The effect of speech situation
on the occurrence of reduced word pronunciation variants

Mirjam Ernestus[ab*], Iris Hanique[ab], Erik Verboom[a]

*Corresponding author: m.ernestus@let.ru.nl, Tel: +31-24-3612970, Fax: +31-24-3521213
[a] Center for Language Studies, Radboud University Nijmegen, P.O. Box 310, 6500 AH Nijmegen, The Netherlands
[b] Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands

Abstract

This article presents two studies investigating how the situation in which speech is uttered affects the frequency with which words are reduced. Study 1 is based on the Spoken Dutch Corpus, which consists of 15 components, nearly all representing a different speech situation. This study shows that the components differ in how often ten semantically weak words are highly reduced. The differences are especially large between the components with scripted and unscripted speech. Within the component group of unscripted speech, the formality of the situation shows an effect. Study 2 investigated segment reduction in a shadowing experiment in which participants repeated Dutch carefully and casually articulated sentences. Prefixal schwa and suffixal /t/ were absent in participants' responses to both sentences types as often as in formal interviews. If a segment was absent, this appeared to be mostly due to extreme co-articulation, unlike in speech produced in less formal situations. Speakers thus adapted more to the formal situation of the experiment than to the stimuli to be shadowed. We conclude that speech situation affects the occurrence of reduced word pronunciation variants, which should be accounted for by psycholinguistic models of speech production and comprehension.

Keywords

Reduced word pronunciation variants; speech situation; corpus-based research; shadowing experiment; segment reduction; psycholinguistic models; Dutch

The effect of speech situation
on the occurrence of reduced word pronunciation variants

1.0 Introduction

In casual conversations, speakers typically do not articulate all words clearly (e.g., Ernestus & Warner, 2011): segments are often weakened or completely absent. The American English word *yesterday* /jɛstɚɾeɪ/, for instance, may sound like *yeshay* /jɛʃeɪ/ and the Dutch word *gewoon* /xəʋon/ 'ordinary' like /xon/. Johnson (2004) reported on the basis of part of the Buckeye corpus (Pitt, Johnson, Hume, Kiesling, & Raymond, 2005) that in American English informal conversations over 60% of the word tokens deviate from their citation forms on at least one segment, and that in approximately 25% of the word tokens at least one segment is absent. Similar numbers have been reported for Dutch (Schuppler, Ernestus, Scharenborg, & Boves, 2011) and French (Adda-Decker, Boula de Mareüil, Adda, & Lamel, 2005). Reduced pronunciation variants appear to be more frequent in informal situations than in formal situations (e.g. Hanique, Ernestus & Schuppler, 2013), but we do not have a clear picture of how the situation in which speech is produced affects reduction degree. We do not even know whether there is indeed an effect of speech situation or whether the main division is between scripted and unscripted speech, as claimed by Shockey (2003: 17). The present article aims at starting to fill this gap.

Nearly all studies investigating speech reduction are based on corpora containing speech collected in one type of situation (e.g., telephone conversations, radio broadcasts, informal face-to-face conversations), which is more often than not different for the different corpora. As a consequence, different reduction phenomena have been investigated for different speech situations (e.g., for American English: Dalby, 1986, studied reduction of schwa in consultative television interview speech; Patterson & Connine, 2001, studied flapping of word-medial /t/ in telephone conversations; Smith, Baker, & Hawkins, 2012, studied the acoustic details of the word-intial syllables *mis* and *dis* on the basis of scripted speech). Since we do not know whether and how speech situation affects reduction, it is unclear how these different findings should be combined and consequently how the combined findings should influence the development of psycholinguistic theories of speech production.

Compare, for instance, the findings reported on the absence of syllables in American English. Johnson (2004) reported that 5.9% of the syllables are absent in the Buckeye corpus of informal face-to-face interviews, while Dalby (1986) reported 9% of absent syllables in his corpus of consultative television interview speech. If we assume that this difference in percentage does not result from difference in transcription method, the question arises whether speakers articulated less clearly in the eighties than fifteen years later or whether interviewees are more likely to articulate less clearly when they are more used to being interviewed.

Similarly, difference in situation may explain differences that have been found between languages. For French, Adda-Decker and colleagues (2005) found that 6% of the syllables were absent in interviews between professional interviewers and politicians or artists. This percentage is comparable to the percentages found for American English reported above. In contrast, Schuppler and colleagues (2011) reported that 19% of words lack at least one syllable in casual face-to-face conversations between friends in Dutch. Do these differences imply that syllables are more likely to be absent in Dutch than in American English and French, or that they are more often absent in interviews between friends than between strangers?

One of the few studies that investigated the effect of speech situation on speech reduction, keeping the transcription method constant, was conducted by Hanique and colleagues (2013). These researchers studied the presence and duration of prefixal schwa and of suffixal /t/ in Dutch past participles in the casual conversations between friends in the Ernestus Corpus of

Spontaneous Dutch (Ernestus, 2000) and in two components of the Spoken Dutch Corpus (Oostdijk, 2000), that is, in the component with interviews broadcasted on TV and in the component with stories read aloud for the blind (Components F and O). They found that schwa was more often absent in the casual conversations (52%) than in the interviews (35%) and that they were the least often absent in the read aloud stories (12%). In contrast, word-final /t/ was absent as often in the spontaneous conversations as in the interviews (29%) and only less often absent in the read speech (12%). These results suggest that small differences in speech situation may substantially affect reduction degree, but also that the size of the effect depends on reduction type.

Another study investigating the effect of speech situation on reduction was conducted by Warner and Tucker (2011). They investigated the realization of the American English flaps and of the stop consonants /b/, /g/, /k/ in telephone conversations between adults who know each other very well, in read aloud stories, and in the reading of word lists. They found clear differences between the three speech situations, among others in the difference between the intensity of the consonant and that of its preceding or following vowel.

A final piece of evidence that speech situation may affect degree of reduction comes from a production experiment conducted by Brouwer, Mitterer and Huettig (2010). These researchers asked participants to shadow sentences extracted from the Spoken Dutch Corpus, half of them containing fully articulated target words (average duration: 490 ms) and half of them containing reduced target words (average duration: 364 ms). The participants produced all target words more clearly than in the stimuli and showed only a small difference between the responses to the fully articulated target words (average duration: 501 ms) and reduced target words (average duration: 480 ms). Most participants (unconsciously) consider production experiments as rather formal speech situations since they have to produce speech out of social context, in the presence of strangers. These results therefore suggest that the speakers adapted their speech above all to the situation rather than to the stimuli.

If speech situation indeed affects reduction, a related question is whether the difference is only quantitative (frequency of occurrence) or also qualitative (type of reduction). Previous studies have claimed that segments may be weakened or even be completely absent due to two completely different mechanisms. On the one hand, segments may be reduced during the articulation process. A segment may, for instance, be inaudible because its articulation is hidden by the articulation of surrounding segments (e.g. Browman & Goldstein, 2009). On the other hand, segments may be weakened or be absent due to higher level processes, which include both phonological reduction processes and the selection of a reduced pronunciation variant of the word from the mental lexicon (e.g. Ranbom & Connine, 2007; Bürki, Ernestus & Frauenfelder, 2010; Bürki, Ernestus, Gendrot, Fougeron, & Frauenfelder, 2011). Possibly, reduction more often results from the articulation process and less often from higher level processes if the speech situation is more formal since in these situations speakers may aim at producing words in full.

The present article addresses the question of how speech situation affects reduction by reporting the results of two very different studies. Study 1 compares the 15 components of the Spoken Dutch Corpus, which nearly all represent different speech situations, ranging from carefully articulated news broadcasted on the radio, to lectures and speeches, to telephone conversations, to spontaneous face-to-face conversations. We studied the pronunciation of ten polysyllabic content words that are known to be often reduced in casual speech (e.g., the word *gewoon* /xəʋon/ 'ordinary', which can be pronounced as /xon/). Most of these content words are semantically weak.

While Study 1 focuses on speech produced in natural situations, Study 2 investigates speech elicited in a production experiment. Following Brouwer, Mitterer and Huettig (2010), we report the results of a shadowing experiment in which Dutch participants shadowed sentences that

were carefully or casually articulated. We focus on the realizations of prefixal schwa and word-final /t/, which also formed the focus of the corpus study by Hanique and colleagues (2013). We did so for two reasons. First, we would like to directly compare the realizations of segments elicited in a shadowing experiment with realizations in corpora of different speech situations. This comparison will show us whether this type of experiment indeed elicits speech that is adapted more to the formal speech situation of the experiment than to the stimuli to be shadowed.

Secondly, we would like to address the question whether there is also a qualitative difference between the speech elicited in a shadowing experiment and the speech produced in more informal situations. Hanique and colleagues (2013) listed the variables that favor the absence and shortening of prefixal schwa and suffixal /t/ in Dutch past participles in three (sub)corpora (informal conversations, interviews and lively read aloud stories). They noticed that the presence and duration of /t/ were affected by roughly the same phonetic variables, which strongly suggests that the absence and shortening of /t/ result from the same mechanisms and that /t/ disappeared during the articulation process. In contrast, the presence of schwa was affected by more and different variables than its duration, which suggests that its absence does not necessarily result from the same process underlying its shortening, but may also result from a higher level categorical deletion process (e.g., a phonological rule of schwa deletion or the selection of a pronunciation variant without schwa from the mental lexicon). In Study 2, we investigated whether the same differences between prefixal schwa and affixal /t/ can be found in speech elicited in a shadowing experiment.

