The processing of reduced word pronunciation variants by natives and learners

Evidence from French casual speech

Max Planck Institute for Psycholinguistics

Series

Sophie Brand

The processing of reduced word pronunciation variants by natives and foreign language learners

Evidence from French casual speech

2017, Sophie Brand

ISBN: 978-90-76203-87-4

Cover image: Jolijn Ceelen

Printed and bound by Ipskamp Drukkers b.v.

The processing of reduced word pronunciation variants by natives and foreign language learners

Evidence from French casual speech

Proefschrift

ter verkrijging van de graad van doctor

aan de Radboud Universiteit Nijmegen

op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,

volgens besluit van het college van decanen

in het openbaar te verdedigen op woensdag 27 september 2017

om 16.30 uur precies

door

Sophie Wijnanda Maria Brand

geboren op 25 augustus 1985

te Roermond

Promotor: Prof. dr. M.T.C. Ernestus

Copromotor: Dr. E. Janse

Manuscriptcommissie:

Prof. dr. R.W.N.M. van Hout

Prof. dr. N. Warner (University of Arizona, Verenigde Staten)

Dr. A. Bürki (Universität Potsdam, Duitsland)

Prof. dr. H.M.G.M. Jacobs

Dr. M.B.P. Starren

The research reported in this dissertation was supported by a Consolidator grant from the European Research Council [grant number 284108] awarded to Prof. dr. M.T.C. Ernestus.

'I don't know what I may seem to the world, but, as to myself, I seem to have been only like a boy playing on the sea-shore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay all undiscovered before me'.

Isaac Newton, 1727.

Contents

Introduction

Chapter 1

In everyday casual speech, words are often produced with fewer or lenited segments compared to their citation forms. In American English, the word *hilarious* [hɪlɛriʌs], for instance, may be pronounced as something like [hlɛrɛs]. This phenomenon, called reduction, is highly frequent (e.g., for Dutch, Ernestus, 2000; for English, Johnson, 2004; Greenberg, 1999; Shattuck-Hufnagel & Veilleux, 2007; for Finnish, Lennes, Alaroty & Vainio, 2001; for French, Adda-Decker, Boula de Mareüil, Adda & Lamel, 2005; for German, Kohler, 1990; for Russian, Van Son, Bolotova, Lennes & Pols, 2004). In French, /ə/ in, for instance, *renard* 'fox', is absent in more than 50.0% of the word tokens (e.g., Fougeron, Goldman & Frauenfelder, 2001; Hansen, 1994), and word-final /ə/ in, for instance, *théâtre* 'theatre', has been reported to be absent in 35.0% of the monosyllabic word tokens and in 90.0% of the polysyllabic word tokens (Adda-Decker, Boula de Mareüil & Lamel, 1999).

Despite the ubiquity of reduction in everyday speech, little is known about how speakers produce and comprehend reduced words. The aim of this dissertation is to examine in more detail how natives and foreign language learners process reduced word pronunciation variants.

1.1 Models of speech production

A much debated question is whether reduced word pronunciation variants are stored in the mental lexicon. There is a continuum of speech production accounts differing in their assumptions about the degree to which phonetic variation is represented in lexical memory. The models at one end of the continuum assume that the pronunciation of each word is stored as a single abstract representation (e.g., Roelofs, 1997). During production, application of a phonological deletion rule (e.g., schwa-deletion rule, see McCarthy & Prince, 1993; Chomsky & Halle, 1968) transforms the lexically stored, unreduced variant (e.g., /ʀəvy/ *revue* 'magazine') into a reduced variant (e.g., /ʀvy/).

A reduced variant may also result from articulatory processes operating on the stored unreduced pronunciation variant. The results of an articulatory study on American English by Browman and Goldstein (1990) illustrate that there are (many) realisations of the word in which a consonant is articulatory present, but where it is acoustically weakly to extremely reduced. Browman and Goldstein assume that this acoustic consonant weakening is the result of the retiming of articulatory gestures resulting in stronger co-articulation. This gradient co-articulation or sound weakening leads to a continuum of pronunciations with the unreduced phoneme on the one extreme and the complete absence of the phoneme on the other.

A different type of speech production models hypothesizes a word to be lexically stored with more than one pronunciation variant (e.g., /ʀəvy/ and /ʀvy/ are both stored for *revue* 'magazine'). Absence of a phoneme results from the selection of a word pronunciation variant without that phoneme (e.g., Bürki & Gaskell, 2012; Bürki, Ernestus & Frauenfelder, 2010; Bybee, 1998). Evidence for storage of more than one pronunciation variant of a word was, for instance, provided by Bürki and Gaskell (2012). These researchers compared English words with schwa in post-stress position (e.g., as in *mackerel*) to words with schwa in pre-stress position (e.g., as in *salami*). They instructed participants to read aloud pseudo-homophones (e.g., with pre-stress schwa: *b*/ə/*loone*; with post-stress schwa: *batt*/ə/*ree*) and pseudo-words (e.g., with pre-stress schwa: *b*/ə/*leen*; with post-stress schwa: *bott*/ə/*ree*). These pseudo-homophones and pseudo-words were either presented with schwa or without schwa. The authors found shorter naming latencies for pseudo-homophones than for pseudo-words for both variants of post-stress schwa words. This suggests that the naming of the two variants of the pseudo-homophones was facilitated by the corresponding real words. In contrast, for pre-stress schwa words, the pseudo-homophone was named faster than the pseudo-word but only when schwa was present. The corresponding homophonic real word thus only facilitated the naming of the pseudo homophone with schwa. Based on these results, they argue that both variants of post-stress schwa words are lexically represented, while only schwa variants of pre-stress schwa words are.

In the two speech production models mentioned above, information about phrasal position, speech rate and phonetic context may be incorporated into the phonological deletion rules or into the lexical representations. In addition, the effects of these variables may arise during articulation.

Models at the other end of the continuum assume that all tokens of a word are lexically stored in the form of exemplars (e.g., Goldinger, 1998). The exemplars contain detailed information about the acoustic and articulatory characteristics of the tokens. The effects of phrasal position, speech rate, and phonetic context may arise from the exact specifications of the exemplars or may arise during articulation.

To investigate which model best describes the production of casual speech, I study, in Chapter 2, which processes underlie the production of reduced phonemes.

1.2 The nature of the processes underlying the absence of phonemes

Only a handful of studies have investigated whether a given reduction pattern only results from gradient reduction or also from categorical reduction. As a consequence, we know very little about the frequency of categorical reduction processes and which phonemes are especially prone to categorical reduction.

We thus do not know whether the categorical absence of phonemes reported in the literature (e.g., Bürki & Gaskell, 2012) represent exceptions or instances of a very general pattern. The answer to this question has consequences for our view of the processing of everyday speech. We therefore need more studies of other reduction phenomena to address this.

Several methods have been used to determine whether phoneme absence only results from gradient or also from categorical processes. First, the phonetic properties of the word have been studied (e.g., Zimmerer, Scharinger & Reetz, 2011). Phoneme absence has been assumed to result from a categorical process when the phoneme is completely absent and has not left any traces in the acoustic signal. In contrast, the absence of a phoneme results from gradient reduction when the phoneme has left traces, for instance, in durational or spectral properties of surrounding phonemes. In their analyses of French vowel reduction, Barnes and Kavitskaya (2002) observed residues of absent vowels in lip rounding. Bürki, Fougeron, Gendrot, and Frauenfelder (2011b) reported that shorter schwas were more co-articulated than longer schwas and took on the spectral characteristics of the surrounding consonants, like the second formant. Meunier and Espesser (2011), and Adda-Decker, Gendrot, and Nguyen (2008) demonstrated that shorter vowels show more centralised formant values than unreduced vowels.

A second method to attribute the absence of a phoneme to either gradient or categorical processes is the inspection of the distribution of the duration of the phoneme (e.g., Bürki, Fougeron & Gendrot, 2007). If the distribution shows two clearly distinct peaks, one of which is around zero ms, representing the absent phoneme, absence of the phoneme is assumed to result from both a categorical and a gradient reduction process. If, in contrast, the distribution is unimodal, without a clear peak around zero ms, absence of the phoneme is assumed to result from gradient shortening. For the duration of French word-initial [s(e)] in *c'était* /sete/ 'it was', Torreira and Ernestus (2011) found a bimodal distribution, with one mode containing most tokens with [e] (38.0% of the tokens, mean duration: 37 ms) and the other containing most tokens without [e] (62.0% of the tokens in which [e] had a duration of zero ms). This pattern suggests that the absence of [e] in French *c'était* especially results from a categorical process rather than from extreme vowel shortening. Similarly, Bürki, Fougeron, and Gendrot (2007) observed a bimodal distribution for the duration of schwa in words like *sera* /səʀa/ 'will be'. Their results support the hypothesis that the processes underlying schwa reduction in French are not only gradient but also categorical in nature. In contrast, another study by Bürki and colleagues (2011b) showed that the distribution of the duration of French schwa in word-initial syllables (e.g., [ə] in /ʀəvy/ *revue* 'magazine') is close to normal, and that schwa duration varies from very short to very long. They therefore concluded that the absence of schwa may mainly be the result of a process of gradual phonetic reduction. Bürki and colleagues (2011b) indicate that the presence of contextual information for a given token (e.g., tokens in isolation, Bürki et al., 2011b; tokens in continuous speech; Bürki et al., 2007) and methodological options (e.g., acoustic examination, analysis of the distribution of phoneme duration, perceptual categorization in Bürki et al., 2011b versus analysis of the distribution of phoneme duration in Bürki et al., 2007) are likely to have influenced the categorization of the reduction of schwa as categorical or gradient.

A third approach addressing the issue of gradient versus categorical reduction is to compare the variables predicting phoneme absence and those predicting phoneme duration. Perfect (or substantial) overlap in predictors would suggest that phoneme absence and duration are driven by the same processes, and that phoneme absence is the endpoint of gradient sound shortening. In contrast, if phoneme absence and phoneme duration are affected by different variables, they are more likely to result from different processes, and the absence of a phoneme may then also result from a categorical process. Largely the same phonetic variables were found to predict presence and duration of suffixal [t] in past participles (e.g., /t/ in /xəmakt/ *gemaakt* 'made') in a study on Dutch informal speech (Hanique, Ernestus & Schuppler, 2013). Hanique and colleagues demonstrated that the absence of [t] may simply be the extreme result of shortening. On the other hand, Torreira and Ernestus (2011) found evidence for categorical absence of [e] in French *c'était* 'it was'. They observed that the presence and duration of [e] are conditioned by different variables, which suggests that presence and duration of [e] are determined by different processes, as was also suggested by the distribution of the duration (the second method).

So far, these three methods (i.e., phonetic analyses, examination of the distribution of phoneme duration, and the comparison of the predictors of phoneme absence and phoneme duration) have not been combined in studies on the processes underlying the absence of phonemes. In this dissertation, I combine these three methods to examine whether phoneme reduction affects the realisation of neighbouring phonemes.

1.3 Perception of reduced words by native listeners

In the two following sections, I review literature addressing the processing advantage of unreduced variants in native speech comprehension and the different cues on which native listeners rely during speech comprehension. The presented studies provide an important basis for follow-up research on the perception of reduced words.

1.3.1 Processing advantage for the unreduced variant

Previous research has shown that natives recognize a noun in isolation less well when it is presented in a reduced variant than when it is presented in an unreduced variant (e.g., Janse, Nooteboom & Quené, 2007; Tucker & Warner, 2007; Ernestus, Baayen & Schreuder, 2002). Tucker and Warner, for instance, showed that American English listeners responded more quickly and accurately to a visual stimulus preceded by an unreduced pronunciation variant of the same word than to a visual stimulus preceded by a reduced pronunciation variant. For example, an unreduced variant of *puddle* produced greater priming than a reduced variant of *puddle*.

The processing advantage for the unreduced variant over the reduced variant of a word may decrease in context. Ernestus, Baayen, and Schreuder (2002), for instance, reported that Dutch listeners recognized highly reduced variants well when these forms were presented in their full context, but less well when presented in isolation or limited context.

The EEG study by Drijvers, Mulder, and Ernestus (2016) partly supports this finding. Drijvers and colleagues reported that alpha power increase, which was interpreted as a reflection of higher auditory cognitive load, was not larger in response to reduced variants than to unreduced variants in mid-sentence position. In contrast, gamma power increases were observed to unreduced but not to reduced variants, also in mid-sentence position, which suggests that it takes more effort for reduced variants than for unreduced variants to activate their semantic networks. The researchers conclude that there is a processing advantage for unreduced variants, especially in non-medial sentence position. In non-medial sentence position, words tend to carry accent which prevents them from being reduced and reduction in these positions is therefore rare (e.g., Kohler, 1990; Dalby, 1984).

Natives thus seem to process the unreduced variant more easily than the reduced variant, especially in sentence non-medial position or in limited context. The processing advantage of unreduced variants will be further investigated in this dissertation.

1.3.2 Different types of cues in perception

Natives may recognize reduced words because they use fine phonetic details in the acoustic signal, such as voicing and lengthening of surrounding phonemes (e.g., Zimmerer & Reetz, 2014; Manual, 1991). Zimmerer and Reetz, for instance, reported that the duration of the [s] in [st] clusters helps German native listeners to reconstruct a reduced [t]. Manuel reported that the laryngeal specification of the [p] in English *s'pport* enables native listeners of English to distinguish reduced *s'pport* from *sport*. Niebuhr and Kohler (2010) investigated which characteristics of a speech signal lead native listeners of German to interpret a stretch of speech as a two-word phrase (*eine Rote* 'one red one') or as a highly reduced version of a three-word phrase (*eigentlich eine Rote* 'actually a red one'). They reported that the duration of the palatal glide section in the word-initial diphthong made it possible to distinguish between *eine* 'one' and *eigentlich'ne* 'actually a'.

In addition to using acoustic cues, natives use semantic and syntactic context following and preceding a word to understand reduced word pronunciation variants. Van de Ven, Ernestus, and Schreuder (2012), for instance, demonstrated that Dutch listeners use both a word's preceding and following context to process reduced words. Similarly, participants were shown to recognize Dutch past participles more quickly when they occurred after their associated auxiliary verbs than when they preceded them in an eye-tracking study (Viebahn, Ernestus & McQueen, 2015). This effect was stronger for casually than for carefully produced sentences. Tuinman, Mitterer, and Cutler (2014) focused on Dutch present tense verb forms (e.g., *rent* 'runs') and examined whether Dutch listeners make use of syntax in the perception of words in which affixal [t] is reduced. Listeners reported more [t]s in sentences in which [t] would be syntactically correct than in sentences in which [t] would be syntactically incorrect.

Words differ in how often they occur in a certain pronunciation variant. Frequency of occurrence of pronunciation variants may also serve as a cue in processing. Native listeners may recognize more frequent variants more quickly than less frequent variants. Ranbom and Connine (2007), for instance, studied how quickly American English listeners recognize words that can be pronounced with both [nt] and a nasal flap (e.g., *gentle*) and found that native listeners recognize the flap variant more quickly if it is the word's most frequent variant (as is the case for *gentle*) than if it is not the most frequent variant (as for *lantern*). Pitt, Dilley, and Tat (2011) reported similar results for the recognition of American English words in which word-medial /t/ occurs in one of four phonological contexts, each favouring one of four realisations of /t/ ([t], [ʔ], [ɾ] or a deleted variant). Pitt and colleagues documented a recognition benefit for the variant that is the most frequent in the given phonological context.

1.4 Comprehension of reduced words by adult language learners

Previous research has shown that learners encounter more difficulties than natives in the comprehension of reduced words. Nouveau (2012) asked advanced Dutch foreign learners of French to perform a transcription task. The task included words with word-medial schwa (like [ʀəvy] *revue* 'magazine'), which was left out in some words. The students correctly transcribed no more than 56.1% of the French words that were produced without schwa, while they correctly transcribed 92.6% of the words produced with schwa. In line with Nouveau (2012), Wong, Mok, Chung, Leung, Bishop, and Chow (2015) found that intermediate to advanced Hong Kong students of English made errors for 40.0% of the words showing regular simple phoneme reduction in an English transcription task. The processing of reduced words by second language learners who learn their second language in the country where it is the native language have also been tested (e.g., Ernestus, Dikmans & Giezenaar, in press). Ernestus and colleagues auditorily presented Dutch reduced and unreduced sentences to advanced second language (L2) learners of Dutch and observed more transcription errors for reduced sentences (70.0%) than for unreduced sentences (38.0%). These results demonstrate that there is a large processing advantage for unreduced variants for learners, at least in transcription tasks. The question is whether the processing advantage is also reflected in other perception/comprehension tasks.

Another question that arises is whether second language learners and foreign learners, like natives, benefit from subtle phonetic, semantic, syntactic and frequency information during the processing of reduced speech. Very few studies have investigated how language learners understand reduced words.

Learners of English with native languages that do not allow word-final /nt/ (Spanish and Mandarin) have been shown to have problems interpreting reduced variants of this word-final cluster because they do not make full use of subtle phonetic cues (Ernestus, Kouwenhoven & van Mulken, 2017). Their results demonstrate that comprehension of casual speech in a foreign language is affected by phonotactic constraints in the listener's native language. Similar results were obtained in a study examining advanced second language learners of Dutch (Ernestus, Dikmans & Giezenaar, in press). Advanced learners of Dutch transcribed, for instance, reduced *eind December* 'end of December' as *één*

December 'first of December'. Unreduced *eind* and *één* differ in the quality of the vowel ([ɛi] versus [e]) and in the presence versus absence of [t]. In the reduced sentence, *eind* was produced without clear [t]. The learners' transcription errors showed that they could not benefit from the difference in vowel quality between *eind* and *één* or from the subsegmental cues to the [t].

Simiarly, the advanced learners of Dutch in the study by Ernestus and colleagues (in press) did not always make full use of the semantic and syntactic information provided by the signal, as was evidenced by the numerous incorrect transcriptions that did not fit the syntactic structures and semantic contents of the sentences. Occasionally, they did make use of words introducing time specifications and place specifications (e.g., *in* 'in' introducing *Nederland* 'the Netherlands').

The findings for advanced second language learners are in line with the results for beginner second language learners reported by Van de Ven, Tucker, and Ernestus (2010). Beginner Asian learners of English were shown to have more difficulties than native listeners in using (subtle) semantic cues in their second language for lexical access, not only when listening to reduced words, but also in their processing of unreduced words (see also, e.g., Bradlow & Alexander, 2007).

Frequency effects of the different variants on language learners' processing of different pronunciation variants have so far not been examined. It is still unknown whether language learners are sensitive to the frequency of occurrence of a word pronunciation variant in a foreign language. Learners may not show sensitivity to the frequency of occurrence of a given pronunciation variant because they rely more on other sources of information (e.g., the structure of the sentence) as they rarely encounter reduced variants (e.g., Gilmore 2004, Fonseca-Greber & Waugh, 2003; Waugh & Fonseca-Greber, 2002; McCarthy & Carter, 1995). This issue will be taken up in this dissertation.

1.5 Reduction in models of speech perception

Reduction can be accounted for in several ways in speech perception models. Some speech perception models hypothesize, like some speech production models (see section 1.1), that a word is only stored with one pronunciation in a single abstract representation. Reduced variants (e.g., French [ʀvy] *revue* 'magazine') can be matched with the lexical representations for the unreduced variants (e.g., /ʀəvy/), by means of general processes (e.g., schwa-insertion), each applying to multiple words. This 'abstractionist' account is compatible with many psycholinguistic models of word recognition, such as TRACE (McClelland & Elman, 1986), the Neighborhood Activation Model (NAM: Luce & Pisoni, 1998), PARSYN (Luce, Goldinger, Auer & Vitevitch, 2000) and Shortlist B (Norris & McQueen, 2008).

An alternative abstract model for the processing of segment reduction is the *naive discriminative learning model*, as proposed by Baayen (2010). In this model, which is a two-layer network with symbolic representations for a word's form and a word's semantics, each input unit (unigram, bigram, trigram, uniphone, biphone, or triphone) is linked to each meaning, and a weight is associated with each link. Baayen states that the low frequencies of segment combinations in some reduced words, like [nt] in Dutch [ntyk] /natyrlək/ *natuurlijk* 'of course', make these combinations excellent cues for discriminative learning. The connection weights from the biphone [nt] to the meaning 'of course' will be relatively high because this biphone is rather rare, and once the reduced form has been heard and interpreted correctly, discriminative learning predicts it will be relatively easy to understand this form on subsequent encounters.

In line with the speech production model presented in section 1.1 (e.g., Bürki & Gaskell, 2012; Bürki, Ernestus & Frauenfelder, 2010; Bybee, 1998), both unreduced and reduced word pronunciation variants can be argued to be stored in the mental lexicon in speech perception models. A French speaker can recognize unreduced [ʀəvy] *revue* ‘magazine’ by matching the acoustic signal with the lexical representation /ʀəvy/. Reduced [ʀvy] can be recognized by matching the acoustic signal with the lexical representation /ʀvy/. This account requires that the word recognition models, such as Shortlist B, extend their lexicons with additional word pronunciation variants (see e.g., Ranbom & Connine, 2007).

Like in speech production models (see section 1.1), fine-grained phonetic details of each occurrence of a word can be argued to be stored in perception models (e.g., Johnson, 2004; Goldinger, 1998). In exemplar-based models, the fine phonetic details for each word token, containing, for instance, speaker-specific and situation-specific information are stored in the form of an exemplar. The many exemplars of every word together are assumed to form a word cloud. Perception of a pronunciation variant involves the direct mapping of the perceived acoustic event on the best matching exemplar(s) in the cloud. The comprehension process thus proceeds without a mediating pre-lexical representation and without speaker normalization.

Alternatively, hybrid theories of lexical representation vary with respect to the number of phonetic variants that are stored and the degree to which phonetic detail is retained in memory. Two hybrid models for speech recognition are: 1. the hybrid model proposed by McLennan, Luce, and Charles-Luce (2003) based on the Adaptive Resonance Theory (ART) developed by

Grossberg and Stone (1986), and 2. Goldinger's Complementary Learning System (CLS, 2007). In the first model, the acoustic input activates chunks of lexical and sublexical representations. These chunks may be abstract (e.g., phonological features) or captured by exemplars (e.g., token specific information). The abstract representations are more frequent than any of the exemplars and these abstract representations therefore establish resonance with the input more easily and more quickly. In the second model, the acoustic signal first passes that part of the brain that is involved in abstract processing. With all fine phonetic detail still present, it is then matched to similar indexical properties. Both models thus assume that processing based on abstract representations precedes processing based on exemplars.

There is no general agreement about whether and how pronunciation variants of a word are stored in the mental lexicon. More research on the processing of unreduced versus reduced pronunciation variants will help us address this theoretical debate and unravel the processes underlying the comprehension of casual speech.

1.6 Reduction in French

Reduction is a frequent phenomenon in French casual speech, as it is in many other languages. Many previous studies have investigated the absence of schwa in Standard French (e.g., Bürki, Ernestus, Gendrot, Fougeron & Frauenfelder, 2011a; Racine & Grosjean, 2002; Fougeron, Goldman & Frauenfelder, 2001; Hansen, 1994). These studies showed that schwa in unstressed word-initial syllables is absent in more than 50.0% of spontaneously produced word tokens. The French word combination /la fənɛtʁə/ 'the window' is thus often pronounced as [laf.nɛtʁə]. Schwa reduction may result in resyllabification: in the example above, the onset of the noun has become the coda of the determiner.

Note that schwa in the initial syllable of French words can be absent even if the words are produced in isolation and in a formal situation (e.g., Bürki, Ernestus & Frauenfelder, 2010). The absence of schwa in a word in isolation is often complete, leaving no traces in the acoustic signal (e.g., Bürki, Fougeron, Gendrot & Frauenfelder, 2011b; Bürki, Fougeron & Gendrot, 2007). In word-final position, schwa has been reported to be absent in 35.0% of the word tokens in which it is preceded by one syllable, and in 90.0% of the word tokens in which it is preceded by several syllables (two, three, or more; Adda-Decker, Boula de Mareüil & Lamel, 1999).

Several studies have examined the reduction of European French single consonants (e.g., Milne, 2014; Adda Decker, Gendrot & Nguyen, 2008; Armstrong, 1998; Su & Basset, 1998; Dell, 1995; Laks, 1977). Laks studied six adolescents from Villejuif (commune in the southern suburbs of Paris) conversing with each other. He observed that the more formal the conversation and the higher the speaker's social status, the less likely [ʁ] was to be absent. Armstrong also examined liquids; he studied speech produced by French girls aged 11-12 years from Dieuze (north-eastern France), and also observed that the more formal the speech style, the less likely the liquid ([l] or [ʁ]) was to be reduced. Su and Basset studied word-internal consonants spontaneously produced by two male natives of European French and reported that only voiced fricatives and the liquid [l] may be absent. This latter finding was confirmed in a corpus study of phone conversations by Adda-Decker, Gendrot, and Nguyen, who showed that [l] is the most likely consonant to be absent (43.0%).

Segmental and prosodic context has been shown to influence the absence of a phoneme. Milne (2014), for instance, reported that in European French, the word-final consonant is more often absent in preconsonantal (44.2%) than in prevocalic (23.7%) or prepausal (31.6%) position, and that schwa following a word-final consonant cluster is more often absent before a vowel-initial word (87.5%) than before a consonant-initial word (63.9%) or before a pause (69.7%). These results confirm those reported for other languages (e.g., Bybee, 2015).

This dissertation provides more information on reduction in European French, investigating both vowel and consonant reduction. In this way, it increases the impact of Roman languages on our knowledge of speech reduction.

1.7 Research questions

This dissertation seeks to provide more information about the production and perception of reduced word pronunciation variants by natives and language learners. I base my studies on French. Most studies on reduction have focused on English, Dutch and German (e.g., for English, Johnson, 2004; Greenberg, 1999; for Dutch, Ernestus, 2000; for German, Kohler, 1990). Nevertheless, we know that reduction is also very frequent in French (e.g., Milne, 2014; Bürki, Fougeron, Gendrot & Frauenfelder, 2011b; Adda-Decker, Gendrot, and Nguyen, 2008; Bürki, Fougeron & Gendrot, 2007; Adda-Decker, Boula de Mareüil & Lamel, 1999). Additional studies on French reduction may be informative as to whether reported findings generalize to other languages.

My first question is: How do speakers produce reduced word pronunciation variants? In order to increase our knowledge of this phenomenon, I examine word-final obstruent-liquid-schwa (henceforth OLS) clusters, like /tʁə/ in *ministre* 'minister'. I investigate the frequencies of occurrence of the pronunciation variants for the OLS cluster, the relationship between the reductions of the individual OLS phonemes, and the nature of the processes underlying the reduction of the OLS phonemes. If the nature of the processes underlying the absence of phonemes turns out to be mainly categorical, this would have consequences for our view of the processing of everyday speech.

My second question is: What is the effect of reduction on listeners' processing? Given that foreign language learners have less often been exposed to reduced speech than natives, and given that they understand unreduced pronunciation variants much better than reduced pronunciation variants (see section 1.4), the processing advantage for the unreduced variant over the reduced variant may be larger in learners than in natives. In addition to natives, both Dutch advanced and beginner learners of French are tested on French reduction. The Dutch advanced learners are undergraduate students of French. Their proficiency levels roughly correspond to C1-C2 level according to the Common European Framework of Reference for Languages (CEFR, Council of Europe, 2011). The Dutch beginner learners are undergraduate students of other studies with proficiency levels roughly corresponding to B1-B2 level.

My third, and most important question is: Which mechanisms underlie the recognition process of reduced speech in natives and language learners? I study whether natives and foreign language learners rely on the variants' frequencies of occurrence during the processing of reduced speech. Moreover, the effect of reduction on lexical competition and on the semantic activation of words will be examined. Finally, I explore whether the learners make use of phonetic, semantic or syntactic information provided by the word or the context.

1.8 Methods

In order to address the research questions introduced above, I use a wide range of research methods. I investigate the nature of the processes underlying the reduction of a sequence of phonemes in a study based on the Nijmegen Corpus of Casual French (NCCFr, Torreira, Adda-Decker & Ernestus, 2010). In contrast to previous studies which restricted their research on the nature of reduction processes to one single method (e.g., Meunier & Espesser, 2011; Torreira & Ernestus, 2011; Bürki, Fougeron & Gendrot, 2007), the corpus study presented in Chapter 2 employs three different methods. The three methods are the ones described in section 1.2: studying the phonetic properties of the phoneme sequences resulting from the absence of the OLS phonemes, inspection of the distributions of the durations of the OLS phonemes, and comparison of the factors that predict phoneme absence and phoneme duration. Comparison of the results obtained with these three methods not only provides more information about the processes underlying OLS phonemes' absence but also shows the advantages and disadvantages of these methods.

To examine the second research question, that is, whether the effect of reduction is larger in language learners' word recognition than in natives' word recognition, I use three different methods. First, French natives and Dutch advanced language learners are asked to perform a transcription task containing words in which single and multiple segments are reduced. Like in everyday communication, transcription tasks will present listeners with speech varying in rate, in degree of casualness, and in disfluencies. In addition, transcription tasks provide information about the processing of *all* the words in a sentence, rather than that of a single word, and can offer important insights into the eventual interpretation of casual speech by natives and language learners.

Second, I ask French natives and Dutch advanced learners of French to perform a visual lexical decision task with cross-modal identity priming. Visual target words are combined with auditorily presented prime sentences, each containing a prime word, which is either: 1. an unreduced variant of the target word (e.g., [ministʁ] or [ministʁə] for *ministre* 'minister'); 2. an infrequent, weakly reduced variant of the target word (e.g., [minisʁ] for *ministre*); 3. a highly frequent, strongly reduced variant of the target word (e.g., [minis] for *ministre*); 4. or a semantically and phonetically unrelated word (e.g., *virage* 'turn' for *ministre*). At the offset of the prime word, the visual stimulus appears on the computer screen for which the participants have to take a lexical decision. This task shows how fast and how accurately the different prime word types are processed and whether unreduced prime words induce faster response times (henceforth RTs) in the visual lexical decision task than reduced prime words.

Third, natives and foreign language learners are asked to perform an auditory lexical decision task on French unreduced and reduced schwa words that are only preceded by their determiners. Participants are asked to indicate whether the presented word is an existing word of the language or not. Participants' RTs and accuracy scores inform us about the recognition of reduced variants relative to unreduced variants in limited context.

Two methods are used to examine the first part of the third research question, that is, whether natives and language learners rely on the frequencies of occurrence of pronunciation variants during the processing of these variants. First, I investigate in the lexical decision task with cross-modal identity priming (mentioned above) whether language learners respond faster and more accurately to pronunciation variants of phoneme sequences that occur more often. I test whether the frequencies of the variants of the OLS clusters that I observed in Chapter 2 predict how easily these variants are processed. The speech that learners hear is likely to be less reduced than the type of speech that natives typically hear, as foreign language learners are less often involved in casual conversations. Dutch learners of French may therefore store non-native-like frequency information, which may not be useful during the processing of French casual speech.

Second, I ask French natives and Dutch learners to perform a relative frequency estimation task in which they estimate the relative frequencies of two pronunciation variants (unreduced and reduced) of a (large) set of words. I then investigate whether participants respond faster in an auditory lexical decision task to variants that according to them (as a group) occur more often.

To investigate the second part of the third research question, that is, the effect of reduction on lexical competition and on the semantic activation of words, I conduct a perception study in which I simultaneously measure participants' eye movements and ERPs. I test French natives and Dutch advanced learners of French. They are presented with sentences that either contain an unreduced or a reduced pronunciation variant of a target word. At the same time, four objects (one representing the target word, one representing a phonological competitor, two representing neutral distractors) appear on the computer screen. The proportion of eye fixations on the different objects is taken as an index of the amount of lexical competition the participant experiences during recognition. The recorded EEG signal (N400) is taken to index the difficulty of lexical-semantic activation of the reduced or the unreduced word pronunciation variant presented in context. In this way, I simultaneously investigate two processing stadia.

Finally, I investigate whether learners make use of the phonetic, semantic and syntactic cues in the word or the context by conducting a transcription task. Participants transcribed sentences extracted from a spontaneous casual conversation.

All studies reported in this dissertation were approved by the local ethics committee.

1.9 Outline

In four studies, I investigate how reduced word pronunciation variants are processed in speech production and perception. Chapter 2 focusses on speech production. In this chapter, I study reduction of word-final obstruent-liquid-schwa clusters in casual French (e.g., /tʁə/ in /fənɛtʁə/ 'window'). The frequencies of occurrence of the clusters' pronunciation variants are reported. I further focus on the phonetic properties of the OLS phonemes, and present the distributions of the durations of the OLS phonemes and the statistical analyses of absence and duration of the OLS phonemes. These three types of data provide information about the type of processes (categorical or gradient) underlying the reduction of the phonemes.

In Chapter 3, I investigate the extent to which reduced speech affects natives' and language learners' processing and on which type of information (i.e., phonetic, semantic, syntactic, frequency information) natives and learners rely during the processing of casual speech. I first present the results of a transcription task, which shows how natives and foreign language learners perceive reduced words in context and whether natives and foreign language learners rely on acoustic, semantic and syntactic cues. I then report the results of a lexical decision task with cross-modal identity priming. The results of this experiment provide information about the use of frequency information of the variants during the processing of variants of OLS clusters in context (see section 1.8).

In Chapter 4, I focus on schwa reduction in words in isolation to investigate the recognition of reduced word pronunciation variants and the role of variant frequency in processing therein. Word-initial schwa can be absent if the words are produced in isolation and in a formal situation (e.g., /ʀəvy/ 'magazine' can be pronounced as [ʀvy]; Bürki, Ernestus & Frauenfelder, 2010). I present RTs and accuracy scores of natives and two learner groups with different proficiency levels on an auditory lexical decision task containing French unreduced and reduced schwa words. I further report whether the participants' estimations of these variants' frequencies obtained in a relative frequency estimation task (see section 1.8) can predict listeners' RTs and accuracy in the auditory lexical decision task. I investigate whether learners' accuracies and RTs are better predicted by their own relative frequency ratings of the unreduced and reduced variants of a given word than by the natives' ratings.

In Chapter 5, like in Chapter 4, I focus on the processing of unreduced and reduced schwa words. I examine the continuous processes that take place in the recognition of unreduced and reduced schwa words presented in context. I report the results of a combined eye-tracking and ERP study, examining the lexical competition that takes place in both natives' and learners' processing, and the difficulty of lexical-semantic activation of a word in its context.

In Chapter 6, I summarize and discuss the results of these four studies. I focus on what these studies tell us about the nature of the processes underlying the reduction of phonemes, the effect of reduction on speech processing by natives and language learners, the role of variant frequency in natives' and learners' processing, and the use of phonetic, semantic or syntactic information during the processing of casual speech. Moreover, I discuss the implications of the results for psycholinguistic models of speech production and speech comprehension.

Reduction of word-final obstruent-liquid-schwa clusters in casual European French

Chapter 2

This chapter is a reformatted version of:
Sophie Brand & Mirjam Ernestus (submitted). Reduction of word-final obstruent-liquid-schwa clusters in casual European French.

Abstract

This corpus study investigated pronunciation variants of word-final obstruent-liquid-schwa clusters in nouns in casual European French. Results showed that at least one phoneme was absent in 80.7% of the 291 noun tokens in the dataset, and that the whole cluster was absent (e.g., [mis] for *ministre*) in no less than 15.5% of the tokens. We demonstrate that phonemes are not always completely absent, but that they may leave traces on neighbouring phonemes. Further, the clusters display undocumented voice assimilation patterns. Statistical modelling showed that a phoneme is most likely to be absent if the following phoneme is also absent. The durations of the phonemes are conditioned particularly by the position of the word in the prosodic phrase. We argue, on the basis of three different types of evidence, that in French word-final obstruent-liquid-schwa clusters, the absence of obstruents is mainly due to gradient reduction processes, whereas the absence of schwa and liquids may also be due to categorical deletion processes.

2.1 Introduction

In casual speech, words tend to be produced with fewer or lenited phonemes compared to their citation forms. In American English, the word *hilarious* [hɪlɛriʌs], for instance, may be pronounced as something like [hlɛrɛs] (e.g., Ernestus & Warner, 2011; Johnson, 2004). This phenomenon, called reduction, is highly frequent and has been studied in a number of languages, such as Dutch (e.g., Ernestus, 2000), German (e.g., Kohler, 1990), Finnish (e.g., Lennes, Alaroty & Vainio, 2001), and Russian (e.g., Van Son, Bolotova, Lennes & Pols, 2004). This study contributes to our knowledge of reduction by providing a detailed analysis of a reduction pattern in the Romance language French. We use the terms *reduction* and *reduced* to refer to both absence and weakening of segments. The term *absent* refers to segmental absence, that is, to a phoneme that is at best present in the details in the pronunciation of the neighbouring segments.

Previous studies have shown that reduction especially affects frequently occurring words (e.g., Pluymaekers, Ernestus & Baayen, 2005a; Bybee & Scheibmann, 1999). For instance, in American English conversational speech (Johnson, 2004), the percentage of phonemes in a word that deviates or is absent from those in the word's citation form is higher in high-frequency function words than in content words (20% deletion and 40% deviation for function words versus 10% deletion and 25% deviation for content words respectively). Similarly, Greenberg (1998) reported that, in American English, pronouns, articles, conjunctions and modal/auxiliary verbs are most susceptible to phoneme reduction. Likewise, in Dutch, function words like articles, pronouns, and conjunctions are most prone to deletion (Van Bael, Baayen & Strik, 2007). As a consequence of this, many studies on reduction have focused on high-frequency words. Bell, Jurafsky, Fosler-Lussier, Girand, Gregory, and Gildea (2003), for instance, investigated the ten most frequent English function words and Pluymaekers, Ernestus, and Baayen (2005b) studied occurrences of the seven most frequent Dutch words ending in the adjectival suffix *–lijk* [lək] *–able.* Less is known about reduction in content words of low to medium frequency.

Another restriction of most previous studies is that they focused on the reduction of single phonemes (consonants: e.g., Raymond, Dautricourt & Hume, 2006; Bybee & Scheibmann, 1999; vowels: e.g., Patterson, LoCasto & Connine, 2003). The few studies that have examined reduction of sequences of phonemes (e.g., Milne, 2014; Côté, 2004) have mainly focused on the reduction of one of the cluster's phonemes. Côté, for instance, reported that liquid absence in French word-final obstruent-liquid-schwa sequences is more constrained in Parisian varieties than in any other variety of French. Milne observed that in European French word-final consonant is often completely deleted in clusters spanning the word boundary that violate the Sonority Sequencing Principle (SSP). Study of the reduction of all phonemes in a sequence may, however, provide important information about the processes underlying speech reduction. When the articulatory movements for a given phoneme are reduced, this may affect the articulation of the preceding and/or following phoneme, which may still be present, but acoustically weakened.

Furthermore, nearly all previous studies on reduction have focused on the presence versus absence of phonemes and on the duration of individual phonemes or words (e.g., Pluymaekers et al. 2005a; Johnson, 2004; Bybee & Scheibmann, 1999). Only a few studies have investigated more detailed phonetic characteristics of reduced word variants, showing that absent phonemes, which have durations of zero ms, are still cued by the detailed acoustic characteristics of neighbouring phonemes. Manuel (1991) reported, for instance, that English *s'pport* (the apostrophe marks the absence of schwa) differs from *sport* due to the presence of an oral (mouth opening) or glottal (vocal fold adduction) gesture for the underlying schwa. Ernestus and Smith (forthcoming) studied the pronunciation of the Dutch discourse marker *eigenlijk* 'actually', and showed that although the [l] frequently has a duration of zero ms, acoustic traces of [l] often remain in the neighbouring fricative, for instance as a dip in the second formant.

In sum, little is known about the influence of phoneme reduction on the (detailed) realisations of sequences of phonemes in (low frequency) content words. In this study, we aim to contribute to filling these gaps by examining obstruent-liquid-schwa (henceforth OLS) clusters at the end of spontaneously produced nouns [1] (e.g., [tʁə] in [ministʁə] *ministre* 'minister') in European French. We restricted our analyses to one word class, that is, nouns, because we aimed to investigate in detail a word class that, so far, has been shown to be least susceptible to phoneme reduction. We selected OLS clusters as previous studies (e.g., Brand & Ernestus, 2015; Milne, 2014; Laks, 1977) have reported that these clusters are frequently reduced. The present study was set up to investigate the possible pronunciation variants of noun-final OLS clusters, their phonetic characteristics, the frequencies of the pronunciation variants, the predictors of these pronunciation variants, and the relationship between the reductions of the individual OLS phonemes.

Previous studies have investigated the absence of schwa in French. We only briefly discuss here some studies on European French as we also study this variety of French (e.g., Bürki, Ernestus, Gendrot, Fougeron & Frauenfelder, 2011a; Racine & Grosjean, 2002; Fougeron, Goldman & Frauenfelder, 2001; Hansen, 1994). These studies have shown that schwa in word-initial syllables is

[1] We consider the obstruent and the liquid as well as the schwa, although according to some literature (e.g., Dell, 1985), the schwa is always absent or only surfaces as the result of a process of schwa insertion.

absent in more than 50.0% of spontaneously produced word tokens. The French word /pəluz/ *pelouse* ‘lawn’ is thus often pronounced as [pluz]. For our study, it is important to note that in word-final position, schwa has been reported to be absent in 35.0% of the word tokens in which it is preceded by one syllable, and in 90.0% of the word tokens in which it is preceded by several syllables (two, three, or more; Adda-Decker, Boula de Mareüil & Lamel, 1999).

Reduction of schwa in French has been shown to be influenced by many socio-indexical factors (e.g., regional accent, social status) and speech style (formal versus informal). Durand and Eychenne (2004) and Coquillon (2007), for instance, have shown that schwa deletion is more wide spread in the North of France than in the South of France. Furthermore, Eychenne and Pustka (2007) demonstrated that young educated speakers from the South of France are more likely to delete schwa in frequent verbs and/or specific constructions than their middle-aged or older counterparts. We studied words produced by young educated speakers from the North of France.

Context has been shown to influence the absence of a phoneme. Milne (2014), for instance, reported that schwa following a word-final consonant cluster is more often absent before a vowel-initial word (87.5%) than before a consonant-initial word (63.9%) or before a pause (69.7%). These results confirm those reported for other languages (e.g., Bybee, 2015).

There are several studies that have examined the reduction of European French single consonants. Examples are Milne (2014), Adda-Decker, Gendrot, and Nguyen (2008), Armstrong (1998), Su and Basset (1998), Dell (1995) and Laks (1977). Laks studied six adolescents from Villejuif (a commune in the southern suburbs of Paris) conversing with each other. He observed that the more formal the conversation and the higher the speaker’s social status, the less likely [ʁ] was to be absent. Armstrong also examined liquids; he studied speech produced by French girls aged 11-12 years from Dieuze (north-eastern France), and also observed that the more formal the speech style, the less likely the liquid ([l] or [ʁ]) was reduced. Su and Basset studied spontaneously produced consonants produced by two male speakers and reported that only voiced fricatives and the lateral approximant [l] may be absent. This latter finding was confirmed in a corpus study by Adda-Decker, Gendrot, and Nguyen, who showed that in phone conversations, [l] is the most likely consonant to be absent (43.0%). Like schwa, the word-final consonant is more often absent in preconsonantal (44.2%) than in prevocalic (23.7%) or prepausal position (31.6%) (Milne, 2014).

As noted above, most studies on reduction, including studies on reduction in French, have focused on the absence versus presence of phonemes. Only a few studies have examined the phonetic characteristics of reduced word variants. This also holds for European French. One of these few exceptions is the corpus-based study by Bürki, Fougeron, Gendrot, and Frauenfelder (2011b). These authors reported that shorter schwas were more co-articulated than longer schwas and took on the spectral characteristics of the surrounding consonants, like the second formant. Another exception is the production study by Barnes and Kavitskaya (2002)[2], who observed that the rounding gestures required for the articulation of schwa were still present when schwa was reduced. Interestingly, Côté and Morrison (2007) did not replicate these results in their study on Québec French. Finally, Meunier and Espesser (2011) and Adda-Decker and colleagues (2008) demonstrated that shorter vowels show more centralised formant values than unreduced vowels.

[2] Barnes and Kavitskaya (2002) do not specify the variety spoken by their participant, but we can presume that it is European French.

The studies by Bürki and colleagues (2011b) and Barnes and Kavitskaya (2002), among others, show that schwa can be strongly co-articulated. This co-articulation may result in the acoustic absence of schwa. Indeed, co-articulation forms one of the two types of main processes that may underlie the absence of phonemes. Gradient co-articulation leads to a continuum of pronunciations with the unreduced phoneme on the one extreme and the absence of the segment on the other. That is, there are (many) realisations of the word in which this phoneme is present, but where it is acoustically weakly to extremely reduced.

The results of an articulatory study on American English by Browman and Goldstein (1990) nicely illustrate this for a consonant. These researchers analyzed realizations of the sequence *perfect memory,* either produced as two independent words in a word list (and thus with an intonation boundary between the two words), or produced as a grammatical unit in a fluent phrase. In the fluent phrase version, the final [t] of *perfect* was often not audible, but the alveolar closure gesture for the [t] was nevertheless always present, with almost the same magnitude as in the word list version, in which [t] was clearly audible. In the fluent phrase version, the alveolar burst was often completely overlapped by the closure of the following bilabial [m]. Hence, Browman and Goldstein assume gradient reduction to be the result of the retiming of articulatory gestures which results in stronger co-articulation. Other studies (e.g., Bybee, 2002) propose that gradient reduction may also be stored as part of a word's lexical representation.

Second, the absence of phonemes may result not only from gradient reduction processes, but also from categorical processes, which result in the complete absence of phonemes without any acoustic traces. The categorical absence of a phoneme could be due to a phonological deletion rule (e.g., McCarthy & Prince, 1993; Chomsky & Halle, 1968) or the storage of a word pronunciation variant without the given phoneme in the mental lexicon (e.g., Bybee, 1998). In the latter case, the categorical absence of a phoneme results from the selection of the word pronunciation variant without that phoneme.

Evidence for lexical storage of categorically distinct pronunciation variants is provided by, among others, Bürki and Gaskell (2012). Bürki and Gaskell compared English words with schwa in poststress position (e.g., as in *mackerel*) to words with schwa in prestress position (e.g., as in *salami*). Previous studies (e.g., Kager 1997; Patterson, LoCasto & Connine, 2003) showed that reduced schwa variants of poststress schwa words (e.g., *mackerel* realized as [mækrəl]) are categorically distinct (word-medial schwa is either present or absent) from the corresponding schwa variants ([mækərəl]). In contrast, reduced schwa variants of prestress schwa words (e.g., *salami* realized as [slami]) show gradient schwa reduction. Bürki and Gaskell investigated the lexical representations underlying the production of categorically and gradient distinct English schwa words, by instructing participants to read aloud pseudohomophones (e.g., with pre-stress schwa: *b*/ə/*loone*; with post-stress schwa: *batt*/ə/*ree)* and pseudowords (e.g., with pre-stress schwa: *b*/ə/*leen*; with post-stress: *bott*/ə/*ree*) with and without schwa. The authors found shorter naming latencies for pseudohomophones than for pseudowords for both variants of poststress schwa words, but only for schwa variants of prestress schwa words. Based on these results, they assumed that both variants of poststress schwa words are lexically represented, while only schwa variants of prestress schwa words are.

Several methods have been used to determine whether phoneme absence only results from gradient or also from categorical processes. First, the phonetic properties of the phoneme sequences resulting from the absence of a phoneme have been studied in detail. Phoneme absence can be assumed to result from a categorical process when the phoneme is completely absent and has not left any traces in the acoustic signal. In contrast, the absence of a phoneme results from gradient reduction when the phoneme has left traces, for instance, in the durational or spectral properties of surrounding phonemes. In their analyses of French vowel reduction, Bürki and colleagues (2011b), Meunier and Espesser (2011), Adda-Decker and colleagues (2008), and Barnes and Kavitskaya (2002) observed residues of absent vowels in lip rounding and in durational and spectral characteristics of the acoustic signal.

A second method for attributing the absence of a phoneme to only gradient or also categorical processes is the inspection of the distribution of the duration of the phoneme (e.g., Bürki, Fougeron & Gendrot, 2007). If the distribution shows two clearly distinct peaks, one of which is around zero ms, representing the absent phoneme, absence of the phoneme is assumed to result from both a categorical and a gradient reduction process. If, in contrast, the distribution is unimodal, without a clear peak around zero ms, absence of the phoneme is assumed to result from gradient shortening. For the duration of French word-initial [s(e)] in *c'était* [sete] 'it was', Torreira and Ernestus (2011) found a bimodal distribution, with one mode containing most tokens with [e] (38.0% of the tokens, mean duration: 37 ms) and the other containing most tokens without [e] (62.0% of the tokens in which [e] had a duration of zero ms). This pattern suggests that the absence of [e] in French *c'était* results from a categorical process rather than from extreme vowel shortening. Bürki and colleagues (2011b) showed that the distribution of the duration of French schwa in word-initial syllables (e.g., [ə] in [pəluz] *pelouse* 'lawn') is close to normal, and that schwa duration varies from very short to very long. They therefore concluded that the absence of schwa, which is traditionally assumed to result from a categorical process of alternation, may especially be the result of a process of gradual phonetic reduction.

A third approach addressing the issue of gradient versus categorical reduction is to compare the variables predicting phoneme absence and those predicting phoneme duration. Perfect (or substantial) overlap in predictors would suggest that phoneme absence and duration are driven by the same processes, and that phoneme absence is the endpoint of gradient sound shortening (possibly as the result of co-articulation). In contrast, when phoneme absence and phoneme duration are affected by different variables, they are more likely to result from different processes, and the absence of a phoneme may then also result from a categorical process. In their study on Dutch informal speech, Hanique, Ernestus, and Schuppler (2013) found approximately the same phonetic variables to predict presence and duration of suffixal [t] in past participles (e.g., [t] in *gemaakt* 'made'), indicating that the absence of [t] may simply be the extreme result of shortening. Torreira and Ernestus (2011) found evidence for categorical absence; they observed that the presence and duration of [e] in French *c'était* 'it was' are conditioned by different variables, which suggests that presence and duration of [e] are determined by different processes, as was also suggested by the distribution of the duration (the second method).

In order to investigate the nature of the processes that underlie the reduction of the word-final OLS phonemes in French nouns, we applied all three methods to see whether their absence was mostly due to gradient or also to categorical reduction processes. Comparison of the results obtained with the three methods may not only provide more information about the underlying processes, but may also show what the best method for analysing reduction in corpus data is.

This chapter is structured as follows. We first provide a brief description of the procedure used to create our dataset. We then present our results in three parts. The first part of the results section discusses the different pronunciation variants of the OLS clusters that we observed at the phonemic level. In the second part, we focus on the phonetic properties of the OLS phonemes. In the third part, we present the distributions of the durations of the OLS phonemes and our statistical analyses of absence and duration of the OLS phonemes. In the general discussion, we summarize our results and discuss the implications of our findings for speech production models.

