–2394  
© Experimental Psychology Society 2017  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/1747021817741859  
qjep.sagepub.com



## Abstract

This study examines the influence of orthography on the processing of reduced word forms. For this purpose, we compared the impact of phonological variation with the impact of spelling-sound consistency on the processing of words that may be produced with or without the vowel schwa. Participants learnt novel French words in which the vowel schwa was present or absent in the first syllable. In Experiment 1, the words were consistently produced without schwa or produced in a variable manner (i.e., sometimes produced with and sometimes produced without schwa). In Experiment 2, words were always produced in a consistent manner, but an orthographic exposure phase was included in which words that were produced without schwa were either spelled with or without the letter <e>. Results from naming and eye-tracking tasks suggest that both phonological variation and spelling-sound consistency influence the processing of spoken novel words. However, the influence of phonological variation outweighs the effect of spelling-sound consistency. Our findings therefore suggest that the influence of orthography on the processing of reduced word forms is relatively small.

## Keywords

Novel-word learning; schwa deletion; orthography; spoken-word production; spoken-word recognition

Received: 20 February 2017; revised: 15 August 2017; accepted: 3 October 2017

## Introduction

An important characteristic of spoken language is the large amount of variability with which words are produced. This variability is to a large extent due to reductions that often occur in spontaneous speech (e.g., Ernestus, 2000; Johnson, 2004). As a result of such reductions, individual sounds and syllables may be either shortened or completely absent. One speech sound that is often affected by reduction processes is the vowel schwa. For example, the English word *sufficient* may be pronounced without schwa as in /sɪfɪjnt/ instead of /səfɪjnt/ and the French word *seringue* “syringe” may be pronounced as /srɛ̃g/ instead of /səɾɛ̃g/. Corpus studies of spontaneous speech have shown that these kinds of reductions are very common across different languages including English, Dutch, French, and German (for a review see Ernestus & Warner, 2011). The large amount of phonological variability that results from these reductions poses a challenge to psycholinguistic theories that attempt

to describe the processes and representations that allow language users to produce and comprehend spoken language (e.g., Luce & McLennan, 2005). The goal of the current study is to examine the role of orthography in the

<sup>1</sup>Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

<sup>2</sup>Centre for Language Studies, Radboud University, Nijmegen, The Netherlands

<sup>3</sup>Laboratoire de Psycholinguistique Expérimentale, University of Geneva, Geneva, Switzerland

<sup>4</sup>Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

<sup>5</sup>Department of Linguistics, University of Potsdam, Potsdam, Germany

### Corresponding author:

Malte C Viebahn, Laboratoire de Psycholinguistique Expérimentale, University of Geneva, Boulevard du Pont d’Arve, 40 1204 Geneva, Switzerland.

Email: malte.viebahn@gmail.com

processing of speech reduction by comparing the influence of spelling-sound consistency with the influence of phonological variation on word production and word recognition.

The variation in the spoken forms of words due to reductions contrasts with the consistency of orthographic forms. Although there is considerable variation in the handwriting of words, the number and order of the letters that words consist of remain invariant. For example, while a speaker of Dutch may produce the word *eigenlijk* “actually” as either /eɪxələk/, /eɪxlək/, or even /eɪx/ (Ernestus & Smith, in press), the orthographic form always remains <eigenlijk>.<sup>1</sup> Studies have shown that orthographic information can influence how spoken words are processed (e.g., Grainger, Muneaux, & Farioli Ziegler, 2005; Perre, Pattamadilok, Montant, & Ziegler, 2009; Seidenberg & Tanenhaus, 1979; Taft, Castles, Davis, Lazendic, & Nguyen-Hoan, 2008 but see also Cutler & Davis, 2012; Cutler, Treiman, & van Ooijen, 2010). Moreover, a word’s orthography has been shown to influence how phonological variation is processed. Using an artificial-word learning paradigm, Bürki, Spinelli, and Gaskell (2012) demonstrated that French listeners who learn novel words without schwa generate schwa variants of these words after having been exposed to orthographic forms that are consistent with the existence of a spoken variant that contains schwa. For example, French-speaking participants learnt the auditory form of the novel word /plur/ by associating it with a picture of a novel object. Afterwards, they were exposed to an orthographic form that either contained the letter <e> in the first syllable (<pelour>), or they saw an orthographic form without the letter <e> (<plour>). Words with an orthographic form that contained the letter <e> were named more slowly and were more likely to be produced with schwa compared to words without the letter <e>. These findings demonstrate an orthographic influence on the production of reduced word forms. Work by Racine, Bürki, and Spinelli (2013) suggest that orthographic knowledge can also influence the recognition of reduced word forms. In this study, the authors showed that the recognition of spoken schwa and non-schwa variants in French changes with orthographic knowledge. Whereas pre-readers recognise the more frequent variants faster, recognition times in readers are influenced not only by frequency but also by spelling.

While these data reveal an orthographic influence on spoken-word production and recognition processes, it is not yet known how large this influence is relative to the influence of phonological information. It is possible that the influence of orthography is quite large, particularly in the processing of reduced speech. For example, it has been proposed that the influence of orthography can, at least in part, explain why unreduced pronunciation variants are recognised more efficiently than reduced variants even if the reduced variants occur more frequently (e.g., Ranbom

& Connine, 2007). This suggests that the influence of orthography is large enough to neutralise the effect of auditory-variant frequency. The present study aims to extend our knowledge about the processing of reduced word forms by comparing the influence of phonological variation with the influence of spelling-sound consistency, both in word production and word recognition.

One reason to assume that orthography might have a strong influence on the processing of words is that the spelling of orthographic forms varies much less than the pronunciation of spoken forms. However, in order to get from an orthographic form to a phonological one, readers have to draw inferences about which sound corresponds to a given grapheme. Depending on the language, these inferences might be quite difficult to make (e.g., Ziegler, Jacobs, & Stone, 1996) because a given grapheme may correspond to multiple phonemes, or it can be silent. Especially in those languages without fully regular grapheme to sound mappings, such as French, phonological information, under good listening conditions, may be much less ambiguous and provide therefore more direct evidence for a particular speech sound. Thus, several possibilities are plausible: Orthography may be more influential than phonological information, orthography may be less influential than phonological information, or the two may have a similar influence. So far, no direct empirical comparison of the two sources of information has been made. Here, we asked whether the impact of phonological and orthographic information on the processing of reduced word forms differ in French. This will help us to gauge how large the role of orthography is relative to that of variability in the phonological input when listeners process reduced speech.

In order to examine this question, we conducted two experiments that focus on schwa deletion in French. In these experiments, we combined three experimental methodologies: a novel-word learning paradigm, a picture-naming task, and visual-world eye-tracking. As discussed above, the novel-word learning paradigm has previously been employed in combination with a picture naming task in order to study the production of reduced pronunciation variants (e.g., Bürki & Frauenfelder, 2012; Bürki et al., 2012). Other studies (e.g., Magnuson, Tanenhaus, Aslin, & Dahan, 2003; Sulpizio & McQueen, 2012) have combined novel-word learning with the visual-world paradigm in order to investigate the time course of the recognition of spoken words that have just been learnt. Combining these methodologies allows us to examine jointly the processing of newly learnt reduced pronunciation variants in language production and comprehension.

In our two experiments, participants learnt over a 3-day period novel French words in which schwa was either present (e.g., /səkɔb/) or absent (e.g., /skɔb/). Schwa presence was categorical, meaning that words that were produced without schwa did not contain any acoustic cues for schwa.

**Table 1.** Experimental conditions in Experiments 1 and 2. The examples shown in parentheses indicate the phonological form using IPA symbols with slashes and the orthographic form in angle brackets. (Note that each item occurred only in one condition for each participant.)

EXPERIMENT 1	EXPERIMENT 2
Experimental conditions	Experimental conditions
1. Spoken consistently without schwa (/skɔb/)	1. Spoken without schwa and spelled without <e> (/skɔb/ and <scobe>)
2. Spoken in a variable manner (/skɔb/ and /səkɔb/)	2. Spoken without schwa but spelled with <e> (/skɔb/ but <secobe>)
Filler condition	Filler condition
Spoken consistently with schwa (/səkɔb/)	Spoken with schwa and spelled with <e> (/səkɔb/ and <secobe>)

Note, however, that spontaneous speech typically contains different kinds of reduction, some of which are considered to be categorical and some gradient. There is currently a debate about the degree to which reductions should be considered to be categorical or gradient and whether they are phonological, phonetic, or acoustic in nature (e.g., Browman & Goldstein, 1992; Hanique, Ernestus, & Schuppler, 2013). French schwa deletion is considered to be categorical (e.g., Bürki, Ernestus, & Frauenfelder, 2010; Côté & Morrison, 2007).

In Experiment 1, we examined the influence of phonological variation on the learning of the novel words. The words were either consistently produced without schwa or produced in a variable manner (i.e., sometimes produced with and sometimes produced without schwa). In Experiment 2, we investigated the effect of spelling-sound consistency on the processing of reduced word forms building on Bürki et al. (2012) and using the same materials as in Experiment 1. In contrast to Experiment 1, words in Experiment 2 were consistently produced without schwa but an orthographic exposure phase was included in which words were either spelled consistently with the phonological form (i.e., *without* the letter <e> in the initial syllable) or inconsistently with the phonological form (*with* the letter <e> in the initial syllable).

This manipulation allowed us to compare the effect of phonological variation in Experiment 1 with a spelling-based effect. In both experiments, we conducted two tests: A picture-naming task in order to examine processing during word production and an eye-tracking task to examine processing during word recognition.

## Experiment 1

In Experiment 1, we taught participants novel French words that were either produced without schwa or in a variable manner. The items that were produced in a variable manner were produced half of the time with schwa and half of the time without schwa. We predicted that participants are more likely to produce a word with schwa if it had been presented in a variable manner compared to a word that had always been presented without schwa. More specifically, we expected that naming responses would follow the exposure proportions, such that words that had

been heard with schwa on 50% of exposure trials would tend to be named with schwa on 50% of naming test trials. In contrast, words that had always been heard without schwa during the exposure phase ought not be produced with schwa during the naming phase. Furthermore, we predicted that naming latencies would be longer for words that had been presented in a variable manner compared to words that had always been presented without schwa. With respect to recognition performance in the eye-tracking task, we expected that listeners would recognise an auditorily presented word more quickly if it had always been heard without schwa during the exposure phase than if it had sometimes been heard with schwa.