2.0 Study 1: reduction in different components of the Spoken Dutch Corpus

2.1 Materials
Our materials were extracted from the Spoken Dutch Corpus, which consists of 15 components, with speech from different speech situations (see Table 1). We restricted ourselves to the recordings made in the Netherlands, since previous research suggests that speakers from the Netherlands and speakers from the Dutch speaking part of Belgium may show different reduction patterns (Keune, Ernestus, van Hout, & Baayen, 2005). We focused on ten words (listed in Table 2) that occur in all components and that are known to be prone to reduction in casual conversations (e.g., Ernestus, 2000).These ten words are semantically weak and therefore seldom accented, which explains why they are so often reduced.

The appendix lists the frequencies of occurrence of these ten words in the Dutch parts of each component of the Spoken Dutch Corpus as well as their frequencies listed in two databases. The first database is CELEX, based on written texts of over 42 million word tokens (Baayen, Piepenbrock, & Gulikers, 1995). The second database is SUBTLEX-NL, based on 44 million of word tokens in film subtitles (Keuleers, Brysbaert, & New, 2010). Our overview shows that both CELEX and SUBTLEX-NL underestimate the frequencies with which most of these words occur in spoken language, although SUBTLEX-NL to a lesser extent. The components of the Spoken Dutch Corpus also show some differences among each other, which are more pronounced for some words (e.g. *gewoon*) than for others (e.g. *precies*). Note that the frequencies in some components are not very informative since these components are rather small (see Table 1).

Table 1: *Components of the Spoken Dutch Corpus, with the total number of word tokens, the number of tokens of the ten words investigated in Study 1, and the percentages of these latter tokens that are highly reduced.*

Component ID	Description	Total word tokens	Tokens of the ten words	Percentage of highly reduced tokens
A	Spontaneous face-to-face conversations	2 626 172	13 790	36.0
B	Face-to-face interviews with teachers of Dutch	565 433	2 207	40.1
C	Telephone dialogues recorded via a platform	1 208 633	6 285	39.4
D	Telephone dialogues recorded via a minidisc recorder	853.371	5 133	43.1
E	Business transactions	136 461	1 136	34.9
F	Interviews and discussions on TV	790 269	4 829	34.8
G	Discussions, debates, meetings	360 328	906	26.9
H	Lessons recorded in the classroom	405 409	2 106	37.6
I	Spontaneous commentary on TV	208 399	434	17.3
J	Current affairs programs on radio and TV	186 072	648	32.4
K	News bulletins on radio and TV	368 153	245	13.5
L	Reflections and commentaries on radio and TV	145 553	422	27.3
M	Masses, lectures, solemn speeches	18 075	14	14.3
N	Lectures, seminars	140 901	359	27.3
O	Read aloud stories	903 043	950	5.5

Table 2: *The ten words investigated in Study 1: their orthographic forms, their full pronunciations, their translations, the number of tokens investigated and the highly reduced pronunciation variants investigated.*

Orthographic form	Citation form	Translation	Number of tokens	Highly reduced variant(s)
allemaal	/'ɑləmal/	'all'	4 986	/'ɑməl/
anders	/'ɑndərs/	'else'	2 627	/'ɑs/
bijvoorbeeld	/bɛi'vorbelt/	'for instance'	1 144	/'vɔlt/
eigenlijk	/'ɛixələk/	'actually'	5 780	/'ɛik/, /'ɛixk/
gewoon	/xə'ʋon/	'normally'	10 973	/'xon/
helemaal	/'heləmal/	'completely'	5 527	/'heməl/
mogelijk	/'moxələk/	'possible'	639	/'mok/, /'moxk/
natuurlijk	/na'tyrlək/	'of course'	5 690	/'tyk/ /nɑ'tyk/, /nə'tyk/
ongeveer	/ɔŋxə'ver/	'approximately'	498	/ɔ'ver/
precies	/prə'sis/	'precisely'	1 600	/'psis/

Since words are less likely to be reduced in sentence initial or sentence final position or when they are preceded or followed by hesitations (e.g., Bell, Jurafsky, Fosler-Lussier, Girand, & Gildea, 2003), we restricted our study to the 39 464 tokens uttered in sentence medial position in stretches of fluent speech. Each component is well represented, by at least 20 speakers (range: 20 - 604, mean: 176), with one exception: Component M (Masses, lectures, solemn speeches) is represented by only 14 tokens produced, by two speakers.

We automatically created broad phonetic transcriptions for all speech in the corpus, using the transcriber and procedure described in Schuppler et al. (2011). This automatic transcriber is based on the Hidden Markov Toolkit HTK (Young et al., 2002), and received as its input the acoustic signals and the corresponding orthographic transcriptions. The recognizer selected for each word in the orthographic transcriptions the pronunciation variant that best matched the acoustic signal, choosing from a lexicon containing for every word both the full pronunciation and several reduced pronunciation variants. For instance, for the word *gewoon*, the lexicon contained the variants /xəˈʋon/, /xʋon/, and /xon/. The selected pronunciation variants were aligned with the acoustic signal, segment by segment, such that the resulting transcriptions not only show the best matching pronunciation variants but also the durations of all containing segments (which is relevant for Study 2).

The translation between the abstract segments in the lexicon and the acoustic signal was based on acoustic models for these segments. These were the 37 32-Gaussian tristate monophone models produced by Hämäläinen, Gubian, ten Bosch and Boves (2009) on the basis of Component O of the Spoken Dutch corpus (see Table 1). The models were trained at a frame shift of 5 ms and a window length of 25 ms, where for each frame 13 MFCCs (i.e. the mel-scaled cepstral coefficients C0-C12) and their first and second derivatives (39 features) were calculated.

Since the acoustic models consisted of three emitting states and the frame shift was 5 ms, the system assigned to each segment a duration of at least 15 ms. If a segment was in reality shorter than 15 ms, its boundaries were placed within the neighbouring segments (and these neighbouring segments were consequently assigned shorter durations than they really had).

The accuracy of the transcriptions generated by this automatic transcriber and transcription procedure has been estimated in two different tests. Both focused on spontaneous speech, which is notoriously hard to transcribe. Schuppler and colleagues (2011) automatically transcribed part of the spontaneous speech incorporated in the IFA corpus (van Son, Binnenpoorte, van den Heuvel, & Pols, 2001) and found that 14% of the segments were transcribed differently in their transcriptions than in the transcriptions created with a different automatic method and corrected by human transcribers. Since disagreements between human transcribers may vary between 5.6% and 58%, depending on the degree of spontaneity of the speech (Kipp, Wesenick, & Schiel, 1996, 1997; Ernestus & Baayen, 2011), these results strongly suggest that the transcriptions automatically generated with their transcriber and procedure are as reliable as human made transcriptions.

Hanique and colleagues (2013) used the same transcriber and procedure for transcribing the Ernestus Corpus of Spontaneous Dutch, and compared the results for 148 schwas in the prefixes of past participles with the transcriptions created by two human transcribers. They found no statistically significant differences between how often the human transcribers agreed with each other on the presence versus absence of schwa (in 82.4% of cases) and how often these human transcribers agreed with the automatic transcriber (in 75.7% and 77.0% of cases, respectively). The authors then compared the durations assigned by the automatic speech recognizer and by the two human transcribers to the present schwas and found that, after removal of outliers (5%), the average differences were as small as 12.8 ms and 3.7 ms. These differences are also similar to those typically obtained between human transcribers. For instance, Wesenick and Kipp (1996) reported that 96% of the segment boundaries determined by three human transcribers for

64 read sentences of German differed by less than 20 ms. Similarly, Raymond and colleagues (Raymond, Pitt, Johnson, Hume, Makashay, Dautricourt, & Hilts, 2002) observed an average alignment difference of 16.4 ms between human transcribers of the Buckeye corpus.

2.2 Results and discussion

The automatic speech recognizer is likely to confuse pronunciation variants that are very similar. We therefore focused on highly reduced variants, which we define as those variants that maximally consist of the first segment, the stressed syllable(s) and the final segment of the word's full pronunciation (cf. Ernestus, 2000). These highly reduced variants can be easily distinguished from less reduced variants. Table 2 lists for each word the variant(s) that we labelled highly reduced. It shows, for instance, that the only highly reduced word pronunciation variant that we attested for the word *bijvoorbeeld* was /ˈvɔlt/. Out of the 39 464 tokens, 14 225 (36.0%) represent these highly reduced pronunciation variants. The percentage of highly reduced tokens per word ranged from 15.6% (*precies*) to 51.0% (*allemaal*).

Table 1 shows the percentages of highly reduced tokens split by component. There are clear differences among the 15 components: the percentage of highly reduced tokens ranges from 43.1% (Telephone dialogues recorded via a minidisc recorder) to 5.5% (lively read aloud stories). In order to investigate whether these differences are statistically significant, we first grouped the 15 components in four groups: Casual face-to-face conversations (Components A and B), Casual telephone conversations (Components C and D), Spontaneous speech produced in formal situations (Components E, F, G, H, I, J, L) and Read speech (Components K, M, N, O), and we investigated the differences among these groups. Since especially the Component groups of Spontaneous speech produced in formal situations and of Read Speech are heterogeneous, we then also investigated whether there are differences among the components in these groups.

We modelled the probability of a highly reduced token by means of a generalized mixed effect model with as fixed predictor Component Group, and with as crossed random predictors Word and Speaker. Since words are more likely to be reduced when speech rate is higher (e.g., Dalby, 1984; Raymond, Dautricourt, & Hume, 2006), we also included speech rate as a predictor. We investigated whether the effects of the fixed predictors differed for the different speakers or words (i.e., we tested for random slopes).

We defined Speech Rate as the number of syllables per second intended in the stretch of fluent speech containing the target word. We chose this definition rather than the number of syllables per second that were actually produced, because the former measure can be estimated more reliably (as it is based on orthographic rather than phonetic transcriptions). The two measures are highly correlated and research by Koreman (2006) showed that they both contribute to listeners' perception of speech rate. Since speech rate was clearly higher in the components with unscripted than with scripted speech, we orthogonalised the predictors Speech Rate and Component Group by replacing Speech Rate by the residuals of a linear model predicting Speech Rate as a function of Component Group. Below we will refer to these residuals just as Speech Rate. (The statistical results are almost identical if we do not apply this orthogonalization.)

