Can hearing puter activate pupil? Phonological competition and the processing of reduced spoken words in spontaneous conversations

Susanne Brouwer [a] , Holger Mitterer [b] & Falk Huettig [b c]

[a] Linguistics Department, Northwestern University, Evanston, IL, USA

[b] Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

[c] Donders Institute for Brain, Cognition, and Behaviour, Radboud University,
Nijmegen, The Netherlands

First

PLEASE SCROLL DOWN FOR ARTICLE

howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Can hearing *puter* activate *pupil*? Phonological competition and the processing of reduced spoken words in spontaneous conversations

Susanne Brouwer[1], Holger Mitterer[2], and Falk Huettig[2,3]

[1]Linguistics Department, Northwestern University, Evanston, IL, USA
[2]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[3]Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, The Netherlands

In listeners' daily communicative exchanges, they most often hear casual speech, in which words are often produced with fewer segments, rather than the careful speech used in most psycholinguistic experiments. Three experiments examined phonological competition during the recognition of reduced forms such as [pjutər] for *computer* using a target-absent variant of the visual world paradigm. Listeners' eye movements were tracked upon hearing canonical and reduced forms as they looked at displays of four printed words. One of the words was phonologically similar to the canonical pronunciation of the target word, one word was similar to the reduced pronunciation, and two words served as unrelated distractors. When spoken targets were presented in isolation (Experiment 1) and in sentential contexts (Experiment 2), competition was modulated as a function of the target word form. When reduced targets were presented in sentential contexts, listeners were probabilistically more likely to first fixate reduced-form competitors before shifting their eye gaze to canonical-form competitors. Experiment 3, in which the original /p/ from [pjutər] was replaced with a "real" onset /p/, showed an effect of cross-splicing in the late time window. We conjecture that these results fit best with the notion that speech reductions initially activate competitors that are similar to the phonological surface form of the reduction, but that listeners nevertheless can exploit fine phonetic detail to reconstruct strongly reduced forms to their canonical counterparts.

Keywords: Spoken word recognition; Reduced forms; Lexical competition; Eye tracking.

The speech we encounter most often in daily life is casual speech. Although there is a growing interest in studying spoken word recognition with this type of speech, it remains a relatively unexplored area. A critical feature of casual speech is that it contains large amounts of variation. Any one word is almost always uttered differently on different occasions, and pronunciations can vary both from one speaker to another and from one situation to another. For example, an American-English speaker might produce the four-syllable word 'apparently'[1] once in its full (henceforth, canonical)

form /əpʰɛrəntli/ and once as [pʰɛɾ̃ɪ]. In the last case, the word is realized in only two syllables and contains fewer phonemes than its canonical transcription in a dictionary would prescribe. This variation in production is called speech reduction, where segments, syllables, and even whole words can be changed and/or deleted (e.g., Johnson, 2004). Reduction processes can thus significantly modify the way words are produced. Speech reductions are also very common. Johnson, for example, found in a corpus of English conversational speech that more than 60% of the words deviated from their citation form by at least one segment, and 28% of the words even deviated on two or more segments. The listener's challenge is to recognize words in spite of this variability. In the present study, we examine phonological competition during the recognition of strongly reduced forms such as [pjutər] from the canonical form [kɔmpjutər] *computer* and [ʋɛs] from the canonical form [ʋɛtstrɛit] *wedstrijd* 'match'.

Past research on speech reductions in spontaneous conversations

Only few studies have investigated listeners' comprehension of strongly reduced forms in spontaneous conversation. Ernestus, Baayen, and Schreuder (2002) examined how listeners recognize highly reduced forms in Dutch such as [ifɑl] for [ɪn idər χəfɑl] *in ieder geval* 'in any case'. They presented such forms in different context sizes. The listeners' task was to write down the form they heard. The results showed that, in isolation, listeners hardly recognized these forms. If the forms were presented in a phonetic context, recognition performance increased, but listeners reached ceiling level only when the context was several words long. Their results suggest that highly reduced forms cannot be recognized on the basis of their acoustic forms alone; only when there is a semantic/syntactic context available can one recognize reduced forms correctly (see also Arai, 1999; Brouwer, Mitterer, & Huettig, in press). In a subsequent study, Kemps, Ernestus, Schreuder, and Baayen (2004) provided evidence that listeners reconstruct (most likely) subconsciously highly

reduced forms—that is, they compute their canonical counterparts. Listeners had to monitor for the phoneme /l/ in highly reduced forms such as [ɛik] from /ɛiχələk/ *eigenlijk* 'actually'. If such forms were embedded in a sentence context, listeners often incorrectly reported the phoneme /l/. However, listeners did not report /l/ if the reduced forms were presented in isolation.

These two studies suggest that people only recognize strongly reduced forms within a context of several words, and when they do so, they also activate the canonical word forms. In particular, the results of Kemps et al. (2004) seem to indicate that reduced forms are linked to the canonical representation in the mental lexicon and not to a more veridical reflection of the actual input, reflecting the acoustic signal itself. It is, however, possible that these results reflect their use of offline tasks. Such tasks require listeners' metalinguistic judgements, which are conscious and controlled and thus take time to develop. Studies using offline tasks are therefore unable to measure whether listeners also actively consider lexical candidates compatible with the acoustic structure of reduced pronunciations early on in the recognition process. For example, listening to the reduced form [pjutər] of the canonical form *computer* may activate lexical candidates that sound similar in onset such as 'pupil' and 'pure'. Similarly, listening to the reduced form [ʋɛs] of the canonical form *wedstrijd* 'match' may activate a lexical candidate such as 'west' more than hearing the canonical form [ʋɛtstrɛit], because of the longer onset overlap of the reduced form with the reduced-form competitor.

Printed-word eye tracking

A useful technique to investigate the *online* processing of strongly reduced forms is visual world eye tracking (Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). In this methodology, listeners' eye movements are measured as they listen to speech and see pictures of objects on a computer screen. The timing and proportion of fixations to pictures of objects are typically taken to be related to which lexical

candidates the listener is entertaining as speech unfolds over time (see Huettig, Rommers, & Meyer, 2011, for further discussion). Allopenna, Magnuson, and Tanenhaus (1998), for instance, showed that listeners fixate pictures with names similar to the target name more often than phonologically unrelated names. In Allopenna et al.'s study, participants saw four pictures on a computer screen (e.g., a 'beaker', a 'beetle', a 'speaker', and a 'carriage') and listened to spoken instructions such as 'Pick up the beaker'. Participants looked at the pictures of both types of competitors, but more often to pictures of onset-match competitors (e.g., the 'beetle') than to pictures of offset-match competitors (e.g., the 'speaker'). Moreover, looks to 'beetle' started to increase at the onset of the word 'beaker', whereas looks to 'speaker' started to increase as the end of the word 'beaker' unfolded. Recently, a printed-word version of the visual world paradigm has been developed (Huettig & McQueen, 2007; McQueen & Viebahn, 2007).

The printed-word variant has been shown to be particularly sensitive for observing fine-grained phonetic and phonological processing. McQueen and Viebahn (2007) replicated—with even tighter experimental control (e.g., controlling for degree of phonetic mismatch, no repetition of stimuli)—the findings of Allopenna et al. (1998): Their participants looked more often to onset-matching (e.g., buffer for buffel 'buffalo') than to offset-matching competitors (e.g., lotje 'lottery ticket' for rotje 'fire-cracker'). It is important to note that McQueen and Viebahn's results are consistent with the results from a variety of other techniques (cross-modal priming, Marslen-Wilson & Zwitserlood, 1989; phoneme monitoring, Connine, Titone, Deelman, & Blasko, 1997), which suggests that printed-word eye tracking is a sensitive and valid method for observing spoken-

word recognition processes. Huettig and McQueen (2007) used both (printed word and picture) versions of the eye-tracking method and found the printed-word variant to be more sensitive to phonological manipulations than the traditional version using pictures. Results obtained by Weber, Melinger, and Lara Tapia (2007) support this view.[2]

In a recent study, Brouwer, Mitterer, and Huettig (2012) used the printed-word variant of the paradigm to examine spoken-word recognition in Dutch conversational speech containing many speech reductions. More specifically, the study investigated whether word recognition processes in casual speech differ from word recognition processes with carefully articulated laboratory speech as used in many psycholinguistic experiments. Following the example of Ernestus and colleagues (2002; Kemps et al., 2004), the stimulus material was compiled from a spontaneous speech corpus. However, whereas Ernestus and colleagues used an offline task to study the effect of spontaneous speech on word recognition, Brouwer and colleagues used an online task to tap directly into the time course of processing. They compared the recognition of reduced and canonical forms of mid-to-high-frequency content words in a four-word display of which one of the words was the target word. Two types of competitors were displayed on the screen. The first type was a "canonical-form competitor" (e.g., [kɔmpɛnjən] 'companion' for [kɔmpjutər]), which phonologically overlapped more at onset with the canonical form than with the reduced form of the spoken word. The second type was a "reduced-form competitor" (e.g., [pjupɪl] 'pupil' for [pjutər]), which phonologically overlapped more at onset with the reduced form than with the canonical form of the spoken word. In Brouwer et al.'s Experiment 1, listeners directed their attention to a similar degree to both competitors on hearing either

---

[2] Note, however, that eye movement behaviour in both versions reflects complex mapping processes between spoken words and printed words/pictures. Past research (see Huettig, Mishra, & Olivers, 2012, for detailed discussion) has shown that with picture displays, fixations can be determined by matches between knowledge retrieved on the basis of information in the linguistic and in the visual input at phonological, semantic, and visual levels of representation. With printed word displays, fixations are determined by online matches at phonological, semantic, and orthographic levels. The exact dynamics of the representational level at which such mapping occurs however is codetermined by the timing of cascaded processing in the spoken word and object/visual word recognition systems, by the temporal unfolding of the spoken language, and by the nature of the visual environment (e.g., which other representational matches are present).

the reduced realization [pjutər] or the canonical realization [kɔmpjutər] of *computer*. In their Experiment 2, they did not include any reduced forms in the experiment and compared the recognition of canonical forms in laboratory speech with the recognition of canonical forms in casual speech. In such a listening situation, listeners directed significantly more overt attention to the canonical-form competitor than the reduced-form competitor both in the laboratory speech condition and in the casual speech condition. In Experiment 3, they intermixed the canonical forms of Experiment 2 with reduced forms and found (as in Experiment 1) that there was no difference between listeners' fixations to canonical-form and reduced-form competitors. Brouwer and colleagues concluded that during casual speech, which includes a great deal of reduced word forms, listeners are more tolerant of acoustic mismatches between input and canonical form. These data therefore showed that speech-intrinsic variation (e.g., the overall reliability and quality of the phonetic input) can modulate phonological competition.

Nevertheless, given the earlier visual world studies showing increased competition effects for phonological onset competitors (Allopenna et al., 1998; McQueen & Viebahn, 2007), it is surprising that there was no preference for canonical-form competitors over reduced-form competitors in the Brouwer et al. (2012) study. There are two possible explanations why such a result was not observed. One possibility is that the task situation affects phonologically mediated eye gaze. In the visual world paradigm, visual and auditory information jointly determine eye gaze (Huettig et al., 2011). In other words, listeners' eye gaze not only reflects the processing of the spoken input, but is also affected by the processing of the stimuli in the visual display. Huettig and Altmann (2011), for instance, have shown that the properties of all the (partly) matching objects in the display affect the magnitude and timing of eye gaze. In the Brouwer et al. study, there were three items in the display for which there was some phonological overlap with the spoken word (the target; the canonical-form competitor; and the reduced-form competitor). Thus, it is possible that the combination of three at least (partly) matching items in the display

and the great number of speech reductions in the spoken stimuli created a task situation in which listeners were more tolerant of phonological mismatches.

