

Speaking Style Influences the Brain's Electrophysiological Response to Grammatical Errors in Speech Comprehension

Malte C. Viebahn^{1,2}, Mirjam Ernestus^{1,2}, and James M. McQueen^{1,2}

Abstract

■ This electrophysiological study asked whether the brain processes grammatical gender violations in casual speech differently than in careful 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 an electrical brain response sensitive to syntactic violations when the talker was speaking in a careful manner. When the talker was

speaking in a casual manner, this response was absent. A control condition showed electrophysiological responses 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. Absent schwas in casual speech are effectively not grammatical gender violations. These changes in syntactic processing are evidence of contextually driven neural flexibility. ■

INTRODUCTION

Spoken language is characterized by an extraordinary amount of variability. The type of variability investigated in this 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., Johnson, 2004; Ernestus, 2000). How does the brain cope with this variability? In the following, we investigate whether the brain uses acoustic–phonetic information provided by the speaking style to adapt the way in which they recognize speech. Critically, we examine whether this kind of neural adaptation can have consequences for syntactic processing.

The notion that listeners' brains 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 & Ernestus, 2011; Van de Ven, Tucker, & Ernestus, 2011; Dillely & Pitt, 2010). 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, Mitterer, & Huettig, 2012) 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., McQueen & Viebahn, 2007; Allopenna, Magnuson, & Tanenhaus, 1998), 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 the brain adapts to casual speech by penalizing acoustic mismatches less strongly than when processing careful speech. In another study showing adaptation to casual speech (Poellmann, Bosker, McQueen, & Mitterer, 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, Mitterer, & Huettig, 2010) showing that hearing reduced

¹Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands, ²Radboud University, Nijmegen, The Netherlands

speech increases the probability of producing reduced word forms.

The ERP experiment presented in this article extends these previous studies on casual speech by asking whether neural adaptation to acoustic–phonetic reductions can have consequences for how the brain processes syntactic markers. More specifically, we investigate how the reduction of inflectional schwa in Dutch (spelled here 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”). Previous studies have shown that grammatical gender plays an important role in language processing. In speech production in Dutch, the gender congruency of adjectives (with or without final schwa) and nouns can affect utterance onset latency (Schriefers, 1993). In comprehension, ERP studies in Dutch investigating gender violations such as incorrectly inflected adjectives have revealed a clear P600/syntactic positive shift effect for the following noun (Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005; Hagoort & Brown, 1999). This late positive deflection (which we label as the P600) is usually associated with syntactic processing and is often assumed to be an indication for syntactic parsing problems and repair processes (e.g., Friederici, Pfeifer, & Hahne, 1993; Hagoort, Brown, & Groothusen, 1993). There is a discussion in the literature about the exact cognitive and neural processes that underlie the P600 effect (compare, e.g., van Herten, Kolk, & Chwilla, 2005; Hahne & Friederici, 1999; Osterhout & Hagoort, 1999; Coulson, King, & Kutas, 1998). There is also evidence that the P600 can arise in response to semantic anomalies (van Herten et al., 2005). Irrespective of what the precise mechanism underlying the P600 might be (e.g., whether it reflects syntactic reanalysis or monitoring processes) and irrespective of whether it is exclusively syntactic in nature, there is general agreement that difficulties in syntactic processing are indexed by the P600 (Gouvea, Phillips, Kazanina, & Poeppel, 2010).

In casual Dutch, the vowel schwa is often either shortened in its duration or completely absent (e.g., Pluymaekers, Ernestus, & Baayen, 2005; 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 forms resulting from common reduction processes that in careful speech would be grammatically inappropriate.

A study by Hanulíková, van Alphen, van Goch, and Weber (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 nonnative speaker. Gender violations in native speech resulted in a larger P600 compared with correct sentences, indicating that the listeners were sensitive to the grammatical errors. In contrast, when the same violations were produced by a nonnative 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 this study, we used a similar design to Hanulíková et al. (2012) and applied it to the domain of casual speech. Is the brain more tolerant of the absence of inflectional schwas when hearing a casual speech style compared with 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. To rule out the possibility that the absence of a P600 effect is due to shallower processing or that listeners were unable to understand the content of the casually produced utterances, we added a control condition (as did Hanulíková et al.) in which listeners were exposed to semantically anomalous utterances (spoken once again in either a careful style or a casual style). Such stimuli have been shown to elicit a negative electrophysiological deflection around 400 msec (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 with the word “sugar” presented in the same sentence (for a review, see Kutas & Federmeier, 2011). We predict that, if the expected absence of a P600 effect in casual speech is due to neural adaptation to speaking style, adaptation should have no influence on how the brain responds to semantic anomalies and N400 effects ought to be the same for casually and carefully produced utterances.

CORPUS STUDY

There is currently a debate about the role of morphology in acoustic–phonetic reduction processes (e.g., Plag, Homann, & Kunter, 2017; Hanique & Ernestus, 2012; Schuppler, van Dommelen, Koreman, & Ernestus, 2012; Hay, 2003). The extent and 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 schwa constitute inflectional affixes. 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 3753 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 5496 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 inflectional schwa can be absent. On the basis of the collected data, it is clear that schwas can indeed be absent in casual Dutch when they function as syntactic markers but also that they are not always deleted.

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 and 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, Brenier, Gregory, Girand, & Jurafsky, 2009; Pluymaekers et al., 2005). 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, Piepenbrock, & Gulikers, 1995). For the unreduced adjectives, the preceding words had a mean frequency of 12.57 ($SD = 2.29$), whereas 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 after the unreduced adjectives had a mean frequency of 7.76 ($SD = 2.37$), whereas the words after the reduced adjectives had a mean frequency of 8.32 ($SD = 2.15$). Overall, the frequencies between the reduced and unreduced adjectives show no striking differences.

