On the resolution of phonological constraints in spoken production: Acoustic and response time evidence

Audrey Buerki, Ulrich H. Frauenfelder, F-Xavier Alario

On the resolution of phonological constraints in spoken production: Acoustic and response time evidence

Audrey Bürki[a] **and Ulrich H. Frauenfelder**
Faculté de Psychologie et des Sciences de l'Education, Université de Genève, Bd du Pont d'Arve 42, 1205 Genève, Switzerland
audrey.buerki@unige.ch, ulrich.frauenfelder@unige.ch

F.-Xavier Alario
Laboratoire de Psychologie Cognitive, Aix-Marseille Université & CNRS, 3, Place Victor Hugo, 13331, Marseille, France
francois-xavier.alario@univ-amu.fr

Abstract: This study examines the production of words the pronunciation of which depends on the phonological context. Participants produced adjective-noun phrases starting with the French determiner *un*. The pronunciation of this determiner requires a liaison consonant before vowels. Naming latencies and determiner acoustic durations were shorter when the adjective and the noun both started with vowels or both with consonants, than when they had different onsets. These results suggest that the liaison process is not governed by the application of a *local* contextual phonological rule; they rather favor the hypothesis that pronunciation variants with and without the liaison consonant are stored in memory.

1. Introduction

A fundamental aim in linguistic and psycholinguistic research is to determine the nature of linguistic knowledge. Key questions concern for instance the nature of the linguistic units stored in speakers' memory (or, in psycholinguistic terms, the content of the mental lexicon, e.g., Hinskens *et al.*, 2014) or, relatedly, whether speakers have internalized sets of rules or whether they use probabilistic computations to combine linguistic units (e.g., sounds) into larger units (e.g., Chater and Manning, 2006). The present study contributes to this debate by examining whether highly constrained (systematic) contextual phonological variation can be accounted for in terms of local contextual rules or, alternatively, is better accounted for in terms of an activation-selection mechanism operating on multiple representations.

The study focuses on a specific case of external sandhi, namely, obligatory French liaison. In obligatory French liaison, a word-final consonant is pronounced when the word is followed by a vowel-initial word (e.g., *grand ami* /gʀɑ̃tami/ "great friend") but is not pronounced when this word is produced in isolation (/gʀɑ̃/ or followed by a consonant-initial word (*grand film* /gʀɑ̃film/ "great film"). Traditional accounts of the phenomenon assume that the liaison consonant is inserted or deleted following the application of a phonological rule (e.g., Tranel, 1981). Notably, in the phonological literature, phonological processes usually apply locally (e.g., Kenstowicz, 1994). Alternative accounts of the liaison process have been discussed as well; some of them assuming for instance that the pronunciation with and without the liaison consonant each have a corresponding underlying representation (e.g., Klausenburger 1984). The empirical data to date do not allow us to distinguish between such distinct proposals.

To address this issue, we examine French liaison as it occurs in determiners. In French, many determiners (e.g., *un* "a," *mon* "my," *les* "the") ending with a vowel are indeed realized with a liaison consonant when followed by a vowel-initial noun or adjective (e.g., *un* "a" realized /ɔ̃/ in *un chien* /ɔ̃ʃjɛ̃/ "a dog" and /ɔn/ in *un âne* /ɔ̃nan/ "a donkey"). To investigate the cognitive processes underlying the production of these determiners, we build on previous work with context-dependent determiners in psycholinguistic research. Spalek *et al.* (2010) asked English native participants to name

[a]Author to whom correspondence should be addressed.

pictures using the English indefinite and definite singular determiners (i.e., *a*/*the* before consonants; *an*/*thee* before vowels). They observed shorter production latencies when the adjective and noun both started with either a consonant or with a vowel as opposed to different onsets. Bürki *et al.* (2014) reported a similar phonological consistency effect with the French definite singular determiners *le* and *la* (realized *l'* before vowels, see Miozzo and Caramazza, 1999 for similar results on determiners the two pronunciations of which vary in more than one segment). Phonological consistency effects reveal an influence of the distant context (the noun) on the production of the determiner that is phonologically constrained by the local context (the adjective). Hence such effects reveal that the constraints on determiner form production are not resolved locally.