There is, however, yet another explanation why the input form may not have influenced differentially the fixations to canonical-form and reduced-form competitors in Brouwer et al. (2012). It is possible that the reduced segments in the reduced forms (e.g., [pjuːtər] for *computer*) carried fine phonetic detail that indicated that the [p] is not word-initial, but that the word starts with a weak syllable starting with /k/. In fact, Brouwer et al. observed for the reduced form [pjuːtər] for *computer* that the closure duration for /p/ was rather long (>100 ms) for connected speech. This may signal that a weak syllable was literally "swallowed" in this closure; consequently, /p/-initial words would not serve as strong competitors for either reduced or canonical word forms because neither carries sufficient evidence for a p-initial word. While the multitude of reductions in the stimulus set of Brouwer et al. makes it difficult to pinpoint all possible phonetic details, other research supports the idea that reduced forms often carry cues to the full form. Manuel (1992) showed subphonemic differences between versions of [sport] when the intended word was *sport* and *support*, even if the schwa in the first syllable was deleted (i.e., there was no voicing between the [s] and the [p]). Similar effects have been reported for French voicing assimilation (Snoeren, Hallé, & Seguí, 2006), place assimilation in English (Gow, 2002, 2003), and voice assimilation in German (Kuzla, Mitterer, & Ernestus, 2010). Even though not all reduced forms may carry cues to the canonical form (see Gaskell & Snoeren, 2008), cues from phonetic details may be rather ubiquitous in reduced forms. This has repercussions for the interpretation of the eye-tracking study by Brouwer et al. Phonetic-detail cues may be the reason why the fixation patterns did not differ between canonical and reduced input forms, because those cues in fact minimize the difference between a full and reduced form, as the reduced form is only a full form "in disguise", so to speak. Note that, if this explanation of the Brouwer

et al. results is correct, changing the task situation should not influence competitor activations.

It is conceivable, however, that when hearing reduced forms in other task situations, words (e.g., *pupil*) matching the phonological surface form of the speech reduction (e.g., [pjuːtər]) compete more strongly than words matching the phonological surface form of the canonical form (e.g., *companion*) of the target word. The target-absent version of the visual world paradigm (Huettig & Altmann, 2005), for instance, in which a fully matching target word referent is excluded from the visual display (e.g., hearing *computer* and not seeing a 'computer' on the screen), has been shown to greatly increase the magnitude of competition effects for related words (cf. Figure 2 in Huettig & Altmann, 2005). Huettig and Altmann (2005) found that participants directed their attention to an object in the visual display (e.g., *trumpet*) when a semantically related target word (e.g., *piano*) was heard, both in a condition in which the target (i.e., piano) was copresent and in a condition in which the target (piano) was absent. Fixation proportions in the target-absent condition, however, were much greater than those in the target-present condition. In the present study, therefore, we used the target-absent version of the paradigm because of its greater sensitivity to reveal online activation of competitor representations.

Present study

Here we investigated whether phonological competition for reduced words in casual speech can ever be influenced by the exact phonetic form of the spoken word. We used the target-absent version of the visual world paradigm, described above, to maximize the likelihood of observing competitor effects and their time course. The visual display in the current study therefore had the following structure: Given the target word *computer* (in either canonical or reduced form), the visual display contained a canonical-form competitor (*companion*), a reduced-form competitor (*pupil*), and two phonologically unrelated distractors (*jewel*; *holiday*). Note that the actual items were in Dutch. We

examined when and to what extent reduced-form and canonical-form competitors play a role in the online recognition of naturally reduced words. If we observe that the overt attention to different competitors is influenced by the input form, this may indicate that reduced forms do not carry sufficient phonetic-detail cues to prevent the activation of words that are similar to the reduced form in the auditory input.

EXPERIMENT 1

Method

Participants

Twenty-four participants from the Max Planck Institute's subject pool, undergraduates at the Radboud University Nijmegen, were paid to participate in this experiment. All were native speakers of Dutch and reported normal hearing and normal or corrected-to-normal vision.

Materials

We selected 32 polysyllabic, mid-to-high-frequency content words from the Spoken Dutch Corpus (Oostdijk, 2000). We took both a canonical (e.g., [bənedə] for *beneden* 'downwards') and a reduced realization (e.g., [məneə]) of every target word (see Appendix). Recordings with background noise, overlapping speech, or unfamiliar dialects such as Flemish were excluded. The target words were transcribed by two independent raters to observe the signal in auditory and visual spectrographic form. The independent transcriptions were compared to verify agreement between the transcriptions. In case of disagreement, the transcribers reached consensus.

The segments of the canonical forms were almost always fully realized, whereas their reduced counterparts missed one or more segments. The critical criterion for a reduced form to be included in the study was that it shared more initial segments with another existing Dutch word than with its own canonical form. To illustrate this, the reduced form [ʋɛtstrɛi] for the canonical form [ʋɛtstrɛit] *wedstrijd* 'match', in which the final /t/

is deleted, does not live up to the criterion. In this case, no other Dutch word exists that phonologically matches the reduced form [ʋɛtstrɛi] except its own canonical form *wedstrijd*. As a consequence, the reduced form [ʋɛtstrɛi] could not be included in our material. An example of a reduced form that would live up to our criterion is [ʋɛs] for *wedstrijd*. In this case, for example, the Dutch word *wesp* 'wasp' matches phonologically better with the reduced form [wɛs] than with the canonical form [ʋɛtstrɛit].

There was considerable variation in the reductions. Reduced forms differed either in the initial part (i.e., first or second segment) such as [mənəə] from [bənedə] for *beneden* 'downwards', or in a later part (i.e., third, fourth or fifth segment) such as [ʋɛs] for [ʋɛtstrɛit] *wedstrijd* 'match' from the canonical form. The Appendix lists all target items including their canonical and reduced transcriptions. Note that such a degree of variability in the materials is the norm in work on spontaneous speech. Such materials often lead to greater variability in results, but it is necessary in order to study real spontaneous speech (see Warner, 2012, for a discussion of this trade-off).

For each trial, the computer screen displayed three different word types: a canonical-form competitor (e.g., *beneden* 'downwards'), a reduced-form competitor (e.g., *meneer* 'mister'), and two phonologically unrelated distractors (e.g., *juweel* 'jewel' and *vakantie* 'holiday'; see Figure 1). Note
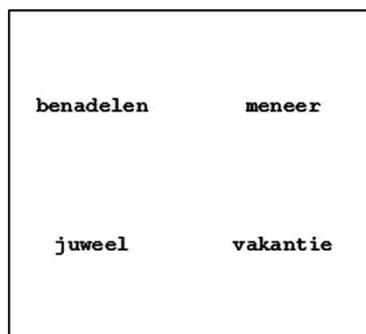

Figure 1. *Example of a printed-word display without a visual target presented to participants (the spoken target word in this example is* beneden *'downwards').*

that in 75% of the cases (24 out of 32, see Appendix) the canonical-form competitor and the reduced-form competitor overlap phonologically at onset. Some reduced-form competitors are therefore also to a certain degree competitors of the target word's citation form. They thus function as onset overlap competitors. However, the canonical-form competitor always overlapped more at onset with the canonical form than with the reduced form, and the reduced-form competitor always overlapped more at onset with the reduced form than with the canonical form. For example, the word *wedstrijd* 'match' was realized canonically as [ʋɛtstrɛit] and in a reduced way as [ʋɛs]. The first three segments of the canonical-form competitor *wetboek* [ʋɛtbuk] are shared with the canonical form, but only the first two segments are shared with the reduced form. The first three segments of the reduced-form competitor *wesp* [ʋɛsp], however, are shared with the reduced form, but only the first two segments are shared with the canonical form. It is therefore important to compare the relative strength of the two types of competitors under different conditions.

We carefully selected competitors that were equally well matched to the canonical form and to the reduced form. To rule out that the canonical-form competitors have more phonemes in common with the intended target word than the reduced-form competitors to their targets, we analysed whether the canonical-form and the reduced-form competitors differ with respect to their total segmental overlap with the target forms. This analysis took the segmental order into account, but did not require an exact match of the position. For example, the reduced-form competitor *persoon* 'person' [pərson]—matching the reduced form [pəsipə]—shares three out of six phonemes with its target form *principe* 'principle' [prɪnsipə]. The shared phonemes between the reduced-form competitor and the target form are [p], [r], and [s], which appear in the same order in both words. If the order of the phonemes were not taken into account, the segment [n] as well as the schwa would also have been included in this calculation. The number of matching phonemes was then divided by the total number of phonemes of the

reduced-form competitor. Similar comparisons were made between the canonical-form competitors and their target forms. A t test showed no differences in segmental overlap between the overlap values for the reduced-form and the canonical-form competitors, $t(62) = -0.18$, $p > .1$.
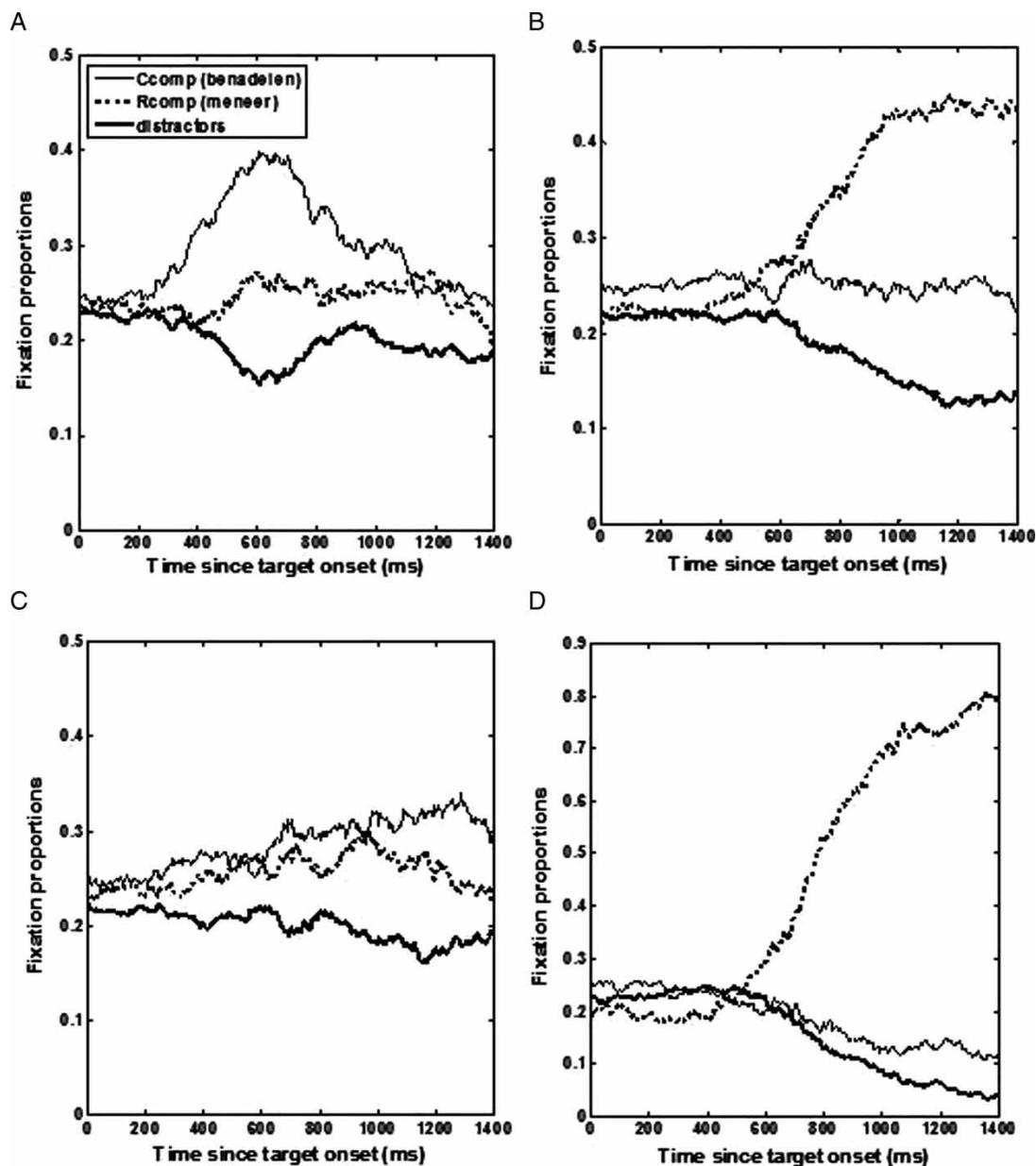

Figure 2. *Mean fixation proportions to the "canonical form" competitor (Ccomp), the "reduced form" competitor (Rcomp), and the averaged distractors, in (A) canonical forms in isolation (all trials); (B) reduced forms in isolation (all trials); (C) reduced forms in isolation (correct trials); and (D) reduced forms in isolation (incorrect trials).*

This result thus excludes the potential confound that there is a difference in the amount of phonological overlap between competitors and their targets.