ERP EXPERIMENT

Method

Participants

Thirty-two native speakers of Dutch were recruited from the participant pool of the Max Planck Institute for Psycholinguistics. All were university students and right-handed. Age ranged from 18 to 24 years (mean = 20.9 years), and five of the participants were men. 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 of Dutch. Each utterance consisted of two or three sentences. Table 1 shows an example of the critical, control, and filler utterances. Each utterance was produced by a male native speaker 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.'s (2012) study in which there was a consistent mapping between speaker and accent (native vs. foreign).

One hundred twenty critical utterances were constructed. These contained a noun phrase consisting of

Table 1. Example Utterances for Each of the Three Utterance Types

Utterance Type		Example
Critical	Dutch	MORGEN ga ik met de trein naar Berlijn. Ik wil wel nog een <u>spannende roman</u> / * <u>spannend roman</u> 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 <u>suspenseful novel</u> 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 <u>hengel</u> / * <u>atleet</u> binnen met een VIS eraan.
	English	“I was walking beside a pond used for FISHING. Then, coincidentally, a fisherman JUST pulled his <u>fishing rod</u> / * <u>athlete</u> out with a FISH on it.”
Filler	Dutch	Mijn oma is de laatste tijd heel warrig en MOE. De dokter zegt dat het een <u>mogelijk effect</u> is van haar MEDICIJNEN.
	English	“My grandma has recently been rather woozy and TIRED. The doctor says that this might be a <u>possible side effect</u> of her MEDICINE.”

Words with sentential stress are written in capital letters. Crucial words are underlined.

the indefinite article “*een*” or “*a*,” an adjective, and a common-gender noun. For each utterance, a correct version and an incorrect version were created. In the correct version, the adjective was inflected, whereas in the incorrect version, 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 neuter-gender noun.

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. To make each noun semantically more predictable, a strong semantic expectation was generated during the phrase that 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 284 utterances. The control and critical utterances were based on Hanulíková et al. (2012), but often with substantial adjustments that 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 drive.google.com/open?id=0B5mtD4tPL57WdFM1MDhVMlZadkk.

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 manner; 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 syntactic error.

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 to avoid spillover 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 “yes,” and in the other half, it was “no.”

Stimulus Recordings

Recordings were made by a male native speaker and a female native speaker of Dutch. Each speaker produced careful and casual versions of each utterance. A correct version and an incorrect version were 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. To determine whether a schwa was present, a phonetically trained listener examined the recordings with audio editing software by looking at the waveforms and the spectrograms and listening to them. 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 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. A similar procedure was used to establish the onset of the critical words. For the critical utterances, we recorded 960 tokens. For the 104 control utterances, 832 tokens were recorded, and for the 60 filler utterances, we recorded 240 tokens. Including the 10 practice items, 2042 tokens were recorded.

As an example, the critical adjective–noun pair “*spannende roman*” or “suspenseful novel” is shown in Figure 1. In some cases, the speaker introduced relatively long pauses between the sentences forming a given utterance. The long pauses made the utterances sound unnatural. To avoid unnaturally sounding utterances, the recordings were adjusted with PRAAT audio editing software (Boersma & Weenink, 2015) such that the maximum duration of a pause was 400 msec. The durations of the adjectives, adjusted sentences, and schwas are shown in Figure 2. 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.

To assess how much the careful and casual productions deviate from their dictionary transcriptions, we computed the phonological overlap between the phonetic transcriptions of 31 of our utterance types (230 recordings) and their corresponding dictionary transcriptions.

The dictionary transcriptions were taken from the CELEX lexical database (Baayen et al., 1995), and the phonological overlap computations were based on the Levenshtein distance measure (Levenshtein, 1966). The average overlap for the careful utterances was 81%, whereas for the casual utterances, it was 72%. When we consider only the words that precede the adjective, the careful utterances have 80% overlap with the dictionary transcription, whereas the casual utterances have only 69% 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. A phonetically transcribed example utterance is shown in Table 2. In this example, the carefully produced sentences overlap with 94% of the utterance’s dictionary transcription, whereas the casually produced sentences overlap only by 73% with the dictionary transcription. When considering only the words that precede the adjective, the careful utterance has 93% overlap with the dictionary transcription, whereas the casual utterance has only 60% overlap with the dictionary transcription.