Speech had a significant main effect: if speech rate increased, the likelihood of highly reduced pronunciation variants increased as well ($F(1, 39\,465) = 81.95$, $p < 0.001$). More importantly, there was a main effect of Component Group ($F(3, 39\,461) = 56.89$, $p < 0.001$). Further analyses showed that the differences between all Component Groups are statistically significant ($ps < 0.01$), except the difference between the Component Groups of casual face-to-face conversations and casual telephone conversations ($p > 0.1$). That these two groups do not show a statistically significant difference may be as expected since the two groups hardly differ in the formality of the speech situation. The largest statistically significant difference was found

between the scripted speech (17.3% of the tokens highly reduced) and the non-scripted speech (36.5%, 41.0% and 33.4% highly reduced tokens in the Component Groups with casual face-to-face conversations, with casual telephone conversations and with spontaneous speech produced in formal situations, respectively).

The likelihood of a highly reduced pronunciation variant varied per Word (standard deviation: 0.81) and per Speaker (0.41). Similarly, the effect of Speech rate on this likelihood varied per Word (0.015) and per Speaker (0.010). The effect of Component Group was different for the different Words (0.13 - 0.25).

We then only focused on the Component Group of spontaneous speech uttered in formal situations, and investigated whether there were differences among the components of this group, following the same procedure as for the comparison of the four component groups (except that we now orthogonalised Speech Rate with Component rather than with Component Group, but again the orthogonalisation hardly affected the statistical results). We found again that the likelihood of a highly reduced pronunciation variant varies per Word (0.42) and per Speaker (0.74). Furthermore, the effect of Speech Rate on this likelihood is again different for the different words (0.12). In addition, there was again a main effect of speech rate (F(1, 10 480) = 68.77, p < 0.001).

More importantly, there was a main effect of Component (F(6, 10 475) = 6.16, p < 0.001). Further analyses showed that all components differ from Component I (Spontaneous commentary on radio), which contains the lowest percentage of highly reduced tokens (differences with Components E, F, E, H, J: ps < 0.001; difference with Component G: p < 0.05). Further, Components L (Reflections and commentaries on radio and TV) and G (Discussions, debates, meetings) contain a lower percentage of reduced tokens than components F (Interviews and discussions on TV) and H (Classroom lessons, ps < 0.01).

Finally, we focused on the Component Group of read speech, following the same statistical procedure as for the other analyses. We found again that the likelihood of a reduced variant was different for the different words (0.65) and speakers (1.10) and that there was a main effect of Speech Rate (F(1, 1567) = 30.08, p < 0.001).

There was also a main effect of Component (F(3, 1564) = 21.21, p < 0.001). Further analysis showed that Component O (lively read aloud stories) differs from all other components of read speech (ps < 0.001, with the exception of Component H, most likely because this component is only represented by 14 tokens). The differences between the other components are not statistically significant.

In conclusion, Study 1 clearly demonstrates that semantically weak words are often highly reduced. Moreover, the frequency with which these highly reduced variants occur depends on speech situation. By far the most important distinction is between scripted speech (17.3% of the tokens highly reduced) and unscripted speech (38.4% of the tokens highly reduced), with the read aloud stories showing the lowest percentage of highly reduced variants (5.5%). In addition, we found statistically significant differences between casual face-to-face and telephone conversations on the one hand (37.0% of the tokens highly reduced) and spontaneous speech produced under formal conditions on the other hand (33.5% highly reduced tokens). From the different formal situations, the spontaneous commentary on TV appeared to elicit highly reduced variants the least frequently (17.3%) and interviews and discussions on TV (34.6%) and classroom lessons (37.3%) the most frequently.

3.0 Study 2: A production study

In Study 2, we investigated the type of speech that is elicited by a shadowing experiment with carefully and casually articulated stimuli. We focused on the schwas in the prefixes of past participles and on these words' affixal /t/s, such that we could directly compare our results with the data discussed by Hanique and colleagues (2013), who studied the pronunciation of past participles in casual face-to-face conversations in the Ernestus Corpus of Spontaneous Dutch and in Components F and O of the Spoken Dutch Corpus.

3.1 Participants
We tested 35 Dutch native speakers, six men and 29 women, aged between 18 and 27 (mean: 20 years). They were all undergraduate students at Radboud University Nijmegen.

3.2 Materials
Our experiment contained 180 target past participles starting with the prefix *ge-*, of which 140 end in /t/ (see the Apptendix). Examples are *gedanst* 'danced' and *gebakken* 'baked'. In addition, the experiment included 100 filler past participles starting with *ver-* or *be-*, of which 88 end in /t/, such as *verlicht* 'enlightened'. Finally, we selected 10 past participles for the practice trials, of which seven started with *ge-* and seven ended in /t/. All past participles consisted of two or three syllables and the schwa in the initial syllable was followed by a consonant. They spanned the entire range of frequency of occurrence, which was established on the basis of the Spoken Dutch Corpus.

Each past participle was embedded in the middle of a sentence, which on average consisted of ten words. See (1) for an example. Sentence accent was never on the past participle. Also, past participles were never preceded by a fricative, and those ending in /t/ were never followed by /t/ or /d/. Most past participles were followed by a function word, especially *in* 'in', *op* 'on', *voor* 'before / in front of', and *na* 'after'.

(1) *Ze heeft de hele avond gedanst op tien-centimeter-hoge hakken.*
'She danced the whole night on high heels of ten centimeters.'

While we created only one sentence for each filler, we created two sentences for each target past participle: in one sentence the past participle was followed by a vowel, and in the other by a consonant. In this way, we could study the effect of the type of following segment on the reduction of suffixal /t/. Thus, sentence (1) was complemented with sentence (2).

(2) *Ze had nooit eerder gedanst voor publiek.*
`She had never before danced in front of an audience.'

We asked a female native speaker of Dutch to record the sentences once in a careful way and once as if uttered in a conversation with friends. The carefully articulated sentences had an average duration of 2 208 ms, while the casually articulated sentences had an average duration of 2 019 ms. We automatically generated phonetic transcriptions of the sentences, with the method described in section 2.1. These transcriptions indicate that schwa was absent in the initial syllables of 120 of the 180 (66.67%) past participles in the casually articulated sentences, while it was never absent in the carefully articulated sentences ($\chi^2 = 177.01$, p < 0.001, df = 1). The present schwas were shorter in the casually (on average 28.1 ms) than in the carefully articulated sentences (on average 49.2 ms; t = -10.68, p < 0.001, df = 111.05). With respect to word final /t/, this segment was less often absent in the casually articulated sentences (in 30 of the 140 sentences) than in the carefully articulated sentences (in 57 of the 140 sentences, $\chi^2 = 11.27$, p < 0.001, df = 1). The present /t/s were approximately equally long in both stimuli types (casually articulated sentences: on average 51.6 ms; carefully articulated sentences: on average 47.0 ms; t = 1.86, p > 0.05, df = 183.82). In the following, we will refer to the audio sentences presented to the participants as stimuli.

The experiment started with the ten practice stimuli, which were followed by four blocks, each consisting of 45 target and 25 filler sentences. The first block after the practice stimuli started with seven fillers, while each of the other three blocks started with three fillers.

Within one block, we presented either casually or carefully articulated sentences. The first two blocks contained stimuli of the same speech register (i.e. either carefully or casually articulated stimuli) and the same was true for Blocks 3 and 4. If the first two blocks contained casually articulated sentences, the third and fourth blocks consisted of carefully articulated sentences, and vice versa. Together, the four blocks presented to a participant contained every past participle exactly once (either carefully or casually articulated and either followed by a vowel or consonant).

In total, there were twelve experimental lists, reflecting three pseudo-randomizations of all stimuli, in which no more than three target sentences occurred in succession. The four lists with the stimuli in the same order differed in the condition in which a given past participle was presented: carefully articulated and followed by a vowel, carefully articulated and followed by a consonant, casually articulated and followed by a vowel or carefully articulated and followed by a consonant. Together, the four lists contained all past participles in all four conditions. Every participant heard one list.

3.3 Procedure
Each participant was tested individually in a sound-attenuated booth and listened to the sentences via headphones. We asked participants to repeat these sentences as quickly as possible, and we stressed that this meant that they started repeating the stimuli before they ended.

We recorded responses by means of a head-mounted microphone on an R-09 Edirol recorder. Each trial started with a fixation point shown for 500 ms on a computer screen, and, after an interval of 100 ms, the stimulus sentence was presented. The next trial started 1500 ms after the end of the respective stimulus.

Participants were offered the possibility to take a break after each block. On average, the experiment lasted approximately 30 minutes.

3.4 Data processing
For all target sentences produced by participants, we created automatic broad phonetic transcriptions following the procedure described in section 2.1. Since this procedure has never been tested with speech elicited in a shadowing experiment, we validated the resulting transcriptions by asking two human transcribers to manually transcribe schwa in 100 tokens of different target past participles. These transcribers agreed with each other on the presence versus absence of schwa for 90% of the tokens, while each of them agreed with the automatically generated transcriptions for 85% and 87% of the tokens, respectively. The average differences between the durations assigned to present schwas by the human transcribers and by the automatic speech recognizer were smaller than 13 ms, which can be considered acceptable (see section 2.1).