## Method

**Participants.** Thirty-one students of the University of Geneva who were native speakers of French were paid for their participation. The mean age was 23 years. Nine of the participants were male. None had known hearing problems, and all had normal or corrected-to-normal visual acuity.

**Design.** The experiment consisted of a learning phase and a test phase. The learning phase was performed over 3 days and the test phase was performed on the third day directly after the end of the learning phase. During the test phase, participants performed an eye-tracking task and a naming task. The design of the experiment is shown in Table 1. There were two experimental conditions: Words in the *without-schwa* condition were never produced with schwa in the initial syllable during the learning phase (e.g., /skɔb/) whereas words in the *variable* condition were presented 50% of the time during learning with schwa in the initial syllable (e.g., /səkɔb/) and 50% of the time without schwa (e.g., /skɔb/). In addition to these two experimental conditions, we added a filler condition in order to increase the overall number of trials in which a novel word was produced with schwa. In this *with-schwa* condition, the novel words were always presented with schwa (e.g., /səkɔb/).

During the eye-tracking task, participants were presented with four objects on each trial and were asked to click on one of them. Of the four objects, two objects belonged to the

same minimal pair (e.g., /səkɔb/—/səkɔf/), while the other two objects belonged to a different minimal pair (e.g., /bəlaʒɛ/—/bəlafɛ/). In the following, the word that the participants were instructed to click on (e.g., /səkɔb/) will be referred to as the target word while the other word that belonged to the same minimal pair (e.g., /səkɔf/) will be referred to as the competitor. The two words belonging to a different minimal pair will be referred to as the distractors.

During the eye-tracking task, the variable condition was split into two sub-conditions: variable words that were presented with schwa and variable words that were presented without schwa. In all three conditions—learnt without-schwa, learnt in a variable manner and presented with schwa, and learnt in a variable manner and presented without schwa—the competitor was a word that had been presented with schwa during learning (e.g., /səkɔf/).

Based on the assumption that words with a variable pronunciation will be harder to recognise than words with a consistent pronunciation, we predicted that listeners would take longer to recognise the target word that had been presented in a variable manner compared to the target word that had consistently been presented without schwa during the learning phase. We also compared the condition in which variable words were presented with schwa with the condition in which variable words were presented without schwa because this comparison allowed us to determine the validity of our experimental paradigm. We predicted that variable words that were presented with schwa would be recognised more slowly compared to variable words presented without schwa. This finding would replicate the phonological competition effect between words with the same acoustic onset that has been demonstrated with real words (e.g., Allopenna, Magnuson, & Tanenhaus, 1998; McQueen & Viebahn, 2007) as well as in other novel-word learning studies (e.g., Creel, Aslin, & Tanenhaus, 2008; Magnuson et al., 2003; Sulpizio & McQueen, 2012). In the condition in which the variable word was presented with schwa, the target word had the same initial syllable as the competitor (e.g., /səkɔb/ vs. /səkɔf/), whereas in the condition in which the variable word was presented without schwa, the target word shared only the initial consonant with the competitor (e.g., /skɔb/ vs. /səkɔf/). The previous studies suggest that the amount of word-initial acoustic overlap has a strong influence on lexical activation. Therefore, we ought to find more target-competitor competition for words presented with than without schwa.

In the filler condition—words that had been presented with schwa during learning—half of the time the competitor was a without-schwa word (e.g., /skɔf/) and half of the time the competitor was a word that had been presented in a variable manner during the learning phase (e.g., as /skɔf/ and /səkɔf/).

**Materials.** For the auditory stimuli, a total of 20 minimal pairs of French pseudowords were created that contained

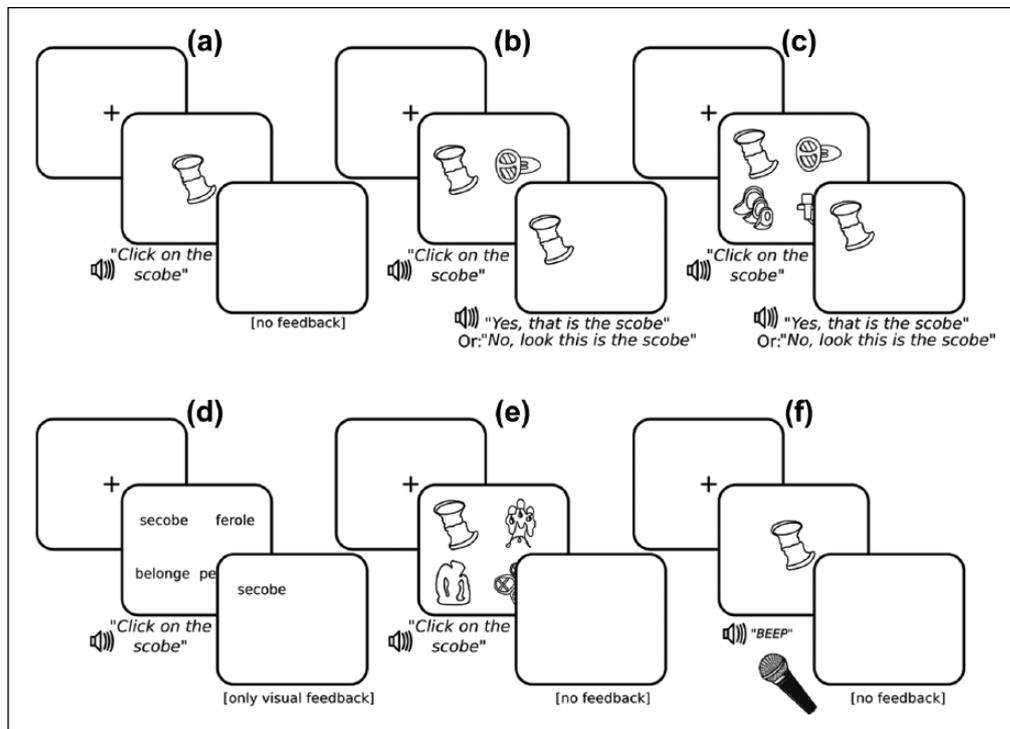
the vowel schwa after the initial consonant (e.g., /səkɔb/—/səkɔf/). These pseudowords were constructed such that the word-initial consonant-vowel-consonant sequences occurs in French words and so that if the schwa after the first consonant is removed, the resulting onset cluster is legal in French. We created 10 pairs of bisyllabic and 10 pairs of trisyllabic novel words. Members of a minimal pair differed only in a single consonant. For the bisyllabic items, this was the last phoneme (e.g., /səkɔb/ vs. /səkɔf/) whereas for the trisyllabic words, this was the penultimate phoneme (e.g. /bəlaʒɛ/ vs. /bəlafɛ/). The complete set of novel words is shown in Supplementary Appendix A.

The stimuli were recorded by a female native speaker of Swiss French. For each novel word, two different versions were recorded—one with and one without schwa. Each of these versions was recorded twice. The first recording of a given version was presented during the learning phase of the experiment, whereas the second recording was presented during the eye-tracking task. The reason for using two different recordings was so that listeners could not recognise the novel words during the eye-tracking task based on the idiosyncratic acoustic properties of a particular recording. Recordings were made with a sampling frequency of 22,050 Hz and scaled to 70 dB. During the recording procedure, the items were read from a pseudo-randomised list such that two items from the same minimal pair did not follow one another. Each novel word was recorded following the French carrier phrase *Cliquez sur le ...* “Click on the ...”.

For the visual stimuli, 40 pictures were taken from the MPI database of non-existing objects. Each picture was assigned to a specific novel word. The complete set of pictures is shown in the Supplementary Appendix.

The novel words were rotated through every experimental and filler condition such that each word occurred in each condition across participants but also such that no word appeared, for any given participant, in more than one condition. The only exception to this is that the words in the variable condition were presented to each participant during the eye-tracking task once with schwa and once without schwa (in order to keep the number of items in the experimental conditions as high as possible). For a given participant, 10 items occurred in the without-schwa condition, ten in the variable condition, and twenty items in the filler (i.e., with-schwa) condition.

**Procedure.** The procedures of the different experiments are illustrated in Figure 1. Before the beginning of the learning blocks, a familiarisation task was performed. During this task, participants saw the picture of a randomly selected object in the centre of the screen and were instructed (in French) to click on it with mention of the object’s name (e.g., *Click on the scobe*). During the learning blocks, there were either two or four pictures on the screen. After a preview of one second, participants were instructed to click



**Figure 1.** Experimental procedures in Experiment 1 (a, b, c, e, f) and Experiment 2 (a through f). See Table 2 for the order of the tasks in each Experiment: (a) Familiarisation, (b) learning with 2 objects, (c) learning with 4 objects, (d) orthographic exposure, (e) eye-tracking task, and (f) naming task.

on one of the objects (e.g., *Click on the scoobe*). After clicking on an object, only the correct one remained on the screen. Two-hundred milliseconds after the incorrect object(s) had disappeared, participants received auditory feedback telling them whether or not their choice was correct and repeating the name of the target object (e.g., *Yes, that's the scoobe* or *No, look this is the scoobe*). Five-hundred milliseconds after the end of the feedback sentence, the next trial was initiated. During the learning blocks, targets were never presented with the other member of the same minimal pair. The trial structure during the eye-tracking task was identical to the learning trials with four objects, with the exception that participants did not receive any feedback (neither auditory nor visual). After the eye-tracking task, participants performed the picture naming task. During this task, participants were presented with the pictures of the novel objects in random order. On each trial, participants had to recall the name of the object and say it out loud into a microphone. On each trial, participants first saw a fixation cross for 800 ms. Then the picture of an object appeared in the centre of the screen while a short (100 ms) beep sound was played. This beep sound was later used to determine the naming latencies by manually measuring the time between the onset of the beep and the onset of the participant's vocal response based on the information provided by a spectrogram and an oscillogram. Participants had 4 s in order to produce the name of the object. Then a blank screen was shown for 1 s after

which the next trial began. The different experimental tasks described above were presented in different blocks across three experimental sessions which were carried out on three consecutive days. Table 2 shows the structure and content of the different blocks and the order of tasks on each day.