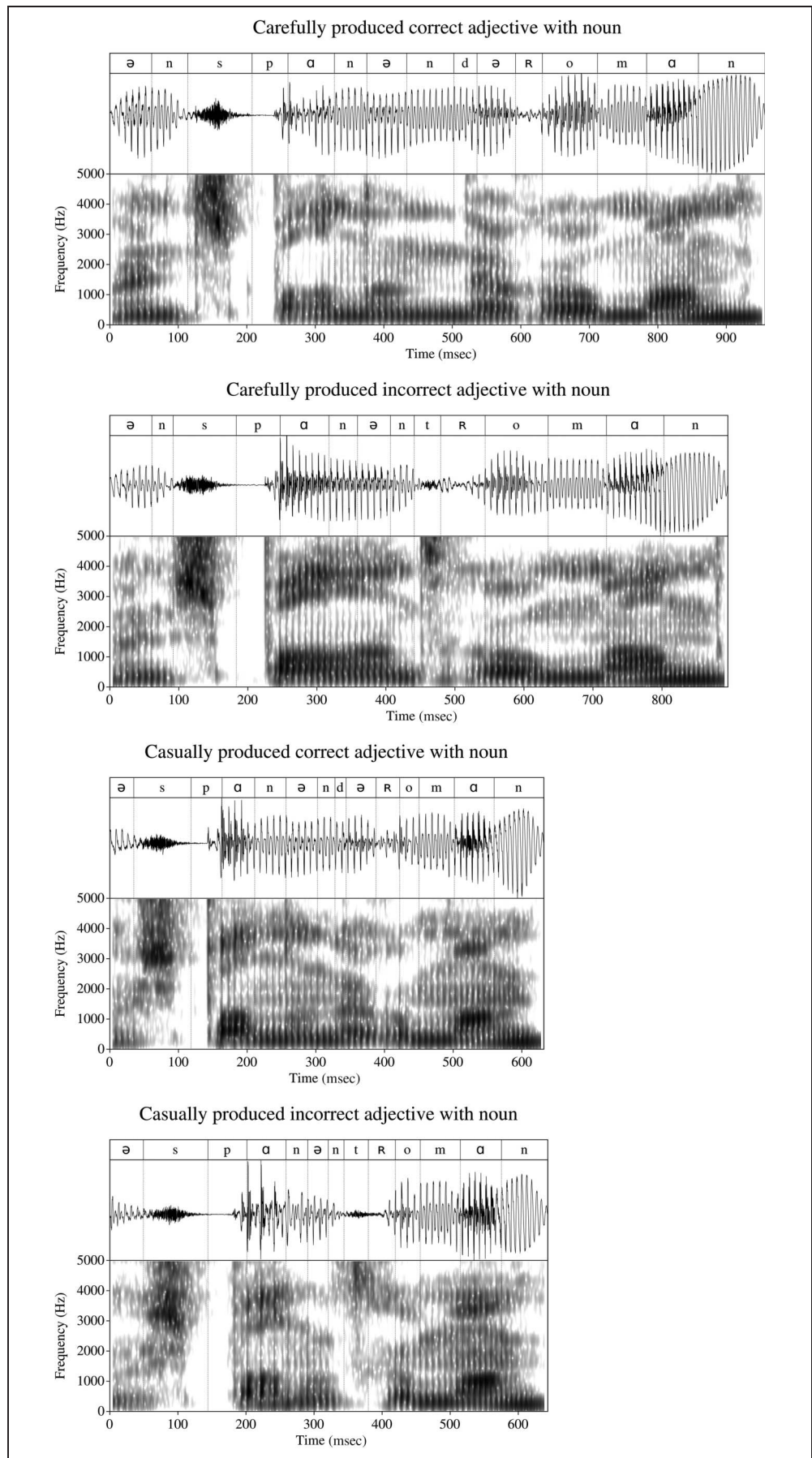
Apparatus

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 system (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. To monitor horizontal EOGs, electrodes FT9 and FT10 were placed on the left and right temples of the participant, respectively. Vertical EOGs 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 15 k Ω . The EEG and EOG signals were recorded and digitized with PyCorder and amplified by a BrainAmp DC amplifier with an online band-pass filter for 0.02–200 Hz. The montage of the electrodes is shown in Figure 3. 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 participant. The utterances were then combined, and a pseudorandom running order was created (with the constraints that a given condition could not occur more than three times in a row and that

Figure 1. The Dutch adjective–noun pair *spannende roman* (“suspenseful novel”) recorded by the male speaker in four experimental conditions: carefully produced and with adjective-final schwa (top), carefully produced and without adjective-final schwa (top center), casually produced and with adjective-final schwa (bottom center), and casually produced and without adjective-final schwa (bottom). The adjectives 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 a /t/.