Most authors take phonological consistency effects to suggest that the two pronunciations of the determiner have a corresponding word form representation to which both the adjective and noun can send activation. The selection of the "to be produced" determiner form is quicker when the adjective and noun send activation to the same form as is the case in consistent utterances (Miozzo and Caramazza, 1999, see Spalek *et al.*, 2010, for an alternative account and Bürki *et al.*, 2014, for empirical evidence against the view of Spalek *et al.*). Note that in all previous experiments but one (Spalek *et al.*, 2010, experiment 2), the two forms of the determiner had different spellings. Spalek *et al.* (2010) suggested that this may influence the way these words are represented in the mental lexicon.

French liaison in determiners differs from the determiner form variations examined in previous studies in that it is both a general phenomenon in the language (many word classes also have words with a liaison consonant) and a highly systematic process (the pronunciation of determiners with a liaison consonant never fails to accommodate the phonological context). In addition, the two variants share a single spelling. These properties make liaison in determiners a very good candidate for being implemented as the application of a local contextual rule.

Here we examine whether the encoding of the determiner *un* is sensitive to the consistency between onsets in determiner + adjective + noun utterances. We examine the influence of the consistency between onsets on both the naming latencies and the acoustic properties of the determiner. The hypothesis that the liaison process is governed by the application of a local contextual rule that transforms the underlying representation into a different surface form predicts no difference in processing time and no difference in the phonetic realization of the determiner between consistent and inconsistent utterances. Our findings appear to go against this prediction.

2. Methods

2.1 Participants

Twenty-three native French speakers, with no hearing or language disorders, took part in the experiment. They were all students at the University of Geneva and aged between 18 and 35 yr. They were given course credit for their participation.

2.2 Materials

We selected two French adjectives frequently used in prenominal position. The first started with a consonant (*demi* "half"), the second started with a vowel (*immense* "huge"). We further selected 40 black and white pictures representing masculine French nouns. All pictures were framed in a black outline square. Half corresponded to consonant-initial nouns, the other half corresponded to vowel-initial nouns. Ten additional pictures representing feminine nouns were used as fillers. We made two versions of each picture. To elicit the production of the adjective *demi*, a semi-opaque white mask was applied to the right-hand half of the original black outline picture. To elicit the production of the adjective *immense*, we enlarged the pictures of the objects inside the black square such that they nearly touched the square.

There were two lists, and each object appeared in both lists, once to be used with the adjective *immense* and once to be used with the adjective *demi*. The pictures were distributed amongst the lists such that the two adjectives were to be used the same number of times in both.

Each participant was presented with the two lists consecutively, the order of the lists being counterbalanced across participants. Items were fully randomized within each list.

2.3 Procedure

The experiment took place in a quiet booth and was programmed using the software DMDX (Forster and Forster, 2003). The participants were first familiarized with the pictures and their corresponding names. They were asked to name the pictures in two tasks. In the noun phrase naming task, they had to use the indefinite determiner (*un*), the appropriate adjective (*demi* or *immense* depending on the picture) and noun. In the bare noun naming task, they had to produce the noun only. In the two tasks, trials started with a 800 ms fixation cross, followed by a 200 ms blank screen interval. The pictures then appeared on the screen and stayed until the participant's response or for a maximum of 4000 ms. The next trial started after a 1000 ms blank screen interval. The two tasks started with ten practice trials. All participants named the pictures with the determiner and adjective first. The whole experimental session lasted for about 45 min. The bare noun naming task was introduced to control for the variation in response times driven by specific properties of the nouns or pictures. We introduced the naming latencies from the bare noun naming task as a fixed predictor in the statistical model for the naming latencies of the noun phrase naming task.

