Acoustic reduction in spoken-word processing

Distributional, syntactic, morphosyntactic, and orthographic effects

ISBN/EAN: 978-90-76203-75-1
Cover design: Waveform of the Dutch word *geluisterd* "listened". This word was used as a stimulus in Chapter 3. The red colour highlights the vowel schwa and the blue colour the consonant /t/. Both sounds are frequently reduced in Dutch and were the focus of the work presented in this thesis.
Printed and bound by Ipskamp Drukkers, Nijmegen

Acoustic reduction in spoken-word processing

Distributional, syntactic, morphosyntactic, and orthographic effects

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus,
volgens besluit van het college van decanen
in het openbaar te verdedigen op maandag 25 april 2016
om 10.30 uur precies

door

Malte Clemens Viebahn

geboren op 1 december 1981
te Bremen, Duitsland

Promotoren:	Prof. dr. Mirjam Ernestus
	Prof. dr. James M. McQueen
Manuscriptcommissie:	Prof. dr. Antje Meyer (Voorzitter)
	Prof. dr. Herbert Schriefers
	Dr. Esther Janse

The research reported in this thesis was supported by a fellowship of the International Max Planck Research School for Language Sciences. The work presented in Chapter 5 was supported by a *Kurzstipendium* grant from the German Academic Exchange Service (DAAD).

Für meine Eltern.
Im Gedenken an meine Mutter.

Contents

Introduction

Chapter 1

Variability is a central property of spoken language. The same word uttered by the same speaker on multiple occasions will differ, for example, in speaking rate, the speaker's voice, and details of the acoustic properties of the vowels and consonants. When this word is uttered by two different speakers, variability increases even more due to differences in the shape and size of the speakers' vocal tract, age, sex, language background, and regional dialect. Whereas these sources of variability still pose a great challenge for Automatic Speech Recognition (ASR) systems, human listeners are able to cope with variability in a way that seems remarkably easy and automatic. How do we get from a highly variable and noisy sensory input to a stable and clear perceptual experience? This thesis sets out to find answers to this question by investigating the processing of one widespread source of acoustic-phonetic variability: speech reduction.

1.1 What is speech reduction?

In natural conversations, speech is typically produced with a casual speaking style. This speaking style (or *register*) is characterized by the omission or acoustic weakening of individual phonemes and even whole syllables (Ernestus and Warner, 2011). The degree to which speech is reduced varies on a continuum. A relatively mild kind of reduction, in which one specific speech sound is omitted, is word-final /t/ reduction (e.g., Ernestus and Warner, 2011). For example, the word *mist* might be produced without the word-final /t/ as in [mɪs]. But reductions can also be quite extensive. An example of an extremely reduced pronunciation is the word *yesterday* produced as [jɛʃeɪ] instead of [jɛstɚdeɪ] (its canonical form). Speech reductions also vary in whether they are categorical or gradient. For example, word-medial schwa in French can be completely absent (e.g., *renard* "fox" produced as [ʁnaʁ] instead of [ʁənaʁ]) or it can be shortened in its duration (Bürki et al., 2011). Such gradient reduction can also affect consonants. This is for example the case for nasal flapping in American English where words like *center* are produced as [sɛ̃ɾɚ] instead of [sɛntɚ] (e.g., Ranbom and Connine, 2007). Another mild form of reduction is nasal place assimilation where words like *lean* are produced as [lim] when followed by a bilabial stop as in *lean bacon* (e.g., Gaskell and Marslen-Wilson, 1996). As a result of such reduction processes, words are often produced in ways that differ from the *canonical* pronunciations that are found in dictionaries. Pronunciation variants that differ from canonical pronunciations as a result of reduction processes are typically referred to as *reduced pronunciation variants* or, for short, *reduced variants*.

Similar types of reduction have been documented across several different languages including Dutch (e.g., Ernestus, 2000; Schuppler et al., 2011), English (e.g., Dalby, 1986; Johnson, 2004), German (e.g., Kohler, 1990), Spanish (e.g., Torreira and Ernestus, 2012), and French (e.g., Adda-Decker and Snoeren, 2011) but also typologically more different languages such as Mandarin (e.g., Cheng and Xu, 2009) and Japanese (e.g., Nakamura et al., 2008). Corpus studies suggest that acoustic-phonetic reduction is a very frequent phenomenon. For example, Johnson (2004) estimates that over 60% of the words uttered in casual English speech deviate from their citation forms by at least one phoneme and 28% deviate by two or more phonemes.

One phoneme that is particularly often reduced is the vowel schwa /ə/. For example the English word *sufficient* may be pronounced as [sfɪʃn̩t] instead of [səfɪʃənt]. Similarly, the Dutch word *geweest* "been" may be pronounced as [xwest] instead of [xəwest], the German word *haben* "to have" may be pronounced as [habm̩] instead of [habən], and the French word *seringue* "syringe" may be pronounced as [sʀɛ̃g] instead of [səʀɛ̃g]. The present thesis will focus primarily on the reduction of schwa. More specifically, this thesis investigates when schwa reduction occurs and how it is processed.

1.2 When does speech reduction occur?

One of the most reliable predictors of speech reduction is speech register. Speech reductions are much more frequent in spontaneous and conversational speech produced in informal contexts than in formal or read speech (e.g., Ernestus et al., 2015; Hanique et al., 2013; Mitterer and Ernestus, 2006). Another powerful predictor of when speech reduction occurs is speaking rate. The probability that a speaker produces reduced pronunciation variants is positively correlated with speaking rate: faster speech is very likely to result in more speech reductions (e.g., Hanique et al., 2013; Raymond et al., 2006). Speaking rate is typically higher in conversational speech than in formal or read speech. This brings up the question of whether the effects of speaking rate and register are additive or interactive. For example, it might be that whether or not register influences speech reduction depends on the speaking rate of the talker. Corpus studies have shown, however, that register influences reduction even if speaking rate is statistically controlled for (Hanique et al., 2013, also see Chapter 2 of this thesis). This suggests that the effect of register on reduction is at least in part independent from speaking rate.

Furthermore, the likelihood with which a given word will be reduced is influenced by its prosodic position within a sentence. Words that carry sentential stress are less likely to be reduced than unaccented words (Aylett and Turk, 2006). Whether or not a word will be produced in a reduced way depends also on how predictable it is. This has been demonstrated in terms of a word's lexical frequency. High-frequency words are more likely to be reduced than low-frequency words (Bell et al., 2009; Pluymaekers et al., 2005b). A similar result has been found for syllables. Segments tend to be more reduced in high- than low-frequency syllables (Aylett and Turk, 2006). Another form of predictability is the conditional probability of a word's occurrence based on the preceding and/or following word. Words that are more predictable based on the words that surround them are more likely to undergo reduction than words that are less strongly associated with the words around them (Bell et al., 2003, 2009; Pluymaekers et al., 2005a). The present thesis will further examine contextual effects on the processing of reduced speech. In particular, the studies presented in the following chapters will investigate the influence of co-occurrence, syntactic context, speaking style, and orthographic information.

1.3 How is reduced speech processed?

Despite the ubiquity of speech reductions in everyday speech, we rarely notice them unless we pay attention and listen very carefully. This suggests that our speech recognition system is well prepared to deal with the acoustic-phonetic variability that is due to reduced speech. This is not surprising given that listeners have to cope with a wide range of different forms of variability when processing speech. How do listeners accomplish this task?

The processing of reduced pronunciation variants has been shown to be influenced by a variety of variables. One important factor is the ability of listeners to perceive fine-grained phonetic detail which allows them to recover gradiently reduced segments based on acoustic remnants of the segments. For example, Dutch listeners have been shown to use fine-grained acoustic information in order to recover reduced /t/ (Mitterer and Ernestus, 2006). In this study, whether or not listeners indicated the presence of a reduced /t/ changed in a continuous fashion depending on the degree of reduction. That listeners are sensitive to fine-grained acoustic-phonetic information has also been demonstrated for coronal place assimilation in English (e.g., Gow, 2001, 2002, 2003; Gow and McMurray, 2007). Gow (2002) compared the frequencies of the second and third formants between unreduced coronal nasals (i.e., [n]), assimilated coronal nasals (e.g., a labialized [n] which is [m]-like), and the canonical segment that is approximated by the assimilated coronal nasal (e.g., canonical [m]). He found that the formant frequencies of the assimilated nasal were in between the unassimilated and the approximated nasal, showing that assimilation was not complete but that it resulted in an acoustic-phonetic amalgam of both segments. A series of recognition experiments suggests that listeners might exploit this acoustic information in order to recover the assimilated segment (Gow, 2001, 2002, 2003) and to predict the approximated segment (Gow and McMurray, 2007). For example, Gow showed that assimilation does not result in lexical ambiguity if the assimilated segment is presented in the appropriate phonological context (e.g., a labialized coronal segment is presented before a labial consonant). For example, the word *line* produced with a labialized /n/ does not prime the word *lime*, showing that listeners did not interpret the labialized /n/ as /m/ and instead recovered the unassimilated segment. These results show that at least in cases in which acoustic reduction is gradient, listeners' sensitivity to phonetic detail allows them to recover reduced segments and can thereby help to process reduced speech.

Furthermore, listeners use probabilistic information from the phonetic context about speech production when processing reduced segments. Often, certain types of reduction are more likely to occur in some phonetic contexts than others. For example, Dutch /t/s are particularly often reduced when preceded by /s/ and followed by bilabial consonants (Mitterer and Ernestus, 2006). Comprehension experiments have shown that listeners are sensitive to this probabilistic information and can apply it when recognizing reduced speech. Listeners are more likely to recover the reduced segment if it occurs in a context in which it is likely to be reduced than if it occurs in a context in which reduction is unlikely (e.g., Mitterer and Ernestus, 2006; Mitterer and McQueen, 2009).

Acoustic-phonetic context is not only beneficial to listeners when recovering individual segments that are mildly reduced (such as reduced /t/) but also when processing massive reductions that affect the whole word (such as *yesterday* produced as [jɛʃeɪ]). Word-recognition experiments have shown that the ability to recognize strongly reduced pronunciation variants increases when the words are presented in phonetic contexts rather than in isolation (e.g., Ernestus et al., 2002; Janse and Ernestus, 2011). In the study by Ernestus et al., participants listened to tokens of strongly reduced words (e.g., [mok], a reduced form of Dutch /moxələk/ "possible") that had been spliced out of connected speech. When these forms were presented in isolation, recognition accuracy was only 52%. In contrast, when the forms were presented together with neighbouring segments (e.g., [ɛlmokna]) accuracy increased to 70%. This suggests that acoustic-phonetic information from the surrounding context can help the listener to process reduced speech.

In addition to acoustic-phonetic factors, lexical knowledge has also been shown to influence the processing of reduced speech. Kemps et al. (2004) investigated the role of lexical knowledge on the perception of reduced speech sounds in the Dutch suffix /lək/. When presented with the reduced variant [k], participants were more likely to report that they heard an /l/ when the suffix was presented in a sentential context than when the suffix was presented in isolation. This suggests that listeners inferred the presence of the reduced segment /l/ based on their knowledge about the word in which the suffix occurred. A similar finding was reported for the perception of reduced /t/ (Janse et al., 2007; Mitterer and Ernestus, 2006). In these studies, Dutch listeners that were presented with words ending in acoustically weak or completely absent /t/s were more likely to indicate the presence of /t/ when the /t/ turned the word into a real word of Dutch (e.g. *orkes(t)* "orchestra") than if it did not (e.g., *moras* is the Dutch word for "swamp" but *morast* is not a Dutch word).

Besides lexical knowledge, listeners' prior experience with particular pronunciation variants influences processing. Highly frequent variants are recognized more easily and accurately than less frequent variants. For example, in American English, effects of variant frequency have been demonstrated for words with reduced /t/ (e.g., *atom* produced as [æɾəm] or *witness* produced as [wɪʔnəs], Pitt et al., 2011; Ranbom and Connine, 2007), and reduced schwa (e.g., *broccoli* produced as [brɑkli], Connine et al., 2008). Studies in French also suggest that the processing of words with reduced schwa (e.g. *chemin* "path" produced as [ʃmɛ̃] instead of [ʃəmɛ̃]) might be influenced by variant frequency in word recognition as well as production (Bürki et al., 2010; Bürki and Frauenfelder, 2012; Racine et al., 2013).

When presented in isolation, reduced word forms have been found to be processed more slowly and less accurately compared to their canonical counterparts (e.g., Pitt, 2009; Ranbom and Connine, 2007; van de Ven et al., 2011). This processing advantage for canonical pronunciations even persists when the reduced variant occurs more frequently than the canonical variant (Ranbom and Connine, 2007). Why is it that reduced pronunciation variants are difficult to process when presented in isolation but in our everyday experience we are hardly ever bothered by them? Part of the answer to this question has to do with the fact that outside the laboratory, we rarely listen to words in isolation. Instead, we typically hear words embedded in connected speech and in informationally rich contexts that provide multiple sources of information.

The properties of the acoustic-phonetic context can also support processing by providing information that allows listeners to adapt to casual speech. For example, a visual-world eye-tracking study by Brouwer et al. (2012b) suggests that listening to casual speech causes the word recognition system to change the way it operates. More specifically, it changes the dynamics of lexical competition during word recognition. In this study, participants listened to sentences extracted from a spontaneous speech corpus and saw four printed words: a target (e.g., *computer*), a competitor that was similar to the canonical form (e.g., *companion*), a competitor that was similar to the reduced form (e.g., *pupil* which is similar to "puter" which is a reduced form of *computer*), and an unrelated distractor. Consistent with previous visual-world studies, Brouwer et al. found lexical competition effects when carefully produced target words were presented (e.g., Allopenna et al., 1998; McQueen and Viebahn, 2007). However, when carefully and casually produced word forms were presented together in the same experiment, lexical competition was weaker and was influenced less by the phonological overlap between the target and the competitor. These results suggest that listeners adapt to casual speech by penalizing acoustic mismatches less strongly.

Further evidence for the adaptation to casual speech has been provided by Poellmann et al. (2014). In this study, Dutch listeners were exposed to segmental and syllabic reductions in spoken Dutch. In the subsequent test phase, participants heard both kinds of reductions but applied to words that they had not heard during exposure. Learning about reductions was applied to these previously unheard words, demonstrating that listeners can adapt to acoustic-phonetic reduction. Adaptation to a casual speaking style is not limited to perception but also extends to the domain of language production. A shadowing experiment by Brouwer et al. (2010) demonstrated that hearing reduced speech increases the probability of producing reduced word forms.

1.4 Aim of this thesis

The aim of this thesis is to investigate how language users process reduced speech on multiple psycholinguistic levels and how listener flexibility aids in this task. More specifically, the studies presented here explore the role of distributional, syntactic, morphosyntactic, and orthographic information in the processing of spoken words with acoustic-phonetically reduced schwa. By adopting a broad perspective on language processing this thesis aims to provide further evidence showing that acoustic reduction influences not only the phonetic level of processing but other components of the psycholinguistic architecture as well. Each chapter investigates the processing of acoustic reduction on a different linguistic level, showing that the way listeners respond to acoustic reductions filters through from acoustic-phonetic to higher levels of processing. This cross-level response to acoustic reduction can only be observed when adopting a wider view that goes beyond a single level of linguistic processing. Adopting such a view will increase our understanding of how speech reductions are processed and contribute empirical findings that will prove useful in building models of human speech processing.

1.5 Methods used in this thesis

In order to pursue the aim described above, it is necessary to employ a variety of different methods. These include traditional psycholinguistic methods such as the collection of decision response times, accuracy, acoustic measurements, and naming latencies. In addition, more advanced methods will be used such as visual-world eye tracking, event-related brain potentials (ERPs), and the computational analysis of corpus data. This section briefly describes eye tracking and ERPs.

The visual-world eye-tracking methodology allows one to measure *where* and *when* a participant is looking in response to an auditory stimulus presented in a psycholinguistic experiment. Participants' eye movements are monitored while several visual stimuli are shown on a computer screen and spoken instructions are presented asking the participant to click on one of these stimuli using a computer mouse. The item that participants are asked to click on is typically referred to as the target. The visual stimuli can consist of pictures of objects (e.g., a picture of a lion; see e.g. Cooper, 1974; Tanenhaus et al., 1995) or they can consist of printed words (e.g., the word "lion"; see e.g. Huettig and McQueen, 2007; McQueen and Viebahn, 2007). By tracking participants' eye movements while the speech input unfolds over time, researchers gain insights into how quickly a word is being recognized and which other words participants were paying attention to before deciding on which word to click. The eye-tracking data are typically presented in terms of a probability distribution over time showing how likely it was that participants looked at a given word in a particular time window measured from the onset of the target word. This methodology will be used in Chapters 3 and 5 of this thesis.

The event-related brain potentials (ERP) technique measures the brain's electrophysiological response to different types of sensory stimulation (such as visually or auditorily presented words or sentences). This measure is based on the electroencephalogram (EEG) which measures the changes in voltage over time that are emitted from the human scalp. As the EEG is extremely noisy, EEG recordings are averaged across multiple trials in order to create the ERP. During the averaging process, most of the EEG signal that is unrelated to the stimulus is removed and the electrophysiological response that is related to the stimulus becomes visible. There are a variety of ERP responses (called *components*), each with a particular latency and polarity (positive or negative), that are related to distinct aspects of linguistic processing. One of the most robust components is the N400 which has a negative polarity and occurs approximately 400 millisecond after the onset of a semantically meaningful stimulus (Kutas and Federmeier, 2011). The amplitude of this component changes depending on how expected the content of a meaningful utterance was. For example, the N400 component will be larger for the sentence *I take coffee with cream and dog* than for the sentence *I take coffee with cream and sugar*. Differences in components between experimental conditions are referred to as ERP *effects*. Chapter 4 of this thesis will present a study in which the ERP technique was used in order to investigate the processing of morphosyntactic violations. In this study, we were particularly interested in the P600 component, which has been shown to be related to syntactic processing.

The aim of this thesis, taking a multi-level approach to the processing of reduced speech, makes it necessary to use different methodologies because each of these methods is useful for investigating particular aspects of linguistic processing. For example, whereas eye tracking is particularly well suited to examine lexical competition between similar sounding words, ERPs are very well suited to investigate the violation of linguistic expectations (e.g., with respect to meaning and morphosyntax). There is no single method that can penetrate all levels of processing equally well. Instead, different levels of processing require different methods of investigation. As the aim of this thesis is to examine the influence of speech reduction on multiple linguistic levels, different methods are needed in order to tap into each level of processing.

In addition to using several experimental methods, this thesis also uses corpus data and combines corpus methodology with controlled experiments. Corpus studies allow the examination of the rich informational content and patterns that occur in language usage under natural conditions. As corpus studies use samples of naturally occurring language, they have high external validity. In contrast, experimental methods allow for a higher degree of control and stronger arguments with respect to conclusions about causal relationships between experimental variables. Experimental methods therefore provide a high degree of internal validity. Experimental methods and corpus studies thus complement each other. In Chapter 4 these methods will be used in combination in order to maximize the internal as well as the external validity of the study.

1.6 Outline of this thesis

Chapter 2 presents a corpus study that examines distributional properties of reduced pronunciation variants. More specifically, it describes an investigation of the likelihood with which reduced Dutch past participles (e.g., *geweest* "been") co-occur in natural speech. This study makes use of phonetic transcriptions that were automatically generated by an automatic speech recognition (ASR) system based on the Hidden Markov Model Toolkit (Young et al., 2002). These transcriptions were originally created by Hanique et al. (2013) and provided the foundation for the computational analyses presented in Chapter 2. Based on these phonetic transcriptions, Dutch past participles were extracted from three different speech registers. I investigated the influence of the time lag between succeeding past participles, speech register, and speaker identity on the presence and duration of schwas in prefixes and /t/s in suffixes. At the end of Chapter 2, the findings of these analyses are discussed with respect to their implications for psycholinguistic models of lexical representation.

Chapter 3 examines the interaction between acoustic-phonetic reductions and syntactic information in the comprehension of spoken words. For this purpose, I conducted a series of eye-tracking experiments that tested whether the recognition of spoken past participles is influenced by how predictable past participles are given their syntactic context. In addition, this study examined whether listeners assign more weight to syntactic predictability information when listening to casual speech compared to careful speech. Syntactic predictability was manipulated by varying the free word order of past participles and auxiliary verbs in Dutch subordinate clauses. For example, the past participle *geleund* "leaned" can be produced before or after its associated auxiliary verb *heeft* "have". Consequently, the sentence *Ik weet zeker dat hij heeft geleund op de houten tafel* "I know for sure that he leaned against the wooden table" can also be produced, without changing the meaning, with the reversed ordering of the auxiliary and the participle as in *Ik weet zeker dat hij geleund heeft op de houten tafel*. Because a past participle must follow an auxiliary verb in subordinate clauses, the past participle is more predictable in the version of the sentence in which the auxiliary precedes the participle than in the version in which the participle precedes the auxiliary. In Chapter 3, we investigated whether listeners are able to use this information in order to help them recognize the past participle.

Whereas Chapter 3 looked at how syntactic information influences the comprehension of reduced speech, Chapter 4 looked at the reverse: how reduced speech influences the processing of morphosyntactic information. An ERP study was conducted that investigated whether listeners tolerate grammatical gender violations in casual speech that are known to disrupt processing in careful speech. Native speakers of Dutch were presented with Dutch utterances that contained adjective-noun pairs in which the adjective was either correctly inflected with a word-final schwa (e.g. *een spannende roman* "a suspenseful novel") or incorrectly uninflected without that schwa (*een spannend_ roman*). Do listeners take information about the talker's speaking style into account when processing grammatically incorrect morphological inflections? How does this influence the morphosyntactic interpretation of phonetic information? Chapter 4 attempts to answer these questions and discusses the implications of the way in which listeners adapt to variable speech input for morphosyntactic processing.

Chapter 5 explores the role of orthography in the learning of pronunciation variants and compares it with the influence of variation in the phonetic input. For this study, a novel-word learning method was employed in which French-speaking participants learnt novel French words over three consecutive days by associating French nonwords (e.g., *le secobe*) with pictures of non-existing objects. The novel-word learning paradigm allows researchers to isolate specific linguistic variables, a goal which is often difficult to accomplish with natural language because the properties of words are highly intercorrelated (e.g., Landauer and Streeter, 1973). This paradigm has previously been employed in order to study language comprehension (e.g., Magnuson et al., 2003; Sulpizio and McQueen, 2012) as well as word production (e.g., Bürki et al., 2012). In Chapter 5, native speakers of French learnt novel French words that either contained the vowel schwa after the first consonant (e.g., *secobe*) or began with a consonant cluster (e.g., *scobe*). In Experiment 1, the pronunciation of some of the words varied (i.e., they were sometimes produced with schwa and sometimes produced without schwa). In Experiment 2, this manipulation was replaced by an orthographic exposure phase in which words that began with a consonant cluster (e.g., *scobe*) were either spelled with the letter <e> (as in *secobe*) or without it. In order to examine the effects of these experimental manipulations on word production, a naming task was employed. In order to examine the effects on language comprehension, the novel-word learning approach was combined with the visual-world eye-tracking paradigm (e.g., Allopenna et al., 1998; Tanenhaus et al., 2000). Chapter 5 discusses the implications of our findings with respect to the relevance of orthographic information compared to phonetic information for the processing of reduced pronunciation variants.

Finally, Chapter 6 summarizes and discusses the findings presented in the previous chapters with respect to the overall goal of this thesis. This work aims to use a multi-method approach in order to investigate how acoustic reduction influences psycholinguistic processing at multiple levels and how flexibility helps listeners to adapt to variation in the reduced speech input. Chapter 6 will evaluate the findings with respect to this goal and discuss their implications for our understanding of the role of acoustic reduction in spoken-word processing.

Co-occurrence of reduced word forms in casual speech

Chapter 2

This chapter is based on:
Malte Viebahn, Mirjam Ernestus and James McQueen (2012).
In *Proceedings of the 13th Annual Conference of the International Speech Communication Association (Interspeech 2012)*, pages 2021-2024.

This paper presents a corpus study that investigates the co-occurrence of reduced word forms in natural speech. We extracted Dutch past participles from three different speech registers and investigated the influence of several predictor variables on the presence and duration of schwas in prefixes and /t/s in suffixes. Our results suggest that reduced word forms tend to co-occur even if we partial out the effect of speech rate. The implications of our findings for episodic and abstractionist models of lexical representation are discussed.

2.1 Introduction

In spontaneous conversations, the way speakers produce words often deviates from dictionary pronunciations. For example, the schwa in the English word *separate* may be shortened or omitted resulting in the reduced form *sep'rate*. In Dutch, schwas are frequently reduced when they occur in the prefix of past participles (Pluymaekers et al., 2005b). For instance, the Dutch word *geweest* may be pronounced as *g'weest*. In the present study, we investigated whether reduced word forms are likely to co-occur in natural speech. The main question we asked was whether the degree of reduction of a given word is influenced by the degree of reduction of a previous word. In addition, we investigated if this effect is influenced by the time lag between the succeeding words and the degree to which words overlap phonologically.

Data on co-occurrences among word forms can inform us about the nature of lexical representation. Theories of the mental lexicon differ with respect to the degree to which phonetic variation is represented in lexical memory. On the one hand, there are abstractionist theories that claim that words are stored as single abstract representations (Levelt et al., 1999). On the other hand, there are episodic theories that assume that fine-grained phonetic detail of each occurrence of a word is stored (Goldinger, 1998). Between these two extreme viewpoints there are hybrid theories that vary with respect to the number of phonetic variants that are stored and the degree to which phonetic detail is retained in memory (Bürki et al., 2010). Whereas extreme abstractionist accounts predict no increase in acoustic similarity between word forms that occur together, extreme episodic theories predict that co-occurring words have similar acoustic properties, especially if they occur closely together and show greater phonological overlap.

2.2 Method

2.2.1 Materials

The data were extracted from the Ernestus Corpus of Spontaneous Dutch (ECSD, Ernestus, 2000) and two components of the Corpus of Spoken Dutch (CGN, Oostdijk, 2002). We collected a total of 3,241 Dutch past participles from three speech registers: spontaneous speech from the ECSD, interviews from the CGN, and read speech from the CGN. We focused on past participles because they are frequent and often subject to reductions in Dutch (Pluymaekers et al., 2005b). Furthermore, they provide two different kinds of reductions: schwa reduction in the prefixes and /t/ reduction in the suffix. Following Hanique et al. (2013), our analyses were restricted to tokens that fulfilled the following criteria: words were only included if they were not followed by a hesitation, participles with the prefix /vər/ were only included if the /r/ in the prefix was not realized, the duration of the following consonant (after the schwa) had to fall between 8 and 478 milliseconds, and, for the /t/ reduction data, words were only included if they were not followed by a word starting with a stop consonant (i.e. /tdpbk/). Each pair of past participles was spoken by the same speaker. The time lag between two succeeding past participles was defined as the interval between the offset of the first and the onset of the second past participle. Our analyses were based on the output of the automatic speech recognition system used by Schuppler et al. (2011). Using a very similar dataset, Hanique et al. (2013) have shown that the agreement between human transcribers and the automatically generated transcriptions is approximately as high as the agreement among the human transcribers.

2.2.2 Variables

We measured speech reduction with two different variables: segment presence (i.e. did the speaker produce a /t/ or a schwa, or not?) and segment duration (i.e. how long was the produced /t/ or schwa?). Our main predictor variables were the realization of the same segment (presence and duration) in the preceding past participle, the time lag between the two words, and their phonological relatedness. The variable relatedness originally had three levels: the two participles were coded as identical, as different words with the same prefix, or as different words with different prefixes. However, because we did not find any differences between different words with or without the same prefix, we collapsed these two levels into one. For the variables segment duration, previous segment duration, and time lag we took the natural logarithm in order to obtain normal distributions. Furthermore, we included register and speech rate as control variables. Speech rate was defined as the number of syllables in the full word forms per second within a continuous chunk of speech containing the past participle. These variables were included in order to rule out that possible effects of the realization of the preceding past participle were due to similarity in speech rate or register.

2.2.3 Statistical analyses

We used linear mixed effects models and included speaker and word type as random variables in our analyses. For categorical predictor variables, we used contrast coding. The baseline level for the variable corpus was spontaneous speech, for presence of the previous segment it was segment absent, and for relatedness it was different words. For models in which the dependent variable was continuous (i.e. segment duration), we excluded data points with standardized residuals larger than 2.5. The proportion of excluded values ranged from 0% to 3.7%.

Predictor	β	t value
prev. schwa	−0.06	−0.35
lag	−0.03	−0.41
identical words	−7.39	−3.25
speech rate	−0.07	−5.52
prev. schwa × lag	0.01	0.30
prev. schwa × identical words	1.87	3.23
lag × identical words	0.74	2.92
prev. schwa × lag × identical words	−0.19	−2.87

Table 2.1: Results for schwa duration predicted by previous schwa duration (prev. schwa).

2.3 Results

2.3.1 Schwa reduction

In the first model our dependent variable was schwa presence and our critical predictor variable was presence of the schwa in the previous token. A total of 1,848 tokens were included in the analysis. We found only significant effects of speech rate and register. Schwas were less likely to be present the higher the speech rate ($\beta = 0.26, t = 5.21$) and less likely to be realized in spontaneous speech than in read speech ($\beta = 2.01, t = 9.66$) or interview speech ($\beta = 0.79, t = 4.08$). Note that absolute t values of equal to or larger that two indicate statistical significance.

In the second model, the critical predictor variable was changed to the duration of the preceding schwa. Tokens in which the schwa was absent in the preceding past participle were removed, leaving 1,259 tokens. We found a similar pattern of results: no effects of the duration of the preceding schwa, time lag, or relatedness, but there were again fewer schwas in faster speech ($\beta = 0.23, t = 3.50$) and fewer schwas in spontaneous speech than in read speech ($\beta = 1.998, t = 8.35$) or interview speech ($\beta = 0.82, t = 3.54$).

In the third model, the dependent variable was schwa duration and the critical predictor variable was the duration of the schwa in the preceding participle. All tokens in which the schwa was absent in either the first or the second past participle were removed, leaving 918 tokens. This model indicates a significant three-way interaction between the duration of the previous schwa, time lag, and relatedness (see Table 2.1 and Figure 2.1).

Separate analyses were then conducted for different and identical words. For different words ($N = 864$), we found no effects of the duration of the previous schwa, time lag, and no interaction. The only significant predictor was speech rate (shorter schwas at higher rates: $\beta = 0.07, t = 5.18$). In contrast, for the tokens that were

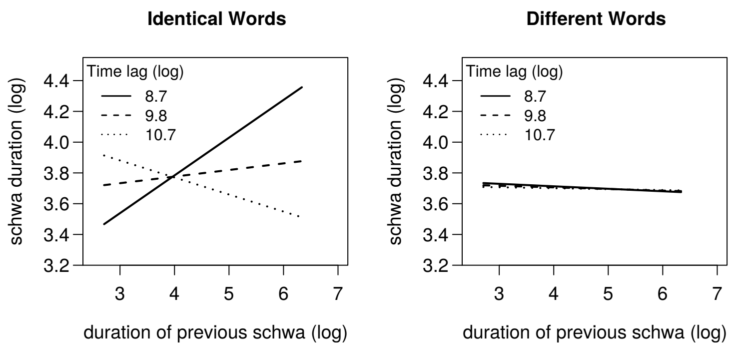


Figure 2.1: Interaction between relatedness, time lag, and duration of previous schwa for schwa duration predicted by previous schwa duration. The different lines for time lag represent the 25th, 50th, and 75th percentiles.

preceded by identical words ($N = 54$), we found effects of the duration of the previous schwa ($\beta = 2.07, t = 3.83$) and lag ($\beta = 0.79, t = 3.33$) and these predictors interacted ($\beta = 0.21, t = 3.43$). We then performed a median split on time lag and ran separate models for tokens with lags shorter or longer than the median. For the short time lag, we found a significant effect of previous schwa duration ($\beta = 0.42, t = 3.07$): schwa durations of succeeding tokens are similar when they occur closely together in time. In contrast, we found no significant effect for the long time lag ($\beta = 0.03, t = 0.22$).

In the fourth model, the dependent variable was schwa duration and the critical predictor variable was the presence of the preceding schwa. Tokens in which schwa was absent were removed, leaving 1,250 tokens. This analysis showed only an effect of speech rate ($\beta = 0.07, t = 6.64$): schwa durations tended to be shorter when speech rate increased.

2.3.2 /t/ reduction

We investigated the co-occurrence of reduced /t/s conducting the same kind of analyses as we did for schwa reduction. First, we ran a model with /t/ presence as the dependent variable and the presence of the /t/ in the preceding past participle as the critical predictor variable. A total of 856 tokens were included. This model showed effects of only speech rate (fewer /t/s in faster speech: $\beta = 0.27, t = 3.49$) and speech register: fewer /t/s in spontaneous speech than in read speech ($\beta = 1.67, t = 5.79$) or interview speech ($\beta = 0.63, t = 2.26$).

In the second model, /t/ presence was the dependent variable and the duration of the previous /t/ was the critical predictor variable. We excluded all tokens

Predictor	β	t value
prev. /t/	−0.09	−2.29
identical words	−0.18	−2.27
speech rate	−0.10	−5.29
read speech (CGN)	0.16	2.08
interview speech (CGN)	−0.19	−2.20

Table 2.2: Results for /t/ duration predicted by previous /t/ duration (prev. /t/).

in which the /t/ was not present in the previous participle, leaving 635 tokens. Significant effects were present only for speech rate (fewer /t/s in faster speech: $\beta = 0.32, t = 3.28$) and speech register (fewer /t/s in spontaneous than in read speech: $\beta = 1.47, t = 4.29$).

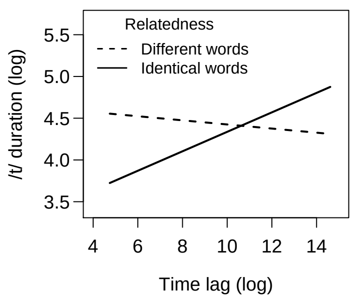


Figure 2.2: Interaction between relatedness and time lag for /t/ duration predicted by previous /t/ presence for spontaneous speech.

In the third model, /t/ duration was the dependent variable and the duration of the preceding /t/ the critical predictor variable. Tokens in which the /t/ was not present either in the first or the second participle were removed, leaving 496 data points. This model showed significant main effects of previous /t/ duration, relatedness, speech rate, and speech register (see Table 2.2). Shorter /t/s were associated with longer /t/s in the previous participle. Furthermore, /t/s were shorter when the preceding word was identical than when it was different.

In the fourth model, /t/ duration was the dependent variable and the presence of the preceding /t/ was included as a predictor variable. We excluded tokens in which the /t/ was not present in the second participle, leaving 642 data points. We found

Predictor	β	t value
lag	−0.03	−1.94
identical words	−1.50	−3.39
speech rate	−0.08	−5.18
read speech (CGN)	0.06	1.07
interview speech (CGN)	−0.24	−3.89
lag × identical words	0.14	3.07

Table 2.3: Results for /t/ duration predicted by previous /t/ presence.

significant effects of speech rate and speech register and an interaction of time lag and relatedness (see Table 2.3 and Figure 2.2). An analysis of the simple effects showed a significant negative slope for different words (shorter /t/s at longer lags: $\beta = 0.03, t = 2.02; N = 606$) and a significant positive slope for identical words (longer /t/s at longer lags: $\beta = 0.12, t = 2.36; N = 36$).

2.4 Discussion and conclusions

In the present study, we investigated if reduced word forms co-occur with one another in natural speech. For the schwas, we found a positive relationship between schwa durations of succeeding tokens when the tokens belong to identical words and occur closely together in time. For the /t/s, we found a negative correlation between /t/ durations of succeeding past participles, and a positive relationship between segment duration and time lag for identical words and a negative relationship for different words.

How do current models of lexical memory account for these results? Episodic accounts could explain the schwa results by assuming that the acoustic properties of a given word are influenced most strongly by the immediately preceding production of this word. Therefore, if the most recent trace had a relatively short schwa, the next occurrence will have a relatively short schwa, too. When the time lag between occurrences is long, the influence of the previous production becomes smaller and the acoustic properties of the next word regresses towards the mean of all traces in memory. An abstract account might explain the similarity among schwa durations by assuming priming at a post-lexical level. The motor commands that underlie the schwa production may be primed by the commands of the preceding schwa. This would account for similar schwa pronunciations without requiring multiple representations.

For the /t/s we found a negative correlation between the durations of succeeding tokens and we did not find an interaction with phonological relatedness. This could indicate that /t/s are processed differently from schwas. However, given that the /t/ dataset was substantially smaller than the schwa dataset (N = 918 vs. N = 496), it is possible that the absence of a significant interaction is due to a lack of power. This possibility is supported by a comparison to the results from the model in which /t/ duration was predicted by /t/ presence (rather than duration). In this model we found a significant interaction between time lag and relatedness which was absent in the model predicting /t/ duration from the previous /t/ duration. Nevertheless, apart from the critical predictor variable, the two models were essentially the same and differed only in the number of data points (N = 496 vs. N = 642). The fact that we did not find this interaction in the former model further supports the notion that it had insufficient power.

Furthermore, we found a positive correlation between time lag and /t/ duration for identical words and a negative one for different words. The positive correlation could be explained in terms of a priming mechanism that facilitates lexical access and speeds up articulation, leading to shorter segment durations. The negative correlation may be due to inhibitory connections between lexical representations. When there is only little time between the production of two different words, producing the second word may take more time due to lingering activation of the first word. As the first word is still activated it will be harder to activate the second word. Consequently, the second word will be articulated more slowly. As time passes by, activation of the first word decays and the time to access and articulate the second word returns to a normal level, explaining the negative correlation between time lag and /t/ duration.