The target word mentioned in the casual speech fragments was absent from the visual display (cf. Huettig & Altmann, 2005). The displays of the experimental trials never contained printed target words corresponding fully to the spoken target words. In these cases, the participants' task was to click in the middle of the screen. To decrease the likelihood that participants would use a strategic form of processing that might be more indicative of postlexical processing, we added filler items in which one of the visual words on the screen matched with the auditory target stimulus. We created twice as many filler (128) as experimental items (64). Thus, in two thirds of the cases participants had to click on one of the four visually presented words on the screen, whereas in one third of the cases listeners had to click in the middle of the screen. Since the display for the experimental trials contained two phonologically related words (i.e., the two competitors), we masked this pattern in the filler items in which a visual target was always present. Participants had to click half of the time on one of the phonologically similar words and half of the time on a word that was phonologically unrelated. In this way, the fillers prevented participants from developing any expectation that items sharing certain phonological attributes would be mentioned.

The experimental and filler items were put into one list, and the order was randomized, so that each participant got a different order of presentation. The position of the three types of printed words was randomized over the four quadrants on the screen. That is, the reduced-form competitor, the canonical-form competitor, and the distractors appeared with equal probability on each of the four screen position over the course of an experimental run. Besides experimental and filler items, we also selected six practice items from the Spoken Dutch Corpus (Oostdijk, 2000) to familiarize participants with the task. Half of the practice items contained a target on the screen; the other half did not.

Procedure

Participants were tested individually. They were seated in a sound-attenuated booth at a comfortable viewing distance from the computer screen. The eye-tracking system was mounted and calibrated. Eye movements were monitored using an SMI EyeLinkII system, sampling at 250 Hz. Auditory stimuli were presented to the participants over headphones using the Nijmegen Experiment Set-Up software.

Participants received written instructions on the screen. Participants had to click with the computer's mouse on the printed word in the visual display representing the word they heard in the auditory stimulus. If none of the printed words matched with the auditory stimulus—as for all experimental trials—participants had to click in the middle of the screen.

Each trial had the following structure. First, a grid with four printed words appeared in a 24-point Courier font on the screen. The centres of the printed words corresponded, independently of the length of the words, to the centres of the quadrants on the screen. After 2,500 ms, the auditory stimulus was presented. The preview time in the current study thus was longer than the one-second preview typically used in eye-tracking studies with lab speech (e.g., McQueen & Viebahn, 2007). There are two reasons why we chose to use this longer preview time. First, our target sentences are more complex than target sentences often used in lab studies (e.g., *'t ergste is nog als de wedstrijd dus afgelopen is* 'the worst thing is if the match is finished' versus *Klik op het woord X* 'Click on the word X'). Second, the position of the target was unpredictable in our sentences, whereas in lab studies the target word often follows the sentence frame 'Click on the word'. We chose the longer preview time to ensure that participants would have enough time to read the four printed words. A longer preview time is less critical in 'Click on the word X'-lab studies because simple carrier sentences and predictable target word positions allow for more concurrent processing of the display. The next trial was initiated after participants clicked with the mouse on the screen. Participants were put under no

time pressure to perform this action. Every five trials, a central fixation cross appeared centred on the screen, permitting for drift correction in the calibration.

After the six practice trials, the 64 experimental and 128 filler items were presented in random order. The experimental session took approximately 15 min.

Design and analysis

The dependent variables were click-responses and eye movements. For the click-responses we calculated the percentage of correct rejections—that is, the percentage of clicks in the middle of the screen and the percentage of incorrect clicks on the three word types. A statistical analysis of the error pattern was carried out with linear mixed-effects models using a logistic linking function (cf. Dixon, 2008). We did not analyse response times. Response times are an uninformative measure because the auditory target word was never present on the screen for the experimental trials.

For the eye-tracking data, we discarded blinks and saccades and analysed the data from the right eye of the participants. As it takes time to programme and initiate a saccadic eye movement (typically about 100 to 200 ms, cf. Matin, Shao, & Boff, 1993), we analysed successive 100-ms time windows from 200 ms after target onset.

For the analysis, we transformed the proportion data with the empirical logit function (cf. Barr, 2008). From these data we constructed two linearly independent measures: (a) overall competition effects: mean of looks to both competitors versus mean of looks to both distractors; and (b) specificity of the competition effects: mean of looks to the canonical-form competitor versus mean of looks to the reduced-form competitor. All measures are difference measures, so that a difference from zero indicates a preference for one type of stimulus.

We tested whether these measures were influenced by word form (i.e., canonical forms versus reduced forms) using linear mixed-effects models (Baayen, Davidson, & Bates, 2008) with participants and items as random effects and in which word form was coded as a numeric contrast (−0.5 and 0.5, cf. Barr, 2008). Canonical forms were

coded as −0.5 and reduced forms as 0.5. A negative beta indicates that the dependent variable has a higher value for the canonical form condition, whereas a positive beta indicates that the dependent variable has a higher value for the reduced form condition. Note that the interpretation of the beta depends on the dependent measure. In the case of the accuracy measure, a positive beta reveals more errors in the reduced form condition than in the canonical form condition—and hence that recognizing reduced forms is more difficult than recognizing canonical forms. In the case of the overall competition measure, a positive beta indicates more overall competition in the reduced form condition than in the canonical form condition. A similar interpretation holds for the specific competition measure: A positive beta implies more specific competition in the reduced-form condition than in the canonical form condition. We estimated p-values by using Markov chain Monte Carlo simulations (Baayen et al., 2008).

Results

Table 1 shows the percentage of correct rejections and the percentage of incorrect click responses to the three word types. An analysis on the error rates showed a main effect of word form ($\beta_{\text{word form}} = -4.26$, $p < .0001$), indicating that listeners made more errors in the reduced-form condition than in the canonical form condition.

Table 1. *Task performance in Experiment 1*

Response type	Forms presented in isolation	
	Canonical forms	Reduced forms
Correct rejections	97.1	65.9
Incorrect responses		
Canonical-form competitor	1.6	2.7
Reduced-form competitor	1.3	31.0
Distractors	0	0.4

Note: Performance is shown as the percentages of click responses. "Correct rejections" correspond to clicks in the middle of the screen. "Incorrect responses" are incorrect responses to the three word types.

Listeners clicked 31% of the time on the reduced-form competitor when listening to reduced forms.

For the eye movement data, we plot the fixation proportions of all trials (including the correct and incorrect responses) as well as for correct responses separately. In typical eye-tracking experiments with careful speech, error rates tend to be low (<5%), and errors are typically discarded. In the current case, however, with more than 30% of errors, simply discarding the errors is problematic, because it would exclude the (apparently) most difficult trials with the most severe reductions. Nevertheless, misidentifications obviously lead to prolonged looks at a competitor, simply because the competitor is clicked on. Therefore, Figure 2 shows the mean fixation proportion to the two types of competitors and the averaged distractors from acoustic target onset (0 ms) to 1,400 ms thereafter for (A) canonical forms (all trials), and in three different plots for reduced forms. Figure 2B shows the data for all trials, Figure 2C for correct trials only, and Figure 2D for incorrect trials. These additional plots give us insight into how the competition pattern changes depending on participants' performance on a trial. Fixation proportions were analysed during an early (400–800 ms) and a late time window (800–1,200 ms).

All trials

We first analysed whether there was an effect of overall competition (competitors − distractors). Figure 3 shows the beta weights in the different time windows. Note that the error bars show confidence intervals, so that a beta weight for which the error bars do not include the x-axis are significant. The top panel of Figure 3 shows that an overall competition effect arises in the time window starting 400 ms after target onset, and participants look more at competitors than at distractors in all later time windows. Overall competition was also dependent on word form. In the early time windows (500–600 ms), overall competition was stronger in the canonical form condition as indicated by the negative regression weights. In the late time windows (900–1,200 ms), however, overall competition was strongest in the reduced-

form condition as indicated by the positive regression weights.

Second, we analysed whether listeners looked more often to canonical-form competitors than to reduced-form competitors (canonical-form competitor − reduced-form competitor). As the bottom panel of Figure 3 shows, we found an overall effect of specific competition in the 300–500-ms time window. This effect was modulated by word form. That is, it varied over conditions whether canonical-form or reduced-form competitors received more looks.

To further investigate this pattern, we analysed the effect of specific competition in each condition separately for those time windows in which there was a significant effect of word form on the specific competition measure. This strategy is analogous to the breaking-down of an interaction in factorial analysis of variance (ANOVA) designs. That is, we tested whether there was a significant preference for one of the competitors in both the reduced and the canonical form condition. Note that the measure is defined as canonical-form competitor − reduced-form competitor. A positive intercept hence means more looks to the canonical-form competitor. The results of these analyses are added as subscripts to Figure 3 (bottom panel). If there was a significant difference in the canonical form condition, this is indicated by a "C"; if there was a significant difference in the reduced-form condition, this is indicated by an "R". Signs are used to indicate the direction of the difference. As an inspection of the figure then reveals, there was a preference for the canonical-form competitor in the canonical form condition in the early time windows (the "C +" subscripts for the time windows 400–700 ms) and a preference for the reduced-form competitor in the reduced-form competition in the late time windows (>700 ms).

Correct trials

For the correct trials, we used the same strategy as that for all trials. First, it was analysed whether listeners looked more often at competitors than at distractors (overall competition). As Figure 4 (top panel) shows, participants looked more at competitors than at distractors from 400 ms onwards. In
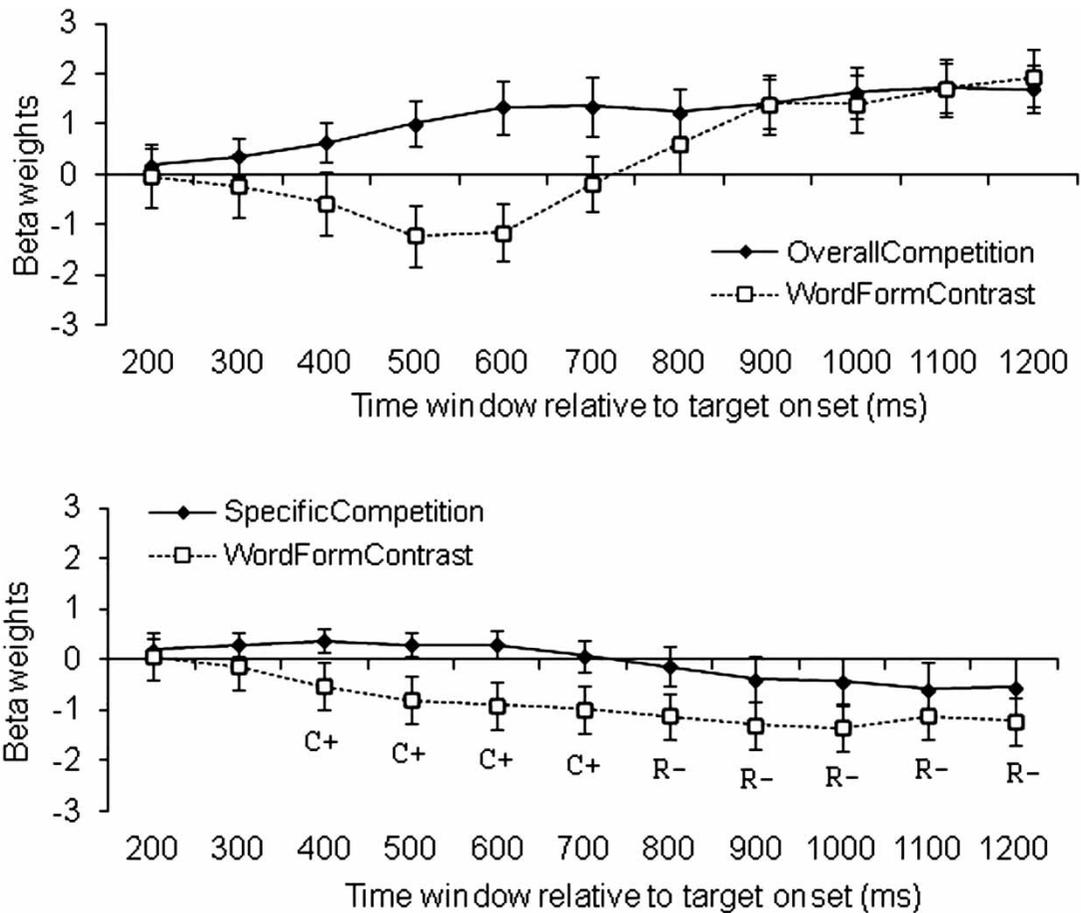
NB: WordFormContrast: canonical targets vs. reduced targets;
OverallCompetition: (canonical-form competitor + reduced-form competitor) vs. averaged distractors;
SpecificCompetition: canonical-form competitor vs. reduced-form competitor.