2.4 Statistical analyses

The dependent variables were the inverse (1/x) of the naming latencies (as indicated by the Box-Cox test) in the noun phrase naming task and the production duration of the determiner relative to that of the whole noun phrase. These variables were entered in mixed-effects regression models, using the statistical software R and library lmerTest. Denominator degrees of freedom and p values were computed based on Satterthwaite's approximations. Phonological consistency and the onset of the adjective (vowel vs consonant) were entered as fixed effects. In the analysis of naming latencies, two covariates were also introduced, namely, the naming latency for the same picture produced with a bare noun (see preceding text) and the position of the trial in the experiment. The latter was introduced to account for the variability in naming latencies arising as a consequence of the same object being presented several times in the course of the experiment. We also tested the interaction between phonological consistency and the onset of the adjective. This was done to ensure that the consistency effect is present for the two forms of the determiner (i.e., with and without the liaison consonant, see Bürki *et al.*, 2014). Two covariates were also introduced for the analysis of the determiner relative duration, i.e., the position of the trial in the experiment and the naming latencies in the noun phrase naming task. All models had random intercepts for participants and items, and random slopes allowing for the effect of phonological consistency to differ across participants and items.

3. Results

3.1 Naming latencies

Naming accuracy and latencies (i.e., time interval between picture onset and articulation onset) were manually checked offline with the software CHECKVOCAL (Protopapas, 2007). Of the 1840 responses in the noun phrase naming task, there were 308 errors (17%, among which 72% dysfluencies, 18% mispronunciations/selection of wrong word, 10% missing responses).

For the statistical analysis, we further removed the 25 trials the pictures of which were not named correctly in the bare noun naming task (participants made 59, 3%, errors in this task). The visual inspection of the distribution of naming latencies further led to the removal of the 50 data points above 1400 ms. Finally, 1457 data points were included in the statistical analyses.

Mean naming latencies were 751 ms and 770, for phonologically consistent and inconsistent trials, respectively (standard error of the difference between means = 9.8). These data are illustrated in Fig. 1.

The statistical analysis revealed a main effect of phonological consistency on naming latencies ($\beta = 2.55 \times 10^{-05}$, $t = 2.17$, $p < 0.05$), these being shorter for consistent than for inconsistent trials. Naming latencies were also shorter for utterances with consonant-initial than vowel-initial adjectives ($\beta = 5.52 \times 10^{-05}$, $t = 4.84$, $p < 0.0001$), and they decreased with the position of the trial in the experiment ($\beta = 5.46 \times 10^{-07}$, $t = 2.78$, $p < 0.01$). Finally, there was a main effect of the naming latencies in the bare noun naming task. Naming latencies for noun phrases decreased when the naming latencies for the same pictures produced with a bare noun decreased ($\beta = 1.30 \times 10^{-07}$, $t = 4.60$, $p < 0.0001$). There was no interaction between phonological consistency and the onset of the adjective ($p = 0.26$).

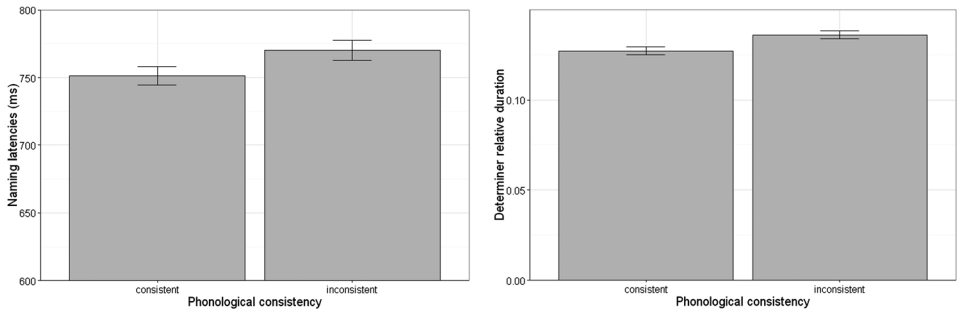

Fig. 1. Mean latencies (left panel) and mean determiner relative duration (i.e., duration of the determiner relative to that of the whole noun phrase, right panel) broken down by phonological consistency. Error bars correspond to the 95% confidence intervals around the means.