**Apparatus.** For the learning phase, the experiment was controlled by PsychoPy, version 1.81.03 running on a Lenovo laptop with Ubuntu GNU/Linux 14.04. The objects were displayed on a 22-inch screen with a resolution of 1680 by 1050 pixels. The auditory stimuli were presented via headphones at a comfortable volume. The eye-tracking task was controlled by E-Prime running on a Dell laptop with Windows 7. Visual stimuli were again shown on a 22-inch display with a resolution of 1680 by 1050 pixels. Eye-movements were recorded with a SMI RED eye-tracker by SensoMotoric Instruments GmbH, Telto, Germany at a sampling rate of 250 Hz. For the naming task, the same equipment was used as in the learning phase. In addition, participants' vocal responses were recorded using a USB microphone (Snowball by Blue Microphones) in combination with the audio software Audacity.

## Results

All of the statistical analyses were conducted in the following way. Accuracy scores were analysed with generalised

**Table 2.** Block structure of Experiments 1 and 2. There was a break between each block. There was also a break after 100 trials in Blocks 1 and 7.

	Block	Experiment 1		Experiment 2	
		Trials	Task	Trials	Task
Day 1	0	40	Familiarisation (1 object)	40	Familiarisation (1 object)
	1	200	Learning with 2 objects	200	Learning with 2 objects
	2	120	Learning with 4 objects	120	Learning with 4 objects
	3	40	Learning with 4 objects	40	Orthographic exposure
	4	80	Learning with 4 objects	80	Learning with 4 objects
Day 2	5	40	Learning with 4 objects	40	Orthographic exposure
	6	80	Learning with 2 objects	80	Learning with 2 objects
	7	200	Learning with 4 objects	200	Learning with 4 objects
Day 3	8	40	Learning with 4 objects	40	Orthographic exposure
	9	40	Learning with 2 objects	40	Learning with 2 objects
	10	80	Learning with 4 objects	80	Learning with 4 objects
	11	40	Learning with 4 objects	40	Orthographic exposure
	12	60	Eye-tracking task	40	Eye-tracking task
	13	40	Naming task	40	Naming task

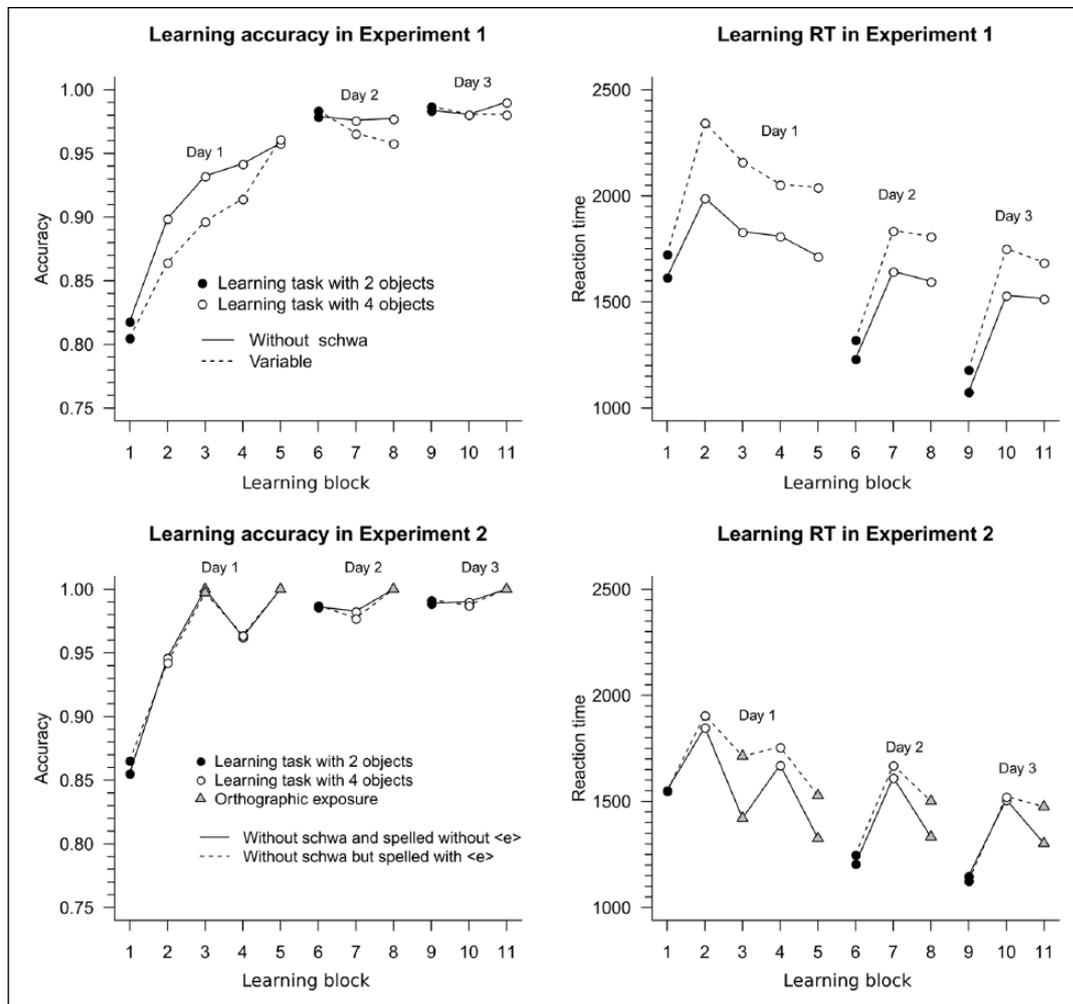
linear-mixed effects models with a binomial link function. Reaction times (RTs) were log-transformed and fitted with linear mixed-effects models. Standardised residuals larger than 2.5 were regarded as extreme values and removed during the modelling procedure. In order to analyse the time course of the target gaze probabilities, we used growth-curve analysis with linear, quadratic, and cubic polynomials in order to model the S-shaped rise in gaze probabilities over time, which is typically observed in visual-world eye-tracking experiments (Mirman, 2014; Mirman, Dixon, & Magnuson, 2008). We estimated  $p$  values by means of the Gaussian distribution. The models were fitted with the maximal random-effects structure unless the model-fitting procedure did not succeed due to convergence errors (Barr, Levy, Scheepers, & Tily, 2013). In cases in which the model-fitting algorithm failed to converge, we removed random effects in a step-wise fashion starting with the random slopes with the smallest standard deviation until convergence was successful.

**Learning phase.** Mean accuracy scores and RTs for the learning phase are shown in Figure 2. Overall, the accuracy results show that participants learnt to select the correct novel objects very quickly. Accuracy was already at ceiling in the second learning session. Generalised linear models showed that accuracy improved over the course of the three learning sessions ( $\beta_{\text{session}}=1.90$ ,  $z=12.28$ ,  $p<.001$ ) and that variable words were responded to less accurately compared to without-schwa words ( $\beta_{\text{variable}}=-0.23$ ,  $z=-2.51$ ,  $p=.01$ ). There was no interaction between condition and day ( $\beta_{\text{session*condition}}=-0.03$ ,  $z=-0.22$ ,  $p=.83$ ). RTs improved continuously over the three learning sessions ( $\beta_{\text{session}}=-0.12$ ,  $t=-10.24$ ,  $p<.001$ ) and participants responded more slowly to novel words

presented in a variable manner compared to novel words consistently produced without schwa ( $\beta_{\text{variable}}=0.11$ ,  $t=6.39$ ,  $p<.001$ ). There was no interaction between condition and day ( $\beta_{\text{session*condition}}=-0.01$ ,  $t=-0.83$ ,  $p=.41$ ).

**Naming task.** For the analysis of the naming results, each vocal response was checked for accuracy. Productions of non-target words, no responses, and mispronunciations were considered as errors and removed from the dataset. Responses were categorised as correct if they corresponded either to the without-schwa or the with-schwa variant of the target word. The classification as with-schwa or without-schwa production was performed by a trained phonetician whose native language is French, on the basis of auditory information. When a token could not be classified clearly as a with-schwa or a without-schwa production, it was marked as unclear and removed from the analyses. Naming latencies were computed using the Praat software package by manually measuring the time from the onset of the beep that occurred when the object appeared on the screen until the onset of the name of the object produced by the participant. Figure 3 summarises the average values for each of the dependent measures.

In the following analyses, we will focus on the comparison between the novel words that had been presented without schwa or in a variable manner during the learning phase. For naming accuracy, there was no significant difference between the without-schwa and variable conditions ( $\beta_{\text{variable}}=0.39$ ,  $z=1.64$ ,  $p=.10$ ). For the naming latencies, we first examined whether latencies differed between variable words that were produced with or without schwa during the naming task. Since there was no significant difference ( $\beta_{\text{variable with schwa}}=-0.02$ ,  $t=-0.36$ ,  $p=.72$ ), we collapsed across both types of responses. An additional



**Figure 2.** Mean accuracy and reaction times (RT in ms) during the learning phase of Experiments 1 and 2 for each learning block and day.

analysis showed longer latencies for variable compared to without-schwa words ( $\beta_{\text{variable}} = 0.11$ ,  $t = 2.51$ ,  $p = .01$ ).