2.2 Methods

2.2.1 Materials

We extracted noun tokens from the Nijmegen Corpus of Casual French (NCCFr; Torreira, Adda-Decker & Ernestus, 2010). The NCCFr contains speech from 46 speakers (24 females, 22 males), who have completed the secondary education cycle in France and have been raised in central/northern France. Except for two female speakers, all speakers are university students, aged between 18 and 27 years. The speakers participated in pairs, coming to the laboratory with a third friend, who was a confederate. After adjusting the recording volume during the first two minutes of the conversation, the confederate was asked to come out of the sound attenuated room under the pretext that his/her microphone was not working properly. The two friends left in the room then spontaneously started a conversation. At this moment, the recording started. In the second part of the recording, the confederate was asked to go back to his/her friends and the three continued discussing topics raised in the first part of the recording. In the third part, the participants received a sheet of paper with several questions about politics and society. They were asked to negotiate unique answers for at least five questions. The corpus contains around 90 minutes of recorded conversation for every pair of speakers.

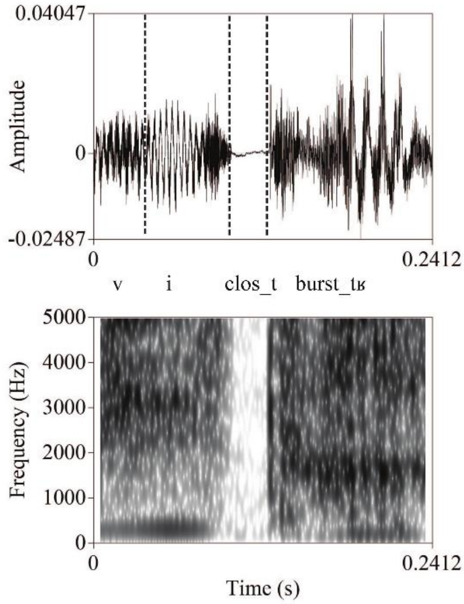
We selected noun tokens with citation forms ending in a voiceless obstruent-liquid-schwa (OLS) cluster. We restricted the analysis to clusters with voiceless obstruents because the acoustic analysis of these clusters is much easier than the acoustic analysis of OLS clusters containing voiced obstruents: the presence of an unreleased stop can be more easily detected if its closure is produced without vocal fold vibration. After we had excluded tokens that occurred in phrases containing laughter or overlapping speech or that occurred in fixed expressions (e.g., *par exemple* ‘for instance’), 291 noun tokens were left (212 singulars and 79 plurals, 41 lemmas; see Appendix). This selection of tokens includes one token ending in /flə/, seven in /fʁə/, 66 in /klə/, one in /kʁə/, 55 in /plə/, and 161 in /tʁə/. OLS clusters were preceded by one syllable in 147 word tokens and by two syllables in 144 tokens. They were produced by 43 different speakers (21 females, 22 males), aged between 18 and 50 years (mean age: 22.7).

2.2.2 Transcriptions

The first author annotated the acoustic signal of the OLS clusters of the 291 word tokens in Praat (Boersma & Weenink, 2014). Segment boundaries were placed to indicate the presence, onsets, and offsets of the obstruents, liquids, and schwas, based on the periodicity in and the acoustic intensity of the speech signal, as well as movements of the second formant in the spectrograms. Figure 2.1 gives an example of a phonetic transcription of [vitʁə] *vitre* ‘window glass’.

The onset of the oral closure was considered as the onset of the voiceless stop ([k], [t] and [p]). Due to the high-frequency fricative noise, the beginning of the [f] could also be distinguished easily. On the basis of auditory inspection, two transcribers determined whether the obstruent was voiced or voiceless.

Figure 2.1 Example of a phonetic transcription of /vitʁə/ *vitre* ‘window glass’. ‘Clo_t’ refers to the closure of the /t/ and ‘burst_tʁ’ refers to the burst of the /t/ and the /ʁ/.



The release of the voiceless stop was not included in the stop segmentation, because it could rarely be distinguished from the following liquid, which was often voiceless (81.8% of the liquids was voiceless). The segmentation of the stop thus only includes its closure. Since liquids have neither a (complete) closure nor a turbulent obstruction of airflow (e.g., Ball & Rahilly, 1999), segmentation of [l] and [ʁ] was also challenging after fricatives. The transcriber placed the boundaries between the fricatives and the liquids taking into account changes in the second formant and in the amplitude. Following previous literature, we also took into account the possibility of cluster-medial vowel epenthesis (e.g., /kʁy/ *cru* ‘raw’ may be pronounced as something like [gəʁy], at least in Québec French, see Colantoni & Steele, 2005, 2007).

We considered an increase of energy (and the start of periodic waveforms if the liquid was voiceless) as cue(s) for the onset of the schwa. Schwa offset was considered to be indicated by a decrease in intensity.

The first author further annotated every word token for variables that were significant predictors of the presence of phonemes in previous studies (e.g., Bell, Brenier, Gregory, Girand & Jurafsky, 2009; Gopal, 1990; Klatt, 1976): we annotated which phoneme followed and preceded each of the three phonemes, and specified its voicing, manner and place of articulation.

One hundred randomly selected tokens were also annotated by a second transcriber. No major disagreements were found with respect to the judgments on the presence versus absence of phonemes (95.0% agreement on presence of the obstruent and the schwa, 96.0% agreement on presence of the liquid). The mean between-transcriber difference in obstruent duration was 7.19 ms, in liquid duration 9.23 ms, and in schwa duration 10.36 ms. Comparisons of the annotations further showed an agreement of 96.0% for the type of phoneme following the word and of 93.0% for the type of phoneme preceding the word. The phonemes on which the transcribers did not agree with respect to presence or duration (disagreement being quantified as a duration difference > 20 ms) were transcribed by a third transcriber, and we adopted the majority vote.

2.2.3 Variables in our dataset

We added the frequency of occurrence of the given word, based on the Nijmegen Corpus of Casual French (NCCFr; Torreira, Adda-Decker & Ernestus, 2010), to our dataset. Since previous studies have shown that a word’s predictability influences its reduction and that of the preceding word (e.g., Pluymaekers et al., 2005a; Bell et al., 2003), we also incorporated the bigram frequencies of the word with the following word and with the preceding word, with bigram frequencies also being derived from the NCCFr. All these frequency variables were logarithmically transformed.

Furthermore, we took the position of the word token in the prosodic phrase into account: we annotated whether it was in phrase-medial or phrase-final position. We also included the number of syllables of the word in its unreduced form. Finally, we incorporated speech rate, which we defined as the number of syllables in the utterance's citation form divided by the duration of the utterance. We defined the utterance following Torreira and colleagues (2010), that is as a stretch of speech with clear syntactic and semantic coherence that contains no pauses. Pauses, which could be filled (e.g., by 'euh') or unfilled, were only annotated if they were minimally 200 ms, because this appears to be the standard threshold for perception (e.g., Zellner, 1994; Grosjean & Deschamps, 1975; Goldman-Eisler, 1968).

2.3. Results

2.3.1 Pronunciation variants

The majority of the OLS clusters (80.7%) lack at least one phoneme. Schwa is absent in 79.4% of all tokens, the liquid is absent in 42.9% of all tokens, and the obstruent is absent in 17.9% of all tokens. Our data show lower percentages of absence for the schwa, the liquid, and the obstruent than found in previous studies, which reported schwa absence rates of 90.0% for polysyllabic words (Adda-Decker, Boula de Mareüil & Lamel, 1999), liquid absence rates of 58.0% (for mono- and polysyllabic words, Villeneuve, 2010), and obstruent absence rates of 23.7% - 44.2%, depending on the following context (for mono- and polysyllabic words, Milne, 2014). Our absence rates are probably lower because our dataset only contains nouns, and does not contain high-frequency verbs (e.g., *être* 'to be') or conjunctions (e.g., *par exemple* 'for example'), which are known to show particularly high reduction rates.

If the phonemes of OLS clusters can be absent independently from each other, we would expect to find eight pronunciation variants. Our data only show seven pronunciation variants, of which three occur only rarely. Table 2.1 lists these seven variants and their relative frequencies.

Table 2.1 Relative frequencies (percentages of occurrence) of the OLS pronunciation variants, illustrated with the variants of the word *ministre* 'minister'. Below the dotted line, we present the three 'rare' variants.

variant type	obstruent	liquid	schwa	example	relative frequency
OLS	present	present	present	[ministʁə]	18.9%
OL	present	present	absent	[ministʁ]	35.7%
O	present	absent	absent	[minist]	27.1%
X	absent	absent	absent	[minis]	15.5%
LS	absent	present	present	[minisʁə]	1.4%
L	absent	present	absent	[minisʁ]	1.0%
OS	present	absent	present	[ministə]	0.3%

As mentioned above, the majority of word tokens are produced with reduced OLS clusters given that the citation form of the OLS cluster only occurs in 18.9% of the tokens. The variant without schwa has often been considered as the most frequent and thus canonical variant (e.g., Adda-Decker et al., 1999). This is also true for our dataset, where the most frequent variant of the OLS cluster is the OL variant. However, this variant only occurs in 35.7% of the tokens (out of which 21.7% had one syllable preceding the OLS cluster; and 14.1% had two preceding syllables), and the difference in frequency with the second most frequent variant is small. This second most frequent variant (27.1% of the tokens) is the variant in which both the liquid and schwa are absent. Strikingly, the variant in which the entire OLS cluster is absent also occurs quite often (15.5%)[3].

Table 2.1 also lists three low-frequency pronunciation variants. In the first of these, only the obstruent is absent (LS), in the second, both the obstruent and schwa are absent (L), and in the third, only the liquid is absent (OS). In the LS and L variants, for instance in [minisʁə] and [minisʁ], the obstruent is absent, but the liquid is still present. In the OS variant, for instance in [ministə], the liquid is absent, but [ə] is still present. These variants thus display the following pattern: a phoneme is absent while at least one of the following phonemes is present.

We see a different pattern for the frequent reduced variants, those above the dotted line in Table 2.1 (OL, O, X). If a phoneme is absent, all following phonemes are absent as well. This difference between the high-frequency and low-frequency variants suggests that a phoneme is more likely to be absent if the following phoneme is also absent. This pattern is confirmed by two logistic mixed-effects models (Faraway, 2006) (with word and speaker as random

---

[3] We tested whether the differences in frequency between the pronunciation variants are statistically significant. We conducted pairwise generalized mixed effects models with only the intercept as fixed effect and speaker and word as random effects for each pronunciation variant and the variant that was most similar in frequency. The model comparing the OL and O variants, the model comparing the O and OLS variants, and the model comparing the OLS and X variants showed no statistically significant difference.

variables) showing that the absence of the obstruent and the absence of the liquid can be predicted as a function of the absence of the following phoneme (*ß* = 2.69*, z* = 5.36*, p* < .001; *ß* = 4.50*, z* = 4.31*, p* < .001, respectively).

Figure 2.2 shows the percentages of tokens with a 'present' obstruent, the percentages of tokens with a 'present' liquid, and the percentages of tokens with a 'present' schwa, split by type of OLS cluster. Since we only studied one token of /flə/ and of /kʁə/, we did not incorporate these clusters in the figure.

Figure 2.2 The presence of obstruents, liquids and schwas in different OLS clusters.

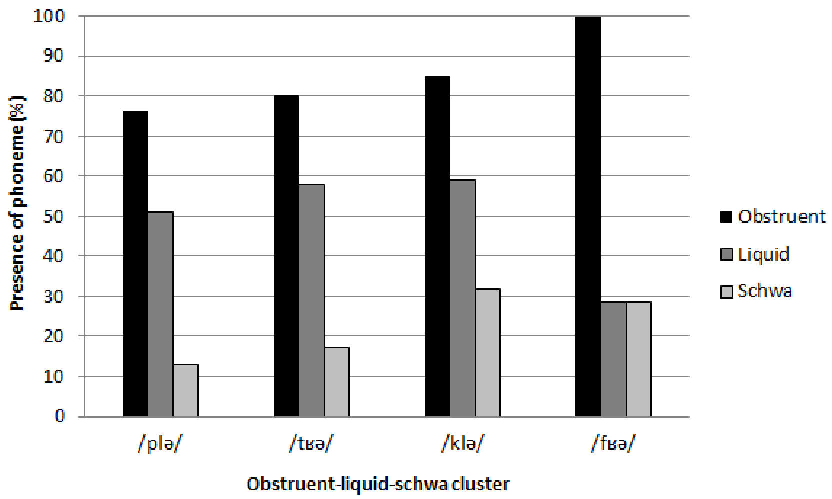


Figure 2.2 shows that, across OLS cluster types, schwa is absent most of the time. Pearson Chi-squared tests show that schwa is more often present in the /klə/ cluster than in the /plə/ and /tʁə/ clusters (differences with /plə/ and /tʁə/, both *ps* < .05, which are also significant after Bonferroni correction). Furthermore, in /plə/ and /tʁə/, the liquid is more often present than the schwa (also significant after Bonferroni correction: for /plə/ *p* < .01; for /tʁə/ *p* < .001;

whereas for /klə/ $p > .05$; and for /fʁə/ $p > .1$).

Regarding following-context effects, the whole cluster can be expected to be absent especially if the initial phoneme of the following word has the same place of articulation as the O, and that LS is especially absent if the following phoneme shares its place of articulation with the L. For instance, in the word sequence *peintre dans* 'painter in', the [t] has the same place of articulation as the initial [d] of *dans* and this may increase the likelihood that [tʁe] is absent. Our data show that only 35.6% of the word tokens with a completely absent OLS cluster are followed by a word that starts with an obstruent with the same place of articulation as the O of the OLS cluster[4], and that in only 41.2% of the tokens in which the liquid and the schwa are absent, the following word starts with a liquid. This shows that which phonemes in the OLS cluster may be absent is not only determined by the type of the initial segment of the following word, although this segment may have an influence.

Finally, the number of syllables preceding the OLS cluster does not seem to influence the likelihood of the absence of the word-final schwa (one versus two syllables preceding the OLS cluster: 78.9% versus 79.9%). In contrast, the two other cluster phonemes seem to be absent more often if the cluster is preceded by two syllables (obstruents: 26.4% versus 9.5%, Pearson Chi-squared test demonstrates $p < .01$; liquids: 49.3% versus 36.7%, but Pearson Chi-squared test demonstrates $p > .05$).

[4] We believe that the obstruent is missing rather than being merged with the following obstruent because if the two had merged we would have expected the following segment to be longer. That is, we then predict the following obstruent to be longer when the OLS obstruent seems absent than when the OLS obstruent is clearly present. This is not what we find. The mean duration of the following [d] is 26.5 ms when the obstruent is absent (N = 15), and 29.0 ms when the obstruent is present (N = 3). We cannot say anything about the absent [k] since absent [k] only once showed overlap with the place of articulation of the following segment.

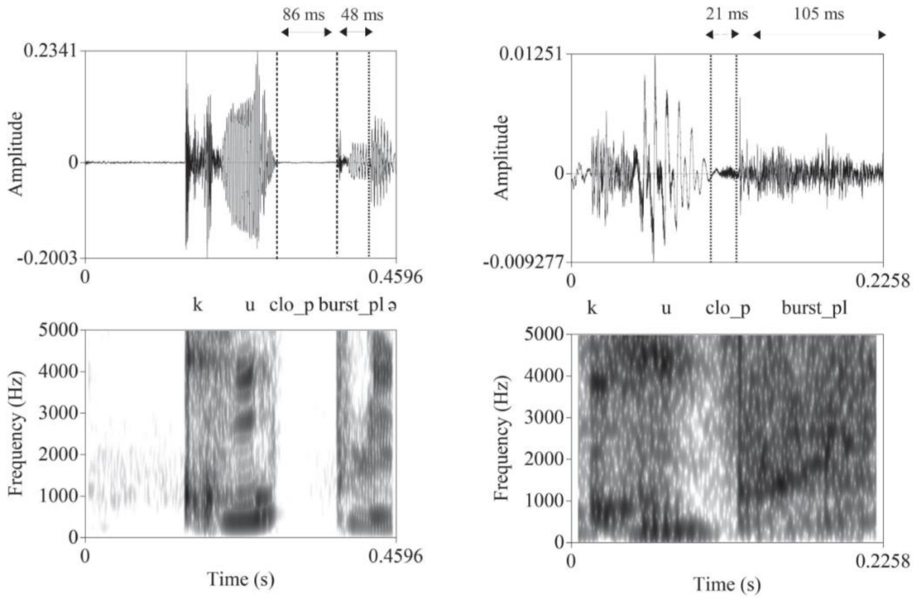
2.3.2. Phonetic properties

In this section, we discuss the phonetic properties of the different pronunciation variants of the OLS clusters in more detail.

2.3.2.1 Reduction of the stop closure

We first focus on the reduction of the stop closure. Figure 2.3 presents two pronunciation variants of *couple* /kuplə/ 'couple'.

Figure 2.3 Two pronunciation variants of *couple* /kuplə/ 'couple'. In the variant on the right, which is followed by *pour*, the burst of the stop and the liquid together (burst_pl) are unexpectedly long compared to the duration of the stop's closure (clo_p) and to the duration of the preceding vowel. This is not the case in the variant on the left, which is followed by *depuis*.



In the variant on the left, which is followed by *depuis*, the final schwa is present. The relatively long closure of the [p] (i.e., 86 ms) is followed by the burst of the [p] and the [l], which together have a duration of 48 ms. In the variant on the right, which is followed by *pour*, the schwa is absent. The relatively short closure of the [p] (i.e., 21 ms) is followed by the burst of the [p] and the [l], which together have a duration of 105 ms. Data like these suggest that if the closure of the [p] is shorter, the total duration of the burst of the [p] and the [l] is longer. The shorter stop closure may be the result of a (partly) incomplete closure and thus of the weakening of articulatory gestures.

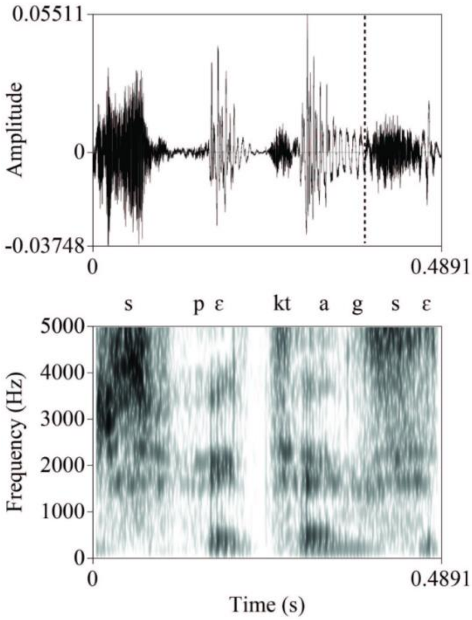
This duration pattern is reflected by a logistic mixed-effects model with word and speaker as random variables (Faraway, 2006), which shows that, for the whole data set, the duration of the stop closure can be predicted as a function of the duration of the burst and liquid: the stop closure is shorter if the burst and liquid are longer ($ß = -0.136$, $t = -2.858$, $p < .01$).

2.3.2.2 Obstruent voicing

While the obstruent is voiceless in the citation forms of the words, it is often realized as voiced in our data. If the liquid and schwa are both absent and the obstruent is in word-final position, the obstruent is realized as voiced in 13.9% of the tokens (i.e., in 11 of the 79 word tokens). The question arises about the source of this voicing. In 18.2% of the tokens (i.e., in 2 of the 11 tokens) in which the OLS cluster is pronounced as the voiced variant of O, the word token was followed by a voiceless consonant. The voicing of the obstruent did thus not exclusively result from the voicing of the following *present* segment (i.e., regressive voice assimilation to the following *present* segment). In Figure 2.4, we see an example of the noun token *spectacle* 'show' in which the OLS cluster /klə/ is pronounced as [g] and the next word *c'est* starts with a voiceless [s].

Progressive voice assimilation, where a phoneme takes on the voicing properties of the preceding one, is not very common in French (e.g., Lodge, Armstrong, Ellis & Shelton, 1997) and, to our knowledge, progressive voice assimilation to French vowels has not been reported. Moreover, all voiced obstruents were in word-final position ([dʁ], [dʁə], [bl], [blə], [gl], [glə], [vʁ] and [vʁə] do not occur), which suggests that the voicing of obstruents does not result from assimilation to the voicing properties of the vowels preceding the obstruents. This pattern of results also confirms our hypothesis formulated above that the voicing does not come from the following present sounds. Otherwise, the obstruents would also have been voiced if followed by the liquids or the schwa.

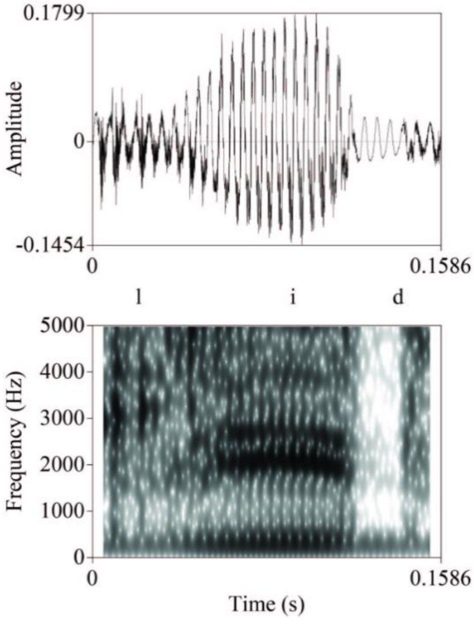
Figure 2.4 Token of *spectacle* 'show' in which the final OLS cluster is pronounced as the voiced variant of the obstruent. The next word starts with a voiceless /s/.



We propose that the voicing of the obstruents originates from the *absent* liquids. Since we showed that the obstruent is neither likely to be voiced as a result of progressive assimilation nor as a result of regressive assimilation to the surrounding *present* phonemes, obstruents, like [k] in Figure 2.4, are likely to be voiced as a result of co-articulation with the voicing properties of the segmentally absent liquid.

Figure 2.5 presents another token of a noun (*litres* 'litres') in which the liquid and the schwa are absent and the obstruent is voiced. This noun token was followed by a nasal from the following word (*mais* [mɛ] 'but'), and the /t/ may therefore be voiced as a consequence of regressive voice assimilation to this nasal as well. However, Hallé and Adda-Decker (2007) pointed out that nasals rarely induce voice assimilation processes in French, and we therefore believe that the absent liquid may have played a role in the voicing of the obstruent in the four O tokens followed by nasals.

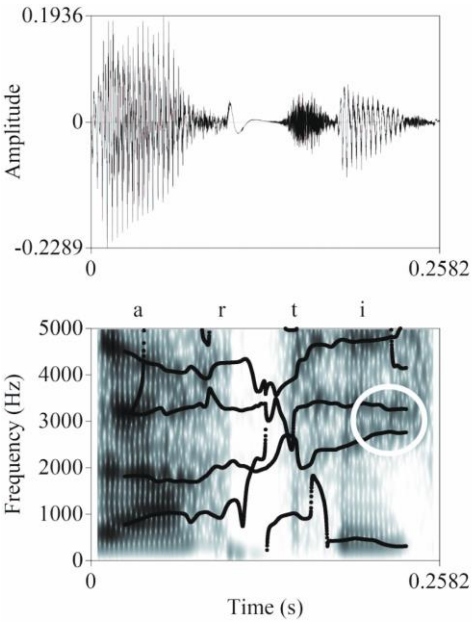
Figure 2.5 A token of *litres* /litʁə/ 'litres' pronounced as /lid/. The noun is followed by the /m/ of *mais*.



2.3.2.3 Acoustic residues of velar obstruents

Figure 2.6 shows an example of a pronunciation variant of *articles* 'articles' followed by *par* in which the [k] surfaces as a velar pinch in the vowel formant structure: during the preceding vowel, the second and third formants come very close to each other, which is typical for a velar context. We observed a similar pattern in 62.5% of the tokens (i.e., for 5 of the 8 tokens) in which /klə/ seems completely absent.

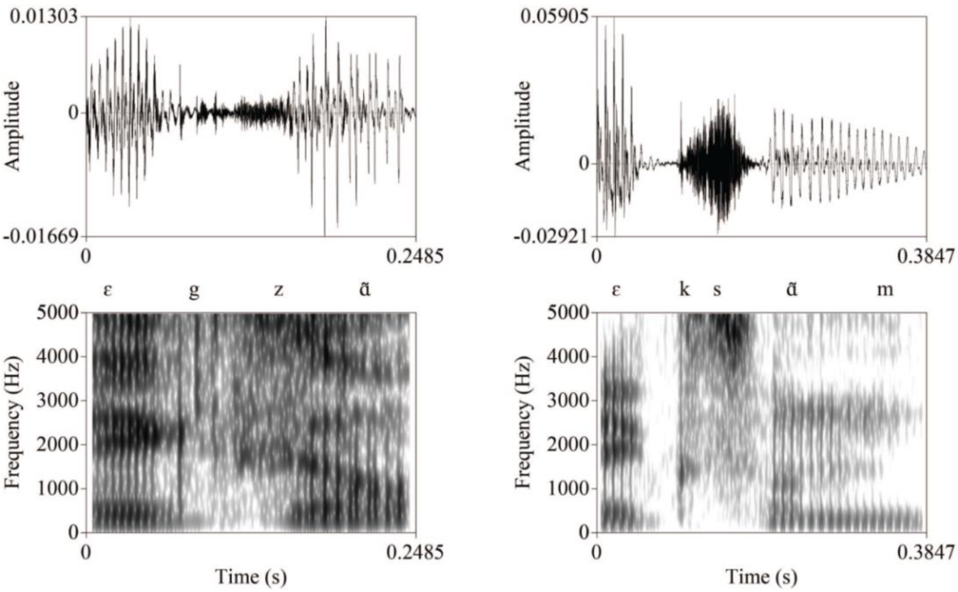
Figure 2.6 Example of a pronunciation variant of *articles* 'articles' in which the OLS cluster is completely absent. The variant is followed by *par*. The black lines in the spectrogram represent the formants. In the white circle, during the /i/, the second and third formants come very close to each other, which is typical for a velar context.



2.3.2.4 Obstruent nasalisation

Twenty-one word tokens in our dataset have a nasalised vowel preceding the OLS cluster (e.g., *exemple* /ɛgzɑ̃plə/ 'example'). In 71.4% (15 out of 21) of these tokens (seven were pronounced by the same speaker), the liquid and schwa were absent and the obstruent was pronounced as [n] or [m] (e.g., /ɛgzɑ̃plə/ was pronounced as [ɛgzɑ̃m]). The speaker kept the velum lowered, needed for the nasalised vowel, during the realisation of the obstruent, as a result of which the consonant was nasalised[5]. Figure 2.7 presents two pronunciation variants of the word *example* ‘example’; one pronounced without the OLS cluster (and thus ending in the nasal vowel) in phrase-final position and one with a nasalised [p], which is followed by *du*. This phenomenon of obstruent nasalisation was also observed by Malécot and Metz (1972) in the context ‘nasal vowel + stop + word juncture + consonant’.

Figure 2.7 Examples of pronunciations of *exemple* /ɛgzɑ̃plə/ ‘example’. In the phrase-final variant on the left, the whole OLS cluster is absent. In the variant on the right, which is followed by *du*, the obstruent has become nasalised.

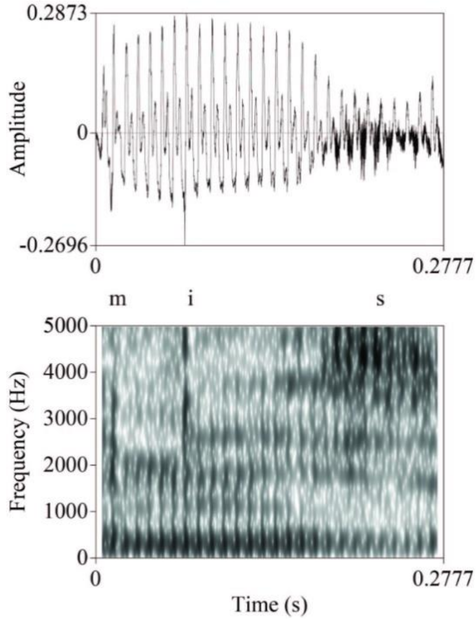


[5] Note that European French differs from North American varieties (i.e., Canada and Louisiana) in that in these former varieties, nasal stops do not follow nasal vowels in formal speech.

2.3.2.5 Reduction of phonemes in addition to the OLS cluster

Most of the 45 word tokens without the OLS cluster also lack other phonemes. For instance, *ministre* 'minister' is realized as [mis] in 52.9% of the word tokens (in 9 out of 17 tokens, see Figure 2.8 for an example). In [mis], the whole OLS cluster is missing as well as the phonemes /ni/, and this realisation thus lacks two complete syllables. The final segment of the token in Figure 2.8, which is followed by *de*, is relatively long ([s] is 97.97 ms), such that reduction of /tʁə/ in *ministre* probably involves lengthening of word-final [s] (see section 2.4.2).

Figure 2.8 Example of a pronunciation variant of *ministre* /ministʁə/ 'minister', which is followed by *de*, and lacks phonemes in addition to the OLS cluster phonemes.



2.3.2.6 Summary

To summarize, the phonetic analyses show that absent or reduced obstruents and liquids may leave durational and phonetic cues on neighbouring segments. If a stop has a short closure, its burst and following liquid are often lengthened in a compensatory fashion, and if the OLS cluster is completely missing, the preceding phoneme may be lengthened. An absent liquid may leave traces in the form of the voicing of the preceding obstruent. The results further demonstrate that the obstruent can become nasalised after a nasalised vowel. We did not find acoustic traces of segmentally absent schwa.

2.3.3 Analyses of the duration and absence of each OLS phoneme

This section presents the distributions of the durations of the individual OLS phonemes. In addition, we analyze the predictors of absence and durations of these phonemes. These data will provide further evidence for the categorical versus gradient nature of the processes underlying the phonemes' absence.

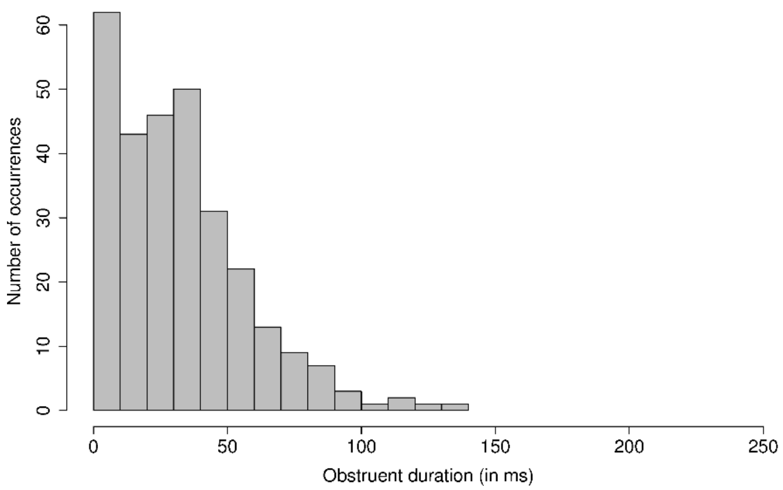
2.3.3.1 The obstruent

Figure 2.9 shows the distribution of the duration of the obstruent. Absent obstruents are represented as tokens with zero ms and the duration of the obstruent thus ranges between zero ms and 136 ms. The distribution shows two peaks (one peak at zero ms, 17.9% of the noun tokens; one peak around 35 ms) with a very shallow dip between them. Because of the shallowness of this dip, we consider the distribution of the duration of the obstruent rather as unimodal than bimodal.

To determine the predictors of absence of the obstruent, we fitted generalized mixed-effects regression models with the binomial link function, with word and speaker as crossed random variables, and with the following types of potential fixed predictors: 1. two boolean variables (final position and medial position) representing the phrasal position of the word token, 2. word frequency based on the NCCFr (log-transformed), 3. the bigram frequencies of the word with the following and with the preceding word (again based on the

NCCFr, and log-transformed), 4. speech rate, quantified as the number of syllables, and thus the number of vowels, realized in the utterance's citation form, divided by the duration of the utterance, 5. the number of syllables in the word, and 6. the nature of the preceding phoneme. Since we observed that a phoneme was more likely to be absent when the following phoneme was also absent, we also incorporated the absence of the liquid as a predictor of obstruent absence. We did not include random slopes because preliminary testing suggested that models including them did not converge. We only retained those predictors in the model that were statistically significant or figured in statistically significant interactions and whose presence resulted in a model with a lower AIC.

Figure 2.9 Distribution of the duration of the obstruent (N = 291).



For the analysis of the log-transformed duration of the obstruent (excluding the burst phase of the stop), we fitted linear mixed-effects regression models following the same fitting procedure and with the same potential predictors as in the analysis of the obstruent absence. We did not include in our analysis the absent obstruents (52 data points, 17.9%). Finally, we discarded all obstruents with durations that deviated more than 2.5 times the standard error from the values predicted by the best statistical model and refitted the model.

Table 2.2 summarizes the final models for obstruent absence and obstruent duration.

Table 2.2 Summary of the regression models predicting obstruent absence and duration. The intercept represents a word token with a present liquid that is in phrase-final position.

Fixed effects		absence		duration	
		ß	*z*	*ß*	*t*
Intercept		-41.55	-3.67***	3.94	22.77
Liquid absent		14.45	3.76 ***	0.26	3.60***
Medial position		6.11	2.29*	-0.15	-2.08*
Following bigram frequency		3.43	1.99*	-	-
Number of syllables		-	-	-0.28	-2.49*
Random effects			*SD*		*SD*
Word	intercept		26.20	intercept	0.27
Speaker	intercept		11.91	intercept	0.16
				residual	0.46

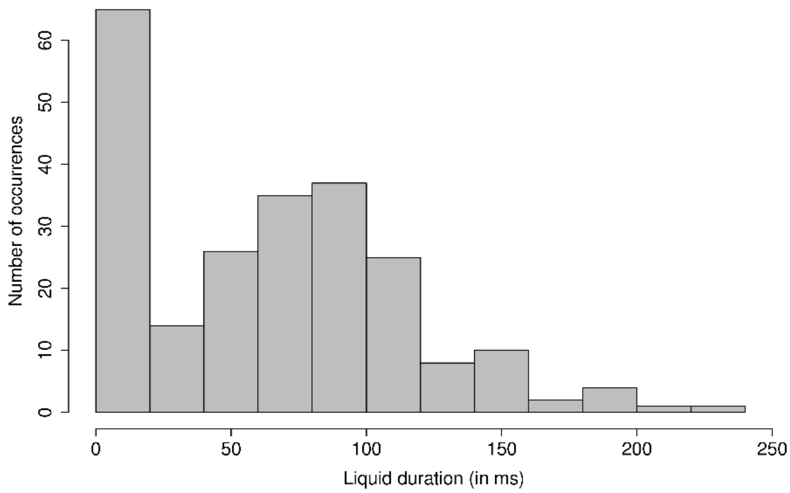
*Note: *** indicates p < .001 and * indicates p < .05*

The two models have two predictors in common (i.e., liquid absence and phrasal position). If the liquid was absent, the obstruent was more likely to be absent as well and tended to be longer. If the word was in medial position, the obstruent was more often absent and tended to be shorter. The model for obstruent absence also shows that the obstruent was more often absent if the following word was more predictable. Finally, the obstruent was shorter in trisyllabic words than in bisyllabic words. Except for the effect of liquid absence on obstruent duration, these effects are in line with previously reported effects on reduction in general (e.g., Torreira & Ernestus, 2009). The longer duration of the obstruent in case the liquid is absent shows that absent liquids may leave durational cues and that 'compensatory' lengthening may take place.

2.3.3.2 The liquid

The distribution of the duration of the liquid is presented in Figure 2.10. Here again, absent segments were assigned durations of zero ms.

Figure 2.10 Distribution of the duration of the liquid (N = 291).



The duration thus ranged from zero ms to 223 ms. The distribution shows two clear peaks (one at zero, 42.9% of the noun tokens; one at around 90 ms) with a deep dip between them.

We analyzed the predictors of liquid absence and duration, including the burst of the preceding stops if present, following the same procedures as for obstruent absence and duration. We also tested the same predictors, except for the predictor of liquid absence which was replaced by the predictor of schwa absence. Liquid durations of zero ms (125 data points, 42.9%) were removed for the duration analysis. We also discarded all liquids with durations that deviated more than 2.5 times the standard error from the values predicted by the best statistical model and refitted the model. Table 2.3 summarizes the best statistical regression models.

Table 2.3 Summary of the regression models predicting liquid absence and duration. The intercept represents a liquid followed by a schwa for the analysis of the absence of the liquid and a liquid in a word in phrase-final position for the duration analysis.

Fixed effects	absence		duration	
	ß	*z*	*ß*	*t*
Intercept	-9.31	-4.51***	5.12	15.85
Speech rate	2.21	2.37*	-0.38	-2.02*
Schwa absent	4.43	4.19***	-	-
Following bigram frequency	0.89	2.65**	-	-
Medial position	-	-	-0.15	-2.47*
Random effects		*SD*		*SD*
Word	intercept	0.39	intercept	0.26
Speaker	intercept	0.59	intercept	0.11
			residual	0.34

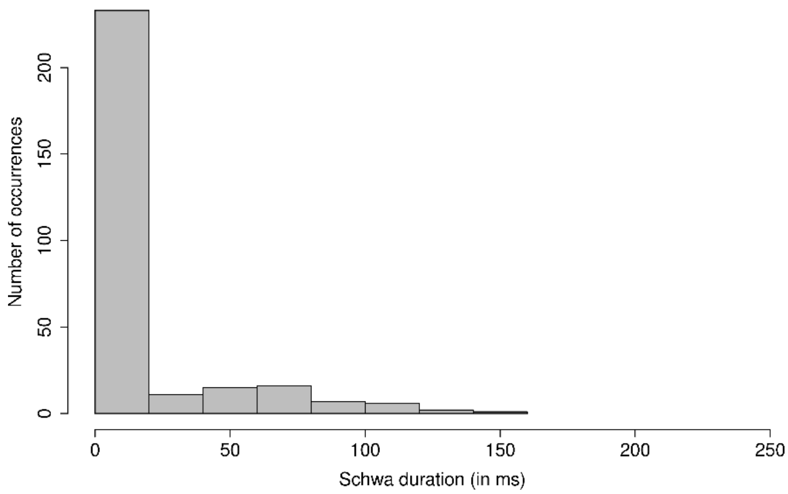
*Note: *** indicates p < .001, ** indicates p < .01 and * indicates p < .05*

Table 2.3 shows that there is only one variable predicting both liquid absence and liquid duration. If speech rate was higher, the liquid was shorter and more often absent. The liquid was also more likely to be absent if schwa was absent and if the following word was more predictable. The liquid was shorter in phrase-medial position. Similar effects have been found for other reduction patterns (e.g., Keating, Cho, Fougeron & Hsu, 2003; Pierrehumbert & Talkin, 1992).

2.3.3.3 The schwa

The distribution of the duration of schwa is presented in Figure 2.11. Schwa duration ranged from zero ms to 160 ms. This figure shows one clear peak at zero ms (79.4% of the data).

Figure 2.11 Distribution of the duration of schwa (N = 291).



We analyzed schwa absence and duration following the same procedure as for obstruent and liquid absence and duration and tested the same potential predictors, except for those of liquid and schwa absence. These latter predictors were replaced by the boolean predictor indicating whether the following segment was a consonant. Schwa durations of zero ms (231 data points, 79.4%) were removed for the duration analysis. Finally, we discarded all schwas with durations that deviated more than 2.5 times the standard error from the values predicted by the best statistical model and refitted the model. Probably due to the small number of remaining data points (60), we did not find significant predictors for schwa duration. The predictors for schwa absence are listed in Table 2.4.

Table 2.4 Summary of the linear mixed-effects model predicting schwa absence. The intercept represents a schwa followed by a vowel.

Fixed effects		absence	
		ß	*z*
Intercept		-4.48	-2.19*
Speech rate		3.26	3.02 **
Word frequency		0.37	2.38*
Following consonant		-1.34	-3.11**
Random effects			*SD*
Word	intercept		0.00053
Speaker	intercept		1.01

*Note: ** indicates p < .01 and * indicates p < .05*

Schwa was significantly more often absent if speech rate was higher and if the word was more frequent. In contrast, schwa was more often present if it was followed by a consonant than by a vowel. The same effects have been found in previous studies (e.g., for schwa in word-final syllable, see Milne, 2014; for schwa in word-initial syllable, see Bürki et al., 2011a; Bürki, Ernestus & Frauenfelder, 2010).

2.4. General discussion

2.4.1 Segmental variation in OLS clusters

This study investigated reduction in French nouns ending in obstruent-liquid-schwa (OLS) clusters. In contrast to most previous studies on reduction, our study focused on the reduction of content words, on the reduction of a *sequence* of phonemes, and on the presence versus absence of phonemes as well as phonetic traces of absent sounds. Our analyses were based on 291 noun tokens from the Nijmegen Corpus of Casual French (NCCFr, Torreira, Adda-Decker & Ernestus, 2010).

Our data show that few word tokens were pronounced with all segments that could be present. The word *ministre* 'minister', for instance, had pronunciation variants ranging from full [ministʁə] to highly reduced [mis]. Reduction thus frequently affects (low frequency) content words. We can distinguish four main pronunciation variants for the OLS cluster at the segmental level. The unreduced form (e.g., *ministre* [ministʁə] 'minister') occurs in 18.9% of the tokens. The variant in which schwa is absent (e.g., [ministʁ]) is generally considered as the most frequent variant (e.g., Adda-Decker, Boula de Mareüil & Lamel, 1999). In our dataset the most frequent variant of the OLS cluster is the OL variant as well, although it only occurs in 35.7% of cases. Following the variant without schwa, the variant in which both the liquid and schwa are absent (e.g., [minist] for /ministʁə/) is the most frequent one (27.1%). Strikingly, the variant in which the whole OLS cluster is absent (e.g., [minis]) also occurs quite often (15.5%). Reduction thus often affects sequences of phonemes like LS and OLS.

In addition to these four main categories, we found three categories represented by only a few tokens: LS (e.g., [minisʁə], 1.4%), L (e.g., [minisʁ], 1.0%) and OS (e.g., [ministə], 0.3%). Our finding that the OL (e.g., [ministʁ]), the O (e.g., [minist]) and the X (e.g., [minis]) variants are substantially more frequent than the LS (e.g., [minisʁə]), the L (e.g., [minisʁ]) and the OS (e.g., [ministə]) variants suggests that a phoneme is more likely to be absent if the following phoneme is also absent.

One possible explanation for the observed pattern is that phonemes are more likely to be absent if they are not followed by other phonemes in the word. This might be because final phonemes are likely to be co-articulated with the following word-initial phonemes, which are typically strongly articulated (e.g., Keating, Cho, Fougeron & Hsu, 2003). This strong co-articulation may lead to the segmental absence of these final phonemes. Alternatively, the pattern may arise simply because schwas are more likely to be absent than liquids and liquids are more likely to be absent than obstruents (e.g., Adda-Decker et al., 2008; Su & Basset, 1998). This hierarchy may be related to the ease with which these phonemes are recovered. Possibly, both tendencies contribute to the emergence of the reduction pattern.

2.4.2 Residual acoustic cues of absent segments

In 13.9% of the OLS tokens in which the liquid and schwa were absent (and the cluster was thus realized as O), the obstruent was voiced, which shows that the voicing did not always result from regressive voice assimilation. We also observed obstruent voicing before words starting with voiceless consonants. In French, regressive assimilation is more common than progressive assimilation (e.g., Lodge, Armstrong, Ellis & Shelton, 1997), and to our knowledge, progressive voice assimilation of a voiceless obstruent to a preceding vowel has not been reported so far. Further, the obstruent is never voiced if the liquid is present. These findings support the idea that obstruent voicing in this context has resulted from the absence of the liquid, at least in some tokens. As such, the obstruent voicing is a signature of this liquid. We assume that the liquid is then co-articulated with the preceding obstruent (as a consequence of the retiming of articulatory gestures).

Segmentally absent phonemes also left durational traces. If the obstruent and/or liquid were/was absent, the preceding phoneme was often longer (e.g., [s] in [mis]). Zimmerer, Scharinger, and Reetz (2011) found similar results for German [t]. Both in French and German, the absence of a phoneme can thus be cued by the lengthening of a neighbouring phoneme.

There are two possible causes for this 'compensatory' lengthening. First, the reduced version of a segment (e.g., /t/) may be indistinguishable from the preceding segment (e.g., /s/). The reduced segment may then seem to be completely absent while the preceding segment seems to be lengthened. Articulatorily, however, both segments are present and the duration of the preceding segment is independent of the reduction of the target segment. This may be the right account of the lengthening of [s] in German /st/ clusters pronounced without [t] observed by Zimmerer, Scharinger and Reetz (2011), and the extraordinary long duration of [s] in [mis], which is the reduced form of *ministre* 'minister' observed in the present study.

A second possible cause for 'compensatory' lengthening accompanying segment reduction in French may relate to the rhythmic structure of the language (syllable-timed; e.g., Colantoni & Steele, 2011). Speakers may tend to lengthen consonants when the following consonants are absent in order to maintain the syllable timed structure of the phrase. For instance, speakers may tend to lengthen the /t/ of [lit] /litʁə/ *litres* ‘liters’ after reduction of the liquid following the /t/ in order to preserve the duration of the word.

Furthermore, when /k/ in the OLS cluster (e.g., /k/ in /spɛktaklə/ *spectacle* ‘show’) was absent, the spectrogram could still reveal its phonetic remnants, in the form of a velar pinch (i.e., second and third formants of the preceding vowel come very close to each other). We observed a velar pinch in 62.5% of the tokens with segmentally absent /klə/. This velar pinch provides clear evidence that /k/ was not categorically absent and that its absence was due to extreme weakening of the articulatory gestures.

Our results further demonstrate that in 71.4% of the tokens with a nasal vowel, the following stop was nasalised. For instance, the word *exemple* /ɛgzɑ̃plə/ ‘example’ was pronounced as [ɛgzɑ̃m]. This may be unexpected because the common assumption about European French is that, except for prefixed words, nasal consonants do not occur after nasal vowels (e.g., Walker, 2001: 64). We believe that the consonant is nasalised as a result of co-articulation: the speaker can keep the velum lowered, a requirement for the articulation of the nasal vowel, as a result of which the following obstruent (e.g., /p/) is nasalised (resulting in e.g., [m] replacing /p/).

2.4.3 Categorical versus gradient reduction processes

Obstruent voicing, ‘compensatory’ lengthening, velar pinches, and nasalised obstruents in tokens with nasal vowels suggest that reduction of the obstruent and the liquid in OLS clusters may be gradient and that absence of these phonemes may be the result of extreme co-articulation (in combination with weakening). The phonetic analyses cannot give information about whether the absence of the OLS phonemes can also be the result of categorical reduction processes. If such processes play a role as well, the absence of the OLS phonemes can be complete, without leaving acoustic traces.

To further investigate the nature of the reduction processes underlying the absence of the OLS phonemes, we made histograms of the durations of the three phonemes in the cluster. If the duration distribution shows two clearly distinct peaks, of which one occurs at zero ms, representing the absent phoneme tokens, absence of the phoneme is likely to result from both a categorical process and a gradient shortening process (e.g., due to co-articulation). If, in contrast, the duration distributions show a unimodal distribution, without a clear peak around zero ms, absence of the phoneme only results from extreme shortening (e.g., Torreira & Ernestus, 2011; Bürki et al., 2011b; Bürki, Fougeron & Gendrot, 2007).

The histograms showed no clear bimodal distribution for the duration of the obstruent, but they did for the durations of the liquid and the schwa. The distributions of the durations of the obstruent and the liquid in OLS clusters have never been studied before, such that our results cannot be compared to earlier findings. Our result for schwa differs from the findings on schwa duration by Bürki and colleagues (2011b), who reported a unimodal, normal distribution for schwa duration. This difference may be due to the position of [ə] in the word: we focused on schwa in word-final position, whereas Bürki and colleagues examined word-medial schwa. Our duration distributions suggest that, in contrast to obstruent reduction, liquid and schwa reduction may also be categorical. However, a note of caution is due here since the shallowness of the dip between the two peaks in the histogram of the obstruent makes this distribution difficult to interpret. It is not completely clear whether the first peak represents a distinct group of its own or not.

We also examined the predictors of phoneme absence and phoneme duration (only for phonemes that were present). Obstruents and liquids were more often absent in phrase-medial position. This finding corresponds to the results from a study by Ernestus and Smith (forthcoming), who reported that Dutch /l/ was more often absent in phrase-medial position (53.0%) than at phrase boundary (16.0%). Further, we found that the higher the probability of the following word, the more often obstruents and liquids were absent. This observation is in line with the results of, for instance, Schuppler, Van Dommelen, Koreman, and Ernestus (2012), who reported that constrictions, bursts and alveolar friction in Dutch /t/ are more likely to be absent in word combinations of higher frequencies.

In the duration analyses, phrasal position predicted both obstruent and liquid durations: liquids and obstruents were shorter in phrase-medial position. These findings are in line with previous results which reported shorter segments in phrase-medial position than in phrase-initial position (e.g., Keating, Cho, Fougeron & Hsu, 2003; Pierrehumbert & Talkin, 1992). Possibly, due to the small number of data points (60), we did not find significant predictors for schwa duration.

If a segment is absent mainly due to extreme shortening (e.g., as a result of co-articulation) and there is hardly any role for categorical reduction processes, its duration and absence should be conditioned by the same variables. Both obstruent absence and duration were predicted by phrasal position: the obstruent was shorter and more often absent in phrase-medial position than in phrase-final position. The absence of the following liquid also affected both the presence and the duration of the obstruent, but this cannot be interpreted as evidence for gradient reduction of the obstruent because when the liquid was absent, the obstruent tended to be either completely absent or relatively long.

In the liquid analyses, only speech rate (one out of four predictors that play a role in either the absence or duration analysis) predicted both liquid absence and duration: if speech rate was higher, the liquid was more often absent and shorter. The overlap of predictors in both the obstruent and liquid analyses is thus not substantial which suggests that the phoneme may be absent as a consequence of reduction processes that do not necessarily lead to sound shortening as well.

The three different methods (phonetic analyses, examination of the distribution of phoneme duration, and the comparison of the predictors of phoneme absence and phoneme duration) do not provide an unequivocal picture of the nature of the reduction processes underlying the absence of the obstruent, the liquid, and the schwa. Based on the phonetic analyses and the distributions of the durations, one may conclude that the reduction process underlying obstruent absence is above all gradient: obstruent absence often leaves phonetic remnants and obstruent duration shows no clear bimodal distribution. However, the comparison of the predictors of obstruent absence and obstruent duration suggests that obstruent absence and duration are mainly sensitive to different variables and therefore result from different processes, which would imply that the absence of the obstruent does often not result from gradient reduction. This inconsistent pattern of results may be because our dataset consisted of only 291 noun tokens. Possibly a larger dataset shows that similar variables predict both the absence and the duration of the obstruent. We can then conclude that the absence of the obstruent is the extreme result from sound shortening, from overlap in articulatory gestures, or articulatory “undershoot” (Lindblom, 1963).

The phonetic analyses suggest that an absent liquid may also leave acoustic traces, for instance, in the form of the voicing of the preceding obstruent. The distribution of liquid duration and the comparison of the predictors of liquid absence and liquid duration suggest that the reduction process underlying the absence of the liquid is often categorical: we saw a clear peak around zero ms in the histogram and no substantial overlap of predictors (one out of four predictors that play a role in either the absence or duration analysis). The absence of the liquid may thus result from both gradient and categorical processes.