3.2 Acoustic analyses

The onset and offset of the determiner were set manually by an expert phonetician who inspected spectrogram and waveform displays using the software PRAAT (Boersma and Weenink, 2015). For utterances with the adjective *immense*, determiner duration was measured from the onset of voicing to the onset of the increase in energy in the third formant at the end of the liaison consonant. For utterances with the adjective *demi*, determiner duration was measured from the onset of voicing to the onset of the decrease in energy in the second formant. Examples of alignments are provided in Fig. 2. The mean duration of the determiner was, respectively, 104 and 113 ms for the consistent and inconsistent sequences (standard error of the difference between means = 2.18).

The dependent variable was the duration of the determiner relative to that of the whole noun phrase (but note that the effect of phonological consistency is also present when we use the effective duration as the dependent variable). Mean relative duration was, respectively, 0.127 and 0.136 for consistent and inconsistent sequences (standard error of the difference between means = 0.0024, see Fig. 1). There was no potentially harmful multicollinearity in this model (all tolerance values above 0.985). The statistical model revealed a main effect of consistency with longer relative durations for determiners in the inconsistent condition ($\beta = 3.41$, $t = 2.53$, $p < 0.05$), a main effect of the onset of the adjective with a longer relative duration for determiners before vowel-initial adjectives ($\beta = 62.8$, $t = 53.54$, $p < 0.0001$), and a main effect of the position of the trial in the experiment, with the duration decreasing with later positions ($\beta = -0.039$, $t = -1.98$, $p < 0.05$). There was no effect of the naming latencies in the noun phrase naming task ($\beta = 0.0037$, $t = 0.079$, $p > 0.9$).

4. Discussion

This study examined whether the production of words the phonological form of which is determined by the phonological context is governed by a local contextual rule. We tested whether the production of adjective-noun phrases starting with the French indefinite determiner *un*, the pronunciation of which involves a liaison consonant before vowels, is sensitive to the consistency in onsets between the adjective and the noun. The hypothesis that the production of an utterance with a liaison consonant results from the application of a local phonological rule predicted no difference in response times or in the acoustic properties of the response between consistent and inconsistent utterances. The pattern of results contradicts these predictions. Speakers needed more time to encode utterances with inconsistent onsets and produced determiners with longer durations in this condition. This reveals that the production of determiners with liaison is sensitive to non-local constraints. This result shows very clearly that the liaison process is not governed by the application of a *local* contextual phonological rule.

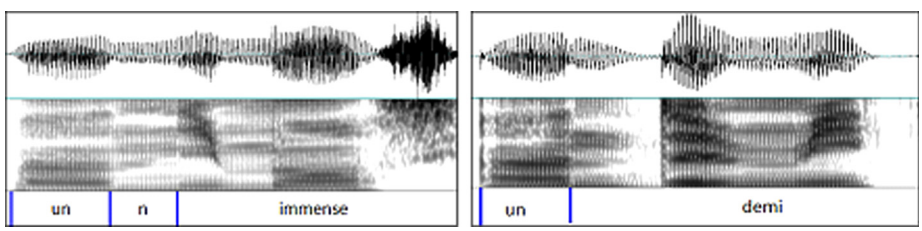

Fig. 2. (Color online) Phonetic alignment of determiner boundaries in a representative utterance with the adjective immense "huge" (left panel) and in a representative utterance with the adjective demi "half" (right panel).

These results may challenge the very idea that the liaison process results from a phonological rule or may only question the scope of this rule. It could be argued, for instance, that French liaison results from a non-local phonological rule. According to this account, the noun, given its frequent association with the determiner, would trigger the application of a rule (e.g., "insert a consonant"). A time-consuming additional rule would then have to be applied to block or reverse the first rule when the adjective does not require its application. Importantly, however, this account predicts that the variants with and without the liaison consonant would be affected differently by the phonological consistency between the adjective and the noun; only the represented variant should be sensitive to the consistency between onsets. The lack of interaction between the onset of the adjective and phonological consistency in our data does not support this view. Moreover, the acoustic difference we observe between phonologically consistent and phonologically inconsistent sequences is difficult to explain within a rule-based account.