In conclusion, our results suggest that reduced word forms tend to occur together. This co-occurrence can not be attributed only to speech rate. Episodic models may account for this pattern of results by assuming multiple representations whereas abstractionist theories may attribute these effects to priming of post-lexical processes.

Syntactic predictability in the recognition of carefully and casually produced speech

Chapter 3

This chapter is based on:
Malte Viebahn, Mirjam Ernestus and James McQueen (2015).
Syntactic predictability in the recognition of carefully and casually produced speech.
Journal of Experimental Psychology: Language, Memory, and Cognition, 41 (6), 1684-1702.

The present study investigated whether the recognition of spoken words is influenced by how predictable they are given their syntactic context and whether listeners assign more weight to syntactic predictability when acoustic-phonetic information is less reliable. Syntactic predictability was manipulated by varying the word order of past participles and auxiliary verbs in Dutch subordinate clauses. Acoustic-phonetic reliability was manipulated by presenting sentences either in a careful or a casual speaking style. In 3 eye-tracking experiments, participants recognized past participles more quickly when they occurred after their associated auxiliary verbs than when they preceded them. Response measures tapping into later stages of processing suggested that this effect was stronger for casually than for carefully produced sentences. These findings provide further evidence that syntactic predictability can influence word recognition and that this type of information is particularly useful for coping with acoustic-phonetic reductions in conversational speech. We conclude that listeners dynamically adapt to the different sources of linguistic information available to them.

3.1 Introduction

Language comprehension is a complex task. Listeners are confronted with two or three words per second (Levelt et al., 1999) and have to choose from many thousands of words in their mental lexicon. This task is complicated by the fact that the pronunciation of words is often quite variable. In natural conversations, speech is typically produced with a casual speaking style leading to the omission or acoustic weakening of individual phonemes and even whole syllables (Ernestus and Warner, 2011). Johnson (2004) estimated that over 60% of the words uttered in casual speech deviate from their citation forms by at least one phoneme and 28% deviate by two or more phonemes. Despite these facts, listeners are able to recognize casual speech quickly and accurately. One of the reasons for this ability is that listeners can use multiple sources of information from the sentence context to predict upcoming words (e.g. Altmann and Kamide, 1999; Arai and Keller, 2013). In the present study, we investigated if listeners can use word order information in Dutch subordinate clauses to predict upcoming past participles. In particular, we explored how this type of information interacts with acoustic-phonetic information by presenting spoken sentences either in a casual or a careful speaking style.

Previous research has demonstrated that listeners are able to use semantic context and discourse-based information to anticipate upcoming words (e.g. Altmann and Kamide, 1999; Brouwer et al., 2013; Chambers et al., 2002). Using a visual-world eye-tracking task, Altmann and Kamide (1999) showed that participants were more likely to look at a picture of a cake after hearing *The boy will eat...* than after hearing *The boy will move...*, demonstrating that listeners can use the semantic content of verbs to anticipate subsequent nouns. Semantic context effects have also been found in event-related potential studies (e.g. van Berkum et al., 2005; Wicha et al., 2004). For example, DeLong et al. (2005) presented sentences starting with, for instance, *The day was breezy so the boy went outside to fly...* followed either by a high-probability noun (e.g., *a kite*) or a low-probability noun (e.g., *an airplane*). Crucially, the form of the article preceding the noun (*a* vs. *an*) differed between the high- and low-probability nouns. DeLong et al. found larger N400 components while participants were reading the article matching the low-probability noun (*an*) compared with the article matching the high-probability noun (*a*). This shows that readers had already predicted the high-probability noun when reading its preceding article.

In addition to semantic and discourse-based information, listeners are also able to use syntactic cues for prediction. For example, Kamide et al. (2003) showed that case marking information in German can be used by listeners to anticipate nouns. Furthermore, Arai and Keller (2013) showed that whether a verb is transitive or intransitive influences what sentence continuations listeners predict. In a visual-world study, the authors found that listeners were more likely to anticipate upcoming objects (e.g., *the artist*) when the verb in the sentence was transitive (as in *The nun punished the artist*) rather than intransitive (as in *The nun agreed and the artist threw the kettle*).

Although there are many studies that investigated prediction in auditory and visual sentence processing, only a few studies have investigated the effect of predictive contextual information on how phonological information is evaluated (e.g. Magnuson et al., 2008; van Alphen and McQueen, 2001). Dahan and Tanenhaus (2004) showed that semantic information can decrease lexical competition among similar sounding words. In a visual-world experiment in Dutch, participants were presented with sentences in which the main verb occurred either before the target noun (as in *Never before climbed a goat so high*) or after the target noun (as in *Never before has a goat climbed so high*). It was found that when the main verb occurred after the target noun, there was competition between the noun (*goat*, Dutch: *bok*) and a similar sounding word (*bone*, Dutch: *bot*). However, when the main verb occurred before the noun, evidence for lexical competition disappeared. This result shows that the predictions that listeners make based on the semantic restrictions of verbs influence how phonological information is evaluated during lexical processing. It is important to note that Dahan and Tanenhaus also showed that semantic context does not make listeners ignore subsequent acoustic information. When the initial part of the target word was cross-spliced with the phonological competitor (e.g., the "bo" of *bot* spliced onto the "k" of *bok*), competition between the target and the phonological competitor was present even when the main verb preceded the target noun.

In the present study, we were concerned with how the predictability of a syntactic word class can influence word recognition. More specifically, we were interested in the predictability of past participles in Dutch subordinate clauses. We took advantage of the fact that the word order of auxiliary verbs and past participles in Dutch subordinate clauses is free: either the auxiliary can precede the participle, or the participle can precede the auxiliary. Take, for instance, the sentence I know for sure that he has leaned against the table. This sentence could be translated into Dutch in two ways:

(a) Ik weet zeker dat hij <u>heeft geleund</u> op de houten tafel.

(b) Ik weet zeker dat hij <u>geleund heeft</u> op de houten tafel.

These two translations differ only in the order in which the auxiliary verb (*heeft* English: *has*) and the past participle (*geleund*, English: *leaned*) occur. In the Dutch linguistics literature, the two word orders have been referred to as the *red* (i.e., auxiliary-first) and the *green* (i.e., participle-first) word orders (e.g. Pardoen, 1991).

There are multiple variables that influence when speakers prefer to use one word order over the other (e.g. de Sutter, 2009; Swerts and van Wijk, 2005). Among these are prosodic, morphosyntactic, semantic, and discourse-related factors. Furthermore, there are regional differences between Dutch and Belgian speakers of Dutch. In Flemish (the variant of Dutch spoken in Belgium), the participle-first word order is used almost exclusively whereas in Dutch spoken in the Netherlands there is more variation (Barbiers et al., 2008). Whereas both word orders occur to a similar extent in the central parts of the Netherlands (especially the east), the participle-first word order is more common in the northern and southern parts.

It is important to note that the two word orders differ in how predictable the past participle is. In the auxiliary-first construction, the participle is more predictable because the auxiliary indicates that a participle must follow immediately. In our study, we compare the recognition of participles in the more predictable (i.e., auxiliary-first) and the less predictable (i.e., participle-first) syntactic construction. If listeners are sensitive to the predictive information provided by the auxiliary, we expect that the auxiliary-first word order will lead to faster processing of the participle than the participle-first word order.

Whether and how much listeners use syntactic information to predict upcoming words may depend on the listening conditions. When listeners attempt to identify spoken words, they rely primarily on the phonological information provided by the speech signal McQueen (2007). This information concerns not only the word currently being recognized, but also its acoustic context (e.g. Lieberman, 1963; Pollack and Pickett, 1963, 1964). Especially under conditions in which the speech stream does not provide reliable cues, contextual cues become an important source of information for identifying the words and segments of speech. This has been demonstrated by studies investigating the recognition of acoustically reduced speech (e.g. Brouwer et al., 2012b; Ernestus et al., 2002; Janse and Ernestus, 2011; van de Ven et al., 2012). For example, Ernestus et al. (2002) presented strongly reduced word forms extracted from a corpus of spontaneous Dutch either in isolation or in context. Listeners' ability to recognize the words was heavily influenced by the amount of context available. Whereas the proportion of correct identifications was only a little more than 50% when words were presented in isolation, it increased to more than 90% when words were presented in full sentence contexts. The strong effect of context suggests that the importance of contextual information depends on the reliability of the acoustic information provided by the words themselves. For reduced speech, in which acoustic information is less reliable, context is therefore more important compared with careful speech.

We investigated the influence of different listening conditions on effects of syntactic predictability by presenting both word orders in a careful and a casual speaking style. A casual speaking style often results in acoustic reductions that can make words more confusable Brouwer et al. (2012a). Among the segments that are particularly often reduced in Dutch are schwas, for instance in prefixes of past participles (e.g. Ernestus, 2000; Hanique et al., 2013). Dutch past participles such as *geleund* (pronounced as [xəlønt], English: *leaned*) consist of a prefix containing a schwa ([xə], [bə] or [fər]), a stem (e.g., [løn]), and a suffix (e.g., [t]). When the schwa in the prefix is reduced or deleted, the participle is more similar to other words that are not past participles. For example, the word *geleund* becomes more similar to the first syllable of the word *gleuven* (English: *grooves*). When confronted with casual speech, listeners can therefore not reliably say if the sequence [xlø] is the onset of the past participle *geleund* or of the noun *gleuven*. There is ample evidence in the literature suggesting that phonological overlap among words increases lexical competition (e.g. Allopenna et al., 1998; Brouwer et al., 2012a; Luce and Pisoni, 1998; McQueen and Viebahn, 2007). We would therefore expect that a casual speaking style should increase lexical competition and slow down word recognition.

There are at least two ways in which a casual speaking style could influence the way in which syntactic information is evaluated. On the one hand, syntactic information may become more important for listeners when they are confronted with casual speech. As there is more phonological ambiguity in casual speech than in careful speech, listeners may benefit more from syntactic predictability. On the other hand, in casual speech the whole utterance is likely to be affected by acoustic reduction. This includes the words that carry syntactic information that could be used to predict upcoming words, such as auxiliary verbs. Syntactic information may therefore play less of a role in the processing of casual speech because it is more difficult to extract from the speech signal. A recent study by van de Ven et al. (2011) using semantic priming suggests that acoustically reduced words only function as effective primes if the time lag between prime and target is relatively long. This suggests that listeners require time to recover from acoustic reductions. We may therefore find that listeners are less able to make use of predictive syntactic information when the words that carry it are acoustically reduced because listeners may not have enough time to recover from the reduction of the auxiliary verb before hearing the participle. The idea that the importance of some linguistic cues depends on the availability of other cues is part of both accounts and is consistent with the frameworks on cue integration in speech segmentation by Mattys et al. (2005) and Norris et al. (1997). According to these approaches, listeners use lexical, segmental, and prosodic cues for speech segmentation but the weight that is assigned to each cue depends on its availability in the signal.

In the following experiments, we used a printed-word variant of the visual-world eye-tracking paradigm (e.g. Huettig and McQueen, 2007; McQueen and Viebahn, 2007). We presented listeners with carefully and casually produced sentences in which the target word was a past participle that either preceded or followed its associated auxiliary verb. While listening to these utterances, participants looked at visual displays of quadruplets of printed words that included the target participle as well as a nonparticiple that overlapped with the target word phonologically. The overlap between target and competitor word was larger if the target word was produced in a reduced rather than a careful way. We measured the reaction times (RTs) with which participants identified the target word as well as how much they paid attention to the similar sounding nonparticiple. If syntactic predictability helps listeners identify spoken words, we expect participants to identify the participle more quickly and to be less distracted by the similar sounding nonparticiple when the participle follows its auxiliary verb compared with when it precedes it. We also expect participants to identify the participle more quickly in careful than in casual speech because the phonological overlap with the nonparticiple will be smaller. Furthermore, we may find that the effect of word order is stronger for casual speech than for careful speech because listeners may rely more on syntactic information when the acoustic cues are less reliable. Alternatively, we may find an interaction that goes into the opposite direction. In the casual speaking-style condition, the whole utterance, including the auxiliary, is produced in an acoustically reduced way. Listeners may therefore have difficulties identifying the auxiliary and may consequently not be able to use it to predict the upcoming past participle. As a result, the effect of word order may be weaker (or even absent) in the casual speaking-style condition.

3.2 Experiment 1

3.2.1 Method

Participants

Forty-eight native speakers of Dutch were recruited from the subject panel of the Max Planck Institute for Psycholinguistics. All were university students. Age ranged from 18 to 27 years. The participants reported no hearing problems and had normal or corrected-to-normal vision. They were informed about the procedure of the experiment before taking part and were paid for their participation.

Materials

Sixty-four pairs of Dutch participles and nonparticiples were selected from the CELEX database (Baayen et al., 1995). The participles had a mean frequency of 191 per million and the nonparticiples of 347 per million. All words started with the letter <g>

(pronounced as /x/). The pairs were chosen such that the two words overlap for the initial three phonemes when the schwa in the past participle is absent. For example, the words of the pair *geleund-gleuven* (*leaned-grooves*) overlap for the initial three phonemes /x/, /l/, and /ø/ when the schwa in *geleund* is absent. All words are listed in the Table 3.1.

For each pair, two carrier sentences were constructed, one that contained the participle and one that contained the nonparticiple. The sentences that contained the participle were used in the experimental trials. On these trials, the participle was the target word and the nonparticiple was the competitor. The sentences that contained the nonparticiple served as filler trials. On these trials, the nonparticiple was the target word. The sentences were identical up until the onset of each target word (see Table 3.2 for an example). The target words were positioned approximately in the center of their carrier sentences. For the experimental sentences, two versions were constructed: in one version, the auxiliary preceded the participle and in the other version, the auxiliary followed the participle (see Table 3.2). For each version, a carefully and a casually produced recording was made. In the casually produced recording, the target word (but also other words in the sentence) was produced in an acoustically reduced way. For the experimental sentences, this resulted in four different versions of each sentence: an auxiliary-first version that was carefully produced, an auxiliary-first version that was casually produced, a participle-first version that was carefully produced, and a participle-first version that was casually produced. For the filler sentences, this resulted in two different versions: one that was carefully and one that was casually produced. In addition, eight practice sentences were created. Four of these sentences contained a participle and four did not. In half of the sentences containing a participle, the auxiliary preceded the participle and in the other half the auxiliary followed the participle.

To investigate whether there was a preference for one of the two word orders, we conducted a rating experiment. Twelve participants (native speakers of Dutch) who did not participate in the other experiments were simultaneously presented with visual representations of both syntactic versions of each sentence. We asked the participants to indicate which version they preferred by using a scale from 1 (auxiliary-first) to 6 (participle-first). The mean rating was 3.39. Independent-samples t tests showed that subject and item means did not differ significantly from 3.5 ($t_1(11) = -0.52$, $p = 0.62$; $t_2(63) = -1.81$, $p = 0.08$), indicating that there was no preference for either word order.

For the construction of the visual displays, the 64 participle-nonparticiple pairs were combined into 32 quadruplets (see Table 3.1). For example, the pair *geleund-gleuven* (*leaned-grooves*) and the pair *gelift-glimlachte* (*lifted-smiled*) were combined into a quadruplet. Each visual display consisted of one quadruplet. On a given

Quadruplet	Participle	Nonparticiple	Participle	Nonparticiple
1	geroosterd (roasted)	grootschalig (large-scale)	geroerd (stirred)	groette (greeted)
2	gerold (rolled)	gromde (growled)	geroepen (called)	groene (green)
3	geraspt (grated)	grappig (funny)	gericht (aimed)	grind (gravel)
4	gerot (rotted)	grotten (caves)	geroest (rusted)	groeven (grooves)
5	gered (saved)	grendeltje (lock)	geregeld (organized)	grenen (pine)
6	gelopen (run)	glooiend (declining)	geluwd (abated)	glurende (peeking)
7	geraadpleegd (consulted)	gracieus (gracefully)	gereikt (reached)	grijpkranen (cranes)
8	gelapt (patched)	glanzend (glossy)	gelikt (licked)	glitters (glitter)
9	geloofd (believed)	glorie (glory)	gelaveerd (maneuvered)	glazuur (gloss/icing)
10	gerangeerd (shunted)	graffiti (graffiti)	geriskeerd (risked)	griffierschap (clerkship)
11	gereisd (traveled)	grijpgrage (grabby)	geruimd (cleared)	gruiskolen (coal dust)
12	gelaagd (layered)	glazig (glassy)	gelijmd (glued)	glijbanen (slides)
13	geranseld (whipped)	grassprieten (blades of grass)	gerinkeld (jingled)	grimmig (grim)
14	geluisterd (listened)	gluiperds (weaselly)	gelicht (shined)	glibberig (slippery)
15	geleid (led)	glijdend (sliding)	gelogen (lied)	globes (globes)
16	geremd (inhibited)	greppels (ditches)	gerammeld (rattled)	grabbeltonnen (grab bags)
17	gelost (unloaded)	glommen (shine)	geloeid (mooed)	gloeiende (glowing)
18	gerezen (raised)	gretig (eagerly)	gerafeld (frayed)	grafisch (graphic)
19	geraapt (picked)	graaide (snatched)	geringd (ringed)	grilde (grilled)
20	geleund (leaned)	gleuven (grooves)	gelift (hitchhiked)	glimlachte (smiled)
21	gerimpeld (wrinkled)	grinnikend (chuckling)	geramd (rammed)	grandioos (magnificently)
22	gerond (circumnavigated)	grofweg (roughly)	geroemd (praised)	groepen (groups)
23	gerouwd (mourned)	grauwe (gray)	gerommeld (rummaged)	grondig (thoroughly)
24	gerept (rushed)	grenzen (limits)	geraakt (hit)	gratis (cost-free)
25	gerangschikt (arranged)	grammen (grams)	geritseld (managed)	grillige (bizarre)
26	gelucht (aired)	glunderen (smile)	gelaten (let)	glaasjes (glasses)
27	gerild (trembled)	griffels (pencils)	geraamd (estimated)	graag (gladly)
28	geraasd (raged)	graanvelden (cornfields)	gereinigd (cleaned)	grijnzend (smiling)
29	gelogeerd (stayed/lodged)	globale (global)	geloerd (lurked)	gloednieuwe (brand new)
30	gerekend (counted)	grepen (holds)	geraden (guessed)	graatmager (skinny)
31	geragd (stuck out)	grafurnen (urns)	gerijpt (matured)	grijs (gray)
32	geronseld (recruited)	grove (coarse)	geroeid (rowed)	groeiende (growing)

Table 3.1: Quadruplets used in the Experiments.

Experiment and stimulus example	Condition
Experiment 1 and 3	
Ik weet zeker dat hij heeft geleund op de houten tafel. *(I know for sure that he has leaned against the wooden table.)*	Auxiliary first
Ik weet zeker dat hij geleund heeft op de houten tafel. *(I know for sure that he leaned has against the wooden table.)*	Participle first
Ik weet zeker dat hij gleuven maakte in de houten tafel. *(I know for sure that he made grooves into the wooden table.)*	Filler
Experiment 2	
Ik weet zeker dat hij heeft geleund *(I know for sure that he has leaned)*	Auxiliary first
Ik weet zeker dat hij geleund heeft *(I know for sure that he leaned has)*	Participle first
Ik weet zeker dat hij gleuven *(I know for sure that he grooves)*	Filler

Table 3.2: Example Stimuli for the Two Word Order Conditions in Experiments 1-3.

trial, the words from one participle-nonparticiple pair served as target and competitor, respectively, while the words from the other pair served as distractors. The words across the two pairs shared the second consonant but differed in the following vowel. For instance, all four words in the example quadruplet contain the consonant /l/ but only the words in the same target-competitor pair share the same vowel following the /l/ (*geleund* and *gleuven* share the /ø/ whereas *gelift* and *glimlachte* share the /ɪ/). Each visual display was presented four times such that each word was the target once. As a result, there were 128 trials and eight practice trials. The words of a given quadruplet were presented pseudorandomly across the four positions on the screen such that the words occurred in different positions each time a quadruplet was repeated.

Three pseudorandomized running orders were created such that each presentation of a given quadruplet was separated by at least three trials. For each running order, experimental sentences were randomly assigned to one of the four conditions with an equal number of sentences per condition. Each running order was then rotated through the remaining three conditions resulting in 12 different experimental lists. An equal number of participants was assigned to each list. For the practice sentences, the quadruplets were not repeated. Each practice trial consisted of a unique quadruplet and a unique target sentence.

Recordings and acoustic analyses

The sentences were recorded in a sound-proof booth by a female native speaker of Dutch. For the casual sentences, the speaker was instructed to speak in a fast and casual way. It was explicitly stated that acoustic reductions were desirable. For the careful sentences, the speaker was instructed to speak in a clear and careful manner and to avoid acoustic reductions. We investigated if the different speaking styles influenced the acoustic properties of the stimuli by analyzing sentence duration, target word duration, auxiliary verb duration, schwa presence, schwa duration, initial /x/ duration, speaking rate until target word onset, and the divergence point between the target and the other words in its quadruplet. We defined the divergence point as the earliest point in time, measured from the beginning of the word, at which a word differs phonologically from the other words in the quadruplet. A schwa was judged to be present if there was a detectable portion of vocalic energy of at least one pitch period. Note that this does not mean that there were absolutely no cues to schwa in the tokens labeled in this analysis as having no schwa; nevertheless, such tokens are more reduced than those with an identifiable schwa. All of these acoustic measures are listed in Table 3.3 for the experimental and the filler items.

To determine whether speaking style and word order influenced the acoustic properties of the stimuli, we employed linear mixed-effects models. Word order and speaking style were entered as fixed effects and random intercepts were included for each participle. For durations and speaking rate we used linear mixed-effects models and for schwa presence we used generalized mixed-effects models with a binomial link function. For the linear mixed-effects models, data points with standardized residuals of two and a half or more were considered outliers and removed from the analysis. All measures except for the divergence points suggested that the casually produced sentences were acoustically more reduced than the carefully produced ones. A casual speaking style resulted in a higher speaking rate and shorter sentence, participle, and critical schwa durations. Furthermore, the probability that a past participle contained a schwa was smaller for casually than for carefully produced sentences (all $|t| < 2.00$ and $p > 0.01$). The absence of an effect for the divergence points suggests that the increase in segmental overlap for casual words in which the schwa was absent (82.8% of the cases) is traded against shorter word durations. As casual words tended to have no detectable schwa in the prefix, the phonological overlap with the competitor increases. However, because they are produced more quickly, the overlapping phonemes are squeezed together temporally such that their divergence points do not differ from those of the carefully produced words.

In addition to the effects of speaking style, we also found effects of word order (all $|t| > 2.00$ and $p < 0.05$). Participles in the auxiliary-first word order had shorter /x/

	Auxiliary first		Participle first		Filler items	
Acoustic property	Careful	Casual	Careful	Casual	Careful	Casual
Sentence duration	2,678	2,365	2,642	2,350	2,661	2,435
Target duration	429	325	425	338	381	343
Auxiliary duration	170	148	141	127	n/a	n/a
Schwa presence	100	15.62	98.44	18.75	n/a	n/a
Schwa duration	50	14	44	13	n/a	n/a
/x/ duration	81	77	97	90	86	86
Speaking rate	6.59	7.45	6.68	7.54	6.88	7.69
Divergence point	187	190	195	201	101	97

Table 3.3: Acoustic Properties of the Stimuli Used in the Three Experiments. Durations and divergence points are given in milliseconds. For the calculation of the average schwa durations, only participles with a schwa duration larger than zero were included. Schwa presence is expressed as a percentage and speaking rate is expressed as number of syllables per second.

durations and earlier divergence points while having longer schwas and sentences. Furthermore, target words in casual sentences were longer in participle-first than in auxiliary-first sentences. These results are consistent with studies showing that words that are more predictable have shorter durations and more reduced prefixes (Bell et al., 2009; Lieberman, 1963; Pluymaekers et al., 2005b). The shorter /x/ durations might be due to the fact that participles are more predictable when following an auxiliary verb than when preceding it. The slightly longer schwa durations (mean: 6 ms) might be a small compensatory effect. However, because the increase in the schwa duration does not fully absorb the decrease in the /x/ duration, the divergence points occur earlier than in the participle-first word order. One may speculate that the fact that participle word durations were affected by word order only when produced casually indicates that predictability effects are stronger in casual than in careful speech production.

In summary, the acoustic measures confirmed that our casually produced stimuli were acoustically more reduced than our carefully produced stimuli. Furthermore, we found evidence suggesting that participles in the auxiliary-first word order are more reduced than participles in the participle-first word order. This finding is consistent with previous studies suggesting that words that are more predictable are more likely to be produced in a reduced way.

At first sight the fact that participles are more reduced in the auxiliary-first word order than in the participle-first word order might appear to be a confound. Differences in how listeners respond to the two word orders could be either due to

Figure 3.1: Example of a visual display containing one of the quadruplets of words presented to participants in Experiments 1, 2, and 3. In this example, the target word is the past participle *geleund* [leaned] and the competitor is the noun *gleuven* [grooves]. The two other words (*gelift* [lifted] and *glimlachte* [smiled]) serve as distractors and occur as targets and competitors during other trials. (Note that the figure is not drawn to scale.)

the different word orders or due to differences in word duration. But the acoustic differences do not pose a problem for the validity of our argument because they work against our hypothesis. To reiterate, we expect participles in the auxiliary-first word order to be easier to recognize than in the participle-first word order. However, the words in the auxiliary-first word order are more reduced, which means that, on signal-based grounds alone, they should be harder to recognize. Therefore, if we do find the expected effect of word order, it cannot be due to differences in reduction because the effect of reduction goes into the opposite direction to the effect of word order.

Procedure

Participants were seated approximately 70 cm from a 47.5 x 30 cm LCD computer screen running at 120 Hz. Monocular eye movements were recorded with a remote desktop-mounted Eyelink 1000 system (SR-Research, Ottawa, Ontario, Canada) at a sampling rate of 500 Hz. Participants were told that they would hear spoken sentences and see four words on a screen. Their task was to click as quickly as possible on the word that they heard in the spoken sentence. The experiment was preceded by a brief calibration session.

On each trial, participants saw a fixation cross for 500 ms followed by a quadruplet of printed words. All words were presented in lower-case Lucida Sans Typewriter font size 20. The horizontal distance between the centers of the words was 512 pixels and the vertical distance was 385 pixels. An example of the type of visual display that participants saw is depicted in Figure 3.1. The spoken sentences were presented through headphones at a comfortable listening level. The time between the visual onset of the printed words and the onset of the spoken sentence was fixed at 2 seconds. After 72 trials, participants could take a break. A drift check was carried out before the experiment was resumed. The complete experimental session took approximately 20 minutes.

3.2.2 Results

For all of the following analyses, we employed mixed-effects modeling with word order and speaking style as fixed factors and past participle and subject as random factors. Model fitting was performed in a stepwise fashion. To determine the fixed-effects structure of the model, we started by including word order, speaking style and interaction terms and subsequently removed terms if they were not significant. Once the fixed-effects structure was determined, we included random slopes for all significant fixed effects and tested whether the inclusion of a random slope improved the model fit using a chi-square test (Baayen, 2008). If a random slope did not improve the model, we removed it. We estimated p values by using the standard normal distribution (Baayen et al., 2008). For the statistical analyses of the RTs, the data were log-transformed and RTs that differed more than two standard deviations from a given participant's mean were discarded as outliers. During the modeling procedure, data points with standardized residuals larger than 2.5 were removed. For the analysis of the accuracy data, generalized mixed-effects models with a binomial link function were used. In all of the following analyses only significant fixed and random effects are reported.

Accuracy

Trials with RTs smaller than 200 or larger than 4,000 ms, as measured from the onset of the participle, were regarded as extreme values and were not included in the analyses (< 0.8%). Trials on which participants clicked on the word that was actually mentioned in the sentence were scored as correct and trials on which other words were clicked on were scored as incorrect. Table 3.4 shows the mean accuracy values for each of the four experimental conditions. Accuracy was very high (all means are higher than 98%). The statistical analyses suggest that there was a small but significant difference between carefully and casually produced stimuli with the latter ones being responded to slightly less accurately ($\beta_{casual} = -1.57$, $z = -2.23$, $p < .05$; *SD* of random intercepts for participants: 0.65, *SD* of random intercepts for

Experiment	Accuracy		Reaction times	
	Auxiliary first	Participle first	Auxiliary first	Participle first
Experiment 1				
Careful	99.87	99.74	1,331	1,403
Casual	98.82	99.61	1,364	1,466
Experiment 2				
Careful	99.35	99.48	1,252	1,343
Casual	98.48	98.96	1,290	1,391
Experiment 3				
Careful	94.27 (0.39, 5.34)	91.41 (1.3, 7.29)	1,122	1,179
Casual	87.11 (1.95, 10.94)	84.24 (1.95, 13.8)	1,106	1,200

Table 3.4: Accuracy and Reaction Times for Experiments 1, 2, and 3. Accuracy values are given in percentages and reaction times are given in milliseconds. For Experiment 3, the values in parentheses indicate the percentages of incorrectly selected words (first value) and the percentage of trials on which participants did not respond within the time limit.

words: 5.73). There was no effect of word order and no interaction between speaking style and word order.

RTs

Only correct trials were included in the analyses. RTs were measured from target word onset. The average RTs are displayed in Table 3.4. To account for differences in duration among the words, we included past participle duration as a control variable. The analyses showed a main effect of word order ($\beta_{participle\ first}$ = 0.06, t = 6.9, $p < 0.001$; *SD* of random intercepts for participants: 0.11, *SD* of random intercepts for words: 0.07), indicating that participles in sentences with the participle-first word order were responded to more slowly than participles in sentences with the auxiliary-first word order. The same model indicated also a main effect of speaking style (β_{casual} = 0.02, t = 2.1, $p < 0.05$), indicating that RTs for casually produced stimuli were longer than for carefully produced ones. The interaction between word order and speaking style was not significant and neither was the effect of word duration.

Fixed effects	β	t	p
Intercept	−1.30	−13.43	
Word order (participle first)	−0.12	−1.02	> 0.1
Speaking style (casual)	−0.12	−0.99	> 0.1
Word order × Speaking style	0.45	2.64	< 0.01
Random effects		SD	
Participant			
Intercept		0.27	
Word			
Intercept		0.45	
Word order (participle first)		0.60	
Speaking style (participle first)		0.58	
Word order × Speaking style		0.88	

Table 3.5: Linear-mixed effects model for competitor fixations in window 3 of Experiment 1.

Gaze probability

Only correct trials were included in the fixation analyses. Fixations were scored as having landed on a particular word when the fixation fell within a rectangular area of 300 by 200 pixels around the center of that word. We coded whether a fixation fell on a given word on the display for one hundred 10-ms time intervals ranging from 200 ms before the onset of the target word until 1,400 ms later. Fixation proportions are shown in Figures 2A and 2B. Before conducting linear mixed-effects modeling, fixation proportions were transformed to empirical logits (Barr, 2008).

To investigate the time course of the fixation behavior, we tested the effects of word order and speaking style across four time windows. The purpose of the first time window analysis was to determine whether listeners use word order information before the onset of the target word. This window ranged from the average onset of the auxiliary verb in the auxiliary-first condition (159 ms before participle onset), offset by a further 200 ms, until the onset of the following past participle (plus the same 200-ms offset). Note that the same time window was used for sentences with the auxiliary-first and the participle-first word order. The additional 200 ms were added to these and all other window boundaries to account for the fact that it takes approximately this amount of time to program and launch a saccade (Matin et al., 1993). Therefore, Time Window 1 ranged from 41 ms after participle onset until 200 ms after participle onset. If listeners use the information provided by the auxiliary verb to predict the upcoming past participle, we would expect to find an effect of word order in this time window. The second time window ranged from the onset of the past participle (plus 200 ms) until the average offset of the past participles (plus 200 ms). This window thus ranged from 200 ms until 579 ms measured from the average past participle onset. This time window covers the period during which the acoustic information of the past participle unfolds. If speaking style and word order influence the efficiency with which listeners access the past participle while acoustic information becomes available, we would expect to find differences in fixation probability across conditions in this window. The remaining two time windows covered the time period (again with the 200-ms offset) from the average offset of the past participle until the time the average RT was recorded. These time windows were of the same length as the second time window (379 ms). We kept the time windows identical in size to compare the same amount of data (i.e., samples) across analyses. Therefore, the window boundaries are as follows. Time Window 1 ranged from 41 ms after word onset until 200 ms; Window 2 ranged from 200 ms until 579 ms; Window 3, from 579 ms until 958 ms; and Window 4 ranged from 958 until 1337 ms. The vertical lines in Figures 2 and 3 illustrate these time windows.

In Time Window 1, there were no effects of word order or speaking style for either the fixations to the target or those to the competitor. These results indicate that participants were equally likely to look at the past participle and the competitor across all conditions. In Time Window 2, a different pattern of results emerges. For target fixations, we found significantly fewer fixations to the target word for sentences with the participle-first compared with the auxiliary-first word order ($\beta_{participle\ first} = -0.14$, $t = -2.60$, $p < 0.01$; *SD* of random intercepts for words: 0.10). There was no significant difference between casually and carefully articulated words and no interaction. This pattern was also reflected in the competitor fixations. We find significantly more fixations to the competitor word in the participle-first condition than in the auxiliary-first condition ($\beta_{participle\ first} = 0.15$, $t = 2.50$, $p < 0.05$; *SD* of random intercepts for participants: 0.09, *SD* of random intercepts for words: 0.12). There was no effect of speaking style and no interaction.

In Window 3, the identical pattern of results was found. There were fewer looks to the target in the participle-first than the auxiliary-first word order ($\beta_{participle\ first} = -0.15$, $t = -2.50$, $p < .05$; *SD* of random intercepts for words: 0.19), no effect of speaking style, and no interaction. For the competitor, however, we found a significant interaction between word order and speaking style. Table 3.5 summarizes the parameters of the model. An analysis of the simple effects indicated that for carefully produced sentences there was no effect of word order. For casually produced sentences there were more fixations toward the competitor in the participle-first condition than the auxiliary-first condition ($\beta_{participle\ first\,|\,casual} = 0.26$, $t = 2.77$, $p < 0.01$; *SD* of random intercepts for participants: 0.32, *SD* of random intercepts for words: 0.57). For sentences with the auxiliary-first word order, there was no ef-

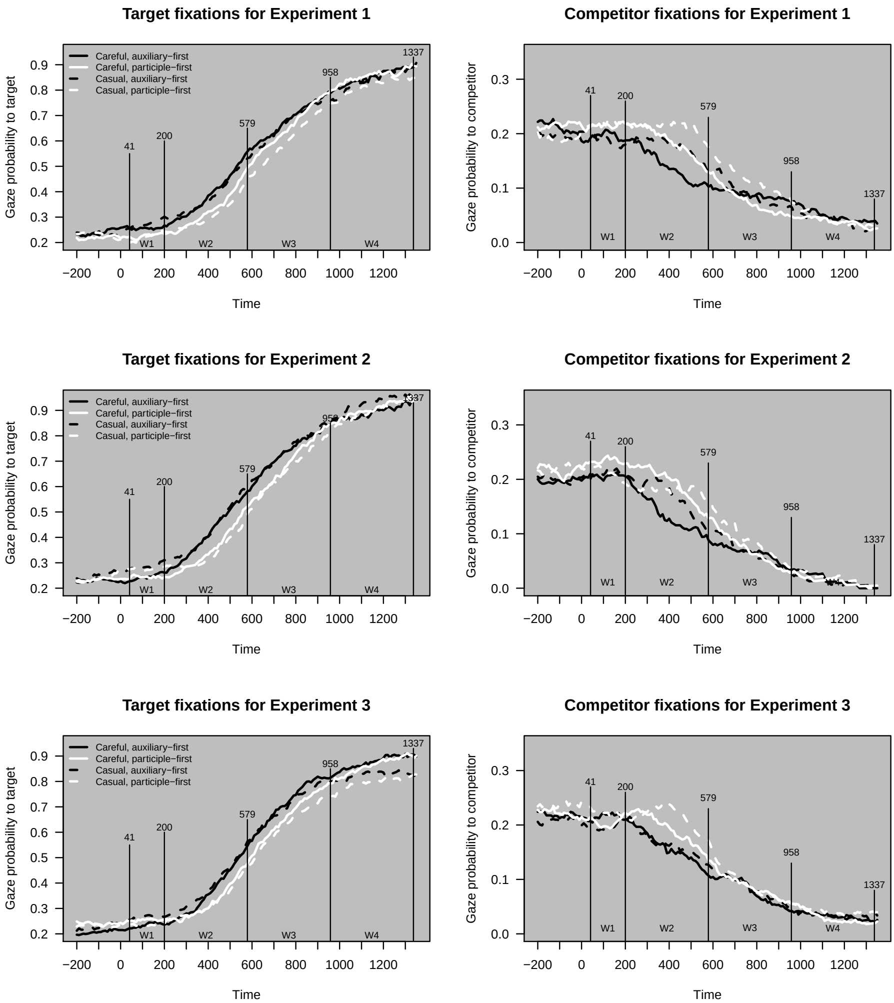


Figure 3.2: Gaze probability over time for target words (i.e., past participles) and competitors in Experiments 1, 2, and 3. Time 0 is aligned to the onset of the target word. An example for a target word is *geleund* (*leaned*) and an example for a competitor is *gleuven* (*grooves*). W1 = analysis Window 1; W2 = analysis Window 2; W3 = analysis Window 3, and W4 = analysis Window 4.

fect of speaking style. For sentences with the participle-first word order, there were more fixations toward the competitor when the past participle was produced casually than when it was produced carefully ($\beta_{casual\,|\,participle\ first}$ = 0.30, t = 2.06, $p < 0.05$; *SD* of random intercepts for participants: 0.35; *SD* of random slopes of the factor speaking style for participants: 0.53; *SD* of random intercepts for words: 0.59; *SD* of random slopes of the factor speaking style for words: 0.66). In Time Window 4 there were no effects of word order or speaking style for either the fixations to the target or those to the competitor.