Figure 3. *Time window analysis of the eye-tracking data (isolation condition, all trials) based on the beta weights (logistic units). The top panel shows the beta weights in different time windows for the measure overall competition, and the bottom panel shows the beta weights for the measure specific competition.*

two time windows (500–700 ms) there was more competition in the canonical form condition, as indicated by the negative regression weights.

Second, we analysed whether there was an effect of specific competition (see Figure 4, bottom panel). These analyses showed an overall effect in the time windows from 400 to 700 ms, which was modulated by word form in three of these four time windows. When analysing this effect in each condition separately, there was only a preference for the canonical-form competitor in the canonical form condition in the time windows from 400 to 600 ms after target onset. The specific competition effect was not significant in the reduced-form condition in any time window.

These results suggests that even in the trials in which participants made the correct decision they may not (or not always) have recognized the

NB: WordFormContrast: canonical targets vs. reduced targets;
OverallCompetition: (canonical-form competitor + reduced-form competitor) vs. averaged distractors;
SpecificCompetition: canonical-form competitor vs. reduced-form competitor.
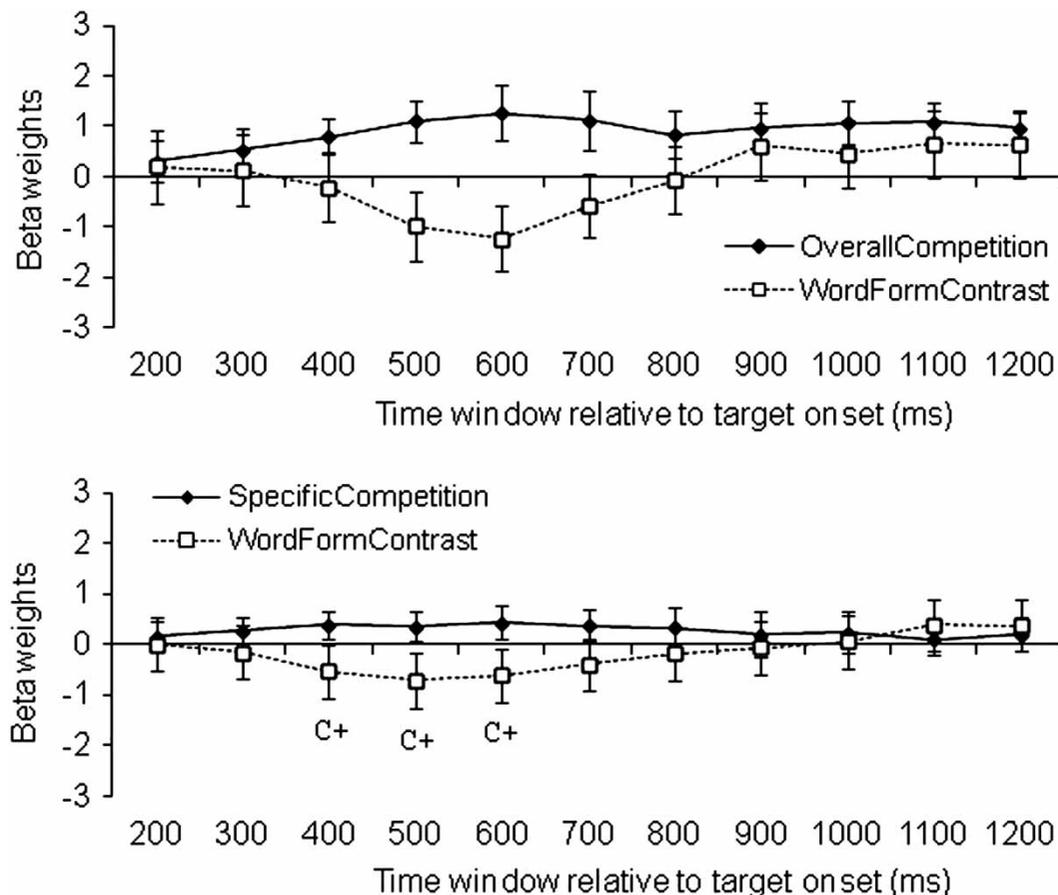
Figure 4. *Time window analysis of the eye-tracking data (isolation condition, correct trials) based on the beta weights (logistic units). The top panel shows the beta weights in different time windows for the measure overall competition, and the bottom panel shows the beta weights for the measure specific competition.*

target word. Participants may instead have based their decision on overlap in perceived phonemes without making contact with the lexicon. By contrast, for the incorrect trials, we observe a clear preference for the reduced-form competitors, suggesting that incorrect trials reveal numerous and sustained fixations to the word type that was chosen (i.e., reduced-form competitor), as people guide the mouse cursor toward the object they intended to click on.

Discussion

In Experiment 1, we observed that overt attention to canonical-form and reduced-form competitors in a printed-word display can be influenced by the exact form of the acoustic input. This therefore suggests that (at least in task situations such as the present, which maximize the opportunity for competition) canonical forms can activate different competitors more strongly than reduced forms.

This suggests that the previous findings by Brouwer et al. (2012) were most likely due to the task situation rather than phonetic detail in the stimuli. There is, however, one alternative interpretation. It is conceivable that by presenting isolated fragments, we limited the listeners' ability to exploit the phonetic detail. Consider the example of the reduced form (e.g., [pjutər] for *computer*) we used in the introduction. We had thought that the long closure duration might be a cue that tells listeners that the /p/ is not (underlyingly) word-initial. However, with a single-word presentation, the closure duration is not even audible, making it unlikely that this phonetic detail could influence the competition process. In addition to this, listeners might need information about the surrounding speech rate, or coarticulation with preceding sounds, in order to interpret the reductions.

Moreover, the click-responses showed that listeners made more errors in the reduced-form condition than in the canonical-form condition. Listeners often clicked on the reduced-form competitor in the reduced-form condition. This offline preference for the reduced-form competitor is in line with the online eye movement data. Listeners looked most often at the reduced-form competitor when listening to reduced forms. These results accord with the findings of Ernestus et al. (2002) and of Arai (1999) who showed that reduced forms are difficult to recognize on the basis of the acoustic form alone.

These two observations pave the way for the subsequent experiments. Given the fact that performance for reduced forms in isolation is seriously compromised (34% error rate), the question arises how much context is necessary to allow recognition. In Experiment 2, we examined phonological competition processes when reduced forms (and canonical forms) are presented in a sentential context. Previous work showed that in such a situation, performance increases considerably (Kemps et al., 2004). This experiment therefore allows us to examine the phonological competition process during the actual *recognition* of strongly reduced forms. For canonical forms, we predict the same competition pattern as in Experiment 1: more looks to the canonical-form competitor than to

the reduced-form competitor (cf. Allopenna et al., 1998; McQueen & Viebahn, 2007). Two outcomes are conceivable for reduced forms. If listeners in Experiment 1 were simply unable to make use of the phonetic detail when they heard short fragments (cf. Ernestus et al., 2002), we should now again replicate the pattern observed in Brouwer et al. (2012): Competitor fixations are not modulated by the input form. If, however, it was the task situation (i.e., the absence of the target word in the display) that resulted in increased competition effects, we should still observe differences in competitor fixations contingent upon the input form.

EXPERIMENT 2

Method

Participants

Twenty-four undergraduates were paid to participate in this experiment. All participants were native speakers of Dutch. They reported no hearing problems and normal or corrected-to-normal vision. None of the listeners participated in the previous experiment.

Materials

We used the same 32 canonical and 32 reduced realizations as those in Experiment 1, but they were now embedded in a context of several words. As a consequence of the nature of these materials, the amount of reduction was also reflected not only in the target word, but throughout the sentence, so that the phoneme deletion rate was higher for the sentences containing reduced forms (22%) than for sentences containing the canonical forms (6%).

Note that the context for a canonical form (e.g., *ook naar beneden die sluit dan aan* 'the one going downwards, as well, this connects then to') always differed from that of a reduced form (e.g., *buigt het zo af en dan valt het naar beneden, dat is echt* … 'it bends like this and then it falls down, that's really…') because they occurred in different corpus utterances. We therefore conducted a

(web-based) cloze test to investigate whether the different contexts induced a preferential bias for certain word types (i.e., target, canonical-form competitor, reduced-form competitor, and one of the distractors), which might have caused confounds in our material. This test measured the predictability of the target word given the preceding context in canonical and reduced utterances. For both types of utterances, the words preceding the target were presented on the screen. In the first part, participants ($n = 35$) had to finish the sentence freely with three to seven words suitable to the context. In the second part, the sentence was again shown on the screen, but now the eventual target, the two types of competitors, and one of the distractors were provided. The participants had to rank these words in the order of how likely they completed the sentence.

In the first open-ended part of the cloze test, participants named the target word on 5.8% of the trials (6.2% in the canonical form sentences, 5.4% in the reduced-form sentences). These results suggest that some target words were indeed somewhat predictable given their preceding linguistic context. The target words were, however, not more predictable in the sentences in which they happened to be reduced. The participants never named a competitor with the exception of one occurrence of a reduced-form competitor ($<1\%$).

In the second part, participants rated the target word as the most likely option (in 81.6% of the trials). The mean rank of the target word was hence close to 1, and this did not differ between sentences with canonical forms (1.25) and sentences with reduced forms (1.30). We compared the mean rank of both competitors for both types of sentences (i.e., sentences with canonical forms and sentences with reduced forms) to test whether there was a difference in terms of semantic compatibility of the canonical-form competitor and the reduced-form competitor. The mean rank in all four cases was around 3 (canonical-form competitors: 3.07 in the canonical form sentences and 2.94 in the reduced-form sentences; reduced-form competitors: 2.94 in the canonical-form sentences and 2.84 in the reduced-form sentences). It is hence unsurprising that there were no significant

differences as evaluated with a two-by-two repeated measures ANOVA with competitor type and sentence type as predictors, $F_{\text{SentenceType}}(1, 30) = 1.67$, $p > .1$ (all other Fs > 1).

Procedure, design, and analysis

The procedure and design were identical to those in the previous experiment. For the analysis, we used the same two measures as those in the previous experiment (i.e., overall competition and specific competition), but we analysed the correct trials instead of all trials. The experimental session took approximately 25 min.

Results

Table 2 shows the percentage of correct rejections and the percentage of incorrect click responses for the three word types. An analysis on the error rates showed a main effect of word form ($\beta_{\text{word form}} = -2.83$, $p < .0001$), indicating that listeners performed better in the canonical form (99.5%) than in the reduced-form condition (93.8%).

For the eye movement data, we plot and analyse the fixation proportions of correct trials only for the canonical and reduced forms, since there were only few errors. Figure 5 shows the mean fixation proportion to the two types of competitors and the averaged distractors from acoustic target onset (0 ms) to 1,400 ms thereafter for (A) canonical

Table 2. *Task performance in Experiment 2*

Response type	Forms presented in a sentential context	
	Canonical forms	Reduced forms
Correct rejections	99.5	93.8
Incorrect responses		
Canonical-form competitor	0.1	1.8
Reduced-form competitor	0.4	4.4
Distractors	0	0

Note: Performance is shown as the percentages of click responses. "Correct rejections" correspond to clicks in the middle of the screen. "Incorrect responses" are incorrect responses to the three word types.

forms and (B) reduced forms in a sentential context. As in the previous experiment, we analysed fixations in successive 100-ms time windows.