We excluded 35% of the 6 300 responses to target stimuli from the analyses in two steps. First, 1 866 responses were excluded on the basis of the automatically generated transcriptions. We excluded responses of which the transcriptions indicated a silence directly before or after the target word, since in these trials the target words are not embedded in stretches of connected speech, and as a consequence the target words are not expected to show reduction (e.g., Bell et al., 2003). In addition, we excluded responses of which the transcriptions are likely to be incorrect. We identified these incorrect transcriptions by means of a set of criteria that we

determined by checking 200 automatically generated transcriptions. Responses were excluded of which the transcriptions contained three or more segments of 15 ms in the target word itself or in the directly preceding or following words. These sequences of very short segments typically indicate that the sentence was produced incompletely or with a wrong word order, or that the automatic speech recognizer had selected a non-suitable pronunciation variant. Further, we also excluded responses with transcriptions in which the longest phoneme was the schwa or in which the longest phoneme was shorter than 50 ms, which are also highly likely to be transcription errors. Finally, we excluded those responses of which the transcriptions contain suspiciously long phonemes, as this often indicates that multiple phonemes were transcribed as one long phoneme. We set the boundary for suspiciously long segments at 175 ms for plosives, at 165 ms for vowels and fricatives, and at 155 ms for other consonants. In the second step, we listened to all remaining responses, and excluded from further analysis those 339 responses in which the speaker had not produced the target word or the directly preceding or following word fluently. The fact that we discarded 2 205 out of 6 300 responses shows that the task was difficult.

Interestingly, the number of discarded responses varied among participants from 23 to 104 trials. In addition, it varied among target words, from 2 (*gedanst* 'danced') to 23 (*gerommeld* 'messed around' and *genaderd* 'approached'). The percentage of discarded responses was similar for the carefully (32%) and for the casually (38%) articulated stimuli and similar for the target words followed by consonants (37%) and for those followed by vowels (33%).

3.5 Analyses

We tested the influences of several predictors on the presence and duration of schwa and /t/ (for an overview, see the Appendix). Three predictors were defined on the basis of the experimental design: (1) the register in which the sentence had been produced (carefully or casually articulated); (2) which block the stimulus had occurred in; and (3) the type of segment following the past participle (only relevant for word-final /t/).

In addition, we tested the influences of several other variables that are known from the literature to affect the presence versus absence of segments and their durations. These variables will show whether the absence of schwa and /t/ can be completely due to extreme co-articulation. We tested the same variables as Hanique and colleagues (2013).

One of these variables is Speech Rate (e.g., Dalby, 1984), which we defined as the number of syllables per second in the whole response given by the participant (assuming that all words had been pronounced without reduction, cf. Study 1). We also included a variable (word length) that indicates whether the past participle consisted of two or three syllables, as segments are often shorter if they are followed by more syllables (e.g., Nooteboom, 1972).

We included three measures of word predictability, since words tend to be more reduced if they are more predictable (e.g., Bell, Brenier, Gregory, & Jurafsky, 2009). One measure was the log-transformed word's frequency of occurrence. The second and third measures were the Conditional Probabilities (CP) of the target word (w_{target}) given the preceding word ($w_{preceding}$) or the following word ($w_{following}$), which were calculated with formulae (3) and (4), respectively. The frequencies on which these predictability measures were based were all taken from the entire Spoken Dutch Corpus.

(3) $CP_{preceding} = \log_2((Frequency\ (w_{preceding}, w_{target}) + 1)/(Frequency(w_{preceding}) + 1))$

(4) $CP_{following} = \log_2((Frequency\ (w_{target}, w_{following}) + 1)/(Frequency(w_{following}) + 1))$

Since the properties of the surrounding segments may affect reduction degree (e.g., Dalby, 1984; Shockey 2003; Mitterer & Ernestus, 2006), we included information about the place and

manner of articulation and about the voicing of these segments. For schwa, the preceding segment was always /x/ and we therefore only focused on the effects of the voicing and place and manner of articulation of the following consonant. In addition, we tested explicitly whether the realization of the schwa was different if it was followed by /x/ than by another segment, since Hanique and colleagues (2013) observed that schwa may be longer if the places of articulation of the preceding and following consonants are identical. Their findings are in line with findings by Walter (2007), who argued that the production of two subsequent consonants with the same place of articulation requires more effort than the production of a non-repetition sequence, which leads to longer vowels between consonants with the same place of articulation. For /t/, we examined the place and manner of articulation as well as the voicing of the preceding segment, and whether the following segment was a vowel or a consonant (as already mentioned above).

Finally, we included as predictors for the presence versus absence of schwa and /t/ and for their durations, the log transformed durations of the directly preceding and the directly following segments. If schwa and /t/ are shortened or completely absent due to co-articulation with neighbouring segments, these surrounding segments may be longer. In line with this, Hanique and colleagues (2013) found that schwa tends to be more often absent after longer consonants, and Zimmerer, Scharinger and Reetz (2011) reported that in German /s/ tends to be longer if the following /t/ is absent.

We used linear mixed effects models in order to analyse segment duration and generalized linear mixed effects models to analyse the presence versus absence of segments. The models contained Target word and Participant as crossed random effects. We also investigated whether the main effects differed for the different target words and participants. Each predictor was added individually to a model, and only remained in that model if it showed a statistically significant main effect or interaction and improved the Akaike Information Criterion value (AIC; Akaike, 1973). Duration analyses were based on present segments only.

3.6 Results and discussion

3.6.1. Schwa

Figure 1 shows the results for prefixal schwa. It consists of two kernel density plots (which can be viewed as smoothed versions of histograms) showing the distributions of schwa duration in responses to the carefully articulated sentences (left plot) and to the casually articulated sentences (right plot). In these plots, absent schwas have a duration of zero. Note that the minimum duration that could be assigned to a present segment was 15 ms.
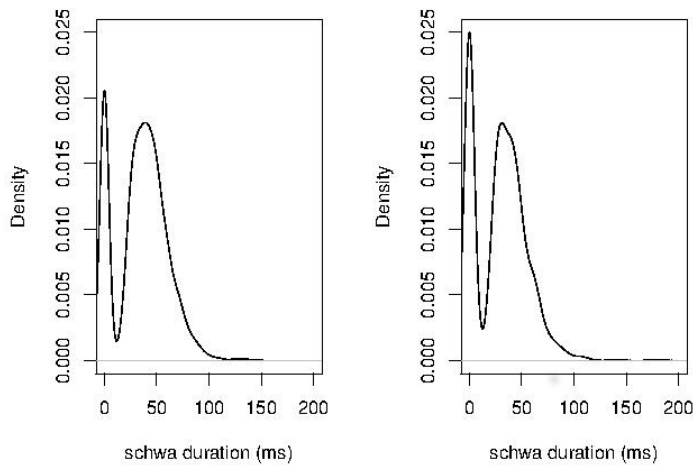
Figure 1: Distributions of schwa duration in the responses to the carefully articulated sentences (left plot) and to the casually articulated sentences (right plot), with absent schwas having a duration of zero.

Out of the 4 095 tokens of target past participles in the responses, 3 019 were produced with schwa in their first syllables. Importantly, the speech style of the stimulus did not have a great influence on the presence of schwa, since schwa was present in 1 661 out of the 2 141 (77.6%) past participle tokens in the responses to carefully articulated sentences and in 1358 out of the 1954 (69.5%) past participle tokens in the responses to casually articulated sentences (compare also the leftmost peaks in the two plots of Figure 1). Schwa was substantially more often present in the responses to the casually articulated stimuli than in these stimuli themselves (35%). Moreover, the percentages of present schwa in the responses are higher than the percentage of present schwas in interviews broadcasted on TV (65% in Component F of the Spoken Dutch Corpus), as reported by Hanique and colleagues (2013).

The responses to the carefully and casually articulated stimuli also do not differ greatly in the duration of present prefixal schwa (compare the rightmost peaks in the two plots of Figure 1): if present, schwa had an average duration of 45.1 ms in the responses to the carefully articulated stimuli and of 42.1 ms in the responses to the casually articulated stimuli. These average durations are similar to the average duration of schwa in the carefully articulated stimuli (49.2 ms). Apparently, participants did not try to mimic the short schwas in the casually articulated stimuli (on average 28.1 ms). Further, these durations are similar to the average duration reported for lively read aloud stories incorporated in the Spoken Dutch corpus (Component O: 43 ms; Hanique et al., 2013).

Tables 3 and 4 show the fixed effects of the two final statistical models for the presence and duration of schwa. For the presence of schwa, the number of degrees of freedom is 13 864, and the estimated standard deviations of the random effects for Target word and Participant are 0.46 and 0.94, respectively. For the duration of schwa, the number of degrees of freedom is 12 678, the estimated standard deviations for the random effects of Target word and Participant are 14.73 and 4.74, respectively, and the estimated random effect of Duration of the preceding consonant for Target word is 2.43.