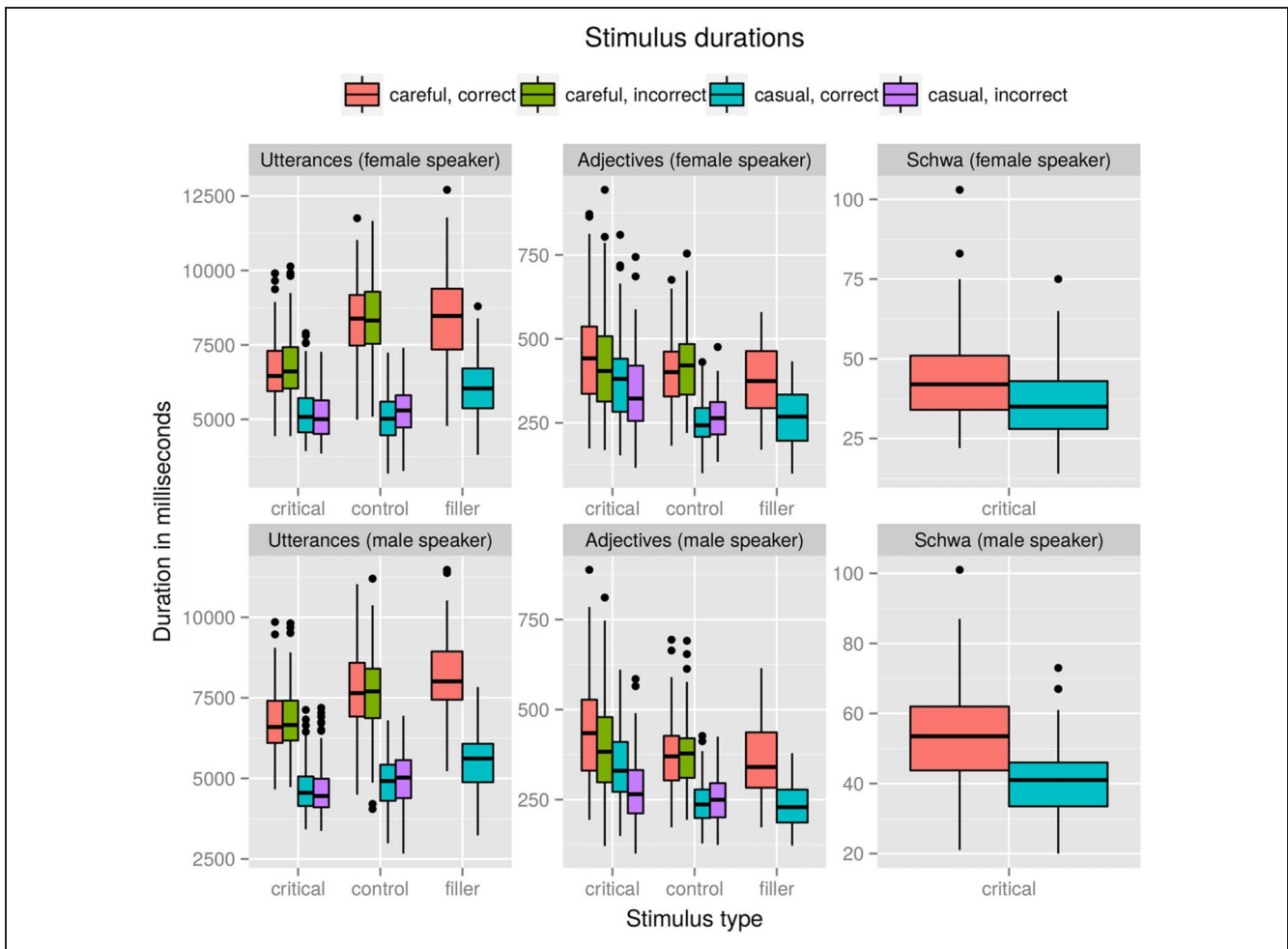


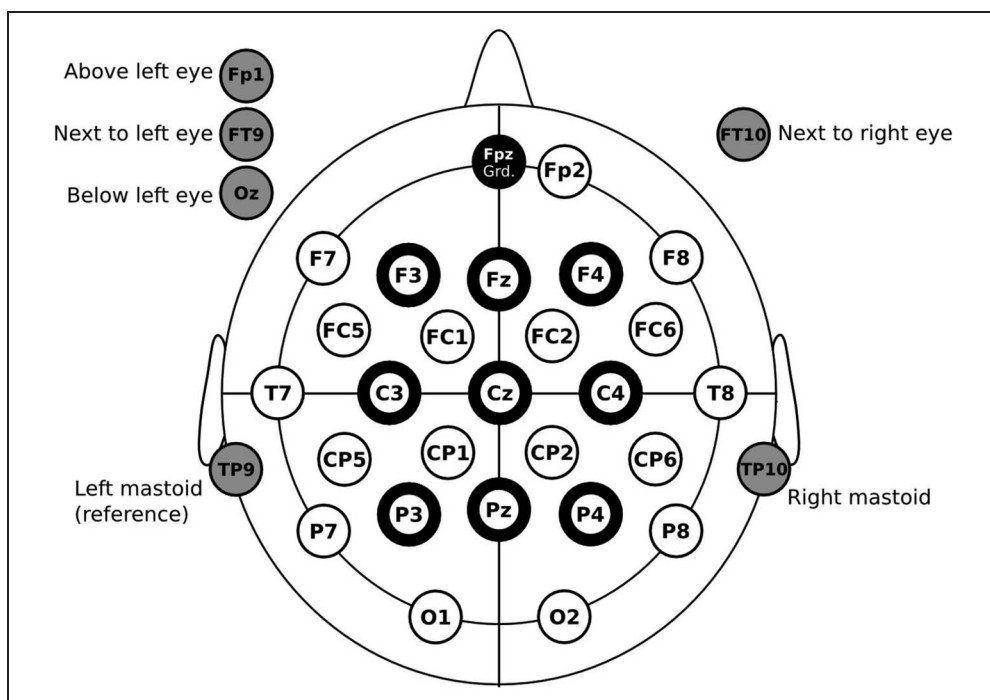
Figure 2. Durations of the (adjusted) utterances, adjectives, and schwas. Critical stimuli refer to utterances that may contain morphosyntactic violations, control stimuli refer to utterances that may contain semantic violations, and filler stimuli refer to utterances that never contain any type of violation.

Table 2. Example Transcriptions of a Carefully and Casually Produced Utterance

Part	Register	Phonetic Transcription	Number of Segments Corresponding to Dictionary Transcription	Percentage of Overlap with Dictionary Transcription
1	Dictionary	[mərɔxə xɑ ɪk mət də trɛɪn nar bərleɪn]	29	-
	Careful	[mərɔxə xɑ ɪʔ mət ə trɛɪn na bərleɪn]	26	90%
	Casual	[mɔ xɑ ʔ mɛ trɛɪn na bəleɪn]	18	62%
2	Dictionary	[ɪk vɪl vɛl nɔx ən spənəndə romɑn kɔpə fɔr ɪk fɛrtɹɛk]	42	-
	Careful	[ɪk vɪl vɛl nɔx ən spənəndə romɑn kɔpə fɔr ɪk fətrɛk]	40	95%
	Casual	[xəvɛl nɔx ə spənəndə romɑn kɔpə fɔ ʔ fətrɛk]	31	74%
3	Dictionary	[dɑn hɛp ɪk ɪts tə lezə]	17	-
	Careful	[dɑn hɛp ɪk ɪts tə lezə]	17	100%
	Casual	[dɑn hɛp ɪts tə lezə]	15	88%

The utterance consists of three parts. See the critical utterance shown in Table 1 for the orthographic transcription and English translation. Phonetic transcriptions are shown in symbols of the International Phonetic Alphabet. A segment was counted as not corresponding to the dictionary transcription if it was either missing or changed (e.g., 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 7272 msec, and the duration of the casual utterance is 4794 msec.

Figure 3. EEG montage. Data from the nine electrodes with a broad black border were entered into the statistical analysis (see Figure 5).



an utterance with a question was always followed by a filler). The utterances in this running order were then rotated through each condition, resulting in eight rotations for the running order (four rotations for the filler utterances). Rotations 5–8 were replications of Rotations 1–4 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 a woman and the casual speaker was a man 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). Listening trials began with the presentation of a blank screen for 500 msec, 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 at 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 pressing a button on a