3.2.3 Discussion

The results of Experiment 1 show that both word order and speaking style can have an influence on how quickly listeners are able to identify words uttered in a sentence context. Participants responded more quickly and were more likely to pay overt visual attention to target words that were syntactically predictable. Similar to the fixations on the target word, fixations on the competitor were influenced by the syntactic predictability of the target word. If the target word was not syntactically predictable, listeners were more likely to consider the competitor as a potential target. In contrast, when the target word was syntactically predictable, listeners were more likely to rule out the competitor because it belonged to a syntactic category that cannot occur after an auxiliary verb.

Furthermore, the analysis of the RTs and the accuracy of the mouse clicks suggest that the speaking style in which the sentences were produced also influenced the listeners' ability to recognize the past participles. In the majority of cases, casually produced target words did not contain a schwa in the prefix, which increased the initial phonological overlap between the targets and the competitors (see Table 3.3). The fact that listeners paid more overt attention to the competitor when hearing casually produced targets reflects that they were sensitive to the increase in phonological overlap. The lack of an interaction between word order and speaking style for target fixations suggests that listeners make use of syntactic information irrespectively of how reliable the acoustic cues in the speech input are. It also shows that the auxiliary verbs were intelligible enough so that listeners could extract syntactic cues from the signal. However, the interaction between word order and speaking style for competitor fixations (in Window 3) showed that participants were more likely to look at the competitor if the participle was less predictable, but only if the sentences were produced in a casual manner. This result supports the hypothesis that syntactic information is more useful if acoustic-phonetic information is less reliable. But the fact that the interaction only emerged for the competitor fixations and not the target fixations is somewhat puzzling. We return to this issue later.

In sum, the results of Experiment 1 support the notion that syntactic predictability can influence word recognition. However, the effect of word order could also be driven by the information given by the following sentence context. It has previously been shown that listeners not only use preceding context to recognize words but also use following context (e.g. Connine et al., 1991; Pollack and Pickett, 1964; van de Ven et al., 2012). If the listeners in our study used the following context to identify the target participle, they could have done so more quickly in the auxiliary-first than the participle-first condition. As in the auxiliary-first word order, the auxiliary verb is positioned to the left of the participle, the distance between the participle and the following noun is exactly one word shorter than in the participle-first condition. It is therefore logically possible that the processing advantage for sentences in which the auxiliary precedes the participle is actually due to the quicker arrival of the following noun. To investigate this possibility we conducted Experiment 2. We presented the stimulus sentences only until the offset of the target word and its associated auxiliary verb (see Table 3.2). If the effect of word order that we found in Experiment 1 is still present in Experiment 2, we can exclude the possibility that it was entirely due to information from the following context becoming available more quickly.

3.3 Experiment 2

3.3.1 Method

Participants

Another 48 native speakers of Dutch were recruited from the subject panel of the Max Planck Institute for Psycholinguistics. Age ranged from 18 to 26 years. The participants reported no hearing problems and had normal or corrected-to-normal vision. They were paid for their participation.

Materials

The materials were based on the stimuli from Experiment 1. New versions were created by removing the portion of the speech signal that followed the target word and its associated auxiliary verb (see Table 3.2). To keep the amount of information carried by the sentence fragments constant across conditions, the auxiliary verb was included for both word order conditions (i.e., even when it followed the participle).

Procedure

The Procedure was the same as in Experiment 1.

3.3.2 Results

Accuracy

The mean values for the four experimental conditions are shown in Table 3.4 was very high (all means are above 98%). There was no effect of speaking style, word order, or an interaction.

RTs

The average RTs are displayed in Table 3.4. As for Experiment 1, we included past participle duration as a control variable to account for differences in word durations. The analyses showed a main effect of word order ($\beta_{participle\ first} = 0.07$, $t = 8.1$, $p < .001$; *SD* of random intercepts for participants: 0.10, *SD* of random intercepts for words: 0.06), indicating that participles in sentences with the participle-first word order were responded to more slowly than participles in sentences with the auxiliary-first word order. There was also a main effect of speaking style ($\beta_{casual} = 0.03$, $t = 2.9$, $p < 0.01$), indicating that RTs for casually produced stimuli were longer than for carefully produced ones. As in Experiment 1, the interaction between word order and speaking style was not significant and neither was the effect of word duration.

Gaze probability

As in Experiment 1, we analyzed the effects of word order and speaking style for each of the four time windows individually. Fixation proportions are shown in Figures 2C and 2D. In Time Window 1, there were no effects of word order or speaking style for the fixations to the target. In contrast, for the fixations to the competitor we found an interaction between word order and speaking style. The parameters of this model are summarized in Table 3.6. An analysis of the simple effects indicated that for carefully produced sentences there was no effect of word order. For casually produced sentences there were more fixations toward the competitor when the participle preceded the auxiliary than when it followed it ($\beta_{participle\ first\,|\,casual} = 0.23$, $t = 2.06$, $p < 0.05$; *SD* of random intercepts for participants: 0.25, *SD* of random intercepts for words: 0.05). There was no effect of speaking style for sentences with the participle-first or the auxiliary-first word order.

In Time Window 2, there were fewer fixations to the target for sentences with the participle-first word order than the auxiliary-first word order ($\beta_{participle\ first} = -0.14$, $t = -2.70$, $p < 0.01$; *SD* of random intercepts for participants: 0.03; *SD* of random intercepts for words: 0.09). There was no significant difference between casually and carefully articulated sentences and no interaction. This pattern was also reflected in the competitor fixations. There were more fixations to the competitor

Fixed effects	β	t	p
Intercept	−1.01	−11.25	
Word order (participle first)	−0.11	−0.98	> 0.1
Speaking style (casual)	−0.16	−1.39	> 0.1
Word order × Speaking style	0.34	2.15	< 0.05
Random effects		SD	
Participant			
Intercept		0.15	
Word			
Intercept		0.25	

Table 3.6: Linear-Mixed Effects Model for Competitor Fixations in Window 1 of Experiment 2.

in sentences with the participle-first word order than the auxiliary-first word order ($\beta_{participle\ first}$ = 0.24, t = 4.57, $p < 0.001$; *SD* of random intercepts for participants: 0.06, *SD* of random intercepts for words: 0.18). There was no effect of speaking style and no interaction.

In Time Window 3, there were also fewer fixations to the target for sentences with the participle-first word order than the auxiliary-first word order ($\beta_{participle\ first}$ = −0.13, t = −2.20, $p < 0.05$; *SD* of random intercepts for participants: 0.05; *SD* of random intercepts for words: 0.10). There was also no effect of speaking style and no interaction. For the competitor, we found no significant effects of word order, speaking style, or the interaction.

In Time Window 4, there were fewer fixations to the target when the past participle was produced casually than when it was produced carefully (β_{casual} = −0.52, t = −3.98, $p < 0.001$; *SD* of random intercepts for participants: 0.57; *SD* of random slopes of the factor speaking style for participants: 0.56; *SD* of random intercepts for words: 0.48; *SD* of random slopes of the factor speaking style for participants: 0.58). For competitor fixations, neither the effect of word order nor of speaking style was significant.

3.3.3 Discussion

The purpose of Experiment 2 was to investigate if the effect of word order that we found in Experiment 1 was due to information that the participants extracted from the following context. The results suggest that this was not the case. Even when the following context was removed, target words embedded in sentences with the participle-first word order were responded to more slowly and fixated less often compared with targets in sentences with the auxiliary-first word order. Similarly, casually produced targets were responded to more slowly compared with targets in carefully produced sentences.

Both Experiments 1 and 2 suggest that listeners are sensitive to both syntactic predictability and changes in phonological overlap as a result of acoustic reductions. Furthermore, the RTs and target fixations suggest that these factors influence word recognition independently from each other. Listeners appear to make use of syntactic information to the same extent when listening to casually produced speech as they do when listening to carefully produced speech. This finding is contrary to the idea that listeners adapt dynamically to the demands of different listening situations (Brouwer et al., 2012b; McQueen and Huettig, 2012). From such a perspective, one might expect that as information from the speech signal becomes less reliable because of phonetic reductions, listeners rely more on other sources of information such as syntactic predictability. The influence of word order should then be stronger for sentences produced in a casually speaking style because under these circumstances phonological information is less reliable. However, as in Experiment 1, we found an interaction between word order and speaking style for the competitor fixations. This interaction suggests that word order information is more useful for casually than for carefully produced sentences. But the fact that we find this interaction only for competitor and not target fixations as well as the fact that it occurs in different time windows across Experiments 1 and 2 remains puzzling. It is possible that this is a spurious effect, a possibility that we investigate in a combined analysis later.

How could the lack of an interaction between speaking style and word order for target fixations and RTs be explained? One potential explanation might be that although the phonetic information in the casual speech signal was less reliable it was nevertheless sufficient when there was enough time for listeners to make a decision. When facing an acoustically ambiguous stimulus, there are (at least) two ways in which one can respond. First, one can wait until more acoustic information becomes available. Second, one can use other sources of information to compensate for the lack of unambiguous acoustic information. As there was no time limit in the previous experiments, participants could take as much time as they needed to process the reduced speech signal. However, when put under time pressure, listeners need to find another way to compensate for the lack of acoustic information. We may therefore find that under time pressure, syntactic predictability has a stronger effect when listening to casual speech than when listening to careful speech. To investigate this hypothesis, we conducted Experiment 3, in which we limited the amount of time listeners had to respond.

3.4 Experiment 3

3.4.1 Method

Participants

Another 48 native speakers of Dutch were recruited from the subject panel of the Max Planck Institute for Psycholinguistics. Age ranged from 18 to 29 years. The participants reported no hearing problems and had normal or corrected-to-normal vision. They were paid for their participation.

Materials

The materials were the same as in Experiment 1.

Procedure

The procedure was identical to that of Experiment 1, with one crucial exception. Participants had only a limited amount of time to respond. Based on informal piloting, we used a time limit of 1,400 ms after the offset of the target word. If participants had not responded within this time period, the trial ended and the words *Te langzaam* (*Too slow*) were displayed in the center of the screen in large red letters. Participants were told about the time limit before the start of the experiment.

3.4.2 Results

Accuracy

For the accuracy analyses, trials during which participants clicked on the wrong word or failed to make a response within the given time limit were scored as incorrect. The mean accuracy values are shown in Table 3.4. The statistical analyses showed that there was a main effect of word order ($\beta_{participle\ first} = -0.33$, $z = -2.73$, $p < 0.01$; *SD* of random intercepts for participants: 0.57, *SD* of random intercepts for words: 0.68), showing that participants were less accurate when responding to sentences with the participle-first word order than vice versa. Furthermore, there was a main effect of speaking style ($\beta_{casual} = -0.85$, $z = -6.82$, $p < 0.001$), indicating that performance was better for items spoken with a careful compared with a casual speaking style. The interaction between word order and speaking style was not significant.

RTs

Average RTs for each experimental condition are listed in Table 3.4. There was a significant main effect of word order, showing that participants took longer to click

Fixed effects	β	t	p
Intercept	−0.84	−12.73	
Word order (participle first)	0.32	3.85	< 0.001
Speaking style (casual)	0.22	2.61	< 0.01
Word order × Speaking style	−0.26	−2.27	< 0.05
Random effects		SD	
Word			
Intercept		0.22	

Table 3.7: Linear-Mixed effects model for competitor fixations in window 2 of Experiment 3.

on the past participle if it preceded the auxiliary verb compared with when it followed it ($\beta_{participle\ first} = 0.07$, $t = 7.7$, $p < 0.001$; *SD* of random intercepts for participants: 0.07, *SD* of random intercepts for words: 0.05). There were no significant effects of speaking style and participle duration. The interaction between word order and speaking style was also not significant.

Gaze probability

As in the previous two experiments, we analyzed the effects of word order and speaking style individually for each of the four time windows. Fixation proportions are shown in Figures 2E and 2F. In Time Window 1, there were no effects of word order or speaking style on the fixations to the target or the competitor. In Time Window 2, there were significantly fewer fixations to the target in sentences with the participle-first word order than the auxiliary-first word order ($\beta_{participle\ first} = -0.20$, $t = -3.83$, $p < 0.001$; *SD* of random intercepts for words: 0.10). There were no effects of speaking style and no interaction. For the fixations to the competitor, we found an interaction between word order and speaking style. The results of this model is summarized in Table 3.7.

An analysis of the simple effects indicated that for carefully produced sentences, there were more fixations toward the competitor if the participle preceded the auxiliary than when it followed it ($\beta_{participle\ first\,|\,careful} = 0.31$, $t = 3.74$, $p < 0.001$; *SD* of random intercepts for words: 0.23). For casually produced sentences there was no effect of word order. For sentences with the auxiliary-first word order, there were more fixations to the competitor if the participle was produced casually than when it was produced carefully ($\beta_{casual\,|\,auxiliary\ first} = 0.23$, $t = 2.81$, $p < 0.01$; *SD* of random intercepts for words: 0.34). For sentences with the participle-first word order there was no effect of speaking style.

In Time Window 3, there were fewer fixations to the target in sentences with the participle-first word order than the auxiliary-first word order ($\beta_{participle\ first} = -0.24$, $t = -4.01$, $p < 0.001$; *SD* of random intercepts for words: 0.12). There were no effects of speaking style and no interaction. For the fixations to the competitor, there were no significant effects either. In Time Window 4, there were no effects of word order or speaking style for the fixations to the target or the competitor.

3.4.3 Discussion

The purpose of Experiment 3 was to investigate how time pressure influences the degree to which listeners make use of syntactic information when confronted with carefully and casually produced speech. Neither the accuracy data, nor the RT data, nor the gaze probability data support the hypothesis that listeners rely more on syntactic information if the speaking style is casual rather than careful. The target fixations and RTs replicated the main effect of word order that we found in the previous two experiments but did not provide any evidence suggesting that word order is more important when listening to casual speech than when listen- ing to careful speech. As in the previous two experiments, there was an interaction between word order and speaking style for the competitor fixations. Whereas in Experiments 1 and 2 this interaction suggested a more important role for syntactic information in casual than in careful speech, the interaction that appeared in Experiment 3 suggested the opposite. Taken together, these findings suggest the possibility that the effect is spurious and not reliable. We therefore compared all three experiments in a combined analysis.

3.5 Comparison of Experiments 1, 2, and 3

3.5.1 Accuracy

The analyses show that participants made fewer correct responses in the participle-first than in the auxiliary-first condition ($\beta_{participle\ first} = -0.27$, $z = -2.35$, $p < 0.05$; *SD* of random intercepts for participants: 0.58, *SD* of random intercepts for words: 0.68; *SD* of random slopes for Experiments 1 vs. 3 for words: 2.31; *SD* of random slopes for Experiments 2 vs. 3 for words: 0.94). Furthermore, they made fewer correct responses when being presented with casual compared with careful stimuli ($\beta_{casual} = -0.83$, $z = -6.88$, $p < 0.001$). The interaction between word order and speaking style was not significant. In addition, accuracy in Experiment 1 was higher than in Experiment 3 ($\beta_{Exp1\ vs.\ Exp3} = 6.02$, $z = 4.81$, $p < 0.001$) and higher in Experiment 2 than in Experiment 3 ($\beta_{Exp2\ vs.\ Exp3} = 3.76$, $z = 7.87$, $p < 0.001$). There was no significant difference between Experiments 1 and 2 and no significant interactions.

3.5.2 RTs

The results of the model fitted to the RTs is shown in Table 3.8. Responses in Experiment 1 and 2 were slower than in Experiment 3. Furthermore, responses in Experiment 2 were faster than in Experiment 1. These differences reflect a speed-accuracy trade-off: while RTs became faster from Experiment 1 to 3, accuracy decreased. Furthermore, there was a significant effect of target word duration: Longer past participles were responded to more slowly than shorter participles. It is important to note that there was a significant two-way interaction between word order and speaking style. An analysis of the simple effects showed that participants responded more slowly to sentences with the participle-first word order than the auxiliary-first word order for both speaking styles. However, as shown by the beta weights, this effect was stronger for casual speech ($\beta_{participle-first\,|\,casual} = 0.08$, $t = 8.8$, $p < 0.001$; *SD* of random intercepts for participants: 0.10, *SD* of random intercepts for words: 0.08) than for careful speech ($\beta_{participle\ first\,|\,careful} = 0.05$, $t = 7.4$, $p < 0.001$; *SD* of random intercepts for participants: 0.10, *SD* of random intercepts for words: 0.06). Furthermore, in sentences with the participle-first word order casually produced past participles were responded to more slowly than carefully produced past participles ($\beta_{casual\,|\,participle\ first} = 0.03$, $t = 4.5$, $p < 0.001$; *SD* of random intercepts for participants: 0.09, *SD* of random intercepts for words: 0.06). In contrast, for the auxiliary-first condition, there was no significant effect of speaking style. There were no two- or three-way interactions of experiment with word order or speaking style.

3.5.3 Gaze probability

The averaged fixation proportions are shown in Figures 3A and 3B. In Time Window 1, there were no effects of word order or speaking style for either target or competitor fixations. In Time Window 2, there were significantly fewer fixations to the target if it occurred in sentences with the participle-first word order compared with the auxiliary-first word order ($\beta_{participle\ first} = 0.16$, $t = -5.33$, $p < 0.001$; *SD* of random intercepts for words: 0.15). There was no significant difference between casually and carefully articulated words and no interaction. Furthermore, there were significantly fewer fixations in Experiment 1 compared with Experiment 3 ($\beta_{Exp1\ vs.\ Exp3} = -0.08$, $t = -2.03$, $p < 0.05$). There were no significant differences in fixation proportions between Experiments 2 and 3 or Experiments 1 and 2 and no interactions. For the fixations to the competitor, we also found a main effect of word order. There were significantly more fixations to the competitor in the participle-first condition than in the auxiliary-first condition ($\beta_{participle\ first} = 0.20$, $t = 5.99$, $p < 0.001$; *SD* of random intercepts for participants: 0.07, *SD* of random intercepts for words: 0.19). Furthermore, there were more competitor fixations when the sentences were pro-

Fixed effects	β	t	p
Intercept	6.89	159.77	
Word order (participle first)	0.06	7.11	< 0.001
Speaking style (casual)	0.03	2, 29	< 0.05
Experiment 1 vs. 3	0.16	7.67	< 0.001
Experiment 2 vs. 3	0.11	5.36	< 0.001
Experiment 2 vs. 1	−0.05	−2.32	< 0.05
Target duration	0.0003	2.77	< 0.01
Word order × Speaking style	0.02	2.14	< 0.05
Random effects		SD	
Participant			
Intercept		0.10	
Word			
Intercept		0.06	
Word order (participle first)		0.02	
Speaking style (casual)		0.04	
Word order × Speaking style		0.02	

Table 3.8: Linear-mixed effects model for reaction times in the combined analysis of Experiments 1, 2, and 3.

duced in a casual speaking style than when they were produced in a careful speaking style ($\beta_{casual} = 0.08$, $t = 2.54$, $p < 0.05$). The interaction between word order and speaking style was not significant and there were no main or interaction effects with experiment.

In Time Window 3, a similar pattern of results emerged. There were significantly fewer fixations to the target if it occurred in sentences with the participle-first word order compared with the auxiliary-first word order ($\beta_{participle\ first} = -0.17$, $t = -5.09$, $p < 0.001$; *SD* of random intercepts for participants: 0.01, *SD* of random intercepts for words: 0.17). There was no significant difference between the careful and casual conditions and no significant interaction between word order and speaking style. Furthermore, there were fewer fixations to the target in Experiment 1 than in Experiment 3 ($\beta_{Exp1\ vs.\ Exp3} = -0.11$, $t = -2.53$, $p < 0.05$) and more in Experiment 2 than in Experiment 1 ($\beta_{Exp2\ vs.\ Exp1} = -0.08$, $t = -1.99$, $p < 0.05$). There was no difference in target fixations between Experiments 2 and 3 and none of the two- or three-way interactions with experiment were significant. The pattern of results for the target fixations matches almost perfectly with the results for the fixations to the competitor. There were more fixations to the competitor for sentences with the participle-first than the auxiliary-first word order ($\beta_{participle\ first} = 0.15$, $t = 2.16$, $p < 0.05$; *SD* of

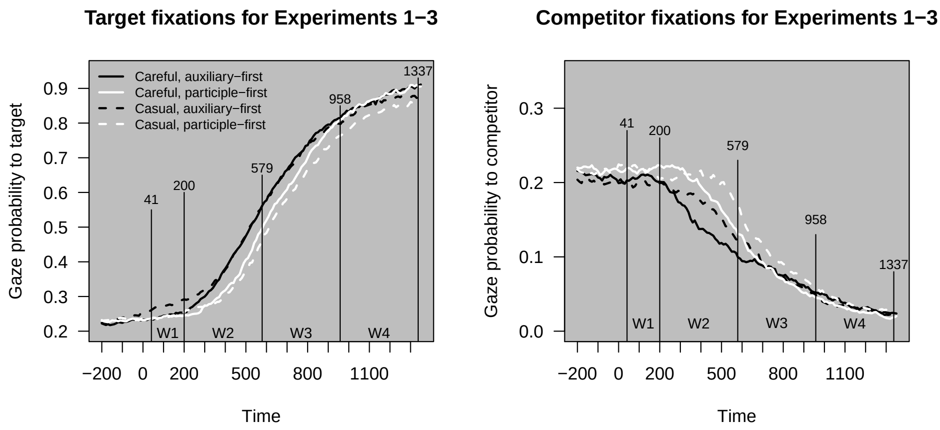


Figure 3.3: Gaze probability over time for target words (i.e., past participles) and competitors collapsed across Experiments 1, 2, and 3. Time 0 is aligned to the onset of the target word. An example for a target word is *geleund* (*leaned*) and an example for a competitor is *gleuven* (*grooves*). W1 = analysis Window 1; W2 = analysis Window 2; W3 = analysis Window 3; and W4 = analysis Window 4.

random intercepts for participants: 0.40, *SD* of random intercepts for words: 0.17; *SD* of random slopes of the factor word order for participants: 0.53; *SD* of random slopes of the factor word order for words: 0.26). There was no difference in the number of competitor fixations between the casual and careful conditions and no significant interaction between speaking style and word order. The same model also indicated that participants were more likely to fixate the competitor in Experiment 1 compared with Experiment 3 ($\beta_{Exp1\ vs.\ 3} = 0.22$, $t = 2.75$, $p < 0.01$) but there were no differences between Experiments 2 and 3 or Experiments 2 and 1. None of the two- or three-way interactions between experiment and speaking style or word order were significant.

In Window 4, the fixations to the target revealed a significant interaction between word order and speaking style. The parameters of this model are summarized in Table 3.9. An analysis of the simple effects showed that for careful sentences, there was no effect of word order. When the sentences were produced casually, there were fewer fixations toward the target word for sentences with the participle-first word order than with the auxiliary-first word order ($\beta_{participle\ first\ |\ casual} = -0.19$, $t = -2.02$, $p < 0.05$; *SD* of random intercepts for participants: 0.47; *SD* of random intercepts for words: 0.22; *SD* of random slopes of the factor word order for words: 0.47).

Furthermore, for sentences with the auxiliary-first word order, there was no effect of speaking style. In contrast, for sentences with the participle-first word order, there were fewer target fixations for casually compared with carefully produced sentences ($\beta_{casual\,|\,participle-first} = -0.42$, $t = -3.65$, $p < 0.001$; *SD* of random intercepts for participants: 0.71; *SD* of random slopes of the factor speaking style for participants: 0.79; *SD* of random intercepts for words: 0.47; *SD* of random slopes of the factor speaking style for words: 0.58).

For the fixations to the competitor, there was no effect of word order or speaking style, and no interaction. However, there was a main effect of experiment, showing that there were more competitor fixations in Experiment 1 than in Experiment 3 ($\beta_{Exp1\ vs.\ Exp3} = 0.65$, $t = 4.85$, $p < 0.001$; *SD* of random intercepts for participants: 0.60, *SD* of random intercepts for words: 0.60) and more fixations to the competitor in Experiment 2 than in Experiment 3 ($\beta_{Exp2\,vs.\,Exp3} = 0.50$, $t = 3.67$, $p < 0.001$).

3.5.4 Control analyses

It remains possible that the effect of word order could be due to other factors: the duration of the sentence after the participle, the participants' preference for one order over the other, or bigram frequency. More specifically, bigram frequency refers here to the frequency with which a given past participle occurs together with its preceding word (be that an auxiliary verb or other types of words). We tested whether the effect of word order remained statistically significant after controlling for these factors by conducting additional analyses for each dependent measure (i.e., accuracy, RT, gaze probability) based on the combined data from all three experiments. We defined sentence remainder duration as the time from the offset of the past participle until the end on the sentence. We used the data from the rating experiment (mean ratings per item) as estimates of preference for one or the other word order for each sentence. For bigram frequency, we used estimates based on the Dutch Internet search engine IxQuick (`https://www.ixquick.com`). The log-transformed bigram frequencies ranged from 1.10 to 11.74 with a mean of 5.89 and a median of 5.37. For each analysis, we first fitted a control model using linear regression in which we regressed the respective dependent variable on the control variables. Subsequently, we used the residuals of this model as the dependent variable in the original models. Correlations between the experimental and control variables are shown in Table 3.10.

For the accuracy scores, the control model suggests that the higher the bigram frequency was, the more accurately participants responded ($\beta_{bigram} = 0.08$, $z = 4.0$, $p < 0.001$). The effect of word order, which we had found earlier, disappeared. For the RTs, the control model indicated increasing RTs with increasing sentence remainder durations ($\beta_{remainder} < 0.001$, $t = 6.12$, $p < 0.001$). Furthermore, RTs were slower for sentences with higher ratings ($\beta_{rating} = 0.01$, $t = 2.14$, $p < 0.05$)

Fixed effects	β	t	p
Intercept	1.72	14.02	
Word order (participle first)	0.05	0.80	> 0.1
Speaking style (casual)	0.06	0.39	> 0.1
Experiment 1 vs. 3	−0.37	−2.38	< 0.05
Experiment 2 vs. 3	−0.20	−1.27	> 0.1
Experiment 2 vs. 1	0.17	1.15	> 0.1
Word order × Speaking style	−0.27	−2.01	< 0.05
Experiment 1 vs. 3 × Speaking style	−0.14	−0.80	> 0.1
Experiment 2 vs. 3 × Speaking style	−0.42	−2.38	< 0.05
Experiment 2 vs. 1 × Speaking style	−0.28	−1.70	> 0.05
Random effects		SD	
Participant			
Intercept		0.62	
Word order (participle first)		0.64	
Speaking style (casual)		0.75	
Word order × speaking style		0.84	
Word			
Intercept		0.38	
Word order (participle first)		0.49	
Speaking style (casual)		0.36	
Experiment 1 vs 3		0.48	
Experiment 2 vs 3		0.55	
Experiment 2 vs 1		0.30	
Word order × Speaking style		0.56	

Table 3.9: Linear-mixed effects model for target fixations in window 4 of the combined analysis of Experiments 1, 2, and 3. During the model fitting procedure, it was not possible to include both random slopes for both interaction terms without R returning an error message. We chose to report the model with random slopes for the interaction between word order and speaking style because our primary interest is in this interaction rather than the interaction between experiment and speaking style. Furthermore, the Akaike information criterion value for the model with random slopes for the interaction between speaking style and word order is lower than for the other model (33,522 vs. 33,602), indicating a better model fit.

Dependent measure	Sentence remainder	Ratings	Bigram frequency	Word order	Speaking style
Accuracy	-0.04	-0.03	0.18	-0.10	-0.29
RT	0.24	0.10	-0.29	0.38	0.13
Target W1	-0.11	-0.07	0.08	-0.12	0.11
Target W2	-0.18	0.01	0.19	-0.30	-0.01
Target W3	-0.17	0.02	0.17	-0.28	-0.12
Target W4	-0.05	0.03	0.11	-0.06	-0.19

Table 3.10: Correlations between the main dependent measures and the control variables (RT = reaction times, Target = target gaze probability, W = window). Bigram frequencies are based on the Dutch Internet search engine IxQuick. Higher ratings reflect a preference for the participle-first word order. Word order is coded as follows: auxiliary first = 0, participle first = 1. Speaking style is coded with careful = 0 and casual = 1.

reflecting that responses were slower when the preference was for the participle-first word order. In addition, participants were faster for stimuli with higher bigram frequencies than lower bigram frequencies ($\beta_{bigram} = -0.01$, $t = -7.29$, $p < 0.001$). It is important to note that despite the significant influence of the control variables, the interaction effect between word order and speaking style remained significant ($\beta_{participle\ first \times casual} = 0.03$, $t = 2.29$, $p < 0.05$; *SD* of random intercepts for participants: 0.1; *SD* of random intercepts for words: 0.06; *SD* of random slopes of the interaction between the factors speaking style and word order for words: 0.02).

For gaze probability, we analyzed Time Windows 2, 3, and 4 because these are the windows in which we had found an effect of word order in the previous analyses. In Windows 2 and 3, the control models indicated a lower gaze probability for longer sentence remainders than for shorter ones (Window 2: $\beta_{remainder} < -0.001$, $t = -4.15$, $p < 0.001$; Window 3: $\beta_{remainder} = -0.001$, $t = -4.06$, $p < 0.001$) and a higher gaze probability for higher bigram frequencies than lower ones (Window 2: $\beta_{bigram} = 0.05$, $t = 4.35$, $p < 0.001$; Window 3: $\beta_{bigram} = 0.04$, $t = 3.72$ $p < 0.001$). It is important to note that when taking the control variables into account, the effect of word order disappeared. However, finding an effect of sentence remainder duration in the early time window is quite puzzling because participants ought not to be able to anticipate the end of the sentence so early. We therefore investigated if the effect of sentence remainder duration is actually due to its correlation with word order ($r = 0.27$, $t = 4.49$, $p < 0.001$). When entering sentence remainder and word order simultaneously into the original model for Time Window 2, sentence remainder is not significant whereas word order is ($\beta_{participle\ first} = -0.15$, $t = -4.43$, $p < 0.001$;

SD of random intercepts for words: 0.15). We therefore fitted a second control model without sentence remainder duration (i.e., only bigram frequency and ratings). For these residuals, word order was still a significant predictor ($\beta_{participle\ first} = -0.08$, $t = -2.49$, $p < 0.05$; *SD* of random intercepts for participants: < 0.001; *SD* of random intercepts for words: 0.18). An analysis for Time Window 3 yielded the same pattern of results. When entering sentence remainder and word order simultaneously into the original LMER, sentence remainder was not significant whereas word order still was ($\beta_{participle\ first} = -0.17$, $t = -4.58$, $p < 0.001$; *SD* of random intercepts for participants: 0.01; *SD* of random intercepts for words: 0.17). As for Time Window 2, for the residuals of a control model without sentence duration remainder, word order was a significant predictor ($\beta_{participle\ first} = -0.10$, $t = -2.90$, $p < 0.01$; *SD* of random intercepts for participants: < 0.001; *SD* of random intercepts for words: 0.2). These results suggest that word order exerted a significant influence on gaze probability above and beyond sentence duration remainder and bigram frequency. In Window 4, the control models did not show effects of sentence remainder, rating, or bigram frequency. It is not surprising that when entering the residuals of the control model into the original LMER, the critical interaction between word order and speaking style remained significant, as shown in Table 3.11.

Taken together, the control analyses suggest that the word order effects remain significant after controlling for potential effects of the duration of the sentence remainder, the preference ratings, and bigram frequency. This does not mean, of course, that bigram frequency does not influence predictive language processing. However, it appears that the present effect of word order cannot be reduced to an effect of bigram frequency.

3.5.5 Discussion

The comparison of the three experiments confirmed that listeners performed better at recognizing the target past participle when it occurred in sentences with the auxiliary-first word order than in sentences with the participle-first word order. Furthermore, the analysis of the different time windows showed that the influence of word order occurred after the onset of the past participle. The fact that the effect did not appear before the onset of the past participle (i.e., in Time Window 1) suggests that listeners did not start to predict the past participle as soon as they heard the auxiliary and instead benefited from word order information only as the target word started to unfold. Furthermore, the analysis of the combined data from all three experiments revealed a finding that did not come out in the individual analyses of the experiments. Both the RTs and the late target fixations showed an interaction between word order and speaking style, suggesting that the effect of word order was stronger for casual speech than for careful speech. In addition, the interaction effects

Fixed effects	β	t	p
Intercept	−1.02	−8.46	
Word order (participle first)	0.12	1.14	> 0.1
Speaking style (casual)	0.08	0.50	> 0.1
Experiment 1 vs. 3	−0.38	−2.58	< 0.01
Experiment 2 vs. 3	−0.22	−1.51	> 0.1
Experiment 2 vs. 1	0.15	1.08	> 0.1
Word order × Speaking style	−0.28	−2.12	< 0.05
Experiment 1 vs. 3 × Speaking style	−0.17	−0.97	> 0.1
Experiment 2 vs. 3 × Speaking style	−0.43	−2.41	< 0.05
Experiment 2 vs. 1 × Speaking style	−0.26	−1.53	> 0.1
Random effects		SD	
Participant			
Intercept		0.64	
Word order (participle first)		0.64	
Speaking style (casual)		0.77	
Word order × speaking style		0.86	
Word			
Intercept		0.33	
Word order (participle first)		0.48	
Speaking style (casual)		0.37	
Word order × Speaking style		0.53	

Table 3.11: Control model for target fixations in window 4 of the combined analysis of Experiments 1, 2, and 3.

between word order and speaking style for the competitor fixations that we found in the individual analyses of the experiments disappeared in the overall analyses. This suggests that it is not reliable and that the effect is probably a Type I error. In contrast, the interaction between word order and speaking style for the target fixations seems to be robust because it occurs in both the late gaze probabilities as well as the RTs. As the interaction effect emerges only in the combined analysis, it appears to be rather small and requires a relatively large amount of data to be detected. The interaction suggests that listeners rely more on syntactic information when the speech input is produced in a casual manner than when it is produced carefully. The fact that we observed the interaction in the RTs and the late time window but not in the earlier time windows suggests that the increased benefit of syntactic information for casual speech manifests itself relatively late in lexical processing.

Furthermore, the comparison of the three experiments sheds light on the question of whether the effect of word order might be at least partially due to the quicker arrival of information from the following context. If so, word order should have had a weaker effect in Experiment 2 compared with Experiments 1 and 3. However, our analyses indicate no interaction between word order and experiment, suggesting that the word order effect is not dependent on information from the following context.

3.6 General Discussion

This study investigated how syntactic predictability influences the recognition of words embedded in carefully and casually spoken sentence contexts. We manipulated syntactic predictability by swapping the order of past participles and their associated auxiliary verbs in Dutch subordinate clauses, where word order is not fixed. The participle is more predictable in the auxiliary-first word order than in the participle-first word order because the auxiliary indicates that a participle must follow immediately. Moreover, we explored whether the influence of syntactic predictability differs depending on whether the sentences were produced in a careful or a casual manner. In a casual speaking style, words typically undergo acoustic reductions, which potentially decreases the amount of information conveyed by the speech signal. We hypothesized that there are two possible ways in which a reduced speaking style could influence the way in which word order information is used. First, listeners might make more use of syntactic cues to compensate for the decrease in acoustic-phonetic information that results from a casual speaking style. Alternatively, listeners may be less able to benefit from syntactic information because it is more difficult to extract it from an acoustically reduced speech input.

We conducted three eye-tracking experiments using a printed-word variant of the visual world paradigm (e.g. McQueen and Viebahn, 2007), in which careful and casual variants of sentences with either the auxiliary-first or the participle-first word order were presented while listeners had to identify the participle on a screen. In addition to the target participle, the display contained a phonological competitor for which phonological overlap with the target word was larger when the target was produced in a reduced way compared with an unreduced way. In Experiment 1, complete sentences were presented and the participants were under no time pressure to respond. In Experiment 2, the sentence context following the target word and its associated auxiliary verb was removed to control for following semantic context effects. Experiment 3 was like Experiment 1 but a time limit was imposed on the responses to put participants under time pressure.

Our first finding is that in all three experiments, participles were recognized more easily when they followed their associated auxiliary verbs compared with when they preceded them. This result provides further evidence for the hypothesis that syntactic predictability can influence the speed with which listeners recognize words (Arai and Keller, 2013; Kamide et al., 2003). Our results extend these findings by showing that listeners can use auxiliary verbs to facilitate the recognition of following past participles. Auxiliary verbs do not contain any semantic information; they merely signal that a participle is more likely to come up than a nonparticiple. In other words, they provide information about the word class that the following word is likely to belong to.