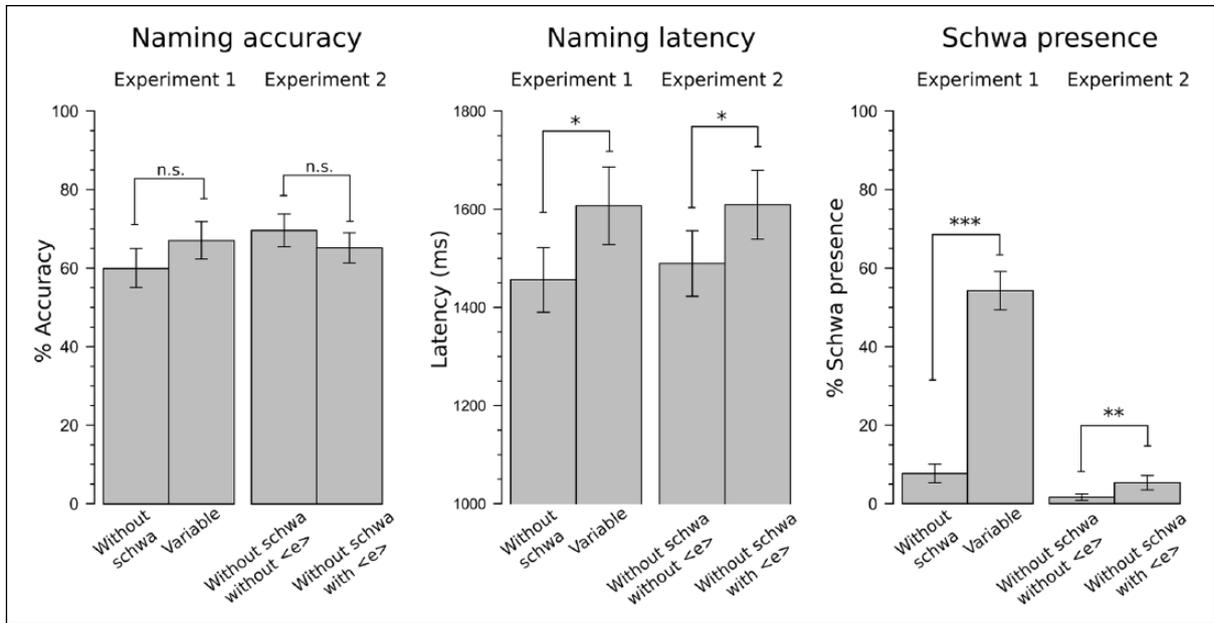
An analysis of the responses indicated a substantial increase in schwa productions for variable words compared to without-schwa words ( $\beta_{\text{variable}} = 55.11$ ,  $z = 4.93$ ,  $p < .001$ ). *T*-tests for participants and items indicated that for the variable condition, the proportion of schwa productions ( $M = 54\%$ , standard deviation [ $SD$ ] = 27%) was not statistically different from the predicted proportion of 50%, which corresponds to the number of times participants were presented with each variant during the learning phase of the experiment,  $t_1(30) = 0.87$ ,  $p = .27$ ;  $t_2(39) = 1.35$ ,  $p = .16$ .

**Eye-tracking task.** The main goal of the eye-tracking task was to examine the influence of phonological variation on the recognition of reduced word forms. For this purpose, we analysed accuracy, RT, and eye movements.

**Accuracy and RT.** Click responses and mean RTs for the different conditions are summarised in Table 3. For the

analysis of RTs, only correct trials were included. When comparing the without-schwa with the variable condition, we analysed only responses to variable targets that were produced without schwa in order to keep the amount of acoustic overlap between target and competitor constant across the two conditions. The results showed no difference in accuracy ( $\beta_{\text{variable}} = -0.22$ ,  $z = -0.30$ ,  $p = .77$ ) but slower responses to variable compared to without-schwa targets ( $\beta_{\text{variable}} = 0.06$ ,  $t = 3.51$ ,  $p < .001$ ). Further analyses showed that variable targets were recognised less accurately ( $\beta_{\text{schwa present}} = -1.49$ ,  $z = -2.38$ ,  $p = .02$ ) and more slowly ( $\beta_{\text{schwa present}} = 0.10$ ,  $t = 3.9$ ,  $p < .001$ ) when they were produced with schwa (and hence when the competitor overlapped more strongly) than when they were produced without schwa.

**Gaze probability.** In order to analyse the eye movements that participants made during the eye-tracking task, gaze probabilities for 50-ms intervals were calculated. For each time bin, gaze probability was calculated by dividing the number of trials during which participants were looking at



**Figure 3.** Results of the naming task in Experiments 1 and 2. Error bars indicate +/- 1 standard error of the mean. (n.s. = not significant, \* indicates that  $p < .05$  and \*\*\* indicates that  $p < .001$ .)

**Table 3.** Mean RTs (in ms) and click responses (in percentages) in the eye-tracking task of Experiment 1. For RTs, only correct trials were included. Standard deviations are shown in parentheses.

Condition	RT	Accuracy (target clicks)	Competitor clicks	Sum of distractor clicks
Without schwa	2,366 (324)	94.84 (6.52)	4.84 (6.26)	0.32 (1.25)
Variable, presented without schwa	2,502 (429)	95.16 (7.24)	4.84 (7.24)	0.00 (0.00)
Variable, presented with schwa	2,751 (441)	85.48 (14.34)	14.19 (14.09)	0.32 (1.80)

RT: reaction time.

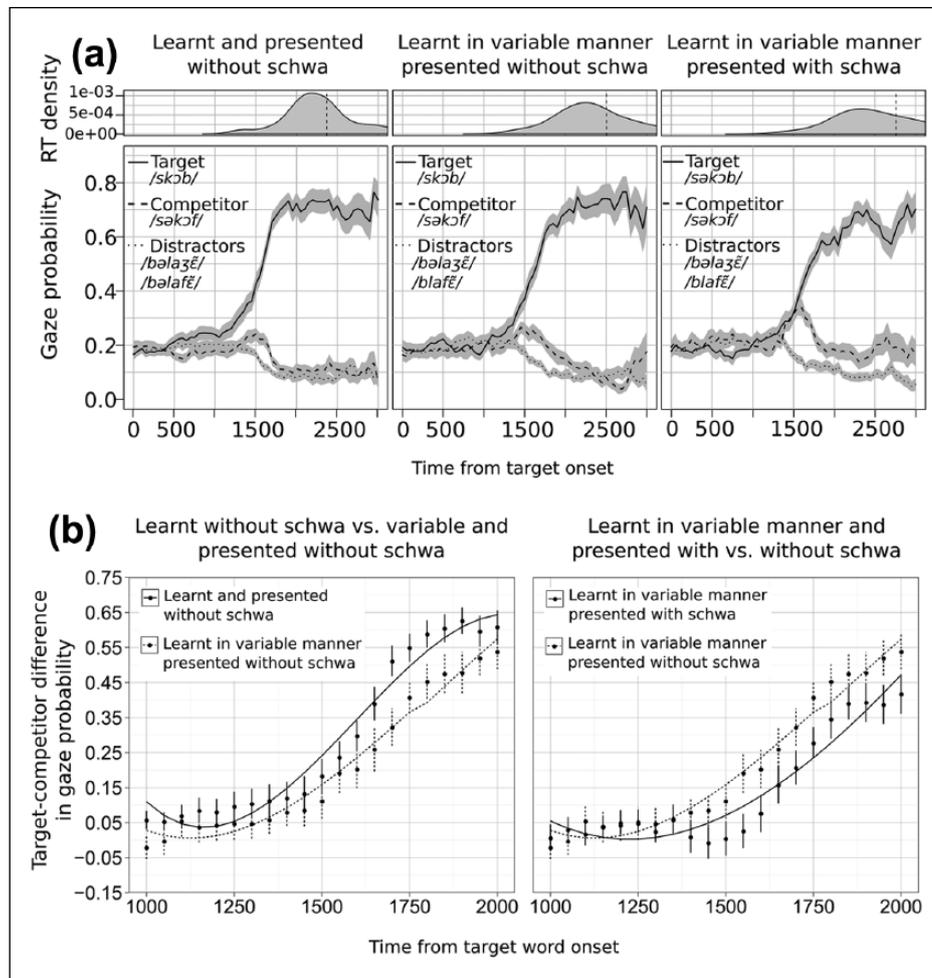
a particular object during this period of time by the total number of trials during which participants had not yet made a mouse click at that time. Fixations were categorised as being directed towards a particular object if they fell into a square of 400 by 400 pixels (approximately 10 by 10 cm) around the centre of the object’s position. The pictures of the objects were approximately 245 by 245 pixels in size (approximately 6.5 by 6.5 cm). Gaze probabilities and RT distributions for each of the five conditions in the eye-tracking task are shown in Figure 4a.

In order to measure the amount of activation of the target word relative to the competitor word, we computed the difference between target and competitor gaze probability. This measure captures fixations to the target and the competitor and allows us to conduct a single analysis for both measures. Growth-curve model plots are shown in Figure 4b, and model parameters are shown in Supplementary Appendix B. The target-competitor difference in gaze probabilities was smaller for variable target words (presented without schwa) than for without-schwa words ( $\beta_{variable} = -0.08, t = -2.74, p = .006$ ), demonstrating that words with a variable pronunciation are recognised less

efficiently compared to those pronounced consistently without schwa. Furthermore, we observed smaller target-competitor differences for variable targets that were produced with schwa compared to those produced without schwa ( $\beta_{without\ schwa} = 0.07, t = 2.70, p < .007$ ). This latter result is in line with previous studies that show that (particularly word-initial) acoustic overlap with competitor words makes word recognition more difficult (e.g., Allopenna et al., 1998), and thus confirms the validity of the visual-world paradigm as used here.

### Discussion

The results of Experiment 1 show that the consistency of the pronunciation with which novel words are presented to listeners influences how these words are learnt and processed. During the learning phase of the experiment, words with a variable pronunciation were identified more slowly than words with a consistent pronunciation. This finding is in line with the results of the eye-tracking task which showed that listeners recognised words that were sometimes produced with and sometimes produced without



**Figure 4.** Eye-tracking results in Experiment 1: (a) time course of gaze probability and RT distributions. The grey areas around the black lines indicate  $\pm 1$  standard errors around the participant means. The dashed vertical line on the RT distributions shows the mean RT for that condition and (b) growth-curve model plots (lines) and data (points with error bars) for target-competitor differences in gaze probability.

schwa more slowly than words that were always produced without schwa.