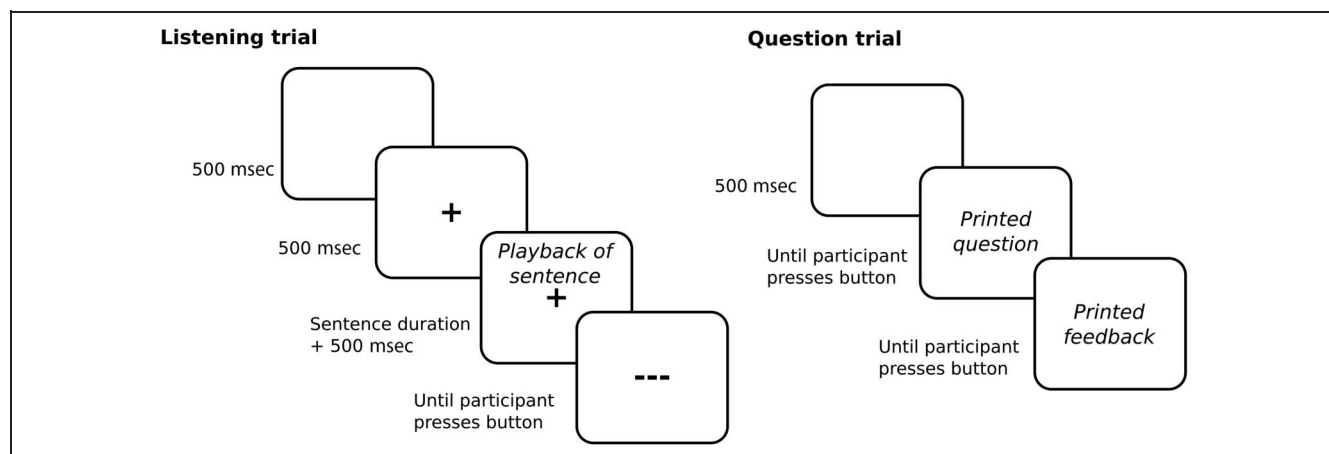


Figure 4. Procedure during the listening and question trials in the experiment. For the question trials, the printed feedback was either the word “correct” printed in green or the word “incorrect” printed in red.

game pad, the word “correct” (in green) or “incorrect” (in red) appeared on the screen providing the participants with feedback as to whether they had answered the question correctly.

One experimental session consisted of 10 practice trials and 284 experimental trials resulting in 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 min. In addition, the fitting of the EEG equipment took another 20–30 min.

Results

We analyzed the question trials to examine if speaking style influenced how well participants were able to respond. Questions after a carefully produced utterance were responded to correctly in 93% of the cases (mean RT = 2954 msec). Questions after casually produced utterances were answered correctly 96% of the time (mean RT = 2958 msec). RT was measured from the time when the printed questions appeared on the screen until participants pressed either the “yes” or “no” button on the response pad. Linear mixed-effects models showed that neither the difference in accuracy nor the difference in RT was statistically significant (both t s < 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 produced in a casual speaking style.

We analyzed the EEG data by computing repeated-measures ANOVAs for participant 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 1500 msec. For the analysis of the N400 component, the time window ranged from 300 to 500 msec. 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 of 200 msec until noun onset was used for baselining.

Before the statistical analysis of the data, each trial was checked for 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. An

ANOVA with the factors Grammaticality, Speaking style, and Electrode (26 electrodes) showed a significant three-way interaction ($F(25, 775) = 1.76, p < .05$) for the critical utterances. In the following analyses, we conducted analyses for individual electrodes focusing on the three frontal electrodes (F3, Fz, and F4), three central electrodes (C3, Cz, and C4), and three parietal electrodes (P3, Pz, and 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 Figure 5.

There were no effects on the P600 component for the frontal electrodes (all F s < 4 and all p s > .05). One of the central electrodes (C3) showed a significant interaction between Speaking style and Grammaticality ($F(1, 31) = 4.71, p < .05$). The parietal electrodes P3 and Pz also showed such an interaction (P3: $F(1, 31) = 6.49, p < .05$; Pz: $F(1, 31) = 5.27, p < .05$). For P4, there was a main effect of Grammaticality showing a larger P600 for incorrect compared with correct utterances ($F(1, 31) = 6.28, p < .05$) but no interaction ($F(1, 31) = 3.42, p = .07$). 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 casual conditions. All three electrodes showed a significant effect of grammaticality for the carefully produced utterances (all F s > 7 and all p s ≤ .01) but not for the casually produced ones (all F s < 0.2 and all p s > .6).

To compare our results more closely to the results reported by Hanulíková et al. (2012), we reran our analysis using the same time window as they did (i.e., 800–1200 msec). An ANOVA with the factors Grammaticality, Speaking style, and Electrode (26 electrodes) showed no significant three-way interaction ($F(25, 775) = 1.09, p > .1$). However, because the analysis of specific individual electrodes was planned before conducting our study, we nevertheless report them. According to these analyses, the interaction between Speaking style and Grammaticality was present at electrode P3 ($F(1, 31) = 5.11, p < .05$) but absent for electrodes C3 and Pz. One-way ANOVAs for electrode P3 showed that there was a significant effect of Grammaticality for careful utterances ($F(1, 32) = 11.39, p < .01$) but not for casual utterances ($F(1, 32) = 1.58, p > .2$).

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. 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 F s < 1 and all p s > .3).