As discussed in the Sec. 1, an alternative account of the liaison process assumes that the variants with and without the liaison consonant are both stored in memory. Such an account has been discussed by phonologists and psycholinguists addressing the allomorphic variation in determiners across languages (e.g., Klausenburger, 1984; Miozzo and Caramazza, 1999). In this account, both the adjective and noun send activation to the corresponding determiner form and the selection is completed faster when the two send activation to the same determiner form. This account predicts shorter naming latencies for consistent sequences and can readily explain their shorter duration. Several lines of evidence converge to suggest that activated representations at a given processing stage can pass on activation to the next processing level, irrespective of whether they are finally selected or not (cascading activation models, e.g., Goldrick, 2006). A similar mechanism could apply here with the articulation of determiners in inconsistent utterances being influenced by the activation of the alternative form during word form access. Note, however, that the present data do not allow us to choose between this hypothesis and an alternative account in which the influence of phonological consistency on the acoustic duration of the determiner is independent from the ease of retrieval for the determiner.

Our conclusion that determiners with a liaison consonant are represented with two variants in the output phonological lexicon complements previous findings with other variation phenomena. In several studies, French and English speakers have been shown to represent nouns with a schwa and a non-schwa variant with two corresponding phonological representations (e.g., Bürki and Gaskell, 2012). These findings call for a re-evaluation of the shared assumption within word production models in the psycholinguistic domain that each word is represented in memory by a single corresponding word form.

To conclude, the present work highlights once again how the study of variation phenomena can be used to constrain models of speech processing. The phonological consistency effects on naming latencies and determiner acoustic duration provide information on the nature of speakers' linguistic knowledge. They suggest that systematic variation is better accounted for in terms of multiple abstract word form representations than in terms of contextual rules.

Acknowledgments

The authors thank Cassandre Mabut and Capucine Du Peloux for their help in collecting the data and aligning the naming latencies. This work was partly supported by the European Research Council under the European Community's Seventh Framework Program (FP7/2007-2013 Grant agreement No. 263575) and by the Swiss National Science Foundation (Grant No. 100014_152793).

References and links

Boersma, P., and Weenink, D. (**2015**). "PRAAT: Doing phonetics by computer (version 5.4.16) [Computer program]," http://www.praat.org (Last viewed 20 August 2015).

Bürki, A., and Gaskell, M. G. (**2012**). "Lexical representation of schwa words: Two mackerels, but only one salami," J. Exp. Psychol. Learn. Mem. Cogn. **38**, 617–631.

Bürki, A., Laganaro, M., and Alario, F.-X. (**2014**). "Phonologically driven variability: The case of determiners," J. Exp. Psychol. Learn. Mem. Cogn. **40**, 1348–1362.

Chater, N., and Manning, C. D. (**2006**). "Probabilistic models of language processing and acquisition," Trends Cognit. Sci. **10**, 335–344.

Forster, K. I., and Forster, J. C. (**2003**). "DMDX: A windows display program with millisecond accuracy," Behav. Res. Methods Instrum. Comput. **35**, 116–124.

Goldrick, M. (**2006**). "Limited interaction in speech production: Chronometric, speech error, and neuropsychological evidence," Lang. Cogn. Process. **21**, 817–855.

Hinskens, F., Hermans, B., and van Oostendorp, M. (**2014**). "Grammar or lexicon. Or: Grammar and lexicon? Rule-based and usage-based approaches to phonological variation," Lingua **142**, 1–26.

Kenstowicz, M. (**1994**). *Phonology in Generative Grammar* (Blackwell, Hoboken, NJ).

Klausenburger, J. (**1984**). *French Liaison and Linguistic Theory* (F. Steiner, Wiesbaden).

Miozzo, M., and Caramazza, A. (**1999**). "The selection of lexical-syntactic features in noun phrase production: Evidence from the picture-word interference paradigm," J. Exp. Psychol. Learn. Mem. Cogn. **25**, 907–922.

Protopapas, A. (**2007**). "CHECKVOCAL: A program to facilitate checking the accuracy and response time of vocal responses from DMDX," Behav. Res. Methods **39**, 859–862.

Spalek, K., Bock, K., and Schriefers, H. (**2010**). "A purple giraffe is faster than a purple elephant: Inconsistent phonology affects determiner selection in English," Cognition **114**, 123–128.

Tranel, B. (**1981**). *Concreteness in Generative Phonology: Evidence from French* (University of California Press, Oakland, CA).