The distribution of the duration of schwa shows a clear peak at zero ms, which suggests that the reduction process of schwa is categorical. In line with other positions of schwa in the word (e.g., Bürki et al., 2011a), we assume that in word-final position the absence of schwa may not only result from gradient shortening processes, but also from categorical processes.

2.4.4 Methodological and theoretical implications

These results raise questions about how to interpret the results from the three methods that we used for investigating the nature of the processes underlying the absence of phonemes. Phonetic analyses can only provide evidence for gradient reduction processes, which nearly always play a role. On the basis of this single method, we cannot determine whether also categorical processes are at play. With respect to distributions of phoneme durations, since there is no clear definition of what counts as a peak, the interpretation of these distributions may be subjective. Finally, the comparison of the predictors of absence and duration of phonemes may not lead to undisputable conclusions because it is not clear how much overlap in predictors is necessary, given the number of data points, to conclude that phoneme absence and phoneme duration are driven by the same processes. Overall, the current data highlight the importance of combining different methods in investigating the nature of processes underlying the absence of phonemes.

The question arises as to how existing speech production models are able to account for our findings. Our results suggest that, if we assume that a word can be lexically stored with more than one pronunciation variant and these variants are stored in the form of strings of phonemes (following, e.g., Bürki & Gaskell, 2012; Bürki, Ernestus & Frauenfelder, 2010), all pronunciation variants that are stored for a given word contain the obstruent of the OLS cluster, since the obstruent is likely to be absent only due to gradient processes. The liquid and schwa, in contrast, may also be absent due to categorical processes, and the mental lexicon may therefore contain word pronunciation variants with or without a liquid or schwa. For instance, *ministre* 'minister' may have the lexical representations /ministʁ/ and /minist/. Alternatively, the liquid and schwa may be categorically absent due to phonological deletion rules (e.g., Levelt, Roelofs & Meyer, 1999).

Finally, some researchers suggest that no word is lexically stored with a final schwa and that if this schwa surfaces, this is due to a categorical schwa insertion rule (e.g., Dell, 1985), which is also in line with our data.

Information about phrasal position, speech rate and phonetic context may be incorporated into the lexical representations or into the phonological deletion rules. In addition, the effects of these variables may arise during articulation, as they do for the obstruent. If so, future research should show that these variables especially affect production when the segment has left acoustic traces.

A different type of speech processing model assumes that all tokens of a word are lexically stored with all their fine phonetic detail (in the form of exemplars; e.g., Bybee, 1998; Goldinger, 1998). The effects of phrasal position, speech rate, and phonetic context may arise from the exact specifications of the exemplars or may arise during articulation.

Our results also raise questions about speech comprehension. The acoustic residues may form cues to the absent sounds, facilitating lexical access. The categorical absence of segments may ask for the use of multiple word pronunciations in the mental lexicon.

2.4.5 Conclusion

In conclusion, our study has shown that European French word-final obstruent-liquid-schwa clusters are often reduced. We observed that in 80.7% of the word tokens extracted from a corpus of casual speech, at least one phoneme was absent and that in no less than 15.5% the entire cluster was absent. Furthermore, the absence of a phoneme was predominantly determined by the presence or absence of the next phoneme. The OLS phonemes not only showed variation in being present or absent and in their durations but also in their subsegmental traces in neighbouring sounds. Finally, on the basis of the corpus data, we argued that whereas the obstruent is likely to be absent due to gradient processes (extreme shortening and co-articulation), the liquid and the schwa may also be absent due to other (categorical) processes. Future research on a larger corpus is needed to confirm the observed reduction patterns in European French word-final OLS clusters and to further compare the three methods for determining the nature of the processes underlying the reduction patterns. For now, we conclude that our results underline the relevance of corpus-based research in the documenting of pronunciation variation in everyday conversations and of the processes underlying this variation.

Appendix

Lemmas (number of occurrences)

ancêtre (1), *article* (17), *boucle* (1), *centre* (16), *cercle* (7), *chantre* (1), *chapitre* (3), *chiffre* (3), *coffre* (2), *couple* (16), *couvercle* (1), *diamètre* (2), *disciple* (1), *exemple* (27), *fenêtre* (12), *filtre* (1), *lettre* (30), *litre* (8), *maître* (2), *meurtre* (4), *ministre* (17), *montre* (2), *muscle* (5), *obstacle* (1), *offre* (2), *oncle* (9), *peintre* (1), *peuple* (11), *plâtre* (1), *prêtre* (1), *rencontre* (4), *semestre* (10), *siècle* (10), *souffle* (1), *spectacle* (15), *sucre* (1), *théâtre* (31), *titre* (4), *trimestre* (2), *ventre* (4), *vitre* (4).

The use of contextual cues and of variant frequency in the comprehension of reduced speech by highly proficient language learners

Chapter 3

This chapter is a reformatted version of:
Sophie Brand & Mirjam Ernestus (submitted). The use of contextual cues and of variant frequency in the comprehension of reduced speech by highly proficient language learners.

Abstract

We investigated how highly proficient learners of a language understand reduced word pronunciation variants, as well as the causes of potential problems. Two experiments (a dictation task and a lexical decision task) showed that Dutch learners of French have difficulties understanding reduction of French words. In the dictation task, representing a variety of reduction patterns, the learners produced an average word error rate of 51.7% (versus 12.0% for native listeners). The errors show that learners experienced word segmentation problems, were unable to reconstruct reduced phonemes, and did not rely much on contextual information. Many errors resulted in their completely misinterpreting the presented sentences. The lexical decision task with cross-modal identity priming, testing the comprehension of words ending in an obstruent, liquid and schwa (e.g. *ministre* 'minister'), showed that, nevertheless, like natives, learners respond faster to more frequently occurring reduced variants, which suggests that they have lexically stored variants' frequencies.

3.1 Introduction

In everyday speech, parts of words are often not fully pronounced by the speaker. In American English, for example, the word *hilarious* [hɪlɛriʌs] may be pronounced as something like [hlɛrɛs] (e.g., Ernestus & Warner, 2011; Johnson, 2004). In his study of 88,000 word tokens produced by 40 native speakers of American English in interviews (part of the Buckeye corpus of Conversational Speech; Pitt, Dilley, Johnson, Kiesling, Raymond, Hume & Fosler-Lussier, 2007), Johnson (2004) found that 25.0% of the content words lacked at least a single sound. This phenomenon of reduction, whereby words are produced with fewer segments compared to their citation forms, is also highly frequent in Germanic languages, such as Dutch (e.g., Ernestus, 2000) and German (e.g., Kohler, 1990), and in non-Germanic languages, such as French (e.g., Adda-Decker, Boula de Mareüil, Adda & Lamel, 2005) and Finnish (Lennes, Alaroty & Vainio, 2001). Despite the ubiquity of reduction in everyday speech, little is known about the processing of reduced word variants by learners of a language. This study contributes to fill this gap by investigating how learners of a language process strongly and weakly reduced words (e.g., strongly reduced [minis] and weakly reduced [minisʁ] for the French word /ministʁə/ *ministre* 'minister') in sentential context.

Studies on the comprehension of reduced word pronunciation variants have mostly focused on natives, and much less on learners. Research has shown that in general, native listeners effortlessly recognize reduced variants in context. In a study by Ernestus, Baayen, and Schreuder (2002), for instance, Dutch native listeners correctly transcribed 92.0% of the highly reduced word tokens (e.g., /mok/ for /moxələk/ 'possible'). Likewise, in an electrophysiological experiment, Drijvers, Mulder, and Ernestus (2016) did not find a higher auditory cognitive load for the processing of reduced variants than for the processing of unreduced variants of Dutch infinitives in sentence-medial position, where these variants occur most frequently.

Several studies (e.g., Manual, 1991; Zimmerer & Reetz, 2014) have suggested that natives recognize reduced words easily because they use fine phonetic details in the acoustic signal, such as voicing and lengthening of surrounding phonemes. Zimmerer and Reetz (2014), for instance, reported that /t/ deletion in German often results in the lengthening of the preceding /s/, which helps German native listeners to reconstruct the absent /t/. Manuel (1991) demonstrated that native listeners of English were able to reconstruct the absent schwa in words like *s'pport* (the apostrophe marks the absence of schwa) and could distinguish *s'pport* from *sport* by relying on the laryngeal specification of the /p/.

In addition to acoustic cues, natives use the following and preceding context to understand reduced word pronunciation variants. Van de Ven, Ernestus, and Schreuder (2012), for instance, demonstrated that Dutch listeners use both the preceding and following context to process reduced words, but that they favour acoustic cues if these conflict with the syntactic and semantic information in the context. In an eye-tracking study, Viebahn, Ernestus, and McQueen (2015) reported that participants recognized Dutch past participles more quickly when they occurred after their associated auxiliary verbs than when they preceded them. This effect was stronger for casually than for carefully produced sentences. Similarly, Tuinman, Mitterer, and Cutler (2014) examined whether Dutch listeners make use of syntax in the perception of words from which final /t/ is absent. They showed that listeners reported more /t/s in sentences in which /t/ would be syntactically correct.

Furthermore, it has been shown that the frequency of occurrence of a variant affects its processing. Ranbom and Connine (2007) studied how quickly

American English listeners recognize words that can be pronounced with both /nt/ and a nasal flap (e.g., *gentle*) and found that natives recognize the flap variant more quickly if it is the word's most frequent variant, (as is the case for *gentle*) than if it is not the most frequent variant (as for *lantern*). Pitt, Dilley, and Tat (2011) reported similar frequency effects for the recognition of American English words in which word-medial /t/ occurs in one of four phonological contexts, each favouring one of four realizations of /t/ (/t/, /ʔ/, /ɾ/ or a deleted variant). Pitt and colleagues documented a recognition benefit for the variant that is the most frequent in the given context.

The few studies that addressed the comprehension of reduced speech by learners of a language (e.g., Nouveau, 2012; Stridfeldt, 2005; Wong, Mok, Chung, Leung, Bishop & Chow, 2015) demonstrated that learners encounter difficulties recognizing reduced word variants. Nouveau performed a dictation task with Dutch university students of French. The task included French nouns with word-medial schwa (like /ʀəvy/ *revue* 'magazine'), which was left out in some words. The students recognized the reduced word variants less often (56.1% correct) than the unreduced variants (92.6% correct). Stridfeldt (2005, Chapter 6) also conducted a dictation experiment with French words that can lose their schwas, focusing on different word types (including e.g., adjectives), embedded in sentences. She found that Swedish university students of French correctly scored 46.0% (speaker 1) or 37.0% (speaker 2) on the transcription of French words in which the schwa was reduced, whereas they scored 61.0% (speaker 1) or 52.0% (speaker 2) on unreduced words. In an English dictation experiment, Wong and colleagues found that Hong Kong university students of English made errors for 40.0% of the word tokens showing regular simple phoneme reduction.

Only a few studies have examined multiple segment reduction. Ten Bosch, Giezenaar, Boves, and Ernestus (2016), for instance, conducted a dictation task in which Dutch reduced and unreduced sentences were presented to beginner second language (L2) learners of Dutch. The reduced sentences contained word combinations in which single or multiple segments were reduced, like in *ik dacht* /ig daxt/ 'I thought' and *zoveel mogelijk* /zovel moxələk/ 'as many as possible', which were pronounced as /gdax/ and /zovel mok/. The L2 learners showed more transcription errors for reduced sentences than for unreduced sentences (percentage of missing syllables or words in the learners' transcriptions when listening to reduced speech relative to unreduced speech: 14.5% versus 4.2%).

Ten Bosch and colleagues (2016) examined the comprehension of reduced word variants by learners with a low proficiency level. Since these learners have not had frequent exposure to reduced word variants, their high error rate in the transcription task is unsurprising. Nouveau (2012), Stridfeldt (2005) and Wong and colleagues (2015) tested more proficient language learners. However, their research on learners' speech processing focused on the reduction of two segments at most (i.e., schwa-reduction or reduction of /il/ or /l/ in *il(s)* 'it'/ 'he'/ 'they' in Stridfeldt, 2005, Chapter 5). Little is still known about how foreign language learners with a high proficiency level, i.e. those who have been exposed to reduced speech, process everyday speech containing words that lack more than one or two phonemes.

In a number of studies, factors underlying the poor comprehension of reduced words in language learners have been investigated. These studies examined whether language learners, like natives, benefit from subtle phonetic, semantic or syntactic information provided by the word or context.

Mitterer and Tuinman (2012) investigated whether L2 learners access knowledge from their native language (L1) when processing casual speech in their L2. In a production experiment, they observed that /t/ reduction is as frequent in Dutch nouns as it is in German proper nouns, but that it is more frequent in Dutch verbal inflections than in German verbal inflections. A follow-up perception study showed that German beginner learners of Dutch process /t/ reduction as native Dutch listeners do when the /t/ was part of a word stem, but not when /t/ was a verbal inflection. The authors concluded that a casual speech process in a second language is problematic for learners if the reduction process is not present in the learner's native language.

Other studies suggest that learners, unlike natives, do not always make full use of the semantic information provided by the signal. Van de Ven, Tucker, and Ernestus (2010) found less semantic priming between semantically related words for beginner Asian learners of English than for native listeners, not only when listening to reduced words, but also when listening to unreduced words (see also, e.g., Bradlow & Alexander, 2007).

We examined how highly proficient Dutch learners of French comprehend French reduced speech. Our first aim was to investigate how well these foreign language learners cope with single and multiple segment reduction in context. The second aim of this study was to examine the mechanisms underlying the learners' processing of reduced speech. We investigated whether highly proficient language learners, like natives, make use of semantic, syntactic or variant frequency information. The use of variants' frequency information in language learners' processing has not yet been examined. The speech that learners hear is likely to be less reduced than the type of speech that natives typically hear, as foreign language learners are less often involved in casual conversations. Learners may therefore store non-native-like frequency information, which may not be useful during the processing of foreign casual speech.

We examined highly proficient Dutch learners of French because their native language has a phonological inventory similar to that of French. Moreover, Dutch listeners are used to schwa-reduction and many other reduction phenomena from their native language also occur in French (e.g., Hanique, Schuppler & Ernestus, 2010). We therefore expected Dutch high proficiency leaners to have little difficulty processing reduced speech in French. If they did, we would expect other language learners (with a lower proficiency level or a different native phonological inventory) to experience even greater difficulties. Foreign language learners with a high proficiency level who have often been exposed to foreign casual speech may be able to reliably estimate the variants' frequencies of occurrence which can help them during the processing of reduced speech.

One of the methods we used was a transcription task. Our transcription task consisted of an informal conversation between two men, extracted from the Nijmegen Corpus of Casual French (Torreira, Adda-Decker & Ernestus, 2010). The conversation contained words in which single and multiple segments are reduced. We used the transcription task for several reasons.

Transcription tasks are relatively simple and natural tasks that are often used in the classroom. As in everyday communication, in a transcription task, listeners can be confronted with speech varying in rate, in degree of casualness, and in disfluencies. Moreover, whereas other tasks (e.g., visual world paradigm, cross-modal priming) often focus on the processing of one single word, transcription tasks can provide information about the processing of all the words in a sentence. Transcriptions tasks therefore offer important insights into the processing of complete sentences. The results of a transcription task provide information on the type of reductions (e.g., single or multiple segments reduction) inducing comprehension errors. Moreover, the transcriptions inform us about the use of semantic or syntactic cues by the listeners.

The use of variants' frequency information in language learners' processing cannot easily be tested in uncontrolled transcription tasks. We therefore used a second method: a lexical decision task with cross-modal identity priming. Our target words ended in an obstruent, liquid and schwa (e.g., *ministre*) and were presented at the offset of the prime words in the auditorily presented prime sentences. The prime word could be an unreduced (e.g., [ministʁ] or [ministʁə][1] *ministre* 'minister') or a reduced (e.g., [minis], [minisʁ]) variant of the target word, or be unrelated (e.g., [viraʒ] *virage* 'turn') to the visual target (e.g., *ministre*). We incorporated two types of reduced identity prime words. The first type was the strongly reduced variant without the entire obstruent-liquid-schwa (henceforth OLS) cluster (e.g., [minis] for *ministre* 'minister'). This variant constitutes 15.5% of the OLS tokens in the Nijmegen Corpus of Casual French (NCCFr, Torreira, Adda-Decker & Ernestus, 2010), which suggests that this variant regularly occurs in casual speech (Brand & Ernestus, 2015). The second type was the weakly reduced variant that lacks the obstruent and the schwa of the OLS cluster, but in which the liquid is still present (e.g., [minisʁ] for *ministre* 'minister'). This variant constitutes 1.4% of the OLS tokens selected from the Nijmegen Corpus of Casual French, which suggests that this variant is uncommon in casual speech. Note that the number of absent phonemes in the OLS cluster correlate inversely with the frequency of occurrence of the variant (the variant in which the complete cluster is absent is more frequent than the variant is which fewer segments are absent).

[1] According to some literature (e.g., Dell, 1985), the word-final schwa of these words is always absent or only surfaces as the result of a process of schwa insertion. Other research (Brand & Ernestus, 2015) demonstrates, however, that the variant with word-final schwa regularly occurs in casual speech (in 18.9% of all cases). Therefore, we considered both variants as unreduced variants.

Our question was to what extent the reduced pronunciation variants activated the intended (prime) word. Activation of this prime word was derived from the amount of priming observed for the visually presented target word. Given the high proficiency level of the learners, we expected that learners, like natives, would show most priming, that is the shortest response times for the highly frequent, strongly reduced variants.

3.2 Experiment 1

3.2.1 Methods

3.2.1.1 Participants

The dictation task was performed by 110 Dutch advanced learners of French, aged between 17 and 44 years (mean age: 21.6; 18 males). The participants were born and raised in the Netherlands, like their parents, and had received French instruction for six years at high school before studying French at university (from five months to three years and seven months). Their Common European Framework of Reference for Languages (CEFR, Council of Europe, 2011) levels correspond to C1-C2 level. All participants tested for this study did not report any hearing problems, and received compensation for their participation.

3.2.1.2 Materials

We extracted a fragment of an informal conversation between two 18-year-old men from Ile-de-France, from the Nijmegen Corpus of Casual French (henceforth NCCFr, Torreira, Adda-Decker & Ernestus, 2010). This fragment consisted of 10 sentences and contained both unreduced and reduced words (137 words in total, see Appendix A). Clitics were considered as independent words. The words all figured in language methods used at Dutch secondary schools. The sentences were produced at an average speech rate of 5.57 syllables per second (excluding pauses). The corpus was digitized at a sampling rate of 48 kHz, a 32-bit quantization. We scaled each sentence to the same average intensity level in PRAAT.

Table 3.1 presents examples of words and word combinations (henceforth referred to as 'items') in the transcription task's sentences with their English translations, their phonetic transcriptions found in dictionaries, the phonetic transcriptions of their actual realizations in the transcription task provided by two transcribers, and the difference between their actual realizations and their dictionary pronunciations in the number of missing phonemes and the number of deviating phonemes (i.e., differently pronounced).

The transcriptions provided by seven French natives aged between 25 and 44 years (mean age: 32.4; *SD*: 7.5, three males) served as a base line. These French native listeners correctly transcribed 88.0% of all 137 words. All words that were missing in the natives' transcriptions were function words, like *et* 'and' (three times, 14.3%), and *pour* 'to' (twice, 9.5%). Other errors were semantically related transcriptions of the auditorily presented words: *terminer* 'to end' for *finir* 'to finish' (once), *non mais moi* 'but not me' / *non ok moi* 'okay not me' for *non que moi* 'not that I' (twice), *tous les jours* 'everyday' for *tous les soirs* 'every evening' (once). The seven French native listeners never provided semantically unrelated transcriptions for the auditorily presented sentences. Further, the natives added chunks and interjections in their transcriptions like *tu sais* 'you know' (once), *mais oui* 'that is right' (once), *ouais* 'yes' (once), and *alors* 'so' (once).

Table 3.1 Properties of several items in the transcription task. The dots in the phonetic transcriptions represent word boundaries.

Item (sentence)	English translation	Phonetic transcription found in dictionaries	Phonetic transcription	Missing phonemes (%)	Deviating phonemes (%)
pouvoir (sentence 1)	‘to be able to’	/puvwar/	[vwa]	3 (50.0%)	0
je travaille (sentence 2)	‘I work’	/ʒə.travaj/	[ʃ.traj]	3 (37.5%)	0
ça m’arrive de (sentence 2)	‘it happens to me’	/sa.m.ariv.də/	[sa.m.ari.də]	1 (11.1%)	1 (11.1%)
passer une nuit blanche (sentence 2)	‘to stay awake’	/pase.yn.nɥi.blɑ̃ʃ/	[pase.ny.blɑ̃ʃ]	3 (23.1%)	1 (7.7%)
j’ai commencé (sentence 2)	‘I have started’	/ʒe.kɔmɑ̃se/	[ʒe.kɔsɛj]	1 (12.5%)	2 (25.0%)
minuit (sentence 2)	‘midnight’	/minɥi/	[məni]	1 (20.0%)	1 (20.0%)
tu vois (sentence 4)	‘you see’	/ty.vwa/	[ty.a]	2 (40.0%)	0
par exemple (sentence 4)	‘for example’	/par.ɛgzɑ̃pl/	[pa.rɑ̃mp]	3 (33.3%)	1 (11.1%)
je rentre (sentence 4)	‘I come home’	/ʒə.rɑ̃tr/	[ʒ.rɑ̃t]	2 (33.3%)	0
tu sais (sentence 7)	‘you know’	/ty.sɛ/	[t.sɛ]	1 (25.0%)	0
c’est bon allez (sentence 10)	‘it is’	/sɛ.bɔ̃.ale/	[sɛ.baj]	2 (28.6%)	2 (28.6%)

3.2.1.3 Procedure

The Dutch participants were tested in sound-attenuated booths at the Max Planck Institute for Psycholinguistics in Nijmegen, at Leiden University, at Utrecht University, at the University of Amsterdam, or at the University of Groningen. The dictation task was presented via WebExp2, a web-based program (Keller, Gunasekharan, Mayo & Corley, 2009). Participants listened to the speech to be transcribed orthographically sentence by sentence via Sennheiser HD 215 headphones. They could replay a sentence as often as necessary within a time frame (range: 10-120 seconds), based on the assumption that one minute of speech takes approximately ten minutes to transcribe. The entire experimental session lasted approximately ten minutes.

3.2.2 Results

The transcribed words were considered correct if they showed that the participant had understood the words in the auditory input. That is, we ignored small spelling errors such as the absence of accents (e.g., *j'ai commence* for *j'ai commencé* 'I have started'). Mean percentage of correctly transcribed words was 48.3% (*SD*: 20.9%, range: 11.6%.-86.2%). Learners' overall performance was thus much worse than the natives' performance (mean accuracy score of 88.0%, *SD*: 10.1%, range: 75.9%-96.4%).

We studied the eleven items listed in Table 3.1 in more detail. The selected items represent weakly to highly reduced words and different types of words, including content words, function words, and fixed expressions. This diversity represents the variety in the speech used in everyday conversations. Table 3.2 demonstrates the learners' accuracy scores for the eleven items, the absence rate of these items in the transcriptions (i.e., how often participants did not provide a transcription), and the three most frequent alternative transcriptions.

Table 3.2 The percentages of Dutch advanced learners who transcribed the items in Table 3.1 correctly or not at all, as well as the three most common incorrect transcriptions they provided. The dots in the phonetic transcriptions represent word boundaries.

Item (English translation)	**Phonetic transcription**	**% Correct**	**% Absent**	**Alternative orthographic transcription (English translation)**	**Frequency of alternative transcription**
pouvoir ('to be able to')	[vwa]	7.3%	82.7%	*pour voir* ('in order to see')	4.6%
				avoir ('to have')	2.7%
				foi/fois ('times')	1.8%
je travaille ('I work')	[ʃ.traj]	16.4%	58.2%	*je trouve* ('I find')	7.3%
				mon travail ('my work')	2.7%
				c'est ('it is')	2.7%
ça m'arrive de ('it happens to me')	[sa.m.ari.də]	11.8%	38.2%	*mon mari(e) de* ('my husband of')	8.2%
				c'est mari(e) de ('it is husband of')	4.6%
				c'est ma(r)ru de	3.6%
passer une nuit blanche ('to stay awake')	[pase.ny.blɑ̃ʃ]	20.9%	42.7%	*une nuit blanche* ('one sleepless night')	5.5%
				nuit blanche ('sleepless night')	4.6%
				nul(le) blanche ('no white')	2.7%
j'ai commencé ('I started')	[ʒe.kɔsɛj]	7.3%	31.8%	*je commence* ('I start')	19.1%
				je conseille / je conseil / je conseils/ je me conseille / me conseille / je me conceille ('I advise (myself)')	17.3%
				j'ai ('I have')	3.6%

Item (English translation)	**Phonetic transcription**	**% Correct**	**% Absent**	**Alternative orthographic transcription (English translation)**	**Frequency of alternative transcription**
minuit (‘midnight’)	[məni]	49.1%	34.6%	*m'ennui(e)* (‘am bored’) *nuit* (‘night’) *midi* (‘noon’)	2.7% 1.8% 1.8%
tu vois (‘you see’)	[ty.a]	20.0%	55.5%	*tu as, tu a* (‘you have’) *tu sais* (‘you know’) *tu* (‘you’)	12.7% 2.7% 2.7%
par exemple (‘for example’)	[pa.rɑ̃mp]	11.8%	60.9%	*plus, plus de* (‘(not) more (than)’) *pas* (‘no’) *parle, je parle* (‘(I) speak’)	16.4% 2.7% 1.8%
je rentre (‘I go home’)	[ʒ.rɑ̃t]	40.9%	44.6%	*je rent / je rend* (‘I make’/ ‘I give back’) *je rente* (‘I private income/pension’) *entre* (‘between’)	2.7% 1.8% 0.9%
tu sais (‘you know’)	[t.sɛ]	0.9%	93.6%	*c’est* (‘it is’) *t’as* (‘you have’) *t’es* (‘you are’)	1.8% 0.9% 0.9%
c’est bon allez (‘it is’)	[sɛ.baj]	0.9%	29.1%	*c’est* (‘it is’) *c'est un bon idée, c'est une bonne idée, c'est bonne idée* (‘it is a good idea’) *c'est mon idée* (‘it is my idea’)	10.9% 8.2% 5.5%

The reduced items were often completely absent in the learners' transcriptions. We found especially high absence rates for *tu sais* 'you know' (93.6%), *pouvoir* 'to be able to' (82.7%), and *par exemple* 'for example' (60.9%). These three items have in common that they do not contribute much to the meanings of the sentences. Words that contribute to the meanings of the sentences, show lower absence rates (absence rate < 50.0%, except for *je travaille* 'I work', which shows an absence rate of 58.2%).

For the items in Table 3.1, the learners also often provided transcriptions that lacked phonemes. For *tu vois* 'you see', which is pronounced as /ty.vwa/ in formal speech but as /ty.a/ in our transcription task, for instance, the onset of the second word was missing in the transcription *tu as* 'you have', which formed 12.7% of the transcriptions. For *je rentre* 'I go home', the coda of the second word was missing in the transcription *je rend(s)* 'I make/give back', which formed 2.7% of the transcriptions. For *je travaille* 'I work', the first word and only the onset of the second word were present in the transcription *je trouve* 'I find', which formed 7.3% of the transcriptions. The learners transcribed *j'ai commencé* 'I have started' in 17.3% of the transcriptions as *je conseille* 'I advise', *par exemple* 'for instance' as *plus, plus de* '(not) more (than)' in 16.4% of the transcriptions, and *c'est bon allez* 'it is' as *c'est (un(e)) bon(ne) idée* 'it is a good idea' in 8.2% of the transcriptions. Transcriptions of these words do not reflect the speakers' intentions, which consisted of longer words. Moreover, the words intended by the speakers had very different meanings: the learners' misunderstandings changed the meanings of the sentences completely, which must have resulted in an incorrect interpretation of the sentences and thus of the conversation.

Learners also encountered word segmentation problems, which resulted in syntactically incorrect and semantically unrelated transcriptions. The learners considered, for instance, the syntactic independent, but prosodic dependent (clitic) *m’* in *ça m’arrive de* ‘it happens to me’ as the onset of *mari* ‘husband’ in 12.7% of the transcriptions. The learners provided two words for the single word *minuit* ‘midnight’, which was transcribed as *m’ennuie* ‘am bored’ in 2.7% of the transcriptions, and also provided two words for the single word *pouvoir* ‘to be able to’, which was transcribed as *pour voir* ‘in order to see’ in 4.6% of the transcriptions. The learners transcribed the segments correctly, but incorrectly split them up into two words.

The alternative transcriptions in Table 3.2 further demonstrate that learners did not always transcribe the word combinations in the correct tense. The word combination *j’ai commencé* ‘I have started’, for instance, was transcribed as *je commence* ‘I start’ in 19.1% of the transcriptions. Learners thus often thought that the French native speaker had started a given action, whereas, actually, he had already completed it.

We further investigated whether the transcriptions provided by the participants were mostly based on just the number of phonemes in the acoustic signal. The number of phonemes in the learners’ transcriptions correlated with the number of phonemes in the input (r = .91, p < .001). The transcriptions provided by the participants were thus mostly based on the number of phonemes in the acoustic signal.

To summarize, the Dutch advanced learners made many more transcription errors than the French natives (mean accuracy score was 48.3% for the learners and 88.0% for the natives). Reduced words were often completely absent in the Dutch learners’ transcriptions. The learners also gave alternative transcriptions which consisted of words of which syllables, onsets or codas showed phonetic overlap with the auditory input, but these transcriptions often lacked segments. The learners were thus not able to reconstruct the absent phonemes. Learners’ transcriptions showed word segmentation problems. Moreover, the learners often gave semantically unrelated and syntactically incorrect transcriptions which resulted in an incoherent sentence that probably blocked sentence comprehension. The high error rate and the alternative transcriptions suggest that the Dutch learners hardly used syntactic or semantic cues provided by the context during processing.

To further investigate the mechanisms underlying learners' processing of reduced speech, we conducted a second experiment: a lexical decision task with cross-modal identity priming. In this second experiment, we tested French natives and a subset of the Dutch advanced learners from Experiment 1 to examine the effect of the frequency with which a given reduction pattern occurs on the processing of content words in sentential context.

3.3 Experiment 2

3.3.1 Methods

3.3.1.1 Participants

We tested fifty-five Dutch undergraduate students of French (aged 17-44 years; mean: 21.9 years; five males) who also participated in Experiment 1. In addition, fifty-seven French native speakers from Paris (13 males), aged between 17 and 35 years (mean age: 22.4), participated in the experiment. They were born and raised in the north of France and their parents were native speakers of French. All participants received compensation for taking part in the experiment.

3.3.1.2 Stimuli

Forty-eight French content words ending in an obstruent-liquid-schwa (henceforth OLS) cluster (28 monosyllabic, 20 bisyllabic; forty singular, eight plural) were selected from vocabulary lists in teaching methods used at Dutch secondary schools, and served as experimental visual targets (e.g., *spectacle* 'show'). In order to have an equal number of required *yes*- and *no*- responses in the lexical decision task and to avoid words with citation forms ending in OLS clusters being the only real words, we added 139 filler words as visual targets, of which 42 were real words and 97 were pseudo words (63 monosyllabic, 76 bisyllabic; 119 singular, 20 plural). The pseudo words consisted of real French syllables and did not resemble real Dutch words.

The visual targets were combined with auditorily presented prime sentences, each containing a prime word which was either: 1. an unreduced variant of the target word (e.g., [ministʁ] or [ministʁə] for *ministre* 'minister'); 2. an infrequent, weakly reduced variant of the target word in which the obstruent and schwa were absent, but the liquid was still present (e.g., [minisʁ] for *ministre*); 3. a highly frequent, strongly reduced variant of the target word in which the whole OLS cluster was absent (e.g., [minis] for *ministre*); 4. or a semantically and phonetically unrelated word (e.g., *virage* 'turn' for *ministre*).

Of the 97 pseudo words functioning as visual targets, 60 were phonetically related to the auditory prime word. We included these pairs to discourage participants from associating phonetic overlap between the auditory prime word and the visually presented target word with a *yes* response in the lexical decision task. For instance, the auditory prime word *lune* [lyn] 'moon' was followed by the visually presented pseudo word *lube*. The 42 real filler words were preceded by phonetically unrelated prime words.

The prime words were in sentence-medial position because reduced word pronunciation variants occur most often in this position. Moreover, pitch accents were never on the prime words, but on the surrounding adjectives (e.g., on *parfait* 'perfect' in *J'ai fait un cercle parfait sans compas* 'I made a perfect circle without compass' in which *cercle* 'circle' is the prime word) or on the following genitive construction (e.g., on *copine* 'girlfriend' in *Je vois que l'oncle de ma copine est aussi venu à la fête* 'I see that the uncle of my girlfriend also came to the party' in which *oncle* 'uncle' is the prime word). Each prime word was followed by a consonant initial word and minimally by one other word. Participants could not predict the prime word based on the main verbs in the sentence (e.g., *penser* 'to think', *faire* 'to do', *être sûr* 'to be sure', *prendre* 'to take').

We recorded the prime sentences, using Adobe Audition 1.5 and a Sennheiser ME 64 microphone, and digitized the speech at a sampling rate of 44.1 kHz, a 16-bit quantization. We asked a 28-year-old French female speaker from the north of France to pronounce the experimental prime sentences first without giving her any instructions. We then asked her to pronounce the same prime sentences three more times, asking her to realize the prime words once in an unreduced way, once without the obstruent-liquid-schwa cluster, and once without the obstruent. The unreduced experimental prime words were spliced out of the "unreduced" recordings and pasted into the recordings obtained without providing instruction, which had a mean speech rate of 5.6 syllables per second and an average duration of 2.31 s. The words were spliced at zero crossings in the wave form. Of the original prime words in the sentences recorded without providing instruction, nine were produced without schwa. The sentences obtained without providing instruction did not contain many reduced words other than the prime words, except for some highly frequent words in which the schwa was reduced like [d] for /də/ *de* 'of', [dvnir] for /dəvnir/ *devenir* 'to become', and [vny] for /vəny/ *venu* 'come' (past participle). The highly reduced experimental prime words (i.e., realized without the OLS cluster) and weakly reduced experimental prime words (i.e., realized without the obstruent and the schwa) were also spliced out of their recordings and pasted into the recordings obtained without providing instruction. We thus ensured that the experimental sentences with the reduced and unreduced prime words only differed in the realization of the obstruent-liquid-schwa cluster and that possible differences in priming effects could not be attributed to differences in the preceding or following context.

We made two recordings of the experimental prime sentences with words that were unrelated to the visual targets and for the filler visual targets. We cross-spliced the prime words from the first recordings into the second recordings. The experimental sentences with unrelated primes had a mean speech rate of 5.9 syllables per second, and the filler sentences had a mean speech rate of 6.0 syllables per second. Each sentence was scaled to the same average intensity level. The first author and an additional native speaker of French checked the experimental and the filler sentences on fluency and casualness.

Figure 3.1 presents the durations of the prime words per prime type. The word-final schwa was absent in 83.3% of the unreduced identity prime words for experimental target words. If present, the /ə/ had a maximum duration of 30 ms.

These unreduced identity prime words (mean duration: 351 ms) were significantly longer (also after Bonferroni correction) than the infrequently occurring, weakly reduced identity prime words (mean duration: 307 ms, $p < .05$) and the highly frequent, strongly reduced identity prime words (mean duration: 232 ms, $p < .001$) for experimental target words. The latter two prime types also differed significantly from each other in their duration ($p < .001$). Finally, the unreduced identity prime words were significantly longer than the unrelated prime words for experimental target words (mean duration: 261 ms, $p < .01$)

because they contained more phonemes than the unrelated prime words (mean number of phonemes: 5.07 versus 3.69).

Figure 3.1 Duration (in ms) of the four different experimental types of prime words.

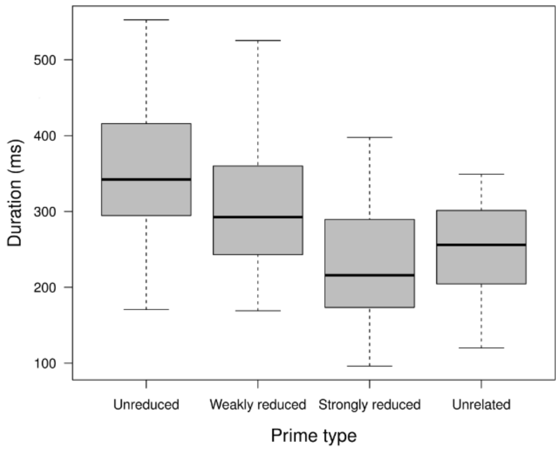
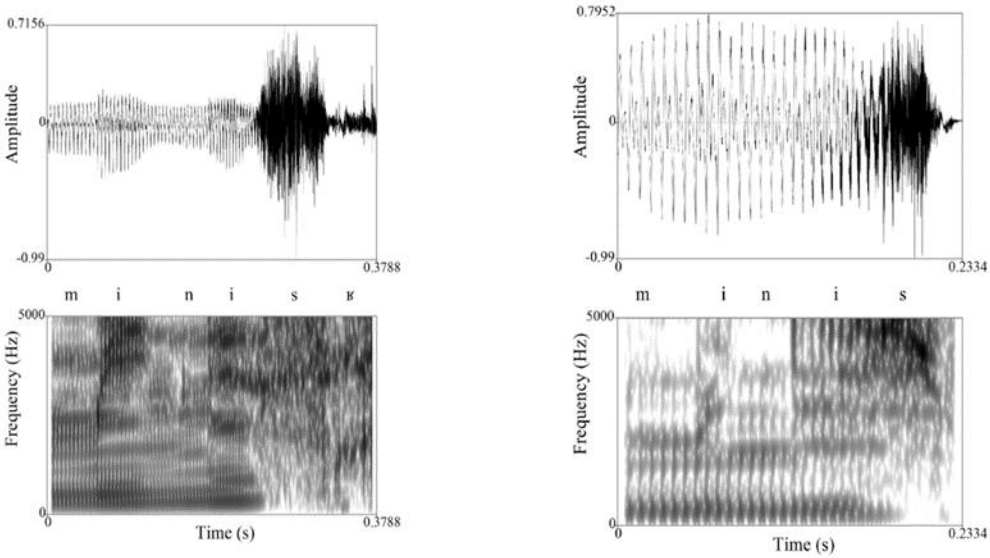


Figure 3.2 presents an example of two realisations of the prime word ministre 'minister' used in the experiment. The panel on the left shows the weakly reduced pronunciation variant, where the obstruent and schwa are reduced, while the panel on the right shows the strongly reduced pronunciation variant, where the complete OLS cluster is reduced. The duration of the prime word on the left is 379 ms, while the duration of the prime word on the right is 233 ms.

We created eight lists each of which contained the 139 filler visual targets (and their prime sentences) and the 48 experimental visual targets, of which 12 were preceded by unreduced variants of the target words, 12 by weakly reduced variants of the target words, 12 by strongly reduced variants of the target words, and 12 by unrelated words. The 48 visual experimental targets occurred equally in the four priming conditions across the eight lists. We then randomized the eight lists. Each list was preceded by four practice trials (two with pseudo words as visual targets and two with real visual target words of which one was preceded by an identical unreduced prime word, and the other one by an unrelated word).

Figure 3.2 The weakly and strongly reduced identity prime types of the visual target *ministre*.



3.3.1.3 Procedure

The Dutch participants were tested in sound-attenuated booths at the Max Planck Institute for Psycholinguistics in Nijmegen, at Leiden University, at Utrecht University, at the University of Amsterdam, or at the University of Groningen. The French participants were tested in a sound-attenuated room in the Laboratoire Charles Bruneau of the *Institut de Linguistique et Phonétique Générales et Appliquées* (ILPGA) in Paris.

The experiment was presented in E-prime 2.0 (Schneider, Eschman & Zuccolotto, 2007) from a laptop. Each participant was presented with one of the eight stimulus lists. Participants listened to the sentences via Sennheiser HD 125 headphones. At the offset of the auditory prime word, the visual stimulus was displayed in lower case, in an 18-point Courier New font size, centered on the computer screen. We presented the visual stimulus at prime word offset because previous research (e.g., Gaskell, Spinelli & Meunier, 2002) has shown that it is likely that the prime word has been recognized by then. The visual stimulus was displayed for 3000 ms.

Participants were asked to indicate for each visually presented stimulus whether it was a real French word or not by pressing a *yes* or *no* key as quickly and as accurately as possible. Right handed participants responded by pressing the key *m* for *yes*, and the key *z* for *no*, on the keyboard. For left-handed participants, the *yes* and *no* keys were reversed.

The next auditory stimulus was played 1000 ms after a key response or 4000 ms after prime word offset if the participant did not respond. Sentences were played in their entirety and did not terminate at the offset of the prime word. The 187 trials were divided in three equal blocks, separated by short breaks in the experiment. The lexical decision task with cross-modal identity priming was followed by the dictation task (see Experiment 1) and a background questionnaire. A complete session lasted approximately 30 minutes.

3.3.2 Results

Both the French natives and the Dutch advanced learners performed well on the visual lexical decision task. The mean accuracy score on the 48 target words was 96.7% for the French natives and 91.0% for the Dutch learners. The mean accuracy score on the pseudo words was 97.3% for the French natives and 78.9% for the Dutch learners.

We analyzed the RTs, measured from the onset of the presentation of the visual stimulus, for the trials in which the target words had correctly been classified as real words (5062 trials). Since the accuracy score and the RT for the target word *chantre* 'cantor' were very different from the average score and average RT for all other words (mean accuracy score of 0.51 for *chantre* versus 0.96 for all other words; mean RT of 1274 ms for *chantre* versus 928 ms for all other words), we excluded this word from our analyses. Furthermore, we excluded two extremely slow Dutch participants to allow a normal distribution of the RTs. Finally, RTs deviating from the new mean (905.89 ms) by more than 2.5 times the standard deviation (1100.65 ms) were considered as outliers and were therefore removed. This resulted in 4916 observations for analysis (91.1% of all data).

We analyzed the RTs by means of mixed-effects regression models, with visual stimulus and participant as crossed random effects and with prime word type and participant group as the main fixed predictors. We also added several control predictors in order to reduce the variance in the data, such as prime word duration, trial number, previous RT and log-transformed word form frequencies for books as listed in Lexique 3.80 (New, Pallier, Ferrand & Matos, 2001). Furthermore, we investigated whether distance (in ms) from prime word offset till sentence offset predicted response times. If so, this could imply that listeners postponed their lexical decision till the end of the sentences. Random slopes were tested for all fixed predictors. We only retained those predictors in the model that showed statistically significant simple effects or figured in statistically significant interactions.

First, we investigated whether the reduced variants of the target words primed the target words at all. The unrelated word was mapped on the intercept of our model. The model, presented in Appendix B, showed that participants responded faster to visual targets preceded by highly reduced variants and unreduced variants than to targets preceded by unrelated variants. We further found a simple effect of the remaining distance, a significant interaction between the weakly reduced prime words and remaining distance, and a three-way interaction between these two variables and participant group. These results show that the longer the time interval between the offsets of auditory prime words and sentence offsets, the more slowly the French participants responded to visual targets, except when these were preceded by weakly reduced words. The Dutch learners showed a different pattern: they only showed an effect of the time interval if target words were preceded by weakly or highly reduced words.

Second, we examined whether the unreduced variants of the target words showed more priming than the other variants. To that end, we reran our model with the unreduced prime word on the intercept. The obtained model is summarized in Table 3.3.

Table 3.3 shows that the Dutch participants were faster than French natives and that they responded faster to target words that occurred more often in books. The random slopes demonstrate that the first effect was larger for some words than for others, and that the second effect held more for some participants than for others. All participants were slower if they had also responded slowly to the previous trial. The random slopes show that these effects differed per word and per participant. Further, our natives showed longer RTs than learners, probably because the number of words that were pre-activated by the prime during the processing of the target word was larger for natives than for learners.

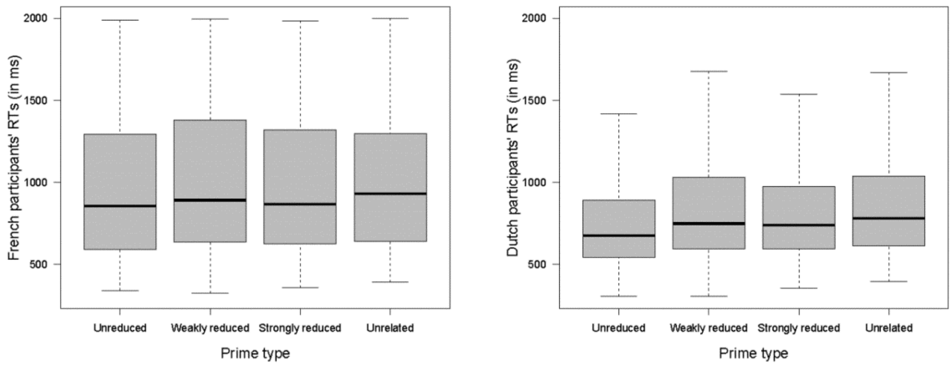
Table 3.3 Summary of the linear mixed-effects model fitted on the lexical decision times from the French natives and the Dutch advanced learners. The intercept represents French participants hearing an unreduced experimental prime word. Remaining time refers to the time interval between prime word offset and sentence offset.

Fixed effects		*ß*	*t*
Intercept		885.52	28.58***
Dutch learners		-180.87	-4.33***
Word form frequency books (Lexique)		-2.670	-0.35
Dutch learners * word form frequency books (Lexique)		-25.06	-2.20*
Previous RT		47.69	7.59***
Trial number		-41.16	-5.21***
Weakly reduced prime word		56.38	4.69***
Strongly reduced prime word		21.94	1.83
Unrelated prime word		64.28	4.45***
Dutch learners * weakly reduced prime word		6.00	0.35
Dutch learners * strongly reduced prime word		34.31	2.012*
Dutch learners * unrelated prime word		38.11	1.86
Remaining time		32.32	2.84**
Dutch learners * remaining time		-34.09	-2.097*
Weakly reduced prime word * remaining time		-33.97	-2.46*
Strongly reduced prime word * remaining time		-11.20	-0.80
Unrelated prime word * remaining time		10.00	0.67
Dutch learners * weakly reduced prime word * remaining time		50.58	2.61**
Dutch learners * strongly reduced prime word * remaining time		38.80	1.97*
Dutch learners * unrelated prime word * remaining time		-0.74	-0.035
Random effects			*SD*
Word	intercept		54.02
	trial number		28.93
	previous RT		23.81
	participant group		66.37
Participant	intercept		224.27
	trial number		59.99
	previous RT		35.64
	word form frequency books (Lexique)		17.98
Residual			189.70

*Note: *** indicates p < .001, ** indicates p < .01 and * indicates p < .05*

More importantly, both French and Dutch participants responded faster to target words preceded by unreduced prime words than to targets preceded by infrequently occurring, weakly reduced identity prime words or by unrelated words (see also Figure 3.3). Unlike the French natives, the Dutch learners also showed shorter RTs to targets preceded by unreduced prime words than to targets preceded by frequently occurring, strongly reduced identity prime words. This result suggests that, unlike the natives, the Dutch advanced learners benefited most from unreduced identity primes.

Figure 3.3 RTs on the visual targets preceded by the four different types of prime words for the French participants (left panel) and the Dutch participants (right panel).



Finally, we again found an effect of the time interval between prime word offset and sentence offset. The larger the time interval (i.e., remaining time in Table 3.3), the longer the French natives' RTs except when the prime word was the infrequent, weakly reduced variant of the word. The Dutch listeners showed an effect of this distance when the prime word was the weakly or highly reduced word.

In order to investigate whether RTs to targets preceded by strongly reduced prime words significantly differed from RTs to targets preceded by weakly reduced prime words, we also ran the model with the weakly reduced prime word on the intercept. This third model shows that both the natives and the learners responded faster to targets preceded by strongly reduced prime words than to targets preceded by weakly reduced prime words (*ß* = -34.43*, t* = -2.85*,*

$p < .01$, while the interaction with participant group was not significant, $p > .05$).

In summary, this experiment produced three important results. First, both the French natives and the Dutch learners of French showed priming from unreduced and highly reduced identity prime words. However, while the French natives showed no difference in priming from the unreduced and highly reduced identity prime words, the Dutch learners of French showed more priming from the unreduced than from the highly reduced identity prime words. This shows that the Dutch learners had more difficulties processing the reduced variants than the native listeners.

Second, both groups showed effects from the time interval between prime word offset and sentence offset. However, the natives showed this effect except when the prime word was the infrequent, weakly reduced identity prime, whereas the Dutch learners only showed these effects when the prime word was reduced (weakly or highly reduced prime word). The natives thus seemed not to take the remainder of the sentence into account when the prime word was infrequent and therefore difficult to process. The learners, in contrast, applied a different strategy and did not take the remainder of the sentence into account if the prime word was easy to process (unreduced identity prime word or unrelated prime word). The learners thus suffered more from reduction than the natives did and applied a different strategy when hearing unreduced rather than reduced identity prime words.

Third, both groups showed more priming from the common highly reduced identity prime words than from the infrequent, weakly reduced prime words, and both groups thus showed a frequency effect of reduction type. This frequency effect is also visible in the effect of the interval between prime word offset and sentence offset in the natives' RTs, which was absent for the infrequent but not for the frequent reduced variants.

3.4 Experiment 1 and Experiment 2

We investigated whether the results of the off-line task presented in Experiment 1 of this study predicted those of the on-line task used in Experiment 2. We examined whether the Dutch learners' accuracy scores in the off-line transcription task predicted their RTs in the on-line lexical decision task with cross-modal identity priming[2]. Thus, we added the participant's overall accuracy score in the transcription task as a predictor for their response speed in Experiment 2 and investigated whether a higher accuracy score correlated with more proficient, that is longer, RTs in the lexical decision task. The final model does not show a significant effect of this accuracy score (t = 1.72, p = .09). Five participants had extreme accuracy scores (below 20.0% or above 70.0% correct). When we reran the model excluding these outliers from our dataset, we found a significant effect of accuracy score on RT: the higher the Dutch learners' accuracy scores in the transcription task, the longer their RTs in the cross-modal identity priming experiment with the lexical decision task ($ß$ = 37.63, t = 2.19, $p < .05$). This suggests that Dutch learners with a higher proficiency level showed more native-like RTs in the lexical decision task, and that learners' conscious, explicit behaviour in the off-line transcription task reflects their processing in real-time. The model showed no interaction of accuracy score with prime type or the duration of the remaining sentence, possibly because of lack of statistical power.

3.5 General discussion

This study investigated how advanced learners of a language cope with single and multiple segment reduction in words in context and whether the relative

[2] We did not take the learners' accuracy score in the lexical decision task with cross-modal identity priming as the dependent variable, because the learners' average accuracy score was close to ceiling level (91%).

frequency of a reduced word pronunciation variant plays a role in the learners' processing of these variants. We used two very different methods to examine our research questions: a transcription task and a lexical decision task with cross-modal identity priming. The types of errors in the first task provide an indication of the difficulties that learners encounter during the comprehension of (casual) speech and demonstrate whether or not learners use phonetic, syntactic or semantic cues provided by the context. The transcriptions show what listeners understand and whether communication is hindered. The lexical decision task also examines the difficulty of processing reduced speech by learners; it investigates whether learners need more time to respond to visual targets preceded by reduced than by unreduced pronunciation variants, and whether all types of reduced variants prime equally, independent of their frequencies. The second task thus also informs us about whether variants' frequencies affect processing.