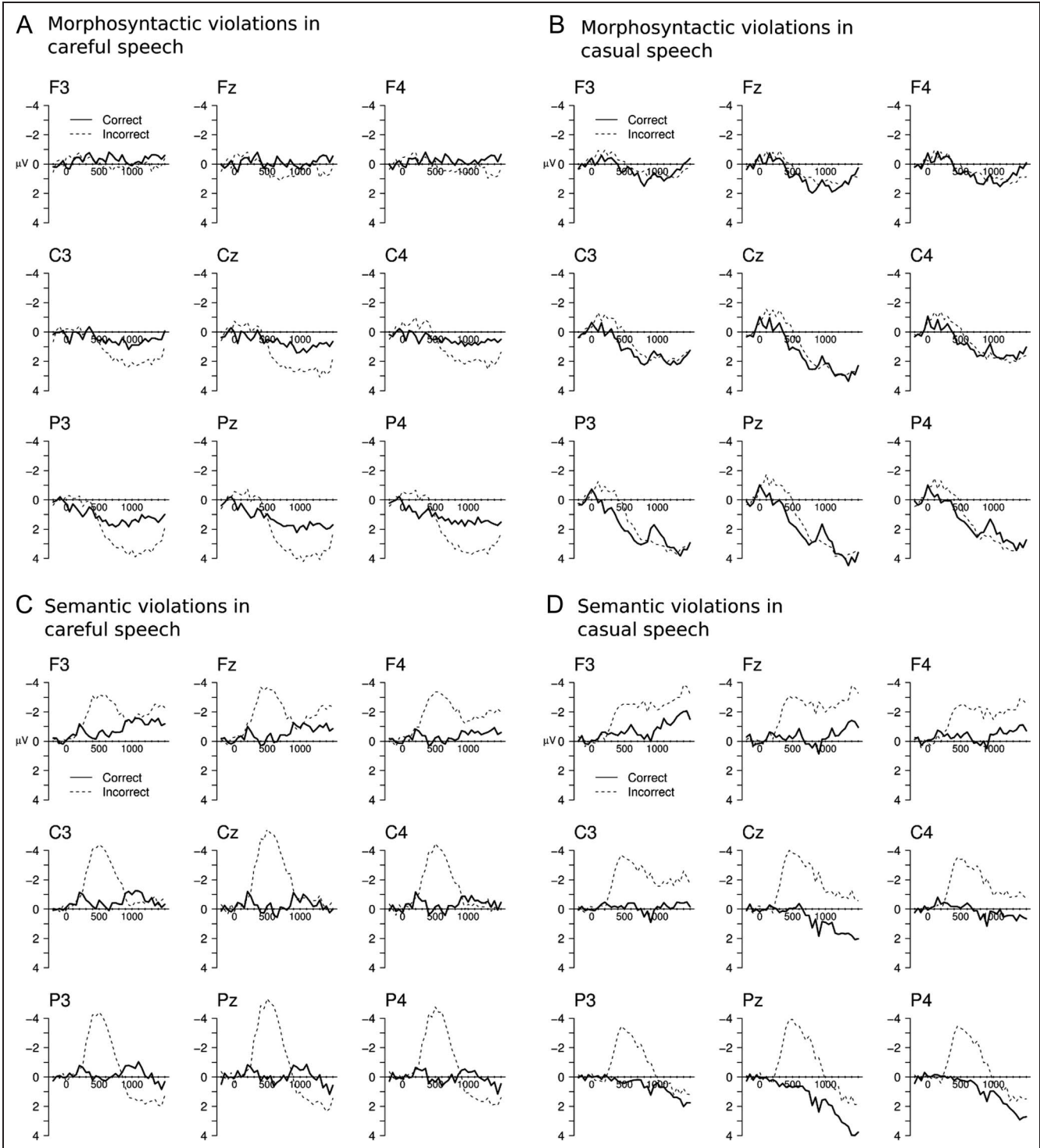


Figure 5. ERPs for critical utterances (morphosyntactic violations, A and B) and control utterances (semantic violations, C and D) for the electrodes entered into the statistical analysis: F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4 (see Figure 3 for topographical distribution).

The ERP plots for the critical utterances in Figure 5 show an early negativity at the central and parietal electrodes for both the casual and careful ungrammatical conditions. To investigate this early negative deflection, additional ANOVAs with the factors Speaking style and Grammaticality and the interaction of these factors were conducted for the time window from 0 to 400 msec. These analyses

showed a significant effect of Grammaticality for electrodes Cz, C4, P3, Pz, and P4, indicating a larger negativity for incorrect compared with correct utterances (all $F_s > 4.3$ and all $p_s < .05$). Furthermore, there was a significant effect of Speaking style for electrodes Cz, Pz, and P4 (all $F_s > 4.2$ and all $p_s < .05$). Crucially, there was no interaction between Speaking style and Grammaticality (all $F_s < 1.05$

and all $ps > .3$). These results show that, although the P600 effect was present only for carefully produced utterances, incorrect adjectives elicited an early negativity in both speaking styles.

Further inspection of the ERP plots for the critical utterances revealed a positive deflection to both correct and incorrect utterances for the casual speech condition in approximately the same time window as the P600 effect in the careful speech condition. We ran additional ANOVAs to examine whether the ERPs for correct adjectives are more positive in casual compared with careful speech using the time window from 500 to 1500 msec. We found significant effects of Speaking style for correct utterances at electrodes C3, P3, and Pz (all $F_s > 5$ and all $ps < .05$) but not for incorrect utterances (all $F_s < 0.45$ and all $ps > .5$). These analyses indicate that ERPs for correct adjectives are more positive in casual speech than in careful speech.

For the control sentences, we investigated whether the effect of semantic validity on the N400 component differed for carefully and casually produced utterances. An omnibus repeated-measures ANOVA including the factors Speaking style, Semantic validity, and Electrode showed a significant interaction between Validity and Electrode ($F(25, 775) = 17.29, p < .001$). We proceeded by analyzing the nine electrodes that we examined in the P600 analysis. These electrodes showed a significant main effect of Semantic validity reflecting a larger N400 component for incorrect compared with correct utterances (all $F_s > 32$ and all $ps < .001$). Crucially, none of the electrodes showed an interaction between Speaking style and Semantic validity (all $F_s < 3.5$ and all $ps > .07$).

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 casual speech. Furthermore, our analyses suggest that the interaction between Speaking style and Grammaticality remained constant across the course of the experiment. Additional post hoc analyses revealed an early negativity for incorrect utterances in both speaking style conditions and a late positive deflection for grammatically correct utterances in casual speech relative to grammatically correct utterances in careful speech. For the control stimuli, there was an N400 effect for casual as well as careful speeches but no interaction between Speaking style and Semantic validity.