Both the RTs and the gaze probabilities show that participants made their responses relatively late compared to previous studies that used similar methods and stimuli (e.g., Creel et al., 2008; Magnuson, Tanenhaus, & Aslin, 2008; Sulpizio & McQueen, 2012). The lateness of the responses is likely to be due to conservative response behaviour that arose because task difficulty increased after the learning phase. During the learning phase, participants were never presented with members of the same minimal pair, whereas during the test phase participants were presented with both members of a minimal pair at once. This may have made participants respond more conservatively. Previous work suggests that the amount of effort that listeners put into a word-learning task depends on the difficulty of the task. The amount of information that listeners encode about a word may be limited to how relevant that information is for the task they have to perform. For

example, Italian participants have been shown to learn to distinguish between words that differ only in lexical stress, but do so only if members of the same minimal stress pair (e.g., /'bi.nu.lo/ vs. /bi.'nu.lo/) are presented as response alternatives during the same trial (Sulpizio & McQueen, 2012), not if they are presented on separate trials (Sulpizio & McQueen, 2011). This demonstrates that word learners do not automatically pay attention to acoustic details that are not relevant for the task. Similarly, in our experiment participants may have put relatively little effort into the learning task because the words were easy to distinguish. In the eye-tracking task, this changed abruptly which may have encouraged participants to become more cautious.

The most important result, however, was provided by the naming task. Participants were significantly more likely to produce a word with a schwa if it had been produced with a variable pronunciation than if it had always been produced without schwa during the learning phase. Furthermore, it took participants more time to produce the

name of an object with a variable pronunciation compared to a without-schwa object. The fact that phonological variation influenced performance in the eye-tracking as well as the naming task indicates that both word recognition and word production are influenced by exposure to variation in the phonological input.

## Experiment 2

In Experiment 2, we investigated how orthographic information influences the way in which listeners produce and recognise novel words so as to compare this influence with the influence of phonological input. This will allow us to compare the effect that phonological variation has on the processing of reduced word forms (Experiment 1) with the influence that spelling has. In Experiment 2, the novel words were always consistently produced with or without schwa. An orthographic exposure phase was added in which participants were presented with the orthographic forms of four novel words on each trial and asked to click on one of the words. Crucially, the spelling of the without-schwa words (e.g., /skɔb/) either contained the letter <e> and thereby indicated the presence of a schwa (as in <secobe>) or it did not contain the letter <e> (as in <scobe>).

According to the results obtained by Bürki et al. (2012), participants should be more likely to produce without-schwa words with a schwa if the word was spelled with <e> than if it was spelled without <e>. If phonological and orthographic information have the same impact on processing, we expect to find that the proportion with which listeners were exposed to each type of information should influence how words are produced to the same extent. For example, in Experiment 1, we observed that words that were presented 50% of the time with schwa during learning were later produced with schwa with the same proportion. If spelling has the same impact, we expect that the proportion of schwa productions in Experiment 2 should also correspond to the proportion with which participants were exposed to orthographic forms that contained the letter <e>. In contrast, if spelling is ignored and participants rely only on phonological information, spelling should have no effect on later production performance and words that were heard without schwa during learning should never be produced with schwa even if they had been spelled with the letter <e>. With respect to recognition performance, we would also expect a similar result as in Experiment 1 if spelling-sound inconsistency is processed in a similar way to phonological variation. Based on the assumption that words with a variable form are harder to recognise than words with a consistent form, we predict that listeners will take longer to recognise the target word when the phonological form and the orthography provide inconsistent information about schwa presence compared to when both forms provide consistent information about schwa.

The number of orthographic exposure trials that could be presented in Experiment 2 was constrained by two factors. First, in order to have comparable learning performance with respect to the associations between the pictures and the auditory stimuli, the number of object-learning trials had to be kept as high as possible (in order to approximate the number of object-learning trials in Experiment 1). Second, due also to practical constraints on the length of a multi-day experiment, the length of the experiment had to be the same as that of Experiment 1. Due to these two constraints, the proportion of trials with inconsistent orthographic forms in Experiment 2 was smaller than the proportion of trials with variable phonological forms in Experiment 1. However, because our primary research question is about the relative difference between input proportions and production proportions (i.e., within each experiment), the input proportions did not have to be identical across both experiments.

Having an orthographic exposure probability of below 50% also allowed us to examine a third possibility regarding the influence of spelling on processing. If participants fully rely on spelling, they might use information about spelling in a categorical way and take the spelling as evidence that both variants are equally permissible. If this is the case, participants should produce each variant 50% even though the actual exposure proportion was lower.

## Method

**Participants.** Participants were 36 students at the University of Geneva that had not taken part in Experiment 1. They were either reimbursed by course credit or a combination of money and course credit. The mean age was 23 years. Seven of the participants were male. All were native speakers of French, none had known hearing problems, and all had normal or corrected-to-normal visual acuity.

**Design and materials.** We used the same auditory stimuli and pictures of non-existing objects as in Experiment 1. In contrast to Experiment 1, however, we included an orthographic exposure phase. During this phase, participants were presented with the orthographic forms of four novel words on each trial and were instructed to click on one of the words (e.g., *Click on the scobe*, see panel D of Figure 1). The design of the learning phase and the naming task consists of the conditions produced without schwa and spelled without <e> and produced without schwa but spelled with <e>. As in Experiment 1, we added a filler condition in which words were produced with schwa and spelled with <e> in order to have an equal number of trials in which the novel word contained a schwa in the first syllable (see Table 1). Note that the condition in which without-schwa novel words were spelled without <e> is comparable to the without-schwa

condition in Experiment 1 and the condition in which a without-schwa novel word was spelled with <e> is equivalent to the variable condition.

The words from the condition produced without schwa but spelled with <e> were heard (without-schwa) 85% of the time and were seen (spelled with the letter <e>) 15% of the time. These exposure proportions were calculated as follows: Participants heard each novel word once in familiarisation, once in the eye-tracking task, once in each of the four orthographic blocks, and twice in each of the seven learning blocks (once in the instruction, once in the feedback), adding up to a total of 46 auditory presentations (see Table 3 and Figure 1). In addition, participants saw each word twice in each of the four orthographic blocks (once in the instruction and once in the feedback), adding up to a total of 8 visual presentations. Thus, the total number of exposures is 54. Consequently, the proportion of times that participants were exposed to orthographic forms indicating the presence of schwa was 15% and the proportion of auditory forms suggesting the absence of schwa was 85%. If participants follow the word-form exposure proportions (as they did in Experiment 1), they ought to produce the without-schwa words that had been spelled with <e> 85% of the time without schwa and 15% of the time with schwa. In contrast, if spelling is fully relied on (i.e. if listeners take even minimal exposure to the spelling to indicate that both forms are permissible), they ought to produce each variant 50% of the time. Finally, if spelling is ignored completely, participants ought to produce words from this condition only without schwa.

In the eye-tracking task, the without-schwa words which were spelled without <e> and the without-schwa words that had been spelled with <e> were each presented with a competitor word which had been learnt with schwa (e.g., /səkɔf/ spelled as <secophe>). The words in the filler condition were presented half of the time with a competitor that had been learnt without schwa and spelled without <e> (e.g., /skɔf/ spelled as <scophe>) and half of the time with a competitor that had been learnt without schwa word and spelled with <e> (e.g., /skɔf/ spelled as <secophe>).

As in Experiment 1, the novel words in all three tasks (learning, naming, and eye-tracking) were rotated through every experimental condition such that each word occurred in each condition, but each participant would encounter a given novel word only in a single condition in a task. Ten items were pronounced without-schwa and spelled without the letter <e>, 10 were pronounced without schwa but spelled with <e>, and 20 items were pronounced with schwa and spelled with <e>.

**Procedure.** The different experimental tasks are identical to those used in Experiment 1. In addition to the tasks from Experiment 1, an orthographic exposure phase was added. The orthographic exposure phase was similar to the learning task with four objects and differed in only two aspects.

First, participants saw the objects' orthographic forms instead of pictures of the objects. Second, there was only visual feedback (i.e., the correct orthographic form remained on the screen while the other forms disappeared), but there was no auditory feedback (see Figure 1). The duration of the print feedback was the same as the pictorial in the two learning tasks. The block structure is similar to Experiment 1 with the exception that the orthographic exposure blocks were added (see Table 2).

**Apparatus.** The apparatus used was the same as in Experiment 1.

## Results

We used the same analysis methods and procedures as in Experiment 1.

**Learning phase.** Mean accuracy scores and RTs for the learning blocks are shown in Figure 2. As in Experiment 1, participants quickly learnt to select the correct novel objects. As the nature of the orthographic task was quite different from the tasks in which participants selected pictures of novel objects, we analysed the results for these types of tasks separately.

In the picture-selection tasks, accuracy improved over the course of the three learning sessions ( $\beta_{\text{session}}=1.55$ ,  $z=17.72$ ,  $p<.001$ ) while there were no effects of spelling ( $\beta_{\text{spelling}}=0.04$ ,  $z=0.32$ ,  $p=.75$ ) and no interaction between spelling and session ( $\beta_{\text{session*spelling}}=-0.19$ ,  $z=-1.08$ ,  $p=.28$ ). In the orthographic task, there were no significant effects on accuracy whatsoever ( $\beta_{\text{session}}=21.37$ ,  $z=0.15$ ,  $p=.88$ ;  $\beta_{\text{spelling}}=-28.68$ ,  $z=-0.42$ ,  $p=.67$ ;  $\beta_{\text{session*spelling}}=18.49$ ,  $z=0.13$ ,  $p=.89$ ).

RTs also improved continuously over the three learning sessions in both the picture-selection ( $\beta_{\text{session}}=-0.09$ ,  $t=-8.12$ ,  $p<.001$ ) and the orthographic tasks ( $\beta_{\text{session}}=-0.04$ ,  $t=-3.25$ ,  $p=.001$ ). In the orthographic task, there was a significant effect of spelling ( $\beta_{\text{spelling}}=0.14$ ,  $t=9.27$ ,  $p<.001$ ), indicating that participants responded more slowly if a without-schwa word was spelled with the letter <e> compared to when it was spelled without it, which suggests that participants were sensitive to the mismatch between orthography and sound. In contrast, in the object-selection task, there was not effect of spelling ( $\beta_{\text{spelling}}=0.02$ ,  $t=1.23$ ,  $p=.22$ ). Furthermore, there was no spelling-by-session interaction in either the object-selection task ( $\beta_{\text{session*spelling}}=0.001$ ,  $t=0.01$ ,  $p=.11$ ) or the orthographic task ( $\beta_{\text{session*spelling}}=-0.02$ ,  $t=-1.65$ ,  $p=.10$ ).