Table 3: *Logistic regression model predicting the presence of prefixal schwa. The intercept represents a past participle in a response to a carefully articulated stimulus in Block 4 and in which schwa is followed by a voiced consonant*

Fixed effect	estimate	z-value	p-value <
Casually articulated	0.57	26.79	0.001
Word frequency	-0.10	-4.87	0.001
Duration $_{\text{following C}}$	-1.19	-7.83	0.001
Duration $_{\text{preceding C}}$	-0.01	-.059	0.1
Following voiceless consonant	-6.32	-5.67	0.001
Block 1	2.64	2.72	0.05
Block 2	1.17	1.23	0.1
Block 3	3.21	3.22	0.01
Duration $_{\text{following C}}$ * Following voiceless consonant	0.86	4.86	0.001
Duration $_{\text{preceding C}}$ * Block 1	-0.47	-2.95	0.005
Duration $_{\text{preceding C}}$ * Block 2	-0.19	-1.23	0.1
Duration $_{\text{preceding C}}$ * Block 3	-0.48	-295	0.001

Table 4: *Regression model predicting the duration of present schwa. The intercept represents a bisyllabic past participle in a response to a carefully articulated stimulus in which schwa is followed by an approximant or nasal.*

Fixed effect	Estimate	t-value	p-value <
Casually articulated	3.68	7.87	0.001
Word frequency	-0.58	-3.93	0.001
Duration $_{\text{following C}}$	-6.33	-11.33	0.001
Duration $_{\text{preceding C}}$	-5.19	-10.39	0.001
Following fricative	-5.03	-4.23	0.001
Following plosive	4.64	-4.41	0.001
Trisyllabic words	-0.47	-0.51	0.1
Speech rate	-1.47	-3.41	0.001
Speech rate * Trisyllabic words	1.57	2.73	0.005

Several variables predict both the presence of schwa and its duration. First, as already mentioned above, schwa was more often present and longer in a response to a carefully articulated (77.6%; 45 ms) rather than a casually articulated stimulus (69.5%; 42 ms). These differences are much smaller than the differences between the casually and carefully articulated stimuli themselves (casual: 32% and 28.1 ms; careful:100% and 49.2 ms).

Second, replicating earlier findings (e.g., Bell et al, 2009), we found that schwa was more often absent and shorter in past participles with higher frequencies of occurrence. Since highly frequent words are produced more often, their production is more automatized and efficient, which typically results in more overlapping articulatory gestures.

Third, schwa tended to be shorter and to be more often absent if the following segment was longer. Since the duration of the following consonant correlated with its manner of articulation, we had orthogonalised these variables: we had replaced the duration of this segment by the residuals of a model that predicted its duration as a function of its manner of articulation. The effect of the duration of the following consonant on the presence of schwa was greater if this consonant was voiced.

Fourth, schwa also tended to be shorter and to be more often absent if the preceding segment was longer. The effect of the duration of the preceding consonant on the presence of schwa was greater in in the first block of a given speech register than in the second block (i.e. greater in

Blocks 1 and 3 than in Blocks 2 and 4). The effect of the duration of the following consonant on the presence of schwa was greater if this consonant was voiced.

Importantly, the effects of the durations of the neighbouring consonants are likely to result from strong co-articulation. A strongly co-articulated schwa is very similar to the neighbouring consonant and therefore difficult to be distinguished from this consonant, both by humans and by automatic speech recognizers. As a consequence, the duration of schwa may be (partly) assigned to this neighbouring segment. This is especially likely if the consonant is voiced because it is more difficult to tease apart (again for both humans and automatic speech recognizers) a short, strongly co-articulated, voiced vowel from a voiced rather than voiceless consonant.

In the model predicting presence of schwa, the block in which the stimulus was presented and whether the following consonant was voiced or voiceless also showed main effects. Schwa was more often present in the first block of a speech register (Blocks 1 and 3), which suggests that speakers co-articulated less strongly in these blocks (as also indicated by the interaction of Block with the duration of the preceding consonant, mentioned above). Further, schwa was more often absent if followed by a voiceless consonant and thus if surrounded by two voiceless consonants. Between two voiceless consonants, short schwas are likely to be voiceless as well and consequently very difficult to recognize.

The duration of schwa was also affected by the manner of articulation of the following consonant and by speech rate. Schwa was longest if the following consonant was a plosive (47 ms), and shortest if it was a fricative (38 ms; other segments: 44 ms). We hypothesize that strongly co-articulated schwa can be less easily distinguished from a following fricative than a following plosive, since fricatives are continuants.

Similarly, the effect of speech rate was as expected: a higher speech rate correlated with shorter schwas. Because speech rate was slightly higher in the responses to the casually articulated rather than the carefully articulated sentences, we had orthogonalised these two variables by replacing Speech Rate by the residuals of a model that predicted Speech Rate as a function of whether the stimulus was carefully or casually articulated. Interestingly, the effect of speech rate was restricted to bi-syllabic past participles. A possible explanation for this restriction is that vowels tend to be longer in the initial syllables of bi-syllabic than of tri-syllabic words (e.g. Nooteboom, 1972), and consequently in bisyllabic words, they can show more variation in their durations.

The presence of schwa and its duration were thus mostly affected by the same variables, which suggests that the absence of schwa resulted from the same mechanisms underlying its shortening, and therefore that schwa disappeared during the articulation process. This hypothesis is supported by the specific effects that we found on the presence of schwa, which all indicate that schwa was mainly absent due to strong co-articulation. This finding contrasts with the findings by Hanique and colleagues (2013) for speech produced in more informal situations. They found that the shortening and absence of schwa in past participles are partly conditioned by different variables, which suggests that these two phenomena do not result from the same mechanism. In more informal situations, the absence of schwa can thus probably also be due to higher level categorical processes (e.g., a phonological rule or the retrieval of a pronunciation variant without schwa from the mental lexicon).

In conclusion, the presence and duration of schwa show that participants did not closely mimic the casually articulated stimuli. Schwa was less often present in the stimuli than in the participants' responses to these stimuli and the differences between the responses to the carefully and casually articulated stimuli were small. Further, we found that if schwa was absent, it most likely disappeared during the articulation process, which implies that the type of

speech elicited by this shadowing experiment is dissimilar in this respect to speech produced in more informal situations.

3.6.2. /t/

The /t/ dataset contains 3 133 target past participle tokens, which is a lower number than the number of past participle tokens analysed for the presence and duration of schwa (4 095), since not all target past participles end in /t/ (see the Appendix). Figure 2 shows the duration distributions of /t/ in the responses to the carefully articulated stimuli (left plot) and in the responses to the casually articulated stimuli (right plot). For these plots, absent /t/s have been assigned a duration of zero.
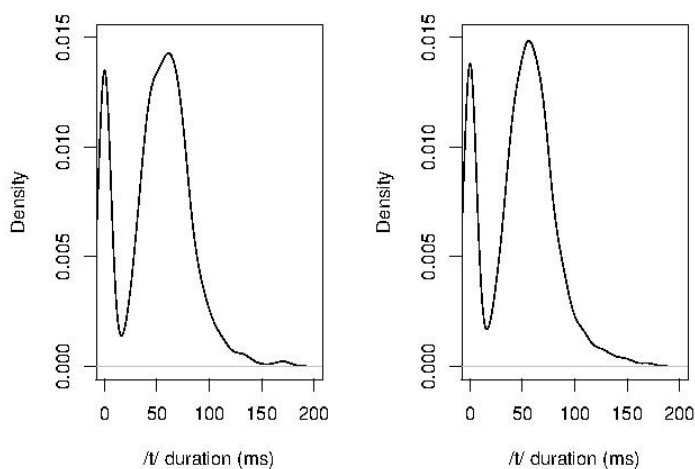

Figure 2: Distributions of /t/ duration in the responses to the carefully articulated sentences (left plot) and to the casually articulated sentences (right plot), with absent /t/s having a duration of zero.

Out of the 3 133 /t/s considered, 2 472 were present: /t/ was absent in 20.7% of the responses to carefully articulated stimuli and in 21.5% of the responses to casually articulated stimuli. These percentages are very similar to each other (compare also the leftmost peaks of the plots in Figure 2). Moreover, these percentages are similar to the percentages of absent /t/ in the stimuli (40.7% in the casually articulated stimuli and 21.4% in the carefully articulated stimuli).

If /t/ was present, it had an average duration of 62.1 ms in the responses to the carefully articulated stimuli and of 61.5 ms in the responses to the casually articulated stimuli. These average durations are very similar to each other (as also appears from the rightmost peaks in the two plots of Figure 2). Importantly, the /t/s in the responses are on average longer than the /t/s in the stimuli (/t/ had an average duration of 47.0 ms in the carefully articulated stimuli and of 51.6 ms in the casually articulated stimuli).

The question rises in which speech situation, as incorporated in the Spoken Dutch Corpus, similar figures can be found. Hanique and colleagues (2013) reported that final /t/ is absent in 11.9% of the past participles in the lively read aloud stories (Component O) and in 29.1% of the past participles in the interviews broadcasted on TV (Component F). In these two components, present /t/ has an average duration of 80 ms and 58 ms, respectively. The past participles elicited

by this shadowing experiment appear therefore somewhat less reduced than past participles produced in formal interviews incorporated in the Spoken Dutch Corpus.

Table 5 and 6 show the fixed effects of the two final statistical models for the presence and duration of /t/. For the presence of /t/, the number of degrees of freedom is 13 123, and the estimated standard deviations of the random effects for Target word and Participant are 1.92 and 0.84, respectively. The estimated random effect for Duration of the preceding consonant for Target word is 0.04, while the estimated random effects for Duration of the following consonant are 0.79 for Target word and 0.39 for Participant. For the duration of /t/, the number of degrees of freedom is 12 389, the estimated standard deviations of the random effects for Target word and Participant are 5.83 and 6.87, respectively. The estimated random effects of Duration of the following consonant are 5.58 for Target word and 2.07 for Participant.