We first analysed whether there was an effect of overall competition (competitors – distractors). Figure 6 (top panel) shows the beta weights and their confidence intervals for all time windows. An inspection of the time course shows that an overall effect of competition was significant in the time windows from 300 ms to 1,000 ms after target onset. The amount of overall competition was not dependent on word form in any of the time windows.

Second, the bottom panel of Figure 6 shows the comparison of looks to the canonical-form competitors versus the reduced-form competitors. This contrast shows no overall significant effect. However, there are significant regression weights for word form, indicating that the preference for one of the competitors over the other differed between conditions. The negative betas indicate that specific competition was strongest in the

canonical form condition. Importantly, the results from the reduced-form condition patterned similarly to the results in the canonical form condition in the later time windows (>800 ms).

As in the previous experiment, we further analysed the effect of specific competition in each condition separately. This showed that, in the canonical-form condition, there was a significant preference for the canonical-form competitor in the time windows 400–600 ms after target onset (indicated by the "C +" subscripts in Figure 6, bottom panel). There was no significant preference for any competitor in the reduced-form condition, although there is a numerical preference for the reduced-form competitor in the earlier time windows and a preference for the canonical-form competitor in the later time windows.

Discussion

Experiment 2 investigated phonological competition during listening to strongly reduced forms
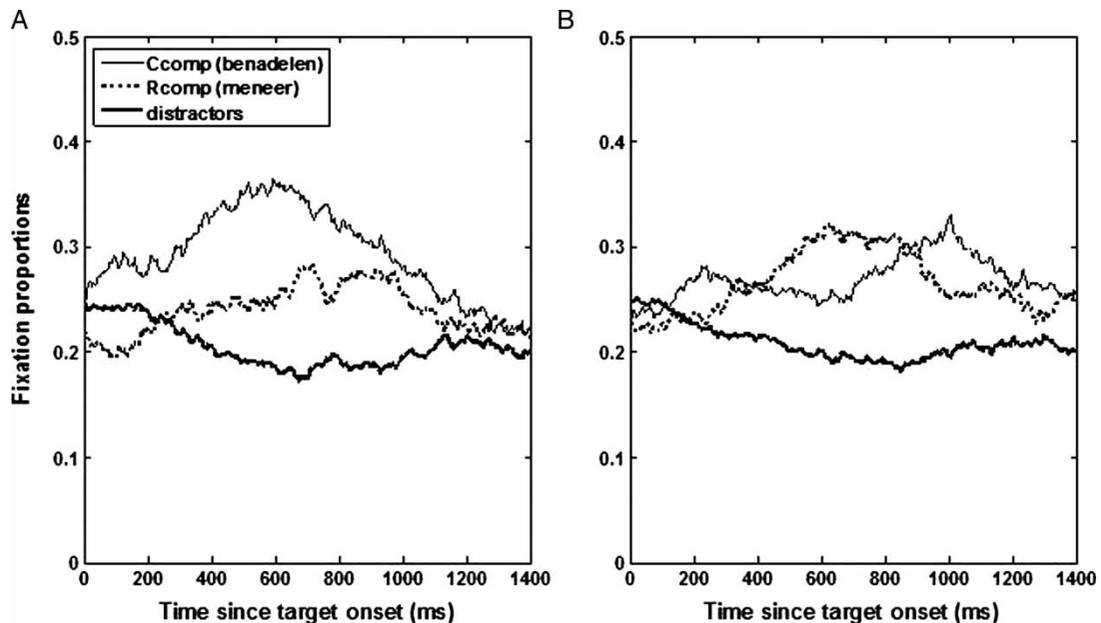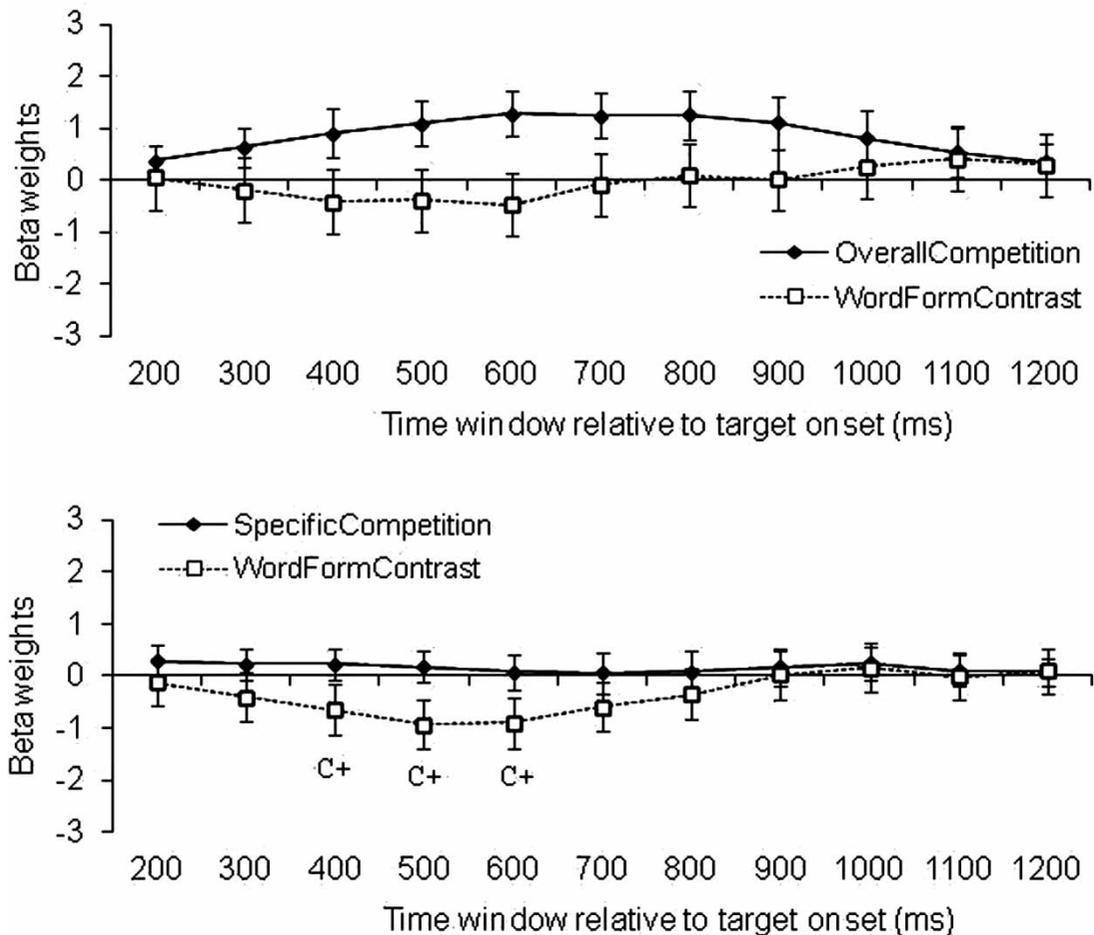

Figure 5. *Mean fixation proportions to the "canonical form" competitor (Ccomp), the "reduced form" competitor (Rcomp), and the averaged distractors, in (A) canonical forms and (B) reduced forms for the correct trials presented in a sentential context.*

Figure 6. *Time window analysis of the eye-tracking data (sentence condition, correct trials) based on the beta weights (logistic units). The top panel shows the beta weights in different time windows for the measure overall competition, and the bottom panel shows the beta weights for the measure specific competition.*

in a sentential context. The error pattern showed that reduction still inhibits word recognition (6% error rate), but to a much lesser extent than in Experiment 1 (34% error rate). This result is in line with the findings by Arai (1999) and Ernestus and colleagues (2002): Listeners benefit from phonetic and semantic/syntactic context during the recognition of strongly reduced forms.

The eye movement data revealed that the competition pattern for canonical forms in Experiment 1 was replicated: The canonical-form competitors attracted more attention than the reduced-form competitors in the early time window. These results are similar to the competitor effects that have been found by Allopenna et al. (1998) and McQueen and Viebahn (2007) because our

canonical-form competitors are similar to the phonological "cohort" competitors in those studies. Over time, as the acoustic form of the canonical-form competitor became inconsistent with the acoustic input, the preference of looks to the canonical-form competitor disappeared.

When the input was a reduced form, however, we observed that, in the early time window, the reduced-form competitors attracted more looks than the canonical-form competitors. Note, however, that the interaction was significant, but that the specific competition effect did not reach significance in the reduced-form condition separately. There was a numerical preference for the reduced-form competitor even though participants heard a complete dialogue fragment —that is, additional phonetic context enabling participants to better exploit any fine phonetic detail cues. This suggests that reduced-form competitors (i.e., unrelated words overlapping in phonemes with the reduced form of the spoken target word) can compete early during the recognition of reduced forms. In other words, there is a better match of the acoustic signal (i.e., [məneə]) with the phonological representation activated from the reduced-form competitor (i.e., *meneer*), than with the canonical-form competitor (i.e., *benadelen*).

In contrast, later in time, canonical-form competitors did attract numerically more attention than phonologically unrelated distractors, similar to the offset-matching (or "rhyme") competitor effects that have been reported by Allopenna et al. (1998) and McQueen and Viebahn (2007). Importantly, however, the offset-matching competitors in the Allopenna et al. and McQueen and Viebahn's studies always attracted fewer looks than the onset overlap competitors. This, however, is not the case in the early time window in our Experiment 2. This suggests that late looks to our canonical-form competitor (i.e., *benadelen*) reflects more than just overlapping phonemes. It suggests that participants reconstructed the reduced form [məneə] to its canonical (or "citation") form (i.e., [bənedə]). Such a reconstruction process may be time consuming, therefore resulting in the later shift in eye gaze to the canonical-form

competitor. This also explains why the previous work using offline tasks always only documented a role for the citation form when listeners were confronted with strongly reduced forms. Kemps et al. (2004), for instance, found that listeners judge phonemes that are phonetically absent in reduced forms as present in a phoneme monitoring task. That is, when asked to monitor for /l/, they responded with "yes" to the phonetic form [tyk], which is a reduced form of /natyrlʏk/ *natuurlijk* 'naturally'. This result is in line with the assumption of a late reconstruction process. The offline task used by Kemps et al. only revealed what happens late in time when the acoustic input form no longer had any influence.

Experiment 2 revealed that reduced-form competitors competed for eye gaze even when participants had the opportunity to make use of additional phonetic context indicating a speech reduction. Does this mean that there is little phonetic detail to exploit? Experiment 3 tests this directly. We examined whether listeners are sensitive to fine phonetic detail information in interpreting whether a reduced form was heard or not. Previous research showed that listeners are good at exploiting the fine phonetic detail of utterances to recognize intended words even when spontaneous speech processes have changed them so that they deviate from their canonical form. Gow (2002), for example, showed that listeners make use of fine phonetic detail to solve the lexical ambiguity produced by place assimilation. He showed, for example, that if the compound noun *right berries* is assimilated to *ripe berries*, the assimilated [p] differs from the unassimilated form [rɑipbɛriz] from 'ripe berries'. The assimilated [p] still bears some cues for an alveolar place of articulation. This finding has two consequences. First, if the listener is presented with a strong [p] in the phrase [rɑipbɛriz], *ripe* is accessed; however, a slightly weaker [p] also activates *right*. In addition, the slightly weaker [p] facilitates the recognition of the upcoming labial segment (see also Gow, 2001, 2003). If listeners make use of fine phonetic detail in strongly reduced forms, this could potentially help them in interpreting whether a reduced form was heard or not and thus whether they should

attempt a reconstruction to a possible canonical form.

In Experiment 3, we take a similar approach as in Gow's (2002) experiments. Analogous to the use of assimilated and intended segments in Gow's experiments, we used cross-splicing to replace the acoustic realization of a "surface" segment in a reduced form with an "intended" segment. For example, in the reduced form [məneə] from the canonical form [bənedə] *beneden* 'downwards' the "underlying" segment [b] has changed into the "surface" segment [m]. We replaced this "surface" segment [m] with an "intended" segment /m/ from the same speaker, which did not arise from reductions (e.g., [mɛt] *met* 'with'). We examined whether listeners are sensitive to the subtle difference between the "surface" segment [m] and an "intended" segment /m/. In other words, will the "surface" segment [m] be comparable to an "intended" segment /m/ or will the "surface" segment [m] still contain traces of the "underlying" segment [b]? We predicted that the cues of the "intended" segment /m/ would bias listeners' interpretation of the reduced form [məneə] toward the reduced-form competitor (e.g., *meneer* 'mister').