DISCUSSION

Unlike in careful speech, which is typically produced in formal social contexts, phonological segments in casual speech are often reduced or absent. The purpose of this study was to examine how the brain responds to these differences in speaking style. Specifically, we asked whether the absence of syntactically relevant schwas in

casual speech disrupts syntactic processing. We conducted an ERP experiment in which Dutch participants listened to carefully and casually produced Dutch 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 (e.g., Hagoort & Brown, 1999), the incorrectly uninflected adjectives in careful speech elicited a positive electrophysiological brain response at approximately 600 msec after noun onset. When occurring in casual speech, however, uninflected adjectives did not elicit such a P600 effect. This suggests that the brain does not treat the absence of the syntactically relevant schwa as a grammatical error when being exposed to casual speech, but it did so when being exposed to careful speech.

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 negative electrophysiological deflection at approximately 400 msec after noun onset 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 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., Ernestus, 2014; Brouwer et al., 2012) 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 to help recognize words in reduced speech (e.g., Viebahn, Ernestus, & McQueen, 2015; Tuinman, Mitterer, & Cutler, 2014). This 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 produced by either a native speaker or a nonnative 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 nonnative speaker. The main difference from our study was that Hanulíková et al. used a condition in which utterances were produced by a nonnative 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 nonnative speakers of Dutch and native speakers who 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.'s (2012) findings 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. This difference across studies could be explained in the following way. 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.'s study, 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.'s study 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 because they involve whole words rather than individual segments. 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 with that in Hanulíková et al.'s study.

Another difference between our study and the study by Hanulíková et al. (2012) is that our data show an early negativity in response to morphosyntactic errors, whereas Hanulíková et al. do not report such an effect. This negativity might indicate that the speech signal contained some kind of cue that signaled to the listeners that an incorrect structure was going to follow. It is possible that the native speakers, when producing an incorrect sentence, inadvertently provided such cues earlier than the critical incorrect word. Although we do not know what these cues might be, there clearly must have been something in the speech signal that elicited this early negativity. Another possibility is that this early negativity reflects a left anterior negativity (LAN) component as described in previous studies on morphosyntactic violations (e.g., Friederici & Weissenborn, 2007; Rossi, Gugler, Hahne, & Friederici, 2005). These studies distinguish between an early automatic process (reflected in the LAN) that involves the detection of a morphosyntactic error and a later process (reflected in the P600) that is associated with a reanalysis of the input. The fact that there was no interaction in the early time window raises the possibility that our speaking style manipulation influenced only later reanalysis processes but did not influence the detection of the morphosyntactic error. However, the negative deflection associated with the LAN is left anterior and not, as in our study, bilateral posterior. Our result is thus not exactly compatible with previous descriptions of the LAN. Therefore, it is more likely that the early negativity indicates the presence of inadvertent cues to grammaticality rather than the presence of a LAN. A visual inspection of Hanulíková et al.'s ERP plots suggests that there may have been a similar early negative deflection in their data as well. Possibly, Hanulíková et al. did not report this finding because the effect is quite small and its topography again differs from that of the LAN.

There are several possible mechanisms that could have 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 phonetic 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 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 would be consistent with the fact that a casually speaking talker is likely to omit individual segments. The early negativity that we observed for both speaking styles would be consistent with

this interpretation. On this view, although listeners detect the morphosyntactic errors in casual speech, they do not perform a reanalysis of the input because such errors are expected given the speaking style of the preceding context. 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 responded to the casual speaking style not by becoming less sensitive to the absence of syntactically relevant schwa but by reanalyzing the speech input even if it did not contain morphosyntactic errors. This possibility is consistent with the finding that the late ERPs to the critical utterances showed an effect of speaking style for both correct and incorrect utterances. The P600 effect may therefore have been absent in casual speech not because there was no reanalysis in casual incorrect sentences but because there was reanalysis in the casual correct sentences as well. Although both explanations are compatible with our data (the explanation that there is reanalysis for both correct and incorrect sentences and the explanation that there is no reanalysis at all), we think that the latter explanation is more likely because it is in line with data that are more reliable. That is, it is based on the direct comparison of correct and incorrect casual utterances, whereas the former explanation is based on a comparison across speaking styles. The comparison across speaking styles is less reliable because it is based on misaligned speech input. The utterances are considerably shorter in the casual speech condition than in the careful speech condition, which makes it impossible to perfectly align the ERP responses. In contrast, for the comparison across the two grammaticality conditions, the ERP response can be compared within each speaking style condition. This allows for more accurate ERP alignment and thus a more reliable analysis.

A third possible explanation for the lack of a difference between correct and incorrect sentences in casual speech is that listeners adapt to the acoustic–phonetic consequences of fast speech, which characterizes casually produced utterances. Because of 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 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 neighboring 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 trying to detect whether a schwa is present. Note, however, that this adaptive process is still likely to be specific to syntactically relevant schwa 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. Future research could further explore this question and examine how the different phonological and acoustic properties that characterize casual speech could be isolated and how their individual effects could be studied. The current results nevertheless indicate that, whether it is triggered by the absence of segments or by speaking rate (or both), the 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 can reveal about them. Although there is widespread consensus that the P600 is correlated with syntactic processing difficulties, there is a debate about the underlying mechanisms. It is still debated whether the P600 occurs exclusively with linguistic stimuli or whether it reflects nonlinguistic cognitive processes. For example, it has been suggested that the P600 is part of the P300 family of ERP components. More specifically, the P600 resembles the P3b component (e.g., Coulson et al., 1998), which is elicited by rare categorical events (so-called “oddballs”), which can be either linguistic or nonlinguistic stimuli. This proposal is consistent with findings that suggest that the P600 reflects late and controlled rather than early and automatic processes (Hahne & Friederici, 1999). These studies call into question the assumption that the P600 is language specific (cf. Osterhout & 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 component does not reflect syntactic processing or, more likely, that the syntactic processes that it reflects are not strictly mandatory.

In conclusion, this 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 suggests that adaptation at an acoustic–phonetic processing level can have consequences at higher (i.e., syntactic) levels in the processing hierarchy. These findings provide further support for the notion that the

brain processes language in an adaptive and flexible manner and that neural changes at an earlier level of processing can trigger changes in later components of the language processing network.