We focused on Dutch learners who had received French instruction at high school for six years and continued upgrading their proficiency level at university, where they were exposed to French on a daily basis. The phonological inventory of Dutch is highly comparable to that of French, and the Dutch learners are also used to reduction in their native language (e.g., schwa reduction, Hanique, Schuppler & Ernestus, 2010). For these reasons, we expected that these learners would encounter few problems when processing French reduced words.

The transcription task presented part of a conversation extracted from a corpus of casual French (NCCFr, Torreira et al., 2010) and thus contained naturally produced reduced word pronunciation variants. The results showed much lower accuracy scores for the highly proficient learners than for natives (48.3% versus 88.0%). The high percentage of errors demonstrates that not only beginner learners but also advanced learners encounter difficulties understanding reduced speech, even if their native phonological inventory is similar to that of the target language.

We conducted a detailed investigation of the learners' performance on eleven items that represent weakly to highly reduced words and word combinations, including content words, function words, and fixed expressions. Whereas French natives encountered no difficulties transcribing these eleven items, the learners often did not transcribe them at all (mean percentage of absence: 52.0%). It is likely that the learners did not recognize these reduced items.

If the learners provided transcriptions, these mostly showed phonetic overlap with the input (e.g., with the onset or coda). Moreover, the learners seldom provided transcriptions that contained more segments than the realizations of the items in the experiment. The learners' transcriptions lacked segments in, on average, 30.7% of the items. The learners transcribed *c'est bon allez* 'it is' pronounced as [sɛ.baj], for instance, in 10.9% of the transcriptions as *c'est* 'it is' /sɛ/. The transcriptions may partly reflect the segments in the auditory input, but do not reflect the speaker's intention, which consisted of longer words. The learners hardly used phonetic cues to reconstruct the input form.

Likewise, the learners did not always make full use of the semantic or syntactic information provided by the context: they provided transcriptions that did not fit into the sentences. The learners transcribed *ça m'arrive de* 'it happens to me' pronounced as [sa.mari.de], for instance, in 8.2% of the transcriptions as *mon mari(e) de* 'my husband of' /mɔ̃.mari.də/. This finding shows that not only beginner learners (e.g., Van de Ven et al., 2010; Bradlow & Alexander, 2007; Kotz & Elston-Güttler, 2004) but also advanced learners have more difficulties than native listeners in using semantic cues, even if they can replay the sentences.

Learners also often encountered word segmentation problems, which resulted in semantically and syntactically incorrect transcriptions. We observed problems in the transcription of clitics: *m'arrive* /m ariv/ in the example fragment *ça m'arrive de* 'it happens to me' already provided above, for instance, was transcribed as *mari de* /mari.də/ 'husband of' in 12.7% of the transcriptions. These errors may be related to earlier observations that Dutch advanced learners of French encounter problems acquiring clitic placement in French (e.g., Sleeman, 2010). Segmentation problems were also observed for content words: learners transcribed *pouvoir* /puvwar/ 'to be able to' pronounced as [vwa], for instance, as *pour voir* /pur vwar/ 'to see' in 4.6% of the transcriptions.

The transcription task thus shows that highly proficient learners have great difficulties understanding reduced word pronunciation variants in a natural casual conversation. The learners' transcriptions often changed the meaning of the sentence completely, which may have led to a misunderstanding of the conversation. The comprehension difficulties have a variety of sources: word segmentation problems and insufficient use of phonetic, syntactic or semantic information.

In the lexical decision task with cross-modal identity priming, we further investigated whether reduced pronunciation variants induced problems in comprehension. Moreover, we examined whether variants' frequencies influenced speech processing: We studied whether listeners responded faster to pronunciation variants that occur more often, focusing on words ending in obstruent-liquid-schwa (OLS) clusters. Visual target words were combined with auditorily presented prime sentences, each containing a prime word, which was either: 1. an unreduced variant of the target word (e.g., [ministʁ] or [ministʁə] for *ministre* 'minister'); 2. an infrequent, weakly reduced variant of the target word in which the obstruent and schwa were absent, but the liquid was still present (e.g., [minisʁ] for *ministre*); 3. a highly frequent, strongly reduced variant of the target word in which the whole OLS cluster was absent (e.g., [minis] for *ministre*); 4. or a semantically and phonetically unrelated word (e.g., *virage* 'turn' for *ministre*). The native listeners recognized visual target words most quickly if these were preceded by unreduced or by highly frequent, strongly reduced prime words. In contrast to previous results (e.g., Racine, Bürki & Spinelli, 2014; Ranbom & Connine, 2007, Janse, Nooteboom & Quené, 2007; Tucker & Warner, 2007), our findings do not show a clear processing advantage for the unreduced variant. It is unlikely that this lack of a clear processing advantage for the unreduced variant is due to lack of statistical power in our data set, because the difference between target words preceded by unreduced variants and infrequent, weakly reduced variants did reach statistical significance. One likely reason for why the unreduced variant did not have a clear processing advantage in our experiment is that we presented our prime words in the middle of sentences, where reduction is most common.

Unlike the native listeners, the learners recognized visual target words most quickly if these were preceded by unreduced identity prime words: They needed more time to respond to visual targets preceded by highly frequent, strongly reduced prime words (mean RT: 801 ms) than to those preceded by unreduced prime words (mean RT: 740 ms). Even advanced learners of a language are thus hindered by reduction, as also shown in Experiment 1.

Learners probably process the unreduced variant the most quickly because, at school, they are more often visually and/or auditorily exposed to the unreduced variant than to the reduced variants of a word. Educational programmes focus on the written language, and audio materials presented in the classroom contain considerably fewer reduced pronunciation variants than everyday casual conversations (e.g., Gilmore 2004, Fonseca-Greber & Waugh, 2003; Waugh & Fonseca-Greber, 2002; McCarthy & Carter, 1995).

The delay in the processing of reduced variants (compared to unreduced variants) may result in problems in the comprehension of words following the reduced words. Learners have less time to process these following words, because at the moment these words are presented, learners are still in the middle of processing the preceding reduced words. This is in accordance with learners' report that they find it difficult to keep up with casual speech.

That learners encountered problems processing both reduced variants is also reflected in the significant statistical interaction between participant group, prime word type and the time interval between the prime word offset and sentence offset: the longer the time interval, the longer the learners' RTs when the prime word was reduced. The natives, in contrast, always showed an effect of this time interval except when the prime word was the infrequent, weakly reduced variant of the target word. Whereas the two reduced variants thus show the same effect of the time interval for the learners, the highly frequent, strongly reduced variant patterns with the unreduced variant for the native listeners.

Furthermore, this interaction suggests that the native listeners and the learners applied different strategies. On the one hand, natives probably had enough cognitive resources available to listen to the complete sentence unless the prime word was an infrequent, weakly reduced variant. As a consequence, they did not show an effect of the time interval between the infrequent, weakly reduced prime word's offset and sentence offset. The learners, on the other hand, probably used most of their cognitive resources to process the unreduced prime words. They, therefore, did not show an effect of the time interval between unreduced prime word offset and sentence offset. When the prime word was reduced, the learners may have taken the following words into account in order to comprehend the prime word and thus showed an effect of the time interval between reduced prime word's offset and sentence onset.

Our interpretation of the interaction that learners tried to extract information from the context when hearing reduced prime words may seem contradictory to the results of the transcription task, which indicate that they could not always use the semantic and syntactic information provided by the context. However, the fact that they could not use this information does not imply that they did not *try* to use it.

Experiment 2 also provides evidence that both native listeners and advanced learners rely on the relative frequencies of pronunciation variants. We found a significant difference in RTs between targets preceded by infrequent, weakly reduced prime words, and targets preceded by highly frequent, strongly reduced prime words. Both participant groups processed the target words preceded by the highly frequent prime variants more quickly; please note that the number of absent phonemes in the prime word's OLS cluster correlated inversely with the frequency of occurrence of the variant. The difference in RT can therefore not be attributed to the amount of reduction in the prime word, as we would then expect the opposite RT pattern, but can only be attributed to the variants' frequencies of occurrence.

This frequency effect is also visible in the effect of the time interval between prime word offset and sentence offset in the natives' RTs: as mentioned above, this effect of the interval was absent for the infrequent but not for the frequent reduced variants. The frequency effect is not visible in the effect of the interval for learners, possibly because the learners still found the highly frequent reduced variant difficult to process, as they were not very familiar with it.

Whereas the high percentage of errors in the learners' transcriptions in Experiment 1 showed that the learners encountered many problems comprehending reduced speech, the relative frequency effect in the lexical decision task of Experiment 2 suggests that learners do have some knowledge of reduced pronunciation variants. Had we only conducted the transcription task, we would have concluded that advanced learners have very little knowledge of reduction. The lexical decision task with cross-modal priming refines this picture: it shows that advanced learners do have knowledge of reduction; they have at least stored the relative frequencies of some reduced pronunciation variants.

The question then arises why learners hardly showed knowledge of reduction in the transcription task. They probably did, but the transcription task not only contained items showing single and multiple segment reduction, as did the lexical decision, but also massive reduction, for instance, *c'est bon allez* 'it is', which is pronounced as [sɛ.bɔ̃.ale] in formal speech, but as [sɛ.baj] in our transcription task. Learners thus may have even greater difficulties processing massively reduced words. Furthermore, learners could more easily rely on the cues in the sentence context in the lexical decision task than in the transcription task, as in the latter task this linguistic context contained many other reduced words.

Finally, we also examined whether learners' accuracy scores in the off-line transcription task predicted their RTs in the cross-modal identity priming study. After excluding five outliers, the learners with higher accuracy scores in the transcription task showed longer RTs in the cross-modal identity priming study and thus showed more native-like RTs. Learners' conscious, explicit behaviour in off-line transcription tasks thus reflects their processing in real-time. This implies that language teachers can simply use dictation tasks to get an indication of the learners’ processing of everyday speech.

To conclude, the combined results of these two experimental studies have extended our knowledge of the processing of casually produced sentences by highly proficient language learners. The learners’ high percentage of errors in the transcription task and the delay in learners’ RTs if a visual target was preceded by a reduced pronunciation variant instead of an unreduced variant, show that even highly proficient learners encounter difficulties processing reduced speech. Learners probably do not make full use of phonetic, semantic and/or syntactic cues provided by the context. They are, nevertheless, sensitive to variants’ frequencies. The lexical decision task showed that the learners responded faster to visual targets preceded by highly frequent prime variants than to targets preceded by infrequent prime variants. The combination of the transcription task and the lexical decision task demonstrates that even though highly advanced learners have great difficulties understanding casual speech, they do have some (frequency) knowledge of reduction.

Appendix A

Sentences in the dictation task

1. *Une fois je me suis couché à quatre heures et demie pour pouvoir finir un travail quoi.*

‘Once, I went to bed at 4.30 p.m. because I had to finish some work.’

2. *En même temps moi je travaille bizarrement ça m’arrive de finir de passer une nuit blanche tu vois mais parce que j' ai commencé à minuit.*

‘At the same time, I hardly work, it happens to me that I stay awake, you see, because I have started at midnight.’

3. *Hm ouais d'accord.*

‘Hm, yes/yeah okay.’

4. *Non que moi je commence tu vois je je par exemple le soir je rentre il est sept heures et demie.*

‘Not that I start, you see, for instance, at night, I come home, it is half past 7.’

5. *Je finis les je finis il est u(ne) enfin je mange et puis après je commence à travailler et je finis à une heure.*

'I finish, the, I finish, it is one, finally I eat and then I start working and I finish at one o'clock.'

6. *Quasiment tous les soirs même.*

'Almost every evening.'

7. *Ouais mais Line tu sais Line elle est en médecine.*

'Yes, but Line, you know Line, she is studying medicine.'

8. *Ouais.*

'Yes.'

9. *Elle c' est euh non stop vingt-quatre sur vingt-quatre quand je te dis.*

'She, it is euh non stop 24/24 when I tell you.'

10. *Même pour m'appeler pour me demander un truc c' est bon allez dépêche toi.*

'Even for calling me to ask me something, it is: hurry up.'

Appendix B

Summary of the linear mixed-effects model fitted on the lexical decision times from the French natives and the Dutch advanced learners. The intercept represents French participants hearing an unrelated experimental prime word. Remaining time refers to the time interval between prime word offset and sentence offset.

Fixed effects		*ß*	*t*
Intercept		949.80	29.87***
Dutch learners		-142.75	-3.31***
Word form frequency books (Lexique)		-2.67	-0.35
Dutch learners * word form frequency books (Lexique)		-25.06	-2.20*
Previous RT		47.69	7.59***
Trial number		-41.16	-5.21***
Weakly reduced prime word		-7.90	-0.55
Strongly reduced prime word		-42.34	-2.90**
Unreduced prime word		-64.28	-4.45***
Dutch learners * weakly reduced prime word		-32.11	-1.56
Dutch learners * strongly reduced prime word		-3.80	-0.19
Dutch learners * unreduced prime word		-38.11	-1.86
Remaining time		42.31	4.47***
Dutch learners * remaining time		-34.83	-2.58**
Weakly reduced prime word * remaining time		-43.97	-3.02**
Strongly reduced prime word * remaining time		-21.20	-1.44
Unreduced prime word * remaining time		-10.00	-0.67
Dutch learners * weakly reduced prime word * remaining time		51.32	2.44*
Dutch learners * strongly reduced prime word * remaining time		39.54	1.87
Dutch learners * unreduced prime word * remaining time		0.74	0.035
Random effects			*SD*
Word	intercept		54.02
	trial number		28.93
	previous RT		23.81
	participant group		66.37
Participant	intercept		224.27
	trial number		59.99
	previous RT		35.64
	word form frequency books (Lexique)		17.98
Residual			189.70

*Note: *** indicates p < .001, ** indicates p < .01 and * indicates p < .05*

Listeners' processing of a given reduced word pronunciation variant directly reflects their exposure to this variant: Evidence from native listeners and learners of French

Chapter 4

This chapter is a reformatted version of:
Sophie Brand & Mirjam Ernestus (2017). Listeners' processing of a given reduced word pronunciation variant directly reflects their exposure to this variant: Evidence from native listeners and learners of French. *Quarterly Journal of Experimental Psychology*. doi: 10.1080/17470218.2017.1313282

Abstract

In casual conversations, words often lack segments. This study investigates whether listeners rely on their experience with reduced word pronunciation variants during the processing of single segment reduction. We tested three groups of listeners in a lexical decision experiment with French words produced either with or without word-medial schwa (e.g., /ʀəvy/ and /ʀvy/ for *revue*). Participants also rated the relative frequencies of the two pronunciation variants of the words. If the recognition accuracy and response times for a given listener group correlate best with the frequencies of occurrence holding for that given listener group, recognition is influenced by listeners' exposure to these variants. Native listeners' relative frequency ratings correlated well with their accuracy scores and RTs. Dutch advanced learners' accuracy scores and RTs were best predicted by their own ratings. In contrast, the accuracy and RTs from Dutch beginner learners of French could not be predicted by any relative frequency rating; the rating task was probably too difficult for them. The participant groups showed behaviour reflecting their difference in experience with the pronunciation variants. Our results strongly suggest that listeners store the frequencies of occurrence of pronunciation variants, and consequently the variants themselves.

4.1 Introduction

In casual speech, native speakers do not pronounce all sounds. In American English, the word *yesterday* /jɛstədeɪ/, for instance, may be pronounced as something like /jɛʃeɪ/. Johnson (2004) studied 88,000 word tokens produced by 40 native speakers of American English in interviews (part of the Buckeye corpus of Conversational Speech, Pitt et al., 2007) and found that 25.0% of the content words lacked at least a single sound. This phenomenon of reduction, whereby words are produced with fewer segments compared to their citation forms, is also highly frequent in Germanic languages, such as Dutch (e.g., Ernestus, 2000) and German (e.g., Kohler, 1990), and in non-Germanic languages such as French (e.g., Adda-Decker, Boula de Mareüil, Adda & Lamel, 2005) and Finnish (Lennes, Alaroty & Vainio, 2001). Despite the ubiquity of reduction in everyday speech, little is known about how listeners comprehend reduced words. This study contributes to filling this gap by investigating whether native listeners and learners of a language rely on lexical representations of reduced word pronunciation variants during the processing of single segment reduction.

To date, there has been little agreement on the processes that underlie the comprehension of reduced word pronunciation variants, and on which word forms are stored in the mental lexicon. We can distinguish two main accounts. The first account assumes that only unreduced word pronunciation variants are stored in the mental lexicon. Reduced words (e.g., French /ʀvy/ *revue* 'magazine') are matched with the representations for the unreduced variants (e.g., /ʀəvy/), by means of general processes (e.g., schwa-insertion), each applying to several words. This account is compatible with most psycholinguistic models of word recognition, such as TRACE (McClelland & Elman, 1986), the Neighborhood Activation Model (NAM: Luce & Pisoni, 1998), PARSYN (Luce, Goldinger, Auer & Vitevitch, 2000), and Shortlist B (Norris & McQueen, 2008).

The second account assumes that both unreduced and reduced word pronunciation variants are stored in the mental lexicon. A French speaker would recognize reduced /ʀvy/ by matching the acoustic signal with the lexical representation /ʀvy/. This account requires that the word recognition models mentioned above extend their lexicons with additional word pronunciation variants.

The debate on the roles of storage versus computation is not restricted to the domain of word pronunciation variants. There have, for instance, also been extensive debates about whether regular morphologically complex words are stored in the mental lexicon or whether they are computed from their morphemes every time they are processed (e.g., Baayen, McQueen, Dijkstra & Schreuder, 2003; Alegre & Gordon, 1999; Clashen, 1999; Sereno & Jongman, 1997; Stemberger & MacWhinney, 1988). Many researchers argue that a class of regular complex words (e.g., plurals) is stored if the members with high frequencies of occurrence are processed more quickly than members of lower frequencies (given similar lemma frequencies).

The debate on storage versus computation of word pronunciation variants also focusses on the role of the variant's frequency of occurrence in processing (e.g., Ranbom & Connine 2007; Pitt, Dilley & Tat, 2011). Ranbom and Connine studied how quickly American English listeners recognize words that can be pronounced with both /nt/ and a nasal flap (e.g., *gentle*) and found that native listeners recognize the flap variant more quickly if it is the word's most frequent variant (as is the case for *gentle*) than if it is not the most frequent variant (as for *lantern*). Pitt and colleagues reported similar results for the recognition of American English words in which word-medial /t/ occurs in one of four phonological contexts, each favouring one of four realisations of /t/ (/t/, /ʔ/, /ɾ/

or a deleted variant). Pitt and colleagues documented a recognition benefit for the variant that is the most frequent in the given context.

These frequency effects reported by Ranbom and Connine (2007) and by Pitt and colleagues (2011) can easily be explained in models that allow for storage of more than one pronunciation variant per word (i.e., the second account discussed above): the variants are stored together with their a-priori probabilities, which determine how easily they are processed (e.g., since they co-determine the variants' resting activation levels).

Frequency effects can also easily be accommodated in models that assume that word pronunciation variants are never stored but always computed (i.e., the first account, e.g., Roelofs, 1997), but only in speech production. During production, the frequency of occurrence of a reduced variant may modulate the ease of application of the rule that transforms the stored, unreduced variant into the corresponding reduced variant. It can be argued that a similar process may account for frequency effects in speech comprehension. Prelexical reconstruction rules may be sensitive to the frequencies with which they are applied to each word. That is, when hearing a word, a listener may apply a reconstruction rule and consider the resulting prelexical representation as likely for the given acoustic input as indicated in the description of the rule. For instance, when a native listener of French hears /ʀvy/ *revue* 'magazine', (s)he may apply schwa-insertion (resulting in /ʀəvy/) and keep the two prelexical representations active until the word is identified. The recognition process would favor the prelexical representation that is most likely given the likelihood that the word is produced with schwa, as indicated by the rule. However, the word-specific probability of an insertion rule can only become available after the word has been identified. If information about the probability of an insertion rule only becomes available after the word has been identified, how then can this probability affect the ease of word recognition?

The assumption that variants are stored together with their frequencies of occurrence predicts that listeners with different experience with the variants react differently on these variants. For instance, if one listener tends to hear one word in its reduced variant and a second word in its unreduced variant while another listener tends to hear the former word in its unreduced variant and the latter word in its reduced variant, we expect that the former listener recognizes the former word best in its reduced variant and the latter word in its unreduced variant while the opposite pattern holds for the latter listener. The aim of this study is to investigate whether this expectation is correct, that is, whether the recognition accuracy and response times for a given listener group correlate best with the frequencies of occurrence of the variants holding for that given listener group. This would indicate that the pronunciation variants are lexically stored.

We focus on schwa in the initial syllable of French words, which can be absent even if the words are produced in isolation and in a formal situation (e.g., /fənɛtʁ/ 'window', which can be pronounced as /fnɛtʁ/; Bürki, Ernestus & Frauenfelder, 2010). The absence of schwa is often complete, leaving no traces in the acoustic signal (e.g., Bürki, Fougeron & Gendrot, 2007). This study focusses on French as spoken in the north of France, where schwa is more often absent than in the south of France (e.g., Durand & Eychenne, 2004; Coquillon, 2007). In the north of France, schwa is absent in more than 50.0% of the word tokens in conversational speech (e.g., Fougeron, Goldman & Frauenfelder, 2001; Hansen, 1994). Sociolinguistic factors influence the likelihood of the absence of schwa in the north of France (e.g., Hansen, 2000) as they do in the south of

France (e.g., Eychenne & Pustka, 2007). This study focusses on Northern French as produced by a young adult, highly educated speaker.

Our study, which concentrates on vowels, thus differs from those of Ranbom and Connine (2007) and Pitt and colleagues (2011), who examined variations in the realisation of consonants. Several studies have shown that listeners process vowels differently from consonants. For instance, Pisoni (1973) and Stevens, Liberman, Studdert-Kennedy, and Öhman (1969) showed that listeners perceive finer distinctions in vowels than in consonants. Cutler, Sebastián-Gallés, Soler-Vilageliu, and Van Ooijen (2000) reported that participants are more likely to change vowels than consonants if they are asked to turn pseudo words into real words. Given this processing difference between consonants and vowels, findings on consonants may not be generalized to vowels. The question thus arises whether listeners are also sensitive to frequencies of word pronunciation variants that differ from their respective full variants only in their vowels.

Two studies (Bürki, Ernestus, Gendrot, Fougeron & Frauenfelder, 2011a; Bürki, Ernestus & Frauenfelder, 2010) investigated whether reduced pronunciation variants of French words with schwa are stored in the mental lexicon, both focusing on speech production. The corpus study by Bürki and colleagues (2011) showed that the duration of schwa in a word like *revue* 'magazine' is physiologically conditioned (e.g., by the voicing specifications of the consonant following schwa), whereas the presence of schwa is particularly affected by prosodic factors. The limited overlap between the predictors for variant choice (i.e., with or without schwa) and for schwa duration, combined with the nature of these variables, suggests that selection of the variant to be pronounced occurs before phonetic implementation, which determines schwa duration. As such, Bürki and colleagues assume that both pronunciation variants of a word, like /ʀəvy/ and /ʀvy/ for *revue*, are lexically stored. Bürki and colleagues (2010) showed that native speakers of French have shorter production latencies for a given variant of a schwa-word if they think this variant occurs more often. The authors argue that speakers store the relative frequencies of the different pronunciation variants of a word in their mental lexicons and therefore these pronunciation variants are themselves also stored. As explained above, however, a mental lexicon that contains only unreduced variants can also account for frequency effects in production: frequency of occurrence of a variant may modulate the ease of application of a rule that transforms the stored variant into the variant to be produced. A study on frequency effects in speech production is therefore not highly informative for answering the question whether word pronunciation variants are stored.

We found one study that investigated the *comprehension* of word pronunciation variants without schwa in French (Racine, Bürki & Spinelli, 2014). The authors studied how the recognition of French reduced schwa words by French native children is affected by spelling. The results show that both readers and pre-readers recognize the variant with schwa (e.g., /ʀənaʀ/ *renard* 'fox') more quickly than the variant without schwa (e.g., /ʀnaʀ/) for words that are most frequently produced with schwa. In contrast, readers and pre-readers react differently to words that are spelled with schwa but never pronounced with this vowel (e.g., *bracelet* /braslɛ/ 'bracelet'): whereas pre-readers recognize the schwa variant more slowly than the non-schwa variant, readers recognize the two variants equally quickly. Racine and colleagues conclude that spelling influences readers' word recognition and that pre-readers' word recognition is only influenced by frequency of occurrence.

We investigated in a more direct way whether listeners are sensitive to variants' frequencies. We first asked participants to perform an auditory lexical decision task in which French schwa words with optional schwa (e.g., *revue*, which can be pronounced as /ʀəvy/ or /ʀvy/) were presented with or without schwa. We then asked them to perform a relative frequency estimation task where they, for each visually presented word, were instructed to indicate the relative frequencies of the two variants based on what they hear in everyday life (following e.g., Bürki et al., 2010; Racine, 2008; Ranbom & Connine, 2007).

We tested native speakers of French (Experiment 1) as well as learners of French (Experiments 2, 3, and 3a in Appendix C), who we expect to have encountered reduced pronunciation variants, of at least some words, less often than natives. It has been shown that there is a mismatch between what tends to be presented to language learners and the actual use of that language outside the classroom (see e.g., Jones & Ono, 2000; McCarthy & Carter 1995; see for French Askildson, 2008; Fonseca-Greber & Waugh, 2003; Waugh & Fonseca-Greber, 2002; O'Connor di Vito, 1991). As a consequence, the relative frequency of the unreduced and the reduced variant of a given word is likely to be different for language learners than for native listeners. We hypothesize that learners' accuracies and response times (henceforth RTs) are better predicted by their own relative frequency ratings of the unreduced and reduced variants of a given word than by the natives' ratings and that is especially true for learners who have had some exposure to both variants.

Advanced learners are likely to have encountered the reduced variants of a (large) number of words. If they have encountered a given reduced variant, we hypothesize that they have stored this variant together with its frequency. Importantly, this frequency is expected to be different from the frequency stored by natives, as advanced learners have had less and different exposure to casual speech. As a consequence, the relative frequencies of the unreduced and reduced variants that hold for native speakers may not correlate well with advanced learners' accuracies and RTs. These dependent variables are predicted to correlate better with the frequency ratings provided by advanced learners themselves.

Beginner learners are only likely to have encountered the reduced variants of a small number of words. Moreover, the frequencies of occurrence of the variants in their input differ from those stored by natives. Thus, we hypothesize that beginner learners' response accuracies and response times also do not correlate with natives' relative frequency ratings. In contrast, because beginner learners and advanced learners received somewhat similar input, their performance may show a (weak) correlation with advanced learners' relative frequency ratings.

It is difficult to predict whether beginner learners' performance will correlate with their own relative frequency ratings. We might expect this correlation because these ratings should reflect beginner learners' input. However, beginner learners' ratings may be unreliable for many words because these learners have very little experience with these words' variants. They may therefore be just guessing. Their relative frequency ratings then do not correlate with their response accuracies and RTs.

We rely on subjective frequency ratings instead of objective frequency measures. Some researchers assume that subjective frequency ratings better predict lexical processing than objective frequency ratings do (e.g., Balota, Pilotti & Cortese, 2001). Others claim the opposite, at least for word frequency. For instance, Ghyselinck, Lewis, and Brysbaert (2004) claim that subjective ratings of word frequency are likely to be influenced by the age the word was acquired. They state that objective word frequency measures are to be preferred above subjective frequency measures if the word frequency measures can be based on reliable large data bases. This statement is supported by Ernestus and Cutler (2015), who showed that a subjective word frequency measure predicts RTs in a Dutch lexical decision task less accurately than several measures of objective word frequency.

We nevertheless relied on subjective relative frequency measures for several reasons. First of all, no objective measures for word pronunciation variants without schwa are available. This holds for French native speakers and above all for language learners. Second, Racine (2008) obtained frequency estimations from Swiss francophone students that correlated well with the variants' frequencies in the productions of 16 different Swiss francophone students who were asked to summarize stories ($r = .44$, $p < .01$). This result demonstrates that native speakers of French are able to provide reliable estimations of the frequencies of occurrence of the schwa and non-schwa variants of French words.

While the effect of relative frequency can thus only be tested with subjective ratings as a predictor for language behaviour, the effect of word frequency can be tested with predictors representing objective counts taken from several databases (e.g., Lexique 3.80, New, Pallier, Ferrand & Matos, 2001). This objective word frequency information may, however, only be representative for the language user group from whose speech these counts were extracted (typically native adult speakers). For other groups, subjective word frequency information may outperform objective word frequency measures in predicting language processing. This study also tests this hypothesis.

We tested learners of French who are native speakers of Dutch. The Dutch phonological system is highly comparable to that of French: the two phoneme inventories are about the same, and Dutch also shows schwa reduction. We would therefore expect that Dutch learners have little difficulty identifying the phonemes in the French stimuli of the experiment and that it will therefore be easy to test the effect of reduction and the role of relative frequency.

We tested native listeners in Experiment 1 to see whether their accuracies and RTs were predicted by their relative frequency ratings. Experiment 2 tested Dutch advanced learners of French (C1-C2 levels according to the Common European Framework of Reference for Languages, CEFR, Council of Europe, 2011), while Experiment 3 tested Dutch beginner learners (B1-B2 levels according to CEFR). We investigated whether the accuracies and RTs in the lexical decision task of the learner groups were better predicted by their own relative frequency ratings of the variants or by those provided by one of the other listener groups. Moreover, because we expect a discrepancy between the words' frequencies as listed in traditional databases, such as Lexique 3.80 (New, Pallier, Ferrand & Matos, 2001), and those experienced by beginner learners, we also tested in Experiment 3, whether the objective word frequency holding for native listeners or the subjective frequency measure holding for the specific listener group better predicts their accuracies and RTs.

4.2 Experiment 1

4.2.1 Methods

4.2.1.1 Participants

Thirty-six native speakers of French from Paris (three males), aged between 19 and 30 years, were paid to participate in the experiment. They were born and raised in the north of France, and their parents were also native French speakers. Like all participants tested for this study, they did not report any hearing problems and did not guess the purpose of the experiment.

4.2.1.2 Materials

We selected 44 French, morphologically simple, bisyllabic target words with a schwa in the first syllable (see Appendix A for the complete list). The target words were selected from vocabulary lists in teaching methods used at Dutch secondary schools. In standard French, these words can be produced with and without schwa in the initial syllable. The absence of schwa results in an illegal consonant cluster in 38 of the 44 words (e.g., /ʀn/ in /ʀnaʀ/ 'fox'). All words tested in this study were presented with their definite determiners.

We added several types of filler words (520 in total, see Table B.1 in Appendix B). First, we created 44 bisyllabic pseudo words with schwa in the first syllable. These pseudo words consisted of real French syllables and did not closely resemble real Dutch words. We determined their definite determiners, *le* or *la*, by adhering to the broad trend that French nouns ending in *-esse*, *-ie*, *-ite*, *ine*, *-té*, *-ure* are feminine, while nouns ending in *-age*, *-aire*, *-al*, *-at*, *-eau*, *-euil, -in*, *-is* are masculine.

Second, we added 149 bisyllabic real words without schwa, in order not to draw the participants' attention to the target words. Lexique 3.80 (New, Pallier, Ferrand & Matos, 2001), a list of 135,000 French words, shows that words with schwa in their initial syllable form only 2.8% of the bisyllabic French words. We further added 149 corresponding bisyllabic pseudo words without schwa. These pseudo words were constructed in the same way as the 44 pseudo words with schwa.

Finally, we selected 89 monosyllabic real words without schwa and we created 89 monosyllabic pseudo words. Due to these filler words, only a few (3.9%) of the monosyllabic real words presented to a given participant resulted from the absence of schwa.

A female French native speaker, aged 21 years, recorded all words in a sound-attenuated booth. She produced the words containing a schwa once with this schwa (unreduced) and once without (reduced). We used Adobe Audition 1.5 and a Sennheiser ME 64 microphone. The stimuli were digitised at a sampling rate of 44.1 kHz, a 16-bit quantization, and were scaled to an average intensity of 70 dB. The unreduced variants of the target words and their determiners (e.g., /laʀəvy/ *la revue* 'the magazine') had a total mean duration of 771.70 ms and the word-medial schwa had a mean duration of 107.07 ms (range: 61 ms - 158 ms). The reduced variants of the target words and their determiners (e.g., /laʀvy/ *la revue* 'the magazine') had a total mean duration of 725.11 ms. In these variants, word-medial schwa was absent, except in the word *seconde*, which contained a short schwa (12.86 ms). Due to absence of schwa, the onset of the noun formed the syllable coda of the determiner (e.g., reduced *la revue* /la. ʀə.vy/ was syllabified as /laʀ.vy/). The schwa in the determiner *le* had a mean duration of 118 ms if followed by an unreduced noun versus 116 ms if followed by a reduced noun. The mean /a/ duration in the determiner *la* was 119 ms if followed by an unreduced variant and 115 ms if followed by a reduced variant.

For the auditory lexical decision experiment, we created four different randomised lists, each containing all 564 words. Each list contained half of the 44 target words and half of the 44 bisyllabic pseudo words in their unreduced variants (i.e., with schwa), and the other halves in their reduced variants. Subsequently, we created the mirror images of these four lists by replacing the reduced variants by the corresponding unreduced variants and vice versa. This resulted in eight lists. Each participant heard one of these lists. The 564 tokens in a list were divided in three equal blocks, which were separated by short breaks in the experiment.

Each list was preceded by eight trials which familiarised the participants with their task. These trials presented three real schwa words, of which two were reduced and one unreduced, and three pseudo words, one with and two without schwa, and two real words without schwa. None of these words occurred in the main experiment. The familiarisation trials were presented in one of two different orders.

The rating experiment only contained the target words, which were presented in alphabetical order. This task also started with a familiarization trial, presenting a schwa word not occurring in the experiment.

4.2.1.3 Procedure

Experiment 1 took place at the Laboratoire Charles Bruneau of the *Institut de Linguistique et Phonétique Générales et Appliquées* (ILPGA) in Paris. Participants were tested individually in a sound-attenuated room. A session consisted of both the auditory lexical decision task and the relative frequency estimation task, always in that order. The two tasks were presented in E-prime 2.0 (Schneider, Eschman & Zuccolotto, 2007) from a laptop. The entire session lasted approximately 45 minutes. In the auditory lexical decision task, the participants heard the stimuli via Sennheiser HD 215 headphones. They were instructed to indicate as quickly as possible for each stimulus whether it was a real French word or not. Right handed participants responded by pressing the key *m* for *yes*, and the key *z* for *no*, on the keyboard. For left-handed participants, the *yes* and *no* keys were reversed. Each trial started with an asterisk shown for 250 ms in the center of the screen. Then, the stimulus was presented auditorily. Participants had to respond within 3000 ms from word onset. After the participant had responded, or after 3000 ms if no answer was given, the screen remained blank for another 300 ms and then the asterisk appeared announcing the next trial. After each block of approximately 10 minutes, the participants took a short break.

For the unspeeded relative frequency estimation task, participants were instructed to rate, for each word, the frequency of the reduced variant compared to the unreduced variant, based on what they commonly hear in daily life. They were asked to choose a value on the Likert scale presented on the computer screen by pressing a number between 1 and 6 on the keyboard. A score of 1 meant that the word was always produced with schwa, whereas a score of 6 meant that it was always produced without schwa. The two pronunciation variants were visually presented at the extreme left end (unreduced variant) and extreme right end (reduced variant) of the scale. The words occurred out of context, but with their definite determiners, as in the lexical decision task.

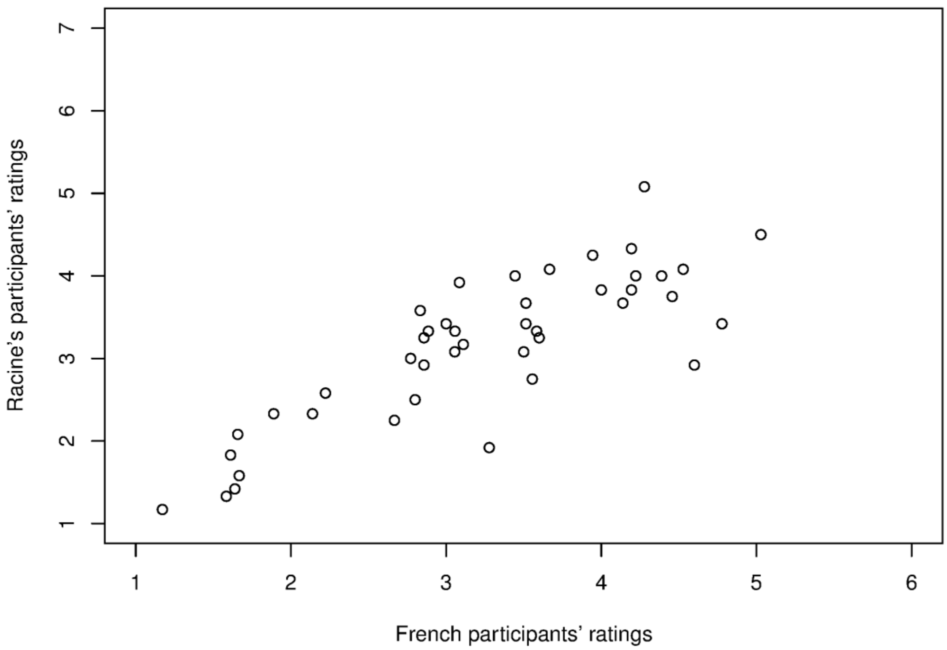
4.2.2 Results

4.2.2.1 Relative frequency estimation task

We first examined the validity of the relative frequency estimations (a total of 1568 due to 16 missing ratings; 99.0% of the trials). For each participant, we calculated a frequency value for each variant of every word by attributing the value entered by the participant to the reduced variant of the word, and subtracting this value from seven to calculate the value for the unreduced variant of that word. For example, if a participant selected 2 on the keyboard for the word *le chemin* 'the way' (meaning that the word was almost always produced with schwa), the reduced variant was assigned the value 2 and the unreduced variant the value 5. The interrater agreement, calculated on the basis of the ratings for the reduced variants, appeared to be slight (*Fleiss' kappa*: .09, $p < .001$; Fleiss, Levin & Paik, 2003). The average by word relative frequency ratings ranged between 1.17 and 5.83. We compared our ratings to those provided by participants from Nantes obtained by Racine (2008) in a similar rating experiment. As illustrated in Figure 4.1, the correlation between the by word average relative frequency ratings for reduced variants given by our

French participants and by Racine's participants' is high (r = .83, p < .0001). This shows that two groups of speakers of more or less the same variant of French produce ratings that are very similar, which strongly suggests that speakers are able to perform the task.

Figure 4.1 The average relative frequency ratings for the reduced variants provided by the native listeners in Experiment 1 plotted against the average relative frequency ratings provided by Racine's (2008) participants. A score of 1 meant that the reduced variant hardly ever occurs, while a score of 6 (for our participants) or 7 (for Racine's participants) meant that the word was always produced without schwa. Every dot represents a word.



4.2.2.2 Lexical decision accuracy

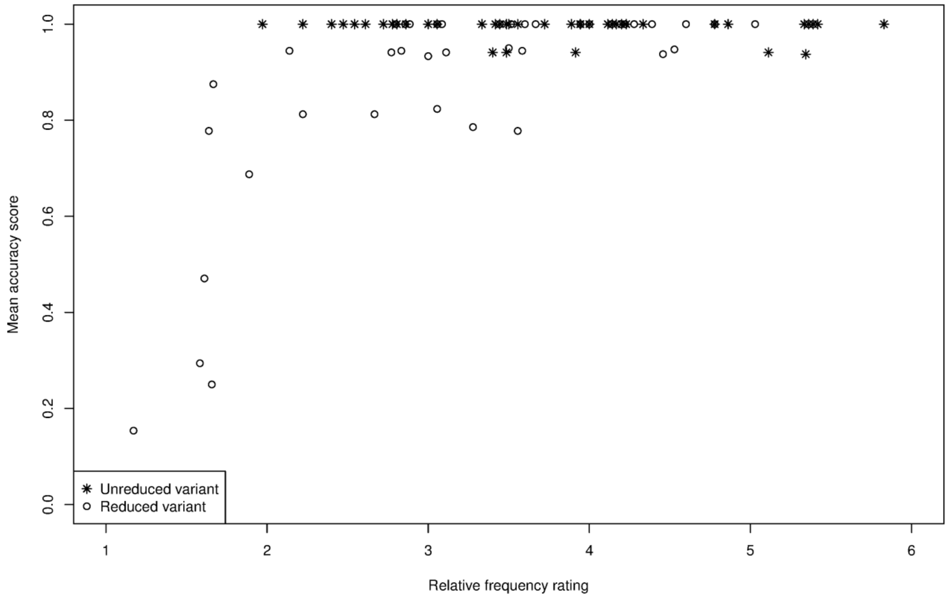
We examined the accuracy of the responses to the target words in the lexical decision task and investigated whether relative frequency rating was a significant predictor. Here and for the other analyses in this chapter, we only report models including relative frequency ratings averaged over participants and not models including the individual ratings as predictor because nearly all these latter models had higher AIC values. We excluded trials where the computer had not registered responses as well as trials with words for which the participant had not entered frequency ratings. The number of responses left for analysis totalled 1558 (98.4% of the data, of which 91.9% correct and 8.1% incorrect).

In Figure 4.2, the natives' mean accuracy score per word variant is plotted against their average relative frequency rating for that variant. The plot shows that the higher the relative frequency rating for the reduced word variant, the higher the French natives' accuracy score for this variant.

This effect of relative frequency on accuracy is supported by statistical analyses. We analyzed the accuracy of the answers by means of logistic mixed-effects models (Faraway, 2006) in R (R Development Core Team, 2015), including the package lme4 (Bates, Maechler, Bolker & Walker, 2015). We included word and participant as crossed random effects, and tested variant relative frequency averaged over participants and variant type (reduced versus unreduced) as the main fixed predictors. We also added several control predictors in order to reduce the variance in the data: trial number and word frequency as listed in Lexique 3.80 for subtitles of films (New et al., 2001). Random slopes were tested for all fixed predictors. We only retained those predictors in the model that were statistically significant or figured in statistically significant interactions. The final model is summarized in Table B.2 (see Appendix B).

French native listeners made fewer errors for variants with higher average relative frequency ratings. Furthermore, they made more errors when responding to reduced than to unreduced variants (see the upper panel of Figure 4.8). As shown by the random slope, the effect of variant type on accuracy differed per word. While the plot suggests that the effect of average relative frequency rating is restricted to the processing of reduced variants, the statistical analysis does not show an interaction between variant type and average relative frequency rating, possibly because of lack of statistical power.

Figure 4.2 The French natives' average relative frequency ratings for the reduced variants (open circles) and for the unreduced variants (asterisks) plotted against the French natives' mean accuracy scores.

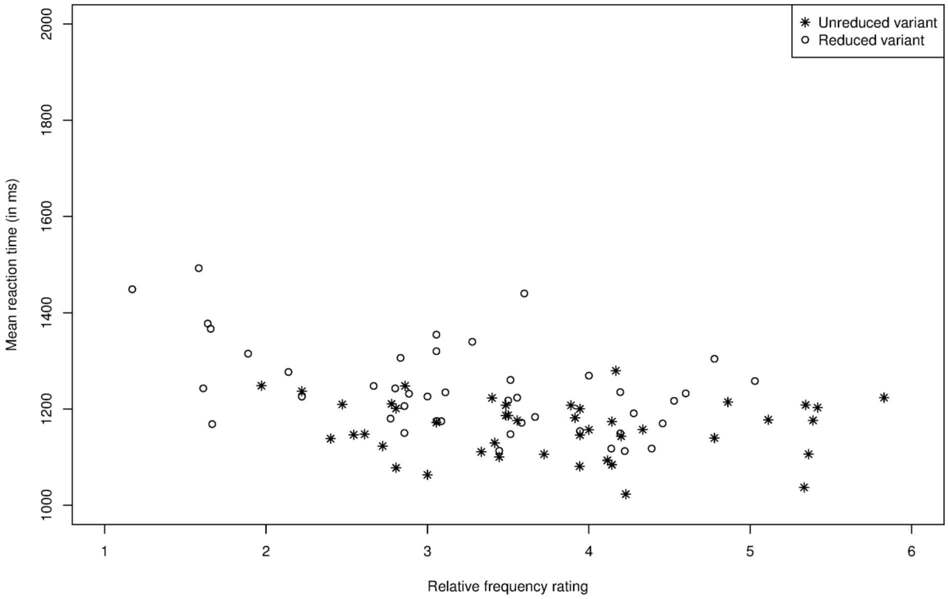


4.2.2.3 Lexical decision RTs

We analyzed the RTs measured from word onset for the trials in the dataset used for the accuracy analysis, discarding trials where the target words had incorrectly been classified as pseudo words (this left 1456 trials, 93.5%). RTs deviating from the grand mean (1214.22 ms) by more than 2.5 times the standard deviation (282.79 ms) were considered to be outliers and removed. The number of observations left for analysis totalled 1390 (87.8% of the data).

In Figure 4.3, the natives' average RT per word variant is plotted against the relative frequency rating averaged over participants for that variant. We see that the higher the relative frequency rating for the (reduced) word variant, the faster the French natives responded.

Figure 4.3 The French natives' average relative frequency ratings for the reduced variants (open circles) and for the unreduced variants (asterisks) plotted against the French natives' mean response times.



We performed statistical analyses to see whether the effect of the average relative frequency rating is statistically significant. We analyzed the log-transformed RTs by means of mixed-effects regression models, with word and participant as crossed random effects and with the same fixed predictors as those used to analyze accuracy. We added two more control predictors: the log-transformed duration of the target word and the log-transformed RT to the previous stimulus. Since the unreduced pronunciation variants were longer (mean duration of the determiner - word combination: 771.70 ms) than the reduced pronunciation variants (mean: 725.11 ms), we orthogonalised target duration and variant type by replacing target duration with the residuals of a linear model predicting target duration by variant type. Random slopes were again tested for all fixed predictors. We only retained those predictors in the model that were statistically significant or figured in statistically significant interactions. Finally, we discarded all RTs that deviated more than 2.5 times the standard error from the values predicted by the best statistical model and refitted the model. The final model is summarized in Table B.3 in Appendix B.

The French native listeners responded significantly faster to shorter words, to stimuli presented later in the experiment, and if they had also responded quickly to the previous trial. More interestingly, we found a significant effect of variant type on RT: the French native listeners took more time to respond to a reduced variant than to an unreduced variant (see the lower panel of Figure 4.8). This effect of variant type was larger for some words than for others, as shown by the significant random slope of variant type for word.

Most importantly with regard to our research question, we found a main effect of relative frequency rating. The French native listeners responded faster to variants that according to this group’s average relative frequency ratings occur more often.

4.2.2.4 Summary and discussion

The average relative frequency ratings provided by the French natives predict both their accuracy and RTs in a lexical decision task. These results support earlier findings suggesting that listeners show sensitivity to the frequency of occurrence of a given pronunciation variant (e.g., Pitt et al., 2011; Ranbom & Connine, 2007). These frequencies must then be stored, which strongly suggests that the pronunciation variants are lexically stored. Furthermore, Experiment 1 provides additional evidence for a privileged status for unreduced variants (e.g., Pitt et al., 2011; Ranbom & Connine, 2007; Tucker & Warner, 2007; Ernestus & Baayen, 2007; Janse, Nooteboom & Quené, 2007): the native listeners made fewer mistakes and responded faster to unreduced variants than to reduced variants.

In order to discover whether participant groups with different experience with the reduced and unreduced variants of a word show behaviour reflecting this difference in experience, we replicated Experiment 1 with Dutch advanced learners of French. Dutch advanced learners of French have less experience with the reduced variants of most words presented in our experiment. As a consequence, we expect that their recognition accuracy and response times better correlate with their own variant frequency ratings than with those provided by the French natives.

The learners performed two additional tests to reveal their proficiencies. The first is the visual lexical decision task LexTALE (Brysbaert, 2013; Lemhöfer & Broersma, 2012), which provides a measure of general lexical proficiency. The second was a dictation task, which provides a good indication of learners' comprehension of connected speech (Kennedy & Blanchet, 2014).

4.3 Experiment 2

4.3.1 Methods

4.3.1.1 Participants

Forty-seven Dutch undergraduate students of French (aged 19-30 years; eleven males) participated in the experiment. All were born and raised in the Netherlands, had taken French classes for five or six years at secondary school, and had studied French at university for at least seven months and at most three years and seven months. Their CEFR-levels roughly corresponded to C1-C2 level (CEFR, Council of Europe, 2011).

4.3.1.2. Materials

The materials for the auditory lexical decision task and the relative frequency estimation task were the same as those used in Experiment 1.

The LexTALE task, the first proficiency test that we administered, includes 56 real French words and 28 phonotactically legal pseudo words.

The dictation task consisted of 10 sentences (137 words in total), extracted from an informal conversation between two men in The Nijmegen Corpus of Casual French (Torreira, Adda-Decker & Ernestus, 2010). The sentences were produced at an average speech rate of 5.57 syllables per second (excluding pauses) and contained high-frequency words, in which many schwas were absent.

4.3.1.3. Procedure

Dutch advanced learners of French were tested in sound-attenuated booths at the Max Planck Institute for Psycholinguistics in Nijmegen and at Leiden University. The procedures of the lexical decision experiment and the rating experiment were the same as in Experiment 1. These two experiments were followed by the two proficiency tests.

The visual lexical decision task LexTALE was run in E-prime, as were the two main experiments. Each trial started with a blank screen shown for 250 ms and once the stimulus was visually presented (in an 18-point Courier New font size), participants had to respond within 3000 ms. As soon as the participant had responded (or after 3000 ms), a blank screen appeared for 100 ms, after which the next trial started.

The dictation task was presented via WebExp2, a web-based program (Keller, Gunasekharan, Mayo & Corley, 2009). Participants listened to the speech to be transcribed orthographically, sentence by sentence. They could replay a sentence as often as necessary. However, after a certain amount of time (range: 10-120 seconds), determined on the assumption that one minute of speech takes approximately ten minutes to transcribe, participants were forced to start on the next sentence. The entire experimental session lasted approximately 60 minutes.