It is interesting that the analyses of the separate time windows suggest that although syntactic predictability facilitated the recognition of the past participle, listeners did not actually predict its occurrence. That is, participants did not start looking at the target word before its acoustic onset. Predictability means that a certain input is, in principle, predictable based on previously occurring information. The predicting information has the potential to be used. However, just because there is information that could be used to predict upcoming input does not necessarily mean that it will or can be used by the listener. Given that previous research has shown that listeners are able to predict upcoming words (e.g. Altmann and Kamide, 1999; Arai and Keller, 2013; Kamide et al., 2003), the question arises as to why participants were not predicting the target word in the present study. One difference between our study and previous work is the time that listeners had to develop a prediction about the upcoming input. In our study, there were no words between the predictive word (i.e., the auxiliary verb) and the target word (i.e., the past participle). In previous studies that found target fixations before the target's acoustic onset, there was at least one word in between the predictive precursor and the target. For example, Altmann and Kamide (1999) presented sentences such as *The boy will eat the cake* in which the word *eat* is predictive of the word *cake*. Thus, in their study, the two words do not follow each other but are separated by the definite article *the*. It is possible that this intervening word was long enough to allow the listeners to predict the upcoming noun. In Dahan and Tanenhaus (2004), the two critical words were also separated by an article. For example, in the sentence *Never before climbed a goat so high*, the predictive verb *climbed* and the target word *goat* are separated by the indefinite article *a*. The results of our study suggest that at least one intermediate word is necessary for predictive eye movements to occur. Unfortunately, this hypothesis cannot be tested with these materials because the syntactic structure that we used does not allow for the insertion of words between auxiliary verbs and past participles.

Another reason participants did not execute anticipatory eye movements in our study could be related to the type of visual display that we used. Finding the target word among the distractors was a quite difficult task. Whereas previous studies investigating predictive processing used mostly displays containing pictures (e.g. Altmann and Kamide, 1999; Arai and Keller, 2013; Dahan and Tanenhaus, 2004), we used printed words. Printed words look more similar to each other than the pictures of existing objects used in the other studies, which makes them more difficult to distinguish from one another. Furthermore, the words that we presented on the screen were spelled in a similar way (they all start with <gel>, <gl>, <ger>, or <gr>), making it even more difficult to find the target word among the distractors. The difficulty of finding the target object on the visual display may have delayed the execution of eye movements to the target. Participants might have needed more time to find the target or they might have been more conservative in their search behavior and waited for more information to appear before starting to look for the target.

The fact that syntactic predictability also improved the recognition of casually produced words contributes to the understanding of how listeners process conversational speech, which often contains acoustically reduced word forms. Previous studies have demonstrated that listeners have difficulty recognizing reduced word forms when these are presented in isolation. Listeners' performance on reduced words when they are embedded in context is almost as good as when recognizing unreduced word forms (Ernestus et al., 2002). This is consistent with the observation that naive listeners are hardly aware of the presence of acoustic reductions (Kemps et al., 2004) despite the fact that they are ubiquitous in everyday conversations (e.g. Ernestus, 2000; Johnson, 2004). There are several contextual cues that have been proposed to facilitate the recognition of reduced word forms. Among these are acoustic (Janse and Ernestus, 2011), semantic (van de Ven et al., 2012), and discourse-based (Brouwer et al., 2013) information. Our results suggest that syntactic information can also be added to this list of cues: Reduced words that are syntactically more predictable are easier to recognize.

Finally, the effect of syntactic predictability also points to the significance of word order variability in Dutch subordinate clauses. Previous studies have investigated the circumstances under which speakers of Dutch prefer either the auxiliary-first or the participle-first word order (e.g. Swerts and van Wijk, 2005). Our study has demonstrated that this choice has consequences for the listener. The auxiliary-first word order leads to faster recognition of the past participle. Crucially, this advantage in recognition performance does not influence the preference for one word order over the other, as was shown in our rating study. Apparently, factors other than the listener's ease of recognition are more important when it comes to the usage of syntactic structures.

The second finding of our study is that participles were recognized more easily when they were produced carefully rather than casually. This result is consistent with previous findings reported by Brouwer et al. (2012b) who showed that massive acoustic reductions that increase the acoustic similarity among words can change the pattern of lexical competition. Our results extend Brouwer et al. (2012b)'s findings by showing that lexical processing is also influenced by a relatively mild form of acoustic reduction (i.e., schwa reduction).

Our third finding is that the effect of syntactic predictability was stronger when the participle was produced in an acoustically reduced way compared with when it was produced in a careful way. This suggests that listeners make more use of syntactic information when acoustic cues are less reliable. The fact that this interaction effect emerged in the RT data and the late time window of the eye-tracking data suggests that it is due to processes that take time to act. At first, acoustic reduction and syntactic predictability influence the processing of the target participle independently from each other. Listeners then appear to recover from acoustic reduction more quickly when the participle is more predictable. This suggests that both syntactic and phonological information are processed in parallel and are integrated not immediately but after a short processing delay. Note that this does not mean that syntactic and phonological processing are themselves delayed (as, indeed, the main effects in the RTs and fixation analyses show), only that their integration takes some time.

The finding that the integration of syntactic and phonological information occurs relatively late does not contradict previous research that has found an early influence of syntax on phonetic processing (e.g. van Alphen and McQueen, 2001). Syntactic information may influence language processing in different ways. van Alphen and McQueen (2001) investigated the effects of syntactic information on phonemic decision-making. This type of process does not necessarily tap into the same kinds of representations and processes that are involved in online word recognition (which is what our study is focusing on). In fact, van Alphen and McQueen assume that phonemic decision making is located in a postlexical phonemic decision module, as proposed in the Merge model (Norris et al., 2000). Furthermore, their design involved multiple repetitions of a small number of highly similar sentence frames. In contrast, in our study, listeners were exposed to a different sentence on each trial making the experimental stimuli much more variable. As the experimental situations between van Alphen and McQueen's study and our own differ quite substantially, comparisons referring to the time course of effects are hard to make. It is, thus, very well possible that knowledge about syntactic structures can bias early judgments about phonological categories, on the one hand, and influence online lexical processing at later stages, on the other hand. The fact that we observe the interaction effect later than the main effects suggests that phonological and syntactic information are processed on independent pathways during early stages of processing (e.g., lexical access). During later stages in which lexical units are integrated into larger structures such as sentences and discourse contexts, both streams are combined.

The fact that the interaction of speaking style and word order was present only in the overall analysis in which we combined the data from the three experiments suggests that it is a relatively small effect. One explanation for why the interaction effect is small is that listeners exploit syntactic cues as much as possible even if the acoustic cues are perfectly reliable. The effect of word order has therefore almost no room to increase when the speech input becomes less reliable. A second possibility is that the acoustic reductions induced by the casual speaking style were not severe enough for listeners to need to rely on the word order cue. Although our results show that the reduced speaking style did decrease recognition performance, this effect may simply not have been strong enough for the listeners to substantially change the way in which they weigh acoustic and syntactic sources of information. A third reason for the small magnitude of the interaction effect may be that listeners could not make much more use of syntactic predictability in the casual speech condition because the auxiliary verbs themselves were acoustically reduced, which made it harder for listeners to process them. Although on the one hand listeners may want to rely more on syntactic information when acoustic cues are less reliable, on the other hand, accessing syntactic information may be harder because the words providing that type of information are themselves acoustically reduced.

So far, no theory of spoken word recognition has explicitly implemented a mechanism that could explain the influence of syntactic predictability on spoken word recognition. However, several theories could be extended to account for the present results. There are at least two possible mechanisms. First, listeners might access representations that contain information about the syntactic category that a word belongs to and use this knowledge to predict upcoming word classes. Depending on which word class is likely to follow (e.g., after an auxiliary verb), words that belong to this word class will be favored whereas words that do not belong to this word class are less likely to be considered. A second mechanism is based on bigram frequencies: Instead of accessing knowledge about word class, listeners might predict upcoming words based on how often two words have occurred together in the past. Our control analyses suggest that bigram frequencies could explain part of the effect of word order that we found but not all of it. This suggests that the word order effect that we found is at least to some extent based on abstract knowledge about the relationships among syntactic categories.

Activation-based models such as TRACE (McClelland and Elman, 1986; Mirman et al., 2006) could learn about word-order-based predictability by encoding the bigram frequencies among words in connection weights. After these connection weights have been established, auxiliary verbs that have become activated would send activation to associated participles and consequently boost their activation levels. More plausibly, given the current results, a connectionist model could include a layer of processing units that represent syntactic classes (e.g., auxiliary verbs, past participles). Activated auxiliary verb units would send activation to past participle nodes, which in turn would activate word forms that belong to the past participle class. These past participles would become preactivated, which would in turn facilitate their recognition. A different framework in which effects of syntactic predictability could be implemented is offered by Bayesian models, such as Shortlist B (Norris and McQueen, 2008). In this model, the recognition of words is based on probabilities rather than levels of activation. For the calculation of the (posterior) probability with which a word is recognized, the model combines the word's prior probability with the probability of the acoustic signal, given that the word was uttered. Shortlist B allows word priors to be influenced by several factors, such as lexical frequency, and could in principle include effects of semantic and syntactic context. According to this account, syntactic predictability could increase the posterior probability of a word by increasing the word's prior probability. Both types of model could be adapted to explain effects of syntactic predictability. However, neither of them specifically predicts that the integration of syntactic and phonological information occurs relatively late. Both types of models would need to address this result.

In conclusion, the results of the present study provide further evidence for the hypothesis that syntactic information facilitates the recognition of words in sentential contexts. Syntactic information becomes even more useful when acoustic cues are less reliable, as when listening to casual speech, suggesting that syntactic context provides useful cues that can help listeners to cope with speech reductions in conversational speech. This supports the notion that listeners dynamically adapt to the different sources of linguistic information that are available to them.

Inflectional schwa in the processing in casual speech

Chapter 4

This chapter is based on:
Malte Viebahn, Mirjam Ernestus and James McQueen (under revision).
Absence of inflectional schwa does not disrupt syntactic processing in casual speech.

This electrophysiological study asked whether listeners tolerate grammatical gender violations in casual speech. Native speakers of Dutch were presented with utterances that contained adjective-noun pairs in which the adjective was either correctly inflected with a word-final schwa (e.g. *een spannende roman* "a suspenseful novel") or incorrectly uninflected without that schwa (*een spannend roman*). Consistent with previous findings, the uninflected adjectives elicited a P600 effect compared to the inflected adjectives when the talker was speaking in a careful manner. When the talker was speaking in a casual manner, this P600 effect was absent. A control condition showed N400 effects for carefully as well as casually produced utterances with semantic anomalies, showing that listeners were able to understand the content of both types of utterance. The results suggest that listeners take information about the speaking style of a talker into account when processing the acoustic-phonetic information provided by the speech signal. This contextually-driven flexibility has consequences for syntactic processing: Absent schwas in casual speech are effectively not grammatical gender violations.

4.1 Introduction

Spoken language is characterized by an extraordinary amount of variability. The type of variability investigated in the present study is determined by a speaker's speech register or speaking style. In spontaneous speech, utterances are often produced in an acoustically reduced manner. Speakers tend to slur, shorten, and omit individual segments and even whole syllables (e.g. Ernestus, 2000; Johnson, 2004). How do listeners cope with this variability? In the following we investigate whether listeners use acoustic-phonetic information provided by the speaking style in order to adapt the way in which they recognize speech. Critically, we examine whether this kind of adaptation can have consequences for syntactic processing.

The notion that listeners adapt to acoustic-phonetic reductions in casual speech is not new. Previous studies have shown that how well listeners can recognize reduced words depends on the surrounding context (e.g. Janse and Ernestus, 2011; van de Ven et al., 2011). This suggests that listeners take contextual information into account when processing reduced word forms. Furthermore, it has been proposed that being exposed to casual speech influences how the word recognition system operates. For example, a visual-world eye-tracking study (Brouwer et al., 2012b) suggests that listening to casual speech changes the dynamics of lexical competition during word recognition. Participants listened to sentences extracted from a spontaneous speech corpus and saw four printed words: a target (e.g., *computer*, with the reduced form *puter*), a competitor similar to the canonical form (e.g., *companion*), a competitor similar to the reduced form (e.g., *pupil*), and an unrelated distractor. Consistent with previous visual-world studies with careful speech (e.g. Allopenna et al., 1998; McQueen and Viebahn, 2007), Brouwer et al. found clear lexical competition effects for phonologically overlapping words when only carefully produced target words were presented (Experiment 2). However, when carefully and casually produced word forms were presented in the same experiment, lexical competition was weaker and less influenced by the phonological overlap between the target and the competitor (Experiments 1 and 3). These results suggest that listeners adapt to casual speech by penalizing acoustic mismatches less strongly than when listening to careful speech. In another study showing adaptation to casual speech (Poellmann et al., 2014), Dutch listeners were exposed to segmental and syllabic reductions during a learning phase. In the subsequent test phase, participants heard both kinds of reductions, but they were applied to words that had not been heard during the previous phase. The results indicated that learning about reductions was applied to previously unheard words, demonstrating that listeners can adapt to acoustic-phonetic reductions. Further evidence for adaptation comes from a shadowing study (Brouwer et al., 2010) showing that hearing reduced speech increases the probability of producing reduced word forms.

The event-related potential (ERP) experiment presented in this paper extends these previous studies on casual speech by asking whether adaptation to acoustic-phonetic reductions can have consequences for how listeners process syntactic markers. More specifically, we investigate how the reduction of inflectional schwa in Dutch (spelled as the letter <e>) influences the interpretation of the resulting ungrammatical forms of adjectives. The schwa functions as an inflectional marker at the end of adjectives indicating the grammatical gender of the following noun. There are two different grammatical genders in Dutch: a common gender and a neuter gender. Common gender is associated with the inflected form of the adjective (e.g. *een spannende roman* "a suspenseful novel"). In contrast, neuter gender is associated with the uninflected form of the adjective and does not end in schwa (e.g. *een spannend verhaal* "a suspenseful story"). ERP studies in Dutch investigating gender violations such as incorrectly inflected adjectives have revealed a clear P600/SPS (syntactic positive shift) effect for the following noun (Hagoort and Brown, 1999; van Berkum et al., 2005). This late positive deflection is associated with syntactic processing and is assumed to be an indication for syntactic parsing problems and repair processes (e.g. Friederici et al., 1993; Hagoort et al., 1993). There is a discussion in the literature about the exact cognitive and neural processes that underlie the P600 effect (compare e.g. Coulson et al., 1998; Hahne and Friederici, 1999; Osterhout and Hagoort, 1999). However, there is general agreement that the P600 is associated with difficulties in syntactic processing (Gouvea et al., 2010).

In casual Dutch, the vowel schwa is often either shortened in its duration or completely absent (e.g. Pluymaekers et al., 2005b; Van Bergem, 1994). This raises the question of how listeners interpret absent schwas that, if present, would function as grammatical markers. If listeners adapt to the acoustic-phonetic reductions that occur typically in casual speech, common-gender adjectives that are produced without the word-final schwa should not be interpreted as ungrammatical. Instead, listeners should take information about the speaking style of the talker into account while listening and assume that upcoming words may be produced in a reduced manner. This adaptation ought to have consequences for syntactic processing and increase tolerance for grammatically inappropriate forms that may result from common reduction processes.

A study by Hanulíková et al. (2012) showed that listeners tolerate ungrammatical forms if spoken by a talker with a foreign accent. ERPs to gender agreement errors in sentences spoken by a native speaker were compared with ERPs to the same errors spoken by a non-native speaker. Gender violations in native speech resulted in a larger P600 compared to correct sentences, indicating that the listeners were sensitive to the grammatical errors. In contrast, when the same violations were produced by the non-native speaker with a foreign accent, no P600 effect was observed. These results demonstrate that listeners take knowledge about speaker identity into account when interpreting the acoustic-phonetic characteristics of the linguistic input.

In the present study, we used a similar design to Hanulíková et al. and applied it to the domain of casual speech. Are listeners more tolerant of the absence of inflectional schwas when hearing a casual speech style compared to a careful speech style? We expect the absence of a P600 effect for casual speech but the presence of a P600 effect for careful speech. In order to rule out the possibility that the absence of a P600 effect is due to shallow processing or that listeners were unable to understand the content of the casually produced utterances, we added a control condition in which listeners were exposed to semantically anomalous utterances (spoken once again in either a careful or a casual style). Such stimuli have been shown to elicit a negative deflection around 400 milliseconds (labeled N400) after the onset of the anomalous word. For example, the word *dog* in the sentence *I take coffee with cream and dog* elicits an N400 effect compared to the word *sugar* presented in the same sentence (for a review see Kutas and Federmeier, 2011). We predict that if the expected absence of a P600 effect in casual speech is due to adaptation to speaking style, that adaptation should have no influence on how listeners respond to semantic anomalies and hence that N400 effects ought to be the same for casually and carefully produced utterances.

4.2 Corpus study

There is currently a debate about the role of morphology in acoustic-phonetic reduction processes (e.g. Bürki et al., 2011; Hanique and Ernestus, 2012; Hay, 2003; Plag et al., 2015; Schuppler et al., 2012). The extent and the way in which phonological segments are influenced by reduction processes may depend on their morphological status and the morphological structure of the words in which they occur. It is therefore not certain that what is known about schwa reduction in general also holds for cases in which schwas constitute inflectional affixes. In order to investigate whether inflectional schwas may in fact be absent in spontaneous speech, we conducted a corpus study based on the Ernestus Corpus of Spoken Dutch (Ernestus, 2000), which contains recordings of conversational speech, and the interview-speech component of the Spoken Dutch Corpus (Oostdijk, 2002). A further goal of this analysis was to determine whether there are segmental or lexical constraints on when schwa reduction occurs. For example, it is possible that schwas are only absent in particular segmental contexts or only in words with a high frequency of occurrence.

Supported by an automatic speech transcription algorithm based on the Hidden Markov Model Toolkit (Young et al., 2002), we selected 3,753 tokens of common-gender adjectives that had been produced without the grammatically prescribed schwa. In addition, another set of common-gender adjectives that had been produced with schwa was selected. The latter set consisted of 5,496 tokens. A check of the data revealed that the automatic transcriptions were more reliable for adjectives transcribed with schwa than for those transcribed without schwa. We therefore manually double-checked the adjectives that had been transcribed without schwa. As this was a time-consuming procedure, only a sample of 215 tokens was analyzed. This analysis revealed that 58 of these tokens had been produced without inflectional schwa. Without a large set of manually transcribed tokens it is impossible to give an estimate of the rate at which inflectional schwa is absent in spontaneous speech. The question here, however, was simply whether or not inflectional schwa *can* be absent. Based on the collected data, it is clear that schwas can indeed be absent in casual Dutch when they function as syntactic markers.

In order to investigate if there are phonological constraints on the absence of schwa, we counted the number of different phonological environments in which schwas were absent and in which they were present. For this analysis, we included only adjectives that were directly preceded and followed by another word (58 reduced adjectives and 3,981 unreduced adjectives). Reduced schwas were preceded by 11 different phonemes and followed by 19 different phonemes. In total, they occurred in 39 different contexts. The schwas in the unreduced adjectives were preceded by 16 different phonemes, followed by 30 phonemes, and occurred in 304 different contexts. The difference in the number of contexts for reduced and unreduced schwas is likely to be due to the substantial difference in sample sizes. A comparison of the phonological contexts showed that 100% of the contexts in which the schwa was absent also occurred in the sample of adjectives in which the schwa was present. This strongly suggests that there is a large (if not complete) overlap in phonological contexts between cases in which schwas are absent and cases in which they are present. Schwa reductions thus occur in many different phonological contexts and there do not seem to be any apparent segmental constraints on where inflectional schwa may be absent.

We also investigated whether there might be an influence of lexical frequency on schwa reduction because previous studies have shown that how frequent and predictable a word is influences how likely it is to be reduced (e.g. Bell et al., 2009; Pluymaekers et al., 2005b). We therefore collected log-transformed word frequencies for the preceding word, the adjectives themselves, and the following word from Celex's Dutch Morphological Word database (Baayen et al., 1995). For the unreduced adjectives, the preceding words had a mean frequency of 12.57 (SD = 2.29) while the preceding words for the reduced adjectives had a mean frequency of 13.27 (SD = 1.99). The unreduced adjectives had a mean frequency of 8.1 (SD = 2.04) and the reduced adjectives had a mean frequency of 7.52 (*SD* = 2.02). The words following the unreduced adjectives had a mean frequency of 7.76 (*SD* = 2.37) whereas the words following the reduced adjectives had a mean frequency of 8.32 (*SD* = 2.15). Overall, the frequencies of the reduced and the unreduced adjectives show no striking differences.

4.3 ERP study

4.3.1 Method

Participants

Thirty-two native speakers of Dutch were recruited from the subject panel of the Max Planck Institute for Psycholinguistics. All were university students and right-handed. Age ranged from 18 to 24 years (mean: 20.9) and five of the participants were male. The participants reported no hearing problems and had normal or corrected-to-normal vision. They were informed about the procedure of the experiment before taking part and were paid for their participation.

Materials and design

Four types of utterances were created: critical, control, filler, and practice utterances. Each utterance consisted of two or three sentences. Table 4.1 shows an example of the critical, control, and filler utterances. Each utterance was produced by a male and a female native speaker of Dutch. During the experiment, each speaking style (careful vs. casual) was mapped consistently with one of the two speakers for a given participant. Across participants, the mapping of speaker and speaking style was balanced. Associating a particular speaking style with a specific speaker makes our study more comparable with Hanulíková et al. (2012) in which there was also a consistent mapping between speaker and accent (native vs. foreign).

One hundred and twenty critical utterances were constructed. These contained a noun phrase consisting of the indefinite article *een* "a", an adjective, and a common-gender noun. For each utterance a correct and an incorrect version was created. In the correct version the adjective was inflected whereas in the incorrect condition it was not inflected. The utterances were created according to the following criteria: First, the sentence accent was not on the adjective-noun pair. Second, there were at least five syllables after the noun before the end of the utterance. Third, there was only one adjective-noun pair in the utterance. Fourth, the noun was never mentioned more than once in the utterance. Fifth, the word preceding the adjective did not give away whether the following noun would be a common- or a neuter-gender noun.

Utterance type		Example
Critical	Dutch	MORGEN ga ik met de trein naar Berlijn. Ik wil nog een spannende roman / *spannend roman kopen VOOR ik vertrek. Dan heb ik iets te LEZEN.
	English	TOMORROW I'm going to travel to Berlin by train. I want to buy a suspenseful novel BEFORE I leave. Then I will have something to READ.
Control	Dutch	Ik liep langs een vijver waar werd GEVIST. Toen haalde een visser toevallig NET zijn hengel / *atleet binnen met een VIS eraan.
	English	I was walking beside a pond used for FISHING. Then, coincidentally, a fisherman JUST pulled his fishing rod / athlete out with a FISH on it.
Filler	Dutch	Mijn oma is de laatste tijd heel warrig en MOE. De dokter zegt dat het een mogelijk effect is van haar MEDICIJNEN.
	English	My grandma has recently been rather woozy and TIRED. The doctor says that this might be a possible side effect of her MEDICINE.

Table 4.1: Example utterances for each of the three utterance types. Words with sentential stress are written in capital letters. Crucial words are underlined.

In addition to the critical stimuli, 104 control utterances were created. These consisted of pairs of utterances that differed only in whether they included a semantically correct or incorrect noun. In contrast to the critical utterances, there were no constraints on the kinds of words that could precede the nouns in the control utterances. In order to make each noun semantically more predictable, a strong semantic expectation was generated during the phrase which preceded the noun. As in the critical utterances, the nouns in the control utterances were not repeated, did not carry sentence accent, and were followed by about five syllables before the end of the utterance.

Furthermore, a set of 60 filler stimuli was constructed. These consisted of utterances containing adjective-noun pairs in which the noun carried the neuter gender. In contrast to the critical items, the filler utterances never contained grammatical errors, that is, the adjective was always correctly uninflected (i.e. did not end in a schwa). These items were constructed in a similar fashion to the critical utterances.

Taking together the critical, control, and filler utterances, we created a total of 284 utterances. The control and critical utterances were based on Hanulíková et al. (2012), but often with substantial adjustments which were intended to make the utterances more likely to be produced in a casual way. Finally, 10 more utterances were created that served as practice stimuli. The complete set of critical, control,

filler, and practice utterances is available at the following location:
http://www.mpi.nl/people/viebahn-malte/tool/stimulus_utterances.txt

For the critical stimuli we used a 2×2 factorial design with the factors speaking style (careful vs. casual) and grammaticality (correct vs. incorrect adjectival form). During the experiment, the 120 critical items were divided equally across the four cells of the design so that 30 utterances occurred per condition. A similar design was used for the control utterances. The factors were speaking style (careful vs. casual) and semantic validity (correct vs. incorrect noun). Each cell of the design was filled with 26 utterances by distributing the 104 control utterances equally across conditions. As there were only correct versions of the filler stimuli, these utterances could only occur in two conditions: 30 of the 60 filler utterances were produced in a careful and 30 in a casual manner. The condition in which a given utterance was presented varied across participants. During the experiment, 60 utterances (21%) contained adjectives without the appropriate inflectional schwa and 52 (18%) contained semantically incongruent nouns. In total, 112 (39%) of the utterances contained either a semantic or a syntactic error.

In order to motivate the participants to stay alert during the experiment and to listen to the content of the experimental utterances, *yes-no* questions were pseudorandomly presented during the experiment. There were two questions during the practice trials (one in which the correct answer was *yes* and one in which it was *no*). For the remaining trials, there were 18 questions, which was one question approximately every 12 sentences. Each question was followed by a filler in order to avoid spill-over effects on the critical or semantic trials. The questions were about the content of the preceding utterance (e.g. *Was I recently on vacation in France?*). Half of the questions followed a casually produced utterance whereas the other half followed a carefully produced utterance. In half of the cases, the correct answer was *no* and in the other half it was *yes*.

Stimulus recordings

Recordings were made by a male and a female native speaker of Dutch. Each speaker produced careful and casual versions of each utterance. A correct and an incorrect version was recorded of each critical and control stimulus. For the careful utterances, the speakers were instructed to speak in a deliberate and careful manner but not so that it would sound like they were reading the words out loud. For the casual utterances, the speakers were asked to produce the words in an informal manner. They were encouraged to reduce segments if this seemed natural to them and to speak with a high speaking rate, which is typical of a casual speaking style. For the incorrect versions of the utterances, speakers were explicitly told to produce incorrect words. In order to determine whether or not a schwa was present we inspected the

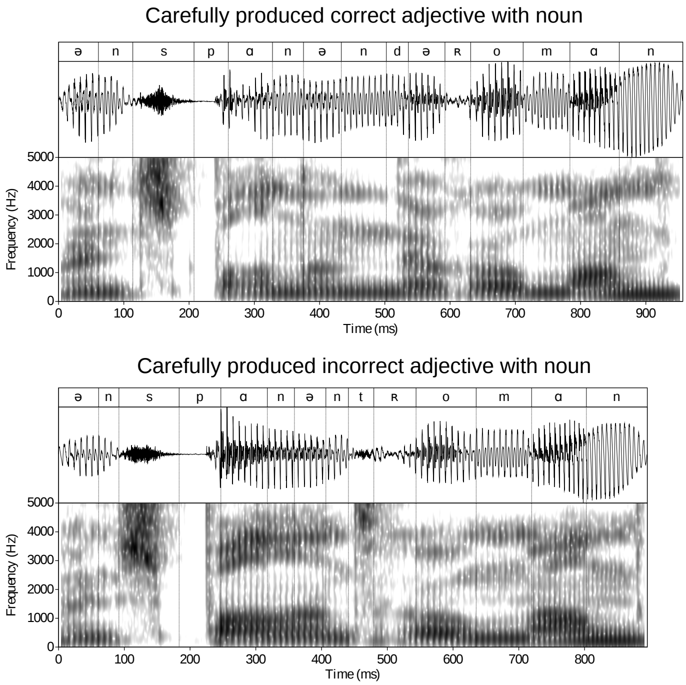


Figure 4.1: The Dutch adjective-noun pair *spannende roman* "suspenseful novel" recorded by the male speaker in a careful speaking style with adjective-final schwa (top) and without adjective-final schwa (bottom). The adjective without schwa is syntactically incorrect because the noun carries the common gender. Note that syllable-final /d/ is devoiced in Dutch and thus produced as [t].

recordings with audio editing software. If there was no vocalic portion at the end of the adjectives we concluded that the inflectional schwa was absent. Otherwise we concluded that it was present. If an adjective that was intended to be uninflected was produced with a schwa, we recorded it again without a schwa so that in the end all of the adjectives in the incorrect condition were produced without schwa and all those in the correct condition were produced with schwa. For the critical utterances, we recorded a total of 960 tokens. For the 104 control utterances, a total of 832 tokens were recorded and for the 60 filler utterances we recorded 240 tokens. Including the 10 practice items, a total of 2,042 tokens were recorded.

As an example, carefully and casually produced versions of the critical adjective-noun pair *spannende roman* "suspenseful novel" are shown in Figures 4.1 and 4.2. In some cases the speaker introduced relatively long pauses between the sentences forming a given utterance. The long pauses made the utterances sound unnatural.

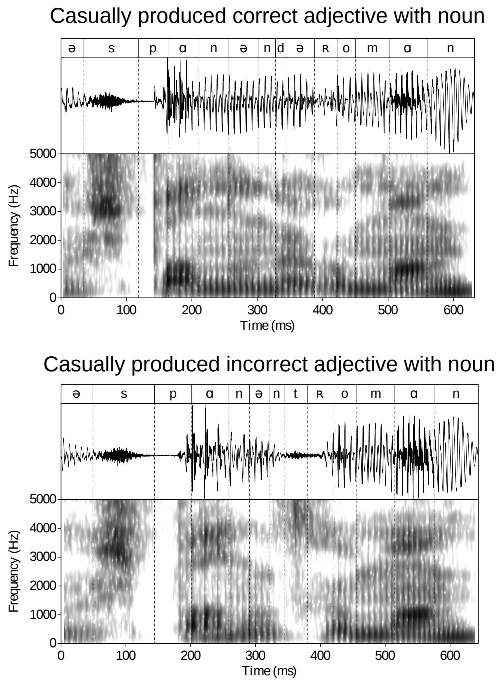


Figure 4.2: The Dutch adjective-noun pair *spannende roman* "suspenseful novel" recorded by the male speaker in a casual speaking style with adjective-final schwa (top) and without adjective-final schwa (bottom). The adjective without schwa are syntactically incorrect because the noun carries the common gender. Note that syllable-final /d/ is devoiced in Dutch and thus produced as [t].

In order to avoid unnaturally sounding utterances, the recordings were adjusted such that the maximum duration of a pause was 400 milliseconds (ms) with PRAAT audio editing software. The durations of the adjectives, adjusted sentences, and schwas are shown in Figure 4.3. Overall, casually produced sentences, adjectives and schwas were clearly shorter than carefully produced ones. Furthermore, adjectives produced with schwa were longer than adjectives produced without schwa.

A phonetically transcribed example utterance is shown in Table 4.2. These transcriptions show the influence of speaking style on the phonological implementation of the utterances. Whereas the carefully produced sentences overlap 94% of the time with a dictionary transcription of the utterance based on the CELEX lexical database (Baayen et al., 1995), the casually produced sentences overlap only by 73% with the dictionary transcription. When we consider only the words that precede the adjective, the careful utterance has 93% overlap with the dictionary transcription whereas

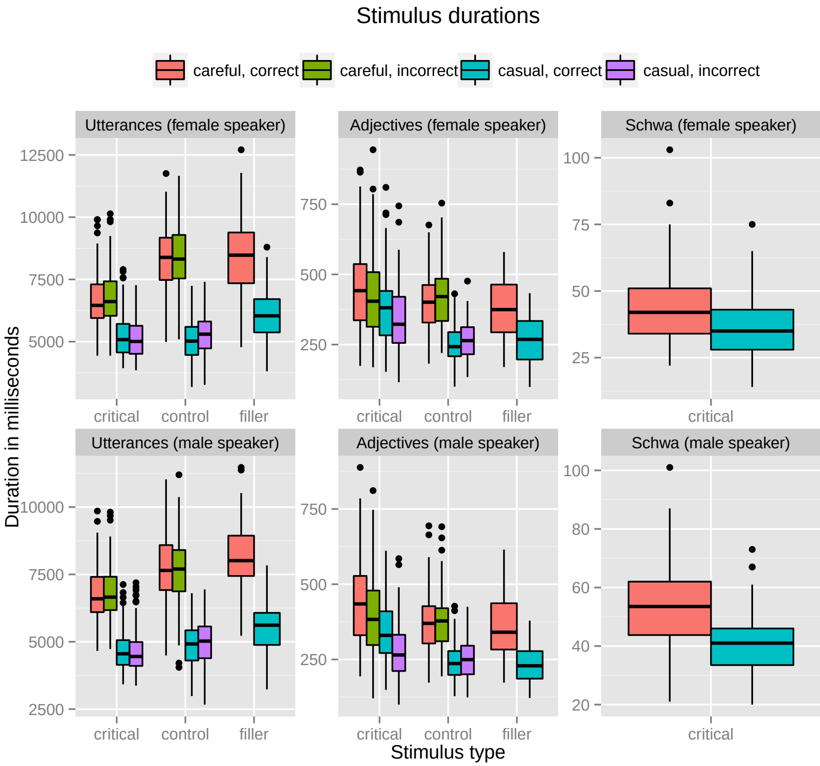


Figure 4.3: Durations of the (adjusted) utterances, adjectives, and schwas. *Critical* stimuli refer to utterances that may contain morphosyntactic violations, *control* stimuli refer to utterances that may contain semantic violations, and *filler* stimuli refer to utterances that never contain any type of violation.

the casual utterance has only 60% overlap with the dictionary transcription. This shows that the words that precede the adjective provide information about the probability with which the following segments (such as the inflectional schwa) will be reduced. As the probability of absent segments is relatively high, the absence of the inflectional schwa may be attributed to the speaking style rather than be treated as a grammatical error.

Apparatus

The EEG was recorded at a sampling rate of 500 Hz with Ag/AgCl electrodes placed at 26 sites according to the International 10-20 system attached on the ActiCap sys-

Register	Phonetic transcription	N dictionary segments	% Overlap
Dictionary	[mɔrxə xa ɪk mɛt də trɛin nar bɛrlɛin. ɪk ʊɪl ʊɛl nɔx ən spɑnəndə romɑn kopə for ɪk fɛrtrɛk. dɑn hɛp ɪk its tə lezə.]	88	–
Careful	[mɔrxə xa ɪʔ mɛt ə trɛin nar bɛrlɛin. ɪk ʊɪl ʊɛl nɔx ən spɑnəndə romɑn kopə for ɪk fətrɛk. dɑn hɛp ɪk its tə lezə.]	83	94%
Casual	[mɔ xa ʔ mɛ trɛin na bɛlɛin. xəʊɛl nɔx ə spɑnəndə romɑn kopə fo ʔ fətrɛk. dɑn hɛp its tə lezə.]	64	73%

Table 4.2: Example transcriptions illustrating differences in acoustic reduction between carefully and casually produced utterances. See Table 4.1 for the orthographic transcription and English translation. A segment was counted as not corresponding to the dictionary transcription if it was either missing or changed (for example a glottal stop [ʔ] instead of a [t]). Both versions were produced by the male speaker using the correct adjectival form. The duration of the complete careful utterance is 7,272 ms and the duration of the casual utterance is 4,794 ms.

tem (Brain Products GmbH, Gilching, Germany). The following 26 electrodes were used as active electrodes: Fp2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1, and O2. In order to monitor horizontal electro-oculograms (EOG), electrodes FT9 and FT10 were placed on the left and right temple of the participant. Vertical electro-oculograms were measured with the electrodes Fp1 and Oz, which were placed above and below the left eye respectively. The ground electrode was placed at Fpz. Electrodes were referenced online to the left mastoid (TP9). An additional electrode (TP10) was attached to the right mastoid for offline referencing. The impedance of the electrodes was kept below 15kO. The EEG and EOG signals were recorded and digitized with PyCorder and amplified by a BrainAmp DC amplifier with an online bad pass filter for 0.02 to 200 Hz. The montage of the electrodes is shown in Figure 4.4. For the registration of the button presses that participants made, a USB game pad was used.

Procedure

The utterances from each stimulus type were divided into sets of equal size and each set was randomly assigned to each of the experimental conditions. The mapping of speaker and speaking style was identical across all utterance types for a given parti-

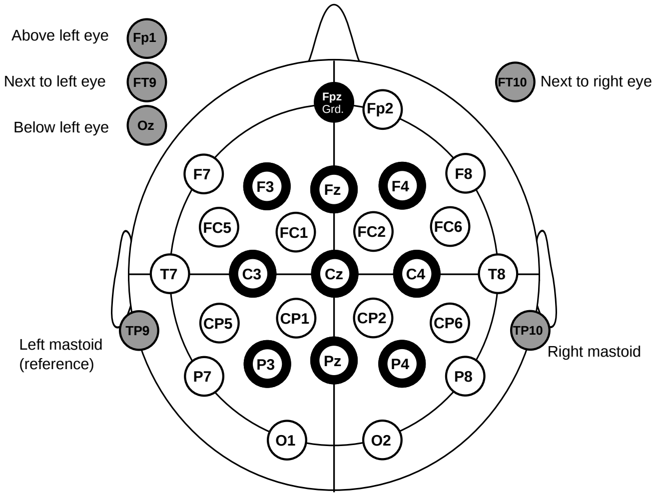


Figure 4.4: EEG montage. Data from the nine electrodes with a broad black border were entered into the statistical analysis (see Figure 4.6)

cipant. The utterances were then combined and a pseudo-random running order was created (with the constraint that a given condition could not occur more than three times in a row). The utterances in this running order were then rotated through each condition, resulting in 8 rotations for the running order (4 rotations for the filler utterances). Rotations 5 to 8 were replications of rotations 1 to 4 with the mapping of speaker (male vs. female) and speaking style reversed. This procedure was repeated four times, resulting in 32 unique lists (one for each participant). The practice stimuli were randomized manually. Two rotations were created, one in which the careful speaker was female and the casual speaker was male and another one with the reversed mapping of speaker and speaking style.