**Naming task.** Vocal responses were coded for accuracy and schwa presence in the same manner and by the same person as in Experiment 1. Figure 3 summarises the average values for each of the dependent measures. With respect to naming accuracy, there was no significant effect of spelling

**Table 4.** Mean reaction times (RT) and click responses in the eye-tracking task of Experiment 2. RTs are given in milliseconds and click responses in percentages. Standard deviations of participant means are indicated in parentheses.

Condition	RT	Target clicks	Competitor clicks	Sum of distractor clicks
Without schwa and spelled without <e>	2,455 (337)	93.89 (9.94)	5.00 (8.45)	1.11 (3.19)
Without schwa but spelled with <e>	2,513 (371)	95.83 (8.74)	3.33 (6.32)	0.83 (3.68)

( $\beta_{\text{Spelling with } \langle e \rangle} = -0.39, z = -1.64, p = .10$ ). For naming latencies, we first investigated whether responses to items from the without-schwa/with-<e> condition differed depending on whether they were produced with or without schwa. This was not the case ( $\beta_{\text{Spelling with } \langle e \rangle, \text{ produced with schwa}} = 0.06, t = 0.49, p = .62$ ). We thus collapsed the data across both types of responses. Next, we compared latencies for without-schwa words that had been spelled with <e> to without-schwa words that had been spelled without <e>. We found that without-schwa words that had been spelled with <e> were named more slowly than without-schwa words without the letter <e> ( $\beta_{\text{Spelling with } \langle e \rangle} = 0.10, t = 2.62, p = .009$ ).

With regard to schwa presence, there is a small but significant increase in schwa productions for without-schwa words spelled with <e> compared to without-schwa words spelled without <e> ( $\beta_{\text{Spelling with } \langle e \rangle} = 1.28, z = 2.19, p = .03$ ). *T*-tests for participants and items indicated that the proportion of schwa productions ( $M = 5\%, SD = 11\%$ ) was significantly below 15%, the proportion based on the exposure during the learning phase,  $t_1(35) = -5.28, p < .001$ ;  $t_2(39) = -6.60, p < .001$ .

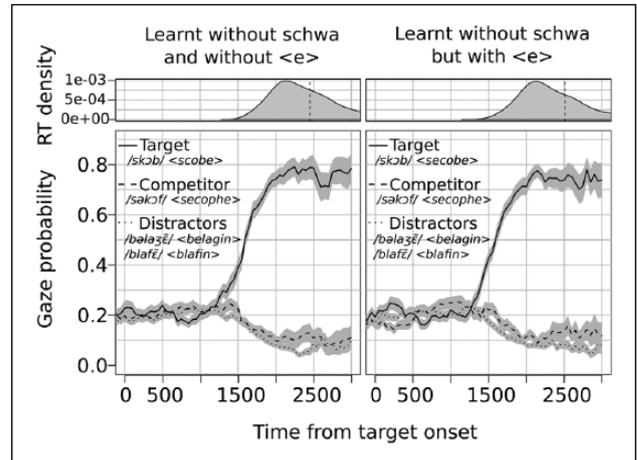
### Eye-tracking task

**Accuracy and RT.** The mean accuracy and RT values during the eye-tracking task are shown in Table 4. Models using spelling (with vs. without <e>) as the predictor variable indicated no effect of spelling for accuracy rates ( $\beta_{\text{Spelling with } \langle e \rangle} = 0.32, z = 0.54, p = .59$ ) or RTs ( $\beta_{\text{Spelling with } \langle e \rangle} = 0.02, t = 1.4, p = .16$ ).

**Gaze probability.** The time course of gaze probability during the eye-tracking task was analysed in the same way as in Experiment 1. See Figure 5 for the average gaze probability in each of the two conditions and Supplementary Appendix B for model parameters. A growth-curve model fitting the differences in gaze probability between the target and the competitor indicated no significant effect of spelling ( $\beta_{\text{spelling with } \langle e \rangle} = -0.02, t = -0.48, p = .63$ ). In other words, there was no difference between the condition in which the without-schwa target had been spelled without <e> and the condition in which it had been spelled with <e>.

### Discussion

The goal of Experiment 2 was to examine the influence that orthography has on the processing of novel reduced



**Figure 5.** Time course of gaze probability and RT distributions in the eye-tracking task of Experiment 2. The grey areas around the black lines indicate  $\pm 1$  standard errors around the participant means. The dashed vertical line on the RT distributions shows the mean RT for that condition.

word forms. We examined whether the presentation of orthographic forms that contain the letter <e> in a position in which it typically signifies the presence of the vowel schwa encourages participants to treat the new words as if they contain a schwa even if the phonological form did not contain one. The results of our naming task indicated that this is indeed the case. Participants were more likely to produce without-schwa novel words with a schwa if the word had been spelled with the letter <e> during the orthographic exposure phase than if it had been spelled without it. Although this effect is quite small (i.e., <5%) it is statistically significant and its size is comparable to the effect reported by Bürki et al. (2012). Furthermore, we found that without-schwa words that were spelled with <e> were produced more slowly than without-schwa words spelled without <e>. Both of these findings replicate Bürki et al.'s results and thus corroborate the claim that orthographic forms can influence the way in which reduced word forms are processed.

In contrast to Experiment 1, the proportion of reduced words produced with schwa was smaller than what we predicted based on the proportion of exposures. Based on the relative amount of exposure to orthographic forms spelled with <e>, we would have expected participants to produce at least 15% of the without-schwa words with schwa. However, the actual proportion of schwa productions was

significantly smaller. This finding suggests that if participants followed the orthographic input statistics in the same way as they did in Experiment 1, they did not rely on these statistics to guide their pronunciation. One possible reason for this is that orthography is not considered as a reliable cue for the phonological form of words. We will return to this issue in the General Discussion.

In addition to examining the influence of orthographic information on reduced word form production, we also investigated its effect on word recognition. Based on the recognition results in Experiment 1, we expected to find that without-schwa target words spelled with <e> would be recognised more slowly compared to without-schwa target words spelled without <e>. However, Experiment 2 suggests that the way in which the without-schwa words had been spelled had no influence on how quickly they were recognised. It is possible that effects of orthography were missed because participants responded very conservatively. As was mentioned earlier, participants responded much more slowly during the eye-tracking task than during the learning tasks. This is likely due to the fact that the objects participants had to choose from during the eye-tracking task were phonologically more similar compared to the objects that they had to distinguish between during learning. Nevertheless, the fact that the results of Experiment 2 differ from the results of Experiment 1 suggests that spelling-sound inconsistency influences word learning differently than phonological variation.

## Comparison of Experiments 1 and 2

In order to compare the results of Experiments 1 and 2 statistically, we conducted a combined analysis of the naming and eye-tracking data for both experiments. For the purpose of these analyses, a new variable was created that coded the consistency with which without-schwa words were presented during the learning phase in both experiments. If a word had a variable pronunciation in Experiment 1 or if it was pronounced without schwa but spelled with <e> in Experiment 2, the word was coded as *inconsistent*. In contrast, if a word was consistently produced without schwa in Experiment 1 or produced without schwa and spelled without <e> in Experiment 2, the word was coded as *consistent*. Furthermore, we added experiment as a predictor variable.

For schwa presence in naming, the analyses showed that inconsistent words were more likely to be produced with schwa than consistent words ( $\beta_{inconsistent}=1.28$ ,  $z=2.19$ ,  $p=.03$ ) and that the probability that a without-schwa word was produced with schwa was higher in Experiment 1 than in Experiment 2 ( $\beta_{Exp1}=1.77$ ,  $z=2.84$ ,  $p=.005$ ). Importantly, the model showed a significant interaction between Consistency and Experiment ( $\beta_{Consistency*Exp}=1.56$ ,  $z=2.32$ ,  $p=.02$ ) indicating that the increase in schwa productions from consistent to

inconsistent words was larger in Experiment 1 than in Experiment 2.

For naming accuracy, the analysis showed no effects of Consistency ( $\beta_{inconsistent}=-0.31$ ,  $z=-1.62$ ,  $p=.11$ ) or Experiment ( $\beta_{Exp1}=0.54$ ,  $z=1.56$ ,  $p=.12$ ). However, there was a significant interaction between Consistency and Experiment ( $\beta_{Consistency*Exp}=0.66$ ,  $z=2.45$ ,  $p=.02$ ). This interaction is likely due to the fact that in Experiment 1 there is a trend suggesting higher accuracy for inconsistent compared to consistent without-schwa words, whereas in Experiment 2 there is a trend in the opposite direction. Neither of these trends was statistically significant.

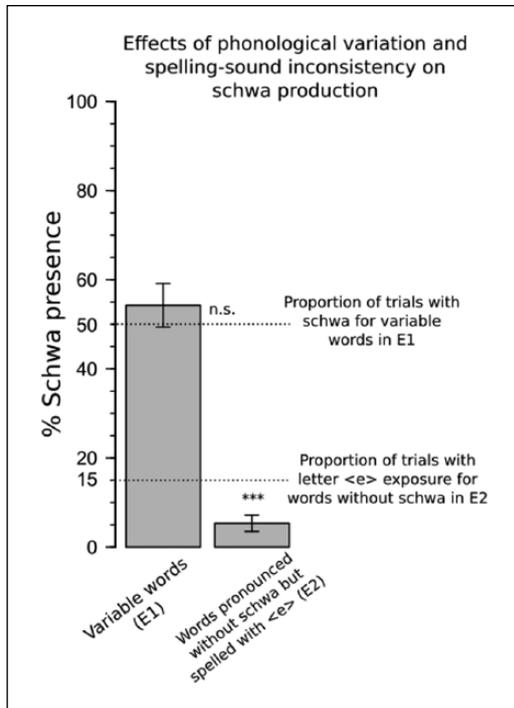
For the naming latencies, the statistical analysis showed longer latencies for inconsistent without-schwa words compared to consistent without-schwa words ( $\beta_{inconsistent}=0.11$ ,  $t=3.68$ ,  $p<.001$ ). There was neither an effect of Experiment ( $\beta_{Exp1}=-0.002$ ,  $t=-0.06$ ,  $p=.95$ ), nor a significant interaction between Experiment and Consistency ( $\beta_{Consistency*Exp}=0.03$ ,  $t=0.57$ ,  $p=.57$ ).