Table 5: *Logistic regression model predicting the presence of /t/. The intercept represents a /t/ followed by a consonant*

Fixed effect	estimate	z-value	p-value <
Duration $_{\text{following segment}}$	-7.50	-8.22	0.001
Duration $_{\text{preceding segment}}$	-0.39	-2.06	0.01
CP $_{\text{following word}}$	-0.04	-0.82	1.0
Following vowel	0.87	0.63	1.0
Duration $_{\text{following segment}}$ * Following vowel	11.80	5.27	0.001
Duration $_{\text{following segment}}$ * Duration $_{\text{preceding segment}}$	0.45	4.13	0.001
Duration $_{\text{following segment}}$ * CP $_{\text{following word}}$	-0.29	-6.49	0.001
Following vowel * CP $_{\text{following word}}$	-0.06	-0.59	1.0
Duration $_{\text{following segment}}$ * Following vowel * CP $_{\text{following word}}$	0.75	4.98	0.001

Table 6: *Regression model predicting the duration of present /t/. The intercept represents a /t/ preceded by an approximant and followed by a consonant*

Fixed effect	Estimate	t-value	p-value <
Duration $_{\text{following segment}}$	-7.99	-9.76	0.001
Duration $_{\text{preceding segment}}$	-5.37	-4.39	0.001
CP $_{\text{following word}}$	-1.50	-6.11	0.001
Following vowel	6.03	-7.75	0.001
Preceding liquid	7.54	4.61	0.001
Preceding nasal	-1.19	-0.61	1.0
Preceding plosive	4.15	2.06	0.05
Preceding vowel	10.80	5.38	0.001
Speech rate	-2.38	-5.23	0.001
Duration $_{\text{following segment}}$ * Following vowel	8.99	8.65	0.001
Duration $_{\text{following segment}}$ * Duration $_{\text{preceding segment}}$	1.61	2.52	0.05
Duration $_{\text{preceding segment}}$ * Preceding liquid	4.28	2.85	0.005
Duration $_{\text{preceding segment}}$ * Preceding nasal	1.91	1.04	1.0
Duration $_{\text{preceding segment}}$ * Preceding plosive	-1.16	-0.62	1.0
Duration $_{\text{preceding segment}}$ * Preceding vowel	6.27	3.29	0.001

Tables 5 and 6 show that the presence and duration of /t/ are largely conditioned by the same variables. First, both the presence of /t/ and its duration correlated with the duration of the following segment. Since the duration of this following segment was correlated with the type of this following segment, we had replaced this duration by the residuals of a linear model predicting this duration as a function of segment type. Word-final /t/ was more likely to be absent and shorter if the following segment was longer. This effect indicates that reduction of /t/ resulted from (strong) co-articulation: /t/ could not be well distinguished from the following segment and (part of) its duration was consequently assigned to the following consonant.

As expected, the effect of the duration of the following segment was restricted to following consonants. Strong co-articulation of /t/ with a following segment is more likely to result in a realization that is difficult to be distinguished from this following segment if this segment is a consonant rather than a vowel, as vowels are produced with an open vocal tract, unlike consonants, including /t/.

Second, both dependent variables showed effects of the duration of the preceding segment. Because this duration was correlated with this segment's manner of articulation, we had orthogonalised them by replacing this duration by the residuals of a linear model predicting this duration as a function of the segment's manner. Segment /t/ tended to be shorter and to be more often absent if the preceding segment was longer. Like the effect of the longer following segment, discussed above, this effect indicates that reduction of /t/ resulted from (strong) co-articulation.

Our finding that /t/ is more often absent after longer consonants is in line with observations for German by Zimmerer and colleagues (2011), who asked participants to produce inflectional forms of infinitives presented on a computer screen and observed that if participants did not (clearly) produce word-final /t/, preceding /s/ tended to be longer. In contrast, our findings differ from those obtained by Mitterer and Ernestus (2006) for Dutch, who analyzed word tokens produced in read aloud stories (Component O of the Spoken Dutch Corpus) and did not find that /s/ was longer if following /t/ was absent. Possibly, Mitterer and Ernestus did not observe a statistically significant difference in /s/ duration between words produced with and without /t/ because of lack of statistical power (they only analyzed 100 words, of which 18 were produced without /t/). In accordance with their null result, Mitterer and Ernestus found in a perception experiment that Dutch listeners interpret word tokens with elongated final /s/ as ending in /s/ rather than in /st/. Listeners may not have shown sensitivity to /s/ duration in this experiment because they adapted to the speech style they were presented with and in which segments are seldom absent.

Importantly, the duration of the preceding segment interacted with the duration of the following segment: /t/ was more likely to be absent and shorter if either the preceding or following segment was longer and the other immediately neighbouring segment was shorter. When both neighbouring segments were longer, their long durations are likely to be due to a lower speech rate. When only one of them is longer, this is less likely the case. Therefore this interaction supports our hypothesis that reduced /t/s co-occur with longer durations of preceding or following segments because strong co-articulation makes /t/ difficult to be distinguished from these neighbouring segments.

Third, both the presence and the duration of /t/ showed effects of the conditional probability of the following word. If this probability was higher, /t/ tended to be shorter and more often absent. For the presence of /t/, this effect was restricted to /t/s followed by long consonants. Speakers probably need less time to plan following words if these words are more likely. As a consequence, they do not have to lower their speech rate. Moreover, they have more experience with more frequent word sequences and may have automatized their articulation. This may result in stronger co-articulation and thus reduction in high frequency word combinations (e.g., Pluymaekers Ernestus, & Baayen, 2005; Bell et al., 2009). As mentioned above, strong co-articulation with a consonant is more likely to result in the apparent absence of /t/ than strong co-articulation with a vowel. This three way interaction shows again that the absence of /t/ was mostly due to strong co-articulation.

The duration of /t/ also showed main effects of speech rate and the type of the preceding consonant. As expected, /t/ was shorter at higher speech rates. Further, /t/ tended to be shorter if the preceding consonant was a nasal or a fricative, especially if this fricative or nasal was longer. Nasals and fricatives are continuants produced with only a small opening in the vocal

tract, and they are therefore very similar to reduced /t/s produced with incomplete closures. Consequently, reduced /t/s are difficult to be distinguished from them and their duration may be assigned to these preceding consonants.

In conclusion, /t/ was as often absent in the participants' responses as in the stimuli themselves. If /t/ was present, however, it tended to be longer in the participants' responses than in the stimuli. As a consequence, /t/ was realized in the responses as little reduced as in the formal interviews incorporated in the Spoken Dutch Corpus. Further, we found that the absence of /t/ was largely conditioned by the same variables as those that favour /t/ shortening. This, in combination with the nature of the effects, strongly suggests that absence of /t/ was mostly due to extreme co-articulation. Similar results have been found by Hanique and colleagues (2013) for past participles produced in more informal situations.

4.0 General Discussion

Previous research suggests that the situation under which words are uttered affects the likelihood that they are reduced (e.g. Warner & Tucker, 2011; Hanique et al. 2013). These studies mostly focus on differences between casual conversations, formal interviews, and read speech. The present article extends this research by comparing more types of speech situations.

In Study 1, we investigated how often ten semantically weak content words are highly reduced in the Dutch parts of the 15 components of the Spoken Dutch Corpus, which nearly all represent different speech situations. Table 1 shows that there are clear differences among the components. The percentage of highly reduced variants ranges from 43% in Component C (Telephone dialogues recorded via a minidisc recorder) to 5% in Component O (read aloud stories).

Statistical analyses showed that, in line with the formality of the situation, the components with casual conversations (either face to face or by telephone) contain higher percentages of highly reduced pronunciation variants (on average 38%) than the components with speech produced in formal situations (on average 33%). The components with read speech contain even substantially fewer highly reduced variants (on average 12%). Possibly, the difference between scripted and unscripted speech plays here a role as well (as hypothesized by Shockey, 2003: 17).

We also examined potential differences among the components with read speech and among the components with speech produced in formal situations. The component containing read speech appeared to be very homogeneous with only Component O (read aloud stories) differing from the other components. Possibly, this component contained fewer highly reduced pronunciation variants because its speakers, in contrast to the speakers in the other components, were obliged to strictly adhere to their text. Since the number of tokens in these components are rather low (ranging from 14 to 950), more statistically significant differences may be found if these components are extended.

Analysis of the components representing speech produced in formal situations showed that there are also differences among these components. Component I (Spontaneous commentary on radio) differs from nearly all other components and contains the lowest percentage of highly reduced tokens of all (17%). Components G (Discussions, debates, meetings) and L (Reflections and commentaries on radio and TV) contain higher percentages of reduced tokens (26% and 27%, respectively), while Components F (Interviews and discussions on TV) and H (Classroom lessons, ps < 0.01) contain the highest percentages (35% and 37%). This shows that reduction degree is highly sensitive to the exact characteristics of the speech situation, and more sensitive than may be concluded on the basis of Tucker and Warner (2011) and Hanique, Ernestus and Schuppler (2013), who only compared very different types of speech situations.

In Study 2, we investigated the effect of speech situation on segment reduction by examining speech elicited in a shadowing experiment. In this experiment, participants repeated Dutch carefully and casually articulated sentences with past participles. These sentences clearly differed in the presence and duration of prefixal schwa in the target past participles (carefully articulated speech: schwa was present in all sentences and had an average duration of 49 ms; casually articulated speech: schwa was present in 31% of the sentences, and if present it had an average duration of 28 ms). Based on the data collected by Brouwer and colleagues (2010), we hypothesized that the speech elicited in this experiment would mainly show the signature of the speech situation rather than showing the effect of the precise acoustic characteristics of the stimuli that had to be shadowed.

With respect to prefixal schwa, we observed that the participants' responses were approximately the same to the carefully and to the casually articulated sentences: prefixal schwa was present in 78% of the responses to the carefully articulated sentences and in 70% of the responses to the casually articulated sentences. Furthermore, the present schwas had an average duration of 45 ms in the responses to the carefully articulated stimuli and of 42 ms in the responses to the casually articulated stimuli. The participants were thus not greatly influenced by the exact acoustic properties of the stimuli, at least not with respect to the properties of schwa. Importantly, the situation was rather formal since the participants were sitting in a sound-attenuated booth, were wearing head-mounted microphones, and produced speech out of social context, in the presence of a stranger. The data therefore strongly suggest that the participants adapted their speech more to the speech situation than to the stimuli.