EXPERIMENT 3

Method

Participants
We tested 33 Dutch native speakers, who were paid for their participation. The participants reported normal hearing and normal or corrected-to-normal vision. None of the participants took part in the previous experiments.

Materials and procedure
We searched for the same segment from the same speaker in the Spoken Dutch Corpus (Oostdijk, 2000) to replace the "surface" segment [m] in the reduced form [məneə] with an "intended"

segment /m/. Additionally, we attempted to find the same segmental context surrounding a "surface" segment as that in the original speech fragment.

For 23 out of the 32 reduced forms, we found appropriate "intended" segments to do the cross-splicing manipulation. This experiment thus used a subset of the same materials as those used in the previous experiments. For example, for the reduced form [məneə] (from the context: *buigt het zo af en dan valt het naar beneden, dat is echt…* 'it bends like this and then it falls down, that's really…') we found an onset /m/ in the word *met* /mɛt/ 'with' in the context *Je kan altijd een keer met korting reizen* 'You can always travel sometime with discount'. If the crucial reduction in a word form occurred in medial or offset position, such as in [ʋɛs] for [ʋɛtstrɛit] *wedstrijd* 'match', we looked for the critical segments in the same position. Thus, for the reduced form [ʋɛs] (from context: *'t ergste is nog als de wedstrijd dus afgelopen is* 'the worst thing is if the match is finished') we found the "intended" segments /ɛs/ in the word *blessuretijd* 'injury time' in the context *en uh en en blessuretijd* 'and uh and and injury time'.

After finding the appropriate "intended" segments, we, for example, deleted the "surface" segment /m/ in [məneə] and replaced this segment by an "intended" onset /m/ from *met* 'with'.[3] Similarly, for the "surface" segments /ɛs/ in [ʋɛs], we replaced this by an "intended" mid /ɛs/ from *blessuretijd* 'injury time'. However, before replacing the "surface" segment with the "intended" segment, we edited the "intended" segment with the PSOLA component of the PRAAT software package (Boersma, 2001) to make the fit as good as possible. First, we made the "intended" segment as long as the "surface" segment. Secondly, we resynthesized the "intended" segment with the original pitch contour of the "surface" segment. Additionally, we gave the "intended" segment the same amplitude as the "surface" segment. In case it

---

[3] Note that it is problematic to do controlled acoustic measurements on the "surface" and the "intended" segments. Obviously, it is possible to do measurements, but there is a need for good control tokens. All segments come out of different contexts; therefore, any obtained measure depends on different speakers, different prosodies, and different quality of the sounds. Most of the sentences contain quite some noise, which also prevented us from doing good controlled measurements on the segments.

was necessary, we also added noise to the "intended" segment to approximate the noise level of the "surface" segment (see Figure 7). The segments were spliced at zero-crossings, and we kept the glottal phases intact to avoid splicing problems. Following these manipulations, we performed informal listening tests on the original and the transformed sound files to ensure that the stimuli maintained their naturalness after signal processing.

As in Experiment 2, the cross-spliced forms were presented in a sentential context. We used the same eye-tracking display as that in the previous experiments. Note that the experimental items consisted only of the 23 cross-spliced forms. We also selected 36 filler items. The procedure was identical to that in the previous experiments. The experimental session took approximately 10 min.

Design and analysis

This experiment consisted of only one condition—that is, the cross-spliced form condition. A comparison was made between this condition and the reduced-form condition of Experiment 2 to investigate whether competition works differently for both forms. Note that the nonspliced reduced-form data in this experiment are a subset (23 items) of the data in Experiment 2 (see Figure 5B). In the analysis, we therefore compared these cross-spliced items with the same 23 reduced forms of Experiment 2. Thus, the analyses are primarily a reanalysis of a subset of Experiment 2's reduced tokens. Word form was coded as a numeric contrasts, in which reduced forms were

coded as –0.5 and cross-spliced forms as 0.5. The same analysis was performed as that in Experiment 2.

Results and discussion

Table 3 shows the percentage of correct rejections and the percentage of incorrect click responses for the three word types. An analysis on the error rates showed a main effect of word form ($\beta_{word\,form} = -1.25$, $p < .0001$), indicating that listeners performed more accurately in the reduced-form condition (93.7%) than in the cross-spliced form condition (88%).

For the eye movement data, we plotted and analysed the fixation proportions of correct trials only for the cross-spliced and the reduced forms, since there were only few errors. Figure 8A shows the mean fixation proportion to the two types of competitors and the averaged distractors from acoustic target onset (0 ms) to 1,400 ms thereafter for the cross-spliced form condition. In addition, we plotted the same 23 items of the reduced-form condition of Experiment 2 in Figure 8B. Note that this subset shows a similar competition pattern as the 32 reduced items in Experiment 2 (see Figure 5B). In the analysis, we compared fixations in the cross-spliced form condition with the reduced-form condition containing only those 23 items. As in the previous experiments, we analysed fixations in successive 100-ms time windows.

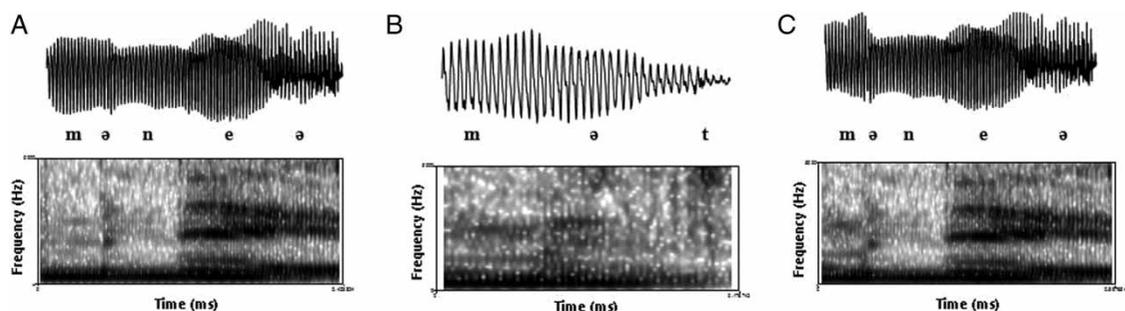We first analysed whether there was an effect of verall competition (see Figure 9, top panel). We

Figure 7. *Realizations of (A) the "surface" segment [m], (B) the "intended" segment /m/, and (C) the cross-spliced [m]. See text for details.*

Table 3. *Task performance in Experiment 3 and comparison with performance in Experiment 2*

	Forms presented in a sentential context	
Response type	Experiment 3: Cross-spliced forms	Experiment 2: Reduced forms (23 items)
Correct rejections	88.0	93.7
Incorrect responses		
Canonical-form competitor	1.8	1.4
Reduced-form competitor	10.2	4.9
Distractors	0	0

Note: Performance is shown as the percentages of click responses. "Correct rejections" correspond to clicks in the middle of the screen. "Incorrect responses" are incorrect responses to the three word types. Comparison was made with the nonspliced reduced-form data, which are a subset of the data from Experiment 2.

found an overall effect from 500 ms onwards. Overall competition was dependent on word form in early time windows, with stronger competition for the

original forms in the time windows starting 300 ms and 400 ms after target onset.

Second, we analysed whether listeners looked more often to canonical-form competitors than reduced-form competitors (see Figure 9, bottom panel). We found no overall effect of specific competition in any time window. Specific competition was, however, modulated by word form in the time window starting 900 ms and 1,000 ms after target onset. Further analysis showed that this was driven by a preference for reduced-form competitors in the cross-spliced form condition and a preference for canonical-form competitors in the original reduced-form condition. Both preferences were, however, not significant by themselves. To further analyse this pattern, we compared eye gaze on the reduced-form competitor across the reduced form and the cross-spliced form condition. We found that splicing significantly increased eye gaze to reduced-form competitors in the time windows starting 900 and 1,000 ms after target onset ($\beta_{word\ form} = 0.42$, $p < .05$, and $\beta_{word\ form} = 0.47$, $p < .05$).
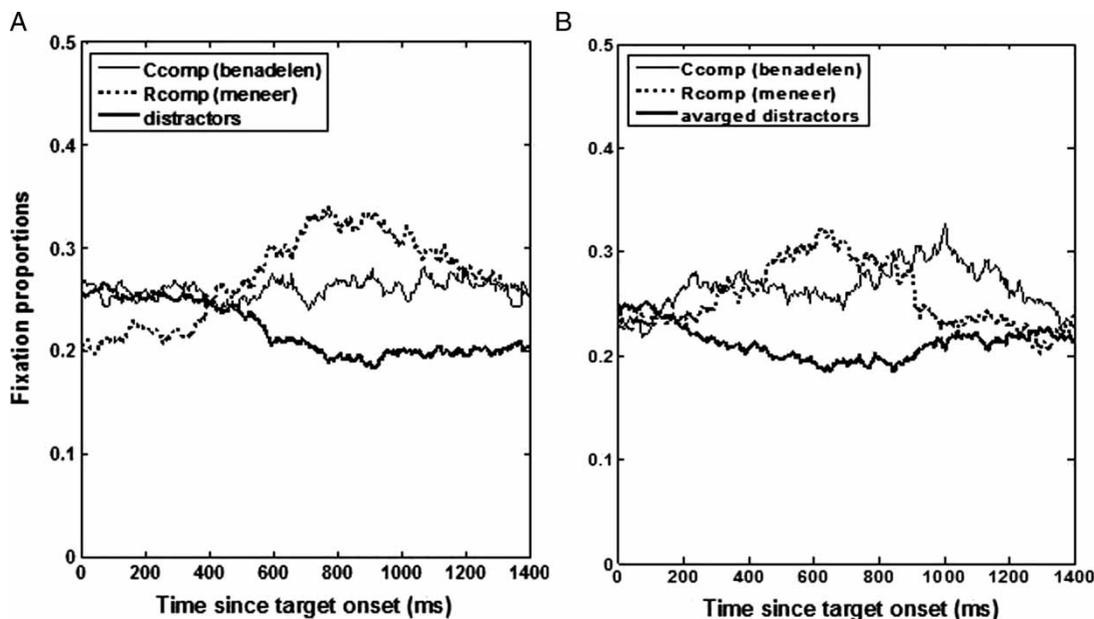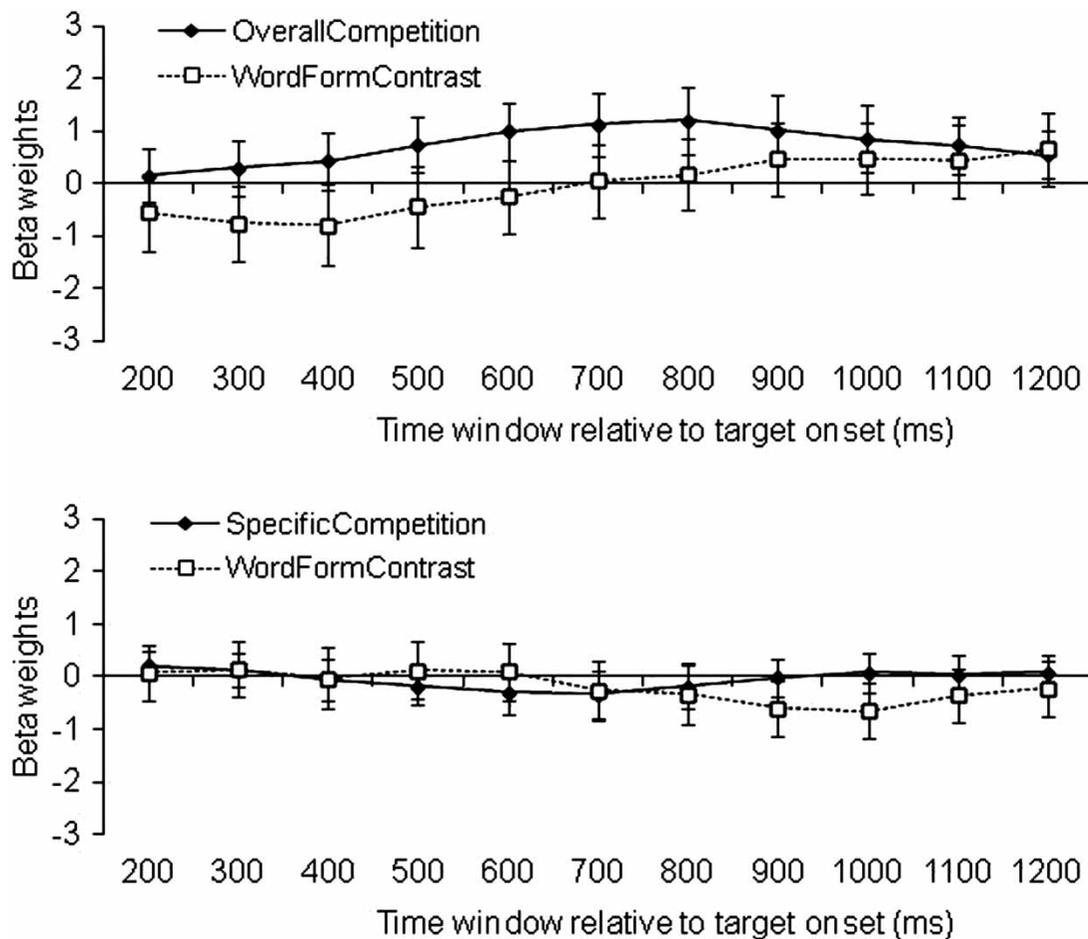