Acknowledgments

We thank Adriana Hanulíková and her colleagues for sharing their stimulus sentences with us. Furthermore, we thank Thera Baayen and Ferdy Hubers for help with the construction and recording of the stimuli, Nadia Klijn and Lisa Rommers for help with data collection, Kimberly Mulder and Florian Hintz for help with data analysis, and two anonymous reviewers for constructive feedback. This study was partly funded by an IMPRS fellowship to M. C. V. and an ERC starting grant (284108) to M. E.

Reprint requests should be sent to Malte C. Viebahn, Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands, or via e-mail: malte.viebahn@gmail.com.

REFERENCES

- Allopenna, P., Magnuson, J., & 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, 419–439.
- Baayen, H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database. Release 2 (CD-ROM)*. Philadelphia: Linguistic Data Consortium.
- Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, 60, 92–111.
- Boersma, P., & Weenink, D. (2015). *Praat: Doing phonetics by computer (Version 5)*. Retrieved from www.praat.org/.
- Brouwer, S., Mitterer, H., & Huettig, F. (2010). Shadowing reduced speech and alignment. *Journal of the Acoustical Society of America*, 128, EL32–EL37.
- Brouwer, S., Mitterer, H., & Huettig, F. (2012). Speech reductions change the dynamics of competition during spoken word recognition. *Language and Cognitive Processes*, 27, 539–571.
- Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes*, 13, 21–58.
- Dilley, L., & Pitt, M. A. (2010). Altering context speech rate can cause words to appear and disappear. *Psychological Science*, 21, 1664–1670.
- Ernestus, M. (2000). *Voice assimilation and segment reduction in casual Dutch: A corpus-based study of the phonology-phonetics interface*. Utrecht, The Netherlands: LOT. Retrieved from repository.ubn.ru.nl/handle/2066/29974.
- Ernestus, M. (2014). Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua*, 142, 27–41.
- Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, 1, 183–192.
- Friederici, A. D., & Weissenborn, J. (2007). Mapping sentence form onto meaning: The syntax-semantic interface. *Brain Research*, 1146, 50–58.
- Gouvea, A. C., Phillips, C., Kazanina, N., & Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and Cognitive Processes*, 25, 149–188.
- Hagoort, P., & Brown, C. (1999). Gender electrified: ERP evidence on the syntactic nature of gender processing. *Journal of Psycholinguistic Research*, 28, 715–728.
- Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, 8, 439–483.
- Hahne, A., & 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., & Ernestus, M. (2012). The role of morphology in acoustic reduction. *Lingue E Linguaggio*, 11, 147–164.
- Hanulíková, A., van Alphen, P. M., van Goch, M. M., & 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, 878–887.
- Hay, J. (2003). *Causes and consequences of word structure*. London: Routledge.
- Janse, E., & 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, 330–343.
- Johnson, K. (2004). Massive reduction in conversational American English. In K. Yoneyama & K. Maekawa (Eds.), *Spontaneous speech: Data and analysis: Proceedings of the 1st Session of the 10th International Symposium* (pp. 29–54). Tokyo: The National International Institute for Japanese Language.
- Kutas, M., & 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.
- Lawrence, M. A. (2013). *ez: Easy analysis and visualization of factorial experiments*. Retrieved from CRAN.R-project.org/package=ez.
- Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, 10, 707–710.
- McQueen, J. M., & 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, 509–517.
- McQueen, J. M., & Viebahn, M. C. (2007). Tracking recognition of spoken words by tracking looks to printed words. *Quarterly Journal of Experimental Psychology*, 60, 661–671.
- Oostdijk, N. (2002). The design of the Spoken Dutch Corpus. In P. Peters, P. Collins, & A. Smith (Eds.), *New frontiers of corpus research* (pp. 105–112). Amsterdam: Rodopi.
- Osterhout, L., & 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, 1–14.
- Plag, I., Homann, J., & Kunter, G. (2017). Homophony and morphology: The acoustics of word-final S in English. *Journal of Linguistics*, 53, 181–216.
- Pluymaekers, M., Ernestus, M., & Baayen, R. (2005). Lexical frequency and acoustic reduction in spoken Dutch. *Journal of the Acoustical Society of America*, 118, 2561–2569.
- Poellmann, K., Bosker, H. R., McQueen, J. M., & Mitterer, H. (2014). Perceptual adaptation to segmental and syllabic reductions in continuous spoken Dutch. *Journal of Phonetics*, 46, 101–127.
- R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from www.R-project.org/.
- Rossi, S., Gugler, M. F., Hahne, A., & Friederici, A. D. (2005). When word category information encounters morphosyntax: An ERP study. *Neuroscience Letters*, 384, 228–233.

- Schriefers, H. (1993). Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 841–850.
- Schuppler, B., van Dommelen, W. A., Koreman, J., & Ernestus, M. (2012). How linguistic and probabilistic properties of a word affect the realization of its final /t/: Studies at the phonemic and sub-phonemic level. *Journal of Phonetics*, *40*, 595–607.
- Tuinman, A., Mitterer, H., & Cutler, A. (2014). Use of syntax in perceptual compensation for phonological reduction. *Language and Speech*, *57*, 68–85.
- 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., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 443–467.
- Van de Ven, M., Tucker, B. V., & Ernestus, M. (2011). Semantic context effects in the comprehension of reduced pronunciation variants. *Memory & Cognition*, *39*, 1301–1316.
- van Herten, M., Kolk, H. H. J., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomaly. *Cognitive Brain Research*, *22*, 241–255.
- Viebahn, M. C., Ernestus, M., & McQueen, J. M. (2015). Syntactic predictability in the recognition of carefully and casually produced speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 1684–1702.
- Young, S., Evermann, G., Hain, T., Kershaw, D., Moore, G., Odell, J., et al. (2002). *The HTK book*. Cambridge: Entropic.