The prefixal schwas in the participants' productions were most comparable to the prefixal schwas in the formal interviews incorporated in Component F of the Spoken Dutch corpus. In this component, schwa is present in 65% of the prefixes of past participles and, if present, it has an average duration of 49 ms (Hanique et al., 2013). This finding supports our hypothesis that the participants in our experiment adapted their speech above all to the rather formal situation.

We also investigated how participants produced the suffixal /t/s of the past participles. The patterns in the presence and duration of this /t/ support our conclusions based on the data for schwa. The evidence from /t/ is, however, less strong, because the stimuli in the experiment already approximated the patterns present in formal interviews for /t/ and our finding that the responses show this pattern as well is therefore less informative.

Together, Study 1 and Study 2 clearly show that degree of reduction is influenced by the formality of the speech situation. This sensitivity to speech situation has theoretical consequences. First, it has consequences for psycholinguistic models of speech production. Production models have to account for how speech situation exactly affects speech production. Our data suggest that speech situation affects segment reduction even if the effect of speech rate is taken into account. This adds to earlier observations that a higher speech rate does not necessarily imply a higher reduction degree (e.g., van Son & Pols, 1990; Shockey 2003: 11 - 13). The sensitivity of speech reduction to speech situation can therefore not only be modelled by the effect of situation on speech rate, and the question rises how exactly speech situation affects the production process.

Second, our data may have implications for psycholinguistic models of speech comprehension. Since the probability of a highly reduced variant is affected by speech situation, the recognition system would be more efficient if it took this effect into account. Future research has to investigate whether this is the case. Preliminary evidence that it is, comes from Brouwer, Mitterer and Huettig (2012), who showed that listeners only take into account that words may be reduced if they hear both full and reduced variants of words. The question is how sensitive the recognition system exactly is and whether it is sensitive, for instance, to the differences between different types of formal spontaneous speech.

We also analysed the data from the shadowing experiment to investigate whether there is a qualitative difference between speech elicited in this type of experiment and more informal speech. Our analyses of which variables favoured the absence and duration of schwa and /t/ strongly suggest that both their absence and shortening arose during articulation. First, this is suggested by the nature of the relevant variables. For instance, the two segments tended to be more often absent and to be shorter if one of the neighbouring segments was longer. As a consequence of strong co-articulation of schwa and /t/ with neighbouring segments, schwa and /t/ are likely to become less distinguishable from these neighbouring segments and (part of) the durations of schwa and /t/ may be assigned to these neighbouring segments. Second, the analyses show that the variables that favoured the shortening of schwa and /t/ also favoured their absence. This strongly suggests that the absence and shortening of these two segments resulted from the same underlying mechanisms and therefore that also the absence of these segments arose during the articulation process.

Importantly, these results differ from those obtained for a combination of casual speech, formal interviews, and lively read aloud speech investigated by Hanique and colleagues (2013). These researchers reported that in their dataset the absence of schwa was conditioned by more and different variables than the shortening of schwa. Moreover, schwa was more often absent than could be expected if the absence and shortening of schwa resulted from the same underlying process. Based on these findings, Hanique and colleagues concluded that in their dataset the absence of schwa could result from a higher level categorical process (i.e. phonological deletion of schwa or the selection of a pronunciation variant without schwa from the mental lexicon). This difference between our data and the data analysed by Hanique and colleagues forms another piece of evidence that the speech elicited in our shadowing task cannot be compared to speech produced in more informal situations: whereas in informal situations schwa may be absent due to a categorical process, the absence of schwa in the speech from our experiment (nearly) always resulted from strong co-articulation. The speech elicited in our shadowing experiment is qualitatively different from casual speech.

It is highly likely that also the highly reduced word pronunciation variants that we examined in Study 1 result from higher level reduction processes. For instance, it is difficult to imagine how extreme coarticulation of the segments in /bɛi'vorbelt/ would lead to the pronunciation variant /'vɔlt/. There may thus be a qualitative difference between the absence of single segments and the occurrence of highly reduced pronunciation variants, and therefore between the phenomena examined in Studies 1 and 2.

This qualitative difference may explain why the lowest percentage of reduction (in Component O of the Spoken Dutch Corpus) is lower for the highly reduced pronunciation variants (5.5%) than for the absence of schwa and /t/ (both 12%, as reported by Hanique and colleagues). Reduction resulting from higher level mechanisms may play a smaller role in more formal speech styles. This is also evidenced by schwa reduction, which appears to result more from co-articulation in more formal speech situations (as appeared from the comparison of the results from our Study 2 with the results from Hanique and colleagues, 2013). A possible reason for this is that speakers may aim at producing words in full in formal speech situations and therefore tend to select the full word pronunciation variants from their mental lexicon. Future research will have to investigate this hypothesis further.

Possibly, as a consequence, reduction phenomena resulting from higher level mechanisms and reduction phenomena resulting from co-articulation may show differences in how much the formality of the situation affects the occurrence of the reduction phenomenon. This hypothesis is supported by the findings by Hanique and colleagues, who showed that the absence of schwa, which may result both from co-articulation and from higher order processes, is more influenced by speech situation than the absence of /t/. Future research has to also investigate this issue in greater detail.

Our results also have implications for how we should investigate speech reduction. Our finding that speech reduction is influenced by the formality of the situation implies that we have to be careful when comparing the results from studies based on different speech corpora. Before we conclude that differences are due to, for instance, differences between the phenomena under investigation (e.g., flap reduction versus schwa reduction), between the languages in the corpora, or between the time periods in which the corpora were collected, we first have to consider the possibility that the attested differences may be due to differences between the corpora in the speech situation they represent. This seriously complicates research into speech reduction. It pleas for the use of one single corpus for the investigation of different phenomena in one language, and for the use of corpora collected with the same procedure for the comparison of languages (e.g. the Nijmegen corpora of casual French, Spanish and Czech, which were all collected with the procedure described in Torreira, Adda-Decker & Ernestus, 2010). At the same time, however, we should investigate reduction phenomena with as many different corpora as possible, because they all have their strengths and weaknesses.

Preferably, the study of reduced speech is not only based on corpora. Corpora have the disadvantage that they typically do not contain enough tokens of segments or words produced in comparable segmental and prosodic contexts, which makes it difficult to investigate which variables affect their acoustic characteristics. This raises the question which experimental paradigm may produce speech that is qualitatively close to spontaneous casual speech while providing researchers the possibility to determine exactly which words the participant utters and under which conditions. Our finding that participants shadowing casually articulated sentences do not produce casual speech are in line with findings by Brouwer and colleagues (2010) and show that a simple shadowing task cannot make participants casually articulate pre-defined sentences. We therefore have to search for another experimental paradigm. Until we have found such a paradigm, tasks precisely describing the topic of the conversation, including map tasks, picture description tasks and story retelling tasks, form the best alternatives (see Warner, 2012).

In conclusion, our two studies have shown that situation affects the frequencies of reduction phenomena. Moreover, speech situation may determine whether segments are mainly absent due to extreme co-articulation or also due to higher level categorical processes. These results have important consequences for psycholinguistic models of speech production and perception and for how we collect data informing these models.

Reference List
Adda-Decker, M., Boula de Mareüil, P., Adda, G., & Lamel L. (2005). Investigating syllabic structures and their variation in spontaneous French. *Speech Communication, 46*, 119-139.
Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Proceedings of isit-1973* (pp. 267-281). Budapest.
Baayen, R.H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database (CD-ROM)*. University of Pennsylvania, Philadelphia, PA: Linguistic Data Consortium.
Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., & Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America, 113*, 1001-1024.
Bell, A., Brenier, J., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language, 60*, 92-111.

Brouwer, S., Mitterer, H., & Huettig, F. (2010). Shadowing reduced speech and alignment. *Journal of the Acoustical Society of America, 128*, E32 - E37.

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.

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

Bürki, A., Ernestus, M., & Frauenfelder, U.H. (2010). Is there only one "fenêtre" in the production lexicon? On-line evidence on the nature of phonological representations of pronunciation variants for French schwa words. *Journal of Memory and Language, 62*, 421-437.

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

Dalby, J. M. (1986). *Phonetic structure of fast speech in American English (Vol. 7).* Reproduced by the Indiana University Linguistics Club.

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

Ernestus, M. & Baayen R.H., (2011). Corpora and exemplars in phonology. In: J. Goldsmith, J. Riggle, & A. Yu (Eds.), *The Handbook of Phonological Theory (2nd ed.)* (pp. 374-400). Chichester, West Sussex: Wiley-Blackwell.

Hämäläinen, A., Gubian, M., ten Bosch, L., & Boves, L. (2009). Analysis of acoustic reduction using spectral similarity measures. *Journal of the Acoustical Society of America, 126*, 2337-3235.

Hanique, I., Ernestus, M., & Schuppler, B., (2013). Informal speech processes can be categorical in nature, even if they affect many different words. *Journal of the Acoustical Society of America, 133*, 1644-1655.

Johnson, K. (2004). Massive reduction in conversational American English. In K. Yoneyama & K. Maekawa (Eds.), *Spontaneous speech: Data and analysis. Proceedings of the 1st session of the 10th international symposium* (pp. 29-54). Tokyo: The National International Institute for Japanese Language.

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, 42, 643-650.

K. Keune, M. Ernestus, R. van Hout, & R.H. Baayen (2005). Social, geographical, and register variation in Dutch: From written "mogelijk" to spoken "mok". *Corpus Linguistics and Linguistic Theory* 1, 183-223.

Kipp, A., Wesenick, M., & Schiel, F. (1996). Automatic detection and segmentation of pronunciation variants in German speech corpora. In *Proceedings of ICSLP* (pp. 106-109). Philadelphia, USA.

Kipp, A., Wesenick, M., & Schiel, F. (1997). Pronunciation modeling applied to automatic segmentation of spontaneous speech. In *Proceedings of Eurospeech* (pp. 1023-1026). Rhodes, Greece.