In order to compare recognition performance, we fitted a growth-curve model with target-competitor differences as the dependent variable and Consistency, Experiment, and the interaction between the two as the independent variables along with the three orthogonal polynomials. The random-effects structure consisted of random intercepts for participants. This model showed a significant interaction between Consistency and Experiment ( $\beta_{Consistency*Exp}=-0.06$ ,  $t=-3.85$ ,  $p<.001$ ), confirming the results reported for the individual experiments showing that only phonological variability influenced recognition performance.

The most important comparison concerns the predicted schwa production proportion based on the input statistics and the actual proportion of schwa productions in each experiment. As we reported previously, whereas the actual and the predicted proportions did not differ in Experiment 1, they did differ in Experiment 2. Figure 6 illustrates the difference in these effects for the two experiments. These analyses confirm that both types of information (phonological and orthographic) influence the production of without-schwa words. Crucially, however, the influence of spelling-sound inconsistency is smaller than the influence of phonological variation on the proportion of schwa productions.

## General discussion

The two experiments presented in this study show that phonological variation and spelling-sound inconsistency can both influence the processing of novel words. In line with previous work (Bürki et al., 2012), Experiment 2 showed that participants were more likely to produce a without-schwa novel word with a schwa (forms they had never heard before) if they had seen the word spelled with the letter <e> than if they had seen the word spelled



**Figure 6.** Comparison of the effects of phonological variation and spelling-sound inconsistency on schwa presence in the picture-naming task in Experiment 1 and 2. Effects are shown relative to the input proportions as indicated by the horizontal lines (E1 = Experiment 1, E2 = Experiment 2, n.s. = not significant, \*\*\* indicates that  $p < .001$ ).

without <e>. This finding is consistent with recent work that has shown that learning about words in the visual domain has consequences for on-line processing in the spoken domain (e.g., Bakker, Takashima, van Hell, Janzen, & McQueen, 2014). Experiment 1 showed that participants were more likely to produce a word with a schwa if it had been presented with a variable pronunciation than if it had been presented consistently without schwa. Crucially, the effect of phonological variation on naming responses was larger than the effect of spelling-sound inconsistency.

Because the absolute amount of exposure to inconsistent input (orthographic or phonological) differed across the two experiments, we compared participants' responses to the relative number of inconsistent exposures within each experiment. By examining the relative number of exposures, we are treating input variability as a continuous variable. This is different from previous studies in which it was treated as a binary variable (e.g., Bürki et al., 2012). In Experiment 1, participants' schwa productions mirrored the proportions of word-form presentations. In contrast, in Experiment 2, the proportion of schwa productions was below the proportion of schwa productions that we would have expected if participants had followed the proportion of exposures to words with inconsistent sound-spelling

patterns. Thus, whereas participants appear to have followed the input statistics of the different pronunciation variants in Experiment 1 quite closely, participants in Experiment 2 were less influenced by the input statistics of the orthographic forms. French learners of novel words seem to rely on phonological information more as a cue to the pronunciation of a novel word than they seem to rely on orthographic information. This suggests that phonological input is more likely to influence learning about the pronunciation variants of spoken words than exposure to orthographic forms.

In principle, it is possible that the difference between the effects of phonological variation and spelling-sound inconsistency is due to the difference between the input proportions. For example, there could be a threshold that has to be reached in order for participants to follow the input statistics and it is possible that the proportion of trials with inconsistent orthographic forms in Experiment 2 was below that threshold. Although this is logically possible, we think it is highly unlikely. As Experiment 2 showed, spelling did clearly have an effect on phonological processing. So, if there is a threshold that has to be reached for spelling to exert an effect, it was reached. It is implausible to assume another threshold that needs to be reached in order for listeners not only to use orthographic information but to use it proportionally to the input statistics. Future research could further explore this issue by comparing the effects of the same exposure proportions between the orthographic and phonological modalities.

In order to examine the influence of spelling-sound inconsistency and phonological variation on word recognition, we employed the visual-world paradigm in which participants' eye movements were monitored while they were selecting a target word (e.g. *scobe*) in the presence of a phonological competitor (e.g. *secophe*) and two distractors (e.g. *belagin* and *blafin*). Similar to the naming results, the eye-tracking results show a difference between the influence of phonological and orthographic input. Whereas the eye-tracking results of Experiment 1 showed an influence of phonological variation on word recognition, the results of Experiment 2 did not show an effect of spelling-sound inconsistency. This finding further supports the notion that orthographic information influences phonological processing to a lesser degree than acoustic information does.

These findings are highly relevant to theories about how listeners process reduced speech. Previous studies have suggested that orthography plays a significant role in the way pronunciation variants are recognised (e.g., Ranbom & Connine, 2007, 2011). In particular, it was proposed that the overlap between spelling and unreduced pronunciation variants could, at least in part, explain why unreduced variants are processed more efficiently than reduced variants even if the reduced variants occur more frequently. According to this view, orthography has a

strong effect that can override the effects of variant frequency. Our study suggests that, at least for pronunciation variants that are the result of schwa deletion in French, the influence of orthography is smaller than expected on this view.

This finding may seem surprising given that speech is inherently more variable than orthographic forms are. However, the crucial point may not be the invariance of orthographic forms but rather the degree of consistency of the mapping between graphemes and phonemes. The grapheme-phoneme mapping in French, like English, is relatively inconsistent (e.g. Ziegler et al., 1996). The grapheme that we focused on in the present study, the letter <e>, is no exception. Although this letter is associated with the vowel schwa in the position in which it was placed in the words in our study, there are many instances in French in which this letter does not correspond to the vowel schwa. For example, in Standard French, in the word *jet* “fountain” the letter <e> corresponds to the open-mid vowel /ɛ/ but in the word *et* “and” it is produced as the close-mid vowel /e/. Moreover, in many words such as the word *bracelet* “bracelet” the first letter <e> is silent. Because the letter <e> has many different phonological correspondences, it is not a very reliable cue for the presence of a vowel. Language users may therefore be reluctant to make strong inferences about the phoneme that the letter <e> corresponds to when learning new words. We believe that it is important to take into consideration the consistency of the grapheme-phoneme mappings when investigating the influence of orthographic information on phonological processing. We predict a larger effect of orthography when examining graphemes with highly consistent mappings (such as French <qu> which always signifies the presence of the phoneme /k/).

Another reason why the influence of phonological variation is larger compared to the influence of spelling might be that phonological forms provide more direct information about how a word sounds. In contrast, orthography requires the word learner to perform an additional step. The sound structure of a newly learnt word has to be inferred from the word’s spelling. It is possible that because of this additional step, it might take more time for orthographic information to develop an effect on phonological processing. Consistent with this notion is the recent word-learning study showing that novel words can influence lexical processing across modalities (Bakker et al., 2014). Words that were learnt in the auditory modality influenced lexical competition in the written modality, and vice versa. However, whereas the influence from sound to print emerged already one day after learning, the influence from print to sound emerged only a week later. While this result shows that there is an exchange of information between the written and the auditory modalities, it is also in line with our finding that speech input has an advantage over print.

Phonological information might also have a larger influence because listeners are typically exposed to auditory word forms before they are exposed to printed word forms. This was also true in our experiment. Because language learners, at least native speakers, are typically exposed to phonological word forms first, phonological information gets a head start which may make it harder for orthographic input to influence already established phonological representations. This is also true for reduced word forms, which has been shown by corpus studies examining infant-directed speech (e.g., Lahey & Ernestus, 2014). The situation is likely to be different for second-language learners who might first read novel words before they are exposed to (reduced) phonological forms of them. For second-language speakers, one might expect therefore a stronger influence of orthographic information on lexical representations compared to native speakers.

As phonological information was given a head start in the experiment (i.e., the new words appeared first in the spoken modality), it may seem not possible to conclude from these findings that speech carries more weight than spelling. However, it is important to note that the crucial manipulation was the source of the information indicating that the new word, initially presented without schwa, could also be produced with schwa. This information came only from speech in Experiment 1 and only from print in Experiment 2. The fact that the words were presented first in spoken form is not relevant for this comparison.

Our results speak to the question of what the nature of orthographic effects in spoken language processing is. Although several studies have demonstrated that orthographic knowledge can influence speech recognition, it is not yet clear what the locus of this influence might be. Some studies suggest that orthography might have a post-lexical influence on decision making that emerges only after listeners have already identified a given word based on the auditory input (e.g., Cutler & Davis, 2012). This hypothesis could explain why spelling effects have primarily been observed in meta-linguistic tasks such as lexical decision (e.g., Ziegler & Ferrand, 1998; Ziegler, Ferrand, & Montant, 2004). However, several other studies have demonstrated spelling effects in tasks that make the use of decisional strategies very unlikely, suggesting that learning to read can have an influence on on-line word recognition at a lexical (or even pre-lexical) locus (e.g., Pattamadilok, Morais, Colin, & Kolinsky, 2014; Perre, Bertrand, & Ziegler, 2011; Taft et al., 2008; Ziegler et al., 2004). Another but related question is whether orthography influences processing because the orthography is called up as the spoken stimulus is being heard or as the result of a longer-term learning and retuning process.

While our study cannot solve the debate on how orthographic information influences spoken word processing in general, it does show that the influence of orthographic information on the processing of reduced word forms is, at

least in French, relatively small compared to the influence of phonological information. This conclusion is consistent with a recent study showing that whether or not a speech sound is orthographically coded does not affect the processing costs that are typically associated with reduced speech (Mitterer & Reinisch, 2015; see also Racine et al., 2013).