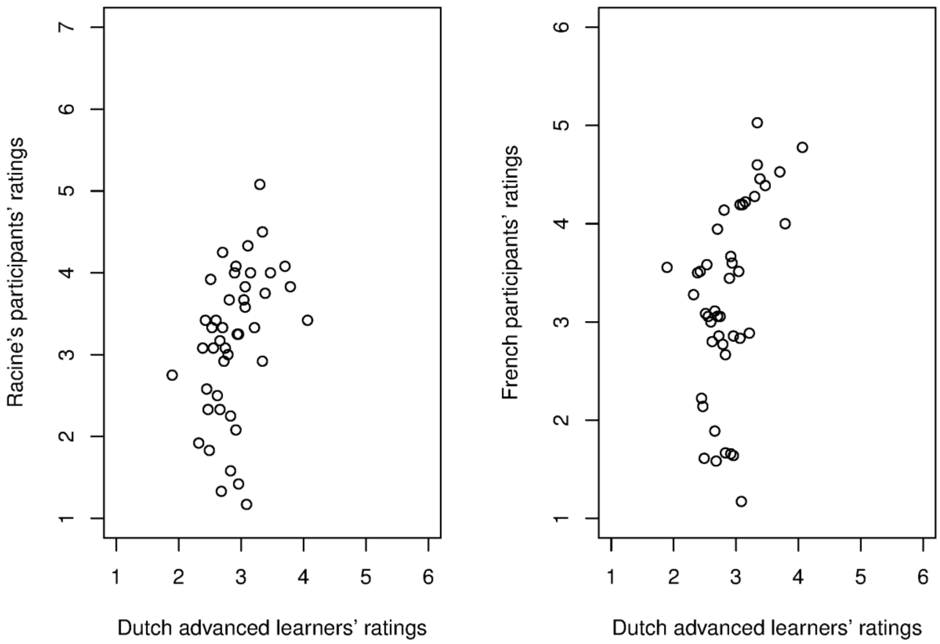
4.3.2 Results

4.3.2.1 Relative frequency estimation task

We obtained 2068 relative frequency ratings, which showed a slight interrater agreement (on the basis of the ratings for the reduced variants *Fleiss' kappa*: .01, $p < .001$). The average by word relative frequency ratings ranged between 1.89 and 4.06 and showed a correlation of .40 ($p < .01$) with the French participants' average ratings obtained by Racine (2008), and of .48 ($p < .001$) with those obtained in Experiment 1 (Figure 4.4).

Interestingly, the advanced learners' relative frequency ratings averaged by word correlated with the word frequencies from Lexique: the more often a word occurs according to the Lexique database, the higher the average relative frequency rating for its reduced word pronunciation variant (and thus the lower the relative frequency rating for its unreduced variant). Possibly, the advanced learners applied the strategy to partly base their relative frequency ratings on the words' frequencies.

Figure 4.4 The average relative frequency ratings for the reduced variants provided by the advanced learners plotted against the average relative frequency ratings provided by Racine's participants (2008, left panel) and by the native listeners from Experiment 1 (right panel). A score of 1 meant that the reduced variant hardly ever occurs, while a score of 6 (for our participants) or 7 (for Racine's participants) meant that the word was always produced without schwa. Every dot represents a word.



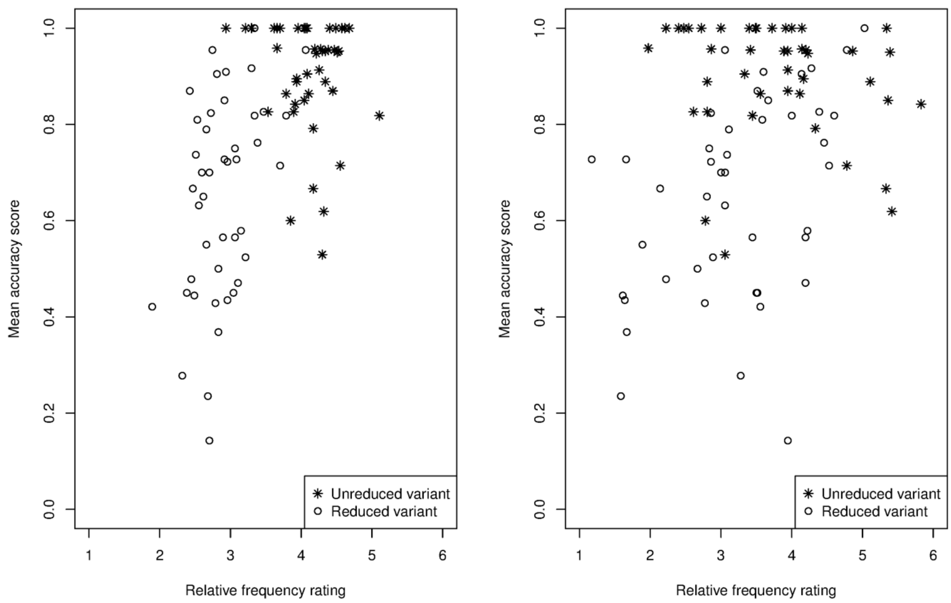
4.3.2.2 Lexical decision accuracy

We obtained responses for 2039 trials (98.6% of the data). In the left panel of Figure 4.5, the Dutch advanced learners' mean accuracy score per word variant is plotted against their mean relative frequency rating for that variant. Like for the French natives, for the Dutch advanced learners, the accuracy scores for the (reduced) word variants increase with the group's average relative frequency ratings.

We also plotted the Dutch advanced learners' accuracy score per word variant against the French natives' average rating for that variant (right panel of

Figure 4.5). We see that the accuracy score for the word variant also increases with this average relative frequency rating for the variant, although the correlation seems smaller in the right panel than in the left panel of Figure 4.5.

Figure 4.5 The advanced learners' mean accuracy scores plotted against the average relative frequency ratings for the reduced variants (open circles) and for the unreduced variants (asterisks) as provided by the advanced learners (left panel) and by the French natives (right panel).



We tested whether the correlation between accuracy score and average relative frequency rating visible in the two figures is statistically significant. We first analyzed the effect of the average relative frequency ratings provided by the advanced learners themselves. We analyzed response accuracy (76.7% of the responses were correct, 23.3% incorrect) by means of logistic mixed-effects models using the same fitting procedure and using the same random and other fixed predictors (i.e., variant type, trial number, word frequency) as in the analysis of the accuracy data of Experiment 1. In addition, we included the percentage of correct words in the orthographic dictation task as a predictor indicating proficiency. We did not include a predictor for proficiency based on the LexTALE task because due to a technical error the data from this task were not available for 21 participants. Because we are interested in variation in the advanced learners' ratings that is not based on their knowledge of word frequency but on exposure to reduced pronunciation variants, we removed the variation in the frequency ratings that can be explained by the variation in word frequency. That is, we replaced the average ratings provided by the advanced learners with the residuals of a linear model predicting average ratings by variant type, the word frequencies from Lexique, and their interaction. Table B.4 (see Appendix B) shows the final model.

The advanced learners made fewer errors to words that occur more often and if they performed better on the dictation task. Like the participants in Experiment 1, they made more errors when responding to reduced variants relative to unreduced variants (see the upper panel in Figure 4.8). As shown by the random slope, this effect was not equally large for all participants.

Importantly, the accuracy scores were also predicted by the average relative frequency ratings from the advanced learners: the higher the average rating for a given variant, the fewer errors they made. In contrast, the accuracy scores could not be predicted by the average ratings provided by the French native listeners, in contrast to what the right panel of Figure 4.5 suggests.

We directly compared whether the effect of relative frequency rating differed between the advanced learners and the native listeners. We pooled the data from the two participant groups and investigated the effect of the natives' average relative frequency ratings and of the advanced learners' average relative frequency ratings. The model with the lowest AIC (2257.8 versus 2260.8;

Akaike, 1973) contained the French participants' average ratings as a simple effect and in interaction with participant group (main effect: *ß* = -0.826, *z =* -3.853, $p < .001$; interaction: *ß* = 0.865, $z = 4.271$, $p < .001$). The interaction shows that the native average relative frequency ratings only predict the natives' accuracies, and not those of the advanced learners. There were no other significant interactions with participant group.

4.3.2.3 Lexical decision RTs

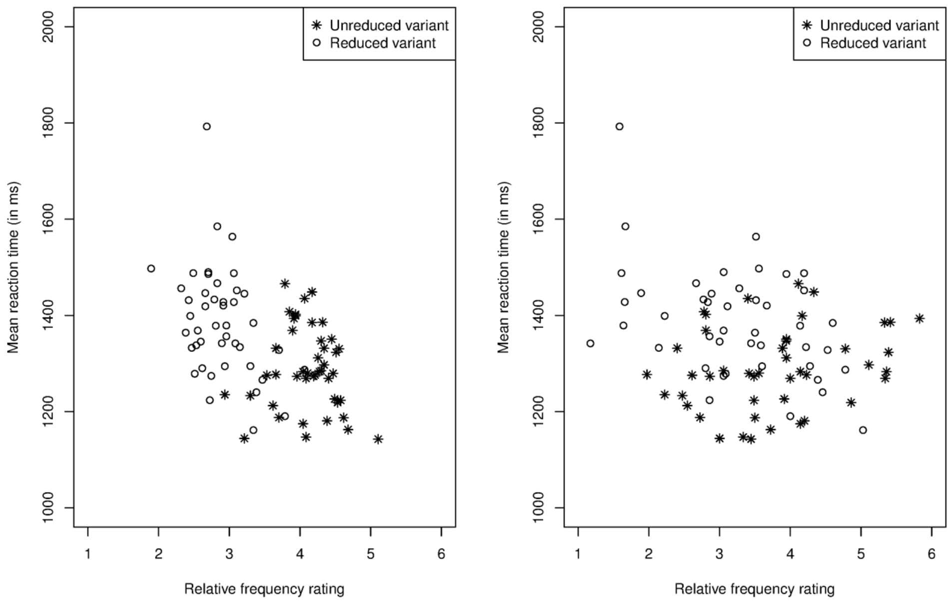
The analyses of the lexical decision RTs were based on correct responses only (1564 trials, 76.7% of the data), and on RTs that were within 2.5 standard deviations (335.91 ms) of the grand mean (1386.53 ms). The number of observations left for analysis totalled 1473 (72.2% of the data).

The left panel of Figure 4.6 presents the Dutch advanced learners' mean RTs per word variant as a function of the relative frequency ratings averaged over participants for that variant. The higher the average relative frequency ratings for the word variants, the faster the advanced learners responded.

The right panel of Figure 4.6 plots the average advanced learners' RTs against the French natives' average relative frequency ratings. The plot does not show a clear relationship between the two.

We performed statistical analyses to further investigate these observations. We analyzed the log-transformed RTs by means of mixed-effects regression models, following the same fitting procedure as in Experiment 1. We tested the same predictors as for the accuracy analysis of this experiment, in addition to log-transformed target word duration and log-transformed previous RT. The final statistical model is summarized in Table B.5 (see Appendix B).

Figure 4.6 The advanced learners' mean response times plotted against the average relative frequency ratings for the reduced variants (open circles) and for the unreduced variants (asterisks) as provided by the advanced learners (left panel) and by the French natives (right panel).



The Dutch advanced learners responded faster to words with higher frequencies of occurrence. Like the French natives in Experiment 1, they speeded up over the course of the experiment, they were faster if they had responded more quickly to the previous trial, and, more interestingly, responded faster to unreduced than to reduced variants (see the lower panel in Figure 4.8). The random slope of variant type by word shows that this latter effect was larger for some words than for others. Finally, the advanced learners responded faster to word variants that, according to the average of the whole group of students, occur more often. A different model with the ratings provided by the French native listeners as predictor showed that this predictor does not reach significance.

We then combined the advanced learners' RTs with those provided by the native listeners to investigate whether the predictors showed different effects for the advanced learners and the natives. The best model included a main effect of the average relative frequency ratings by the French participants (AIC = -2514.6). This model also contains two interactions with participant group. The first is expected, given the main effects in the separate models; it is the interaction with word frequency (main effect: $ß = -0.00305$, $t = -0.614$; interaction: $ß = -0.0265$, $t = -3.474$), which shows that the word frequency effect is substantially larger for the advanced learners than for the French natives. The second interaction is less expected. We found a group by target duration interaction, which shows that target duration effects were larger for the natives than for the learners (main effect: $ß = 0.235$, $t = 6.207$; interaction: $ß = -0.143$, $t = -2.538$).

4.3.2.4 Summary and discussion

The advanced learners' average relative frequency ratings predict both their accuracy and RTs. Importantly, the learners' accuracies were not predicted by the natives' average ratings. This strongly suggests that how well a listener recognizes a pronunciation variant is determined by how much exposure this listener has had with the given variant and thus that advanced learners have lexical representations for these variants specified for their variants' frequency in their own exposure.

In Experiment 3, we tested beginner learners. These learners may have encountered the reduced variants of only a few words, and with different frequencies than native listeners. We therefore expect that they have only lexically stored a few reduced variants. Moreover, the variants that they have stored are specified for these variants' frequencies of occurrence in the learners' own input rather than in the natives' input.

We therefore do not expect a correlation between the learners' performance and the natives' relative frequency ratings, as these ratings reflect the variant frequencies in the natives' input rather than in the beginner learners' input. The advanced learners ratings may better reflect the beginner learners' input, and there may therefore be a (weak) correlation between the beginner learners' performance and the advanced learners' variant frequency ratings.

We also tested whether the ratings of the relative frequencies of the unreduced and reduced variants provided by the beginner learners themselves predict their performance. On the one hand, following the rationale of Experiments 1 and 2, we might expect that this is the case since these ratings would reflect the beginner learners' input. On the other hand, the beginner learners' ratings may be unreliable for many word variants because they have very little experience with these variants. These learners may therefore be just guessing. Their relative frequency ratings then do not correlate with their response accuracies and RTs.

Finally, we investigated in Experiment 3 which measure of word frequency best predicts beginner learners' performance. Since their input is very different from the input that a native listener receives, the word frequencies listed in Lexique 3.80 (New et al., 2001) for film subtitles probably do not well reflect the frequencies with which these learners encounter words. The learners' ratings of these word frequencies may therefore better predict their performance than the word frequencies extracted from Lexique.

4.4 Experiment 3

We performed two experiments with Dutch beginner learners. One experiment was an exact replication of Experiment 2 (Appendix C; Experiment 3a). In the other experiment (Experiment 3, reported below), we changed the instruction for the relative frequency estimation task. The participants were explicitly asked to rate the variants' frequencies in their own language input instead of in everyday French. This change in instruction, however, did not substantially affect the results. We therefore only report here the results of Experiment 3 and describe Experiment 3a in Appendix C, focusing on the differences between the two experiments[1]. Experiment 3 also incorporated a word frequency rating task.

4.4.1 Methods

4.4.1.1 Participants

Fifty-six Dutch university students (aged 18-26 years; ten males) were paid to participate in the experiment. All were born and raised in the Netherlands and had taken a maximum of 2 hours of French classes a week, for a maximum of six years at secondary school. They had spent no more than 15 days a year in France, and were not exposed to French films, books, poems or music on a weekly basis. Their proficiency levels roughly corresponded to the B1-B2 level (CEFR, Council of Europe, 2011).

To verify whether these Dutch beginner learners are able to perceive the difference between the reduced and unreduced variants of French schwa words, we asked them to perform an AXB task. In this task, the participant heard three different tokens of the same word in a sequence (in every trial a different word) and had to say whether the second (X) token was more similar to the first

[1] We also conducted all analyses comparing the beginners with the other listener groups pooling the data from Experiments 3 and 3a, instead of with just the data from Experiment 3, in order to increase statistical power. The obtained results were similar to those reported in this section.

(A) or to the third (B) token. If the first token was unreduced, the third token was reduced and vice versa. There were 17 trials: ten trials with reduced tokens and seven trials with unreduced tokens in the middle of the sequence. Eighteen participants gave incorrect responses to all trials; it is very likely that they reversed the response buttons. We therefore interpreted their responses as correct responses. The complete dataset (96.8% correct) shows that these Dutch beginner learners of French are able to distinguish reduced from unreduced pronunciation variants of French schwa words.

4.4.1.2 Materials

The materials for the lexical decision task and the relative frequency estimation task were the same as in Experiment 2.

The word frequency rating task comprised all target words and 28 real word fillers.

4.4.1.3 Procedure

The experiment was conducted in sound-attenuated booths at the Max Planck Institute for Psycholinguistics in Nijmegen. The procedure was the same as for Experiment 2, except for the question asked with regards to the relative frequency rating task. We adapted our question to the Dutch translation of "How often have you heard the two pronunciation variants?".

After the relative frequency estimation task, we asked our participants to perform the word frequency rating task. This rating task started with two high frequency filler words (*le matin* 'the morning' and *la vie* 'the life') and two low frequency filler words (*le péage* 'the toll (gate)' and *la gendarmerie* 'the gendarmerie') to give the participant the opportunity to interpret the extreme values of the scale. The fillers were followed by six target words and then by four new fillers with extreme frequencies of occurrence. This combination of target and filler words was repeated until all target words were presented. To rate the frequency of each word, participants were asked to choose a value on the Likert scale presented on the computer screen by pressing a number between 1 and 7 on the keyboard. A score of 1 reflected that the word had a very low frequency, whereas a score of 7 reflected that the word had a very high frequency.

4.4.2 Results

4.4.2.1 Relative frequency estimation task

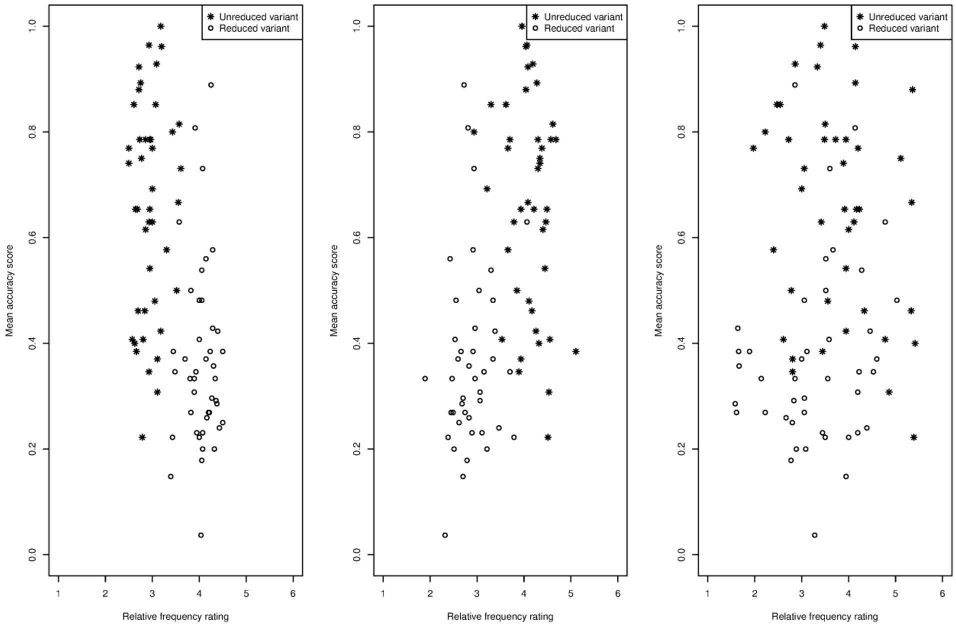
Due to a technical error the relative frequency ratings for the target word *la vedette* were not registered. We obtained 1151 ratings, which showed a very poor interrater agreement in terms of *Fleiss* (on the basis of the ratings for the reduced variants *Fleiss' kappa*: .002, $p > .1$). Although we again used a six-point Likert scale, all average relative frequency ratings were between 3.39 and 4.50. These average relative frequency ratings did not correlate with any of the other relative frequency ratings obtained in this study ($ps > .05$). Surprisingly, the reduced variants were rated as occurring more frequently than the corresponding unreduced variants (see the left panel of Figure 4.7).

4.4.2.2 Lexical decision accuracy

In the left panel of Figure 4.7, the beginner learners' mean accuracy score per word variant is plotted against their average relative frequency rating for that variant. The plot shows that the beginner learners' ratings do not predict the learners' accuracy score.

We further examined whether the beginner learners' mean accuracy scores could be predicted by the advanced learners' (middle panel) or the French natives' (right panel) average relative frequency ratings. The middle and right panels of Figure 4.7 suggest that if there is an effect of average relative frequency ratings of either group on beginner learners' mean accuracy, this effect is small.

Figure 4.7 The beginner learners' mean accuracy scores plotted against the average relative frequency ratings for the reduced variants (open circles) and for the unreduced variants (asterisks) as provided by the beginner learners (left panel), by the advanced learners (middle panel) and by the French natives (right panel).

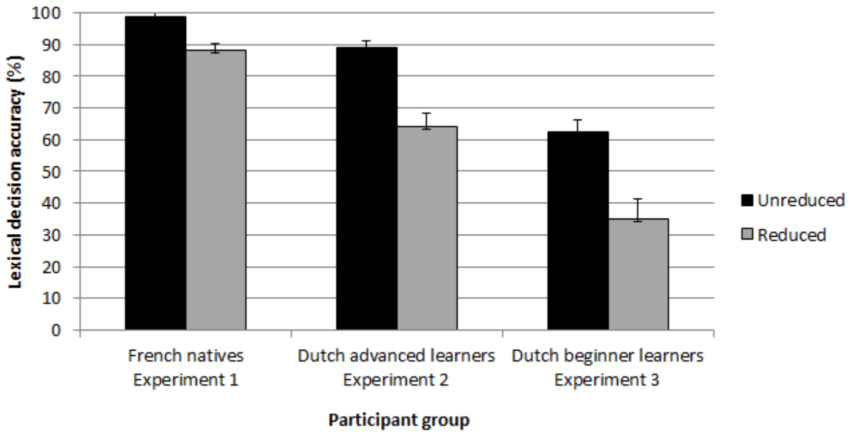


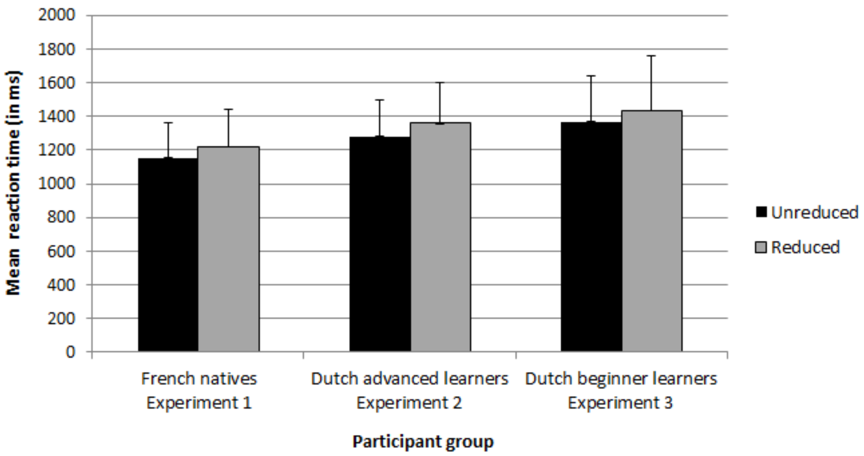
We performed statistical analyses to see whether the effect of any of the different average relative frequency ratings is, nevertheless, statistically significant. We analyzed the 2464 responses (of which 48.8% were correct) by means of logistic mixed-effects models, using the same random predictors as in the analysis of the accuracy data of Experiment 2. We created different models with the relative frequency ratings provided by the native listeners, the advanced learners, and the beginner learners tested in this experiment. In addition, we included as fixed predictors variant type, trial number, dictation task performance, and the Ghent score (the number of correctly classified real words minus twice the number of pseudo words classified as real words in the LexTALE task by the given participant, Brysbaert, 2013). Furthermore, we tested separate models in which we replaced the word frequencies taken from Lexique by the average word frequency ratings as indicated by the participants on a scale from one to seven. The average word frequency ratings did not correlate with the word frequencies in Lexique as reported for subtitles (Lexique 3.80, New et al., 2001, $p > .05$).

Table B.6 (see Appendix B) shows a summary of the best statistical model. Participants made more errors for reduced than for unreduced variants (see the upper panel of Figure 4.8), and again the effect of variant type was larger for some words than for others, as shown by the significant random slope of variant type for word.

Moreover, participants who were better at the dictation task made fewer errors. Participants made fewer errors for words that occur more frequently. Interestingly, we find this latter effect with both predictors reflecting word frequency (the one derived from Lexique and the one derived from the participants' ratings), but the predictor resulting in the best statistical model is based on the word frequency ratings provided by the participants. Most importantly, we did not find an effect of the average relative frequency rating of any participant group.

Figure 4.8 The top figure shows the lexical decision accuracy to unreduced and reduced words by French native participants (Experiment 1), by advanced learners (Experiment 2) and by beginner learners (Experiments 3). The error bars show standard deviations. The bottom figure shows the mean response times to unreduced and reduced words by French native participants (Experiment 1), by advanced learners (Experiment 2) and by beginner learners (Experiments 3). The error bars show standard deviations.





We combined the data from Experiment 3 with the data from Experiment 1 to investigate whether the beginner learners and the natives showed statistically significant differences. The best model contained an interaction between learner group and the relative frequency ratings averaged over the natives, demonstrating that only the natives responded significantly more accurately to variants that, according to them, on average, had higher relative frequencies (main effect: *ß* = 0.143, *z* = 2.183; interaction: *ß* = -0.217, *z* = -2.629).

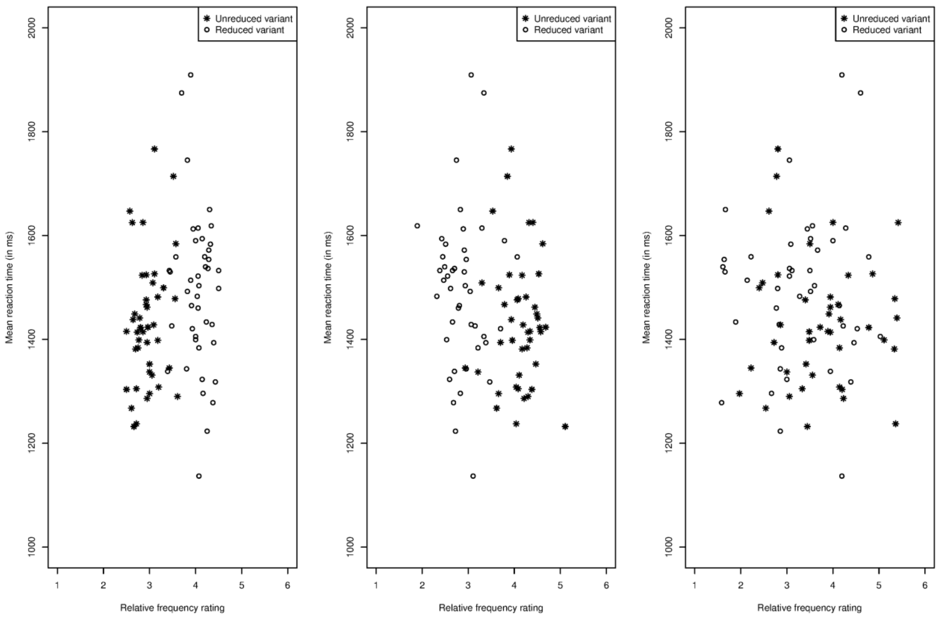
We further combined the data from Experiment 3 with the data from Experiment 2 in order to see whether there were statistically significant differences between the beginner and advanced learners. The best model contained an interaction between the relative frequency ratings averaged over the advanced learners and variant type, demonstrating that the advanced learners' ratings could only well predict the beginner and advanced learners' accuracy scores for the reduced variants (simple effect of relative frequency for unreduced variants: *ß* = -0.571, *z* = -1.818; $p > .05$; interaction with pronunciation variant: *ß* = 1.188, *z* = 2.0194, $p < .05$).

4.4.2.3 Lexical decision RTs

We restricted the analyses of the RTs to those trials where the participants had provided the correct responses (1202 trials, 48.8% of the data). RTs deviating from the grand mean (1442.57 ms) by more than 2.5 times the standard deviation (364.48 ms) were considered to be outliers and removed, which left 1145 observations (46.5% of the data). In the left panel of Figure 4.9, the beginner learners' mean RTs per word variant are plotted against their average relative frequency ratings for that variant. There seems to be no clear correlation between the two variables.

We also plotted the beginner learners' mean RTs per word variant against the advanced learners' (middle panel of Figure 4.9) and natives' (right panel of Figure 4.9) average relative frequency ratings per word variant. Again we see no clear relationship between the relative frequency ratings and the beginner learners' mean RTs.

Figure 4.9 The beginner learners' mean response times plotted against the average relative frequency ratings for the reduced variants (open circles) and for the unreduced variants (asterisks) as provided by the beginner learners (left panel), by the advanced learners (middle panel) and by the French natives (right panel).



This is confirmed by statistical analyses. We analyzed the log-transformed RTs by means of mixed-effects regression models, following the same fitting procedure as in Experiment 2. We tested the same predictors as for the accuracy analysis of this experiment, in addition to log-transformed target word duration and previous log RT. Table B.7 (see Appendix B) summarizes the best statistical model.

Participants responded more slowly to longer words and if they had responded slowly in the preceding trial. More interestingly, they also responded more slowly to reduced variants relative to unreduced variants (see the lower panel of Figure 4.8). As shown by the random slope, this effect of variant type was not equally large for all participants. No predictor reflecting the frequencies of the words emerged as significant. Importantly, none of the variant relative frequency ratings (neither the participants' own average ratings, nor the average ratings from the other groups tested in this study) predicted the participants' RTs.

We combined the data from Experiment 3 with the data from Experiment 1 to investigate whether the predictors showed different effects for the beginner learners and the natives. The best model contains an interaction between participant group and the average relative frequency ratings provided by the natives, demonstrating that only the natives responded significantly faster to variants that according to them, on average, had higher relative frequencies (main effect: $ß = -0.0194$, $t = -4.048$; interaction: $ß = 0.0288$, $t = 4.174$).

We further combined the data from Experiment 3 with the data from Experiment 2 in order to see whether there were statistically significant differences between the beginner and advanced learners. The best model contains an interaction between the average ratings provided by the advanced learners and learner group, showing that only the advanced learners responded significantly faster to variants that according to them, on average, had higher relative frequencies (main effect: $ß = -0.0301$, $t = -3.236$; interaction: $ß = 0.0274$, $t = 2.769$).

4.4.2.4 Summary and discussion

The beginner learners' accuracy scores and RTs can neither be predicted by the advanced learners' nor by the natives' average relative frequency ratings if the data from the beginners are analyzed just by themselves. However, when we combine the data from the beginner and advanced learners, the advanced learners' ratings appear to predict the beginner learners' accuracy. In contrast, the natives' ratings do not predict the beginner learners' accuracy in a combined dataset. This suggests that the frequencies for the reduced pronunciation variants stored in their mental lexicons differ between beginner learners and natives, while there may be some similarity between beginner and advanced learners.

Furthermore, we found that variant relative frequency ratings provided by beginner learners do not correlate with their own word recognition accuracy, nor with their RTs. There may be no variant frequency effect because the relative frequency rating task is too hard for beginner learners. Due to their low exposure to reduced pronunciation variants, they may not be able to provide different ratings for the different words, and may often just be guessing.

The participants in Experiment 3 also rated how often they encountered the target words (independently of the variant). This word frequency measure outperformed the word frequency measure based on Lexique in the accuracy analysis. This suggests that word frequency ratings better reflect learners' stored word frequencies than lexical databases based on language produced by native speakers.

4.5 General discussion

The aim of this study was to further investigate the hypothesis that listeners rely on lexical representations of reduced word pronunciation variants. These lexical representations would be specified for the variants' frequencies of occurrence. If so, how easily a listener recognizes a given word pronunciation variant should directly reflect the listener's exposure to this variant. We investigated this by exploring whether the recognition accuracy and response times for a word pronunciation variant in the lexical decision task correlate best with the frequencies of occurrence of that variant holding for the group to which the listener belongs.

We conducted an experiment with several groups of listeners. First, participants performed an auditory lexical decision task with schwa words that were pronounced with schwa (e.g., /laʀəvy/ *la revue* ‘the magazine’) or without schwa (e.g., /laʀvy/). Second, they indicated the relative frequencies of the two pronunciation variants of each word by choosing a value on a scale from 1 to 6. We examined whether the recognition accuracy and response times (RTs) in the lexical decision task correlated with the participants’ relative frequency ratings. We tested native listeners of French (Experiment 1) and beginner and advanced learners of French (Experiments 2, 3, and 3a in Appendix C).The learners of French are expected to have encountered reduced word pronunciation variants less often than the natives have, and the unreduced and reduced variant of a word may therefore have different relative frequencies for them than for natives.

The natives in Experiment 1 provided relative frequency ratings that, on average, correlated well (r = .83) with other francophone participants’ ratings collected by Racine (2008). This shows that two groups of speakers of more or less the same variant of French produce ratings that are very similar, which suggests that speakers are able to indicate their exposure to the two variants of a schwa word.

The Dutch advanced students’ ratings from Experiment 2 correlated with those provided by native listeners in Racine (2008) and those obtained in our Experiment 1, but these correlations were substantially lower (r = .40, .48, respectively) than the correlation between our group of native listeners and the natives in Racine (2008; r = .83). This suggests that advanced learners’ exposure to reduced variants is different from that of native listeners.

Interestingly, the advanced learners’ relative frequency ratings correlated with the word frequencies of the words listed in Lexique. A higher frequency in

Lexique correlated with a lower average relative frequency rating for the unreduced variants. This correlation is in line with the general finding that more frequent words tend to be more often reduced and to a greater extent (e.g., Pluymaekers, Ernestus & Baayen, 2005a; Jurafsky, Bell, Gregory & Raymond, 2000; Bybee, 1998). Possibly, the advanced learners are aware of this generalisation and applied it when providing their frequency ratings.

Experiment 3 tested beginner learners of French. It could be expected that these learners indicated that they knew most words only in their unreduced variants. This, however, was not the case: their average relative frequency ratings on a scale from 1 (the unreduced variant is the most frequent variant) to 6 (the unreduced variant hardly ever occurs) range between 3.39 and 4.50 (i.e., a range of 1.1). The ranges were larger for the advanced learners' ratings (2.17) and for the natives' ratings (4.66).

The fact that the beginner learners mostly chose values in the middle of the continuum indicates that these beginner learners hardly differentiated among words and may just have been guessing. They may have adopted a guessing strategy because they had encountered the vast majority of words (nearly) only in their unreduced variants and were thus unfamiliar with the reduced ones. They will not have lexically stored the reduced pronunciation variants that they had not encountered yet. The frequencies for the variants that they have encountered and stored will hardly differ among the words.

In the lexical decision task, the French native listeners responded faster and more accurately to (reduced) pronunciation variants with higher relative frequency ratings. The advanced learners' accuracies and RTs only correlated with the average frequency ratings provided by these learners themselves. These results suggest that advanced learners are sensitive to relative frequencies during word recognition, but that these frequencies differ from those for the native listeners. These findings support the hypothesis that the recognition of French reduced pronunciation variants is influenced by the listener's exposure to these variants.

The beginner learners' accuracies and RTs were neither predicted by the natives' nor by the advanced learners' ratings. Since beginner learners receive input that is very different from the input to natives, and that also deviates from the input to advanced learners, this finding is also in line with our hypothesis that listeners lexically store the variants' frequencies reflecting their own input.

The beginner learners' performance could neither be predicted by their own relative frequency ratings. We believe that this is the case because the beginner learners were just guessing in the relative frequency estimation task, as explained above, and that these ratings were thus unreliable. We assume that when reduced pronunciation variants are not stored at all or stored with only low frequencies, as is the case for the beginner learners, listeners may not or hardly use these lexical representations during recognition but use other mechanisms and accept, for instance, imperfect matches of the acoustic signal with the lexically stored unreduced variants.

Thus, we find differences between the listener groups in which relative frequency rating (if any) best predicts their performance in the lexical decision experiment. These differences are supported by the pairwise comparisons of the data from the different groups, which show statistically significant differences between natives and beginner learners in the effects of the ratings provided by the different groups on their accuracy and RTs, between beginner and advanced learners on their RTs, and between advanced learners and natives on their accuracy.

The beginner and advanced learners did not differ significantly in their sensitivity to relative frequency ratings provided by the advanced learners in the accuracy analysis. Similarly, the advanced learners and the natives did not differ in their sensitivity to relative frequency ratings provided by the natives in the RT analysis. The lack of significant differences in sensitivity to relative frequency ratings in the combined datasets have to be considered with care. There may simply be power issues, especially because the learner groups show large within-group variation (see, e.g., the error bars in Figure 4.8): some learners are better (or worse) than suggested by their group's average proficiency level, and therefore performed more similarly to participants in the higher (or lower) proficiency level group.

Our finding that native listeners and advanced learners are sensitive to frequencies of occurrence of reduced variants for French schwa words adds to previous research on the comprehension of pronunciation variants (e.g., Pitt et al., 2011; Ranbom & Connine, 2007) in several respects. First, while Ranbom and Connine documented listeners' sensitivity to the frequencies of word pronunciations resulting from the substitution of one segment by another, we show that this may also hold for variants from which a segment is lacking. Second, while Ranbom and Connine and Pitt and colleagues reported evidence concerning different pronunciations for a consonant, we provide evidence concerning the presence versus absence of a vowel. Frequency effects in comprehension are thus not limited to words in which a consonant is reduced; they can also occur in words in which a vowel is reduced. Third, our data show that listeners not only rely on the frequencies of different pronunciation variants of words in American English, but also in French. Fourth, our study shows that advanced learners of a language may also show sensitivity to variants' frequencies during speech comprehension. Like natives, learners keep frequency counts for pronunciation variants and use them during recognition.

In sum, we showed that listeners of advanced proficiency levels are sensitive to the frequency with which a given word lacks a vowel in their own speech input.

Our results thus show that exposure to a given variant of a given word predicts how well a listener processes that variant. Listeners appear to store the frequencies with which they hear each word form. This makes them very efficient listeners to the type of speech they are trained on, but less efficient listeners to other speech registers.

We obtained these results even though we only had subjective measures of the relative frequencies of the variants. Objective measures may better reflect language exposure, but in the absence of objective measures, subjective measures may also provide insight into speech processing.

Suppose we had only tested native listeners. We could then have explained the relative frequency effect in their processing in a different way. Previous studies have shown that more frequent words tend to be more reduced and to show stronger co-articulation than less frequent words (e.g., Wright, 2004; Jurafsky, Bell, Gregory & Raymond, 2000; Bybee, 1998; Lindblom, 1990). These acoustic differences among words and word pronunciation variants as a function of frequency could then have explained the attested relative frequency effects in processing: listeners could have recognized more frequent variants more easily because these variants tend to show more co-articulation, which could have facilitated recognition (e.g., Salverda, Kleinschmidt & Tanenhaus, 2014; Mattys, White & Melhorn, 2005). However, if the frequency effects had resulted from the words' acoustic properties and had thus been *indirect* consequences of the variants' frequencies of occurrence, all listener groups should have been sensitive to similar relative frequencies, which is contrary to fact. Our results show that the relative frequency effects that we observe do not result from the acoustic properties of the words’ variants. Groups that differ in their experience with the language show sensitivity reflecting their own exposure.

In Experiment 3, we collected subjective word frequency ratings and investigated whether these or a word frequency measure derived from a general database for French (e.g., Lexique, New et al., 2001) better predict beginner learners' behaviour. We found that the subjective word frequency ratings outperformed the objective frequency counts for predicting learners' accuracy. In the analysis of the RTs, neither word frequency measure showed a statistically significant effect. The subjective word frequency ratings probably outperformed the objective word frequency measure in predicting accuracy because the former measure better reflects the learners' experience with the words. This experience may be very different from the experience that native listeners generally have and that is reflected in the frequency database that we consulted. Objective frequency measures may generally outperform subjective frequency measures, as claimed by Ghyselinck and colleagues (2004), but only if the objective frequency measure faithfully reflects the given participants' exposure.

Our data further contribute to the question of whether unreduced word pronunciation variants have a special status in word recognition (e.g., Ernestus & Baayen, 2007; Janse, Nooteboom & Quené, 2007; Tucker & Warner, 2007). The native listeners in Experiment 1 performed well in the lexical decision task, but nevertheless showed a clear processing advantage for the unreduced variants. They processed the variants with schwa more accurately (99.0% correct) and more quickly (mean RT: 1153.96 ms) than the variants without schwa (88.4% correct; mean RT: 1217.99 ms). This privileged status of the unreduced variant may be due to the role of orthography or of context. In our experiment, context may have indeed favoured the unreduced variant, as the unreduced variant is more likely than the reduced variant when the word is only preceded by its determiner. Moreover, the duration of the vowel of the determiner was typical of a determiner followed by an unreduced word (mean /ə/ and /a/ duration in the determiner was 118 and 119 ms, respectively, if followed by an unreduced variant and 116 and 115 ms if followed by a reduced variant). Findings by Bürki and colleagues (2010) suggest that unreduced variants may lose their privileged status in speech production when presented in more natural context. Future research has to show whether this also holds for speech comprehension.

Since learners have less experience with reduced word variants than natives, these learners are expected to show a greater bias for the unreduced variants than the natives. This is exactly what we found. While the difference in lexical decision accuracy for the unreduced and reduced variants was 10.6% for the native listeners, it was 24.9% for the advanced learners and 29.3% for the beginner learners. This shows that the focus on unreduced variants in the classroom (e.g., Fonseca-Greber & Waugh, 2003) is harmful for learners' comprehension of casual speech. Interestingly, the beginners categorized more than half of the reduced variants (63.3%) as pseudo words, possibly because the majority of these variants have illegal consonant clusters (e.g., /ʀv/ in reduced /ʀvy/ 'magazine').

Furthermore, the privileged status of the unreduced variant appears from the interaction between variant type and the relative frequency rating provided by the advanced learners in the combined dataset of the beginner and advanced learners (see accuracy analysis). The unreduced variant is well recognized, independent of its relative frequency. Nearly all accuracy and RT plots seem to show this interaction but it does not surface as statistically significant in the other analyses (possibly because of lack of statistical power).

The question arises how to account for these results in models of word recognition. As explained in the Introduction, a rule-based model of word recognition cannot easily account for effects of relative frequency for reduced word pronunciation variants, since in this type of model, word-specific variant frequencies can only become available (and thus affect the word recognition process) after the word has been identified. In contrast, relative frequency effects can easily be accounted for in models assuming lexical storage of multiple word pronunciation variants. For instance, if Shortlist B (Norris & McQueen, 2008) is adapted in such a way that it stores these variants, the relative frequencies can co-determine the variants' priors and thus how easily these variants are processed.

In line with this, we assume that French natives and advanced learners of French store the relative frequencies of pronunciation variants together with the lexical representations of these variants in the mental lexicon, for instance, in the form of resting activations or in the form of priors. Similarly, beginner learners also store the reduced variants they have encountered, together with their (low) frequency counts. The variants stored and their frequencies change as a function of the type and amount of exposure. We leave open the possibility that variants with very short schwas are also stored, which we view as a topic for future research.

An alternative account for the processing of segment reduction is *naive discriminative learning* as proposed by Baayen (2010). Although this model can account for reduction, it cannot easily simulate the observed relative frequency effects. In the *naive discriminative learning* model, a two-layer network with symbolic representations for a word's form and a word's semantics, each input unit (unigram, bigram, uniphone, biphone or triphone) is linked to each meaning, and a weight is associated with each link. Effects of the relative frequencies of the two pronunciation variants of a word have to be expressed in the weights of the connections between the biphones or triphones of the two variants of the word and the word's meaning. These weights also have to indicate the likelihoods that these input units represent different word types. That is, word-specific relative frequencies cannot be separately represented from other (relative) frequencies, and therefore these relative frequency effects cannot simply emerge. The *naive discriminative learning* model cannot easily explain the observed relative frequency rating effect.

Finally, the two tests added to the experiment to determine the learners' proficiency levels provide interesting insights in the relation between different types of proficiencies. Learners who processed reduced variants more accurately were also better at the dictation task, while they were not necessarily better at visual lexical decision (the LexTALE task). This result suggests that the ease with which learners process spontaneous speech is not strongly correlated with their vocabulary knowledge, but more specifically with their listening skills.

In conclusion, this study provides evidence that listeners lexically store reduced pronunciation variants with their relative frequencies. We compared listener groups differing in their proficiency levels and demonstrated that their processing of unreduced and reduced pronunciation variants is best predicted by their exposure to these variants, as captured by subjective frequency ratings. Together, these results reveal that listeners easily comprehend the speech that they hear in daily life because they are fine-tuned to the frequencies of occurrence of reduced pronunciation variants in their own daily life input.

Appendices

Appendix A

The target words in the experiment.

la besogne ‘the task’, *le besoin* ‘the need’, *la cerise* ‘the cherry’, *le chemin* ‘the path’, *la chemise* ‘the shirt’, *le cheval* ‘the horse’, *la demande* ‘the request’, *la demeure* ‘the residence’, *le degré* ‘the degree’, *le devoir* ‘the duty’, *la fenêtre* ‘the window’, *la gelée* ‘the frost’, *le genou* ‘the knee’, *la leçon* ‘the lesson’, *le melon* ‘the melon’, *la menace* ‘the threat’, *le menu* ‘the menu’, *la mesure* ‘the measure’, *le neveu* ‘the nephew’, *la pelouse* ‘the lawn’, *le rebelle* ‘the rebel’, *la recette* ‘the recipe’, *la recherche* ‘the search’, *le record* ‘the record’, *le reçu* ‘the receipt’, *le recueil* ‘the collection’, *le refrain* ‘the chorus’, *le refuge* ‘the refuge’, *le refus* ‘the refusal’, *le regard* ‘the look’, *le remède* ‘the remedy’, *le renard* ‘the fox’, *le repas* ‘the meal’, *le repos* ‘the rest’, *le retard* ‘the lateness’, *le retour* ‘the return’, *la retraite* ‘the pension’, *la revanche* ‘the revenge’, *la revue* ‘the magazine’, *la seconde* ‘the second’, *le secret* ‘the secret’, *la semaine* ‘the week’, *la tenue* ‘the outfit’, *la vedette* ‘the star’

Appendix B

Table B.1
Fillers in the lexical decision experiment.

	bisyllabic	monosyllabic
real words	149	89
	(e.g., *la boisson*)	(e.g., *le lit*)
pseudo words	44 with schwa: one of two variants	89
	(e.g., *le beseuil*, *le b'seuil*)	(e.g., *le tuit*)
	149 without schwa	
	(e.g., *la bimise*)	

Table B.2
Summary of the regression model predicting the logit of an incorrect answer by the French native listeners. The intercept represents an unreduced variant.

Fixed effects		*ß*	*z*	*p* <
Average relative frequency rating		-1.818	-5.769	.001
Variant type (reduced)		6.151	2.638	.01
Random effects				*SD*
Word	intercept			6.986
	variant type (reduced)			7.155
Participant	intercept			.995

Table B.3
Summary of the linear mixed-effects model fitted on the log-transformed lexical decision times from the French native listeners. The intercept represents an unreduced variant.

Fixed effects		*ß*	*t*
Target duration		. 221	5.748
Trial number		-.0000545	-2.396
Previous RT		.0824	4.576
Variant type (reduced)		.0627	6.0477
Average relative frequency rating		-.0132	-2.743
Random effects			*SD*
Word	intercept		0.0265
	variant type (reduced)		0.0485
Participant	intercept		0.0715
Residual			0.128

Table B.4
Summary of the regression model predicting the logit of an incorrect answer by the advanced learners. The intercept represents an unreduced variant.

Fixed effects		*ß*	*z*	*p* <
Word frequency from Lexique		-.425	-3.482	.001
Dictation task performance		-.0188	-2.336	.05
Variant type (reduced)		1.935	10.160	.001
Average rating by the advanced learners		-.408	-2.272	.05
Random effects				*SD*
Word	intercept			0.921
Participant	intercept			0.814
	variant type (reduced)			0.762

Table B.5
Summary of the linear mixed-effects model fitted on the log-transformed lexical decision times of the advanced learners. The intercept represents an unreduced variant.

Fixed effects		*ß*	*t*
Trial number		-.0000496	-2.0211
Word frequency from Lexique		-.0284	-4.595
Previous RT		.0724	4.611
Variant type (reduced)		.0616	5.484
Average rating by the advanced learners		-.0421	-3.102
Random effects			*SD*
Word	intercept		0.0467
	variant type (reduced)		0.0543
Participant	intercept		0.0601
Residual			0.135

Table B.6
Summary of the regression model predicting the logit of an incorrect answer by the group of beginner learners in Experiment 3. The intercept represents an unreduced variant.

Fixed effects		*ß*	*z*	*p* <
Variant type (reduced)		1.474	9.580	.001
Word frequency		-0.118	-4.298	.001
Dictation task performance		-0.0289	-2.460	.05
Random effects				*SD*
Word	intercept			0.984
	variant type (reduced)			0.757
Participant	intercept			0.671

Table B.7
Summary of the linear mixed-effects model fitted on the log-transformed lexical decision times of the group of beginner learners in Experiment 3. The intercept represents an unreduced variant.

Fixed effects		*ß*	*t*
Target duration		0.167	5.758
Previous RT		0.196	11.808
Variant type (reduced)		0.0164	1.961
Random effects			*SD*
Word	intercept		0.0196
Participant	intercept		0.0951
	variant type (reduced)		0.0333
Residual			0.164

Appendix C

Experiment 3a

Methods

Participants

Forty-seven Dutch university students (aged 19-32 years; five males) were paid to participate in the experiment. They had the same experience with French as the participants in Experiment 3.

Materials

The materials were the same as in Experiment 3.

Procedure

The experiment was administered in sound-attenuated booths at the Max Planck Institute for Psycholinguistics in Nijmegen. The procedure was the same as for Experiment 2.

Results

In this results section, we focus on the differences between Experiments 3 and 3a.

Relative frequency estimation task

We obtained 2068 ratings, which show a slight interrater agreement of .02 ($p < .001$) in terms of *Fleiss' kappa* (calculated on the basis of the ratings for the reduced variants). All by word average relative frequency ratings were between 2.02 and 3.68. These average relative frequency ratings do not correlate with the beginner learners' ratings from Experiment 3. Unlike the ratings obtained in

Experiment 3, those obtained in Experiment 3a correlate with those provided by the advanced learners (correlation of .53, $p < .0001$).

Lexical decision accuracy

Like the beginners of Experiment 3, the beginners of Experiment 3a made more errors for reduced (62.0%) than for unreduced variants (26.9%). We analyzed the combined dataset of the participants from Experiments 3 and 3a to examine whether this larger dataset showed an effect of relative frequency ratings on accuracy. We tested the effects of different predictors, including the average ratings provided by the beginners from Experiment 3, those provided by the participants from Experiment 3a, and the average ratings based on both Experiments, in addition to the average ratings obtained in Experiment 1 and those obtained in Experiment 2. None of the predictors showed statistically significant effects. There were no interactions with participant group.

Lexical decision RTs

We again analyzed the combined dataset of the participants from Experiments 3 and 3a. We only included trials from Experiments 3 and 3a where the participants had provided correct responses (2284 trials, 53.1% of the data). RTs deviating from the grand mean (1433.39 ms) by more than 2.5 times the standard deviation (358.19 ms) were considered to be outliers and removed, which left 2218 observations (51.6% of the data). Only one model showed the expected negative correlation between RTs and participants' ratings of the relative frequencies, but this model had a higher AIC than most other models. We therefore conclude that there is no statistically reliable effect of any relative frequency rating.

Summary and discussion

The accuracy results and RT results are very similar to those obtained in Experiment 3: in the combined dataset of Experiments 3 and 3a, we found no interaction of participant group and any predictor. For instance, neither the relative frequency ratings provided in Experiment 3 nor those provided in Experiment 3a show interactions. This implies that the difference in instruction for the relative frequency rating task between Experiments 3 and 3a does not affect how well the relative frequency ratings predict beginner learners' processing.