During the experiment, participants were seated in front of a computer screen in a sound-attenuated booth. Auditory stimuli were presented via headphones at a comfortable listening level. There were two types of trials: listening trials and question trials (see Figure 4.5). Listening trials began with the presentation of a blank screen for 500ms, followed by a fixation cross for the same amount of time before an utterance was presented via headphones. Five hundred milliseconds after the end of the utterance, the fixation cross disappeared and instead three dashes appeared in the center of the screen. Participants were instructed to blink only when the dashes were present on the screen. The next trial began after participants had pressed a button. The question trials also began with a blank screen, which was followed by a question printed on the screen. After participants had indicated their answer by

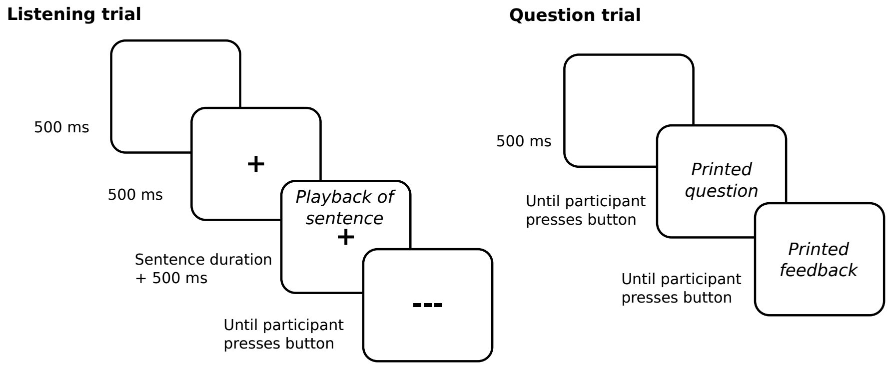


Figure 4.5: Procedure during the listening and question trials in the experiment. For the question trials, the printed feedback was either the word correct printed in green or the word incorrect printed in red.

pressing a button on a game pad, the word correct (in green) or incorrect (in red) appeared on the screen providing the participant with feedback as to whether or not they had answered the question correctly.

One experimental session consisted of 10 practice trials and 284 experimental trials resulting in a total of 294 trials (excluding the 24 question trials). The 284 experimental trials were divided into four blocks consisting of 71 trials each. In between each block, participants were allowed to take a short break. One experimental session took approximately 50 minutes. In addition, the fitting of the EEG equipment took another 20-30 minutes.

4.3.2 Results

We analyzed the question trials in order to examine if speaking style influenced how well participants were able to respond. Questions following a carefully produced utterance were responded to correctly in 93% of the cases (mean Reaction Time [RT]: 2,954 ms). Questions following casually produced utterances were answered correctly 96% of the time (mean RT: 2,958 ms). RT was measured from the time when the printed questions appeared on the screen until participants pressed either the *yes* or the *no* button on the response pad. Subsequent linear-mixed effects models showed that neither the difference in accuracy nor the difference in RT was statistically significant (both $|t| < 0.4$). The analysis of the question trials therefore does not provide evidence that would suggest that participants had any difficulty in comprehending the sentences that were produced in a casual speaking style.

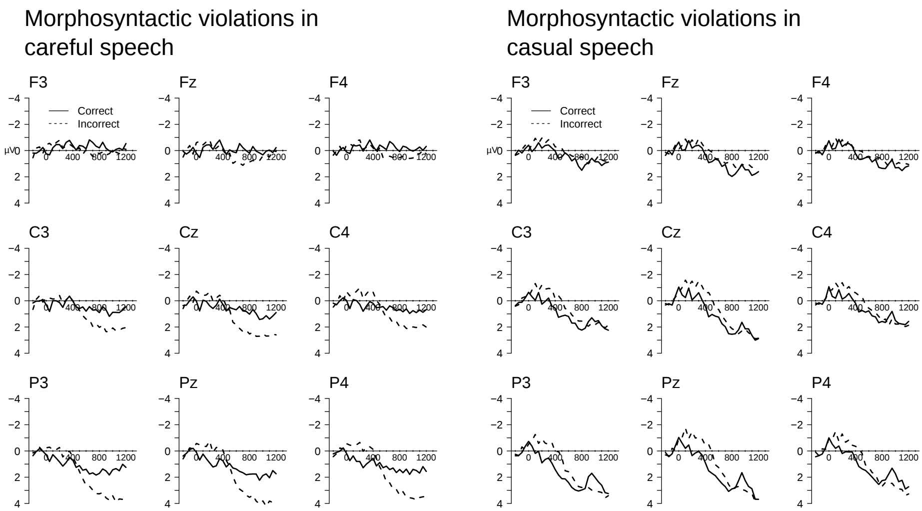


Figure 4.6: ERPs for critical utterances (morphosyntactic violations) for the electrodes entered into the statistical analysis: F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4 (see Figure 4.4 for topographical distribution).

We analyzed the EEG data by computing repeated-measures ANOVAs for subject means with the statistical software R (R Core Team, 2015) and the package ez (Lawrence, 2013). For the critical utterances, the statistical factors were speaking style (careful vs. casual) and grammaticality of the adjective (correct vs. incorrect). For the control sentences, the factors were speaking style (careful vs. casual) and semantic validity (correct vs. incorrect). The components of interest were the N400 and the P600. The P600 was measured from the onset of the noun that followed the grammatically correct or incorrect adjective. The N400 was measured from the onset of the semantically correct or incorrect noun. For the statistical analysis of the P600 component, the time window ranged from 500 to 1,500 ms. For the analysis of the N400 component, the time window ranged from 300 to 500 ms. These time windows were chosen in line with previous research (e.g. Hanulíková et al., 2012) and on the basis of visual inspection of the averaged data. The time period 200 ms until noun onset was used for baselining. We focused on the three frontal electrodes (F3, Fz, F4), three central electrodes (C3, Cz, C4), and three parietal electrodes (P3, Pz, P4). These electrodes were chosen for comparison with the study by Hanulíková et al. (2012). Plots of the ERPs for these electrodes are shown in Figures 4.6 and 4.7. The topographies of the differences between correct and incorrect utterances are shown in Figure 4.8. Before the statistical analysis of the data, each trial was checked for

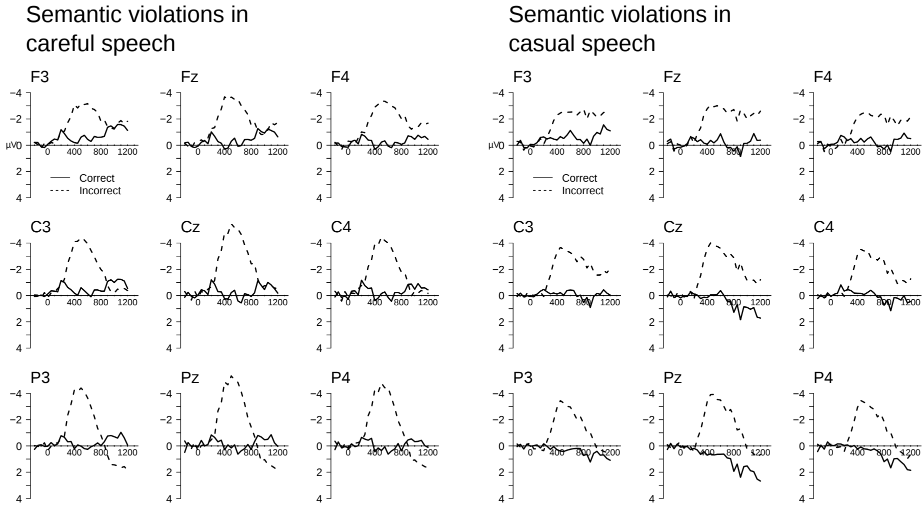


Figure 4.7: ERPs for control utterances (semantic violations) for the electrodes entered into the statistical analysis: F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4 (see Figure 4.4 for topographical distribution).

artifacts due to head movements, eye movements or blinks. Trials in which such artifacts occurred during the baseline period or the time windows in which an N400 or P600 effect was expected were discarded.

There were no effects on the P600 component for the frontal electrodes (all $Fs < 4$ and all $ps > 0.05$). One of the central electrodes (C3) showed a significant interaction between speaking style and grammaticality ($F(1, 31) = 4.71$, $p < 0.05$). The parietal electrodes P3 and Pz also showed such an interaction (P3: $F(1, 31) = 6.49$, $p < 0.05$; Pz: $F(1, 31) = 5.27$, $p < 0.05$). For P4, there was a main effect of grammaticality showing a larger P600 for incorrect compared to correct utterances ($F(1, 31) = 6.28$, $p < 0.05$) but no interaction ($F(1, 31 = 3.42$, $p = 0.07$). In order to examine the interactions between grammaticality and speaking style at the electrodes C3, P3, and Pz, separate one-way ANOVAs with the factor grammaticality were run for the careful and the casual conditions. All three electrodes showed a significant effect of grammaticality for the carefully produced utterances (all $Fs > 7$ and all $ps \leq 0.01$) but not for the casually produced ones (all $Fs < 0.2$ and all $ps > 0.6$).

In order to compare our results more closely to the results reported by Hanulíková et al. (2012), we re-ran our analysis using exactly the same time window as they did (i.e., 800-1,200ms). According to these analyses, the interaction between speaking style and grammaticality was present at electrode P3 ($F(1, 31) = 5.11$, $p < 0.05$) but

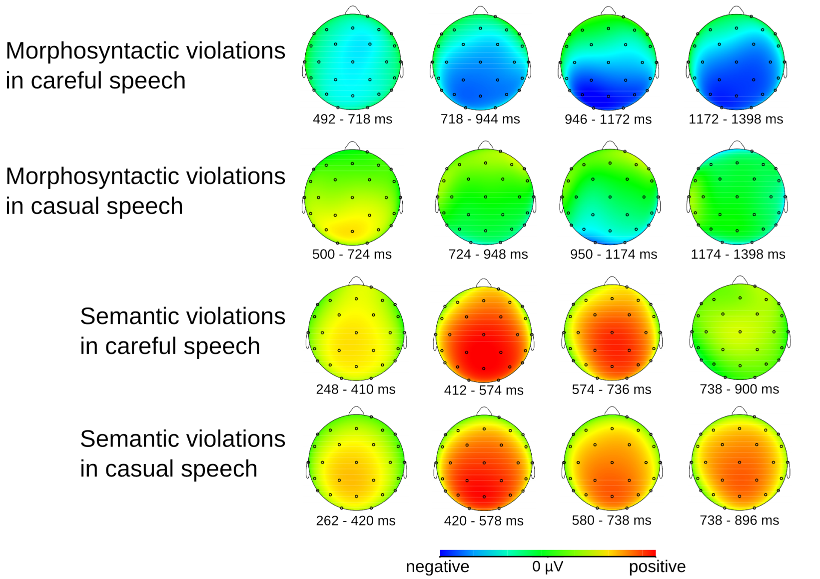


Figure 4.8: Topographies of difference values between correct and incorrect utterances.

absent for electrodes C3 and Pz. One-way ANOVAs for electrode P3 showed that there was a significant effect of grammaticality for careful utterances ($F(1, 32) = 11.39$, $p < 0.01$) but not for casual utterances ($F(1, 32) = 1.58$, $p > 0.2$). The fact that the effect remained significant in this narrower time window at only one out of three electrodes suggests that casual speech takes more time to process than careful speech. This finding is consistent with previous studies that have shown that the recognition of words is slower in casual compared to careful speech (e.g., van de Ven et al., 2011).

Hanulíková et al. (2012) found that listeners' sensitivity to morphosyntactic violations changed over the course of the experiment. More specifically, sentences with gender violations produced by native speakers elicited a P600 effect in the first half of the experiment but not in the second half. In order to investigate whether our listeners' showed the same pattern of results, we conducted additional ANOVAs for the electrodes for which we had found an interaction between grammaticality and speaking style (C3, P3, and Pz) and added part of the experiment (first half vs. second half) as an additional factor. None of these analyses showed a three-way interaction between speaking style, grammaticality, and part of experiment (all $Fs < 1$ and all $ps > 0.3$).

For the control sentences, we investigated if the effect of semantic validity on the N400 component differed for carefully and casually produced utterances. We used repeated-measures ANOVAs with the factors speaking style and semantic validity as well as the interaction of these factors. All of the nine electrodes showed a significant main effect of semantic validity reflecting a larger N400 component for incorrect compared to correct utterances (all $Fs > 32$ and all $ps < 0.001$). Crucially, none of the electrodes showed an interaction between speaking style and semantic validity (all $Fs < 3.5$ and all $ps > 0.07$).

To summarize, the ERP data for the critical stimuli show an interaction between speaking style and grammaticality for the P600 component. More specifically, there was an effect of grammaticality on the amplitude of the P600 for careful but not for casual speech. Furthermore, our analyses suggest that the interaction between speaking style and grammaticality remained constant across the course of the experiment. For the control stimuli there was an N400 effect for casual as well as careful speech but no interaction between speaking style and semantic validity.

4.4 Discussion

Unlike in careful speech, which is typically produced in formal social contexts, phonological segments in casual speech are often times reduced or absent. The purpose of the present study was to examine whether the absence of syntactically relevant schwas disrupts syntactic processing in casual speech. We conducted an experiment in which participants listened to carefully and casually produced utterances which contained either a correctly inflected adjective (i.e., one ending with inflectional schwa) or an incorrectly uninflected adjective (i.e., without the final schwa). Consistent with previous studies (Hagoort and Brown, 1999), the incorrectly uninflected adjectives in careful speech elicited a P600 effect. When occurring in casual speech, however, uninflected adjectives did not elicit a P600 effect. This suggests that listeners did not treat the absence of the syntactically relevant schwa as a grammatical error when listening to casual speech, but they did so when listening to careful speech. In order to control for the possibility that the absence of a P600 effect in casual speech was due to listeners not understanding the content of the utterances, we included a control condition in which participants listened to utterances that contained semantic violations. We found a clear N400 effect for semantically incorrect nouns that occurred in careful as well as casual speech. Crucially, there was no interaction between speaking style and semantic validity. This suggests that speaking style had no influence on how well listeners understood the content of the utterances. The notion that listeners recognized the careful and the casual sentences equally well is further corroborated by the fact that participants answered questions about the content of casually produced utterances as quickly and accurately as they answered questions about the carefully produced utterances.

Our findings extend previous research on the processing of casual speech (e.g. Brouwer et al., 2012b; Ernestus, 2014) by showing that the way in which listeners adapt to reduced word forms can have consequences for syntactic processing. The absence of grammatically necessary schwas in a casual speaking style does not disrupt syntactic processing because the absence is consistent with the speaking style in which it occurs. This suggests that absent schwas in casual speech are effectively not grammatical gender violations. Previous studies suggest that listeners can use syntactic information in order to help recognize words in reduced speech (e.g. Tuinman et al., 2014; Viebahn et al., 2015). The present study shows that adapting to acoustic-phonetic reductions influences syntactic processing and thus highlights the importance of the interplay between acoustic-phonetic and syntactic information in speech processing.

The design of our study was deliberately chosen to be very similar to that of Hanulíková et al. (2012). In that study, participants were exposed to Dutch sentences containing syntactic gender violations that were either produced by a native speaker or a non-native speaker with a foreign accent. Whereas the violations elicited a P600 effect for the native speaker, such an effect was absent for syntactic violations produced by a non-native speaker. The main difference from our study was that Hanulíková et al. used a condition in which utterances were produced by a non-native talker whereas we used a condition in which utterances were produced by a casually speaking native talker. In both conditions the absence of grammatical markers can be expected because non-native speakers of Dutch and native speakers that are talking in a casual way both regularly omit these markers. Both studies thus share a condition in which the absence of syntactic gender markers is unexpected (i.e. a carefully speaking native talker) and a condition in which the absence of syntactic gender markers can be expected (i.e. a speaker with a foreign accent or a casual speaking style). Our main results thus parallel the results found by Hanulíková et al.: Listeners respond to the ungrammatical absence of a gender-marking schwa with a P600 effect if the absence is unexpected, but they do not show a P600 effect if the absence of the schwa could be expected given the available information about the talker and the type of speech he or she produces.

However, our results do not match Hanulíková et al. completely. In their study, listeners' sensitivity to syntactic violations changed over the experiment (the P600 effect for the native speaker was limited to the first half). We did not find evidence for such a change. How could this difference be explained? Previous research has shown that the P600 component is influenced by the proportion of trials during which errors occur within an experiment. For example, Hahne and Friederici (1999) found a P600 response to phrase structure violations if they occurred in 20% of the trials but not if they occurred in 80% of the trials. The difference between the results of our study and the study by Hanulíková et al. is likely to be due to differences in the proportion of trials containing grammatical errors. In Hanulíková et al., a grammatical violation occurred in 35% of the trials whereas in our study an error occurred in only 21% of the trials. A further important difference between our study and Hanulíková et al. is that our grammatical violations were limited to adjectival inflections whereas Hanulíková et al. also included incorrect determiners (the common-gender determiner *de* instead of the neuter-gender determiner *het*). These violations were more likely to be detected by the listeners than the absence of adjective-final schwas. If we add in that half of the violations in our experiment were possibly undetected (the absent schwas in the casual speaking style), the proportion of trials containing an error becomes even smaller (only 10.5%). It is therefore quite likely that we did not find a change in the P600 response over the experiment because the proportion of trials with noticeable grammatical errors was considerably smaller compared to that in Hanulíková et al.

The notion that the participants in our study may not have noticed the absence of inflectional schwas in the casually produced utterances raises the question of what exactly the mechanism is that allowed listeners to adapt to absent schwas in casual speech. One possibility is that listeners change the way in which they interpret the absence of inflectional schwa based on the preceding phonological context. As a result of the casual speaking style, many words in a casually produced utterance contain acoustic-phonetic reductions. This is illustrated in the example utterance shown in Table 4.2. In this utterance, the proportion of realized segments is considerably smaller if the utterance was produced with a casual speaking style than when it was produced with a careful speaking style. Listeners might have kept track of the probability with which the speaker produced (or omitted) individual segments and they might have taken this probability into account when interpreting the absence of speech sounds. The absence of inflectional schwa would therefore not be interpreted as a grammatical error but instead it would be consistent with the fact that a casually speaking talker is likely to omit individual segments. This explanation is consistent with previous studies that have shown that listeners are sensitive to probabilistic information about speech sounds. For example, McQueen and Huettig (2012) found that listeners changed the way they used phonological information when recognizing spoken words if the words appeared in sentences that were disrupted by intermittent bursts of noise. This suggests that the perceptual weight assigned to acoustic information during speech recognition can change as a function of the context in which that information is heard.

Another possibility is that listeners adapt to the acoustic-phonetic consequences of fast speech which characterizes casually produced utterances. Due to the high speaking rate, all segments in the utterance become shortened and compressed. As a result, it becomes difficult to distinguish sounds from one another and to determine whether or not a given segment is present. Furthermore, a high speaking rate is characterized by increased coarticulation of segments, which leads to the spread of phonetic features across neighbouring segments. This means that in cases in which the inflectional schwa was preceded by a voiced segment (e.g. *blauwe wieg*, "blue cradle"), the voiced portion of the sound preceding the schwa might be coarticulated with the following segments. Listeners might have adapted to this situation by stopping to try to detect whether or not a schwa is present. Note, however, that this adaptive process is still likely to be specific to syntax because the ability to comprehend the content of the sentences did not suffer from the casual speaking style.

Because a casual speaking style is characterized by both a high speaking rate and the absence of phonemic segments, it is difficult to tease apart which of these factors is crucial for the absence of the P600 that we observed in our study. Future research could further explore this question and examine how the different phonological and acoustic properties that characterize casual speech could be isolated and their individual effects studied. The current results nevertheless indicate that, whether it is triggered by absence of segments or by speaking rate (or both), adaptation results in syntactic processing of casual speech that is not disrupted by absent schwas.

Our results also advance understanding of the nature of syntactic processing and what the P600/SPS ERP component can reveal about them. Although there is general consensus that the P600/SPS is correlated with syntactic processing difficulties, there is a debate about the underlying mechanisms. It is still debated whether the P600/SPS occurs exclusively with linguistic stimuli or whether it reflects non-linguistic cognitive processes. For example, it has been suggested that the P600/SPS is part of the P300 family of ERP components. More specifically, the P600/SPS resembles the P3b component (e.g. Coulson et al., 1998) which is elicited by rare categorical events (so-called *odd-balls*) which can be either linguistic or non-linguistic stimuli. This proposal is consistent with findings that suggest that the P600/SPS reflects late and controlled rather than early and automatic processes (Hahne and Friederici, 1999). These studies call into question the assumption that the P600/SPS component is language-specific (cf. Osterhout and Hagoort, 1999). Our findings contribute to this debate by showing that the processes that underlie the P600 are not likely to be purely automatic. One hallmark feature of automatic processes is that they are mandatory. Our results, however, suggest that the processing of syntactic violations is flexible and can be adapted quickly in different phonological contexts. This implies either that the P600/SPS component does not reflect syntactic processing or, more likely, that the syntactic processes that it taps into are not strictly mandatory.

In conclusion, the present study shows that morphosyntactic violations that are the result of schwa omissions do not disrupt syntactic processing if they occur in a casual speaking style. This provides further support for the notion that listeners process language in an adaptive and flexible manner. Adaptation at an acoustic-phonetic processing level can have consequences at higher (i.e. syntactic) levels in the processing hierarchy.

The influence of phonetic variability and orthography on the production and recognition of novel French words

Chapter 5

This chapter is based on:
Malte Viebahn, Audrey Bürki, Mirjam Ernestus,
James McQueen, and Ulrich Frauenfelder (in preparation).
The influence of phonetic variability and orthography on the production and recognition of novel French words.

The present study investigates how phonetic variability and orthography influence the processing of words that may be pronounced with or without the vowel schwa. It has previously been suggested that orthography might play a crucial role in how listeners deal with variation in speech. Here we asked: How large is the influence of phonetic information compared to the influence of spelling? We also asked if these two factors influence both spoken-word production and spoken-word recognition. Participants learnt novel French words in which schwa was unreduced (i.e. present, e.g. *secobe*) or reduced (i.e. absent, e.g. *scobe*). In Experiment 1, the reduced words were either consistently reduced (i.e., never produced with schwa) or inconsistently reduced (i.e., sometimes produced with and sometimes produced without schwa). In Experiment 2, words were always consistently reduced or unreduced but an orthographic exposure phase was included in which reduced words were either spelled with the letter <e> or without it. Results suggest that both phonetic and orthographic information can influence the processing of phonological input. However, the influence of phonetic variability outweighs the effect of spelling. Furthermore, evidence from naming and eye-tracking tasks indicates that exposure to both types of information influences not only word production but also word recognition. The influence of orthography on the processing of reduced speech may be smaller than expected because spelling provides less reliable phonological information than phonetic input.

5.1 Introduction

In spontaneous speech, words are often produced in an acoustic-phonetically reduced way (e.g., Ernestus, 2000; Johnson, 2004). As a result of such reductions, individual sounds and syllables may be either shortened or completely absent. One speech sound that is often affected by reduction processes is the vowel schwa. For example, the English word *sufficient* may be pronounced without schwa as in /sfɪʃnt/ and the French word *seringue* "syringe" may be pronounced as /sʀɛ̃g/ instead of /səʀɛ̃g/. Corpus studies of spontaneous speech have shown that these kinds of acoustic reductions are very common across different languages including English, Dutch, French, and German (for a review see Ernestus and Warner, 2011). The large amount of phonetic variability that results from these reductions poses a challenge to psycholinguistic theories that attempt to describe the processes and representations that allow language users to produce and comprehend spoken language. The present study aims to extend our knowledge about the processing of reduced word forms by comparing the influence of phonetic variability with the influence of orthographic information. Furthermore, we investigate whether exposure to variability in the input during word learning influences both word production and word recognition.

The variation in the spoken forms of words due to acoustic-phonetic reductions contrasts with the consistency of orthographic forms. Although there is considerable variation in the handwriting of words, the number and order of the letters that words consist of remain invariant. For example, while a speaker of Dutch may produce the word *eigenlijk* "actually" as either [ɛixələk], [ɛixlək], or even [ɛixə] (Ernestus et al., 2002), the orthographic form always remains <eigenlijk>. Studies have shown that orthographic information can influence how spoken words are processed (e.g., Grainger et al., 2005; Perre et al., 2009; Seidenberg and Tanenhaus, 1979; Taft et al., 2008, but see also Cutler et al., 2010; Cutler and Davis, 2012). Moreover, a word's orthography has been shown to influence how phonetic variation is processed (e.g. Bürki et al., 2012; Racine et al., 2013). Using an artificial-word learning paradigm, Bürki et al. (2012) demonstrated that French listeners who learn novel words without schwas generate schwa variants of these words after having been exposed to orthographic forms that are consistent with the existence of a spoken variant that contains schwa. For example, French-speaking participants learnt the auditory form of the novel word [pluʀ] by associating it with a picture of a novel object. Afterwards, they were exposed to an orthographic form that either contained the letter <e> in the first syllable (<pelour>), or they saw an orthographic form without the letter <e> (<plour>). Words with an orthographic form that contained the letter <e> were named more slowly and were more likely to be produced with schwa compared to words without the letter <e>. These findings demonstrate an orthographic influence on phonological processing.

It is not yet known, however, how large the influence of orthographic information is relative to the influence of acoustic-phonetic information. It is possible that the influence of orthography is quite large and plays a crucial role in the processing of reduced speech. For example, it has been proposed that the influence of orthography can, at least in part, explain why canonical (i.e., unreduced) pronunciation variants are recognized more efficiently than reduced variants even if the reduced variants occur more frequently (e.g., Ranbom and Connine, 2007). This would mean that the influence of orthographic information is large enough to neutralize the effect of variant frequency. A strong influence of orthography on the learning of new words is also plausible considering that the spelling of orthographic forms varies much less than the pronunciation of phonetic forms. However, in order to get from an orthographic form to a phonological one, readers have to draw inferences about which phoneme corresponds to a given grapheme. Depending on the language, these inferences might be quite difficult to make because orthography is not always a reliable cue (e.g., Ziegler et al., 1996). A given grapheme can correspond to multiple phonemes, or it can be silent. In contrast, acoustic-phonetic information is, under good listening conditions, much less ambiguous and provides thus more direct evidence for a particular sound. Thus, both possibilities are plausible: Orthography may be more influential than phonetic information or phonetic information may be more influential than orthography. So far, no direct empirical comparison of the two sources of information has been made. Here, we asked whether the effects of phonetic and orthographic information on the processing of reduced words differ in French. This will allow us to gauge how large the role of orthographic information is relative to that of variability in the acoustic-phonetic input.

The second question of this study concerns the processing of reduced word forms in language production compared to comprehension. We ask whether exposure to phonetic variability and orthographic information influences not only word production but also word recognition. This question has implications for the more general question about how distinct the language production and comprehension systems are. Although both systems have mostly been studied and modeled separately from each other, there is currently a debate about the degree to which both systems make use of the same processes and representations (e.g., Pickering and Garrod, 2013; Schiller and Meyer, 2003). The present study will inform this debate by examining whether the acquisition of pronunciation variants has consequences for both systems.

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

In both experiments, participants learnt over a three-day period novel French words in which schwa was either unreduced (i.e. present, e.g. [səkɔb] or reduced (i.e. absent, e.g. [skɔb]. Reduction was categorical, meaning that reduced words were produced without schwa such that no phonetic cues for schwa were left. The label "reduced" is therefore somewhat arbitrary because "reduced novel words" could also be referred to as "novel words without schwa". However, we use the term "reduced" in order to be consistent with our research question which is about how schwa-reduced words are being processed. In Experiment 1, we examined the influence of phonetic variability on the learning of the novel words. In Experiment 2, we attempted to replicate the effect of spelling on the phonological processing of reduced pronunciation variants found by Bürki et al. (2012) using the same materials as in Experiment 1. In contrast to Experiment 1, words in Experiment 2 were consistently reduced or unreduced. Furthermore, an orthographic exposure phase was included in which reduced words were either spelled with the letter <e> or without it. This allowed us to compare the spelling-based effect with the effect of phonetic variability in Experiment 1. In both experiments, we conducted two tests in order to measure phonological processing: A picture naming task in order to examine processing during word production and an eye-tracking task to examine processing during word recognition.

5.2 Experiment 1

In Experiment 1, we taught participants novel French words that were either produced in a reduced way (i.e., without schwa) or in an unreduced way (i.e., with schwa). Crucially, half of the reduced items were presented in an inconsistent manner: they were sometimes produced with schwa (i.e. unreduced) and sometimes without schwa. We predict that participants ought to be more likely to produce a word with schwa if it had sometimes been presented with schwa and sometimes without it compared to a word that had never been presented with schwa. If phonetic information has a larger impact on phonological processing than orthographic information, this effect ought to be larger than the spelling effect reported by Bürki et al. (2012) (i.e., larger than an approximately 5% increase in productions containing schwa). However, if orthographic information is the more important of the two, the effect of phonetic variation ought to be smaller.

Our second question was whether exposure to a variable input does not only influence word production but also word recognition. If exposure to phonetic variation also influences word recognition, we ought to find that it takes participants more time to recognize words that were sometimes produced with and sometimes produced without schwa compared to words that were never produced with schwa.

5.2.1 Method

Participants

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

Design

During the learning phase and the naming task, the design consisted of one factor with three levels: *consistently reduced*, *inconsistently reduced*, and *unreduced*. These labels refer to the way in which the words were presented during the learning phase. Consistently reduced words were never produced with a schwa (e.g., *scobe*). In contrast, inconsistently reduced words were presented half of the time during learning with a schwa (e.g., *secobe*) and the other half of the time without a schwa. Unreduced words were always produced with a schwa. Note that we could also refer to *inconsistently reduced* words as *inconsistently unreduced* words. However, we chose the label *inconsistently reduced* because the focus of this study is on the processing of reduction. See Table 5.1 for an overview of the experimental conditions. We predict that participants will be more likely to produce inconsistently reduced words with schwa compared to consistently reduced words.

During the eye-tracking task the design was slightly different. In this task, participants were presented with four objects on each trial and were asked to click on one of them. Of the four objects, two objects belonged to the same minimal pair (e.g., *secobe-secophe*) while the other two objects belonged to a different minimal pair (e.g., *belagin-belafin*). In the following, the word that the participants were instructed to click on (e.g., *secobe*) will be referred to as the target word while the other word that belonged to the same minimal pair (e.g., *secophe*) will be referred to as the competitor. The two words belonging to a different minimal pair will be referred to as the distractors. There are five conditions in the eye-tracking task. Note that in each of the five conditions the pronunciation variant of the target depends on

	Experiment 1		Experiment 2	
	Conditions in learning and naming tasks			
1	Consistently reduced: 100% of tokens without schwa		Reduced and spelled without letter <e>	
2	Inconsistently reduced: 50% of tokens without schwa		Reduced but spelled with letter <e>	
3	Unreduced: 100% of tokens with schwa		Unreduced and spelled with letter <e>	
	Conditions in eye-tracking task			
	Target	Competitor	Target	Competitor
1	Consistently reduced (*scobe*)		Reduced, spelled without <e>(*scobe*)	Unreduced, spelled with <e>(*secophe*)
2	Inconsistently reduced, presented without schwa (*scobe*)	Unreduced (*secophe*)	Reduced, spelled with <e>(*scobe*, <*secobe*>)	Unreduced, spelled with <e>(*secophe*)
3	Unreduced (*secobe*)	Consistently reduced (*scophe*)	Unreduced, spelled with <e>(*secobe*)	Reduced, spelled without <e>(*scophe*)
4	Unreduced (*secobe*)	Inconsistently reduced (*s(e)cophe*)	Unreduced, spelled with <e>(*secobe*)	Reduced, spelled with <e>(*secophe*, <*secophe*>)
5	Inconsistently reduced, presented with schwa (*secobe*)	Unreduced (*secophe*)		

Table 5.1: Conditions in the learning, naming, and eye-tracking tasks in Experiments 1 and 2. Example items are shown in parentheses.

the pronunciation variant of the competitor (see Table 5.1 for an overview). In Condition 1, the target was a word that had been consistently reduced during the learning phase (e.g., *scobe*) and the competitor was one that had always been unreduced (e.g., *secophe*). In Condition 2, the target was a word that had been inconsistently reduced during learning and was presented in its reduced variant (i.e., without schwa, e.g. *scobe*) and the competitor was again a word that had always been unreduced during learning (e.g., *secophe*). In Condition 3, the target was an unreduced word (e.g., *secobe*) and the competitor was a consistently reduced word (e.g., *scophe*). In Condition 4, the target was an unreduced word (e.g., *secobe*) while the competitor was an inconsistently reduced word (e.g., *s(e)cophe*). In Condition 5, the target was again an inconsistently reduced word but this time presented in its unreduced variant (i.e., with schwa, e.g., *secobe*) while the competitor was an unreduced word (e.g., *secophe*). Based on the assumption that words with a variable pronunciation will be harder to recognize than words with a consistent pronunciation, we predicted that listeners would take longer to recognize the target word in Condition 2 compared to Condition 1. Similarly, assuming that words with a variable pronunciation will be activated less strongly, inconsistently reduced words should be weaker competitors and we therefore predict faster target recognition in Condition 4 than 3.

Conditions 2 and 5 served as control conditions that allowed us to determine the validity of our experimental paradigm. We predicted that these conditions would replicate the phonological competition effect between words with the same phonological onset that has been demonstrated with real words (e.g., Allopenna et al., 1998; McQueen and Viebahn, 2007) as well as in artificial-lexicon studies (e.g., Creel et al., 2008; Magnuson et al., 2003; Sulpizio and McQueen, 2012). In Condition 5, the variant of the target that was presented had the same initial syllable as the competitor. But in Condition 2, the presented target variant shared only the initial consonant with the competitor. The previous studies suggest that the amount of word-initial phonological overlap has a strong influence on lexical activation. Therefore, we ought to find more target-competitor competition in Condition 5 than in Condition 2.

Materials

For the auditory stimuli, a total of 20 minimal pairs of French pseudowords were created that contained the vowel schwa after the initial consonant (e.g., *secobe-secophe*). For the construction of these novel words, we adopted the following criteria: First, the differences across word pairs had to be as large as possible in order to make it easier for participants to learn (and distinguish between) the words. Second, the word-initial consonant-vowel-consonant sequences had to occur in French words in order to increase the likelihood that listeners would process the vowel as a schwa and not as another vowel (e.g., the close-mid vowel /e/). Third, if the schwa after the first consonant is removed, the resulting onset cluster had to be legal in French (i.e., both the unreduced and the reduced variant had to be phonotactically legal in French). This criterion is important because previous research has suggested that listeners insert a vowel (particularly a schwa) if they hear illegal phoneme sequences (Spinelli and Gros-Balthazard, 2007). We created 10 pairs of bisyllabic and 10 pairs of trisyllabic novel words. The bisyllabic words with schwa have the following structure: $C_1V_S - C_2V_2C_3$ (e.g. *secobe*, produced as [səkɔb]). In this coding scheme, C stands for consonant, V stands for vowel, and the subscript S signifies a schwa. The trisyllabic words with schwa have the structure $C_1V_S - C_2V_2 - C_3V_3$ (e.g. *belagin*, produced as [bəlaʒɛ̃]). The complete set of novel words is shown in the Appendix.

Members of a minimal pair differed only in a single consonant. For the bisyllabic items, this was the last phoneme (e.g. *secobe* vs. *secophe*, produced as [səkɔb] vs. [səkɔf]) whereas for the trisyllabic words, this was the penultimate phoneme (e.g. *belagin* vs. *belafin*, produced as [bəlaʒɛ̃] vs. [bəlafɛ̃]).

The set of novel words began with the consonants /k/, /b/, /f/, /p/, or /s/ (see the Appendix). An important consequence of choosing these initial consonants is that there is a difference between words starting with /k/ and the other words. In French, the grapheme corresponding to /k/ changes depending on whether the following sound is a vowel or a consonant. If it is a vowel, /k/ is represented by <qu>, but if it is a consonant, /k/ is represented by <c>. For the novel words starting with the other consonants, the word-initial grapheme did not change depending on whether or not the following sound was a vowel. In order to test if a change in the word-initial grapheme influenced the results, we included initial consonant as an additional variable in the analyses.

The stimuli were recorded by a female native speaker of Swiss French. For each novel word two different versions were recorded – one with and one without schwa. Each of these versions was recorded twice. The first recording of a given version was presented during the learning phase of the experiment whereas the second recording was presented during the eye-tracking task. The reason for using two different recordings was so that listeners could not recognize the novel words during the eye-tracking task based on the acoustic properties of a particular recording. Recordings were made with a sampling frequency of 22,050 Hz and scaled to 70dB. During the recording procedure, the items were read from a pseudo-randomized list such that two items from the same pair did not follow one another. Each novel word was recorded following the carrier phrase *Cliquez sur le…* "Click on the…". Table 5.2 contains the average word and sentence durations for each recording of the items with and without schwa. Note that the carrier sentences were not produced in a reduced manner. Furthermore, the reduced (i.e., schwa-less) variants were only

		Recording 1		Recording 2	
		Unreduced	Reduced	Unreduced	Reduced
Novel words	Bisyllabic	649 (56)	574 (51)	619 (49)	544 (48)
	Trisyllabic	642 (67)	567 (76)	615 (59)	532 (76)
Carrier phrase	Bisyllabic	648 (51)	677 (65)	654 (52)	681 (52)
	Trisyllabic	635 (44)	639 (47)	655 (60)	665 (71)
Sentences	Bisyllabic	1297 (73)	1252 (71)	1273 (74)	1225 (56)
	Trisyllabic	1277 (66)	1206 (56)	1270 (60)	1197 (91)

Table 5.2: Mean durations (in milliseconds) of the stimuli used in Experiments 1 and 2 with standard deviations shown in parentheses. Recording 1 was presented in the learning phase whereas recording 2 was presented during the eye-tracking task.

reduced in the sense that they were produced without schwa. Speaking rate and articulatory accuracy were the same among the reduced and the unreduced recordings.