Koreman, J. (2006). Perceived speech rate: The effects of articulation rate and speaking style in spontaneous speech. *Journal of the Acoustical Society of America*, 119, 582-596.

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

Nooteboom, S. (1972). *Production and perception of vowel duration: a study of durational properties of vowels in Dutch.* University of Utrecht: PhD thesis.

Oostdijk, N. (2002). The design of the Spoken Dutch Corpus. In P. Peters, P. Collins, & A. Smith, *New Frontiers of Corpus Research* (pp. 105–112). Amsterdam: Rodopi.

Patterson, D. & Connine, C.M. (2001). Variant frequency in flap production. *Phonetica, 58*, 254-275.

Pitt, M. A., Johnson, K., Hume, E., Kiesling, S., & Raymond, W. (2005). The Buckeye corpus of conversational speech: labeling conventions and a test of transcriber reliability. *Speech Communication, 45*, 89-95.

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

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

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

Raymond, W.D., Pitt, M., Johnson, K., Hume, E., Makashay, M., Dautricourt, R., & Hilts, C. (2002). An analysis of transcription consistency in spontaneous speech from the Buckeye corpus. In *Proceedings of ICSLP* (pp. 1125-1128). Denver, Colorado, USA.

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

Shockey, L. (2003). Sound patterns of spoken English. Blackwell Publishing.

Smith, R., Baker, R., & Hawkins, S. (2012). Phonetic detail that distinguishes prefixed from pseudo-prefixed words. *Journal of Phonetics, 40*, 689-705.

van Son, R.J.J.H. & Pols, L.C.W. (1990). Formant frequencies of Dutch vowels in a text, read at normal and fast rate. *Journal of the Acoustical Society of America, 88*, 1683-1693.

van Son, R.J.J.H., Binnenpoorte, D., van den Heuvel, H., & Pols. L. (2001). The IFA corpus: A phonemically segmented Dutch 'open source' speech database. In *Proceedings of Eurospeech* (pp. 2051-2054). Aalborg, Denmark.

Torreira, F., Adda-Decker, M., & Ernestus, M. (2010). The Nijmegen corpus of casual French. *Speech Communication, 52*, 201-212

Walter, M. (2007). *Repetition avoidance in human language*. MIT: PhD thesis. On-line available at: http://faculty.wcas.northwestern.edu/~maw962/docs/walter-dissertation.pdf.

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.

Warner, N. & Tucker, B.V. (2011). Phonetic variability of stops and flaps in spontaneous and careful speech. *Journal of the Acoustical Society of America, 130*, 1606-1617.

Wesenick, M. & Kipp, A. (1996). Estimating the quality of phonetic transcriptions and segmentations of speech signals. In *Proceedings of ICSLP* (pp. 129-132). Philadelpia, USA.

Young, S., Evermann, G., Hain, T., Kerschaw, D., Moore, G., Odell, J,. Ollason, D., Povey, D., Valtchev, V., & Woodland P. (2002). *The HTK book 3.2*. Cambridge: Entropic.

Zimmerer, F., Scharinger, M. & Reetz, H. (2011). When BEAT becomes HOUSE: Factors of word final /t/-deletion in German. *Speech Communication*, 53, 941-954.

Appendix

Table A.1: *Natural logarithms of the frequencies (number of tokens per million word tokens) of five words investigated in Study 1, in the Dutch part of each component of the Spoken Dutch Corpus, in SUBTLEX-NL and in CELEX.*

Corpus	allemaal	anders	bijvoorbeeld	eigenlijk	gewoon
CGN Component A	7.50	6.86	6.01	7.39	8.41
CGN Component B	7.25	6.90	7.06	8.14	7.90
CGN Component C	7.51	6.89	5.45	7.45	8.19
CGN Component D	7.61	6.80	5.73	7.66	8.66
CGN Component E	6.43	7.20	5.65	6.78	8.36
CGN Component F	7.09	6.59	6.92	7.80	7.63
CGN Component G	6.27	6.64	6.53	7.21	6.64
CGN Component H	7.26	6.93	6.97	7.58	7.87
CGN Component I	6.28	5.82	4.81	7.20	7.00
CGN Component J	7.33	5.60	6.36	7.84	7.68
CGN Component K	4.66	5.26	4.99	4.47	5.31
CGN Component L	7.02	6.02	6.30	7.47	7.28
CGN Component M	7.27	7.27	5.19	6.80	5.88
CGN Component N	7.07	6.28	6.18	7.80	7.23
CGN Component O	6.25	6.52	5.15	6.02	6.01
SUBTLEX	6.70	6.65	3.56	5.74	6.92
CELEX	5.94	6.45	5.64	5.80	5.45

Table A.2: *Natural logarithms of the frequencies (number of tokens per million word tokens) of five words investigated in Study 1, in the Dutch part of each component of the Spoken Dutch Corpus, in SUBTLEX-NL and in CELEX.*

Corpus	helemaal	mogelijk	natuurlijk	ongeveer	precies
CGN Component A	7.54	4.42	7.28	5.11	6.58
CGN Component B	7.29	5.38	7.66	5.10	5.73
CGN Component C	7.48	4.44	7.74	4.99	7.11
CGN Component D	7.66	4.26	7.61	4.89	6.99
CGN Component E	7.24	4.99	8.18	4.76	6.59
CGN Component F	6.98	5.85	7.87	5.37	6.49
CGN Component G	6.44	6.55	7.23	4.50	6.14
CGN Component H	7.13	5.31	7.29	5.30	6.59
CGN Component I	6.92	5.37	7.22	4.67	5.40
CGN Component J	6.98	6.00	7.65	5.86	6.00
CGN Component K	5.66	6.14	3.45	5.63	5.02
CGN Component L	6.96	5.78	7.40	5.23	5.82
CGN Component M	6.57	5.19	5.88	5.19	5.88
CGN Component N	6.28	6.37	7.70	5.62	5.35
CGN Component O	6.37	5.34	6.06	4.90	5.43
SUBTLEX	6.31	5.16	6.33	4.51	5.78
CELEX	6.05	6.22	6.14	5.11	5.46

A.3 Target words in Study 2

ge'babbeld, ge'bakken,ge'balsemd, ge'barsten, geba'seerd, ge'beden, ge'beld, ge'boden, ge'boekt, ge'bogen, ge'bonden, ge'bouwd, ge'daald, ge'daan ge'dacht, ge'danst, ge'deeld, ge'dekt, ge'doken, ge'dompeld, ge'durfd, ge'duurd, ge'duwd, ge'feest, ge'fietst, ge'filmd,

ge'filterd, gefi'xeerd, ge'fokt, ge'fopt, gefor'ceerd, gefu'seerd, ge'gaan ge'geten, ge'geven, ge'giecheld, ge'gild, ge'gokt, ge'golfd, ge'gooid, ge'goten, ge'gund, gekal'meerd, gekam'peerd, ge'keerd, ge'keken, ge'kend, ge'ketend, ge'kocht, ge'koesterd, ge'komen, ge'kookt, ge'koppeld, ge'kost, ge'kozen, ge'lachen, ge'laden, ge'lakt, ge'laten, ge'leden, ge'leefd, ge'leend, ge'leerd, ge'legd, ge'legen, ge'legerd, ge'leid, ge'leken, ge'lekt, ge'lepeld, ge'leund, ge'leverd, ge'lezen, ge'logen, ge'lokt, ge'loodst, ge'lopen, ge'lucht, ge'luisterd, ge'lukt, ge'lummeld, ge'maaid, ge'maakt, gemar'keerd, ge'matigd, ge'meend, ge'mengd, ge'merkt, ge'meten, ge'mist, ge'mogen, ge'molken, ge'mompeld, ge'morst, ge'naaid, ge'naderd, ge'nageld, gene'geerd, ge'neigd, ge'niesd, ge'noemd, ge'nomen, geno'teerd, ge'nummerd, ge'nuttigd, ge'pakt, gepar'keerd, gepas'seerd, ge'past, ge'pind, ge'poetst, ge'pompt, ge'puzzeld, ge'raakt, ge'raapt, ge'raden, ge'ramd, ge'rangschikt,ge'red, ge'reden, ge'regeld, ge'regend, ge'reisd, ge'rekend, ge'rekt, ge'remd, ge'rend, ge'rept, ge'richt, ge'roddeld, ge'roemd, ge'roepen, ge'rommeld, ge'roofd, ge'roosterd, ge'ruild, ge'rukt, ge'ruzied, ge'seind, geser'veerd, ge'sierd, gesor'teerd, ge'sukkeld, ge'surfd, ge'sust, ge'tankt, ge'tekend, ge'teld, ge'test, ge'tikt, ge'tild, ge'timmerd, ge'toetst, ge'toond, ge'vallen, ge'vangen, ge'varen, ge'veegd, ge'veild, ge'vestigd, ge'vierd, ge'voeld, ge'voerd, ge'volgd, ge'vonden, ge'vorderd, ge'vormd, ge'vouwen, ge'vuld, ge'zakt, ge'zegd, ge'zegend, ge'zet, ge'zeten, ge'zeurd, ge'zien, ge'zocht, ge'zogen, ge'zongen, ge'zorgd

A.4: Predictors tested in Study 2.

Predictors for schwa presence, schwa duration, /t/ presence and /t/ duration:
Register (carefully versus casually produced); Block; Speech Rate; Word length (two or three syllables); log Word frequency; Conditional probability of the target word given the preceding word; Conditional probability of the target word given the following word; log duration of preceding segment; log duration of following segment.

Predictors for schwa presence and schwa duration only:
Voicing, place, manner of the following consonant; whether the following segment was /x/.

Predictors for /t/ presence and /t/ duration only:
Type of following segment (vowel or consonant); place, manner, voicing of preceding segment.