Since there is also no correlation between the relative frequency ratings provided by the participants in this Experiment and in Experiment 3, we believe that the task was too difficult for them. Beginner learners have too little experience with reduced pronunciation variants to provide reliable relative frequency information. An additional explanation for the absence of the correlation is that there were substantial differences among and between the participants in the two groups. This possibility is supported by the correlation between the beginner learners' ratings of Experiment 3a (but not of Experiment 3) and the advanced learners' ratings, which suggests that participants of Experiment 3a had a higher proficiency than those of Experiment 3. Finally, the combined data set of Experiments 3 and 3a shows no statistically significant effects of any variant relative frequency measure. This shows that these ratings do not even predict beginner learners' processing if tested on 103 participants.

Processing reduced speech in the L1 and L2: A combined eye-tracking and ERP study

Chapter 5

This chapter is a reformatted version of:
Kimberley Mulder, Sophie Brand & Mirjam Ernestus (submitted). Processing reduced speech in the L1 and L2: A combined eye-tracking and ERP study.

Abstract

Native listeners generally understand reduced variants effortlessly, in contrast to learners of a language. We investigated which mechanisms may underlie the poor comprehension of reduction by advanced learners. We focused on how natives and learners process schwa reduction in French nouns (e.g., /ʀkɛ̃/ for /ʀəkɛ̃/ *requin* 'shark') occurring in sentences.

Participants performed a passive listening visual world task. They heard a sentence with either a reduced or an unreduced word variant together with pictures representing the target word (e.g., *fenêtre* /fənɛtʁ/ 'window'), a phonological competitor (e.g., *fourchette* /fuʀʃɛt/ 'fork'), and two neutral distractors. Eye movements and EEG were recorded simultaneously.

In contrast to the natives, the learners showed no N400 for reduced variants. The eye-tracking data show that the learners considered phonological competitors of both reduced and unreduced variants longer than the natives. As a result, learners have more problems than natives in accessing lexical-semantic representations of reduced words.

5.1 Introduction

One of the most challenging aspects of using a foreign language is to understand everyday speech. The words to which we are exposed in everyday life often lack sounds. This phenomenon of reduction, whereby words are produced with fewer segments compared to their citation forms (from now on referred to as reduced and unreduced variants, respectively), is highly frequent in Germanic languages, such as Dutch (e.g., Ernestus, 2000) and German (e.g., Kohler, 1990), and in non-Germanic languages such as French (e.g., Adda-Decker, Boula de Mareüil, Adda & Lamel, 2005) and Finnish (Lennes, Alaroty & Vainio, 2001). While native listeners generally understand reduced variants effortlessly, language learners typically encounter problems understanding these variants (e.g., Wong, Mok, Chung, Leung, Bishop & Chow, 2015; Nouveau, 2012; Stridfeldt, 2005). This study examines why advanced learners suffer more from reduction than native listeners, and investigates which mechanisms may be responsible for their problems.

Previous research has shown that native listeners do not encounter great difficulties processing reduced variants in context. Ernestus, Baayen, and Schreuder (2002), for instance, demonstrated that Dutch listeners correctly recognized 92.0% of the highly reduced variants when these variants were presented in their sentence context. Likewise, ERP analyses reported by Mulder, Drijvers, and Ernestus (in prep.) and referred to in Mulder, Ten Bosch, and Boves (2016) showed no significant effect of reduction on the ERPs to words in mid-sentence position. This suggests that the effect of reduction on semantic processing in mid-sentence position is likely to be small, and that, in general, native listeners process reduced variants as easily as unreduced variants in the middle of sentences.

The EEG study on Dutch natives by Drijvers, Mulder, and Ernestus (2016) partly supports this finding. Alpha power increase, which was interpreted as a reflection of higher auditory cognitive load, is shown not to be larger in response to reduced variants than to unreduced variants in mid-sentence position. In contrast, gamma power increases were observed to unreduced but not to reduced variants, also in mid-sentence position, which suggests that it takes more effort for reduced variants than for unreduced variants to activate their semantic networks (c.f., for words presented in isolation Van de Ven, Tucker & Ernestus, 2011).

Natives may recognize reduced words because they use fine phonetic details in the acoustic signal, such as voicing and lengthening of surrounding phonemes (e.g. Ernestus, Kouwenhoven & Van Mulken, 2017). Likewise, Zimmerer and Reetz (2014) demonstrated, for instance, that the length of the preceding /s/ plays a role in the reconstruction of a reduced /t/ by German natives. Similarly, in a visual world paradigm study, Brouwer, Mitterer, and Huettig (2012a) demonstrated that Dutch natives initially activate competitors that are similar to the auditorily presented reduced variants, but that in the end they identify target words by reliance on fine phonetic detail.

Context may provide syntactic and semantic information that can aid target word identification. Van de Ven, Ernestus, and Schreuder (2012), for instance, demonstrated that Dutch listeners use both the preceding and following context to process reduced words. Evidence for the use of syntactic cues in speech processing by natives comes from the eye-tracking study by Viebahn, Ernestus, and McQueen (2015). They reported that Dutch native listeners recognized past participles more quickly if they occurred after their associated auxiliary verbs than if they preceded them. This order effect was stronger for casually than for carefully produced sentences.

Much less is known about learners' perception of reduced variants. Previous research shows that language learners encounter difficulties with reduced variants in isolation and in limited context. Nouveau (2012), for instance, asked advanced Dutch learners of French to perform a French transcription task including words with word-medial schwa reduction (like /ʀvy/ for /ʀəvy/ *revue* 'magazine'). Students recognized the reduced variants less often (56.1% correct) than the unreduced variants (92.6% correct). Wong, Mok, Chung, Leung, Bishop, and Chow (2015) found that Hong Kong students with intermediate to advanced levels in English made errors for 40.0% of the words showing regular simple phoneme reduction in an English transcription task. Whereas Nouveau and Wong and colleagues tested foreign language learners in a country where it is not that country's language, Ernestus, Dikmans, and Giezenaar (in press) tested second language learners who learned their second language in the country where it is the native language. Ernestus and colleagues auditorily presented Dutch reduced and unreduced sentences to advanced second language (L2) learners of Dutch in the Netherlands and observed more transcription errors for reduced sentences (70.0%) than for unreduced sentences (38.0%). Learners thus show a processing advantage for the unreduced variant.

Language learners have been shown to not always make full use of the semantic information provided by the context while processing reduced word variants (e.g., Van de Ven, Tucker & Ernestus, 2010). In two auditory lexical decision experiments with implicit semantic priming, Van de Ven and colleagues showed that Asian late learners of English have more difficulties than native listeners in using semantic cues for lexical access in their second language, not only when listening to reduced words, but also in their processing of unreduced words (see also, e.g., Ernestus, Dikmans & Giezenaar, in press; Bradlow & Alexander, 2007).

Casual speech processes in a second language have been shown to be problematic for learners if the reduction process is not known from the learner's native language. Mitterer and Tuinman (2012), for instance, investigated whether L2 learners access knowledge from their native language (L1) when processing casual speech in their L2. In a production experiment, Mitterer and Tuinman observed that /t/ reduction similarly occurred in Dutch and German nouns, but differed between the two languages if /t/ was a verbal inflection. In a perception experiment, they found that German beginner learners of Dutch process /t/ reduction as native Dutch listeners do when the /t/ is part of a noun stem, but not when /t/ is a verbal inflection. Likewise, Ernestus, Kouwenhoven, and Van Mulken (2017) observed that learners of English with native languages that do not allow word-final /nt/ (i.e., Spanish and Mandarin) have problems interpreting reduced variants of this word-final cluster.

Little is known about how exactly natives and learners process reduced variants. It is still unclear why even advanced learners encounter great difficulties processing reduced word variants. This study sought to obtain data which help to address this research gap. We compared the processing of reduced variants by advanced learners to that by natives in a combined eye-tracking and ERP study. We combined two different techniques because each provides different types of insights into speech processing which can complement the other and paint a more complete picture of how listeners process reduced speech over time. We examined which competitors are considered by natives and learners when a reduced instead of an unreduced variant (eye-tracking experiment) is presented, and whether or not the two participant groups have difficulties in activating the meanings of reduced and unreduced variants presented in context (N400 in the EEG experiment).

We presented our stimuli within sentences in a passive listening task which better reflects situations in daily life than, for instance, the lexical decision task used by Van de Ven and colleagues (2010). We tested both French natives and advanced Dutch learners of French who had some exposure to reduced variants, and were therefore expected to be at least somewhat familiar with our reduced stimuli. We focused on schwa in the initial syllable of French words, which can be absent even if the words are produced in isolation and in a formal situation (e.g., /fənɛtʁ/ 'window', which can be pronounced as /fnɛtʁ/; Bürki, Ernestus & Frauenfelder, 2010). The absence of schwa is often complete, leaving no traces in the acoustic signal (e.g., Bürki, Fougeron & Gendrot, 2007). This study focused on Northern French as produced by a young adult, highly educated speaker.

Like French natives, Dutch learners are familiar with the concept of schwa reduction from their native language (e.g., Hanique, Schuppler & Ernestus, 2010). The Dutch word *gekocht* 'bought', for example, has the citation form /xəkɔxt/, but the word-initial schwa may be very short, and occasionally absent, as in [xkɔxt].

As advanced Dutch learners of French may have difficulties processing French word variants without schwa as evidenced by results obtained with transcription tasks (e.g., Nouveau, 2012) and lexical decision tasks (e.g., see Chapter 4), we expect that during word recognition, learners consider phonological competitors more than natives, and that semantic activation of the target word proceeds more slowly in learners than in natives. Since even natives show more effortful activation of semantic networks for reduced variants than unreduced variants (e.g., Van der Ven et al., 2011; Drijvers et al., 2016), we also expected to find a larger negative N400 for reduced variants than for unreduced variants in learners. The difference between reduced and unreduced variants may even be larger in learners than in natives because learners have had less exposure to reduced variants.

5.2 Methods

5.2.1 Participants

Twenty-eight right-handed native speakers of French, mostly undergraduate students, (mean age: 22.8 years, 7 males) participated in the experiment. They were all raised in the northern region of France (around Ile de France) by native speakers of French. Twenty-seven right-handed Dutch students of French (mean age: 20.6 years, 3 males) also participated in the experiment. They had all learned French in high school for six years and continued upgrading their proficiency level at university, where they were exposed daily to (formal) French. Their proficiency level roughly corresponds to C1-C2 according to the Common European Framework of Reference for Languages (CEFR, Council of Europe, 2011). All participants had normal hearing ability, no language disability or psychological or neurological impairments, and had normal or corrected-to-normal vision. The study was approved by the local ethics committee, and written consent was obtained prior to participation. Participants were paid to take part in the experiment.

5.2.2 Stimulus materials

We selected 54 French bisyllabic nouns with a schwa in the first syllable and stress on the second syllable (e.g., *requin* /ʀəkɛ̃/ 'shark'). In order not to draw participants' attention to the experimental target words, we added 59 filler nouns: one four-syllable word (*aéroport* 'airport'), 11 three-syllable words (e.g., *boulanger* 'baker'), 31 two-syllable words (e.g., *coiffeur* 'hairdresser') and 16 one-syllable words (e.g., *pont* 'bridge'). Finally, we selected three nouns (1 bisyllabic noun, 2 three-syllable nouns) for the practice trials. All nouns were concrete nouns that could be depicted in a line drawing.

Each of the nouns was embedded in a sentence, in mid-sentence position. The sentences were structured as followed: subject + verb form (+ *que* 'that') + target / filler /practice trial noun (+ verb form) + prepositional phrase / noun (e.g., *J'ai appris que le requin vit dans l'océan* 'I have learned that the shark lives in the ocean'). To avoid predictability effects, we kept the semantic context as neutral as possible up until the first verb form. In contrast, the noun in the prepositional phrase was always semantically related to the preceding experimental target or filler noun.

The sentences were recorded by a female native speaker of French from the north of France. The target sentences were recorded with three different instructions: the speaker was not given any instructions on how to produce the noun; we asked the speaker to produce the noun with its full realization; and we asked the speaker to produce the noun without the schwa in the first syllable. The reduced and unreduced nouns were then spliced from their original sentences and pasted into the sentences recorded without giving the speaker any instructions. This was done to make sure that the sentences with reduced and unreduced target nouns only differed with respect to the realization of the schwa in the noun, and that possible effects in EEG and eye-tracking could not be attributed to a difference in context. The sentences' average intensities were scaled to the same intensity level in PRAAT.

We acoustically analyzed the target nouns using the speech analysis software package *Praat* (Boersma & Weenink, 2014). The spliced reduced variants had a mean schwa duration of 5.3 ms, whereas the spliced unreduced variants had a mean schwa duration of 69.0 ms. The mean durations of the entire target words were 370.4 ms for the reduced variants and 460.3 ms for the unreduced variants. In the sentences produced by the speaker before having received any instructions, schwa was present most of the time (i.e., absent in 11.1% of the noun tokens). The experimental target words occurred, on average, 832 ms after sentence onset.

For the visual world paradigm, the 54 target nouns, the 59 filler nouns, and the three practice trial nouns (i.e., the 116 experimental words) were each paired with three distractor nouns, such that 116 quadruplets were created. We ensured that all distractors were semantically unrelated to the experimental words (in the target trials, the filler trials, and the practice trials) and to each other. The three distractors for a target noun included a phonological distractor and two neutral distractors. The onset of the phonological distractor (e.g., *fourchette* /fuʀʃɛt/ 'fork') overlapped with that of the target noun (e.g., *fenêtre* /fənɛtʁ/ 'window'). The amount of phonological overlap with the phonological competitor was equally large for the unreduced and reduced variants of the target (only the first segment). The neutral distractors did not overlap in onset phonemes (e.g., *coccinelle* /kɔksinɛl/ 'ladybird' and *montre* /mɔ̃tʁ/ 'watch' for *fenêtre* /fənɛtʁ/).

In order to prevent listeners from only considering pictures of the objects that showed overlap in their onsets, we created different distractor combinations for the filler nouns and the practice trial nouns. We presented 1 practice trial word and 38 filler target words together with three neutral distractors of which two shared their onset phonemes (e.g. /tʀɔ̃pɛt/ *trompette*, /tɔmat/ *tomate*, and /muʃ/ *mouche* for the filler /dɑ̃/ *dent*), and 2 practice trial words and 21 filler targets word with three distractors whose onsets showed no phonological overlap (e.g., /pʀiz/ *prise*, /tɑ̃t/ *tente* and /salad/ *salade* for the filler noun /balɛn/ *baleine*).

For every experimental word and for each of its distractor nouns, the participants saw a picture on the screen. The colored line drawings for 264 of our 464 pictures were taken from Rossion and Pourtois (2004) which were based on the original 260 black-and-white pictures produced by Snodgrass and Vanderwart (1980) and are freely available. We chose to use these colored versions because some of the original black-and-white drawings were of poor quality, and because Rossion and Pourtois have shown that basic-level object recognition is facilitated by the presence of color information. Nine of the objects depicted by Rossion and Pourtois were duplicated, so the resulting pictures represented plural nouns.

For the 218 words where no picture was available in the Rossion and Pourtois (2004) set, new drawings were created by a semi-professional graphics artist. Like the pictures of Rossion and Pourtois, these pictures were digitized at a high spatial resolution (600 dpi), resized, cropped, reduced to a resolution of 72 dpi (screen resolution) and centered on a rectangle of 281 by 197 pixels. Coloring and texture processing were mainly carried out with the graphic tools (e.g., pen, paintbrush) available in Photoshop CS6. Fifteen of the objects depicted by the semi-professional graphics artist were duplicated to represent plural nouns.

To test the entire picture set for clarity, eight native speakers of Dutch were asked to participate in a self-paced naming task. Each participant saw a random combination of half of the pictures, and typed their names in French. If participants did not know the object depicted by an image, they were asked to type *pas clair* 'unclear'. The grant mean accuracy score was 87.2% (*SD*: 4.4%, range: 79.1%-93.3%), while the average accuracy score per item was 87.1% (*SD*: 23.4%, range 0%-100%). The ten incorrectly named pictures were replaced by new pictures that were correctly named by all the participants. The final set consisted of 464 pictures.

Finally, for each of the 116 trials (113 experimental trials and 3 practice trials), we selected a photo that did or did not depict the situation described by the auditorily presented sentence. A matching sentence-photo pair is, for instance, the sentence *L'enfant a posé la cerise au sommet du gâteau* 'The child has put the cherry on top of the cake' followed by a photo of a cherry on top of a cake. A non-matching sentence-photo pair is the sentence *Elle a pris la tenaille pour enlever le clou* 'She has taken the pincers to pull out the nail' followed by a picture of someone holding a hammer above a nail. We took 55 photos with a Canon EOS 450 D camera. In addition, 61 right-free photos were selected from the internet. All photos were resized to 300 x 350 pixels and had a resolution of 72 dpi.

Thirty experimental lists were created based on these materials. The two versions of each noun (i.e., one unreduced and one reduced variant) were counterbalanced across lists in such a way that each participant heard only one version of the noun. Each of the 30 lists contained a different random combination of 27 reduced nouns and 27 unreduced nouns. In addition, we created 30 complementary lists by replacing the reduced and unreduced variants of the target words in every list by the corresponding unreduced and reduced variants, respectively. The fillers were the same in all lists and all lists started with the same practice trials. For each participant, the order of the trials in the list was pseudo-randomized in such a way that no more than three target trials occurred in a row. Each list was divided in two blocks of equal size.

5.2.3 Procedure

The French natives were tested individually in a sound-attenuated cabin in the Laboratoire Charles Bruneau of the *Institut de Linguistique et Phonétique Générales et Appliquées* (ILPGA) in Paris. The Dutch participants were tested individually in sound-attenuated booths either at the Max Planck Institute for Psycholinguistics in Nijmegen, at Leiden University, at Utrecht University, or at the University of Amsterdam. The participants were seated at approximately 60 cm distance from a 47.5 x 30 cm LCD computer screen, wearing an elastic cap to measure EEG, and Sennheiser HD215 headphones.

Eye movements were recorded at a sample rate of 500 Hz with an SR Research Ltd. Eyelink 1 eye tracker. Data points from the eye tracker were coded as fixations, saccades, or blinks using the Eyelink algorithm. Fixations were assumed to have a minimum duration of 100 ms. The timing of a fixation was established relative to the onset of the target word in the spoken sentence. Gaze position was categorized by object quadrant. Fixations were coded as directed to the experimental target object, to the phonological competitor, or to the neutral distractors.

Participants wore an elastic cap (Acticap) with 64 active electrodes to measure EEG. Electrode positions were a subset of the international 10-20 system, consisting of 8 midline electrodes and 50 lateral electrodes. Moreover, we placed an electrode on each of the mastoids and each electrode was referenced online to the left mastoid. The electro-oculogram (EOG) was recorded by two vertical electrodes placed above and below the right eye and by two horizontal electrodes with a right to left canthal montage. We kept the electrode impedance below 15 kΩ. The EEG and EOG signals were amplified (band pass = 0.02Hz-100 Hz), and digitized online with a sampling frequency of 500 Hz. Before data analysis, the signal was re-referenced to the average of the left and right mastoids and digitally filtered with a high cut-off filter of 30 Hz. Stimulus markers were sent to the EEG and eye tracker systems via Presentation version 16.5 (www.neurobs.com). To ensure temporal synchronization of the simultaneously recorded EEG and eye data within Presentation, an eye-tracking extension was developed in-house especially for this purpose.

The participants were asked to listen carefully to the spoken sentences they heard, and not to take their eyes off the screen (e.g., Huettig & Altmann, 2005). Before each trial, a fixation cross was presented at the center of the screen for 500 ms, to which the participants were asked to fixate. Each trial started with a short (1 second) preview of the picture display showing the pictures for the experimental word and for the three distractors, followed by the spoken sentence presented binaurally over the Sennheiser HD215 headphones. The visual display remained on the screen until the end of the sentence. Then, a blank white screen was presented for 500 ms. Participants pressed a *yes*-button or a *no*-button to determine whether the scene depicted by the image matched the scene described in the auditory sentence. The trial ended with a 500 ms empty white screen. Eye movements and EEG were recorded simultaneously throughout the experiment. The experiment lasted approximately 17 minutes. After the experiment, we asked the learners to fill out a language background questionnaire.

The learners performed two additional tests to reveal their proficiencies. The first was the visual lexical decision task LexTALE (Brysbaert, 2013; Lemhöfer & Broersma, 2012), which provides a measure of general lexical proficiency. The second was a dictation task, which provides a good indication of learners' comprehension of connected speech (Kennedy & Blanchet, 2013). Due to technical errors, the data of the LexTALE test were not available for two participants, and the data for the dictation task were not available for 13 participants. We found a raw mean GhentScore of 7.64 for the participants who performed the LexTALE task which correlates with a percentile of about 56.5% for learners of French at universities in Belgium, where French is one of the official languages. The accuracy score of the learners' transcriptions in the dictation task, representing a variety of reduction patterns, was on average 53.6%. Based on these results, we were able to classify our participants as advanced learners.

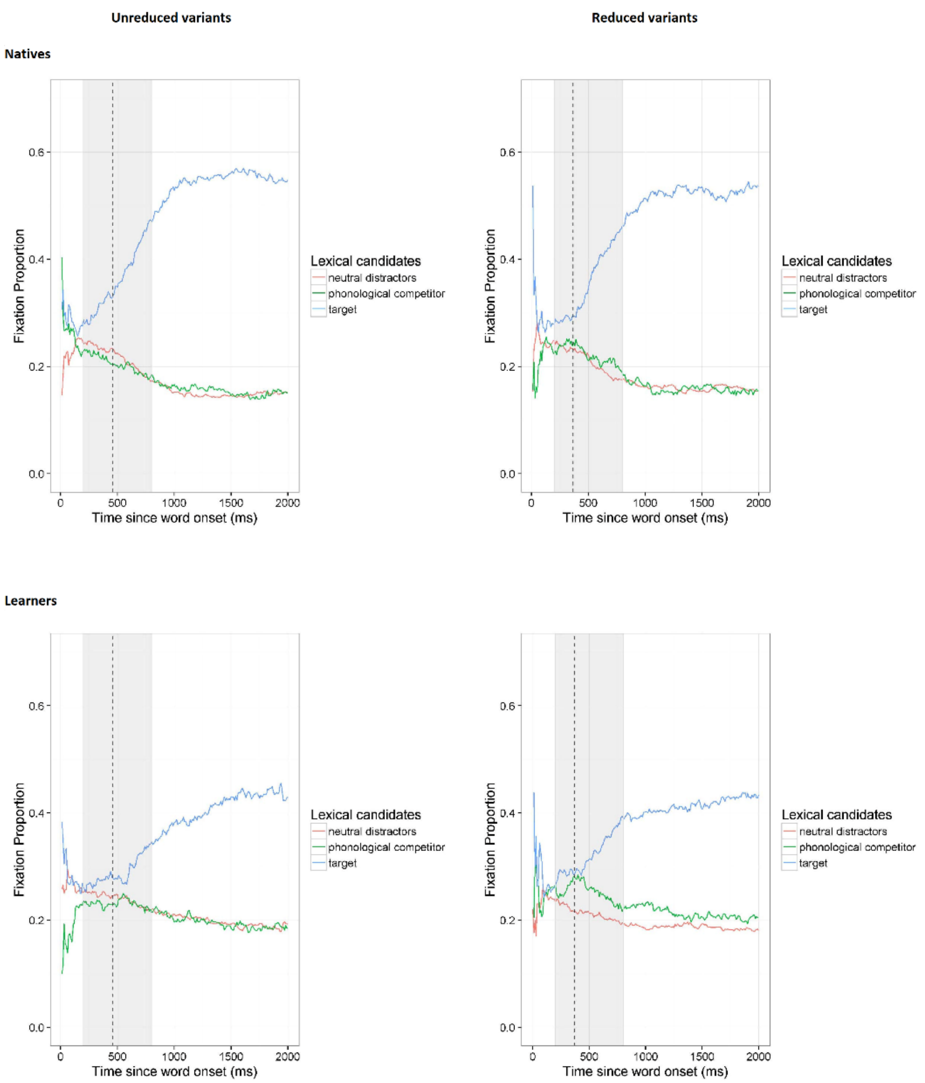
5.3 Results

5.3.1 Eye-tracking

The mean accuracy score on the matching sentence-photo task was 86.0% (*SD*: .35) for the French natives and 82.0% (*SD*: .38) for the advanced Dutch learners. We computed the average proportion of fixation to the four pictures (target, phonological competitor, two neutral distractors) in a given time bin of 10 ms. The proportion of fixations for the two neutral distractors was averaged. Figure 5.1 depicts the time course of the proportions of fixations by the French natives (upper panels) and learners (lower panels) to the unreduced variants (left panels) and the reduced variants (right panels). The y-axis represents the proportion of fixations (> 100 ms); the x-axis represents the time from target word onset; the dashed vertical line represents average target word offset (reduced variants: 370.4 ms, unreduced variants: 460.3 ms), and the gray box represents the analyses window (200-800 ms post target word onset). Given the 200 ms generally assumed necessary for programming and initiating an eye movement (e.g., Dahan, Magnuson, Tanenhaus & Hogan, 2011; Altmann & Kamide, 2004; Allopenna, Magnuson & Tanenhaus, 1998; Matin, Shao & Boff, 1993), it is not surprising that only 200 ms after target word onset, the curves representing the different lexical candidates start to diverge systematically. The proportions of fixations to the neutral distractors and the phonological competitor converged at around 800 ms after target word onset. Consequently, we analyzed our data in a time-window extending from 200 ms to 800 ms. Like

Huettig and McQueen (2007), we did not ask the participants to perform an explicit task, which may explain why the fixation proportion of the target word curve never exceeds 0.5.

Figure 5.1 Native (upper panel) and learner (lower panel) proportions of fixations to the different lexical candidates for the sentences with unreduced variants (left panel) and reduced variants (right panel).



For the natives, both the unreduced and reduced target words received the majority of fixations before their offsets (i.e., the dashed vertical line). The phonological competitor and the two neutral distractors received more fixations before target word offset for the learners than for the natives.

We used growth curve analysis (Mirman, 2014) to analyze the proportion of the fixations to the different lexical candidates in the combined dataset of natives and learners. The curves were modeled with cubic orthogonal polynomials and fixed effects of *Group* (natives or learners)*, Pronunciation variant* (reduced or unreduced variant) and *Lexical candidate* (target, phonological competitor, neutral distractors). The quartic polynomial function did not improve the model fit. The fixed predictors were added individually and their effects on model fit were evaluated by comparing the AIC values of the different models. We used contrast coding (e.g., Baayen, 2008; Baayen, Davidson & Bates, 2008; Bates, 2005; Dalgaard, 2002) for the lexical candidates, resulting in two variables: *neutral* (target versus neutral distractors) and *phonological* (target versus phonological competitor). All analyses were carried out in R version 3.1.2. using the lme4 package (built under R version 3.1.3).

Table 5.1 shows the fixed effect parameter estimates and their standard errors along with p-values. The interaction between *Lexical candidate* (phonological competitor versus target word) and *Group* demonstrates that the difference in fixations between the target word and the phonological competitor is larger for the French natives than for the learners. This finding indicates that the learners more often fixated on the phonological competitor than did the natives. This pattern is also visible in Figure 5.1, where the lines representing the target and the phonological competitor are much closer in the two lower panels for the learners, than for the native listeners in the upper panels.

Table 5.1 Parameter estimates for the analysis of the effects of *Group*, *Pronunciation variant*, and *Lexical candidate* on the proportion of fixations by French natives and learners. The intercept represents a French participant listening to an unreduced variant.

Fixed effects		*ß*	t
Intercept		0.060	2.64**
Group(learners)		0.12	3.72***
Pronunciation variant (reduced)		0.012	1.45
Pronunciation variant (reduced)*Group(learners)		-0.052	-4.53***
Lexical candidate (neutral)		0.044	4.49***
Lexical candidate (phon comp)		0.056	6.43***
Lexical candidate (neutral)*Group(learners)		-0.031	-2.21*
Lexical candidate (phon comp)* Group(learners)		-0.034	-2.70**
Lexical candidate (neutral)*Pronunciation variant (reduced)		0.0099	3.84***
Lexical candidate (phon comp)*Pronunciation variant (reduced)		-0.015	-5.62***
Lexical candidate (neutral)*Pronunciation variant (reduced)* Group(learners)		0.027	7.42***
Lexical candidate (phon comp)*Pronunciation variant (reduced)* Group(learners)		0.0020	0.55
Linear		-0.71	-24.00***
Quadratic		-0.10	-6.87***
Cubic		-0.044	-2.07*
Linear*Group(learners)		0.46	10.85***
Linear*Pronunciation variant (reduced)		-0.046	-1.08
Linear*Pronunciation variant (reduced)* Group(learners)		-0.13	-2.10*
Linear*Lexical candidate (neutral)		0.22	15.41***
Linear*Lexical candidate (phon comp)		0.17	11.96***
Linear*Lexical candidate (neutral)*pronunciation variant (reduced)		0.017	0.88
Linear*Lexical candidate (phon comp)*pronunciation variant (reduced)		0.0075	0.38
Linear*Lexical candidate (neutral)* Group(learners)		-0.11	-5.59***
Linear*Lexical candidate (phon comp)*Group(learners)		-0.13	-6.63***
Linear*Lexical candidate (neutral)*pronunciation variant (reduced)* Group(learners)		-0.011	-0.38
Linear*Lexical candidate (phon comp)*pronunciation variant (reduced)* Group(learners)		0.072	2.56*
Quadratic*Lexical candidate (neutral)		0.019	2.75**
Quadratic*Lexical candidate (phon comp)		0.035	4.89***
Cubic*Lexical candidate (neutral)		0.0043	0.43
Cubic*Lexical candidate (phon comp)		0.018	1.82
Cubic*Pronunciation variant (reduced)		0.13	4.22***
Cubic*Lexical candidate (neutral)*Pronunciation variant (reduced)		-0.0090	-0.64
Cubic*Lexical candidate (phon comp)*Pronunciation variant (reduced)		-0.055	-3.91***
Random slopes			*SD*
Participant	intercept		0.12
	lexical candidate (neutral)		0.051
	lexical candidate (phon comp)		0.045
	pronunciation variant (reduced)		0.031
Residual			0.091

*Note: *** indicates p < .001, ** indicates p < .01 and * indicates p < .05*

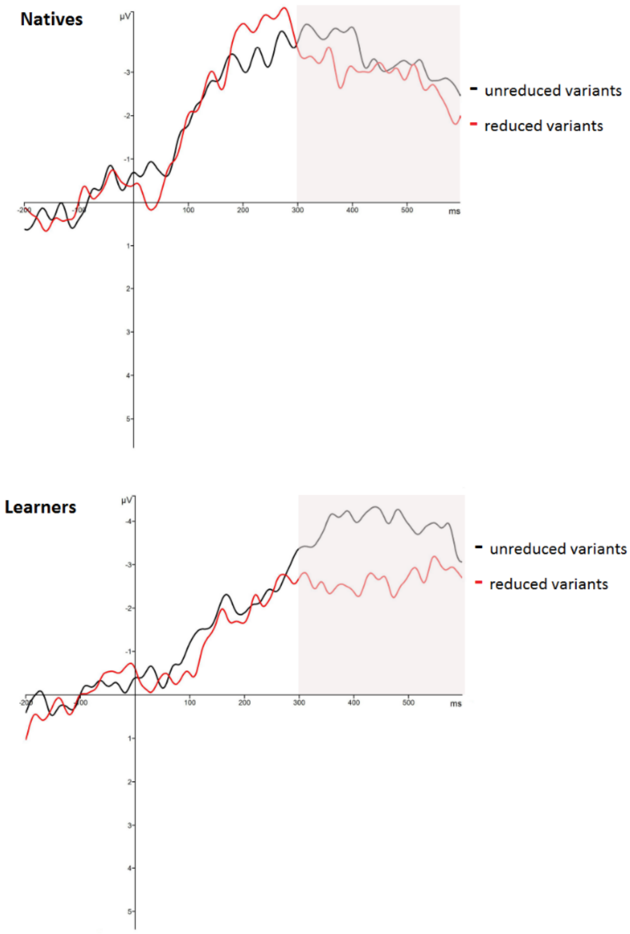
The model in Table 5.1 also demonstrates an interaction between *Lexical candidate* (neutral distractors versus target word), *Pronunciation variant*, and *Group*. The difference in fixations between the target word and the neutral distractors is noticeably larger for the French natives than for the learners if the target word is reduced. This shows that upon hearing a reduced instead of an unreduced variant, the learners considered more (unrelated) lexical candidates and had more difficulties in selecting the appropriate lexical candidate than did the natives.

The random slopes further show that participants were not equally sensitive to the neutral distractors, the phonological competitor, and the reduced variants.

5.3.2 ERP data

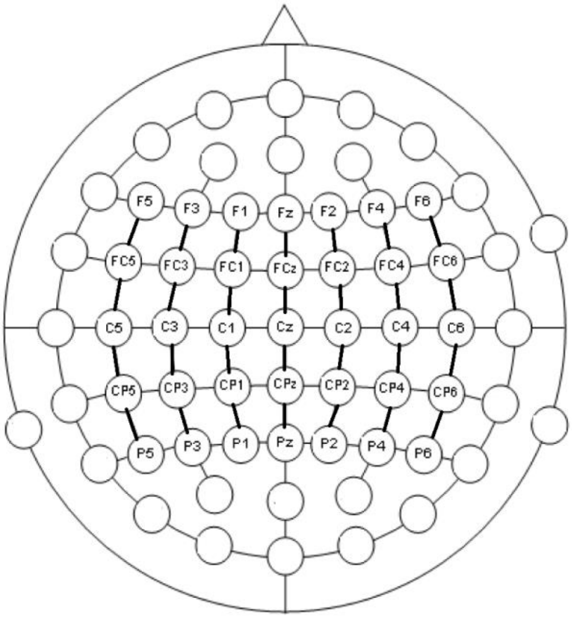
The EEG was time-locked to the onset of the target word. The continuous EEG was segmented into stimulus-time-locked epochs, starting from 200 ms before target word onset up to 600 ms after word onset. The epoch length was chosen to capture the N400, which generally peaks 400 ms after target word onset. Ocular artefacts were identified and removed using Independent Component Analysis (ICA). Residual artifacts were removed from the data with a semi-automatic inspection routine. The interval of 200 ms preceding the target word was used for baseline correction. Average ERPs were calculated for each participant and each electrode, and separate ERPs were computed for each level of the main predictor *Pronunciation variant* (unreduced and reduced). Figure 5.2 displays the grand averages to unreduced and reduced variants in natives and learners.

Figure 5.2 Grand averages of the ERPs to unreduced (black line) and reduced (red line) variants in natives (upper panel) and learners (lower panel) for a representative electrode F3. The grey area corresponds to the N400 analysis window. Negativity is plotted upwards.



To analyze the effect of *Pronunciation variant* on the N400 amplitude, we computed mean amplitudes between 300-600 ms after target word onset for each pronunciation variant, channel and participant (e.g., Nieuwland & Van Berkum, 2006; Kutas & Hillyard, 1980). Mean amplitudes were entered in a repeated measures ANOVA. For the analysis, we divided the head into seven columns of five electrodes along the anterior-posterior axis of the head (one midline column and three lateral columns for each hemisphere; see Figure 5.3; cf. Holcomb & Grainger, 2006).

Figure 5.3 Montage of the electrodes used in the analyses.



We conducted a separate ANOVA for the data from each of the three lateral columns and the midline column. The fixed predictors were *Pronunciation variant* (two levels: unreduced and reduced) and *Site* (5 levels per midline or lateral column, see description below) as within-subject factors and *Group* (two levels: natives and learners) as between-subject factor. The analyses on the lateral columns also included the predictor *Hemisphere* (2 levels: left and right). For both the midline and lateral columns, the anterior/posterior electrode factor *Site* contained five levels (midline column: Fz, FCz, Cz, CPz, and Pz; lateral column 1: F1/F2, Fc1/Fc2, C1/C2, Cp1/Cp2 and P1/P2; lateral column 2: F3/F4, Fc3/Fc4, C3/C4, Cp3/Cp4 and P3/P4; and lateral column 3 F5/F6, FC5/FC6, C5/C6, CP5/CP6 and P5/P6).

Greenhouse and Geisser (1959) correction was applied to all analyses with more than one degree of freedom in the numerator (corrected p-values and degrees of freedom are reported). Table 5.2 shows the interactions of *Pronunciation variant* with *Group* and *Site* and/or *Hemisphere* for the midline and lateral analyses. *Pronunciation variant* did not have a statistically significant main effect across analyses.

Table 5.2 All interactions figuring *Pronunciation variant* and *Group* in the ANOVA analyses of the mean N400 amplitudes on midline and lateral columns.

Column	Interaction	F-value
Midline	*Pronunciation variant * Group*	0.49
	*Pronunciation variant * Group * Site*	0.16
Column1	*Pronunciation variant * Group*	0.55
	*Pronunciation variant * Group * Site*	0.53
	*Pronunciation variant * Group * Hemisphere*	2.73
	*Pronunciation variant * Group * Hemisphere * Site*	0.55
Column 2	*Pronunciation variant * Group*	0.79
	*Pronunciation variant * Group * Site*	1.86
	*Pronunciation variant * Group * Hemisphere*	5.37*
	*Pronunciation variant * Group * Hemisphere * Site*	0.61
Column 3	*Pronunciation variant * Group*	1.58
	*Pronunciation variant * Group * Site*	0.47
	*Pronunciation variant * Group * Hemisphere*	6.26*
	*Pronunciation variant * Group * Hemisphere * Site*	0.06

*Note: * indicates p < .05*

Table 5.2 shows a significant interaction between *Pronunciation variant*, *Group* and *Hemisphere* in the two most lateral columns. Because of these interactions, subsequent ANOVA analyses were performed on the data of the native listeners and the learners, respectively.

The ANOVAs on mean N400 amplitudes in the two most lateral columns for the native listeners showed neither a significant main effect nor interaction of *Pronunciation variant.* Analyses of the data for the learners showed a significant interaction between *Pronunciation variant* and *Hemisphere* in lateral column 2 [$F(1, 26) = 4.71$, $MSE = 4.28$, $p < .05$], showing a larger difference in negative amplitudes between unreduced variants and reduced variants in the left than in the right hemisphere. Further, we found a marginally significant interaction between *Pronunciation variant* and *Hemisphere* in lateral column 3 [$F(1, 26) = 3.54$, $MSE = 14.07$, $p = .07$]. This two-way interaction reveals more negative N400 amplitudes for unreduced than for reduced variants in the left hemisphere, and no difference in the right hemisphere.

5.4 General discussion

We investigated processing of French reduced and unreduced variants in sentence context by native French speakers and Dutch advanced learners of French in a combined ERP and eye-tracking study. In a passive listening visual world task, French natives and Dutch advanced learners of French listened to French sentences containing reduced or unreduced target variants while watching four pictures on the screen. During this task, their brain responses and eye movements were simultaneously recorded.

Our aim was to obtain more insights into the mechanisms that underlie the comprehension of reduced speech in both natives and learners. This will increase, among other things, our understanding of why learners have more problems than natives in processing reduced variants. The ERP data will inform us about the lexical-semantic activation of a given word while the eye tracking data will show the amount and the type of competition between possible lexical candidates.

The analyses of the eye tracking data showed that native listeners fixated target words more when these were presented in their unreduced variants relative to when these were presented in their reduced variants. This finding is in line with previous literature showing a processing advantage for the unreduced variant (e.g., Pitt et al., 2011; Ernestus & Baayen, 2007; Janse, Nooteboom & Quené, 2007; Ranbom & Connine, 2007; Tucker & Warner, 2007). Moreover, native listeners considered other lexical candidates more often when encountering a reduced variant. This result is in line with Brouwer, Mitterer, and Huettig (2012a and see also 2012b), who observed that presence of reduction in speech modulates phonological competition. Noteworthy is that the phonological competitor only overlapped in the first phoneme with the phonological variant of the target word (e.g., target *requin* /ʀəkɛ̃/ 'shark' and phonological competitor *rideau* /ʀido/ 'curtain'). Apparently, this minimal overlap was enough for the phonological competitor to compete with the target word at the form level. This is in line with results reported in previous studies on natives (e.g., Magnuson, Tanenhaus, Aslin & Dahan, 2003; Vitevitch, 2002; Allopenna, Magnuson & Tanenhaus, 1998).

In general, reduced variants were processed as easily by natives as unreduced variants in the middle of sentences, as evidenced by the ERP data. The absence of an effect of reduction on the N400 in the ERP data from native listeners suggests that, although competitors are activated for a short amount of time (as shown by the eye tracking data), they do not seem to hinder the selection of the appropriate meaning representation. This is in line with Mulder and colleagues (in prep.), who did not find any effects of reduction in the N400 time window in sentence context in Dutch listeners listening to their native language. The gamma effects to reduced and unreduced variants in mid-sentence position as observed by Drijvers and colleagues (2016) showed that it takes time to activate semantic networks of reduced variants compared to unreduced variants. Our findings suggest that these gamma effects are only small effects.

In addition, our results suggest that the finding that it may take longer for reduced variants than for unreduced variants to activate their semantic networks (Van der Ven, Tucker & Ernestus, 2011) cannot be generalized to words in context. In an auditory priming experiment, Van de Ven and colleagues observed that reduced variants only primed target words around 1500-1600 ms after the presentation of the target word, in contrast to unreduced variants which primed targets as early as 1000 ms after target presentation. Our results suggest that this delay in activation of the semantic network by reduced variants may be restricted to words presented in isolation.

Like the native listeners, the learners fixated target words more often when presented with unreduced variants compared to reduced variants, but the effect of reduction was larger in the learners than in the native listeners. This is in line with the observation that the effect of reduction on the processing of words presented in a lexical decision experiment is larger for learners than for natives (Chapter 4). In addition, learners showed more competition between lexical candidates than natives upon hearing reduced variants. In the learners, this increased competition effect was reflected in proportionally more fixations to the neutral distractors upon hearing reduced variants.

In contrast to the natives, the learners did show a significant effect of reduction in the N400 time window. The ERP for the unreduced variants shows a clear increase in negativity in the analysis window. In contrast, the ERP for the reduced variants hardly shows an increase. This effect of reduction in the EEG data reflects what we found in the eye movements in the same time window; that there is more competition for reduced than for unreduced variants, and that many reduced variants are not recognized before word offset (note that eye movements are delayed by approximately 200 ms with respect to the brain response, as it takes approximately 200 ms to launch a saccade). The ERP response to reduced variants also suggests that between 300 and 600 ms after word onset, learners are still processing the reduced variant at the form level. In other words, we argue that the ERP to reduced variants shows no clear N400 response, which would reflect semantic processing.

The absence of an N400 waveform to reduced variants in the learners is partially in line with findings by Sims, Bolger, and Tucker (2010) who observed that in contrast to unreduced variants and mildly reduced variants, highly reduced variants presented in isolation did not elicit an N400 response in native listeners of English. The authors explain the absence of an N400 by assuming the listeners did not activate the lexical representation of the reduced variant. While the French words in the present study were not highly reduced (only one segment -schwa- was reduced), encountering these variants in connected speech may have had a similar impact on learners.

Our data suggest that for the learners, the processing of the reduced variant is already hindered at the form level and not only at the semantic level, as the eye movements show increased lexical competition for reduced variants while there is no N400 for these variants. The lexical competition appears to impede semantic processing: as learners are still processing the reduced variant at the form level, activation does not spread (detectably) to semantic representations. The finding that phonological competitors contribute most to the competition suggests that it is the interface of acoustic input and form level that is affected by reduction rather than lexical-semantic processing.

Lexical competition is larger in learners than in natives both for the unreduced and the reduced variants, as shown by the eye movement data. It does not come as a surprise that learners need more time than natives to process words in French; learners have been observed to be slower in identifying spoken words in the second language (e.g., Scarborough, Gerard & Cortese, 1984; Soares & Grosjean, 1984). One possible cause for this processing delay is that learners, even though they know fewer words in their second language, consider more lexical candidates than monolinguals as they are checking more than one lexicon (cf. Rueschemeyer, Nojack & Limbach, 2008).

Another reason for this processing delay for reduced variants might be a direct result of how reduced variants are represented in the mental lexicon. Learners have had little exposure to reduced variants and may not or at best have weakened lexical representations for these reduced variants (Chapter 4). Natives, on the other hand, are assumed to have representations for at least some reduced variants (e.g., Bürki, Ernestus & Frauenfelder, 2010; Chapter 4). Both groups may also identify the word matching a reduced variant to the lexical representation of the unreduced variant by means of general processes (e.g., schwa insertion, Brouwer, Mitterer & Huettig, 2012b), which may take learners more effort.

In sum, this combined eye tracking EEG study shows that learners, more than native listeners, have problems in processing reduced variants in connected speech. Learners consider lexical competitors more often when hearing a reduced variant. As a consequence, they take longer to activate the word's meaning, resulting in the absence of a clear N400 for reduced variants in the learners' brain responses. We draw this conclusion because the interpretation of our EEG data is informed by the eye movement data, and vice versa. Overall, our data highlight the value of combining different methods when investigating the mechanisms underlying the processing of reduced speech.

General discussion and conclusions

Chapter 6

In this chapter, I will discuss what the studies presented in this dissertation have taught us about pronunciation variation in casual speech and the processing thereof by natives and foreign language learners. I will discuss the pros and cons of the different methods I applied for addressing my research questions. Moreover, I will sketch the implications of my results for psycholinguistic models of speech production and speech perception and what they tell us about language teaching methods.

6.1 Gradient versus categorical reduction

It is unclear whether reduction phenomena tend to be categorical or gradient in nature. The few descriptions of categorical processes in the literature leave open the question whether these categorical processes are exceptions or part of a general pattern (e.g., Bürki & Gaskell, 2012; Bürki, Ernestus & Frauenfelder, 2010; Bybee, 1998). The study in Chapter 2 sought to obtain data which helps to answer this question and aimed at providing more information about the frequency of occurrence of reduction patterns.

I investigated whether phoneme absence in pronunciation variants of French nouns ending in obstruent-liquid-schwa (OLS) clusters (e.g., /tʁə/ in *ministre* 'minister') results from gradient co-articulation and articulatory weakening, or from categorical processes. To that end, I extracted nouns from the Nijmegen Corpus of Casual French (NCCFr, Torreira, Adda-Decker & Ernestus, 2010) and applied three methods: phonetic analyses, examination of the distributions of the durations of the OLS phonemes, and comparison of the predictors of absence and duration of the OLS phonemes.

I distinguished seven pronunciation variants of the OLS clusters, of which three occurred only rarely. The citation form of the OLS cluster only occurred in 18.9% of the tokens (e.g., [ministʁə] for *ministre* ‘minister’). The variant without schwa was the most frequent variant of the OLS cluster and occurred in 35.7% of the tokens (e.g., ([ministʁ]). The second most frequent variant (27.1% of the tokens) was the variant in which both the liquid and schwa were absent (e.g., [minist]). Strikingly, the entire OLS cluster was also absent quite often (15.5%; e.g., [minis]). In these three high-frequency variants of the OLS cluster in which a phoneme was absent, all following phonemes were absent as well. The three low-frequency variants displayed a different pattern: a phoneme was absent while at least one of the following phonemes was present (e.g., [minisʁə], [minisʁ], [ministə] for *ministre* ‘minister’).

A phoneme is thus more likely to be absent if the following phoneme is also absent. This might be because final phonemes are likely to be co-articulated with the following word-initial phonemes, which are typically strongly articulated (e.g., Keating, Cho, Fougeron & Hsu, 2003) or because some types of phonemes are more likely to be absent than other types of phonemes (e.g., schwas are more likely to be absent than liquids and liquids are more likely to be absent than obstruents, Adda-Decker et al., 2008; Su & Basset, 1998). The hierarchy may be related to the ease with which phonemes are recovered.

Phoneme absence can be assumed to result from a categorical process when the phoneme is completely absent and has not left any traces in the acoustic signal. In contrast, the absence of a phoneme rather results mainly from gradient reduction when the phoneme has left traces, for instance, in durational or spectral properties of surrounding phonemes.

The phonetic analyses reported in Chapter 2 show that absent obstruents and liquids may leave durational and phonetic cues in neighbouring segments.

For instance, if the OLS cluster was completely missing, the preceding phoneme was often lengthened. Zimmerer, Scharinger, and Reetz (2011) found similar results for German /t/. Both in French and German, the absence of a phoneme can thus be cued by the lengthening of a neighbouring phoneme.

When /k/ in the OLS cluster (e.g., /k/ in *spectacle*) was absent, the spectrogram might still reveal its phonetic remnants, in the form of a velar pinch. Furthermore, I observed obstruent voicing before nouns starting with voiceless consonants. In French, progressive assimilation hardly occurs (e.g., Lodge, Armstrong, Ellis & Shelton, 1997). This supports the idea that obstruent voicing in this context resulted from regressive assimilation to the voicing properties of the segmentally absent liquid ([l] or [ʁ]), and thus that the obstruent voicing was a signature of this liquid.

Inspection of the distribution of the duration of a phoneme can also inform us about the nature of the processes underlying phoneme reduction. If the distribution shows two clearly distinct peaks, one of which is around zero ms representing the absent phoneme, absence of the phoneme can be assumed to result from both a categorical and a gradient reduction process. If, in contrast, the distribution is unimodal, without a clear peak around zero ms, absence of the phoneme can be assumed to only result from gradient shortening.

The histogram of the duration of the obstruent showed one peak at zero ms (17.9% of the noun tokens) and one peak around 35 ms, but no clear dip between these two peaks. The duration of the obstruent showed no clear bimodal distribution. The distribution of the duration of the liquid showed two clear peaks, one at zero ms (42.9% of the noun tokens) and one at around 90 ms, with a deep dip between them. Finally, the distribution of the duration of the schwa showed one clear peak at zero ms (79.4% of the data). The distributions of the durations of the obstruent and the liquid in OLS clusters have never been studied before, such that my results could not be compared to earlier findings. My result for schwa differed from the findings on schwa duration by Bürki, Fougeron, Gendrot, and Frauenfelder (2011b), who reported a unimodal, normal distribution for schwa duration. This difference may be due to the position of [ə] in the word: I focused on schwa in word-final position, whereas Bürki and colleagues examined word-medial schwa. Hence, whether the reduction process for the same sound should be considered categorical or gradient may also depend on the position in the word.

An alternative approach to decide between gradient versus categorical reduction is a comparison of the variables predicting phoneme absence and those predicting phoneme duration. Perfect (or substantial) overlap in predictors would suggest that phoneme absence and duration are driven by the same processes, and that phoneme absence is the endpoint of gradient sound shortening. In contrast, when phoneme absence and phoneme duration are affected by different variables, they are more likely to result from different processes, and the absence of a phoneme may then also result from a categorical process.

The results in Chapter 2 showed that obstruent absence and duration were each predicted by three of the variables that I tested. They were both predicted by phrasal position: the obstruent was shorter and more often absent in phrase-medial position than in phrase-final position. The presence and the duration of the obstruent were also both predicted by the absence of the following liquid. However, this could not be interpreted as evidence for gradient reduction of the obstruent because when the liquid was absent, the obstruent tended to be either completely absent or relatively long. There was one variable that only predicted obstruent absence (i.e., following bigram frequency) and one variable that only predicted obstruent duration (i.e., number of syllables).

In the liquid analyses, only speech rate predicted both liquid absence and duration: if speech rate was higher, the liquid was more often absent or shorter. There were two variables that only predicted liquid absence (i.e., schwa absence and following bigram frequency) and one variable that only predicted liquid duration (phrase-medial position).