For the visual stimuli, 40 pictures were taken from the MPI database of non-existing objects. Each picture was assigned to a specific novel word. The complete set of novel words and their associated pictures is shown in the Appendix. The novel words in all three tasks (learning, naming, and eye-tracking) were rotated through every experimental condition such that each word occurred in each condition across participants but also such that no word appeared, for any given participant, in more than one condition. The only exception to this were the control conditions 2 and 5 in the eye-tracking task in which the same items were used within participants (in order to keep the number of items in the experimental conditions as high as possible). In the learning and naming tasks for a given participant, 10 items were consistently reduced, 10 were inconsistently reduced, and 20 items were unreduced.

The eye-tracking block consisted of two parts that followed each other without interruption. Part 1 consisted of items from conditions 1 through 4. In Part 2 the items from Condition 1 were repeated and presented together with the items from Condition 5. The order of trials within each part was randomized. There were 10 items in each of the five conditions.

Procedure

The procedures of the different tasks are illustrated in Figure 5.1. Before the beginning of the learning blocks, a familiarization task was performed. During this task, participants saw the picture of a random object in the center of the screen and were instructed to click on it by reference to the object's name (e.g., *Click on the secobe*). During the learning blocks, there were either two or four pictures on the screen. After a preview of one second, participants were instructed to click on one of the objects. After clicking on an object, only the correct one remained on the screen. Two-hundred milliseconds after the incorrect object(s) had disappeared, participants received auditory feedback telling them whether or not their choice was correct and repeating the name of the target object (e.g., *Yes, that's the secobe* or *No, look this is the secobe*). Five-hundred milliseconds after the end of the feedback sentence the next trial was initiated. During the learning blocks, targets were never presented with the other member of the same minimal pair. The trial structure during the eye-tracking task was identical to the learning trials with four objects, with the exception that participants did not receive any feedback (neither auditory nor visual). After the eye-tracking task participants performed the picture naming task. During this task, participants were presented with the pictures of the novel objects in random order. On each trial, participants had to recall the name of the object and say it out loud into a microphone. On each trial, participants first saw a fixation cross for 800 ms. Then the picture of an object appeared in the center of the screen while at the same time a short (100 ms) beep sound was played. Participants had four seconds in order to produce the name of the object. Then a blank screen was shown for one second after which the next trial began. The different types of trials described above were presented in different blocks across three experimental sessions which were carried out on three consecutive days. Table 5.3 shows the structure and content of the different blocks.

Apparatus

For the learning phase, the experiment was controlled by PsychoPy running on a Lenovo laptop with Ubuntu GNU/Linux 14.04. The objects were displayed on a 22-inch screen with a resolution of 1680 by 1050 pixels. The auditory stimuli were normalized to 70 dB and presented via headphones at a comfortable volume. For the eye-tracking task, the experiment was controlled by E-Prime running on a Dell laptop with Windows 7. Visual stimuli were again shown on a 22-inch display with a resolution of 1680 by 1050 pixels. Eye-movements were recorded with a SMI RED eye-tracker by SensoMotoric Instruments GmbH, Telto, Germany at a sampling rate of 250 Hz. For the naming task, the same equipment was used as in the learning phase. In addition, participants' vocal responses were recorded using a USB microphone (Snowball by Blue Microphones) in combination with the audio software Audacity.

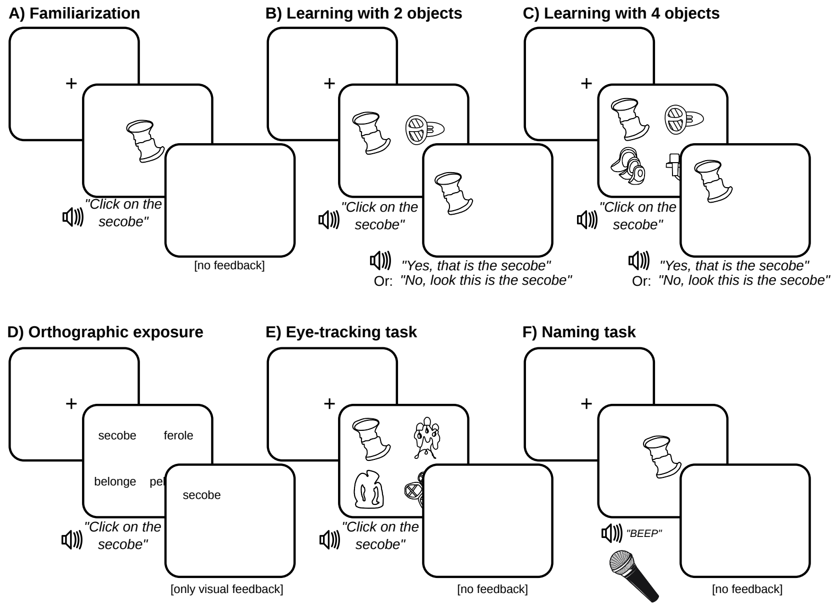


Figure 5.1: Experimental procedures in Experiment 1 (panels A, B, C, and D) and Experiment 2 (panels A,B, and D).

5.2.2 Results

All of the statistical analyses were conducted in the following way. Accuracy scores were analysed with generalized linear-mixed effects models with a binomial link function. RTs were log-transformed and fitted with linear mixed-effects models. Standardized residuals larger than 2.5 were regarded as extreme values and removed during the modeling procedure. In order to analyze the time course of the target gaze probabilities, we used growth curve analysis (Mirman, 2014; Mirman et al., 2008). We estimated p values by means of the Gaussian distribution. Models were fitted with the maximal random effects structure (Barr et al., 2013) unless otherwise noted.

Learning phase

Mean accuracy scores and RTs for the learning phase are shown in Figure 5.2. Overall, the accuracy results show that participants learnt to select the correct novel objects very quickly. Accuracy was already at ceiling in the second learning session. RTs improved continuously over the three learning sessions. Statistical analyses

		Experiment 1		Experiment 2	
	Block	Trials	Task	Trials	Task
Day 1	0	40	Familiarization	40	Familiarization
	1	200	Learning: 2 objects	200	Learning: 2 objects
	2	120	Learning: 4 objects	120	Learning: 4 objects
	3	40	Learning: 4 objects	40	Orthography
	4	80	Learning: 4 objects	80	Learning: 4 objects
	5	40	Learning: 4 objects	40	Orthography
Day 2	6	80	Learning: 2 objects	80	Learning: 2 objects
	7	200	Learning: 4 objects	200	Learning: 4 objects
	8	40	Learning: 4 objects	40	Orthography
Day 3	9	40	Learning: 2 objects	40	Learning: 2 objects
	10	80	Learning: 4 objects	80	Learning: 4 objects
	11	40	Learning: 4 objects	40	Orthography
	12	60	Eye tracking task	40	Eye tracking task
	13	40	Naming task	40	Naming task

Table 5.3: Block structure of Experiments 1 and 2. There was a break between each block. There was also a break after 100 trials in blocks 2 and 7.

showed that participants responded more slowly to inconsistently reduced compared to consistently reduced and unreduced novel words (all $|t|s > 2$ and $ps < 0.05$).

Naming task

For the analysis of the naming results, each vocal response was checked for accuracy. Productions of non-target words, no responses, and mispronunciations were considered as errors and removed from the dataset. Responses were categorized as *correct* if they corresponded either to the unreduced or the reduced variant of the target word. The classification as unreduced or reduced was performed by a trained phonetician whose native language is French. When a token could not be classified clearly as an unreduced or a reduced variant, it was marked as unclear and removed from the analyses. Naming latencies were analysed using the software Praat by manually measuring the time from the onset of the beep that occurred when the object appeared on the screen until the onset of the name of the object produced by the participant. Figure 5.3 summarizes the average values for each of the dependent measures.

In the following analyses, we will focus on the comparison between the novel words that had been learnt as consistently vs. inconsistently reduced. In accu-

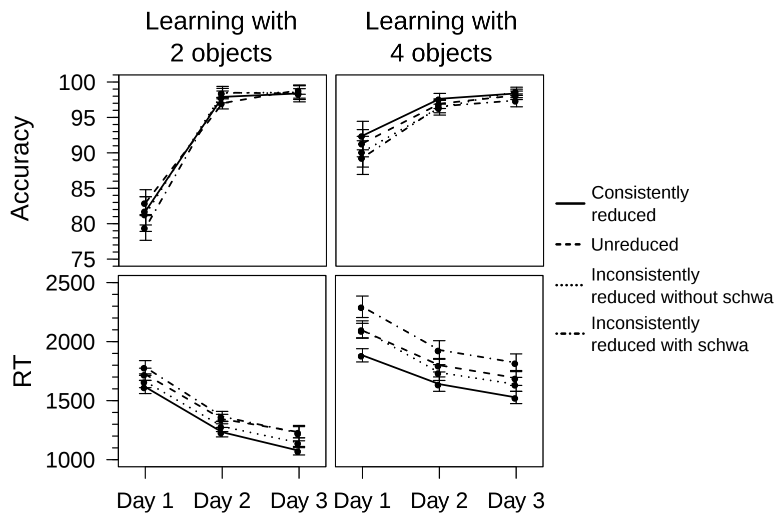


Figure 5.2: Accuracy and reaction times (RT in ms) during the learning blocks of Experiment 1 for each day. Error bars indicate +/- 1 standard error of the mean.

racy, there was no significant difference between consistently and inconsistently reduced words ($\beta_{inconsistent} = 0.39, z = 1.64, p = 0.10$). However, naming latencies were longer for inconsistently compared to consistently reduced words ($\beta_{inconsistent} = 0.11, t = 2.51, p < 0.05$). Importantly, there was a substantial increase in schwa productions for inconsistently reduced words compared to consistently reduced words ($\beta_{inconsistent} = 55.11, z = 4.93, p < 0.001$). In fact, as can be seen in the third panel of Figure 5.3, the probability of producing a word with schwa was approximately as high as the probability of producing a word without schwa for inconsistently reduced words.

Eye-tracking task

The main goal of the eye-tracking task was to examine the influence of phonetic variability on the recognition of reduced word forms. For this purpose, we again focus on the comparison between responses to consistently and inconsistently reduced words.

Accuracy and RT. Click responses and mean RTs for the different conditions are summarized in Table 5.4. For the analysis of RTs, only correct trials were included. The comparison between consistently and inconsistently reduced targets (Condition 1 vs. 2) shows no difference in accuracy ($\beta_{inconsistent} = -0.22$, $z = -0.30$, $p > 0.1$) but slower responses to inconsistently compared to consistently reduced tar-

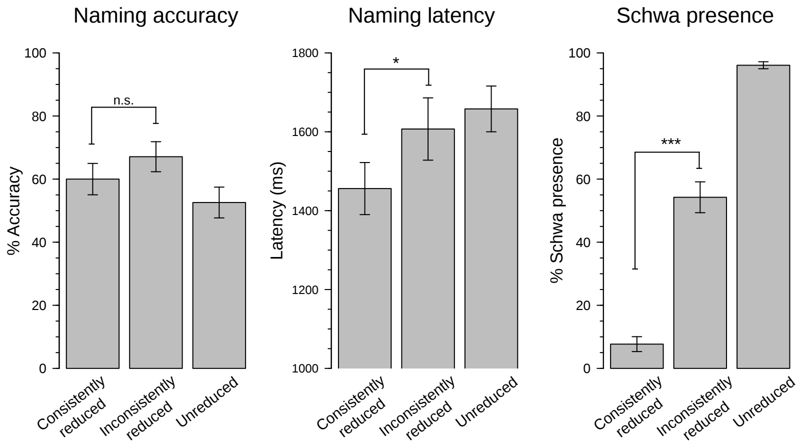


Figure 5.3: Results of the naming task in Experiment 1. Error bars indicate +/- 1 standard error of the mean (n.s.=not significant, * indicates that p<0.05 and *** indicates that p<0.001.)

gets ($\beta_{inconsistent} = 0.06$, $t = 3.51$, $p < 0.001$). Note that we only analysed responses to inconsistently reduced targets that were produced without schwa in order to keep the amount of phonetic overlap between target and competitor constant across conditions. The comparison between consistently and inconsistently reduced competitors (Condition 3 vs. 4) showed that accuracy was lower for targets presented together with inconsistently reduced competitors than for targets presented together with consistently reduced competitors ($\beta_{inconsistent} = -1.35$, $z = -2.71$, $p < 0.01$). However, there was no difference in RT ($\beta_{inconsistent} = 0.01, t = 0.36, p > 0.1$). Furthermore, the comparison between Conditions 2 and 5 showed that inconsistently reduced targets were recognized less accurately ($\beta_{schwapresent} = -1.49$, $z = -2.38$, $p < 0.05$) and more slowly ($\beta_{schwapresent} = 0.10$, $t = 3.9$, $p < 0.001$) when they were produced with schwa (and hence where the competitor overlapped more strongly) than when they were produced without schwa.

Gaze probability. In order to analyze the eye movements that participants made during the eye-tracking task, gaze probabilities for 50-millisecond intervals were calculated. For each time bin, gaze probability was calculated by dividing the number of trials during which participants were looking at a particular object during this period of time by the total number of trials during which participants had not yet made a mouse click at that time. Fixations were categorized as being directed towards a particular object if they fell into a square of 400 by 400 pixels (approximately 10 by 10 cm) around the center of the object's position. The pictures of the objects were

	Condition Target	Competitor	RT	Target clicks	Competitor clicks	Distractor clicks
1	Consistently reduced	Unreduced	2,366 (324)	94.84 (6.52)	4.84 (6.26)	0.32 (1.25)
2	Inconsistently reduced, presented without schwa	Unreduced	2,502 (429)	95.16 (7.24)	4.84 (7.24)	0.00 (0.00)
3	Unreduced	Consistently reduced	2,542 (406)	93.23 (8.71)	5.81 (7.20)	0.97 (3.01)
4	Unreduced	Inconsistently reduced	2,588 (374)	84.84 (9.96)	14.84 (9.96)	0.32 (1.80)
5	Inconsistently reduced, presented with schwa	Unreduced	2,751 (441)	85.48 (14.34)	14.19 (14.09)	0.32 (1.80)

Table 5.4: Mean RTs (in ms) and click responses (in percentages) in the eye-tracking task of Experiment 1. For RTs, only correct trials (i.e. target clicks) were included. For the distractor clicks the sum of both distractors was computed. Standard deviations are shown in parentheses.

approximately 245 by 245 pixels in size (approximately 6.5 by 6.5 cm). Gaze probabilities and RT distributions for each of the five conditions in the eye-tracking task are shown in Figure 5.4.

In order to measure the amount of activation of the target word relative to the competitor word, we computed the difference between target and competitor gaze probability. This measure captures fixations to the target and the competitor and allows us to conduct a single analysis for both measures. The results of the statistical analysis are shown in Table 5.5 and model plots are shown in Figure 5.5. The target-competitor gaze-probability differences were smaller for inconsistently reduced target words than for consistently reduced target words (Condition 1 vs. 2), demonstrating that inconsistently reduced words are recognized less efficiently compared to consistently reduced ones. There was no significant effect on target-competitor differences in the conditions in which the competitor words were either consistently or inconsistently reduced (Condition 3 vs 4). Furthermore, we observed smaller target-competitor differences for inconsistently reduced targets that were produced with schwa compared to those produced without schwa (Condition 2 vs. 5). This latter result is in line with previous studies that show that (particularly word-initial) phonological overlap with competitor words makes word recognition more difficult,

Fixed effects	Consistently vs. inconsistently reduced targets (Condition 1 vs. 2) β	t	p	Consistently vs. inconsistently reduced competitors (Condition 3 vs. 4) β	t	p	Inconsistently reduced targets produced with vs. without schwa (Condition 2 vs. 5) β	t	p
Intercept	0.29			0.20			0.14		
Linear	0.96	10.48	<0.001	0.48	4.77	<0.001	0.61	6.10	<0.001
Quadratic	0.20	2.96	<0.01	0.16	2.19	<0.05	0.29	4.42	<0.001
Cubic	-1.19	-4.50	<0.001	-0.03	-0.55	>0.10	-0.03	-0.57	>0.10
Consistency (inconsistent)	-0.08	-2.74	<0.01	0.06	1.83	>0.05	–	–	–
Schwa presence (without)	–	–	–	–	–	–	0.07	2.70	<0.01
Linear × consistency/presence	-0.11	-0.97	>0.10	0.26	1.95	>0.05	0.24	1.79	>0.05
Quadratic × consistency/presence	0.02	0.24	>0.10	0.01	0.10	>0.10	-0.07	-0.81	>0.10
Cubic × consistency/presence	0.12	2.21	<0.05	-0.02	-0.27	>0.10	-0.04	-0.70	>0.10
Random effects	SD_{PP}	$SD_{PP\times Cons}$		SD_{PP}	$SD_{PP\times Cons}$		SD_{PP}	$SD_{PP\times Cons}$	
Intercept	0.13	0.11		0.12	0.13		0.14	0.10	
Linear	0.24	0.44		0.21	0.50		0.15	0.52	
Quadratic	0.23	0.28		0.19	0.35		0.16	0.30	
Cubic	0.09	0.18		0.10	0.29		0.08	0.19	

Table 5.5: Parameters of the growth-curve models for target-competitor differences in gaze probability in Experiment 1. PP = Participant, PP×cons = PP×consistency. Models plots are shown in Figure 5.5.

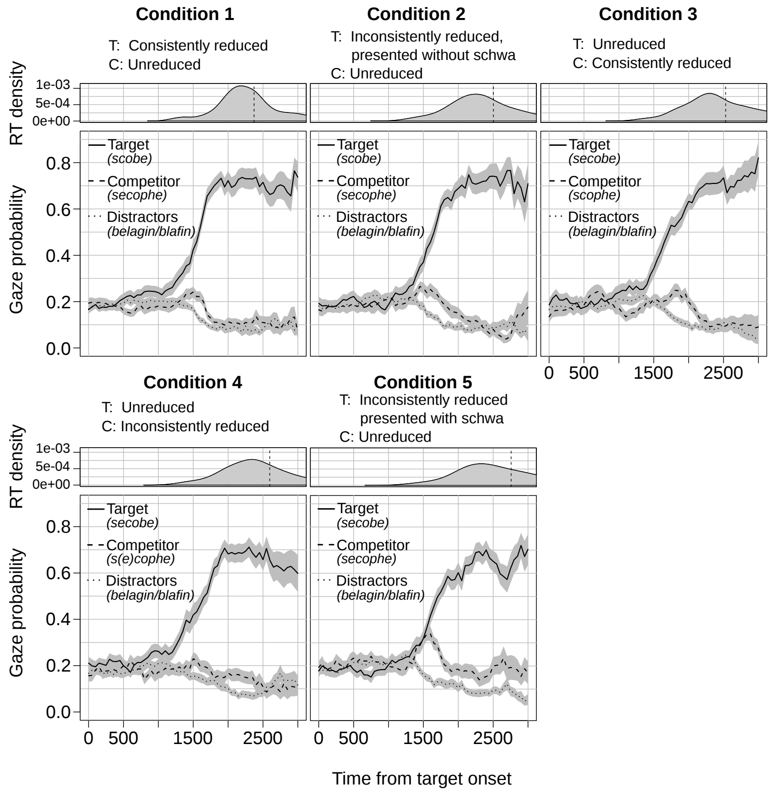


Figure 5.4: Time course of gaze probability and RT distributions in the eye-tracking task (day 3) of Experiment 1. T = target, C = competitor. The grey areas around the black lines indicate +/- 1 standard errors around the participant means. The dashed vertical line on the RT distributions shows the mean RT for that condition.

and thus confirms the validity of the visual-world paradigm as used here.

5.2.3 Discussion

The results of Experiment 1 show that the consistency with which phonetically reduced pronunciation variants are presented to listeners influences how words are processed. The influence of phonetic variability already emerged during the learning phase of the experiment. Inconsistently reduced words were identified more slowly than consistently reduced or unreduced words. This finding is in line with the results of the eye-tracking task which showed that listeners recognized inconsistently reduced words more slowly than consistently reduced words.

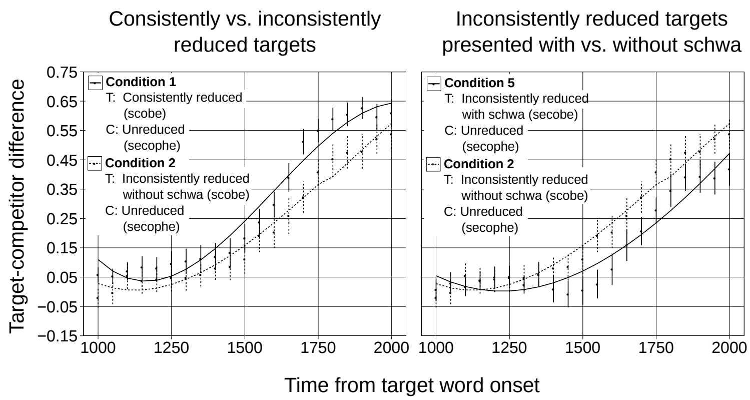


Figure 5.5: Growth-curve model plots (lines) and data (points with error bars) for target gaze probabilities in Experiment 1. The parameters of the models are summarized in Table 5.5. T = target, C = competitor.

Both the RTs and the gaze probability plots show that participants made their responses relatively late. The lateness of the responses is likely to be due to conservative response behaviour that arose because task difficulty increased after the learning phase. Whereas members of the same minimal pair were never presented together during the learning phase, during the test phase participants were presented with both members of a minimal pair at once. This may have made participants more conservative. Previous work suggests that the amount of effort that listeners put into a word-learning task depends on the difficulty of the task. The amount of information that listeners encode about a word may be limited to how relevant that information is for the task they have to perform. For example, Italian participants have been shown to learn to distinguish between words that differ only in lexical stress, but do so only if members of the same minimal stress pair (e.g., [ˈbi.nu.lo] vs. [bi.ˈnu.lo]) are presented as response alternatives during the same trial (Sulpizio and McQueen, 2012) but not if they are presented on separate trials (Sulpizio and McQueen, 2011). This shows that word learners do not automatically pay attention to acoustic details that are not relevant for the task. Similarly, in our experiment participants may have put relatively little effort into the learning tasks because the words were easy to distinguish. In the eye-tracking task, this changed abruptly which may have encouraged participants to become more cautious.

The most important result, however, was provided by the naming task. Participants were significantly more likely to produce a word with a schwa if it was inconsistently reduced during learning than if it had been consistently reduced. Furthermore, it took participants more time to produce the name of an inconsistently reduced object compared to a consistently reduced one. The fact that reduction consistency influenced performance in the eye-tracking as well as the naming task indicates that both word recognition and word production are influenced by exposure to phonetically variable input.

The effect of phonetic variation that we observed on word production was considerably larger than the effect of spelling reported in previous work (Bürki et al., 2012). In that study, the effect of spelling on the percentage of reduced words produced with schwa varied between 2–4%. In our study, in contrast, the effect of phonetic consistency on naming performance was almost 50%. This suggests that the influence of phonetic information outweighs the influence of orthographic information when participants learn phonological forms of new words. In contrast to orthography, phonetic information provides direct evidence for the presence of schwa. Orthographic information requires listeners to draw inferences about how a word sounds based on grapheme-phoneme rules that may not be very reliable (e.g., Ziegler et al., 1996).

However, it is also possible that the difference in the size of the effects is due to differences in the materials and the experimental method that were used. In particular, whereas Bürki et al. used both legal and illegal onset clusters in their novel words, we used only legal onset clusters. Experiment 2 makes it possible to make a tighter comparison between the effects of phonetic variability and orthographic information.

5.3 Experiment 2

In Experiment 2, we investigated how orthographic information influences the way in which listeners produce and recognize reduced pronunciation variants of novel words. The goal of this experiment was to replicate the effect of spelling on production found in Bürki et al. (2012). This allowed us to compare the effect of phonetic variability that we observed in Experiment 1 with the influence that spelling has on the phonological processing of new words.

In Experiment 2, the novel words were always consistently reduced or unreduced. An orthographic exposure phase was added. During this task participants were presented with four orthographic forms of novel words in each trial and asked to click on one of the words. Crucially, the spelling of the reduced words (e.g., *scobe*) either contained the letter <e> and thereby indicated the presence of a schwa (as in <secobe>) or it did not contain the letter <e> (as in <scobe>). According to the results obtained by Bürki et al. (2012), participants ought to be more likely to produce reduced words with a schwa if the word was spelled with <e> than if it was spelled without <e>. If differences between phonetic and orthographic exposure do not have a significant impact, we expect to find a spelling effect of approximately 50% with the same materials as in Experiment 1. If there are differences between the two forms of exposure, however, then we ought to find a much smaller effect similar in size to the effect found by Bürki et al. (who report an effect of approximately 5%).

5.3.1 Method

Participants

Participants were 36 students at the University of Geneva. They were either reimbursed by course credit or a combination of money and course credit. The mean age was 23 years. Seven of the participants were male. All were native speakers of French, none had known hearing problems, and all had normal or corrected-to-normal visual acuity.

Design and Materials

We used the same auditory stimuli and pictures of non-existing objects as in Experiment 1. In contrast to Experiment 1, however, we included an orthographic exposure phase. During this phase, participants were presented with the orthographic forms of four novel words on each trial and were instructed to click on one of the words (e.g., *Click on the secobe*, see panel D of Figure 5.1). The design of the learning phase and the naming task consists of the conditions *reduced and spelled without <e>*, *reduced but spelled with <e>*, and *unreduced and spelled with <e>* (see Table 5.1). Note that words that were always heard as unreduced during training were always spelled with <e> in the orthographic exposure phase.

In the eye-tracking task, there were four conditions. In Condition 1, the target was a reduced word and spelled without <e> (e.g., *scobe* spelled as <scobe>) while the competitor was an unreduced word (e.g., *secophe* spelled as <secophe>). In Condition 2, the target was again a reduced word but this time spelled with <e> (e.g., *scobe* spelled as <secobe>) while the competitor was again an unreduced word. In Condition 3, the target was an unreduced word (e.g. *secobe* spelled as <secobe>) and the competitor was a reduced word that was spelled without <e> (e.g., *scophe* spelled as <scophe>). In Condition 4, the target was also an unreduced word while the competitor was a reduced word that was spelled with <e> (e.g., *scophe* spelled as <secophe>). See Table 5.1 for an overview of the experimental conditions and a comparison of the conditions in Experiments 1 and 2. Note that the unreduced condition in Experiment 2 is equivalent to the unreduced condition in Experiment 1. The condition in which reduced novel words were spelled without <e> is comparable to the consistently reduced condition in Experiment 1. The condition in which a reduced novel word was spelled with <e> is equivalent to the inconsistently reduced condition. Based on the assumption that words with a variable form are harder to recognize than words with a consistent form, we predicted that listeners would take longer to recognize the target word in Condition 2 compared to Condition 1 and be faster at recognizing the target word in Condition 4 compared to Condition 3.

As in Experiment 1, the novel words in all three tasks (learning, naming, and eye-tracking) were rotated through every experimental condition such that each word occurred in each condition but each participant would encounter a given novel word only in a single condition. In the learning and naming tasks, 10 items were reduced and spelled without the letter <e>, 10 were reduced but spelled with <e>, and 20 items were unreduced and spelled with <e>. In the eye-tracking task, there were 10 items in each of the four conditions. Within the experimental blocks, items were presented in a random fashion.

Procedure

The procedures of the different trial types are identical to those used in Experiment 1. In addition to the trial types from Experiment 1, an orthographic exposure phase was added. The orthographic exposure phase was similar to the learning task with four objects and differed in only two aspects. First, participants saw the objects' orthographic forms instead of pictures of the objects. Second, there was only visual feedback (i.e., the correct orthographic form remained on the screen while the other forms disappeared), but there was no auditory feedback (see Figure 5.1). The duration of the print feedback was the same as the pictorial feedback in the two learning tasks. The block structure is similar to Experiment 1 with the exception that the orthographic exposure blocks were added (see Table 5.3).

Apparatus

The apparatus used was the same as in Experiment 1.

5.3.2 Results

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

Learning phase

Mean accuracy scores and RTs for the learning blocks are shown in Figure 5.6. Overall, the accuracy results show that participants quickly learnt to select the correct novel objects. Accuracy was already at ceiling after the first learning session. RTs improved continuously over the three learning sessions.

In order to investigate if participants were sensitive to the mapping between spelling and sound, we analysed the RTs in the orthographic exposure task for reduced words. There was a significant effect of target spelling on the recognition of reduced words,

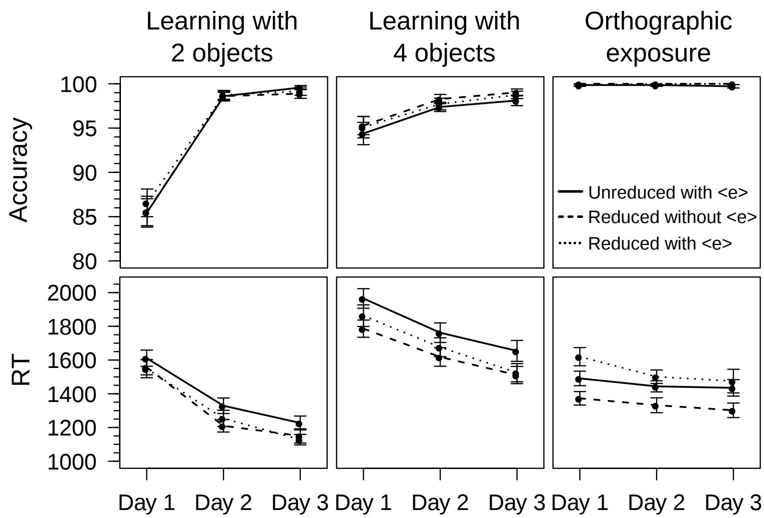


Figure 5.6: Accuracy and reaction times (RT in ms) during the learning and orthographic exposure blocks of Experiment 2 for each day. Error bars indicate +/- 1 standard error of the mean.

indicating that participants responded more slowly if the target was spelled with the letter <e> compared to when it was spelled without it ($\beta_{spelling\ with\ <e>} = 0.13$, $t = 8.71$, $p < 0.001$). In order to further examine this effect, we conducted an analysis in which we included the initial consonant as an additional predictor variable. Initial consonant was coded as a binary variable with the levels /k/ and *other* (consisting of the phonemes /b,f,p,s/). The reason for this coding was that the grapheme corresponding to /k/ changes depending on whether or not the word contains a schwa. If the word contains a schwa, /k/ is mapped onto the grapheme <qu> (e.g. *querage* [kəʀaʒ]) but if the word does not contain a schwa /k/ is mapped onto the grapheme <c> (e.g. *crage* [kʀaʒ]). The phoneme-grapheme mapping for the other consonants does not depend on whether or not the word contains a schwa. We found a significant interaction between spelling and initial consonant ($\beta_0 = 7.20$, $\beta_{spelling \times initial\ consonant} = 0.15$, $t = 4.21$, $p < 0.001$). Separate analyses for the two types of initial consonants showed that there was an effect of spelling for novel words that start with /k/ ($\beta_0 = 7.17$, $\beta_{spelling\ with\ <e>} = 0.25$, $t = 6.51$, $p < 0.001$) as well as for the novel words that start with different consonants ($\beta_0 = 7.20$, $\beta_{spelling\ with\ <e>} = 0.10$, $t = 8.03$, $p < 0.001$). The difference in beta values indicates that the spelling effect is larger for words starting with /k/ than for words that do not. These results suggest that participants were sensitive to the mismatch between orthography and sound. The interaction between spelling and initial consonant shows that participants were influenced more strongly by the mismatch if not only the letter <e> but also the letter corresponding to the first consonant changed.

Naming task

Vocal responses were coded for accuracy and schwa presence in the same manner and by the same person as in Experiment 1. Table 5.6 and Figure 5.7 summarize the average values for each of the dependent measures. In all of the following analyses, we included initial consonant as an additional variable. In none of the naming analyses did this variable have a significant effect.

In the following analyses, we will focus on the comparison between reduced novel words that had either been spelled with or without the letter <e>. With respect to naming accuracy, there was no significant effect of spelling ($\beta_{spelling\ with\ <e>} = -0.39$, $z = -1.64$, $p > 0.05$). For naming latencies we found that reduced words that had been spelled with <e> were named more slowly than reduced words without the letter <e> ($\beta_{spelling\ with\ <e>} = 0.10$, $t = 2.62$, $p < 0.01$). With regard to schwa presence, there is a small but significant increase in schwa productions for reduced words spelled with <e> compared to reduced words spelled without <e> ($\beta_{spelling\ with\ <e>} = 1.28$, $z = 2.19$, $p < 0.05$[1]).

Eye-tracking task

The goal of the eye-tracking task was to examine the influence of the orthographic forms presented during the orthographic exposure phase on the recognition of reduced word forms. For this purpose, the following analyses will focus on the comparison between responses to reduced words that were either spelled with or without <e>.

Accuracy and RT. The mean accuracy and RT values during the eye-tracking task are shown in Table 5.7. We fitted models by including spelling (with vs. without <e>) and initial consonant (/k/ or *other*) and the interaction between the two factors as predictor variables. For accuracy, there were no significant effects either for the conditions in which the target word was reduced or for the conditions in which the competitor was reduced.

In the RT analyses, we found no effects of spelling or initial consonant for the trials in which the target word was reduced. However, for the trials in which the competitor was reduced, we found a significant interaction between competitor spelling and

[1]For this model, the only random variable kept in the model was the random intercept for participants. The other random variables were removed because of convergence errors during the model-fitting procedure.

	Accuracy		Schwa productions		Naming latency: Produced with schwa		Naming latency: Produced without schwa	
	/b,f,p,s/	/k/	/b,f,p,s/	/k/	/b,f,p,s/	/k/	/b,f,p,s/	/k/
Reduced and spelled	68.40	75.00	1.48	1.56	2,217	2,273	1,449	1,515
without <e>	(25.61)	(34.85)	(5.00)	(8.84)	(1,265)	(0)	(387)	(667)
Reduced but spelled	65.58	63.89	6.63	1.67	1,621	2,072	1,595	1,701
with <e>	(22.78)	(37.05)	(14.26)	(9.13)	(632)	(0)	(470)	(944)
Unreduced and	57.14	61.81	94.36	99.26	1,566	1,607	1,786	1,429
spelled with <e>	(29.05)	(31.33)	(12.57)	(4.29)	(498)	(526)	(914)	(0)

Table 5.6: Results of the naming task in Experiment 2. Standard deviations are shown in parentheses. Accuracy and schwa production rates are given in percentages, mean latencies are shown in milliseconds. See Figure 5.7 for the results collapsed across initial consonant.

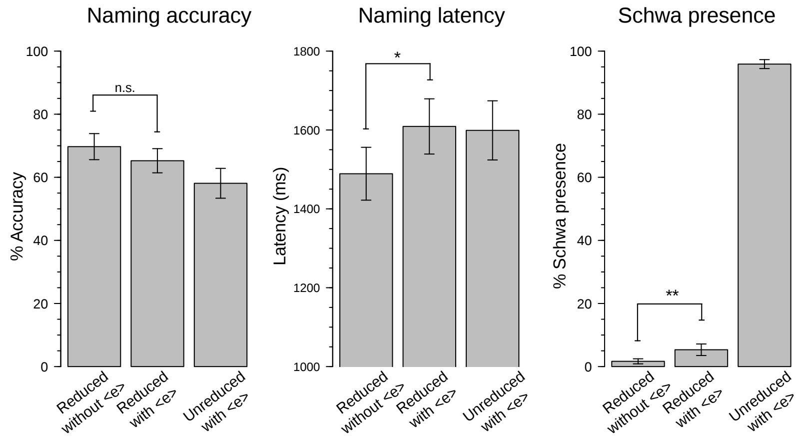


Figure 5.7: Results of the naming task in Experiment 2. Error bars indicate +/- 1 standard error of the mean. For a more detailed division of the results see Table 5.6 (n.s.=not significant, * indicates that p<0.05 and ** indicates that p<0.01).

the initial consonant ($\beta_0 = 7.84$, $\beta_{spelling \times initial\ consonant} = 0.08$, $t = 2.2$, $p < 0.05$[2]). In order to examine the interaction effect, separate models were fitted for competitor words starting with /k/ and competitor words starting with a different consonant. These models show that reduced competitors that had been spelled with the letter <e> slowed down responses compared to competitors that had been spelled without <e> if they started with /k/ ($\beta_0 = 7.76$, $\beta_{spelling\ with\ <e>} = 0.09$, $t = 2.23$, $p < 0.05$). For example, the unreduced target [kəʀal] (spelled as <querale>) was recognized more slowly if the reduced competitor [kʀaʒ] had been spelled with <e> (as in <querage>) than if it had been spelled without <e> (as in <crage>). In contrast, there was no effect of spelling if the competitor started with a different consonant ($\beta_0 = 7.85$, $\beta_{spelling\ with\ <e>} = -0.01$, $t = -0.5$, $p > 0.1$). For example, the unreduced target [səkɔb] (that had been spelled as <secobe>) was recognized equally fast whether the reduced competitor [skɔf] had been spelled with <e> (as in <secophe>) or spelled without <e> (as in <scophe>).