As we employed both a production task (picture naming) and a recognition task (visual-world eye tracking), our results are also informative with respect to the relationship between the recognition and production systems. Experiment 1 showed that learning novel words with variable pronunciations influenced behaviour in the production task as well as in the eye-tracking task. This finding is inconsistent with a cognitive architecture that makes a strict division between language production and comprehension systems and treats them as informationally encapsulated modules. Instead, the learning of novel words appears to require that production and comprehension processes work closely together and exchange information (e.g., Mitterer & McQueen, 2009). Even the relatively passive type of learning procedure that we used in the present study seems to affect both modalities. Although our results do not allow us to decide whether production and comprehension use different (e.g., Levelt, Roelofs, & Meyer, 1999) or the same processes and representations (e.g., Pickering & Garrod, 2013), the similar findings across the production and comprehension tasks in Experiment 1 show that both systems work closely together.

In conclusion, this study provides further support for the notion that orthographic information can influence the phonological processing of schwa-deleted word forms. However, the influence of orthography is, at least in French, outweighed by the influence of phonological information. This finding has implications for the study of the acquisition of phonological knowledge in general and the processing of reduced pronunciation variants in particular. While orthographic information can influence the processing of reduced speech, this influence is rather small and may depend on the consistency of the grapheme-phoneme mapping.

### Acknowledgements

We would like to thank Dan Mirman for helpful advice regarding the application of growth curve analysis to our eye-tracking data.

### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

This work was supported by a Kurzstipendium stipend for PhD students of the German Academic Exchange Service (DAAD) to the first author. The third author was supported by her vici project awarded by the Netherlands Organisation for Scientific Research. This work was further supported by a grant from

the Swiss National Science Foundation (grant number 100014\_159374) to the last author.

### Supplementary material

Supplementary Appendix A and B is available at: journals.sagepub.com/doi/suppl/10.1177/1747021817741859. Stimulus materials and data files are available on the Open Science Framework via the following link: <https://osf.io/2edua/>.

### Note

1. In the following, we will use IPA symbols in slashes in order to describe the phonological forms of words and angle brackets when explicitly referring to the orthographic form.

### References

- Alloppenna, P., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38, 419–439. doi:10.1006/jmla.1997.2558
- Bakker, I., Takashima, A., van Hell, J. G., Janzen, G., & McQueen, J. M. (2014). Competition from unseen or unheard novel words: Lexical consolidation across modalities. *Journal of Memory and Language*, 73, 116–130. doi:10.1016/j.jml.2014.03.002
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278. doi:10.1016/j.jml.2012.11.001
- Browman, C., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49, 155–180.
- Bürki, A., Ernestus, M., & Frauenfelder, U. H. (2010). 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. *Journal of Memory and Language*, 62, 421–437. doi:10.1016/j.jml.2010.01.002
- Bürki, A., & Frauenfelder, U. H. (2012). Producing and recognizing words with two pronunciation variants: Evidence from novel schwa words. *Quarterly Journal of Experimental Psychology*, 65, 796–824. doi:10.1080/17470218.2011.634915
- Bürki, A., Spinelli, E., & Gaskell, M. G. (2012). A written word is worth a thousand spoken words: The influence of spelling on spoken-word production. *Journal of Memory and Language*, 67, 449–467. doi:10.1016/j.jml.2012.08.001
- Côté, M., & 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.
- Creel, S. C., Aslin, R. N., & Tanenhaus, M. K. (2008). Heeding the voice of experience: The role of talker variation in lexical access. *Cognition*, 106, 633–664. doi:10.1016/j.cognition.2007.03.013
- Cutler, A., & Davis, C. (2012). An orthographic effect in phoneme processing, and its limitations. *Language Sciences*, 3, 18. doi:10.3389/fpsyg.2012.00018
- Cutler, A., Treiman, R., & van Ooijen, B. (2010). Strategic deployment of orthographic knowledge in phoneme detection. *Language and Speech*, 53, 307–320. doi:10.1177/0023830910371445

- Ernestus, M. (2000). *Voice assimilation and segment reduction in casual Dutch: A corpus-based study of the phonology-phonetics interface*. Utrecht, The Netherlands: LOT. Retrieved from <http://repository.uibn.ru.nl/handle/2066/29974>
- Ernestus, M., & Smith, R. (in press). Qualitative and quantitative aspects of phonetic variation in Dutch eigenlijk. In F. Cangemi, M. Clayards, O. Niebuhr, B. Schuppler & M. Zellers (Eds.), *Rethinking reduction*. Berlin, Germany: De Gruyter Mouton.
- Ernestus, M., & Warner, N. (2011). An introduction to reduced pronunciation variants. *Journal of Phonetics*, *39*, 253–260. doi:10.1016/S0095-4470(11)00055-6
- Grainger, J., Muneaux, M., & Farioli Ziegler, J. C. (2005). Effects of phonological and orthographic neighbourhood density interact in visual word recognition. *The Quarterly Journal of Experimental Psychology*, *58*, 981–998. doi:10.1080/02724980443000386
- Hanique, I., Ernestus, M., & Schuppler, B. (2013). Informal speech processes can be categorical in nature, even if they affect many different words. *Journal of the Acoustical Society of America*, *133*, 1644–1655. doi:10.1121/1.4790352
- Johnson, K. (2004). Massive reduction in conversational American English. In K. Yoneyama & K. Maekawa (Eds.), *Spontaneous speech: Data and analysis*. Proceedings of the 1st Session of the 10th International Symposium. Tokyo, Japan: The National International Institute for Japanese Language (pp. 29–54).
- Lahey, M., & Ernestus, M. (2014). Pronunciation variation in infant-directed speech: Phonetic reduction of two highly frequent words. *Language Learning and Development*, *10*, 308–327.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–75.
- Luce, P. A., & McLennan, C. T. (2005). Spoken word recognition: The challenge of variation. In D. B. Pisoni & R. E. Remez (Eds.), *The handbook of speech perception* (pp. 591–609). Malden, MA: Blackwell.
- Magnuson, J. S., Tanenhaus, M., Aslin, R., & Dahan, D. (2003). The time course of spoken word learning and recognition: Studies with artificial lexicons. *Journal of Experimental Psychology: General*, *132*, 202–227. doi:10.1037/0096-3445.132.2.202
- Magnuson, J. S., Tanenhaus, M. K., & Aslin, R. N. (2008). Immediate effects of form-class constraints on spoken word recognition. *Cognition*, *108*, 866–873. doi:10.1016/j.cognition.2008.06.005
- McQueen, J. M., & Viebahn, M. C. (2007). Tracking recognition of spoken words by tracking looks to printed words. *Quarterly Journal of Experimental Psychology*, *60*, 661–671. doi:10.1080/17470210601183890
- Mirman, D. (2014). *Growth curve analysis and visualization using R*. Abingdon, UK: Taylor & Francis Group.
- Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*, *59*, 475–494. doi:10.1016/j.jml.2007.11.006
- Mitterer, H., & McQueen, J. M. (2009). Processing reduced word-forms in speech perception using probabilistic knowledge about speech production. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 244–263. doi:10.1037/a0012730
- Mitterer, H., & Reinisch, E. (2015). Letters don't matter: No effect of orthography on the perception of conversational speech. *Journal of Memory and Language*, *85*, 116–134. doi:10.1016/j.jml.2015.08.005
- Pattamadilok, C., Morais, J., Colin, C., & Kolinsky, R. (2014). Unattentive speech processing is influenced by orthographic knowledge: Evidence from mismatch negativity. *Brain and Language*, *137*, 103–111.
- Perre, L., Bertrand, D., & Ziegler, J. (2011). Literacy affects spoken language in a non-linguistic task: An ERP study. *Language Sciences*, *2*, 274. doi:10.3389/fpsyg.2011.00274
- Perre, L., Pattamadilok, C., Montant, M., & Ziegler, J. C. (2009). Orthographic effects in spoken language: On-line activation or phonological restructuring? *Brain Research*, *1275*, 73–80. doi:10.1016/j.brainres.2009.04.018
- Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*, 329–347. doi:10.1017/S0140525X12001495
- Racine, I., Bürki, A., & Spinelli, E. (2013). The implication of spelling and frequency in the recognition of phonological variants: Evidence from pre-readers and readers. *Language, Cognition and Neuroscience*, *29*, 893–898. doi:10.1080/01690965.2013.832784
- Ranbom, L. J., & Connine, C. M. (2007). Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language*, *57*, 273–298. doi:10.1016/j.jml.2007.04.001
- Ranbom, L. J., & Connine, C. M. (2011). Silent letters are activated in spoken word recognition. *Language and Cognitive Processes*, *26*, 236–261. doi:10.1080/01690965.2010.486578
- Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, *5*, 546–554. doi:10.1037/0278-7393.5.6.546
- Sulpizio, S., & McQueen, J. M. (2011). When two newly-acquired words are one: New words differing in stress alone are not automatically represented differently. In Proceedings of the 12th Annual Conference of the International Speech Communication Association (Interspeech 2011), Florence, Italy (pp. 1385–1388).
- Sulpizio, S., & McQueen, J. M. (2012). Italians use abstract knowledge about lexical stress during spoken-word recognition. *Journal of Memory and Language*, *66*, 177–193. doi:10.1016/j.jml.2011.08.001
- Taft, M., Castles, A., Davis, C., Lazendic, G., & Nguyen-Hoan, M. (2008). Automatic activation of orthography in spoken word recognition: Pseudohomograph priming. *Journal of Memory and Language*, *58*, 366–379. doi:10.1016/j.jml.2007.11.002
- Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review*, *5*, 683–689.
- Ziegler, J. C., Ferrand, L., & Montant, M. (2004). Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks. *Memory & Cognition*, *32*, 732–741.
- Ziegler, J. C., Jacobs, A. M., & Stone, G. O. (1996). Statistical analysis of the bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods, Instruments, & Computers*, *28*, 504–515. doi:10.3758/BF03200539