Due to the small number of remaining data points for schwa (N = 60), I did not find significant predictors for schwa duration. As a consequence, the predictors of schwa absence and schwa duration could not be compared.

The three different methods (phonetic analyses, examination of duration distribution, and the comparison of the predictors of phoneme absence and phoneme duration) do not provide an unequivocal picture of the nature of the reduction processes underlying the absence of the obstruent. Based on the phonetic analyses and the distributions of the durations, one may conclude that the reduction process underlying the absence of the obstruent is, above all, gradient: obstruent reduction often left phonetic traces and obstruent duration showed no clear bimodal distribution. However, the comparison of the predictors of obstruent absence and obstruent duration showed that the overlap of predictors was not substantial, which would imply that the absence of the obstruent does not just result from gradient reduction. This inconsistent pattern of results may be due to the fact that my dataset consisted of only 291 noun tokens.

The phonetic analyses suggest that an absent *liquid* may also leave acoustic traces, for instance, in the form of the voicing of the preceding obstruent. This would argue for gradient reduction However, the distribution of liquid duration, in combination with the comparison of the predictors of liquid absence and liquid duration, suggest that the reduction process underlying the absence of the liquid can be categorical. The reduction process underlying the absence of the liquid is thus likely to be both gradient and categorical.

The distribution of the duration of schwa showed a clear peak at zero ms, which suggests that the reduction process of schwa is categorical. In line with other positions of schwa in the word (e.g., Bürki, Ernestus, Gendrot, Fougeron & Frauenfelder, 2011a), I assume that in word-final position the absence of schwa may not only result from gradient shortening processes, but also from categorical processes.

The results in Chapter 2 sketch a detailed picture of the nature of the reduction of OLS phonemes and demonstrate that reduction does not necessarily result from gradient articulatory overlap processes, as argued by Browman and Goldstein (1990), but also from categorical processes (as was the case for the liquid and the schwa in the observed data). I encourage more research addressing the same question for other reduction phenomena, also in other languages, such that we will obtain a more detailed picture of the relevance of categorical reduction processes in casual speech.

Although processes underlying the absence of neighbouring phonemes may be different for the different phonemes, the absence of one phoneme clearly affects the presence of a neighbouring phoneme. A phoneme is more likely to be absent if the following phoneme is also absent. As mentioned above, this might be because final phonemes are likely to be co-articulated with the following word-initial phonemes, or because some types of phonemes are more likely to be absent than other types of phonemes. Future studies are needed to show which of these accounts is most likely. Possibly, both tendencies contribute to the emergence of the reduction pattern.

6.2 The effect of reduction on natives' and learners' word recognition

In Chapters 3, 4, and 5, I investigated natives' and learners' processing of reduced variants presented in isolation and in context.

6.2.1 Natives

Chapter 3 presents the results of a transcription task and a visual lexical decision task with cross-modal identity priming. In both tasks, reduced word pronunciation variants were presented in their phonological and sentence context.

The transcription task consisted of sentences extracted from an informal conversation between two men, who were 18 years old and came from Ile-de-France, from The Nijmegen Corpus of Casual French (henceforth NCCFr, Torreira, Adda-Decker & Ernestus, 2010). The sentences contained single (e.g., [tsɛ] for /tysɛ/ *tu sais* 'you know') and multiple (e.g., [ʒtʀaj] for /ʒətʀavaj/ *je travaille* 'I work') segment reduction. The participants listened to the speech to be transcribed orthographically, sentence by sentence.

The French natives scored 88.0% correct on the transcription task. Most errors were semantically related transcriptions of the auditorily presented words, such as *terminer* 'to end' for *finir* 'to finish'.The natives also added chunks and interjections to their transcriptions, like *tu sais* 'you know', *mais oui* 'that is right', *ouais* 'yes', and *alors* 'so', which did not change the meaning of the sentence. The natives thus did not encounter great difficulty understanding casual speech.

In the cross-modal identity priming task, a sentence with an auditory prime word was played and at the offset of the prime word the visual target was presented on the computer screen. The prime word was either identical or unrelated to the visual target. The identical prime word was either unreduced, ending in an OL cluster (e.g., [ministʁ] for *ministre* 'minister'), weakly reduced, ending in a liquid ([minisʁ]), or strongly reduced, not containing any of the OLS cluster phonemes ([minis]). The participants had to indicate whether or not the visual target was a real French word.

The natives' mean accuracy score in the visual lexical decision task was 96.7%. The natives responded significantly faster to visual targets preceded by unreduced variants and strongly reduced variants than to targets preceded by weakly reduced variants and unrelated words. These findings suggest that reduced variants presented in context do not necessarily inhibit processing. This finding is in line with research on production by, for instance, Bürki, Ernestus, and Frauenfelder (2010), who observed that the processing advantage for unreduced over reduced variants may disappear when reduced words are presented in context.

Chapter 4 focused on the perception of words with simple schwa reduction in limited sentential context. Participants were asked to perform an auditory lexical decision task with unreduced and reduced pronunciation variants of schwa words preceded by their definite determiners (e.g., [la ʀəvy] versus [la ʀvy] for *la revue* 'the magazine'). In contrast to OLS clusters, schwa in the initial syllable of French words can be absent even if the words are produced in isolation and in a formal situation (e.g., Bürki, Ernestus & Frauenfelder, 2010).

The natives made more errors when responding to reduced (88.4% correct) than to unreduced variants (99.0% correct) and took more time to respond to a reduced variant (mean response time, henceforth RT: 1217.99 ms) than to an unreduced variant (mean RT: 1153.96 ms). These results provide additional evidence for a privileged status for unreduced variants in minimal sentential context. The absence of natural context may have favoured the unreduced variant, as the unreduced variant is more likely than the reduced variant when the word is only preceded by its determiner. Moreover, the duration of the vowel of the determiner was typical of a determiner followed by an unreduced word (average vowel duration of the determiner was 119 ms if followed by an unreduced variant, and 115 ms if followed by a reduced variant).

6.2.2 Dutch learners of French

Chapter 3 reported that Dutch advanced learners, undergraduate students of French, scored 48.3% correct on the transcription task. The learners thus made many more transcription errors than the natives (88.0% correct). The learners' accuracy rate is especially low if I take into account that the phonological system in Dutch is highly comparable to that of French: the two phoneme inventories are about the same, and the Dutch learners are also familiar with the same types of reduction (e.g., schwa reduction, Hanique, Schuppler & Ernestus, 2010). In addition, the learners had a high proficiency level of French: they had all followed French for six years at secondary school and continued upgrading their proficiency level at university, where they were exposed to French on a daily basis.

I focused on the transcriptions of eleven words and word combinations (henceforth items) with single and multiple reductions. The learners' average accuracy scores on these items were below 50.0%. The eleven items were often absent in the learners' transcriptions. In addition, the learners correctly transcribed the actually produced phonemes, but they failed to reconstruct the absent phonemes of the item that was intended by the speaker. This resulted in semantically and syntactically incorrect transcriptions, such as *je conseille* 'I advise' for *j'ai commencé* 'I started', which probably blocked sentence comprehension.

The advanced learners also participated in the lexical decision task with cross-modal identity priming which was described in section 6.2.1. The learners' average accuracy score in the visual lexical decision task was close to ceiling level (91.0%). Unlike the French natives, the Dutch learners responded more quickly to a visual target when the prime was the unreduced variant of the target word than when it was any other of the prime types. This suggests that the learners' processing of the reduced pronunciation variants was too slow to facilitate the processing of the immediately following targets.

I examined whether the Dutch learners' accuracy scores in the off-line transcription task predicted their RTs in the on-line lexical decision task with cross-modal identity priming. Running a model with normally distributed accuracy scores, I found a significant effect of accuracy score on RT: the higher the Dutch learners' accuracy scores in the transcription task, the longer their RTs in the lexical decision task with cross modal identity priming. Dutch learners with a higher proficiency level showed more native-like, that is, longer, RTs in the lexical decision task. Learners' conscious, explicit behaviour in the off-line transcription task thus reflects their processing in real-time. This finding is in line with previous studies (e.g., Kotz & Elston-Güttler, 2004). Kotz and Elston-Güttler reported native-like, that is longer, associative (e.g., *boy-girl*) RT priming effects for advanced learners than for low proficiency learners.

Both Dutch beginner and advanced learners of French participated in the auditory lexical decision task containing unreduced and reduced pronunciation variants of schwa words (see Chapter 4). The Dutch beginner learners had taken a maximum of two hours of French classes a week, for a maximum of six years at secondary school. Their proficiency levels roughly corresponded to B1-B2 according to The Common European Framework of Reference for Languages

(CEFR, Council of Europe, 2011). The Dutch advanced learners had the same proficiency levels as the advanced learners tested in Chapter 3 (i.e., C1-C2).

As expected, Dutch beginner learners made more errors for reduced (35.0% correct) than for unreduced variants (73.1% correct) and their responses were slower for reduced variants (mean RT: 1404.49 ms) than for unreduced variants (mean RT: 1355.25 ms). Accuracy scores on the reduced variants were even below chance level. The beginners categorized more than half of the reduced variants as pseudo words, possibly because the majority of these variants had illegal consonant clusters (e.g., /ʀv/ in reduced /ʀvy/ 'magazine'). The Dutch advanced learners also made more errors when responding to reduced variants (64.3% correct) relative to unreduced variants (89.2% correct) and responded faster to unreduced (mean RT: 1281.66 ms) than to reduced (mean RT: 1358.19 ms) variants. These results show that even if the reduced words are presented in limited sentence context, with simple phoneme reduction, both beginner and advanced learners encountered way more difficulty than natives recognizing these variants (for unreduced variants 99.0% correct, mean RT: 1153.96 ms; for reduced variants 88.4% correct, mean RT: 1217.99 ms). The results suggest a larger advantage for unreduced over reduced variants in learners' than in natives' processing.

The learners' performance in the auditory lexical decision task and in the lexical decision task with cross-modal identity priming demonstrates that recognition of reduced variants was impaired. This is probably because learners are not exposed to reduced speech on a daily basis and because they do probably not use cues provided by the word or the context to the same extent as natives during speech processing. Future research might explore whether mere increased exposure to casual speech induces a better performance in learners.

6.3 Variants' frequency effects

Words differ in how often they lack a given phoneme. The frequency of occurrence of word pronunciation variants may serve as a cue in processing. Native listeners recognize more frequent pronunciation variants of words that differ from their full variants only in their consonants more quickly than less frequent variants (e.g., Pitt, Dilley & Tat, 2011; Ranbom & Connine, 2007). More research is needed to demonstrate how wide-spread this phenomenon of variant frequency effects is. It is still unclear whether natives are also sensitive to frequencies of word pronunciation variants differing from their respective full variants only in their vowels, instead of in their consonants. Given the processing difference between consonants and vowels (e.g., Cutler, Sebastián-Gallés, Soler Vilageliu & Van Ooijen, 2000; Pisoni, 1973; Stevens, Liberman, Studdert-Kennedy & Öhman, 1969), findings on consonants may not be generalized to vowels. Moreover, little is known about the use of frequency information during speech processing by foreign language learners. Learners have probably encountered reduced pronunciation variants, of at least some words, less often than natives. The frequencies of reduced variants may be very low in their speech input, as a consequence of which all reduced variants have similar frequencies. As a result, frequency may not form a reliable cue for learners. I examined the use of variants' frequency information by natives and learners in more detail.

6.3.1 The use of variant frequency in the processing of reduced speech by natives

The lexical decision study with cross-modal identity priming presented in Chapter 3 contained prime sentences with prime words in four different conditions. Visual targets were combined with auditorily presented prime sentences, each containing a prime word. As mentioned above, the prime word was either an unreduced variant of the target word, a reduced variant of the target word, or an unrelated word. Importantly, I tested two types of reduced variants of the target word: a highly frequent, strongly reduced one (e.g., [minis] for *ministre* 'minister') and an infrequent, weakly reduced one (e.g., [minisʁ]). The frequencies of occurrence of the word pronunciation variants in this study were derived from Chapter 2 and were based on the cluster, averaged over all word types.

I found that French participants responded faster to visual targets preceded by highly frequent variants than to targets preceded by infrequent variants. This result adds to earlier findings by suggesting that native listeners are sensitive to the frequency of occurrence of a given pronunciation variant (e.g., Pitt et al., 2011; Ranbom & Connine, 2007). It demonstrates that natives are not only sensitive to word-specific frequencies (e.g., Ranbom & Connine, 2007) or frequencies based on the phonological context (e.g., Pitt et al., 2011), but also to the frequencies of a word-final cluster averaged over different words.

I further found an effect of the time interval between prime word offset and sentence offset. The larger the time interval, the longer the French natives' RTs, except when the prime word was the infrequent variant of the word. The natives thus seemed not to take the remainder of the sentence into account when the prime word was infrequent and therefore difficult to process.

Moreover, my results demonstrate that frequency information plays a more important role in speech processing than the amount of reduction. The strongly reduced, but highly frequent words produced more priming than the weakly reduced, but infrequent words. Strongly reduced pronunciation variants did not inhibit speech processing.

In Chapter 4, participants indicated on a scale from 1 to 6 the relative frequencies of two pronunciation variants of schwa words (e.g., [la ʀəvy] versus [la ʀvy] for *la revue* 'the magazine'). I studied whether the variant's relative frequencies predicted participants' RTs and accuracy scores in the auditory lexical decision task. The average relative frequency ratings provided by the French natives predicted both their accuracy and RTs for the two pronunciation variants of a schwa word.

The studies presented in Chapters 3 and 4 showed that variants that occur more often are recognized faster and more accurately. This result suggests that native listeners use frequency information during speech processing not only for word pronunciation variants differing from their respective full variants only in their vowels, but also for variants differing from their full variants only in their consonants.

6.3.2 The effect of reduction on natives' speech processing mechanisms

The transcription task in Chapter 3 informs us about listener's use of syntactic and semantic cues provided by the context and acoustic cues provided by neighbouring phonemes. The high score of the natives in the transcription task in Chapter 3 suggests that natives benefit from subtle phonetic, semantic, and syntactic information in the sentences presented. The natives seem to use this information to reconstruct the absent phonemes.

I conducted a combined eye-tracking and ERP study to examine the effect of reduction on the activation of the words' semantics in word recognition (Chapter 5). I combine eye-tracking and ERP because these methods provide different data which together may offer a detailed picture of the time course of the processing of these variants. The proportion of eye fixations on the different objects/words gives an indication of the lexical competition that participants are facing. The EEG signal informs us about the difficulty of lexical-semantic activation of a word in participants' ongoing sentence processing.

In the combined eye-tracking and ERP study, participants heard a sentence either containing an unreduced or a reduced pronunciation variant of a target word. At the same time, four objects appeared on the computer screen: one representing the auditorily presented target word (e.g., *fenêtre* 'window'), one representing a phonological competitor (e.g., *fourchette* 'fork'), two representing neutral distractors (e.g., *coccinelle* 'ladybird', *montre* 'watch') which did not overlap in onset phonemes and were semantically not related to the target noun and the phonological competitor.

The eye-tracking data demonstrated that the natives' proportions of fixations on the phonological competitors were higher when the target word was reduced than when it was unreduced. Nevertheless, both in the unreduced and the reduced condition, out of the four words on the screen, the target received the vast majority of gazes, which demonstrates that the competition between the target and the phonological competitor was resolved before target word offset. There was no effect of reduction on the N400 in the ERP data, which suggests that semantic representations of reduced variants were activated as easily as those of unreduced variants. Reduced variants thus induced more competition during processing, but did not slow down comprehension. Competition could be resolved before target word offset most likely because natives relied on phonetic, semantic or syntactic cues provided by the word or the context.

6.3.3 The use of variant frequency in the processing of reduced speech by Dutch learners of French

Dutch learners also participated in the visual lexical decision task with cross-modal identity priming in Chapter 3. As was observed for the French natives, the highly frequent variant primed the visual target better than the infrequently occurring variant for the Dutch advanced learners. The advanced learners listened till the end of the sentence when they heard a reduced pronunciation variant (the frequent, strongly reduced, or the infrequent, weakly reduced variant), as if they waited for more cues (e.g., phonetic, syntactic or semantic information) to facilitate recognition of the presented reduced variants. In contrast, the natives always waited till the end of the sentence except when the prime word was the infrequent, weakly reduced identity prime. The natives thus seemed not to take the remainder of the sentence into account when the prime word was infrequent and therefore difficult to process. The learners, in contrast, applied a different strategy and did not take the remainder of the sentence into account when the prime word was easy to process (unreduced identity prime word or unrelated prime word). Although both participant groups showed a variant frequency effect, different strategies and underlying mechanisms were involved in the processing of variants by both groups.

The variants' frequencies in Chapter 3 were averaged over all observed word types and were extracted from native French casual speech. I did not have word-specific information about the frequencies of the possible variants. Moreover, there was no information about whether the natives and the Dutch learners differed in their expectation of how often the different variants occur. I conducted the studies in Chapter 4 to investigate the role of word-specific frequency information in recognition of reduced words and to investigate whether differences between the frequencies of the different variants in the speech input for natives versus language learners may affect their processing.

The averaged advanced learners' ratings predicted both their accuracy and RTs. Importantly, the advanced learners' accuracy scores were not predicted by the average ratings of the natives. This strongly suggests that how well a listener recognizes a pronunciation variant is determined by how much exposure this listener has had with the given variant. The lexical decision RTs and accuracy of the beginner learners could not be predicted by their own variant frequency ratings, the advanced learners' ratings, nor the French natives' ratings.

Variant frequency may thus serve as a cue in the recognition of reduced speech, even in the speech recognition processes by language learners. Advanced learners probably store frequencies that are not native-like, and beginner learners store only low frequencies which they may not use, or do not store any frequencies at all, because they have not yet encountered the variant.

The findings in Chapter 4 suggest that frequencies reflecting the advanced learners' average exposure to pronunciation variants better predict their speech processing behaviour than the natives' frequencies. In future research, the evidence for sensitivity to variants' frequencies based on a group's exposure may be extended with a visual lexical decision task with cross-modal identity priming containing pronunciation variants with frequencies based on advanced learners' exposure.

6.3.4 The effect of reduction on learners' speech processing mechanisms

The transcription task in Chapter 3 also tested advanced learners of French. The learners' transcriptions partly reflected the segments in the auditory input, but did not reflect the speaker's intended words, which were longer. The learners hardly used phonetic cues to reconstruct the input form. Likewise, the learners did not always make full use of the semantic or syntactic information provided by the context: they provided transcriptions that did not fit into the sentences semantically. The learners transcribed *ça m'arrive de* 'it happens to me' pronounced as [sa.mari.de], for instance, in 8.2% of the transcriptions as *mon mari(e) de* 'my husband of' /mɔ̃.mari.də/. Learners also often encountered word segmentation problems, which resulted in semantically and syntactically incorrect transcriptions. I observed problems in the transcription of clitics: *m'arrive* /m ariv/ in the example fragment *ça m'arrive de* 'it happens to me' already provided above, for instance, was transcribed as *mari de* /mari.də/ 'husband of' in 12.7% of the transcriptions. The transcription task thus shows that highly proficient learners have great difficulty understanding reduced word pronunciation variants in a natural casual conversation. The learners' transcriptions often changed the meaning of the sentence completely, which may have led to a misunderstanding of the conversation. Thus, the comprehension of the learners either reflect their word segmentation problems or insufficient use of phonetic, syntactic or semantic information.

I also tested learners in the combined eye-tracking and ERP experiment of Chapter 5. The eye-tracking data demonstrated that Dutch advanced learners' proportions of fixations on neutral distractors were higher when they heard a reduced target word than when they heard an unreduced target word. This result suggests that there was more competition between the lexical candidates when the listeners heard a reduced variant instead of an unreduced variant. In contrast to lexical competition in the natives, the competition in the learners was not resolved before target word offset. Furthermore, the advanced learners showed more negative amplitudes in the N400 time window for unreduced variants than for reduced variants. My interpretation of the ERP data is informed by the eye movement data. As there was no N400 for the reduced variants in the ERP data and the eye movements showed increased lexical competition for reduced variants, the eye-tracking and ERP data together suggest that the processing of the reduced form is already hindered at the form level for the learners, and not necessarily only at the semantic level. The lexical competition then appears to impede semantic processing: as learners were still processing the reduced form at the form level, activation did not spread to semantic representations.

The results of the eye-tracking and ERP study support those of the transcription task: Dutch learners of French have more problems in accessing lexical-semantic representations of reduced words and seriously consider competing lexical candidates during lexical search.

6.4 Psycholinguistic models

6.4.1 Speech production models

In Chapter 2, I investigated the nature of the processes underlying the reduction of word-final OLS phonemes (see also section 6.1). On the basis of corpus data, I argued that whereas the obstruent is likely to be absent due to gradient processes (extreme shortening), the liquid and the schwa may also be absent due to other (categorical) processes. The question then arises which of the existing speech production models is able to account for these findings.

Since the obstruent is likely to be absent only due to articulatory overlap processes, it is very likely that all pronunciation variants stored for a given word ending in an OLS cluster contain the obstruent of the cluster. Words are probably stored with single abstract representations for the obstruents (e.g., Roelofs, 1997). Since the liquid and schwa, in contrast, may be absent either due to categorical or gradient processes, the mental lexicon may contain word pronunciation variants with and without these segments. In this latter account, a word may thus be lexically stored with more than one pronunciation variant (following, e.g., Bürki & Gaskell, 2012; Bürki, Ernestus & Frauenfelder, 2010). For instance, *ministre* 'minister' may have the lexical representations /ministʁ/ and /minist/.

Alternatively, the liquid and schwa may be categorically absent due to phonological deletion rules (e.g., Dell, 1985; Levelt, Roelofs & Meyer, 1999). During production, application of a phonological deletion rule (e.g., liquid- or schwa-deletion rule) transforms the stored, unreduced variant into a reduced variant. The frequency of occurrence of a reduced variant may modulate the ease of application of the rule that transforms the stored, unreduced variant into the corresponding reduced variant.

The corpus study showed that phrasal position, speech rate and phonetic context affect the absence of segments. This may be because the lexical representations are specified for these variables or because these variables license the phonological deletion rule. The effects of these variables may also arise (in a gradient fashion) during articulation, as they do for the obstruent. If so, future research should show that the variables of phrasal position, speech rate and phonetic context especially affect production when the segment has left acoustic traces.

The abstract rule-based speech production model is not a very plausible account, if I assume that production and perception are related to each other (e.g., Meyer, Huettig & Levelt, 2016). Whereas in speech production, the frequency of occurrence of a reduced variant may modulate the ease of application of the rule that transforms the stored, unreduced variant into the corresponding reduced variant, a similar process cannot account for frequency effects in speech comprehension (see Chapters 3, 4 and Pitt et al., 2011; Ranbom & Connine, 2007). One might argue that in perception, prelexical reconstruction rules may be sensitive to the frequencies with which listeners would apply them themselves to each word in production. That is, upon hearing a word, a listener may apply a reconstruction rule and consider the resulting prelexical representation as likely for the given acoustic input as indicated in the description of the rule. For instance, when a native listener of French hears /ʀvy/ *revue* 'magazine', (s)he may apply schwa-insertion (resulting in /ʀəvy/) and keep the two prelexical representations active until the word is identified. The recognition process would favour the prelexical representation that is most likely given the likelihood that the word is produced with schwa, as indicated by the rule. However, the word-specific probability of an insertion rule can only become available after the word has been identified. If information about the probability of an insertion rule only becomes available after the word has been identified, how then can this probability affect the ease of word recognition?

A different type of speech processing model assumes that all tokens of a word are lexically stored with all their fine phonetic detail (in the form of exemplars, e.g., Bybee, 1998; Goldinger, 1998). Different presentations of pronunciation variants of the OLS phonemes for different voices and situations are lexically stored. The effects of phrasal position, speech rate and phonetic context may again arise from the exact specifications of the exemplars or may arise during articulation.

6.4.2 Speech perception models

The results of the studies presented in Chapters 3, 4, and 5 provided information about the processing of casual speech by natives and foreign language learners which has implications for psycholinguistic models of speech comprehension. These chapters showed, for instance, that the unreduced variant has a special status in learners' processing of words that were only preceded by their determiners and of words in context. The natives only showed a clear processing advantage for the unreduced variants in isolation. Moreover, both participant groups showed variant frequency effects: Natives and advanced language learners recognized more frequent variants faster and more accurately. The question now is whether there is a speech perception model that is able to account for these findings.

Some speech perception models assume that words are stored as single abstract representations. Reduced words (e.g., French [ʀvy] *revue* 'magazine') are matched with the representations for the unreduced variants (e.g., /ʀəvy/), by means of general processes (e.g., schwa-insertion), each applying to multiple words. This abstractionist account is compatible with many psycholinguistic models of word recognition, such as TRACE (McClelland & Elman, 1986), the Neighborhood Activation Model (NAM: Luce & Pisoni, 1998), PARSYN (Luce, Goldinger, Auer & Vitevitch, 2000) and Shortlist B (Norris & McQueen, 2008). Although a rule-based model that only stores unreduced variants can easily account for the privileged status of the unreduced variant in isolation, it cannot easily account for the observed variant frequency effects in Chapters 3 and 4 (see section 6.4.1 above).

An alternative abstract account for the processing of segment reduction is the naive discriminative learning model, as proposed by Baayen (2010). In the naive discriminative learning model, which is a two-layer network with symbolic representations for a word's form and a word's semantics, each input unit (unigram, bigram, trigram, uniphone, biphone or triphone) is linked to each meaning, with specific weights being associated with each link. In order to have an effect on speech comprehension, relative variant frequencies have to be expressed in the weights of the connections between the biphones or triphones of the two variants and the word's meaning. A word pronunciation variant can only have a strong connection with its meaning if (some of) its biphones or triphones do not occur often in other variants of the word, nor in other words. That is, the connection weights of the biphones and triphones of a variant are not only determined by this variant's relative frequency as compared to other variants of the same word, but also by the frequencies of other words and their pronunciation variants. To give an example, the weight of the connection of the biphone /pə/ with the semantics of *pelouse* will also be determined by the frequencies of the unreduced variants of, for instance, *pelote*, *peluche* and *pelure*. Similarly, the connection between the biphone /pl/ and the semantics of *pelouse* will also be determined by the frequencies of the reduced variants of these same words as well as by other words starting with the same biphone (e.g., *plouf*, *ploutocrate*, *pluie* and *plumage*). Hence, in the naive discriminative learning model, word-specific relative frequencies cannot be represented separately from other (relative) frequencies, and therefore these relative frequency effects cannot (simply) emerge. As a result, the naive discriminative learning model cannot easily explain the observed relative frequency effect.

The relative frequency effects can easily be accounted for in models hypothesizing lexical storage of multiple word pronunciation variants. For instance, if Shortlist B (Norris & McQueen, 2008) is adapted in such a way that it stores these variants, the relative frequencies can co-determine the variants' priors and thus how easily these variants are processed. French natives may store the frequencies of pronunciation variants together with the lexical representations of these variants in their mental lexicons. Dutch learners may do so as well, or, if they have not encountered the reduced variant before, they may not store any representation for this variant at all. The presence of the variants' representations and the variants' frequencies change as a function of the type and amount of exposure.

Exemplar-based models assume that fine-grained phonetic detail of each occurrence of a word is stored in the form of exemplars (e.g., Johnson, 2004; Bybee, 2001; Goldinger, 1998). Perception of a pronunciation variant involves the direct mapping of the perceived acoustic event on the exemplars in the mental lexicon. Hawkins and Smith (2001) argue that the acoustic signal is a rich source of linguistic and of non-linguistic information, for instance, the speaker's identity, attitudes, and current state of mind. Following Hawkins and Smith (2001) by assuming that exemplars are specified for contextual information, an exemplar-based model can easily account for the finding in Chapter 3 that unreduced variants lose their privileged status in natives' speech processing when presented in context. Variant frequency effects may result from the frequency of occurrence of a given variant in the episodic lexicon of the listener.

Finally, hybrid theories vary with respect to the degree to which phonetic detail is retained in memory (e.g., Pierrehumbert, 2002). The hybrid model proposed by McLennan, Luce, and Charles-Luce (2003) based on the Adaptive Resonance Theory (ART) assumes abstract representations for lexical and sublexical units while speaker-specific and situation-specific information is stored in the form of exemplars. The frequency effects observed in Chapter 2 could be accounted for by this model by assuming that frequency information is stored together with the abstract representations. Like lexical representations used in speech production (see section 6.4.1), those used in speech comprehension may also be specified for context such that we can explain a possible effect of context on processing speed (see sections 6.2.1 and 6.2.2). The exemplars then would not play a role in accounting for the observed variant frequency effect.

6.5 Methodology

To answer my research questions, different methods were applied. Previous research (e.g., Bürki and colleagues, 2011b) has shown that methodological choices and the selection of a specific task may influence the outcome. I therefore combined different methods to obtain a more complete picture regarding the processing of reduced word pronunciation variants. Below, I discuss how reliable the different methods are and whether the different methods give comparable results.

Chapter 2 combined three different methods to investigate the nature of the processes underlying the production of reduced word-final OLS clusters: phonetic analyses, examination of the distributions of the durations of the OLS phonemes, and comparison of the predictors of absence and duration of the OLS phonemes. Phonetic analyses can only provide evidence for gradient reduction processes, which nearly always play a role. On the basis of this single method, it is impossible to obtain convincing evidence that processes underlying reduction may also be categorical. I therefore also made histograms of the durations of the phonemes. However, since there is no clear definition of what counts as a clearly distinct peak, the interpretation of these histograms may be subjective. The comparison of the predictors of absence and duration of phonemes may also not lead to undisputable conclusions because it is not clear how much overlap in predictors is necessary to conclude that phoneme absence and phoneme duration are driven by the same processes. Since all these methods have their own shortcomings, it is important not to restrict research on speech production to either one of them. The combination of research methods provides a framework for the exploration of reduction patterns in future research.

Chapter 3 reported the results of a transcription task and a lexical decision task with cross-modal identity priming. Both tasks provided information about the processing of reduced speech by natives and foreign language learners. In addition, the results of the first task informed us about whether listeners used phonetic, semantic and/or syntactic information, whereas the results of the second task informed us about whether listeners used frequency information during speech processing. The two methods nicely complemented each other and indicated whether the different cues were used in processing by natives and learners.

Whereas for the experiment in Chapter 3, I only had information on the variants' frequencies averaged over word types produced by natives, I could test the effect of variant frequency based on word-specific and group-specific (learners versus natives) estimations in Chapter 4. I obtained these estimations by conducting a relative frequency estimation task. French natives' average relative frequency ratings for reduced variants correlated well with the ratings given by participants from Nantes collected by Racine (2008) in a similar rating experiment ($r = .83$, $p < .0001$). Racine (2008) obtained frequency estimations from Swiss francophone students that correlated well with the variants' frequencies in the productions of 16 different Swiss francophone students who were asked to summarize stories ($r = .44$, $p < .01$). Two groups of speakers of more or less the same variant of French thus produced highly similar ratings. This finding suggests that the participants tested in Chapter 4 were able to estimate the relative frequencies of occurrence of the two pronunciation variants of a word and that this relative frequency estimation task was a reliable method to use.

Chapter 5 combined two different methods to investigate the on-line processing of unreduced and reduced schwa words by natives and language learners: eye-tracking and ERP. Eye-tracking and ERP are complimentary techniques in that eye movements inform us about the lexical competition in listeners and the N400 ERP component inform us about the lexical-semantic activation of a word in context. The interpretation of the EEG data was informed by the eye movement data, and vice versa.

In sum, Chapters 2, 3, 4 and 5 highlight the importance of combining different methods in the research on speech production and speech perception. One single method cannot provide a complete picture of speech production or speech perception processes.

6.6 Recommendations for foreign language teachers

My findings have important implications for foreign language teaching. Based on these results, I recommend that students are trained on the frequencies of occurrence of pronunciation variants and the use of phonetic, syntactic and semantic cues provided by the word and the context. Students have to be frequently exposed to casual speech in order to learn to use the richness of information contained in the word itself and in its context. Since existing educational programmes often focus on the written language, and existing audio materials presented in the classroom contain considerably fewer reduced pronunciation variants than everyday casual conversations (e.g., Gilmore 2004, Fonseca-Greber & Waugh, 2003; Waugh & Fonseca-Greber, 2002; McCarthy & Carter, 1995), teachers are highly recommended to use more spontaneous speech in the classroom.

Listeners are often unaware of the frequencies of occurrence of reduced pronunciation variants. A meta-analytic study (Goo, Granena, Yilmaz & Novella, 2015) showed that in general explicit instruction may be effective in L2 learning.

Future research could be conducted to determine the effectiveness of explicit instruction in the learning of the relative frequency with which variants occur.

6.7 Conclusion

This dissertation focused on the processing of reduced words by natives and foreign language learners. I underlined the relevance of corpus-based research in the documentation of pronunciation variation in everyday conversations and of the processes underlying this variation. The results of the perception studies (Chapters 3, 4, 5) show that learners encounter more difficulty during the processing of everyday speech than natives. They do not make full use of phonetic, semantic or syntactic cues provided by the word or the context.

Nevertheless, learners are sensitive to variant frequency in their speech input, like native listeners are. However, leaners' speech input differs from native listeners' speech input in the frequencies of reduced variants, which makes that learners are not well tuned to listening to casual speech. The difficulty that learners experience with reduced words thus has several sources.

Nederlandse samenvatting

In alledaagse spraak worden klanken vaak ingeslikt. Dit fenomeen noemen we 'reductie'. Zo kan het Engelse woord *probably* 'waarschijnlijk' worden uitgesproken als *probly*, *proly* of zelfs *pro* (zie bijv., Johnson, 2004). In het Frans kunnen sprekers het woord *fenêtre* 'raam' verkorten tot *fnêtr* (zie bijv., Racine, 2007).

De studies in dit proefschrift hebben tot doel meer inzicht te krijgen in de productie en de perceptie van de verschillende uitspraakvarianten van een woord door moedertaalsprekers en taalleerders. Drie onderzoeksvragen staan centraal. Allereerst wordt nagegaan welke processen ten grondslag liggen aan het verdwijnen van een klank. Er wordt gekeken of een klank volledig afwezig is of dat er restjes zijn achtergebleven in het akoestische signaal. Het antwoord op deze eerste vraag kan evidentie vormen voor een bepaald spraakproductie-model.

De tweede onderzoeksvraag heeft betrekking op het begrip van woorden waarvan een of meerdere klanken afwezig zijn. De vraag is of moedertaalsprekers evenveel last hebben van het wegvallen van klanken als taalleerders.

De derde onderzoeksvraag betreft de mechanismes die ten grondslag liggen aan spraakperceptie-processen van moedertaalsprekers en leerders. Ik onderzoek of beide groepen luisteraars gebruikmaken van dezelfde informatie (bijv., informatie over de frequentie van een uitspraakvariant) bij het interpreteren van gereduceerde woorden. Er wordt ook gekeken of luisteraars meer woorden overwegen wanneer ze een gereduceerd woord hebben gehoord dan wanneer ze een ongereduceerd woord hebben gehoord.

Daarnaast kijk ik of reductie de activatie van de betekenis van woorden verstoort. Het antwoord op de derde onderzoeksvraag maakt duidelijk waarom taalleerders mogelijk meer moeite hebben met het begrijpen van gereduceerde spraak dan moedertaalsprekers en kan ons helpen bij het definiëren van een model waarmee de perceptie van alledaagse spraak door moedertaalsprekers en leerders kan worden beschreven. In dit proefschrift focus ik me op het Frans, zodat we meer te weten komen over reductie in Romaanse talen.

Uitspraakvarianten voor woord-finale obstruent-liquid-schwa clusters

Hoofdstuk 2 rapporteert de verschillende uitspraakvarianten voor woord-finale obstruent-liquid-schwa clusters (voortaan OLS, bijv., *tre* in *ministre* 'minister'). Ik vond zeven uitspraakvarianten, waarvan er drie zelden voorkomen. De vorm waarin alle OLS klanken nog aanwezig zijn, kwam in 18.9% van de 291 bestudeerde woord tokens voor (bijv., *ministre* voor *ministre* 'minister'). De variant zonder schwa was de meest frequente variant en kwam voor in 35.7% van de tokens (bijv., *ministr*). De variant waarin zowel de liquid als de schwa afwezig zijn was daarna het meest frequent (27.1%, bijv. *minist*). Opvallend genoeg viel de complete OLS cluster ook vaak weg (15.5%; bijv., *minis*). In de varianten die regelmatig voorkomen (bijv., *ministr*, *minist*, *minis*) is een OLS-klank afwezig en alle volgende klanken ook. In de varianten die zelden voorkomen (bijv., *minisre*, *minisr*, *ministe*) is een OLS-klank afwezig en is minimaal een van de volgende klanken aanwezig.

Een mogelijke verklaring voor dit geobserveerde reductiepatroon is dat klanken aan het einde van een woord samensmelten met de eerste klanken van het volgende woord die over het algemeen heel duidelijk worden gearticuleerd (zie bijv., Keating, Cho, Fougeron & Hsu, 2003). Een andere verklaring is dat sommige klanken sneller wegvallen dan andere klanken (schwas vallen sneller weg dan liquida en liquida verdwijnen sneller dan obstruenten, zie bijv., Adda-Decker et al., 2008; Su & Basset, 1998). Deze hiërarchie hangt mogelijk samen met het gemak waarmee de klanken kunnen worden gereconstrueerd door de luisteraar.

Processen die ten grondslag liggen aan reductie

We kunnen grofweg twee typen processen onderscheiden die ervoor zorgen dat klanken niet uitgesproken worden. Het eerste proces is een gradueel proces dat leidt tot een continuum aan uitspraakvarianten met een ongereduceerde klank aan de ene kant van het continuum (bijv., *e* in *renard* 'vos') en de afwezigheid van deze klank aan de andere kant (bijv., *e* is afwezig in *rnard*). Een woord kan met een licht tot sterk gereduceerde klank worden geproduceerd. Gradueel gereduceerde klanken kunnen sporen achterlaten in het akoestische signaal, bijvoorbeeld in de duur of de spectrale eigenschappen van de omliggende klanken.

Het tweede proces is een categoriaal proces dat zich kenmerkt door de volledige afwezigheid van een klank. Hierbij worden geen sporen achtergelaten in het akoestische signaal. Het proces dat ten grondslag ligt aan een gereduceerde klank is zelden enkel categoriaal van aard is. Een klank kan ook volledig afwezig zijn doordat zijn articulatie extreem verkort is en samenvalt met die van een vorige of volgende klank.

In hoofdstuk 2 zijn drie verschillende methoden gebruikt om na te gaan of de processen die ten grondslag liggen aan de reductie van de woord-finale obstruent, liquid en schwa (bijv., *tre* in *ministre* 'minister') gradueel of categoriaal zijn. Ten eerste is gekeken of de gereduceerde obstruent, liquid en schwa restjes achterlieten in het signaal. Ten tweede is de verdeling van de klankduren van de obstruent, liquid en schwa nader bekeken. Wanneer de verdeling twee pieken vertoont, waarvan een rond de 0 ms (staat gelijk aan de afwezigheid van de klank), en er een duidelijk dal tussen de pieken zit, is het waarschijnlijk dat de afwezigheid van de klank het gevolg is van een categoriaal reductieproces. Ten slotte heb ik gekeken of de duur en de afwezigheid van een klank door dezelfde variabelen konden worden voorspeld. Als dat het geval is, is de afwezigheid van de klank waarschijnlijk het gevolg van het extreem verkorten van zijn articulatie en dus van een gradueel reductieproces.

De verschillende methoden geven een beeld van het proces dat ten grondslag ligt aan de reductie van de obstruent, de liquid en de schwa. Op basis van de analyse van het akoestische signaal en van de klankduren, zou ik kunnen concluderen dat het reductieproces dat ten grondslag ligt aan de afwezigheid van de obstruent met name gradueel is: de obstruent liet vaak sporen achter in het akoestische signaal en liet geen duidelijk dal zien tussen de pieken in de verdeling van de klankduren. De vergelijking van de predictoren van de afwezigheid van de obstruent met die van de duur van de obstruent suggereert dat de afwezigheid en de duur van de obstruent voornamelijk gevoelig waren voor verschillende variabelen en daarom uit verschillende processen resulteerden. Deze bevindingen impliceren dus dat de afwezigheid van de obstruent niet enkel resulteerde uit graduele reductie. Het inconsistente patroon van resultaten is wellicht te wijten aan het feit dat de dataset enkel uit 291 woord tokens bestaat waardoor niet met zekerheid vast te stellen is welke factoren de afwezigheid en duur van de obstruent beïnvloeden.

De akoestische analyses lieten zien dat een afwezige liquid ook sporen kon achterlaten in het signaal. De voorafgaande obstruent kon bijvoorbeeld stemhebbend worden. De verdeling van de duur van de liquid en de vergelijking van de predictoren van de afwezigheid en de duur van de liquid suggereren dat de processen die ten grondslag liggen aan de reductie van de liquid categoriaal zijn: er waren twee pieken met een duidelijk dal ertussen in de verdeling van de klankduren en geen substantiele overlap van predictoren (maar één van de vier predictoren kon zowel de duur als de afwezigheid van de liquid voorspellen). De reductieprocessen die ten grondslag lagen aan de afwezigheid van de liquid waren vermoedelijk dus zowel gradueel als categoriaal.

In mijn data liet de afwezige schwa geen restjes achter in het akoestische signaal. Door het geringe aantal datapunten (60) kon geen significante predictor van de duur van de schwa worden gevonden, maar wel voor de aanwezigheid van de schwa. De schwa was vaker afwezig als de spreeksnelheid hoger was, het woord frequenter en de schwa werd gevolgd door een klinker. De verdeling van de duur van de schwa liet een duidelijke piek rond 0 ms zien. Dit beeld suggereert dat het reductieproces van de schwa categoriaal was. Aan de afwezigheid van de schwa in woord-finale positie liggen dus net als in andere posities in het woord niet alleen graduele processen ten grondslag, maar ook categoriale processen.

Alhoewel de processen die ten grondslag liggen aan de afwezigheid van opeenvolgende klanken kunnen verschillen, tonen onze resultaten aan dat de afwezigheid van één afzonderlijke klank wel de aanwezigheid van een omringende klank kan beïnvloeden. Een klank is sneller afwezig als de volgende klank ook afwezig is. Zoals hierboven al werd gesuggereerd, komt dit wellicht omdat klanken aan het einde van het woord samensmelten met de eerste klanken van het volgende woord die over het algemeen heel duidelijk worden gearticuleerd (zie bijv., Keating, Cho, Fougeron & Hsu, 2003) of omdat sommige klanken vaker afwezig zijn dan andere. Toekomstig onderzoek zal moeten uitwijzen welke verklaring het meest waarschijnlijk is. Een combinatie van deze twee verklaringen kan ook tot het geobserveerde reductiepatroon leiden.

Het verstaan van gereduceerde varianten door moedertaalsprekers en leerders

In dit proefschrift zijn verschillende methodes gebruikt om het verstaan van gereduceerde uitspraakvarianten te onderzoeken. Hoofdstuk 3 rapporteert de resultaten van een studie waarin Franse moedertaalsprekers en Nederlandse studenten Frans werden onderworpen aan een dictee en een lexicale decisie taak. Het dictee bevatte zinnen die uit spontane conversaties waren geknipt. De zinnen bestonden uit woorden waarvan enkele of meerdere klanken weggevallen waren. De leerders van het Frans maakten veel meer fouten in het dictee dan de Fransen (48.5% versus 88.0% correct). Hun transcripties kwamen overeen met de gepresenteerde zinnen, maar misten segmenten. Leerders waren vaak niet in staat de gereduceerde klanken te reconstrueren. Ze transcribeerden *tu vois* 'jij ziet' bijvoorbeeld als *tu as* 'jij hebt' in 12.7% van de gevallen. De leerders hadden ook moeite met het opsplitsen van spraak in afzonderlijke woorden en gaven transcripties die de betekenis van de zin compleet veranderden. Ze interpreteerden bijvoorbeeld *m* in *ça m'arrive de* 'het overkomt me' als de eerste klank van het woord *mari* 'echtgenoot' in 12.7% van de transcripties.

In de tweede methode, die bestond uit een lexicale decisie taak, moesten deelnemers telkens aangeven of ze dachten dat een woord dat werd gepresenteerd op een computerscherm wel of niet bestond. Het visuele woord eindigde in een obstruent-liquid-schwa cluster (bijv., *tre* in *ministre* 'minister'). Voorafgaand aan elk woord kregen de deelnemers telkens een zin te horen die een woord bevatte dat wel of niet (bijv., *virage* 'bocht' voor *ministre*) overeenkwam met het visuele woord in de lexicale decisie taak. Wanneer het overeenkwam met het visuele woord (bijv., *ministre*), werd het woord ofwel ongereduceerd (bijv., *ministre* of *ministr*), ofwel sterk gereduceerd (bijv., *minis*), ofwel licht gereduceerd (bijv., *minisr*) aangeboden. Terwijl de leerders in de lexicale decisie taak het snelst reageerden op het visuele woord nadat ze een ongereduceerde variant (bijv., *ministre* of *ministr*) van het betreffende woord hadden gehoord, leidden bij de Fransen zowel de ongereduceerde variant als de veel voorkomende, sterk gereduceerde variant (bijv., *minis*) tot snellere reactietijden op het visuele woord in de lexicale decisie taak. Deze gereduceerde variant beïnvloedde het perceptieproces van de taalleerders, maar niet dat van de moedertaalsprekers.

Hoofdstuk 4 bevat de resultaten van een studie waarin Franse moedertaalsprekers en leerders van het Frans met twee verschillende beheersingsniveaus een auditieve lexicale decisie taak hebben uitgevoerd. Deze taak bevatte ongereduceerde en gereduceerde varianten van woorden met een schwa (*e*) in de eerste lettergreep (bijv., *renard* vs. *rnard* voor *renard* 'vos'). Elk woord werd voorafgegaan door een bepaald lidwoord (*le* of *la*). Franse moedertaalsprekers maakten meer fouten bij de gereduceerde varianten (88.4% correct) dan bij de ongereduceerde varianten (99.0% correct) en hadden meer tijd nodig om te antwoorden op een gereduceerde variant (gemiddelde reactietijd, voortaan RT: 1217.99 ms) dan op een ongereduceerde variant (gemiddelde RT: 1153.96 ms). De Nederlandse beginnende leerders van het Frans, met een B1-B2 niveau volgens het Common European Framework of Reference for Languages (CEFR, Council of Europe, 2011), maakten ook meer fouten bij gereduceerde (35.0% correct) dan bij ongereduceerde varianten (73.1% correct) en ze antwoordden langzamer op gereduceerde varianten (gemiddelde RT: 1404.49 ms) dan op ongereduceerde varianten (gemiddelde RT: 1355.25 ms). De Nederlandse gevorderde leerders, met een C1-C2 niveau volgens het CEFR, hadden ook meer moeite met het verstaan van gereduceerde uitspraakvarianten (64.3% correct) dan met het verstaan van ongereduceerde uitspraakvarianten (89.2% correct) en ze reageerden sneller op ongereduceerde varianten (gemiddelde RT: 1281.66 ms) dan op gereduceerde varianten (gemiddelde RT: 1358.19 ms). Opvallend is dat alle groepen deelnemers meer fouten maakten en langzamer reageerden bij de gereduceerde woorden dan bij de ongereduceerde woorden. De leerders ondervonden duidelijk meer problemen met het verstaan van gereduceerde uitspraakvarianten dan moedertaalsprekers.

Ik heb de mechanismes die schuilgaan achter het verstaan van gereduceerde spraak door Franse moedertaalsprekers en Nederlandse gevorderde leerders van het Frans ook proberen te ontrafelen aan de hand van een gecombineerde eye-tracking en Event-Related-Potential (ERP) studie (hoofdstuk 5). Eye-tracking is een onderzoeksmethode waarbij oogfixaties en oogbewegingen worden gemeten waardoor het kijkgedrag van een proefpersoon gedurende een experiment kan worden bestudeerd. De visuele resultaten laten zien welke delen van een computerscherm door een proefpersoon worden bekeken en hoe lang ze worden bekeken. Tijdens mijn experiment kregen de deelnemers zinnen te horen met ongereduceerde en gereduceerde varianten van woorden (bijv., *fenêtre* of *fnêtre* voor *fenêtre* 'raam') die overeenkwamen met één van de vier afbeeldingen van de woorden gepresenteerd op een computerscherm. De deelnemers kregen verder een plaatje te zien van een woord dat met dezelfde klank begon (bijv., *fourchette* 'vork') en twee plaatjes van neutrale woorden die qua klanken en betekenis niets met het gepresenteerde woord te maken hadden (bijv., *montre* 'horloge', *coccinelle* 'lieveheersbeestje'). Door de oogfixaties tijdens het luisteren te meten kon ik een idee krijgen van de mate van competitie tussen de verschillende kandidaten. De eye-tracking data lieten zien dat de Nederlandse gevorderde leerders vaker naar de neutrale (ongerelateerde) kandidaten (bijv., *montre* 'horloge', *coccinelle* ‘lieveheersbeestje’) keken wanneer ze een geredueerde uitspraakvariant hadden gehoord (bijv., *fnêtre*) dan wanneer ze een ongereduceerde uitspraakvariant (*fenêtre*) hadden gehoord. Er vond dus meer competitie plaats tussen de verschillende kandidaten wanneer leerders een gereduceerde uitspraakvariant te horen kregen. Dit resultaat toont aan dat zelfs gevorderde leerders problemen ondervinden bij het verstaan van gereduceerde woorden in zinnen.

Naast de oogfixaties werd de hersenactiviteit van de deelnemers via Elektro-encefalografie (EEG) gemeten. De N400 is een negatieve golf in het EEG die ongeveer 400 ms optreedt na presentatie van een woord dat qua betekenis afwijkt. In de zin *The pizza was too hot to cry* ‘De pizza was te heet om te huilen’ zal het woord *cry* een N400 veroorzaken, omdat het qua betekenis niet aansluit bij de strekking van de zin (zie bijv., Kutas & Hillyard, 1980). In de zin *The pizza was too hot to eat* ‘De pizza was te heet om te eten’ zal het woord *eat* geen N400 veroorzaken. In mijn gecombineerde experiment lieten de *event-related potentials* (ERPs) die de elektrofysiologische reacties van de hersenen op de gepresenteerde stimuli weergeven, zien dat de leerders, in tegenstelling tot de moedertaalsprekers, de betekenissen voor de gereduceerde varianten niet activeerden.

Deze resultaten sluiten aan bij die in hoofdstuk 4: Nederlandse gevorderde leerders van het Frans hebben meer moeite met het verstaan van gereduceerde varianten dan met het verstaan van ongereduceerde varianten van een woord. Het dictee en de eye-tracking taak laten zien dat dit waarschijnlijk komt doordat andere woorden worden geactiveerd wanneer gereduceerde uitspraakvarianten worden gepresenteerd dan wanneer ongereduceerde varianten worden gepresenteerd. Dit zorgt ervoor dat de betekenis van het bedoelde woord niet kan worden geactiveerd.