Figure 8. *Mean fixation proportions to the "canonical form" competitor (Ccomp), the "reduced form" competitor (Rcomp), and the averaged distractors, in (A) cross-spliced forms, and (B) reduced forms for the correct trials presented in a sentential context (a subset of 23 reduced tokens of Experiment 2).*

NB: WordFormContrast: canonical targets vs. reduced targets;
OverallCompetition: (canonical-form competitor + reduced-form competitor) vs. averaged distractors;
SpecificCompetition: canonical-form competitor vs. reduced-form competitor.

Figure 9. *Time window analysis of the eye-tracking data (splicing condition, correct trials) based on the beta weights (logistic units). The top panel shows the beta weights in different time windows for the measure overall competition, and the bottom panel shows the beta weights for the measure specific competition.*

In sum, Experiment 3 examined whether phonological competition is modulated if we purposely change segmental information in the reduced form itself. The error pattern showed that listeners made more errors in the cross-spliced form condition than in the reduced-form condition. Participants often clicked on the reduced-form competitors during listening to cross-spliced forms, indicating that listeners interpret the cross-spliced [m] more

often as the "intended" segment [m] than as the "surface" segment [m].

The eye movement data showed that the late rise of the canonical-form competitor in the reduced-form condition was present less in the cross-spliced form condition. Instead, the reduced-form competitor attracted numerically more looks in the cross-spliced form than in the reduced-form condition. This may indicate that the cross-

spliced segments were interpreted as real segments, and that reduced forms contain residual cues with fine phonetic information of the canonical form. Thus listeners are likely to be more sensitive of fine phonetic information in interpreting whether a reconstruction process is likely to be involved or not. The initial onset /m/ suggests that *beneden* is unlikely to have been intended, and thus no (or a reduced) late reconstruction process takes place.

Note that it is difficult to establish what the listeners were exactly picking up from the cross-spliced segments. The differences between the original and the spliced stimuli were very small because we matched their duration, pitch, and amplitude. We therefore only showed that these cues do not contribute significantly to the present findings. It is most likely that spectral differences influenced the results. Such differences also appear to be important for the interpretation of assimilated segments (e.g., Gow, 2002). For further research it would be interesting to investigate the role of duration more explicitly. This is difficult to achieve for the current purposes because the duration differences between the reduced and canonical utterances not only differed by being reduced and unreduced, but also by position in the sentence, speaker differences, and many other factors. From an experimental point of view, it was therefore best to keep the duration similar between the original and the cross-spliced segments.

GENERAL DISCUSSION

Three eye-tracking experiments were conducted to investigate the nature of phonological competition during the recognition of strongly reduced forms. Competition processes were measured using a printed-word, target-absent variant of the visual world paradigm (Huettig & McQueen, 2007). Participants' looks were tracked to four printed words on a computer screen—a canonical-form competitor (e.g., *benadelen* 'to disadvantage', phonologically similar to the canonical form); a reduced-form competitor (e.g., *meneer* 'mister', phonologically similar to the reduced form); and two phonologically unrelated distractors—while they listened to canonical (e.g., [bənedə]) and strongly reduced forms (e.g., [məneə]) of a spoken target word (e.g., *beneden* 'downwards').

A recent study (Brouwer et al., 2012) demonstrated that listeners penalize acoustic mismatches less strongly when listening to reduced speech than when listening to fully articulated speech. When faced with a listening situation in which phonological reductions frequently occurred, listeners directed their eye gaze to a similar degree to both competitors ("canonical form" or reduced-form competitors) independent of the target's *exact* spoken form. In the present research, we examined whether phonological competition during casual speech, containing many phonological reductions, can *ever* be modulated by the exact phonetic form of the spoken word.

In contrast to the experiments in Brouwer et al. (2012), the printed target word was removed from the visual display in the experimental trials of the experiments in the current study: During filler trials the (printed) target word was present, but in one third of the trials (the experimental trials) the target word was absent from the visual displays. When the target word was absent, participants were required to click on the middle of the screen. Such an experimental set-up has been shown to greatly increase competition effects in the visual world paradigm (Huettig & Altmann, 2005; Huettig et al., 2011, for discussion).

By creating a task situation that maximized the likelihood of observing phonological competition, we found that eye gaze to the different competitors was modulated by the input form. In Experiment 1, when a canonical input form was presented in isolation, participants made few errors, and the canonical-form competitor attracted more overt attention than the reduced-form competitor. When a reduced input form was heard in isolation, participants were incorrect on 34% of the trials. Thus, recognizing isolated reduced forms was, unsurprisingly, harder than recognizing isolated canonical forms. Overall, participants looked more at the reduced-form competitor than at the canonical-form competitor. On trials in which participants made no errors, there was no significant

difference in overt attention between reduced-form and canonical-form competitors, probably because these trials contain forms in which the phonetic distance between the reduced form and the reduced-form competitor is larger. In sum, when listeners hear isolated canonical forms, they look more to canonical-form competitors, but when they hear isolated reduced forms (and made the correct mouse click response), they do not differ in their looks to any of the competitor types.

In Experiment 2, we presented the spoken target words in a context of several surrounding words. As in Experiment 1, when canonical forms were heard, canonical-form competitors attracted more overt attention than reduced-form competitors. The pattern of competition was statistically different in the reduced-form condition, where the reduced-form competitors attracted (numerically but not significantly) more attention. In a later time window, however, as in Experiment 1, there was no significant difference in overt attention between reduced-form and canonical-form competitors.

Therefore, when participants encountered reduced forms in the present experimental set-up, early on during the acoustic duration of the reduced form, phonological competitors with a quite different surface form from the canonical counterpart competed for visual attention. In other words, we observed that canonical forms of other words, which are acoustically similar to the reductions occurring in casual speech, were activated. Such activation may underlie the delays reported for recognition of reduced forms in prior studies using offline techniques (Ernestus et al., 2002; Kemps et al., 2004). The results in our study are additional evidence for such "reduction costs". Listeners made more errors in the reduced-form condition than in the canonical form condition. Performance, however, improved considerably compared to performance in Experiment 1.

"Late" in time, however, the competition pattern changed when participants had heard reduced forms. We observed that during 800 to 1,200 ms after the acoustic onset of the reduced form, the canonical-form competitor attracted numerically

more visual attention than the reduced-form competitors. One way to interpret this pattern of results is that reduced forms can be successfully reconstructed to their canonical forms. We conjecture that this process is time consuming due to the early, momentary activation of unwanted competitors: competitors that in contrast are not strongly activated by a canonical pronunciation of the same word.

The results of Experiment 2 thus suggested that phonological reductions do not carry sufficient fine phonetic detail to block such competition. Experiment 3 examined whether fine phonetic detail nevertheless plays a role in the recognition of reduced forms. To this end, the "surface" segment /m/ from [məneə] was replaced with an "intended" segment /m/ from a canonical form, and these cross-spliced forms were presented to listeners. The early eye movement data showed the same rise of fixations to the reduced-form competitors as that in Experiment 2. However, the late eye movements were influenced by the cross-splicing: The late rise of the canonical-form competitor observed in Experiment 2 was less pronounced in Experiment 3. This seems to indicate that the cross-splicing impeded the reconstruction process and therefore the recognition of the intended word. Consistent with this account, participants also made more errors with the cross-spliced than with the original stimuli. Note that the later result cannot be explained as a splicing artefact, as participants more often (falsely) recognized the cross-spliced reduced form as a different intended word. The results of Experiment 3 then revealed that there are subtle phonetic differences between a given phoneme in a reduced form (e.g., the /m/ in the reduced form of *beneden*) and the same phoneme when produced as part of an intended canonical form (e.g., the /m/ in the canonical form of *met*). Listeners thus appear to be sensitive to these differences when listening to cross-spliced forms.

The results of Experiment 3 also rule out an alternative explanation for the pattern of results in Experiment 2 (i.e., early rise of the reduced-form competitor, and late rise of the canonical-form competitor). It could be argued that this pattern

simply reflects a "form matching" strategy. It is conceivable that listeners just match the strongly reduced form they hear with the first best matching word they see on the screen (i.e., reduced-form competitor), and then they look at the second best matching word on the screen (i.e., canonical-form competitor). In other words, listeners may have strategically cut the lexicon down to the four items on the screen. In Experiment 3, we only slightly manipulated the acoustic input, but this manipulation had a great influence on the results. Looks to the canonical-form competitor did not increase over time, whereas this did happen in the reduced-form condition of Experiment 2. It is difficult to see how a simple "form-matching" strategy would be influenced by such a subtle phonetic manipulation that it could explain the difference in results between Experiments 2 and 3. Additionally, the results of Experiment 2 themselves also provide an argument that invalidates such a strategic account of the results of Experiment 2. If participants were using a simple "form-matching" strategy, we would have found a similar pattern in the canonical form condition, with looks to the reduced-form competitor rising late in time when the canonical-form competitor has been ruled out as a potential target. In the canonical form condition, we observed instead that the canonical-form competitor attracted more overt attention than the reduced-form competitor in both time windows.

How do our results relate to the ongoing debate of how pronunciation variants are recognized? Different views on how listeners recognize reduced forms are postulated in the literature. Two main classes of accounts focus on different mechanisms. One class of accounts proposes that a reconstruction process occurs at a prelexical level, which mediates between the speech signal and the lexicon, on the basis of fine phonetic detail in the signal or phonological context (e.g., Gaskell, 2003), or by top-down restoration (McClelland & Elman, 1986; Warren, 1970). For example, upon hearing the reduced pronunciation [mənеə], listeners may reconstruct the corresponding canonical pronunciation *beneden*. This full form then activates the representation of the

word *beneden* in the lexicon and competes with other /b/-initial words for recognition.

Given the ubiquity of reductions in spontaneous speech, one may wonder if it is useful to even entertain the notion of "canonical forms" in spoken word recognition. After all, the definition may be tied to visual word recognition only, where words tend to be perceived in a canonical form. Experimental evidence with spoken materials has, however, repeatedly shown that recognition is easiest if words are presented in their full (and hence assumed canonical) form, even if this form is not the most frequently heard form (e.g., Pitt, Dilley, & Tat, 2011).

A second class of accounts assumes that phonological variants are stored in the mental lexicon. Two different versions of this account exist. According to the episodic view, the entry for a given word in the mental lexicon consists of detailed and concrete episodic memories of pronunciations of that word that have been encountered previously (e.g., Bybee, 2001; Goldinger, 1998; Johnson, 1997; Pierrehumbert, 2001). More precisely, a "grainy spectrogram" of such variants would be stored in the mental lexicon. Such episodic traces of phonological variants are stored in the mental lexicon next to traces of canonical forms of those words. For example, episodic traces such as [mənеə] for *beneden* are stored in the mental lexicon next to traces of canonical forms of those words. Proponents of the second lexical-storage account argue that different pronunciation variants are stored as abstract phonological forms (e.g., Connine, 2004; Ranbom & Connine, 2007). According to this view, both the phonological variants (e.g., [mənеə]) and the canonical form (e.g., [bənedə]) would be stored, but as abstracted variants of the canonical representation that do not include indexical properties of spoken words such as voice quality, speech rate, pitch, and so on, as would be the case for episodic traces of each variant.