Gaze probability. The time course of gaze probability is depicted in Figure 5.8. For the trials in which the target word was reduced (and the competitor unreduced), we found no effects of spelling or initial consonant. In contrast, for the trials in which the competitor was reduced (and the target unreduced), there was a significant interaction between competitor spelling and initial consonant (see Table 5.8). In

[2]For this model, we removed the random slopes for words for the initial consonant and the interaction between initial consonant and spelling because of convergence errors during the model-fitting procedure.

	Condition		RT		Target clicks		Competitor clicks		Distractor clicks	
	Target	Competitor	/b,f,p,s/	/k/	/b,f,p,s/	/k/	/b,f,p,s/	/k/	/b,f,p,s/	/k/
1	Reduced, without e	Unreduced, with e	2,460 (337)	2,446 (485)	93.40 (10.13)	95.83 (14.02)	5.21 (8.11)	4.17 (14.02)	1.39 (3.98)	0 (0.0)
2	Reduced, with e	Unreduced, with e	2,533 (398)	2,432 (545)	95.14 (10.47)	98.61 (8.33)	3.82 (7.21)	1.39 (8.33)	1.04 (4.60)	0 (0.0)
3	Unreduced, with e	Reduced, without e	2,651 (389)	2,445 (450)	95.14 (7.49)	98.61 (8.33)	3.47 (5.68)	1.39 (8.33)	1.39 (3.98)	0 (0.0)
4	Unreduced, with e	Reduced, with e	2,680 (435)	2,726 (576)	92.71 (14.13)	97.22 (11.62)	5.90 (9.68)	2.78 (11.62)	1.39 (6.53)	0 (0.0)

Table 5.7: Mean reaction times (RT) and click responses in the eye-tracking task of Experiment 2. Unreduced novel words were always spelled with <e>. RTs are given in milliseconds and click responses in percentages. Standard deviations of participant means are indicated in parentheses.

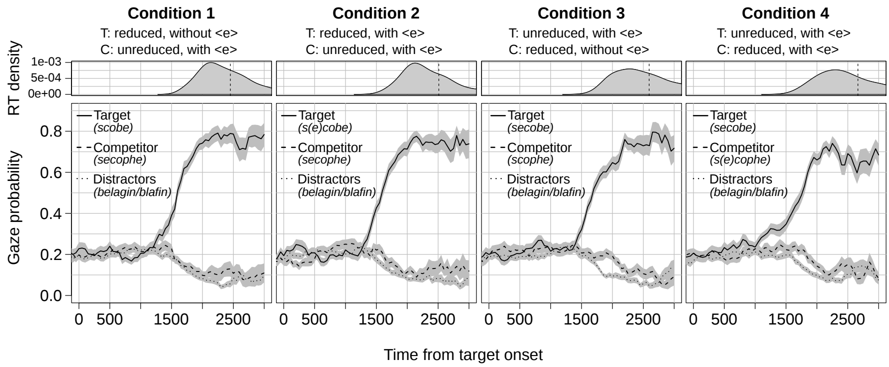


Figure 5.8: Time course of gaze probability and RT distributions in the eye-tracking task (day 3) of Experiment 2. T = target, C = competitor. The grey areas around the black lines indicate +/- 1 standard errors around the participant means. The dashed vertical line on the RT distributions shows the mean RT for that condition.

order to examine this interaction, separate models were fitted for items starting with /k/ and items starting with another consonant (see Table 5.9 for the model parameters and Figure 5.9 for the model plots). For /k/ items (e.g., *querale-querage*) there was a main effect of spelling indicating fewer target fixations relative to the competitor when the competitor had been spelled with the letter <e> compared to when it had been spelled without <e>. In contrast, for items not beginning with /k/ (e.g., *secobe-secophe*) there was no effect of spelling[3].

5.3.3 Discussion

The first goal of Experiment 2 was to replicate the effect of spelling on the production of schwa-reduced words found previously by Bürki et al. (2012) in order to compare it with the influence of phonetic variation which we found in Experiment 1. For this purpose, we examined whether the presentation of orthographic forms that contain the letter <e> in a position in which it typically signifies the presence of the vowel schwa encourages participants to treat the new words as if they contain a schwa even if the phonetic form did not contain one. The results of our naming task indicated that this does indeed appear to be the case. Participants were more likely to produce reduced novel words with a schwa (i.e., in an unreduced manner) if the word had been spelled with the letter <e> during the orthographic exposure phase than if it had been spelled without it. Although this effect is quite small (i.e., <5%) it

[3]Model fitting with three time components did not succeed due to convergence errors, suggesting overfitting. Therefore, the third (cubic) component was removed.

Fixed effects	β	t	p
Intercept	0.14		
Linear	0.71	4.51	<0.001
Quadratic	0.20	1.89	>0.05
Cubic	-0.13	-1.68	>0.05
Competitor spelling (with e)	0.07	1.16	>0.10
Initial consonant (/k/)	0.15	2.47	<0.05
Linear × spelling	-0.11	-0.54	>0.10
Quadratic × spelling	0.07	0.47	>0.10
Cubic × spelling	0.24	2.19	<0.05
Linear × initial consonant	0.27	1.29	>0.10
Quadratic × initial consonant	-0.10	-0.70	>0.10
Cubic × initial consonant	-0.08	-0.75	>0.10
Competitor spelling × initial consonant	-0.22	-2.63	<0.01
Linear × competitor spelling × initial consonant	-0.08	-0.28	>0.10
Quadratic × competitor spelling × initial consonant	0.11	0.50	>0.10
Cubic × competitor spelling × initial consonant	-0.08	-0.52	>0.10
Random effects	SD_{PP}	$SD_{PP \times spell. \times cons}$	
Intercept	0.09	0.25	
Linear	0.32	0.87	
Quadratic	0.03	0.59	
Cubic	0.04	0.41	

Table 5.8: Parameters of the growth-curve models for target-competitor differences in gaze probability in Experiment 2. PP=Participant, spell=spelling, cons=consonant.

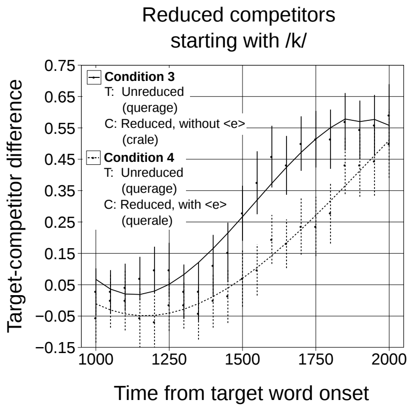


Figure 5.9: The effect of competitor spelling on target-competitor differences in Experiment 2 for words beginning with /k/. Points indicate participant means (error bars indicate +/- 1 standard error) and lines indicate growth-curve model plots. The parameters of the models are summarized in Table 5.9. T=target, C=competitor.

is statistically significant and its size is comparable to the effect reported by Bürki et al. Furthermore, we found that reduced words that were spelled with <e> were produced more slowly than reduced words spelled without <e>. Both of these findings replicate Bürki et al.'s results and thus corroborate the claim that orthographic forms can influence the way in which reduced word forms are processed.

The second goal of Experiment 2 was to investigate if exposure to orthographic forms does not only influence word production but also word recognition. Although we found some evidence suggesting that orthography influences the way in which participants recognize words, the results are not as straightforward as the naming results and differ from the pattern of results that we found in Experiment 1. Based on the naming results, we would have expected to find that reduced target words spelled with <e> are recognized more slowly compared to reduced target words spelled without <e>. However, we found that the way in which reduced targets had been spelled did not influence how quickly they were recognized. Instead, we found that the spelling of reduced *competitors* influenced how quickly (unreduced) target words were recognized. This was demonstrated by the interaction between reduced competitor spelling and initial consonant. For words starting with a /k/, target recognition was slower if the competitor was spelled with the letter <e> compared to when it was spelled without it. In contrast, for the novel words starting with a different consonant than /k/, the factor spelling had no effect. The interaction between initial consonant and spelling is likely to be related to the fact that the grapheme

	Novel words with /k/			Novel words with /b,f,p,s/		
Fixed effects	β	t	p	β	t	p
Intercept	0.29			0.17		
Linear	0.99	4.92	<0.001	0.76	8.35	<0.001
Quadratic	0.10	0.73	>0.10	0.17	2.89	<0.01
Cubic	-0.21	-2.24	<0.05	-0.04	-1.11	>0.10
Spelling (with e)	-0.15	-2.15	<0.05	0.02	0.71	>0.10
Linear × spelling	-0.20	-0.75	>0.10	-0.10	-1.03	>0.10
Quadratic × spelling	0.17	0.93	>0.10	0.08	1.01	>0.10
Cubic × spelling	0.16	1.28	>0.10			
Random effects	SD_{PP}	$SD_{PP\times spell}$		SD_{PP}	$SD_{PP\times spell}$	
Intercept	0.13	0.30		0.10	0.14	
Linear	0.43	1.09		0.36	0.39	
Quadratic	0.15	0.73		0.07	0.32	
Cubic	0.22	0.44				

Table 5.9: Growth curve model parameters for target-competitor differences in gaze probability for words beginning with /k/ and /b,f,p,s/ in Experiment 2. PP=Participant, spell=spelling. See Figure 5.9 for model plots.

corresponding to /k/ changes depending on whether or not the following sound is a vowel. If the following sound is a vowel, the phoneme /k/ is represented by the grapheme <qu> (as in *querage*). In contrast, if the following sound is a consonant, /k/ is represented by the grapheme <c> (as in *crage*). As a result, the change in orthography that went along with the insertion of the letter <e> was considerably larger for words beginning with /k/ compared to words beginning with a different consonant. It is possible that participants were more likely to notice the mismatch between orthography and sound if a word started with a /k/ than if it did not. This hypothesis is supported by the RTs during the learning phase. Listeners slowed down if there was a mismatch between orthography and sound suggesting that they noticed it. Importantly, participants slowed down more for /k/ words than for words starting with a different consonant.

The finding that /k/-initial competitor words made the recognition of unreduced targets harder if spelled with <e> than without it might suggest that listeners stored a schwa-containing representation of the reduced word that competes with the unreduced target. In Condition 4, the target is unreduced (e.g., [kəʀal], spelled <querale>) and the competitor is reduced but spelled with <e> (e.g., [kʀaʒ], spelled <querage>). If participants used the orthographic form of the competitor to infer the presence of a schwa, they may have stored the competitor as /kəʀaʒ/ instead of (or in addition to) /kʀaʒ/. In that case, word-initial overlap between the unreduced target and the unreduced representation of the competitor word would be four segments: /kəʀal-kəʀaʒ/. Therefore, during recognition participants would have had to wait longer (at least until the fifth segment) before a decision about the target word could be made. In contrast, in the condition in which the target word was unreduced and the competitor was reduced and had been spelled without <e> (Condition 3), participants may have stored only a reduced version of the competitor word. In that case, the word-initial overlap would be only one segment: /kəʀal-kʀaʒ/. The decision about the target word could therefore be made as soon as the second segment (the schwa) has been heard. This could explain why we observed a slower rise in target gaze probability in Condition 4 compared to Condition 3 for the /k/ items.

In the conditions in which the target was reduced and the competitor was unreduced (Condition 1 and 2) there was no difference between the two spelling conditions with respect to word-initial segmental overlap between the speech signal and the competitor. In both conditions the information in the speech signal is [kʀal] and the phonological representation of the competitor word is [kəʀaʒ]. In both conditions the overlap between the speech signal and the competitor word is only the first segment (/k/) and the competitor word can already be ruled out when the second segment of the target word (/ʀ/) is being heard.

In summary, the pattern of results that we observe in the eye-tracking task of Experiment 2 differs quite a lot from the pattern of results that we found in Experiment 1. Furthermore, in Experiment 2 the naming and recognition results are inconsistent. The analysis that took initial consonant into account suggests that recognition might be influenced by orthography if the mismatch between sound and spelling is relatively large (e.g., <qu> vs. <c>). However, this observation is based on a post-hoc analysis, which makes it difficult to interpret.

It is possible that clear effects of orthography were missed because participants responded very conservatively. As was mentioned earlier, participants responded much more slowly during the eye-tracking task than during the learning tasks. This is likely due to the fact that the objects presented during the eye-tracking task were phonologically more similar compared to the objects presented during learning. The word-recognition results of Experiment 2 are thus inconclusive. Nevertheless, the fact that they differ from the results of the first experiment suggests that the processes by which word learning is influenced by orthography are different from those by which it is influenced by phonetic information.

5.4 General Discussion

The two experiments presented in this study show that both orthographic and phonetic forms can influence the phonological processing of reduced words. However, orthographic forms influence phonological processing of reduced variants to a lesser degree than phonetic forms do. In line with previous work (Bürki et al., 2012), Experiment 2 showed that participants were more likely to produce reduced novel words with a schwa (forms they had never heard before) if the word had been spelled with the letter <e> than if it had been spelled without it. This finding is consistent with work that has shown that learning about words in the visual domain has consequences for on-line processing in the spoken domain, again for words that had never been heard before (Bakker et al., 2014). Experiment 1 showed that participants were more likely to produce a word with a schwa if it had been inconsistently reduced (i.e., it was sometimes produced with and sometimes without schwa) than if it had been consistently reduced. Crucially, the effect of phonetic consistency on naming responses was ten times as big as the effect of spelling. This strongly suggests that variable phonetic input is more likely to influence phonological processing than exposure to orthographic forms.

Previous studies have suggested that orthography plays a significant role in the way pronunciation variants are processed (e.g., Ranbom and Connine, 2007, 2011). In particular, it was proposed that the overlap between spelling and unreduced pronunciation variants could, at least in part, explain why unreduced variants are processed more efficiently than reduced variants even if the reduced variants occur more frequently. Our study suggests that, at least for pronunciation variants that are the result of schwa reduction, the influence of orthography may be smaller than expected. This finding may seem surprising given that speech is inherently more variable than orthographic forms are. However, the crucial point may not be the invariance of orthographic forms per se but rather the consistency of the mapping between graphemes and phonemes. The grapheme-phoneme mapping in French, like English, is relatively inconsistent (e.g., Ziegler et al., 1996). The grapheme that we focused on in the present study, the letter <e>, is no exception. Although this letter is associated with the vowel schwa in the position in which it was placed in the words in our study, there are many instances in French in which this letter does not correspond to the vowel schwa. For example, in Standard French, in the word *jet* "fountain" the letter <e> corresponds to the open-mid vowel /ɛ/ but in the word *et* "and" it is produced as the close-mid vowel /e/. Moreover, in many words such as the word *bracelet* "bracelet" the first letter <e> is silent. Because the letter <e> has many different phonological correspondences, it is not a very reliable cue for the presence of a vowel. Language users may therefore be reluctant to make strong inferences about the phoneme that the letter <e> corresponds to when learning new words. It is thus important to take into consideration the consistency of the grapheme-phoneme mappings when investigating the influence of orthographic information on phonological processing. Possibly, one would observe a larger effect of orthography when examining graphemes with highly consistent mappings (such as <qu> which always signifies the presence of the phoneme /k/).

Our results also speak to the question of what the nature of orthographic effects in spoken language processing is. Although several studies have demonstrated that orthographic knowledge can influence speech recognition, it is not yet clear what the locus of this influence might be. Some studies suggest that orthography might have a post-lexical influence on decision making that emerges only after listeners have already identified a given word based on the auditory input (e.g., Cutler and Davis, 2012). This hypothesis could explain why spelling effects have primarily been observed in meta-linguistic tasks such as lexical decision (e.g., Ziegler and Ferrand, 1998; Ziegler et al., 2004). However, several other studies have demonstrated spelling effects in tasks that make the use of decisional strategies unlikely, suggesting that orthography can influence on-line word recognition at a lexical (or even pre-lexical) locus (e.g., Perre et al., 2011; Taft et al., 2008). While our study cannot solve the debate on how orthographic information influences spoken word processing, it does suggest that the influence of orthographic information on the learning of new words is, at least in French, relatively small compared to the influence of phonetic information. Moreover, the effect of orthographic knowledge on the processing of reduced pronunciation variants might be smaller than expected. This conclusion is consistent with a recent study showing that whether or not a speech sound is orthographically coded does not affect the processing costs that are typically associated with reduced speech (Mitterer and Reinisch, 2015).

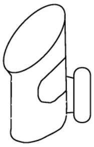
The second question we asked in this study was whether spelling and phonetic variation influence phonological processing during word recognition in addition to word production. In order to investigate this question, we employed the visual-world paradigm in which participants' eye movements were monitored while they were selecting a target word (e.g. *secobe*) in the presence of a phonetic competitor (e.g. *secophe*) and two distractors (e.g. *belagin* and *belafin*). The eye-tracking results of Experiments 1 and 2 both show effects on word recognition. However, the patterns of results were quite different across the two experiments. Whereas in Experiment 1, the recognition results are consistent with the production results, the recognition results in Experiment 2 do not match the pattern of results found in the naming task. The finding that the production and recognition results are consistent in Experiment 1 but not in Experiment 2 suggests that orthographic and phonetic information influence phonological processing by means of different mechanisms. This is plausible because phonetic forms provide direct information about how a word sounds. In contrast, orthography requires the word learner to perform an additional step. The sound structure of a newly-learnt word has to be inferred from the word's spelling. This may not be a simple task as the mapping of graphemes to sounds is often inconsistent. Because of this additional step, it might also take more time for orthographic information to develop an effect on phonological processing. Consistent with this notion is a recent word-learning study that has shown that novel words can influence lexical processing across modalities (Bakker et al., 2014). Words that were learnt in the auditory modality influenced lexical competition in the written modality, and vice versa. However, whereas the influence from sound to print emerged already one day after learning, the influence from print to sound only emerged a week later. While this result shows that there is an exchange of information between the written and the auditory modalities, it also confirms our finding that speech input has an advantage over print.

In addition to showing that there is an exchange of information between the written and the auditory modality, our results also suggest an exchange of information between the production and recognition systems. Experiment 1 showed that learning novel words with variable pronunciations influenced behaviour in the production task as well as behaviour in the eye-tracking task. This finding does not assort well with a psycholinguistic architecture that makes a strict division between the language production and comprehension systems and treats them as informationally incapsulated modules. Instead, the learning of novel words appears to require that production and comprehension processes work closely together and exchange information. Even the relatively passive type of learning procedure that we used in the present study seems to affect both types of processes, which is reflected in our naming and eye-tracking results. Although our results do not allow us to decide whether production and comprehension use different (e.g., Levelt et al., 1999) or the same processes and representations (e.g., Pickering and Garrod, 2013), the similar findings across the production and comprehension tasks in Experiment 1 show that both systems work closely together.

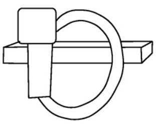
In conclusion, this study provides further support for the notion that orthographic information can influence the phonological processing of schwa-reduced word forms. However, the influence of orthography is, at least in French, outweighed by the influence of phonetic information. This finding has implications for the study of the acquisition of phonological knowledge in general and the processing of reduced pronunciation variants in particular. While orthographic information can influence phonological processing, this influence is comparatively small and may depend on the consistency of the grapheme-phoneme mapping. Furthermore, our study suggests that the influence of orthographic and phonetic information is not limited to word production but also affects word comprehension, suggesting that both processes are closely related.

5.5 Appendix

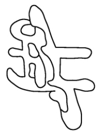
The following shows the pictures and novel words used in Experiments 1 and 2. Each row contains the quadruplets of items used together in a given trial during the eye-tracking experiment. For cases in which the no-schwa variant of an item is spelled differently from the schwa variant (in addition to removing the letter <e>), the orthographic form of the no-schwa variant is shown in parentheses.



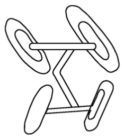
belonte
[bəlɔ̃t]



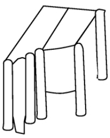
belonge
[bəlɔ̃ʒ]



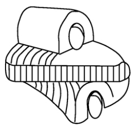
secudeau
[səkydo]



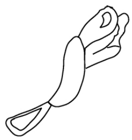
secumeau
[səkymo]



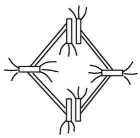
felube
[fəlyb]



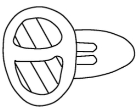
feluge
[fəlyʒ]



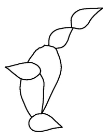
queroussien
(croussien)
[kəʀusjɛ̃]



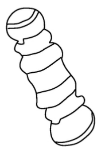
queroumien
(croumien)
[kəʀumɛ̃]



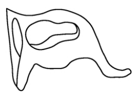
ferole
[fəʀɔl]



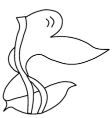
ferone
[fəʀɔn]



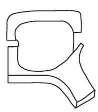
pelouvoi
[pəluva]



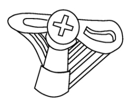
pelouroi
[pəluʀwa]



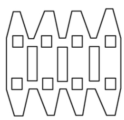
pelèque
[pəlɛk]



pelêve
[pəlɛv]



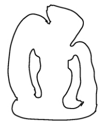
selanquet
[səlɑ̃kɛ]



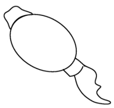
selanret
[ləlɑ̃ʀɛ]

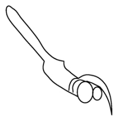

secobe
[səkɔb]

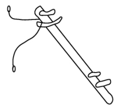

secophe
[səkɔf]

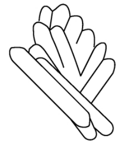

belagin
[bəlaʒɛ̃]


belafin
[bəlafɛ̃]

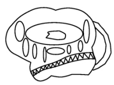

semoude
[səmud]


semour
[səmuʀ]

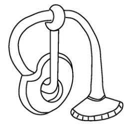

belichou
[bəliʃu]

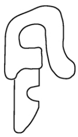

belimou
[bəlimu]

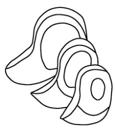

peliteu
[pəlitø]

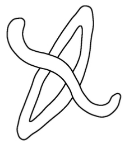

pelimeu
[pəlimø]

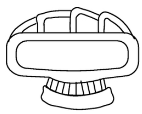

selibe
[səlib]

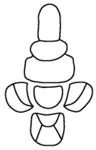

selige
[səliʒ]

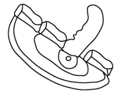

pelingue
[pəlĩg]

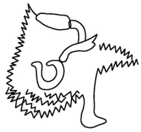

pelinche
[pəlĩʃ]

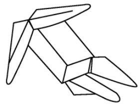

queridan
(cridan)
[kəʀidã]

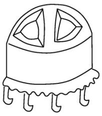

queripan
(cripan)
[kəʀipã]


querale
(crale)
[kəʀal]


querage
(crage)
[kəʀʒ]

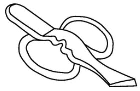

ferimbon
[fəʀĩbõ]


ferinton
[fəʀĩtõ]

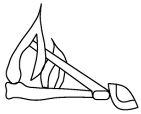


felami
[fəlami]

felaqui
[fəlaki]



querambe
(crambe)
[kəʀãb]



queranle
(cranle)
[kəʀãl]

Conclusions

Chapter 6

This chapter summarizes this thesis and puts its findings in a broader context. In particular, I will discuss implications of the findings presented in the previous chapters and draw conclusions with respect to what these findings can tell us about the role of acoustic reduction in spoken-word processing.

6.1 Summary of results

6.1.1 Chapter 2

Chapter 2 investigated the distribution of reduced word forms in conversational speech. More specifically, the focus was on two types of acoustic reduction that occur frequently in Dutch past participles: schwa reduction in prefixes and /t/ reduction in suffixes. We investigated two types of acoustic reduction: cases in which the schwa or /t/ was completely absent (i.e., categorical reduction) and cases in which the schwa or /t/ was reduced in its duration (i.e. gradient reduction). Our question was whether and under which circumstances reduced word forms co-occur with one another in natural speech. Although the results for schwa and /t/ reduction differed, both analyses suggested that reduced word forms co-occur and showed that speech reductions happen in a systematic way. More specifically, we found that schwas in prefixes of succeeding past participles had similar durations if the time lag between the two participles was short and the past participles were identical. Importantly, this result cannot be attributed to speaking rate because we controlled for this variable in the statistical analysis. For /t/s in suffixes, we found shorter durations if the past participle was preceded by an identical past participle shortly before.

These findings have interesting implications for theories of both speech recognition and speech production. Listeners could, in principle, exploit this systematicity in order to adapt to a reduced speaking style. Co-occurrence of reduced forms can provide listeners with statistical information that might allow them to predict reduced forms based on the occurrence of previously encountered reduced forms. Knowing that upcoming words are likely to be produced in a reduced way would allow listeners to prepare themselves to cope with reduced speech.

Furthermore, the finding that reduced word forms co-occur (even if speaking rate is controlled for) is informative with respect to the mechanisms that underlie the production of reduced speech. Episodic accounts could explain the results in terms of priming of reduced lexical representations. An abstract account might explain the similarity among productions by assuming priming at a post-lexical level, that is, at the level at which motor commands are generated. However, current models have not yet implemented explicit mechanisms that could account for the current results. Advances in modeling work are required in order to provide better computational accounts for how language users process reduced speech in natural conversations.

6.1.2 Chapter 3

Whereas Chapter 2 looked at the production of Dutch past participles, Chapter 3 focused on the recognition of Dutch past participles. We also changed the focus from prediction based on acoustic-phonetic context to prediction based on syntactic information. More specifically, Chapter 3 investigated whether the recognition of spoken past participles is influenced by how predictable they are given their syntactic context. Furthermore, it was investigated whether listeners assign more weight to syntactic predictability in reduced speech compared to carefully produced speech. In three experiments, syntactic predictability was manipulated by varying the word order of past participles and auxiliary verbs in Dutch subordinate clauses. The past participle (e.g., *geleund* "leaned") either preceded the auxiliary (e.g., *heeft* "has") as in *Ik weet zeker dat hij* <u>*heeft geleund*</u> *op de houten tafel* ("I know for sure that he <u>has leaned</u> against the wooden table") or the past participle followed the auxiliary, as in *Ik weet zeker dat hij* <u>*geleund heeft*</u> *op de houten tafel*. Sentences like these were presented to listeners either in a reduced way in which the schwa in the past participle was often times missing, or they were presented in a careful way.

In three eye-tracking experiments, participants recognized past participles more quickly when they occurred after their associated auxiliary verbs than when they preceded them. Response measures tapping into later stages of processing suggested that this effect was stronger for reduced speech than for carefully produced sentences. These findings show that syntactic information helps listeners recognize words. Furthermore, they suggest that syntactic information helps listeners even more when listening to reduced speech. However, our analyses suggest that word order information was not used in order to *predict* the past participle. Although the past participle becomes more predictable when it is preceded by an auxiliary verb, participants did not start looking at the participle before hearing it. This suggests that the auxiliary verb facilitated the recognition of the participle but did not lead participants to predict it in an anticipatory way. The absence of predictive behaviour might be due to the relatively short time interval between the auxiliary verb and the past participle.

6.1.3 Chapter 4

Whereas Chapter 3 investigated how syntactic information influences the processing of reduced speech, Chapter 4 investigated how exposure to reduced speech influences how we process syntactic information. In order to examine whether the way listeners adapt to casual speech influences syntactic processing, this study used the Event-Related-Potential (ERP) technique. More specifically, the question was whether listeners tolerate grammatical gender violations in casual speech. In order to investigate this question, native speakers of Dutch were presented with utterances that contained adjective-noun pairs in which the adjective was either correctly inflected with a word-final schwa (e.g. *een spannende roman* "a suspenseful novel") or incorrectly uninflected without that schwa (*een spannend_ roman*). The incorrectly uninflected adjectives elicited a P600 effect compared to the correctly inflected adjectives when the talker was speaking in a careful manner. This finding is consistent with the ERP literature (e.g., Hagoort and Brown, 1999) and confirms that listeners are sensitive to morphosyntactic violations when listening to careful speech. In contrast, when the talker was speaking in a casual manner, this P600 effect was absent. In order to make sure that the P600 effect was not absent because listeners were not able to understand the content of the utterances, a control condition was included in which participants were presented with semantically anomalous sentences that are known to elicit an N400 effect (e.g., Hanulíková et al., 2012). Carefully and casually produced anomalous sentences both elicited N400 effects, showing that listeners were able to understand the content of both types of utterance. The results show that listeners adapt to the speaking style of a talker when processing the acoustic-phonetic information provided by the speech signal. This adaptation is not limited to the way the acoustic-phonetic information is processed but extends to other levels of processing, specifically it influences morphosyntactic processing.

This finding brings up the question of how exactly this adaptation was accomplished. One possibility is that listeners change the way in which they interpret the absence of inflectional schwa based on the preceding phonological context. As a result of the casual speaking style, many words in a casually produced utterance contain acoustic-phonetic reductions. Listeners might be able to keep track of the probability with which speakers produce and omit individual segments. Based on this knowledge they might interpret the absence of speech sounds differently. The absence of inflectional schwa would therefore not be interpreted as a grammatical error but instead it would be consistent with the fact that a casually speaking talker is likely to omit individual segments. Listeners might therefore change the way they assign perceptual weight to acoustic information during speech recognition depending on the context in which that information is heard. Another possibility is that listeners adapt to the acoustic properties of casual speech. As casual speech is characterized by a high speaking rate and a large amount of coarticulation, it becomes very difficult to determine whether or not certain individual segments (in particular schwas) were actually produced. Listeners might have adapted to this situation by stopping to try to detect whether or not a schwa is present. A future study could investigate which of these possible accounts is correct by manipulating speaking rate, speaking style, and grammaticality in a single experiment.

6.1.4 Chapter 5

Chapter 5 investigated how a word's spelling influences the processing of schwa reduction in a different language than Dutch in which schwa reduction is also very common: French. More specifically, the influence of spelling was compared with the influence of variation in the phonetic input. We asked how large the influence of spelling is in comparison to the influence of phonetic information. Furthermore, we asked if these two factors influence both spoken-word production and spoken-word recognition. In order to investigate these questions, two experiments were conducted using a novel-word learning paradigm. In this paradigm, participants learnt novel French words in which a schwa was either unreduced (i.e. present, e.g. *secobe*) or reduced (i.e. absent, e.g. *scobe*). Learning took place on three consecutive days and after the final learning session on the third day, an eye-tracking task and a picture naming task were performed. The eye-tracking task measured word recognition performance and the naming task measured word production performance. In Experiment 1, listeners were presented with novel words that varied in their pronunciation. Some of the words were produced half of the time with schwa and the other half of the time without schwa. In Experiment 2, the pronunciation of words was consistent but an orthographic exposure phase was included in which the reduced words were either spelled with the letter <e> or without it.

The results from the naming and eye-tracking tasks suggest that both phonetic variation and spelling influenced the way in which schwa reduction was processed and the novel words were learnt. In particular, we found that participants were more likely to produce reduced words with a schwa if these had been presented in an inconsistent way (i.e., sometimes with and sometimes without schwa as in Experiment 1) and if these words had been spelled with the letter <e> (Experiment 2). However, the influence of phonetic variability outweighed the effect of spelling. This might be because orthography provides less reliable cues for the identity of sounds than the sounds do themselves.

It is also possible that there was not enough time for orthographic knowledge to develop a stronger effect on phonological processing. This notion would be consistent with studies that suggest that it takes a considerable amount of time for newly-learned orthographic forms to exert an influence on phonological processing (e.g., Bakker et al., 2014). In this study, participants learned novel words in either the printed or the auditory modality. Subsequently, participants were tested in both modalities for the emergence of lexical competition effects. Whereas the influence from sound to print emerged already one day after learning, the influence from print to sound emerged a week later. It is therefore possible that the effects of spelling that we observed in Chapter 5 would become larger if participants were tested several days, or a week, later. Future work could address this question by testing participants repeatedly across a longer period of time. This way one could determine whether the influence of orthography gains more strength with time and the difference between the effects of sound and spelling becomes smaller.

6.2 Theoretical implications

The findings summarized above are relevant for several topics that are currently debated in the psycholinguistic literature such as prediction in language processing, listener flexibility, and the processing advantage for unreduced speech.

6.2.1 Prediction in language processing

One of the major topics in current cognitive science is the role that prediction plays in cognition in general and in language processing in particular (e.g., Clark, 2013; Pickering and Garrod, 2013). This thesis points to the importance of distinguishing between the a priori predictability of words and actual prediction during on-line processing. Although there is a vast amount of systematic information in the speech signal that could potentially be used in order to predict upcoming input, this does not necessarily mean that listeners *will* or *are able to* exploit this information. Chapter 3 showed that listeners benefit from predictive word-order information: they recognized past participles more quickly if these words were put in a more predictable word order than in a less predictable word order. However, listeners did not use word-order information in order to *predict* the past participles. This was shown by the absence of anticipatory eye movements after the occurrence of predictive auxiliary verbs and prior to the arrival of the past participles. However, whether or not predictive information will be used may not be a categorical issue. That is, it may not be the case that listeners will always use some kind of information for prediction and never use other types of information for prediction. Rather, whether or not a particular piece of information can lead to prediction may depend on the amount of time that is available for listeners. Generating predictions may be a time-consuming process that listeners can only perform if given enough time between the predictive and the predictable word.

This brings up the question of how much time listeners need in order to come up with predictions about what the speaker might say. Future work could investigate this question by varying the amount of time (or words) in between the predictive and the predicted word. A related question is whether listeners differ in the amount of time that they need for prediction. Previous work has shown that individual differences in working memory and processing speed influence how much listeners predict during language processing (e.g., Huettig and Janse, 2015). It would be interesting to investigate whether individual differences in such cognitive abilities also influence how much time listeners need in order to make predictions about upcoming linguistic input.

6.2.2 Speed of adaptation

Another theoretical topic in the current literature concerns the ability of listeners to adapt flexibly to variability in the speech input (e.g., Hanulíková et al., 2012; Poellmann et al., 2014; Samuel and Kraljic, 2009; Witteman et al., 2013). Chapter 4 of this thesis provides additional support for the notion that listeners adapt to reduced speech by showing that the way grammatical gender violations are processed depends on the phonetic context. Furthermore, it corroborates the claim that listeners can adapt quite rapidly to acoustic-phonetic reductions: reduced and unreduced utterances were presented to participants in a randomized fashion, suggesting that listeners adapted to the different speaking styles on a trial-by-trial basis. As the utterances were only a few seconds long, this suggests that listeners are able to tune morphosyntactic processing quite rapidly in response to the acoustic-phonetic properties of the speech input.

6.2.3 Cascaded processing of reduced speech

Chapter 4 also shows that adaptation on one level of linguistic processing (i.e., the acoustic-phonetic level) can have consequences for other levels (i.e., the morphosyntactic level). This finding is important to take into account when thinking about how to model the processing of reduced speech. For example, one could model linguistic processing in a modular system in which acoustic-phonetic reductions are processed in one particular module (e.g., Fodor, 1983). On such an account one might think that the computational problems that arise as a result of acoustic reduction would have to be solved by the phonological module and thus do not influence other modules (e.g., the syntactic module). After having solved the problems associated with acoustic variation, the output of the phonological module would then be the same for reduced and unreduced speech and the operations during the following processing steps would be independent from the type of acoustic input. The results of Chapter 4, however, suggest that this is not the case. Instead, acoustic-phonetic reductions influence the linguistic processing system in a rather cascaded manner (for a review of other evidence for cascaded flow of information in speech recognition see McQueen, 2005). Variation in the acoustic-phonetic properties of reduced speech is not completely "absorbed" at the phonological level but instead influences the way morphosyntactic violations are processed. This suggests that adaptation to reduced speech demands flexibility across multiple levels of linguistic processing. Thus, when modeling the processing of reduced speech, information about reduction is not limited to one particular level. Instead, information is passed forward in cascade and can be used on subsequent levels to optimize higher levels of processing. Processing reduced speech might therefore require flexible adjustment across multiple levels of processing.

6.2.4 Flexible use of different sources of information

Chapter 3 points to another way in which listeners show flexibility when recognizing reduced speech. This study showed that word-order information can help in the recognition of spoken words. Past participles that followed their associated auxiliary verb were recognized more quickly compared to past participles that preceded their auxiliary verb. Importantly, the effect of word order was larger in reduced speech than in careful speech. This suggests that listeners assign more importance to word order information if the acoustic-phonetic information is less reliable. Listeners are therefore flexible in the way in which different sources of information (e.g., phonetic vs. syntactic) are weighted during on-line language processing. This interpretation has parallels with theories of speech segmentation in which listeners use lexical, segmental, and prosodic cues for segmentation but weight these sources of information depending on their availability in the speech signal (e.g., Mattys et al., 2005; Norris et al., 1997).

6.2.5 The processing advantage for unreduced speech

As was pointed out in the introduction, previous studies have shown that reduced word forms are typically processed less efficiently than unreduced words even if the reduced pronunciation variants occur more frequently than the unreduced variants. It has been proposed that this *processing advantage* for unreduced forms is due to the fact that the overlap between sound and spelling is larger for unreduced compared to reduced forms (e.g., Ranbom and Connine, 2007). Another factor that might explain why reduced variants are processed less efficiently than unreduced variants even if the reduced variants are more frequent is that reduced variants are phonetically less distinct. For recognizing spoken words, phonetic distinctiveness might be more important than frequency of occurrence (Pitt, 2009). Furthermore, reduced variants are not only phonetically less distinct but also lexically less distinct. This is because reduced words are, by definition, shorter than unreduced words. Longer words are generally recognized more efficiently than shorter words (Pitt and Samuel, 2006). There are two reasons for this: First, there is more bottom-up evidence in the speech signal that can support the activation of longer words compared to shorter words. And second, shorter words overlap with a larger number of other words compared to longer words. Thus, shorter words have larger lexical neighbourhoods and therefore receive more lexical inhibition during lexical access (Luce and Pisoni, 1998).