De frequenties van uitspraakvarianten

In de hoofdstukken 3 en 4 is onderzocht of moedertaalsprekers en taalleerders gebruikmaken van de relatieve frequenties van uitspraakvarianten tijdens het perceptieproces. In hoofdstuk 3 werd gekeken of luisteraars sneller reageren op visuele woorden die voorafgegaan worden door uitspraakvarianten die vaak voorkomen in Franse informele spraak (bijv., *minis* voor *ministre*) dan op visuele woorden die voorafgegaan worden door laag freqente varianten (bijv., *minisr* voor *ministre*). Voor de frequenties van de uitspraakvarianten baseerde ik me op de resultaten uit hoofdstuk 2 waarbij de frequenties van verschillende uitspraakvarianten van OLS clusters werden gemiddeld over alle 291 geobserveerde woorden.

Ik vond dat Franse luisteraars sneller reageren op visuele woorden waaraan hoog frequente varianten voorafgaan dan op visuele woorden waaraan laag frequente varianten voorafgaan. De Nederlandse gevorderde taalleerders lieten ook een effect zien van de frequenties van uitspraakvarianten. In tegenstelling tot de moedertaalsprekers, luisterden de leerders ook tot en met het einde van de zin wanneer ze licht gereduceerde uitspraakvarianten hoorden. Wellicht wachtten ze op meer informatie (fonetische, syntactische of semantische informatie) om de gepresenteerde gereduceerde uitspraakvariant te kunnen duiden. Alhoewel beide groepen een effect lieten zien van frequentie, gebruikten ze dus verschillende strategieën en gingen verschillende onderliggende mechanismes schuil achter het verstaan van gereduceerde varianten.

Hoofdstuk 4 beschrijft een studie waarin Franse moedertaalsprekers en twee groepen Nederlandse leerders (beginnende en gevorderde) twee taken uitvoerden: een auditieve lexicale decisie taak en een taak waarbij de relatieve frequenties van uitspraakvarianten moesten worden geschat. De relatieve frequenties van twee uitspraakvarianten van een woord met een schwa in de eerste lettergreep moest worden geschat (bijv., *la revue* vs. *la rvue* voor *la revue* ‘het tijdschrift’) op een schaal van 1 tot 6. Vervolgens werd onderzocht of de relatieve frequenties de reactietijden van de deelnemers en hun scores in de lexicale decisie taak konden voorspellen. Het gemiddelde van de door de Fransen geschatte frequenties voorspelde zowel hun RTs als hun scores voor de twee uitspraakvarianten in de lexicale decisie taak: hoe hoger de geschatte frequenties, des te beter en sneller reageerden de Fransen. Het gemiddelde van de relatieve frequenties van de Nederlandse gevorderde leerders van het Frans voorspelde ook zowel hun RTs als hun scores. Het gemiddelde van de relatieve frequenties van de moedertaalsprekers kon de RTs en scores van de leerders echter niet voorspellen. De woord-specifieke relatieve frequenties zijn gerelateerd aan de mate van blootstelling aan de verschillende uitspraakvarianten en verschillen daarom tussen groepen luisteraars met verschillende beheersingsniveaus. Deze resultaten impliceren dat luisteraars die zijn blootgesteld aan gereduceerde spraak gevoelig zijn voor de relatieve frequenties van uitspraakvarianten en ook van deze informatie gebruikmaken tijdens het perceptieproces. Het is zeer waarschijnlijk dat luisteraars de verschillende varianten van een woord samen met hun frequenties opslaan. Gevorderde leerders slaan waarschijnlijk frequenties op die afwijken van die van de moedertaalsprekers. De RTs en de scores van de beginnende leerders van het Frans konden niet worden voorspeld door het gemiddelde van hun eigen relatieve frequenties, noch door het gemiddelde van de frequenties van de gevorderde leerders of die van de Franse moedertaalsprekers. Beginnende leerders slaan gereduceerde varianten alleen op met lage frequenties of helemaal niet als ze deze varianten nog nooit gehoord hebben.

Reductie in spraakproductie-modellen

In hoofdstuk 2 observeerde ik dat de afwezigheid van de obstruent in een woord-finale OLS cluster vaak het resultaat is van graduele reductieprocessen. De liquid en de schwa kunnen ook verdwijnen door categoriale reductieprocessen. De vraag rijst nu hoe bestaande spraakproductie-modellen deze bevindingen kunnen beschrijven.

In abstracte modellen zouden alle opgeslagen uitspraakvarianten de obstruent moeten bevatten. In een informele situatie kan de motoriek van de spraakorganen afnemen (bijv., in het midden van de zin of door verhoogde spreeksnelheid) waardoor de obstruent kan gaan samensmelten met een omringende klank of uiteindelijk zelfs kan verdwijnen. De liquid en de schwa kunnen ook afwezig zijn door categoriale reductieprocessen. Het mentale woordenboek bevat mogelijk zowel uitspraakvarianten met als zonder liquid en schwa. Het woord *ministre* kan bijvoorbeeld de volgende lexicale representaties hebben: *ministr* en *minist*.

Een ander plausibel model veronderstelt dat alle uitspraakvarianten van een woord worden opgeslagen in het mentale woordenboek met al hun fonetische details (in de vorm van *exemplars*, Bybee, 1998; Goldinger, 1998). Volgens *exemplar-based* modellen bevat het mentale woordenboek voor verschillende stemmen en situaties afzonderlijke representaties van de uitspraakvarianten van de obstruent, liquid en schwa.

Reductie in spraakperceptie-modellen

De vraag rijst welk spraakperceptie-model kan verklaren dat de ongereduceerde uitspraakvariant minder problemen oplevert in het perceptieproces van leerders en dat ook de frequentie-effecten in het perceptieproces van leerders en moedertaalsprekers kan beschrijven. Shortlist B (Norris & McQueen, 2008) kan zodanig worden aangepast dat het verschillende uitspraakvarianten met hun frequenties opslaat. Beginnende leerders die veel minder moeite hebben met het begrijpen van ongereduceerde uitspraakvarianten dan van gereduceerde varianten, zullen de ongereduceerde uitspraakvarianten met hogere relatieve frequenties opslaan dan de gereduceerde varianten. Het is ook mogelijk dat ze de gereduceerde varianten nog helemaal niet hebben opgeslagen omdat ze deze nog niet gehoord hebben.

Net als bij productie (zie hierboven) kunnen bij perceptie de fonetische details van elke uitspraakvariant worden opgeslagen in de vorm van *exemplars* (bijv., Johnson, 2004; Bybee, 2001; Goldinger, 1998). Tijdens de perceptie van een uitspraakvariant van een woord wordt het akoestische signaal direct gekoppeld aan bijbehorende *exemplars* in het mentale woordenboek. In dit model zullen de beginnende leerders ongereduceerde uitspraakvarianten opslaan in de vorm van exemplars met hogere relatieve frequenties dan gereduceerde varianten (die ze mogelijk helemaal niet opgeslagen hebben). Hawkins en Smith (2001) beschouwen het akoestische signaal als een rijke bron van linguïstische en niet-linguïstische, bijvoorbeeld contextuele, informatie. De resultaten in hoofdstuk 3 lieten zien dat moedertaalsprekers ongereduceerde uitspraakvarianten in context niet sneller verwerken dan hoog frequente, gereduceerde varianten. Een *exemplar-based* model kan dit resultaat verklaren door aan te nemen dat de wolk van *exemplars* niet alleen gespecificeerd is voor relatieve frequentie, maar ook voor contextuele informatie.

Algemene conclusies

Dit proefschrift beschrijft de variëteit aan uitspraakvarianten in Franse spontane spraak en hoe moedertaalsprekers en taalleerders deze varianten produceren en verstaan. De resultaten in de gepresenteerde studies laten zien dat zelfs gevorderde leerders problemen ondervinden bij het verstaan van gereduceerde spraak. Leerders maken minder vaak gebruik van fonetische, semantische en syntactische informatie die het woord of de context verschaffen. Wanneer een gereduceerde uitspraakvariant wordt gepresenteerd, activeren leerders vaak andere woorden. Hierdoor kan de betekenis van het bedoelde woord niet worden geactiveerd. Niettemin zijn leerders, net als moedertaalsprekers, gevoelig voor informatie over de frequenties van uitspraakvarianten van een woord in hun eigen input. De input van leerders verschilt van die van moedertaalsprekers, hetgeen leerders geen efficiënte luisteraars van informele spraak maakt. De problemen die leerders ervaren tijdens het begrijpen van informele spraak kent dus verschillende oorzaken.

References

Adda-Decker, M., Boula de Mareüil, P., Adda, G., & Lamel, L. (2005). Investigating syllabic structures and their variation in spontaneous French. *Speech Communication*, *46*, 119–139.

Adda-Decker, M., Boula de Mareüil, P., & Lamel, L. (1999). Pronunciation variants in French: schwa & liaisons. *Proceeding of the ICPhS-1999*, *14*, 2239–2242.

Adda-Decker, M., Gendrot, C., & Nguyen, N. (2008). Contributions du traitement automatique de la parole à l'étude des voyelles orales du français. *Revue TAL*, vol.49, *n°3*, 13–46

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. *Proceedings of isit-1973*, 267–281. Budapest, Hungary.

Alegre, M., & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language*, *40*, 41–61.

Allopenna, P.D., 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(4)*, 419–439.

Altmann, G.T.M., & Kamide, Y. (2004). Now you see it, now you don't: Mediating the mapping between language and the visual world. In: F. Ferreira (Ed.), *The integration of language, vision, and action: Eye movements and the visual world* (pp. 347–386). New York: Psychology Press.

Armstrong, N. (1998). The sociolinguistic gender pattern in French: a comparison of two linguistic levels, *Journal of French Language Studies*, *8*, 139–158.

Askildson, V. (2008).*What do teachers and students want from a foreign language textbook?* (Doctoral dissertation). ProQuest Information and Learning Company, Ann Arbor.

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

Baayen, R.H. (2010). Assessing the processing consequences of segment reduction in Dutch with naive discriminative learning. *Lingue & Linguaggio*, *9(2)*, 95–112.

Baayen, R.H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59(4)*, 390–412.

Baayen, R.H., McQueen, J., Dijkstra, A., & Schreuder, R. (2003). Frequency effects in regular inflectional morphology: Revisiting Dutch plurals. In R.H. Baayen, & R. Schreuder (Eds.), *Morphological Structure in Language Processing* (pp. 355–390). Berlin: Mouton de Gruyter (Trends in linguistics. Studies and monographs, 151).

Ball, M., & Rahilly, J. (1999). *Phonetics. The science of speech*. New York: Oxford University Press.

Balota, D.A., Pilotti, M., & Cortese, M.J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition*, *29*, 639–647.

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

Bates, D.M. (2005). Fitting linear mixed models in R. *R News*, *5*, 27–30.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67(1)*, 1–48. doi:10.18637/jss.v067.i01.

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

Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory, M., & Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America*, *113(2)*, 1001–1024.

Boersma, P., & Weenink, D. (2014). *Praat: doing phonetics by computer* (Computer program). Version 5.4, retrieved from http://www.praat.org.

Bradlow A.R., & Alexander, J.A (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *Journal of the Acoustical Society of America*, *121(4)*, 2339–2349.

Brand, S., & Ernestus, M. (2015). Reduction of obstruent-liquid-schwa clusters in casual French. *Proceedings of the ICPhS-2015*, *18*. Scotland, United Kingdom.

Brouwer, S., Mitterer, H., & Huettig, F. (2012a). Can hearing *puter* activate *pupil*? Phonological competition and the processing of reduced spoken words in spontaneous conversations. *Quarterly Journal of Experimental Psychology*, *65*, 2193–2220.

Brouwer, S., Mitterer, H., & Huettig, F. (2012b). Speech reductions change the dynamics of competition during spoken word recognition. *Language and Cognitive Processes*, *27(4)*, 539–571.

Browman, C.P., & Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston & M.E. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech*, 341–376. Cambridge: Cambridge University Press.

Brysbaert, M. (2013). LEXTALE_FR: A fast, free, and efficient test to measure language proficiency in French. *Psychologica Belgica*, *53*, 23–37.

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.

Bürki, A., Ernestus, M., Gendrot, C., Fougeron, C., & Frauenfelder, U.H. (2011a). What affects the presence versus absence of schwa and its duration: A corpus analysis of French connected speech. *Journal of the Acoustical Society of America*, *130*, 3980–3991.

Bürki, A., Fougeron, C., Gendrot, C., & Frauenfelder, U.H. (2011b). Phonetic reduction versus phonological deletion of French schwa: some methodological issues. *Journal of Phonetics*, *39*, 279–288.

Bürki, A., Fougeron, C., & Gendrot, C. (2007). On the categorical nature of the process involved in schwa elision in French. *Proceedings of Interspeech-2007*, 1026–1029. Antwerp, Belgium.

Bürki, A., & Gaskell, G. (2012). Lexical representation of schwa words: two mackerels, but only one salami. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38 (3)*, 617–631.

Bybee, J. (1998), The Phonology of the lexicon: Evidence from lexical diffusion. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of language.* Stanford, California: CSLI Publications.

Bybee, J. (2002). Phonological evidence for exemplar storage of multiword sequences. *Studies in Second Language Acquisition*, *24(2),* 215–221.

Bybee, J. (2015). *Langauge Change*. Cambridge: Cambridge University Press.

Bybee, J., & Scheibmann, J. (1999). The effect of usage on degree of constituency: the reduction of *don't* in American English. *Linguistics*, *37(4)*, 575–596.

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper Row.

Clashen, H. (1999). Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavorial and Brain Sciences*, *22*, 991–1060.

Colantoni, L., & Steele, J. (2005). Phonetically-driven epenthesis asymmetries in French and Spanish obstruent-liquid clusters. In R.S. Gess & E.J. Rubin (Eds.), *Theoretical and experimental approaches to Romance linguistics* (pp. 77–96). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Colantoni, L., & Steele, J. (2007). Voicing-dependent cluster simplification asymmetries in Spanish and French. In P. Prieto, M.J. Sol & J. Mascar (Eds.), *Segmental and prosodic issues in Romance linguistics* (pp. 109–129). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Colantoni, L. & Steele, J (2011). Synchronic evidence of a diachronic change: voicing and duration in French and Spanish stop-liquid clusters. *The Canadian Journal of Linguistics*, *56(2)*, 147–177.

Coquillon, A. (2007). Le français parlé à Marseille : exemple d'un locuteur PFC. In S. Detey & D. Nouveau (Eds.), *PFC: enjeux descriptifs, théoriques et didactiques* (Bulletin PFC 7, pp. 145–156). HAL Id: hal-00250273.

Côté, M.-H. (2004). Consonant cluster simplification in Québec French. *Probus*, *16*, 151–201.

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

Council of Europe (2011). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Council of Europe. Retrieved from http://www.coe.int/t/dg4/linguistic/cadre1_en.asp

Cutler, A., Sebastián-Gallés, N., Soler-Vilageliu, O., & van Ooijen, B. (2000). Constraints of vowels and consonants on lexical selection: Cross-linguistic comparisons. *Memory & Cognition*, *28*, 746–755.

Dahan, D., Magnuson, J.S., Tanenhaus, M. K., & Hogan, E.M. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes*, *16*, 507–534.

Dalby, J.M., 1984. *Phonetic structure of fast speech in American English*. PhD dissertation Indiana University.

Dalgaard, P. (2002). *Introductory Statistics with R*. USA: Springer.

Dell, F. (1995). Consonant clusters and phonological syllables in French. *Lingua*, *95*, 5–26.

Dell, G.S. (1985). Positive feedback in hierarchical connectionist models: applications to language production. *Cognitive Science*, *9*, 3–23.

Drijvers, L., Mulder, K., & Ernestus, M. (2016). Alpha and gamma band oscillations index differential processing of acoustically reduced and full forms. *Brain and Language*, *153-154*, 27–37.

Durand, J., & Eychenne, J. (2004). Le schwa en français : pourquoi des corpus? *Corpus*, *3*, 311–356.

Ernestus, M. (2000). *Voice assimilation and segment reduction in casual Dutch. A corpus-based study of the phonology-phonetics interface* (Doctoral dissertation, LOT, Utrecht).

Ernestus, M., & Baayen, R.H. (2007). The comprehension of acoustically reduced morphologically complex words: The roles of deletion, duration and frequency of occurrence. *Proceedings of the 16th International Congress of Phonetic Sciences-2007*, 773–776. Saarbrücken, Germany.

Ernestus, M., Baayen, H., & Schreuder, R. (2002). The recognition of reduced word forms. *Brain and language*, *81*, 162–173.

Ernestus, M., & Cutler, A. (2015). BALDEY: A database of auditory lexical decisions. *Quarterly Journal of Experimental Psychology*, *68*, 1469–1488.

Ernestus, M., Dikmans, M., & Giezenaar, G. (in press). Advanced second language learners experience difficulties processing reduced word pronunciation variants. *Dutch Journal of Applied Linguistics*.

Ernestus, M., Kouwenhoven, H., & Van Mulken, M. (2017). The direct and indirect effects of the phonotactic constraints in the listener's native language on the comprehension of reduced and unreduced word pronunciation variants in a foreign language. *Journal of Phonetics*, *62*, *50–64*.

Ernestus, M., & Smith, R. (forthcoming). 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/Boston: De Gruyter Mouton.

Ernestus, M., & Warner, N. (2011). An introduction to reduced pronunciation variants [Editorial]. *Journal of Phonetics*, *39(SI)*, 253–260.

Eychenne, J., & Putska, E. (2007). The initial position in Southern French: elision, suppletion, emergence. In J.-P. Angoujard, & O. Crouzet (Eds.), *Proceedings of JEL-2007* (pp. 199–204). Nantes: Université de Nantes.

Faraway, J. (2006). *Extending linear models with R: generalized linear mixed effects and nonparametric regression models*. Boca Raton, FL: Chapman and Hall/CRC.

Fleiss, J.L., Levin, B., & Paik, M.C. (2003). *Statistical Methods for Rates and Proportions*, 3rd Edition. New York: John Wiley & Sons.

Fonseca-Greber, B., & Waugh, L. (2003). On the radical difference between the subject personal pronouns in written and spoken European French. In P. Leistyna, & C. Meyer (Eds.), *Corpus Analysis: Language Structure and Language Use* (pp. 225–240). Amsterdam: Editions Rodopi B.V.

Fougeron, C., Goldman, J.-P., & Frauenfelder, U.H. (2001). Liaison and schwa deletion in French: an effect of lexical frequency and competition? *Proceedings of Eurospeech-2001*, 639–642. Aalborg, Danemark.

Gaskell, G., Spinelli, E., & Meunier, F. (2002). Perception of resyllabification in French. *Memory and Cognition*, *30*, 798–810.

Ghyselinck, M., Lewis, M.B., & Brysbaert, M. (2004). Age of acquisition and the cumulative-frequency hypothesis: A review of the literature and a new multi-task investigation. *Acta Psychologica*, *115*, 43–67.

Gilmore, A. (2004). A comparison of textbook and authentic interactions. *English Language Teaching Journal*, *58(4)*, 363–374.

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

Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. London: Academic Press.

Goo, J., Granena, G., Yilmaz, Y., & Novella, M (2015). Implicit and explicit instruction in L2 learning. Norris & Ortega (2000) revisited and updated. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages, Studies in Bilingualism* (pp. 443–482). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Gopal, H. (1990). Effects of speaking rate on the behavior of tense and lax vowel duration. *Journal of Phonetics*, *18*, 497–518.

Greenberg, S. (1998). Speaking in shorthand - A syllable-centric perspective for understanding pronunciation variation. *Proceedings of the ESCA Workshop on Modeling Pronunciation Variation for Automatic Speech Recognition-1998*, 47–56. Kerkrade, the Netherlands.

Greenhouse, S.W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112.

Grosjean, F., & Deschamps, A. (1975). Analyse contrastive des variables temporelles de l'anglais et du français: vitesse de parole et variables composantes, phénomène d'hésitation. *Phonetica*, *31*, 144–184.

Hallé, P., & Adda-Decker, M. (2007). Voicing assimilation in journalistic speech. *Proceedings of the 16th ICPhS-2007*, 2325–2328. Saarbrücken, Germany.

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.

Hanique, I., Schuppler B., & Ernestus M. (2010). Morphological and predictability effects on schwa reduction: The case of Dutch word-initial syllables. *Proceedings of Interspeech-2010*, 933–936. Makuhari, Japan.

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

Hansen, A.B. (2000). Le E caduc interconsonnantique en tant que variable sociolinguistique - une étude en région parisienne. *LINX*, *42(1)*, 45–48.

Hawkins, S., & Smith, R. (2001), Polysp: a polysystemic, phonetically-rich approach to speech understanding. *Italian Journal of Linguistics – Rivista di Linguistica*, *13(1)*, 99–188.

Holcomb P.J., & Grainger, J. (2006). On the time-course of visual word recognition: An ERP investigation using masked repetition priming. *Journal of Cognitive Neuroscience*, *18*, 1632–1643.

Huettig, F., & Altmann, G.T.M. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition*, *96*, 23–32.

Huettig, F., & McQueen, J.M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language*, *57(4)*, 460–482.

Janse, E., Nooteboom, S.G., & Quené, H. (2007). Coping with gradient forms of /t/-deletion and lexical ambiguity in spoken word recognition. *Language and Cognitive Processes*, *22*, 161–200.

Johnson, K. (2004). Massive reduction in conversational American English. *Proceedings of the 10th international symposium on spontaneous speech: Data and analysis-2010*, 29–54. Tokyo, Japan.

Jones, K., & Ono, T. (2000). Reconciling textbook dialogues and naturally occurring talk: What we think we do is not what we do. *Arizona Working Papers in SLAT*.

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W.D. (2000). Probabilistic relations between words: Evidence from reduction in lexical production. In: J. Bybee, & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 229–254). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Kager, R.W.J. (1997). Rhythmic vowel deletion in optimality theory. In I. Roca (Ed.), *Derivation and constraints in phonology* (pp. 463–499). Oxford: Oxford University Press.

Keating, P., Cho, T., Fougeron, C., & Hsu, C.S. (2003). Domain-initial strengthening in four languages. In J. Local, R. Ogden, and R. Temple (Eds.), *Phonetic interpretation: Papers in Laboratory Phonology VI*, 143–161. Cambridge: Cambridge University Press.

Keller, F., Gunasekharan, S., Mayo, N., & Corley, M. (2009). Timing Accuracy of Web Experiments: A Case Study Using the WebExp Software Package. *Behavior Research Methods*, *41(1)*, 1–12.

Kennedy, S., & Blanchet, J. (2014). Language awareness and perception of connected speech in a second language. *Language awareness*, *23(1–2)*, 92–106.

Klatt, D. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, *59*, 1208–1221.

Kohler, K. J. (1990). Segmental reduction in connected speech in German: Phonological facts and phonetic explanations. In W.J. Hardcastle, & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 21–33). Dordrecht: Kluwer Academic Publishers.

Kotz, S.A., & Elston-Güttler, K.E. (2004). The role of proficiency on processing categorical and associative information in the L2 as revealed by reaction times and event-related brain potentials. *Journal of Neurolinguistics*, *17(2–3)*, 215–235.

Kutas, M., & Hillyard, S.A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207(4427)*, 203–205.

Laks, B. (1977). Contribution empirique à l'analyse socio-différentielle de la chute de /r/ dans les groupes consonantiques finals. *Langue française*, *34*, 109–125.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*, 325–343.

Lennes, M., Alaroty, N., & Vainio, M. (2001). Is the phonetic quality of unaccented words unpredictable? An example from spontaneous Finnish. *Journal of the International Phonetic Association*, *31*, 127–138.

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

Lindblom, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, *35*, 1773–1781.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H & H theory. In W.J. Hardcastle, & A. Marchal (Eds.), *Speech Production and speech modelling* (pp. 403–439). Dordrecht: Kluwer Academic Publishers.

Lodge, R.A., Armstrong, N., Ellis, Y.M., & Shelton, J.F. (1997). *Exploring the French language*. New York: Oxford University Press.

Luce, P.A., Goldinger S.D., Auer, E.T., Vitevitch, M.S. (2000). Phonetic priming, neighborhood activation, and PARSYN. *Perception & Psychophysics*, *62*, 615–625.

Luce, P.A., & Pisoni, D.B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*, 1–36.

Magnuson, J.S., Tanenhaus, M.K., Aslin, R.N., & Dahan, D. (2003). The microstructure of spoken word recognition: Studies with artificial lexicons. *Journal of Experimental Psychology*, *132(2)*, 202–227.

Malécot, A., & Metz, G. (1972). Progressive nasal assimilation in French. *Phonetica*, *26*, 193–209.

Manuel, S.Y. (1991). Recovery of “deleted” schwa. In O. Engstrand, & C. Kylander (Eds.), *Perilus XIV: Papers from the symposium on current phonetic research paradigms for speech motor control* (pp. 115–118). Stockholm: University of Stockholm.

Matin, E., Shao, K., & Boff, K. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics*, *53*, 372–380.

Mattys, S.L., White, L., & Melhorn, J.F. (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology*, *134 (4)*, 477–500.

McCarthy, M., & Carter, R. (1995). Spoken grammar: What is it and how can we teach it? *ELT Journal 49(3)*, 207–218.

McCarthy, J., & Prince, A. (1993). Prosodic Morphology I: constraint interaction and satisfaction. *Linguistics Department Faculty Publication Series*, Paper 14.

McClelland, J.L., & Elman, J.L. (1986). The TRACE Model of Speech Perception. *Cognitive Psychology*, *18*, 1–86.

Meunier, C., & Espesser, R. (2011). Vowel reduction in conversational speech in French: The role of lexical factors. *Journal of Phonetics*, *39*, 271–278.

Meyer, A.S., Huettig, F., & Levelt, W.J.M. (2016). Same, different, or closely related: What is the relationship between language production and perception. *Journal of Memory and Language*, *89*, 1–7.

Milne, P. (2014). *The variable pronunciations of word-final consonant clusters in a force aligned corpus of spoken French* (Doctoral dissertation, University of Ottawa).

Mirman, D. (2014) *Growth Curve Analysis and Visualization Using R*. Boca Raton: CRC Press.

Mitterer, H., & Tuinman A, (2012). The role of native-language knowledge in the perception of casual speech in a second language. *Frontiers in Psychology*, *3*, article 249.

Mulder, K., Drijvers, L., Ernestus, M. (in preparation).

Mulder, K., Ten Bosch, L., & Boves, L. (2016). Comparing different methods for analyzing ERP signals. *Proceedings of Interspeech-2016,* 1373–1377. San Francisco, California.

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE 3.80. *L'Année Psychologique*, *101*, 447–462. Retrieved from http://www.lexique.org

Nieuwland, M.S., & Van Berkum, J.J. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of cognitive neuroscience*, *18(7)*, 1098–1111.

Norris, D., & McQueen, J.M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, *115(2)*, 357–395.

Nouveau, D. (2012). Limites perceptives de l'e caduc chez des apprenants néerlandophones. *Revue Canadienne de Linguistique Appliquée*, *15*, 60–78.

O'Connor Di Vito, N. (1991). Incorporating native speaker norms in second language materials. *Applied Linguistics*, *12(4)*, 383–396.

Patterson, D., LoCasto, P.C., & Connine, C.M. (2003). Corpora analyses of frequency of schwa deletion in conversational American English. *Phonetica*, *60*, 45–69.

Pierrehumbert, J., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In G. Docherty, & D. R. Ladd (Eds.), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody* (pp. 90–117). Cambridge: Cambridge University Press.

Pisoni, D.B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception and Psychophysics*, *13*, 253–260.

Pitt, M., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E., & Fosler-Lussier, E. (2007). *Buckeye Corpus of Conversational Speech* (2nd release) [www.buckeyecorpus.osu.edu] Columbus, OH: Department of Psychology, Ohio State University (Distributor).

Pitt, M., Dilley, L., & Tat, M. (2011). Exploring the role of exposure frequency in recognizing pronunciation variants. *Journal of Phonetics*, *39*, 304–311.

Pluymaekers, M., Ernestus, M., & Baayen, H. (2005a). Lexical frequency and acoustic reduction in spoken Dutch. *Journal of the Acoustical Society of America*, *118*, 2561–2569.

Pluymaekers, M., Ernestus, M., & Baayen, H. (2005b). Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica*, *62*, 146–159.

R Development Core Team (2015). *R: A language and environment for statistical computing* [http://www.R-project.org/]. Vienna: R Foundation for Statistical Computing.

Racine, I. (2008). *Les effets de l'effacement du schwa sur la perception et la production de la parole en français*. (Doctoral dissertation, Université de Genève, Genève, no. L. 652).

Racine, I., Bürki, A., & Spinelli, E. (2014). The implication of spelling and frequency in the recognition of phonological variants: evidence from pre-readers and readers*. Language, Cognition and Neuroscience*, *29(7)*, 893–898.

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

Racine, I., & Grosjean, F. (2002). La production du E caduc facultatif: Est-elle prévisible? *Journal of French Language Studies*, *12*, 307–326.

Raymond, W., Dautricourt, R., & Hume, E. (2006). Word-internal /t,d/ deletion in spontaneous speech: modeling the effects of extra-linguistic, lexical, and phonological factors. *Language Variation and Change*, *18*, 55–97.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, *64(3)*, 249–284.

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: the role of surface detail in basic-level object recognition. *Perception*, *33*, 217–236.

Rueschemeyer, S.-A., Nojack, A., & Limbach, M. (2008). A mouse with a roof? Effects of Phonological Neighbors on Processing of Words in Sentences in a Non-Native Language. *Brain and Language*, *104*, 132–144.

Salverda, A.P., Kleinschmidt, D., & Tanenhaus, M.K. (2014). Immediate effects of anticipatory coarticulation in spoken-word recognition. *Journal of Memory and Language*, *71(1)*, 145–163.

Scarborough, D.L., Gerard, L., & Cortese, C. (1984). Independence of lexical access in bilingual word recognition. *Journal of Verbal Learning & Verbal Behavior*, *23(1)*, 84–99.

Schneider, W., Eschman, A., & Zuccolotto, A., (2007). *E-Prime: User's Guide, Version 2.0*. Pittsburgh, PA: Psychology Software Tools.

Schuppler, B., Van Dommelen, W.A., Koreman, J., & Ernestus, M. (2012). How linguistic and probabilistic properties of a word affect the realization of its final /t/: Studies at the phonemic and sub-phonemic level. *Journal of Phonetics*, *40*, 595–607.

Sereno, J.A., & Jongman, A. (1997). Processing of English inflectional morphology. *Memory & Cognition*, *25*, 425–437.

Sims, M., Bolger, P., & Tucker, B.V., (2010). Processing of reduction in spontaneous speech: An event related potential study. *The Mental Lexicon Conference*, poster presentation. Montreal, Quebec.

Snodgrass, J.G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology Human Learning and Memory*, *6*, 174–215.

Soares, C., & Grosjean, F. (1984). Bilinguals in a monolingual and a bilingual speech mode: The effect on lexical access. *Memory and Cognition*, *12(4)*, 380–386.

Stemberger, J.P., & MacWhinney, B. (1988). Are inflected forms stored in the lexicon? In: M. Hammond, & M. Noonan (Eds.). *Theoretical morphology: Approaches in modern linguistics* (pp. 101-116). San Diego, CA: Academic Press.

Stevens, K.N., Liberman, A.M., Studdert-Kennedy, M., & Öhman, S. (1969). Crosslanguage study of vowel perception. *Language and Speech*, *12*, 1–23.

Stridfeldt, M. (2005). *La perception du français oral par des apprenants suédois* (Doctoral dissertation, Umea Universitet, Sweden).

Su, T.-T. & Basset, P. (1998). Language dependent and independent spontaneous speech phenomena. *SPoSS*, 55–58.

Ten Bosch, L., Giezenaar, G., Boves, L., & Ernestus, M. (2016). Modeling language-learners' errors in understanding casual speech. In G. Adda, V. Barbu Mititelu, J. Mariani, D. Tufiş, & I. Vasilescu (Eds.), Errors by humans and machines in multimedia, multimodal, multilangual data processing. *Proceedings of Errare-2016* (pp. 107–121). Bucharest, Romania.

Torreira, F., Adda-Decker, M., & Ernestus, M. (2010). The Nijmegen corpus of casual French. *Speech Communication*, *52*, 201–212.

Torreira, F., & Ernestus, M. (2009). Probabilistic effects on French [t] duration. *Proceedings of Interspeech-2009*, 448–451. Brighton, United Kingdom.

Torreira, F., & Ernestus, M. (2011). Vowel elision in casual French: the case of vowel /e/ in the word "c'était". *Journal of Phonetics*, *39*, 50–58.

Tucker, B.V., & Warner, N. (2007). Inhibition of processing due to reduction of the American English flap. *Proceedings of the 16th International Congress of Phonetic Sciences*-2007, 1949–1952. Saarbrücken, Germany.

Tuinman, A., Mitterer, H., & Cutler, A. (2014). Use of syntax in perceptual compensation for phonological reduction. *Language and Speech*, *57*, 68–85.

Van Bael, C., Baayen, H., & Strik, H. (2007). Segment deletion in spontaneous speech: A corpus study using mixed effects models with crossed random effects. *Proceedings of Interspeech-2007*, 2741–2744. Antwerp, Belgium.

Van de Ven, M., Ernestus, M., & Schreuder, R. (2012). Predicting acoustically reduced words in spontaneous speech: The role of semantic/syntactic and acoustic cues in context. *Laboratory Phonology*, *3*, 455–481.

Van de Ven, M., Tucker, B.V., & Ernestus, M. (2010). Semantic facilitation in bilingual everyday speech comprehension. *Proceedings of Interspeech-2010*, 1245–1248. Makuhari, Japan.

Van de Ven, M., Tucker, B.V., & Ernestus, M. (2011). Semantic context effects in the comprehension of reduced pronunciation variants. *Memory & Cognition*, *39*, 1301–1316.

Van Son, R.J.J.H., Bolotova, O., Lennes, M., & Pols, L.C.W (2004). Frequency effects on vowel reduction in three typologically different languages (Dutch, Finnish, Russian). *Proceedings of Interspeech-2004*, vol. II, 1277–1280. Jeju Island, Korea.

Viebahn, M., Ernestus, M., & McQueen, J.M. (2015). Syntactic predictability in the recognition of carefully and casually produced speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41(6)*, 1684–1702.

Villeneuve, A.-J. (2010). Word-final cluster simplification in Vimeu French: a preliminary analysis. *Linguistics*, *15(2)*, article 15.

Vitevitch, M.S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28(4)*, 735–747.

Walker, D.C. (2001). *French sound structure*. Canada: University of Calgary Press.

Waugh, L.R., & Fonseca-Greber, B. (2002). Authentic Materials for Everyday Spoken French: Corpus Linguistics vs. French Textbooks. *Arizona Working Papers in SLAT*, *9*, 114–127.

Wong, S.W.L., Mok, P.P.K., Chung, K. K.-H., Leung, V.W.H., Bishop, D.V.M., & Chow, B. W.-Y. (2015). Perception of native English reduced forms in Chinese learners: Its role in listening comprehension and its phonological correlates. *TESOL QUARTERLY*.

Wright, R. (2004). A review of perceptual cues and cue robustness. In: B. Hayes, R. Kirchner, & D. Steriade (Eds.). *Phonetically-based phonology* (pp. 34–57). Cambridge: Cambridge University Press.

Zellner, B. (1994). Pauses and the temporal structure of speech. In: E. Keller (Ed.), *Fundamentals of speech synthesis and speech recognition* (pp. 41–62). Chichester: John Wiley.

Zimmerer, F., & Reetz, H. (2014). Do listeners recover “deleted” final /t/ in German? *Frontiers in Psychology*, *5*, article 735.

Zimmerer, F., Scharinger, M., & Reetz, H. (2011). When BEAT becomes HOUSE: Factors of word final /t/-deletion in German. *Speech Communication*, *53(6)*, 941–954.

Acknowledgements

This dissertation would not have been possible without the help of many people. First and foremost, I would like to thank my promotor, Mirjam Ernestus, for the continuous support, guidance and interesting discussions during my PhD journey. Mirjam, your enthusiasm for research is inspiring and encouraged me to stay curious and open minded. I also had the opportunity to attend several national and international conferences, experiences that have broadened my horizon, both scientifically and personally. Without your support this would not have been possible. I am also thankful to my co-promotor, Esther Janse for her timely support and encouragement. Thank you also for your insightful comments and feedback on my projects.

I express my gratitude to my colleagues of the CLS-M research group and the Max Planck Institute who were always in for a short chat or extensive discussion, and gave helpful feedback: Rian, Marco, Malte, Xaver, Huib, Annika, Martijn, Robert, Kimberley, Mark, Laura, Lou, Louis, Ellen, Mybeth, Thordis, Kim, Linda, Lisa, Sascha, Joe, Chén, Katherine, Emily, Aurora, Mirjam Broersma, Rob Schoonen. Kimberley, we spent many hours discussing and designing the eye-tracking and ERP study. The project was logistically challenging, but it was fun to conduct this experiment with you. You also inspired me with your ability to distil complex concepts into simple terms. I want to extent my appreciation to Roger Staats for English proofreading my manuscripts, to Els den Os, Dirkje van der Aa, and Kevin Lam (IMPRS) for organizing interesting and useful courses and meetings, to Karin Karstens and Meggie Uijen (Library of the Max Planck Institute) for assisting me with the literature search, and to the technical team of the Max Planck Institute for keeping all systems up and running. Without the help and effort of Reiner Dirksmeyer, Ronald Fischer, and Johan Weustink, the combined eye-tracking and ERP experiment would not have been possible.

Data collection is a crucial step in the research process and I am grateful to all people who helped me at this stage of my PhD. First of all, I would like to thank Cécile Fougeron and Martine Adda-Decker. I felt very welcome at your lab at the *Institut de Linguistique et Phonétique Générales et Appliquées* (ILPGA) in Paris. You and your co-workers created the ideal environment for my experiments. I am also grateful to those who assisted with the data collection at more nearby sites in the Netherlands, including Audrey Rousse-Malpat and Jeanette den Toonder (University of Groningen), Jan Hulstijn and Margarita Steinel (University of Amsterdam), Esther de Vrind (ICLON), Jenny Doetjes, Willemijn Heeren, Daan van de Velde, and Jos Pacilly (Leiden University), Iris Mulders (Utrecht University), Mattie Lagarrigue, Evelyne Vos-Fruit, Marc Smeets, Haike Jacobs, Dominique Nouveau, and Mariska Koopman-Thurlings (Radboud University, Nijmegen). A special word of appreciation goes to Janine Berns (Radboud University) who helped me with the recruitment of participants. And more importantly, as a friend you were always around for discussing frustrations and successes in science, and chatting about anything else not related to research. I also thank all student assistants who facilitated the conduct of my experiments, and Thera Baayen in particular, who provided excellent support during the eye-tracking and ERP experiment in Paris.

I thank Suzanne and Heleen for agreeing to be my paranymphs. I am very lucky to have you among my friends. Our unforgettable day outs with visits to museums, theatres, cultural festivals, cinemas and our delicious dinners made a welcome change in my life as a PhD student. Suzanne, my “buurman” from Haarlem, and Heleen, my fellow student from Amsterdam, you both have a great sense of humour and this is exactly what kept me going all these years.

My sisters, Charlotte, Judith, and Victorine, many thanks for the much-appreciated distractions from work. Since we almost read one another's minds, you knew exactly when I needed a break. Thanks also for putting my worries aside and all the encouraging words. I am extremely grateful to my parents, Els and Guus, who have taught me the value of hard work and have offered me wonderful opportunities to explore my curiosity. This work would not have been possible without your unconditional love and unending support. And last but not least, Pieter, I am very grateful for all your patience and support, and for always bringing fun and laughter during this journey. I am still wearing that smile you gave me in 2012 when I first met you and started my PhD.

Curriculum Vitae

Sophie Brand was born in Roermond, the Netherlands, on 25 August 1985. She studied French Linguistics (MA) at Radboud University Nijmegen and French Literature (MA) at the Free University in Amsterdam. She then moved to Leiden University to take a master's degree in teacher-training. From 2009 till 2012, Sophie worked as a high-school teacher in French in Haarlem. She combined teaching with a research project at the Meertens Institute in Amsterdam where she investigated how Turkish-Dutch bilingual children acquire the Dutch article *het* and the Dutch pronoun *het*. In 2012, Sophie started her PhD project at the Centre for Language Studies at Radboud University. Her research was part of the project *Foreign Casual Speech*, which was funded by a Consolidator grant [284108] from the European Research Council awarded to Mirjam Ernestus.

MPI series in psycholinguistics

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing.
Miranda I. van Turennout

2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography.
Niels O. Schiller

3. Lexical access in the production of ellipsis and pronouns.
Bernadette M. Schmitt

4. The open-/closed class distinction in spoken-word recognition.
Alette Petra Haveman

5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach.
Kay Behnke

6. Gesture and speech production.
Jan-Peter de Ruiter

7. Comparative intonational phonology: English and German.
Esther Grabe

8. Finiteness in adult and child German.
Ingeborg Lasser

9. Language input for word discovery.
Joost van de Weijer

10. Inherent complement verbs revisited: towards an understanding of argument structure in Ewe.
James Essegbey

11. Producing past and plural inflections.
Dirk J.Janssen

12. Valence and transitivity in Saliba: an Oceanic language of Papua New Guinea.
Anna Margetts

13. From speech to words.
Arie H. van der Lugt

14. Simple and complex verbs in Jaminjung: a study of event categorisation in an Australian language.
Eva Schultze-Berndt

15. Interpreting indefinites: an experimental study of children's language comprehension.
Irene Krämer

16. Language-specific listening: the case of phonetic sequences.
Andrea Christine Weber

17. Moving eyes and naming objects.
Femke Frederike van der Meulen

18. Analogy in morphology: the selection of linking elements in Dutch compounds.
Andrea Krott

19. Morphology in speech comprehension.
Kerstin Mauth

20. Morphological families in the mental lexicon.
Nivja Helena de Jong

21. Fixed expressions and the production of idioms.
Simone Annegret Sprenger

22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria.
Birgit Hellwig

23. Paradigmatic structures in morphological processing: Computational and crosslinguistic experimental studies.
Fermín Moscoso del Prado Martín

24. Contextual influences on spoken-word processing: an electrophysiological approach.
Danielle van den Brink

25. Perceptual relevance of prevoicing in Dutch.
Petra Martine van Alphen

26. Syllables in speech production: effects of syllable preparation and syllable frequency.
Joana Cholin

27. Producing complex spoken numerals for time and space.
Marjolein Henriëtte Wilhelmina Meeuwissen

28. Morphology in auditory lexical processing: sensitivity to fine phonetic detail and insensitivity to suffix reduction.
Rachèl Jenny Judith Karin Kemps

29. At the same time…: the expression of simultaneity in learner varieties.
Barbara Schmiedtová

30. A grammar of Jalonke argument structure.
Friederike Lüpke

31. Agrammatic comprehension: an electrophysiological approach.
Marijtje Elizabeth Debora Wassenaar

32. The structure and use of shape-based noun classes in Miraña (North West Amazon).
Frank Seifart

33. Prosodically-conditioned detail in the recognition of spoken words.
Anne Pier Salverda

34. Phonetic and lexical processing in a second language.
Mirjam Elisabeth Broersma

35. Retrieving semantic and syntactic word properties: ERP studies on the time course in language comprehension.
Oliver Müller

36. Lexically-guided perceptual learning in speech processing.
Frank Eisner

37. Sensitivity to detailed acoustic information in word recognition.
Keren Batya Shatzman

38. The relationship between spoken word production and comprehension.
Rebecca Özdemir

39. Disfluency: interrupting speech and gesture.
Mandana Seyfeddinipur

40. The acquisition of phonological structure: distinguishing contrastive from non-contrastive variation.
Christiane Dietrich

41. Cognitive cladistics and the relativity of spatial cognition.
Daniel Haun

42. The acquisition of auditory categories.
Martijn Bastiaan Goudbeek

43. Affix reduction in spoken Dutch: probabilistic effects in production and perception.
Mark Plumaekers

44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence.
Valesca Madalla Kooijman

45. Space and iconicity in German sign language (DGS).
Pamela M. Perniss

46. On the production of morphologically complex words with special attention to effects of frequency.
Heidrun Bien

47. Crosslinguistic influence in first and second languages: convergence in speech and gesture.
Amanda Brown

48. The acquisition of verb compounding in Mandarin Chinese.
Jidong Chen

49. Phoneme inventories and patterns of speech sound perception.
Anita Eva Wagner

50. Lexical processing of morphologically complex words: an information-theoretical perspective.
Victor Kuperman

51. A grammar of Savosavo: a Papuan language of the Solomon Islands.
Claudia Ursula Wegener

52. Prosodic structure in speech production and perception.
Claudia Kuzla

53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: investigating knowledge of forms and functions in production and comprehension.
Sarah Schimke

54. Studies on intonation and information structure in child and adult German.
Laura de Ruiter

55. Processing the fine temporal structure of spoken words.
Eva Reinisch

56. Semantics and (ir)regular inflection in morphological processing.
Wieke Tabak

57. Processing strongly reduced forms in casual speech.
Susanne Brouwer

58. Ambiguous pronoun resolution in L1 and L2 German and Dutch.
Miriam Ellert

59. Lexical interactions in non-native speech comprehension: evidence from electroencephalography, eye-tracking, and functional magnetic resonance imaging.
Ian FitzPatrick

60. Processing casual speech in native and non-native language.
Annelie Tuinman

61. Split intransitivity in Rotokas, a Papuan language of Bougainville.
Stuart Payton Robinson

62. Evidentiality and intersubjectivity in Yurakaré: an interactional account.
Sonja Gipper

63. The influence of information structure on language comprehension: a neurocognitive perspective.
Lin Wang

64. The meaning and use of ideophones in Siwu.
Mark Dingemanse

65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants.
Marco van de Ven

66. Speech reduction in spontaneous French and Spanish.
Francisco Torreira

67. The relevance of early word recognition: insights from the infant brain.
Caroline Mary Magteld Junge

68. Adjusting to different speakers: extrinsic normalization in vowel perception.
Matthias Johannes Sjerps

69. Structuring language: contributions to the neurocognition of syntax.
Katrien Rachel Segaert

70. Infants' appreciation of others' mental states in prelinguistic communication: a second person approach to mindreading.
Birgit Knudsen

71. Gaze behavior in face-to-face interaction.
Federico Rossano

72. Sign-spatiality in Kata Kolok: how a village sign language of Bali inscribes its signing space.
Connie de Vos

73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning.
Attila Andics

74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation.
Marijt Witteman

75. The use of deictic versus representational gestures in infancy.
Daniel Puccini

76. Territories of knowledge in Japanese conversation.
Kaoru Hayano

77. Family and neighbourhood relations in the mental lexicon: A cross-language perspective.
Kimberley Mulder

78. Contributions of executive control to individual differences in word production.
Zeshu Shao

79. Hearing speech and seeing speech: Perceptual adjustments in auditory-visual processing.
Patrick van der Zande

80. High pitches and thick voices: The role of language in space-pitch associations.
Sarah Dolscheid

81. Seeing what's next: Processing and anticipating language referring to objects.
Joost Rommers

82. Mental representation and processing of reduced words in casual speech.
Iris Hanique

83. The many ways listeners adapt to reductions in casual speech.
Katja Poellmann

84. Contrasting opposite polarity in Germanic and Romance languages: Verum Focus and affirmative particles in native speakers and advanced L2 learners.
Giuseppina Turco

85. Morphological processing in younger and older people: Evidence for flexible dual-route access.
Jana Reifegerste

86. Semantic and syntactic constraints on the production of subject-verb agreement.
Alma Veenstra

87. The acquisition of morphophonological alternations across languages.
Helen Buckler

88. The evolutionary dynamics of motion event encoding.
Annemarie Verkerk

89. Rediscovering a forgotten language.
Jiyoun Choi

90. The road to native listening: Language-general perception, language-specific input.
Sho Tsuji

91. Infants' understanding of communication as participants and observers.
Gudmundur Bjarki Thorgrímsson

92. Information structure in Avatime.
Saskia van Putten

93. Switch reference in Whitesands.
Jeremy Hammond

94. Machine learning for gesture recognition from videos.
Binyam Gebrekidan Gebre

95. Acquisition of spatial language by signing and speaking children: a comparison of Turkish sign language (TID) and Turkish.
Beyza Sumer

96. An ear for pitch: on the effects of experience and aptitude in processing pitch in language and music.
Salomi Savvatia Asaridou

97. Incrementality and Flexibility in Sentence Production.
Maartje van de Velde

98. Social learning dynamics in chimpanzees: Reflections on (nonhuman) animal culture.
Edwin van Leeuwen

99. The request system in Italian interaction.
Giovanni Rossi

100. Timing turns in conversation: A temporal preparation account.
Lilla Magyari

101. Assessing birth language memory in young adoptees.
Wencui Zhou

102. A social and neurobiological approach to pointing in speech and gesture.
David Peeters

103. Investigating the genetic basis of reading and language skills.
Alessandro Gialluisi

104. Conversation electrified: The electrophysiology of spoken speech act recognition.
Rósa Signý Gisladottir

105. Modelling multimodal language processing.
Alastair Charles Smith

106. Predicting language in different contexts: The nature and limits of mechanisms in anticipatory language processing.
Florian Hintz

107. Situational variation in non-native communication
Huib Kouwenhoven

108. Sustained attention in language production.
Suzanne R. Jongman

109. Acoustic reduction in spoken-word processing: Distributional, syntactic, morphosyntactic, and orthographic effects.
Malte Viebahn

110. Nativeness, dominance, and the flexibility of listening to spoken language
Laurence Bruggeman

111. Semantic specificity of perception verbs in Maniq.
Ewelina Wnuk

112. On the identification of FOXP2 gene enhancers and their role in brain development
Martin Becker

113. Events in language and thought: The case of serial verb constructions in Avatime.
Rebecca Defina

114. Deciphering common and rare genetic effects on reading ability.
Amaia Carrión Castillo

115. Music and language comprehension in the brain.
Richard Kunert

116. Comprehending comprehension: Insights from neuronal oscillations on the neuronal basis of language.
Nietzsche H. L. Lam

117. The biology of variation in anatomical brain asymmetries.
Tulio Guadalupe

118. Language processing in a conversation context.
Lotte Schoot

119. Achieving mutual understanding in Argentine Sign Language
Elizabeth Manrique

120. Talking Sense: the behavioural and neural correlates of sound symbolism
Gwilym Lockwood

121. Getting under your skin: The role of perspective and simulation of experience in narrative comprehension
Franziska Hartung

122. Sensorimotor Experience in Speech Perception
Will Schuerman

123. Explorations of beta-band neural oscillations during language comprehension: Sentence processing and beyond
Ashley Lewis

124. Influences on the magnitude of syntactic priming
Evelien Heyselaar

125. Lapse organization in interaction
Elliott Hoey

List of publications

Brand, S., & Ernestus, M. (2015). Reduction of obstruent-liquid-schwa clusters in casual French. In Scottish consortium for ICPhS 2015, M. Wolters, J. Livingstone, B. Beattie, R. Smith, M. MacMahon, J. Stuart-Smith, & J. Scobbie (Eds.), *Proceedings of the 18th International Congress of Phonetic Sciences* [ICPhS 2015]. Glasgow: University of Glasgow.

Brand, S., & Ernestus, M. (2017). Listeners' processing of a given reduced word pronunciation variant directly reflects their exposure to his variant: evidence from native listeners and learners of French. *Quarterly Journal of Experimental Psychology*. doi: 10.1080/17470218.2017.1313282

ISBN 978-90-76203-87-4