It may well be the case that both mechanisms play roles in the recognition of reduced forms. Gaskell and Marslen-Wilson (1998) already found that phonological reconstruction of pronunciation variants is more efficient for words than for nonwords. Recently, more evidence is accruing that

even for the same pronunciation variant, phonological and lexical processes may operate together. Snoeren, Gaskell, and Di Betta (2009) showed that phonological reconstruction for variants with place assimilation works more efficiently on known words than for nonwords. In a similar vein, Pitt (2009) provided evidence that variants with nasal flaps (*center* pronounced as *cenner*) are recognized by a combination of lexical and phonological processes. He taught participants new words with medial /t/ that could be flapped (e.g., *senty*). Participant did not accept flapped variants (*senny*) as instances of the same word unless they had previously been exposed to the variant form. While this highlights the importance of lexical storage of variant forms, an additional experiment showed that phonological processing plays a role as well. Nasal flapping is much more likely to occur if the /t/ is followed by a reduced vowel (e.g., *center*) than if it is followed by a full vowel (e.g., *content*). Pitt showed that this phonological conditioning matters. Variants with nasal flaps were only accepted if followed by a reduced vowel, so that exposure to a variant form was not sufficient for recognition.

Although the current data cannot distinguish conclusively among these alternative accounts, they appear to be more in line with a reconstruction mechanism. The late looks to the canonical-form competitors in Experiment 2 fit best with the notion that listeners reconstruct canonical forms from reduced forms. In addition, Experiment 3 revealed that this late rise of the canonical-form competitor only appears when fine phonetic detail is preserved in the signal. It is important to note here that our results cannot distinguish whether reduced forms are stored in the mental lexicon or not. Such storage may greatly depend on, for example, how strongly reduced a word form is or how frequent it is. We cannot rule out, for instance, that phonological reduction increases the competitor space (in line with exemplar-storage accounts).[4] Reduced forms may be more likely to activate more competitors than canonical forms because they contain somewhat ambiguous segments. In other

words, listeners could be activating veridical representations of reduced forms, but experiencing more competition when this happens. Further research is required to clarify the contributions of the two mechanisms during the comprehension of strongly reduced forms.

To conclude, the current study addressed the implications of reduction processes for phonological competition in spontaneous conversation. Using an experimental set-up that maximizes the opportunity to measure phonological competition, we observed that strongly reduced forms in casual speech can activate competitors that are similar to the phonological surface form of the reduction. These same competitors are not strongly activated by a canonical pronunciation of the same word. We conjecture that this added competition is one of the causes of the delay during the recognition of strongly reduced forms. Although this delay demonstrates that processing speech reductions is cognitively costly, our results also show that listeners can exploit fine phonetic detail to reconstruct strongly reduced forms to their canonical counterparts. This provides further evidence for the efficiency of the spoken word recognition system.

REFERENCES

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38,* 419–439.

Arai, T. (1999). A case study of spontaneous speech in Japanese. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. Bailey (Eds.), *Proceedings of the International Congress of Phonetic Sciences (ICPhS)* (pp. 615–618). Berkeley, CA: Department of Linguistics, University of California.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects

---

[4] We thank an anonymous reviewer for this suggestion.

for subjects and items. *Journal of Memory and Language: Special Issue on Emerging Data Analysis Techniques, 59,* 390–412.

Barr, D. J. (2008). Analyzing "visual world" eyetracking data using multilevel logistic regression. *Journal of Memory and Language: Special Issue on Emerging Data Analysis Techniques, 59,* 457–474.

Boersma, P. (2001). PRAAT, a system for doing phonetics by computer. *Glot International, 5*(9/10), 341–345.

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.

Brouwer, S., Mitterer, H., & Huettig, F. (in press). Discourse context and the recognition of reduced and canonical spoken words. *Applied Psycholinguistics.*

Bybee, J. (2001). *Phonology and language use.* Cambridge, UK: Cambridge University Press.

Connine, C. M. (2004). It's not what you hear but often how often you hear it: On the neglected role of phonological variant frequency in auditory word recognition. *Psychonomic Bulletin and Review, 11*(6), 1084–1089.

Connine, C. M., Titone, D., Deelman, T., & Blasko, D. G. (1997). Similarity mapping in spoken word recognition. *Journal of Memory and Language, 37,* 463–480.

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, 6,* 84–107.

Dixon, P. (2008). Models of accuracy in repeated-measures design. *Journal of Memory and Language, 59,* 447–456.

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

Gaskell, M. G. (2003). Modelling regressive and progressive effects of assimilation in speech perception. *Journal of Phonetics, 31,* 447–463.

Gaskell, M. G., & Marslen-Wilson, W. D. (1998). Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance, 24*(2), 380–396.

Gaskell, M. G., & Snoeren, N. D. (2008). The impact of strong assimilation on the perception of connected speech. *Journal of Experimental Psychology: Human Perception and Performance, 34* (6), 1632–1647.

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

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

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

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

Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition, 96,* B23–B32.

Huettig, F., & Altmann, G. T. M. (2011). Looking at anything that is green when hearing "frog"—How object surface colour and stored object colour knowledge influence language-mediated overt attention. *Quarterly Journal of Experimental Psychology, 64*(1), 122–145.

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

Huettig, F., Mishra, R. K., & Olivers, C. N. (2012). Mechanisms and representations of language-mediated visual attention. *Frontiers in Psychology, 2,* 394.

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica, 137,* 151–171.

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 145–165). San Diego, CA: Academic Press.

Johnson, K. (2004). Massive reduction in conversational American English. In K. Yoneyama & K. Maekawa (Eds.), *Casual speech: Data and analysis. Proceedings of the 1st session of the 10th International Symposium* (pp. 29–54). Tokyo, Japan: The International Institute of Japan Language.

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

Kuzla, C., Mitterer, H., & Ernestus, M. (2010). Compensation for assimilatory devoicing and prosodic

structure in German fricative perception. In C. Fougeron, B. Kühnert, M. D'Imperio, & N. Vallée (Eds.), *Laboratory Phonology 10* (pp. 731–758). Berlin, Germany: Mouton de Gruyter.

Manuel, S. Y. (1992). *Recovery of "deleted" schwa. Perilius: Papers from the Symposium on Current Phonetic Research Paradigms for Speech Motor Control* (pp. 115–118). Stockholm, Sweden: University of Stockholm.

Marslen-Wilson, W. D., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 576–585.

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

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

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

Oostdijk, N. (2000). The Spoken Dutch Corpus project. *The ELRA Newsletter, 5,* 4–8.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In J. Bybee & P. Hopper (Eds.), *Frequency effects and the emergence of lexical structure* (pp. 137–157). Amsterdam, The Netherlands: John Benjamins.

Pitt, M. A. (2009). How are pronunciation variants of spoken words recognized? A test of generalization to newly learned words. *Journal of Memory and Language, 61,* 19–36.

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

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

Snoeren, N. D., Gaskell, M. G., & Di Betta, A. M. (2009). The perception of assimilation in newly learned novel words. *Journal of Experimental Psychology: Learning, Memory & Cognition, 35,* 542–549.

Snoeren, N. D., Hallé, P. A., & Seguí, J. (2006). A voice for the voiceless: Production and perception of assimilated stops in French. *Journal of Phonetics, 34* (2), 241–268.

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

Warner, N. (2012). Methods for studying spontaneous speech. In A. Cohn, C. Fougeron, & M. Huffman (Eds.), *The Oxford Handbook of Laboratory Phonology* (pp. 621–633). Oxford: Oxford University Press.

Warren, R. M. (1970). Restoration of missing speech sounds. *Science, 167,* 392–393.

Weber, A., Melinger, A., & Lara Tapia, L. (2007). The mapping of phonetic information to lexical presentations in Spanish: Evidence from eye movements. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS 2007)* (pp. 1941–1944). Dudweiler: Pirrot.

APPENDIX

Experimental items: Canonical and reduced realizations with their canonical-form and reduced-form competitors, respectively

Target items	Canonical form	Canonical-form competitor	Reduced form	Reduced-form competitor
afspraak 'appointment'	[ɑfspraːk]	[ɑfspreːkə]	[ɑspʋaː]	[ɑspiraːtsi]
apparaat 'apparatus'	[ɑpaːraːt]	[ɑpəritif]	[ɔpraːt]	[ɔpraːpə]
beneden 'downwards'	[bəneːdə]	[bənaːdeːlə]	[məneːə]	[məneːr]
bijvoorbeeld 'for example'	[bəvoːrbelt]	[bəvoːreχtən]	[vɔlt]	[vɔlt]
computer	[kɔmpjutər]	[kɔmpətɛnt]	[pjutər]	[putsə]
concert	[kɔnsɛrt]	[kɔnjak]	[kɔzɛr]	[kɔstbaːr]
concurrent 'competitor'	[kɔnkyrɛnt]	[kɔnkur]	[kɔ̃krɛnt]	[kɔnkreːt]
constant 'constant'	[kɔnstɑnt]	[kɔnsɛntraːtsi]	[kɔ̃zən]	[kɔzɛin]
cultuur 'culture'	[kʏltyr]	[kʏltʏs]	[kɔmtym]	[kɔmst]
december	[deːsɛmbər]	[deːkaːn]	[eːsɛmər]	[eːtaːʒə]
dinsdag 'Tuesday'	[dɪnsdɑχ]	[dɪŋən]	[dɪzɑ]	[dɪzɑjn]
directeur 'director'	[dɪrɛktør]	[dɪriχeːrən]	[dɪktø]	[dɪktaːtɔr]
kweekschool 'school'	[kʋeːksχɔl]	[kʋeːkən]	[kʋeːsχɔl]	[kʋeːst]
maandag 'Monday'	[maːndɑχ]	[maːnd]	[maːnz]	[maːnzaːt]
ogenblik 'moment'	[oːχəblɪk]	[oːχkɑs]	[blɪk]	[blɪk]
oktober 'October'	[ɔktoːbər]	[ɔktoːpʏs]	[toːʋər]	[toːʋərə]
overheid 'government'	[oːʋərhɛit]	[oʋərhɛmt]	[oːʋərɛi]	[oːʋərɛint]
parlement 'parliament'	[pɑrləmɛnt]	[pɑrkeːrə]	[pɑləmɛn]	[pɑlɛt]
plaatsen 'to place'	[plaːtsə]	[plaːtsnaːm]	[plaːs]	[plaːseːbo]
positie 'position'	[poːzitsi]	[poːzeːrə]	[psitsi]	[psiχə]
prestatie 'performance'	[prəstaːtsi]	[prɛstiʒə]	[pəstaːsi]	[pɛsimɪst]
principe 'principle'	[prɪnsipə]	[prɪns]	[pəsipə]	[pərson]
publiek 'audience'	[pyblik]	[pyblizeːrə]	[ʋlik]	[ʋil]
redelijk 'reasonable'	[reːdələk]	[reːdərɛi]	[reːlək]	[reːlikʋi]
rekenen 'to count'	[reːkənə]	[reːks]	[reːχən]	[reːχən]
rotzooi 'garbage'	[rɔtsoːj]	[rɔts]	[rɔsɪ]	[rɔsɪχ]
standaard 'default'	[stɑndaːrt]	[stɑndplaːts]	[stɑ̃əd]	[stɑŋ]
standpunt 'point of view'	[stɑntpʏnt]	[stɑntpɑstəχ]	[stɑmpʏ]	[stɑmpɔt]
station 'station'	[statʃɔn]	[staːtʏs]	[saʃɔn]	[saːtɛin]
tandarts 'dentist'	[tɑndɑrts]	[tɑndpɑstaː]	[tɑz]	[tɑs]
wedstrijd 'match'	[ʋɛtstrɛit]	[ʋɛtbuk]	[ʋɛs]	[ʋɛsp]
winter	[ʋɪntər]	[ʋɪntstil]	[ʋɪndə]	[ʋɪndə]