Although studies suggest that orthographic knowledge can in fact influence the processing of reduced speech (e.g., Bürki et al., 2012), it is not known how large this influence actually is. If orthographic overlap is the reason for the processing advantage for unreduced pronunciation variants, this would imply that the influence of orthographic information is quite large because it could neutralize the effect of variant frequency. Chapter 5 investigated how large the influence of orthographic information is by comparing it with the influence that variability in the phonetic input has on the learning of novel words. The results showed that the influence of phonetic variation far outweighed the influence of spelling. This suggests that spelling has a rather negligible influence on how reduced speech is processed. It is thus unlikely that only differences in the overlap between sound and spelling can account for the processing advantage of unreduced over reduced pronunciation variants. And it is therefore more plausible that the processing advantage is due to other factors.

6.3 Conclusion

This thesis has shown that spoken-word processing of acoustically reduced speech involves a great deal of flexibility on the part of the language user. This flexibility involves multiple levels of linguistic processing, which suggests that the way we process the acoustic-phonetic properties of reduced speech interacts with how we process other types of linguistic information. Furthermore, this flexibility includes the ability to weight different sources of information (such as acoustic, syntactic, and orthographic information) depending on their availability and reliability. These findings are relevant for our understanding of how we process the great deal of acoustic variability that results from speech reductions. Moreover, these findings are also relevant for our understanding of how humans process acoustic-phonetic variability in general. Understanding how language users process reduced speech might uncover mechanisms that are also relevant for the processing of other types of systematic variation such as, for example, foreign accents or regional dialects. This thesis has made a step into that direction by showing that speech processing in response to variability requires flexible adjustments across several processing stages.

Nederlandse samenvatting

Alledaagse spontane spraak kenmerkt zich door een informele spreekstijl. In deze spreekstijl (of *register*) worden klanken, en soms zelfs hele lettergrepen, verkort of helemaal weggelaten. Sprekers kunnen het woord *geweest* bijvoorbeeld uitspreken zonder de klinker *e* als in *g'weest* of zonder de *t* als in *gewees*. Alle studies in dit proefschrift onderzoeken hoe taalgebruikers deze akoestisch gereduceerde woorden verwerken.

In de eerste studie werd een groot aantal opnames van spontane spraak bestudeerd. Het doel van deze corpusstudie was om meer te weten te komen over de spreiding van de gereduceerde woorden. We onderzochten of gereduceerde varianten van voltooid deelwoorden (bijv., *geweest*) individueel of samen voorkomen. Daarbij focusten we op 2 typen akoestische reducties die veel voorkomen in Nederlandse voltooid deelwoorden: schwa reductie in prefixen en /t/ reductie in suffixen. We onderzochten gevallen waarin de schwa of de /t/ volledig afwezig was (*categoriale reductie*) en gevallen waarin de schwa of /t/ gereduceerd werd in duur (*graduele reductie*). De vraag was of en onder welke omstandigheden gereduceerde woorden samen voorkomen in natuurlijke spraak. Alhoewel de resultaten voor schwa en /t/ reductie verschilden, suggereren beide analyses dat gereduceerde woorden samen voorkomen. Dit laat zien dat reducties op een systematische manier voorkomen. Schwas in prefixen van opeenvolgende voltooid deelwoorden duurden ongeveer even lang als de afstand tussen de twee geproduceerde voltooid deelwoorden kort was en ze identiek waren. Dit resultaat kan niet worden toegekend aan spreeksnelheid, want voor deze voorspeller werd gecontroleerd in de statistische analyses. De /t/ in suffixen was korter als het voltooid deelwoord kort daarvoor voorafgegaan werd door een identiek voltooid deelwoord.

Deze bevindingen hebben belangrijke implicaties voor spraakherkennings- en ook voor spraakproductietheorieën. Luisteraars zouden aan de hand van deze systematiek hun manier van luisteren aan de gereduceerde spreekstijl kunnen aanpassen. Combinaties van gereduceerde woorden in conversaties kan luisteraars statistische informatie verschaffen en stelt hen in om gereduceerde woorden te voorspellen op basis van al eerder gereduceerde woorden. Wanneer luisteraars weten dat de volgende woorden gereduceerd zullen worden uitgesproken, kunnen zij zich voorbereiden op de verwerking ervan.

In hoofdstuk 2 werd ingegaan op de productie van Nederlandse voltooid deelwoorden, terwijl in hoofdstuk 3 werd gekeken naar de perceptie van Nederlandse voltooid deelwoorden. In hoofdstuk 3 werd onderzocht of de perceptie van voltooid deelwoorden wordt beïnvloed door hoe voorspelbaar ze zijn gegeven de syntactische context. Verder werd onderzocht of luisteraars meer gewicht toekennen aan syntactische voorspelbaarheid in gereduceerde spraak dan in nauwkeurig geproduceerde spraak. In drie experimenten werd de syntactische voorspelbaarheid gemanipuleerd door de woordvolgorde van de voltooid deelwoorden en hulpwerkwoorden in de Nederlandse bijzinnen te variëren. Het voltooid deelwoord (bijv. *geleund*) volgde op het hulpwerkwoord (bijv. *heeft*, als in *Ik weet zeker dat hij heeft geleund op de houten tafel*) of ging eraan vooraf (*Ik weet zeker dat hij geleund heeft op de houten tafel*). Zinnen werden gepresenteerd in gereduceerde vorm, waarbij de schwa in het voltooid deelwoord ontbrak of in ongereduceerde vorm waarbij de schwa wel werd geproduceerd.

Met behulp van eye-tracking, een methode die meet waar mensen naar kijken terwijl ze luisteren, werd gemeten hoe deze zinnen werden verwerkt door luisteraars. De luisteraars herkenden de voltooid deelwoorden sneller als deze achter de bijbehorende hulpwerkwoorden stonden dan wanneer ze ervoor stonden. Resultaten met betrekking tot de verwerking op een later moment suggereren dat dit effect sterker was voor gereduceerde spraak dan voor zorgvuldig geproduceerde zinnen. Deze bevindingen laten zien dat syntactische informatie de luisteraar helpt woorden te herkennen. Verder suggereren ze dat syntactische informatie de luisteraar zelfs meer helpt wanneer ze naar gereduceerde spraak luisteren. De analyses laten zien dat woordvolgorde niet gebruikt wordt bij het voorspellen van het voltooid deelwoord. Hoewel het voltooid deelwoord voorspelbaarder is wanneer het voorafgegaan wordt door een hulpwerkwoord, kijken deelnemers niet naar het voltooid deelwoord alvorens ze het gehoord hebben. Dit suggereert dat het hulpwerkwoord de herkenning van het voltooid deelwoord vergemakkelijkte, maar dat het de deelnemers niet aanzette tot het voorspellen ervan. Dat de luisteraars niet gingen voorspellen heeft waarschijnlijk te maken met de relatief korte tijdsduur tussen het hulpwerkwoord en het voltooid deelwoord.

Terwijl in hoofdstuk 3 werd onderzocht hoe syntactische informatie de verwerking van gereduceerde spraak beïnvloed, onderzochten we in hoofdstuk 4 hoe blootstelling aan gereduceerde spraak de verwerking van syntactische informatie beïnvloedt. Hiervoor werd de Event-Related-Potential (ERP) techniek gebruikt. De vraag was of luisteraars grammaticale fouten met betrekking tot geslacht in spontane spraak tolereren. Om deze vraag te bestuderen, kregen Nederlandse moedertaalsprekers bijvoeglijke naamwoorden gevolgd door zelfstandige naamwoorden te horen. De bijvoeglijke naamwoorden waren ofwel correct vervoegd, met een woordfinale schwa (bijv. *een spannende roman*) ofwel incorrect vervoegd, zonder een schwa (bijv. *een spannend roman*). Ten opzichte van de correct, vervoegde bijvoeglijke naamwoorden, veroorzaakten de incorrecte, niet vervoegde bijvoeglijke naamwoorden een P600 effect in zorgvuldig geproduceerde spraak. Dit resultaat is conform de ERP literatuur en bevestigt dat luisteraars gevoelig zijn voor morfosyntactische fouten wanneer ze naar zorgvuldig geproduceerde spraak luisteren. Wanneer de spreker op een spontane manier sprak, was dit P600 effect echter afwezig. Om er zeker van te zijn dat het P600 effect niet afwezig was omdat luisteraars de zinnen niet begrepen, werd een controle conditie toegevoegd waarin de deelnemers semantisch incorrecte zinnen te horen kregen. Van deze zinnen weten we dat ze een N400 effect teweegbrengen. Zorgvuldig en informeel geproduceerde semantisch incorrecte zinnen veroorzaakten een N400 effect. Dit laat zien dat luisteraars beide typen zinnen konden begrijpen. De resultaten laten zien dat luisteraars zich aanpassen aan de spreekstijl van een spreker tijdens het verwerken van de akoestische-fonetische informatie uit het spraaksignaal. Deze aanpassing beperkt zich niet tot de manier waarop de akoestische-fonetische informatie wordt verwerkt maar breidt zich uit over meerdere niveaus van verwerking, in het bijzonder morfosyntactische verwerking.

Hoofdstuk 5 onderzocht hoe de spelling van een woord de verwerking van schwa reductie beïnvloedt in een andere taal dan Nederlands, een taal waarin schwa reductie ook heel gebruikelijk is: Frans. De invloed van spelling werd vergeleken met de invloed van gevarieerde fonetische input. We onderzochten hoe groot de invloed van spelling is in vergelijking met de invloed van fonetische informatie. Verder bestudeerden we of deze twee factoren zowel productie als perceptie van gesproken woorden beïnvloedden. Om deze vraag nader te bekijken werden twee experimenten uitgevoerd. Er werd gebruik gemaakt van een methode waarbij nieuwe woorden moeten worden geleerd. De deelnemers leerden nieuwe Franse woorden waarin de schwa ongereduceerd (bijv. *secobe*) of gereduceerd (bijv. *scobe*) was. De deelnemers leerden de woorden gedurende drie opeenvolgende dagen. Na de laatste sessie op de derde dag voerden de deelnemers zowel een eye-tracking taak uit als een andere taak waarin ze plaatjes benoemden. Tijdens de eerste taak werd woordherkenning gemeten en in de tweede taak woordproductie. In het eerste experiment kregen luisteraars nieuwe woorden te horen die varieerden in uitspraak. Woorden werden de helft van de tijd met schwa uitgesproken en de andere helft van de tijd zonder schwa. In het tweede experiment was de uitspraak consistent maar was er een fase opgenomen waarin de deelnemers werden blootgesteld aan gereduceerde woorden die of met of zonder de letter <e> werden gespeld.

De resultaten van beide taken suggereren dat zowel fonetische variatie als spelling voorspellen hoe schwa reductie wordt verwerkt en nieuwe woorden worden geleerd.

Deelnemers waren vaker geneigd om gereduceerde woorden te produceren met een schwa als ze op een inconsistente manier waren gepresenteerd (soms met en soms zonder schwa, als in het eerste experiment) en als deze woorden met de letter <e> waren gespeld (als in het tweede experiment). De invloed van fonetische variatie woog zwaarder dan het effect van spelling. Dit komt waarschijnlijk doordat orthografie een minder betrouwbare leidraad is bij het identificeren van klanken dan de klanken zelf.

Dit proefschrift heeft laten zien dat het verwerken van akoestisch gereduceerde spraak een grote mate van flexibiliteit vergt van de taalgebruiker. Deze flexibiliteit brengt verschillende niveaus van taalverwerking met zich mee, hetgeen suggereert dat de manier waarop we akoestische-fonetische eigenschappen van gereduceerde spraak verwerken een wisselwerking aangaat met de manier waarop we andere typen talige informatie verwerken. Deze flexibiliteit houdt ook in dat de taalgebruiker in staat is om de verschillende bronnen van informatie tegen elkaar af te wegen (zoals akoestische, syntactische en orthografische informatie) afhankelijk van hun aanwezigheid en betrouwbaarheid. Deze bevindingen zijn relevant voor ons begrip van het verwerken van de grote akoestische variatie die voortkomt uit spraakreducties. Verder zijn deze bevindingen ook relevant voor ons begrip van het verwerken van akoestische-fonetische variatie door mensen in het algemeen. Wanneer we weten hoe taalgebruikers gereduceerde spraak verwerken, komt er misschien ook informatie beschikbaar over mechanismes die ook relevant zijn voor het verwerken van andere typen systematische variatie, zoals, bijvoorbeeld, buitenlandse accenten of regionale dialecten. Met dit proefschrift heb ik een eerste stap gezet in het verkrijgen van informatie door te laten zien dat het verwerken van informele spraak flexibiliteit vereist gedurende meerdere verwerkingsprocessen.

References

Adda-Decker, M. and Snoeren, N. D. (2011). Quantifying temporal speech reduction in french using forced speech alignment. *Journal of Phonetics*, 39(3):261–270.

Allopenna, P., Magnuson, J., and Tanenhaus, M. (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. and Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*, 73(3):247–264.

Arai, M. and Keller, F. (2013). The use of verb-specific information for prediction in sentence processing. *Language and Cognitive Processes*, 28(4):525–560.

Aylett, M. and Turk, A. (2006). Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei. *The Journal of the Acoustical Society of America*, 119(5):3048–3058.

Baayen, H., Piepenbrock, R., and Gulikers, L. (1995). *The CELEX lexical database. Release 2 (CD-ROM)*. Linguistic Data Consortium, Pennsylvania.

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

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

Bakker, I., Takashima, A., van Hell, J. G., Janzen, G., and McQueen, J. M. (2014). Competition from unseen or unheard novel words: Lexical consolidation across modalities. *Journal of Memory and Language*, 73:116–130.

Barbiers, S., van der Auwera, B. J., Bennis, H., Boef, E., De Vogelaer, G., and van der Ham, M. (2008). *Syntactische atlas van de Nederlandse Dialecten Deel II [Syntactic Atlas of the Dutch Dialects II]*. Amsterdam: Amsterdam University Press.

Barr, D. J. (2008). Analyzing ‘visual world’ eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, 59(4):457–474.

Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3):255–278.

Bell, A., Brenier, J. M., Gregory, M., Girand, C., and 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., and Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in english conversation. *The Journal of the Acoustical Society of America*, 113(2):1001–1024.

Brouwer, S., Mitterer, H., and Huettig, F. (2010). Shadowing reduced speech and alignment. *Journal of the Acoustical Society of America*, 128(1):EL32–EL37.

Brouwer, S., Mitterer, H., and 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(11):2193–2220.

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

Brouwer, S., Mitterer, H., and Huettig, F. (2013). Discourse context and the recognition of reduced and canonical spoken words. *Applied Psycholinguistics*, 34(3):519–539.

Bürki, A., Ernestus, M., and Frauenfelder, U. H. (2010). Is there only one “fênetre” 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(4):421–437.

Bürki, A., Ernestus, M., Gendrot, C., Fougeron, C., and Frauenfelder, U. H. (2011). 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(6):3980–3991.

Bürki, A. and Frauenfelder, U. H. (2012). Producing and recognizing words with two pronunciation variants: Evidence from novel schwa words. *Quarterly Journal of Experimental Psychology*, 65(4):796–824.

Bürki, A., Spinelli, E., and Gaskell, M. G. (2012). A written word is worth a thousand spoken words: The influence of spelling on spoken-word production. *Journal of Memory and Language*, 67(4):449–467.

Chambers, C., Tanenhaus, M., Eberhard, K., Filip, H., and Carlson, G. (2002). Circumscribing referential domains during real-time language comprehension. *Journal of Memory and Language*, 47(1):30–49.

Cheng, C. and Xu, Y. (2009). Extreme reductions: contraction of disyllables into monosyllables in taiwan mandarin. In *INTERSPEECH*, pages 456–459.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3):181–204.

Connine, C. M., Blasko, D. G., and Hall, M. (1991). Effects of subsequent sentence context in auditory word recognition: Temporal and linguistic constrainst. *Journal of Memory and Language*, 30(2):234–250.

Connine, C. M., Ranbom, L. J., and Patterson, D. J. (2008). Processing variant forms in spoken word recognition: The role of variant frequency. *Perception & Psychophysics*, 70(3):403–411.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, 1(6):84–107.

Coulson, S., King, J. W., and Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes*, 13:21–58.

Creel, S. C., Aslin, R. N., and Tanenhaus, M. K. (2008). Heeding the voice of experience: The role of talker variation in lexical access. *Cognition*, 106(2):633–664.

Cutler, A. and Davis, C. (2012). An orthographic effect in phoneme processing, and its limitations. *Frontiers in Psychology: Language Sciences*, 3:1–7.

Cutler, A., Treiman, R., and Ooijen, B. v. (2010). Strategic deployment of orthographic knowledge in phoneme detection. *Language and Speech*, 53(3):307–320.

Dahan, D. and Tanenhaus, M. (2004). Continuous mapping from sound to meaning in spoken-language comprehension: Immediate effects of verb-based thematic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(2):498–513.

Dalby, J. M. (1986). *Phonetic structure of fast speech in American English*. Bloomington: Indiana University Linguistic Club.

de Sutter, G. (2009). Towards a multivariate model of grammar: The case of word order variation in dutch clause final verb clusters. In Dufter, A., Fleischer, J., and Seiler, G., editors, *Describing and modeling variation in grammar*, pages 225–254.

DeLong, K., Urbach, T., and Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, 8(8):1117–1121.

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

Ernestus, M. (2014). Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua*, 142:27–41.

Ernestus, M., Baayen, H., and Schreuder, R. (2002). The recognition of reduced word forms. *Brain and Language*, 81(1-3):162–173.

Ernestus, M., Hanique, I., and Verboom, E. (2015). The effect of speech situation on the occurrence of reduced word pronunciation variants. *Journal of Phonetics*, 48:60–75.

Ernestus, M. and Warner, N. (2011). An introduction to reduced pronunciation variants. *Journal of Phonetics*, 39(3):253–260.

Fodor, J. (1983). *The modularity of mind: An essay on faculty psychology*. MIT Press.

Friederici, A. D., Pfeifer, E., and Hahne, A. (1993). Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, 1:183–192.

Gaskell, M. G. and Marslen-Wilson, W. D. (1996). Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 22(1):144–158.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105(2):251–279.

Gouvea, A. C., Phillips, C., Kazanina, N., and Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and Cognitive Processes*, 25(2):149–188.

Gow, D. W. (2001). Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language*, 45(1):133–159.

Gow, D. W. (2002). Does english coronal place assimilation create lexical ambiguity? *Journal of Experimental Psychology: Human Perception and Performance*, 28(1):163–179.

Gow, D. W. (2003). Feature parsing: Feature cue mapping in spoken word recognition. *Perception & Psychophysics*, 65(4):575–590.

Gow, D. W. and McMurray, B. (2007). Word recognition and phonology: The case of english coronal place assimilation. *Papers in Laboratory Phonology*, 9:173–200.

Grainger, J., Muneaux, M., Farioli, F., and Ziegler, J. C. (2005). Effects of phonological and orthographic neighbourhood density interact in visual word recognition. *The Quarterly Journal of Experimental Psychology*, 58(6):981–998.

Hagoort, P. and Brown, C. (1999). The interplay between syntax and semantics: evidence from P600/SPS and N400. *International Journal of Psychophysiology*, 33(1):41.

Hagoort, P., Brown, C., and Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, 8(4):439–483.

Hahne, A. and Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, 11:194–205.

Hanique, I. and Ernestus, M. (2012). The processes underlying two frequent casual speech phenomena in dutch: A production experiment. In *13th Annual Conference of the International Speech Communication Association 2012 (INTERSPEECH 2012), Vols 1-3*, pages 2011–2014.

Hanique, I., Ernestus, M., and 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(3):1644–1655.

Hanulíková, A., van Alphen, P. M., van Goch, M. M., and Weber, A. (2012). When one person's mistake is another's standard usage: The effect of foreign accent on syntactic processing. *Journal of Cognitive Neuroscience*, 24(4):878–887.

Hay, J. (2003). *Causes and consequences of word structure*. New York and London: Routledge.

Huettig, F. and Janse, E. (2015). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, 31(1):80–93.

Huettig, F. and 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. and Ernestus, M. (2011). The roles of bottom-up and top-down information in the recognition of reduced speech: Evidence from listeners with normal and impaired hearing. *Journal of Phonetics*, 39(3, SI):330–343.

Janse, E., Nooteboom, S. G., and Quene, H. (2007). Coping with gradient forms of /t/-deletion and lexical ambiguity in spoken word recognition. *Language and Cognitive Processes*, 22(2):161–200.

Johnson, K. (2004). Massive reduction in conversational american english. In Yoneyama, K. and Maekawa, K., editors, *Spontaneous Speech: Data and Analysis*, Proceedings of the 1st Session of the 10th International Symposium, pages 29–54. The National International Institute for Japanese Language. Tokyo, Japan.

Kamide, Y., Scheepers, C., and Altmann, G. T. M. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research*, 32(1):37–55.

Kemps, R., Ernestus, M., Schreuder, R., and Baayen, H. (2004). Processing reduced word forms: The suffix restoration effect. *Brain and Language*, 90(1-3):117–127.

Kohler, K. J. (1990). Segmental Reduction in Connected Speech in German: Phonological Facts and Phonetic Explanations. In Hardcastle, W. J. and Marchal, A., editors, *Speech Production and Speech Modelling*, pages 69–92. Springer Netherlands.

Kutas, M. and Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62:621–647.

Landauer, T. K. and Streeter, L. A. (1973). Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. *Journal of Verbal Learning and Verbal Behavior*, 12(2):119–131.

Lawrence, M. A. (2013). ez: Easy analysis and visualization of factorial experiments.

Levelt, W., Roelofs, A., and Meyer, A. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1):1–75.

Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech*, 6(3):172–187.

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

Magnuson, J., Tanenhaus, M., Aslin, R., and Dahan, D. (2003). The time course of spoken word learning and recognition: Studies with artificial lexicons. *Journal of Experimental Psychology: General*, 132(2):202–227.

Magnuson, J. S., Tanenhaus, M. K., and Aslin, R. N. (2008). Immediate effects of form-class constraints on spoken word recognition. *Cognition*, 108(3):866–873.

Matin, E., Shao, K., and Boff, K. (1993). Saccadic overhead – information-processing time with and without saccades. *Perception & Psychophysics*, 53(4):372–380.

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

McClelland, J. and Elman, J. (1986). The TRACE model of speech-perception. *Cognitive Psychology*, 18(1):1–86.

McQueen, J. M. (2005). Speech perception. In Lamberts, K. and Goldstone, R., editors, *The Handbook of Cognition*, pages 255–275. London: Sage Publications.

McQueen, J. M. (2007). Eight questions about spoken-word recognition. In Gaskell, M. G., editor, *The Oxford handbook of psycholinguistics*, pages 37–53. Oxford: Oxford University Press.

McQueen, J. M. and Huettig, F. (2012). Changing only the probability that spoken words will be distorted changes how they are recognized. *Journal of the Acoustical Society of America*, 131(1):509–517.

McQueen, J. M. and Viebahn, M. C. (2007). Tracking recognition of spoken words by tracking looks to printed words. *Quarterly Journal of Experimental Psychology*, 60(5):661–671.

Mirman, D. (2014). *Growth curve analysis and visualization using R*. Taylor & Francis.

Mirman, D., Dixon, J. A., and Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*, 59(4):475–494.

Mirman, D., McClelland, J. L., and Holt, L. L. (2006). An interactive Hebbian account of lexically guided tuning of speech perception. *Psychonomic Bulletin & Review*, 13(6):958–965.

Mitterer, H. and Ernestus, M. (2006). Listeners recover /t/s that speakers reduce: Evidence from /t/-lenition in dutch. *Journal of Phonetics*, 34(1):73–103.

Mitterer, H. and McQueen, J. M. (2009). Processing reduced word-forms in speech perception using probabilistic knowledge about speech production. *Journal of Experimental Psychology: Human Perception and Performance*, 35(1):244–263.

Mitterer, H. and Reinisch, E. (2015). Letters don't matter: No effect of orthography on the perception of conversational speech. *Journal of Memory and Language*, 85:116–134.

Nakamura, M., Iwano, K., and Furui, S. (2008). Differences between acoustic characteristics of spontaneous and read speech and their effects on speech recognition performance. *Computer Speech & Language*, 22(2):171–184.

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

Norris, D., McQueen, J. M., and Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, 23(3):299–325.

Norris, D., McQueen, J. M., Cutler, A., and Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology*, 34(3):191–243.

Oostdijk, N. (2002). The design of the spoken dutch corpus. In Peters, P., Collins, P., and Smith, A., editors, *New Frontiers of Corpus Research*, pages 105–112. Amsterdam: Rodopi.

Osterhout, L. and Hagoort, P. (1999). A superficial resemblance does not necessarily mean you are part of the family: Counterarguments to coulson, King and Kutas (1998) in the P600/SPS-P300 debate. *Language and Cognitive Processes*, 14(1):1–14.

Pardoen, J. (1991). De interpretatie van zinnen met de rode en de groene volgorde [The interpretation of sentences with the red and the green order]. *Forum Der Letteren*, 32:1–22.

Perre, L., Bertrand, D., and Ziegler, J. (2011). Literacy affects spoken language in a non-linguistic task: An ERP study. *Frontiers in Psychology: Language Sciences*, 2:1–8.

Perre, L., Pattamadilok, C., Montant, M., and Ziegler, J. C. (2009). Orthographic effects in spoken language: On-line activation or phonological restructuring? *Brain Research*, 1275:73–80.

Pickering, M. and Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36(4):329–347.

Pitt, M. (2009). The strength and time course of lexical activation of pronunciation variants. *Journal of Experimental Psychology: Human Perception and Performance*, 35(3):896–910.

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

Pitt, M. A. and Samuel, A. G. (2006). Word length and lexical activation: Longer is better. *Journal of Experimental Psychology: Human Perception and Performance*, 32(5):1120–1135.

Plag, I., Homann, J., and Kunter, G. (2015). Homophony and morphology: The acoustics of word-final s in english. *Journal of Linguistics*, 1:1–36.

Pluymaekers, M., Ernestus, M., and Baayen, R. (2005a). Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica*, 62(2-4):146–159.

Pluymaekers, M., Ernestus, M., and Baayen, R. (2005b). Lexical frequency and acoustic reduction in spoken dutch. *Journal of the Acoustical Society of America*, 118(4):2561–2569.

Poellmann, K., Bosker, H. R., McQueen, J. M., and Mitterer, H. (2014). Perceptual adaptation to segmental and syllabic reductions in continuous spoken dutch. *Journal of Phonetics*, 46:101–127.

Pollack, I. and Pickett, J. M. (1963). The intelligibility of excerpts from conversational speech. *Language and Speech*, 6:165–171.

Pollack, I. and Pickett, J. M. (1964). Intelligibility of excerpts from fluent speech: Auditory vs. structural context. *Journal of Verbal Learning and Verbal Behavior*, 3:79–84.

R Core Team (2015). R: A Language and Environment for Statistical Computing.

Racine, I., Buerki, A., and Spinelli, E. (2013). The implication of spelling and frequency in the recognition of phonological variants: Evidence from pre-readers and readers. *Language, Cognition and Neuroscience*, 29(7):893–898.

Ranbom, L. J. and Connine, C. M. (2007). Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language*, 57(2):273–298.

Ranbom, L. J. and Connine, C. M. (2011). Silent letters are activated in spoken word recognition. *Language and Cognitive Processes*, 26(2):236–261.

Raymond, W. D., Dautricourt, R., and 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.

Samuel, A. G. and Kraljic, T. (2009). Perceptual learning for speech. *Attention, Perception & Psychophysics*, 71(6):1207–1218.

Schiller, N. and Meyer, A. (2003). Introduction to the relation between speech comprehension and production. In Schiller, N. and Meyer, A., editors, *Phonetics and phonology in language comprehension and production: Differences and similarities*, pages 1–8. Berlin: Mouton de Gruyter.

Schuppler, B., Ernestus, M., Scharenborg, O., and Boves, L. (2011). Acoustic reduction in conversational dutch: A quantitative analysis based on automatically generated segmental transcriptions. *Journal of Phonetics*, 39(1):96–109.

Schuppler, B., van Dommelen, W. A., Koreman, J., and 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(4):595–607.

Seidenberg, M. S. and Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, 5(6):546–554.

Spinelli, E. and Gros-Balthazard, F. (2007). Phonotactic constraints help to overcome effects of schwa deletion in French. *Cognition*, 104(2):397–406.

Sulpizio, S. and McQueen, J. M. (2011). When two newly-acquired words are one: New words differing in stress alone are not automatically represented differently. In *12th Annual Conference of the International Speech Communication Association 2011 (INTERSPEECH 2011)*, pages 1396–1399.

Sulpizio, S. and McQueen, J. M. (2012). Italians use abstract knowledge about lexical stress during spoken-word recognition. *Journal of Memory and Language*, 66(1):177–193.

Swerts, M. and van Wijk, C. (2005). Prosodic, lexico-syntactic and regional influences on word order in dutch verbal endgroups. *Journal of Phonetics*, 33(2):243–262.

Taft, M., Castles, A., Davis, C., Lazendic, G., and Nguyen-Hoan, M. (2008). Automatic activation of orthography in spoken word recognition: Pseudohomograph priming. *Journal of Memory and Language*, 58(2):366–379.

Tanenhaus, M., Magnuson, J., Dahan, D., and Chambers, C. (2000). Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research*, 29(6):557–580.

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., and Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268(5217):1632–1634.

Torreira, F. and Ernestus, M. (2012). Weakening of intervocalic /s/ in the nijmegen corpus of casual spanish. *Phonetica*, 69(3):124–148.

Tuinman, A., Mitterer, H., and Cutler, A. (2014). Use of syntax in perceptual compensation for phonological reduction. *Language and Speech*, 57(1):68–85.

van Alphen, P. and McQueen, J. (2001). The time-limited influence of sentential context on function word identification. *Journal of Experimental Psychology: Human Perception and Performance*, 27(5):1057–1071.

Van Bergem, D. (1994). A model of coarticulatory effects on the schwa. *Speech Communication*, 14:143–162.

van Berkum, J., Brown, C., Zwitserlood, P., Kooijman, V., and Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(3):443–467.

van de Ven, M., Ernestus, M., and Schreuder, R. (2012). Predicting acoustically reduced words in spontaneous speech: The role of semantic/syntactic and acoustic cues in context. *Laboratory Phonology*, 3(2):455–481.

van de Ven, M., Tucker, B. V., and Ernestus, M. (2011). Semantic context effects in the comprehension of reduced pronunciation variants. *Memory & Cognition*, 39(7):1301–1316.

Viebahn, M. C., Ernestus, M., and 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.

Wicha, N., Moreno, E., and Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in spanish sentence reading. *Journal of Cognitive Neuroscience*, 16(7):1272–1288.

Witteman, M. J., Weber, A., and McQueen, J. M. (2013). Foreign accent strength and listener familiarity with an accent codetermine speed of perceptual adaptation. *Attention, Perception & Psychophysics*, 75(3):537–556.

Young, S., Evermann, G., Gales, M., Hain, T., Kershaw, D., Liu, X., Moore, G., Odell, J., Ollason, D., Povey, D., Valtchev, V., and Woodland, P. (2002). *The HTK Book*. Cambridge: Entropic.

Ziegler, J. C. and Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review*, 5(4):683–689.

Ziegler, J. C., Ferrand, L., and Montant, M. (2004). Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks. *Memory & Cognition*, 32(5):732–741.

Ziegler, J. C., Jacobs, A. M., and Stone, G. O. (1996). Statistical analysis of the bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods, Instruments & Computers*, 28(4):504–515.

Acknowledgments

Many people have contributed to the work presented in this thesis and they have done so in many different ways. I will not be able to thank everyone individually or to thank them for every individual way in which they helped. But I am grateful to each of them for everything they did.

First and foremost, I would like to thank my supervisors Mirjam Ernestus and James McQueen who supported my application to the IMPRS and who gave me the opportunity to come to Nijmegen to work with them. Over the years, they have provided continuous and consistent support and I have learned a lot from them. This dissertation would not have been possible without their help.

I would also like to thank the IMPRS and the CLS for providing financial support and for making it possible for me to attend international conferences, Summer/Winter Schools, and for organizing interesting and useful courses and meetings. In particular, I would like to thank Els, Dirkje, and Rachel for always having kept an open door and for their friendly and helpful advice.

Furthermore, I would like to thank my colleagues in the CLSM for creating a positive atmosphere at the workplace during all these years: Annika, Arina, Ellen, Helen, Huib, Iris, Kim, Kimberley, Laura, Lisa, Louis, Mark, Martijn, Mybeth, Nienke, Sascha, Sho, Sophie, Thordis, and Xaver. In particular, I would like to thank my office mates Ellen, Iris, Kim, Mybeth, and Sophie. For helping me to learn how to analyse ERP data, I would like to thank Florian Hintz, Kimberly Mulder, and Mark Noordenbos. Special thanks go to my paranimfs Sophie and Xaver who have been invaluable in helping me to submit my thesis and organize my defence from abroad. Many thanks go to the student assistants for testing participants and for their indispensable help in the creation, recording, and analysis of many stimulus materials. Thank you Corine, Erik, Ferdy, Ilse, Jens, Lisa, Martijn, Nadia, Thera, and Tijl.

I would like to thank the members of the MaRCS/CS group, the Sound Learning group, the CLS, and the Speech Comprehension PI group for sharing their work and for helpful feedback.

In addition, I thank the MPI Volleyball team, in particular Add Verbunt for organizing it and Thordis for friendly chats while giving me rides with her car. I also would like to express gratitude to the MPI library staff and the MPI Technical Group.

I am indebted to Antje Meyer, Herbert Schriefers, and Esther Janse who formed my reading committee and who took the time to read and evaluate this thesis.

Furthermore, I would like to thank Uli Frauenfelder and Audrey Bürki for welcoming me in their lab at the University of Geneva during my six-month research stay. In their *équipe*, I enjoyed stimulating debates and a friendly atmosphere, and I benefited from a fruitful collaboration. Many thanks also go to Natalia Kartushina and Jean-Laurent Astier for having been such good office mates during that time. I am very happy to be back in this group now, and I appreciate the support I got while I was still writing the final pages of this thesis.

Finally, I thank my family and friends for their love and support. M – zum Schluss möchte ich mich bei dir für deine Liebe, Geduld und Unterstützung bedanken.

Geneva, February 2016

Curriculum vitae

Malte Viebahn was born on the 1st of December 1981 in Bremen (Germany). He studied Psychology at the Westfälische Wilhelms-Universität in Münster (Germany) where he worked in the Psycholinguistics laboratory of Pienie Zwitserlood. During this time he also studied for six months under the supervision of Christoph Scheepers at the University of Dundee in Scotland (UK) and spent a four-month internship in the Speech Comprehension Group at the Max Planck Institute for Psycholinguistics in Nijmegen (the Netherlands) working with James McQueen. He received his Diploma from Münster in 2007 and went on to study Cognitive Psychology at SUNY Buffalo (USA) under the supervision of Paul Luce and Jim Sawusch where he graduated in 2011. In the same year he was awarded a 3-year PhD fellowship by the International Max Planck Research School for Language Sciences that allowed him to move to Nijmegen to start the work that resulted in this thesis. He was supervised by Mirjam Ernestus and James McQueen. During this time he spent six months as a visiting researcher in the laboratory of Uli Frauenfelder at the University of Geneva (Switzerland) where he conducted a research project together with Audrey Bürki. In the same group, he is currently working as a postdoctoral researcher.

Author publications

Viebahn, M. C., McQueen, J. M., Ernestus, M., Frauenfelder, U. H., and Bürki, A. (in preparation). The influence of phonetic variability and orthography on the processing of novel French words: Evidence from spoken-word production and recognition.

Viebahn, M.C., Ernestus, M., and McQueen, J.M. (under revision). Absence of inflectional schwa does not disrupt syntactic processing in casual speech.

Viebahn, M. C., Ernestus, M., and 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.

Viebahn, M. C., Ernestus, M., and McQueen, J. M. (2012). Co-occurrence of reduced word forms in natural speech. In *Proceedings of Interspeech 2012*, 2019-2022.

McQueen, J.M. and Viebahn, M.C. (2007). Tracking recognition of spoken words by tracking looks to printed words. *The Quarterly Journal of Experimental Psychology, 60* (5), 661-671.

MPI Series in Psycholinguistics

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing. *Miranda 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 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 Janssen*
12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie 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 Weber*
17. Moving eyes and naming objects. *Femke 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 H. de Jong*
21. Fixed expressions and the production of idioms. *Simone A. 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 cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological approach. *Daniëlle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra M. 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 Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. 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. *Marlies 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 Broersma*

35. Retrieving semantic and syntactic word properties. *Oliver Müller*

36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*

37. Sensitivity to detailed acoustic information in word recognition. *Keren B. 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 B.M. Haun*

42. The acquisition of auditory categories. *Martijn Goudbeek*

43. Affix reduction in spoken Dutch. *Mark Pluymaekers*

44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Kooijman*

45. Space and iconicity in German Sign Language (DGS). *Pamela 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 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 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 electro-encephalography, 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 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 Junge*

68. Adjusting to different speakers: Extrinsic normalization in vowel perception. *Matthias J. Sjerps*

69. Structuring language. Contributions to the neurocognition of syntax. *Katrien R. 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. *Conny 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 Sümer*

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 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 Jongman*

109. Acoustic reduction in spoken-word processing: Distributional, syntactic, morphosyntatic, and orthographic effects. *Malte Viebahn*