Informal speech processes can be categorical in nature, even if they affect many different words

Iris Hanique[a)] and Mirjam Ernestus
Centre for Language Studies, Radboud University Nijmegen and Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, the Netherlands

Barbara Schuppler
Signal Processing and Speech Communication Laboratory, Graz University of Technology, Inffeldgasse 16c/EG, A-8010 Graz, Austria

(Received 4 April 2012; revised 2 November 2012; accepted 22 January 2013)

This paper investigates the nature of reduction phenomena in informal speech. It addresses the question whether reduction processes that affect many word types, but only if they occur in connected informal speech, may be categorical in nature. The focus is on reduction of schwa in the prefixes and on word-final /t/ in Dutch past participles. More than 2000 tokens of past participles from the Ernestus Corpus of Spontaneous Dutch and the Spoken Dutch Corpus (both from the interview and read speech component) were transcribed automatically. The results demonstrate that the presence and duration of /t/ are affected by approximately the same phonetic variables, indicating that the absence of /t/ is the extreme result of shortening, and thus results from a gradient reduction process. Also for schwa, the data show that mainly phonetic variables influence its reduction but its presence is affected by different and more variables than its duration, which suggests that the absence of schwa may result from gradient as well as categorical processes. These conclusions are supported by the distributions of the segments' durations. These findings provide evidence that reduction phenomena which affect many words in informal conversations may also result from categorical reduction processes. © *2013 Acoustical Society of America*.
[http://dx.doi.org/10.1121/1.4790352]

PACS number(s): 43.70.Fq, 43.70.Jt [AL] Pages: 1644–1655

I. INTRODUCTION

In conversational speech (CS), words are often not produced in their full forms, that is, in the form that is used in formal situations (e.g., Ernestus, 2000; Johnson, 2004). Segments may be very short, altered (e.g., a voiced plosive may be pronounced as a glide), or even be completely absent. The English word *particular*, for example, has the full form /pətɪkjələ˞/ but in CS may also be produced as /ptɪkə˞/. Similarly, the Dutch word *gewenst* "wished" may be produced as /xwɛns/ instead of /xəwɛnst/. Recently, researchers have started to investigate the processes underlying this type of pronunciation variation. This paper contributes to this line of research by studying the processes underlying the two most frequent reduction phenomena in Dutch, which affect many words but only if they occur within a sentence in informal speech.

Browman and Goldstein (1990) hypothesized that all pronunciation variants that typically occur in casual speech result from a gradient increase in overlap and a gradient decrease in the magnitude of articulatory gestures. Gestural overlap may result in blending if two neighboring segments share an articulator. One or both segments are then pronounced incompletely. An example of this blending process is the pronunciation of English *this shop* as /ðɪʃːɒp/, with a long /ʃ/ but without /s/. Gestural overlap may result in the complete hiding of a segment if two neighboring segments are pronounced with different articulators. For instance, the closure of a bilabial stop may hide a preceding /t/ by making the release of the alveolar constriction inaudible.

Browman and Goldstein found support for their hypothesis in several articulatory studies (see Browman and Goldstein, 1992, for an overview). In addition, their hypothesis is supported by several acoustic studies. Davidson (2006) for instance studied the absence of pretonic schwa in consonant-schwa-consonant sequences in American English (e.g., in the words *tomato* and *support*). She found that schwa is less often absent if the preceding segment is voiced or is a voiceless stop which can only overlap partially with schwa. In addition, if schwa is absent in /#səp/ sequences, /p/ is often aspirated, which suggests that the articulatory gestures for schwa are still present. Furthermore, Davidson (2006) also showed that preceding voiceless fricatives and following /l/s tend to be longer if schwa is absent. All these results indicate that the absence of schwa in the acoustic signal is due to co-articulation with surrounding consonants.

Another acoustic study supporting Browman and Goldstein's hypothesis was conducted by Torreira and Ernestus (2010). They studied vowel devoicing in French, a reduction phenomenon that is mostly restricted to fast and casual speech. They showed that partial devoicing is affected by the same predictors as complete devoicing (speech rate, manner of articulation of the preceding consonant, distance to the

[a)]Author to whom correspondence should be addressed. Electronic mail: iris.hanique@ mpi.nl

following Accentual Phrase boundary). These results suggest that both partial and complete devoicing result from the same reduction process, which is consequently gradient in nature.

The pronunciation variants resulting from co-articulation may be stored in the mental lexicon. These representations then have to specify the acoustic details of the variants, for instance the exact durations of their segments. They therefore differ from the abstract representations in the form of phonemes, which are assumed by most models of speech production (e.g., Levelt *et al.*, 1999). Since gradient reduction by definition results in a continuum of pronunciation variants, the storage of these variants also implies a complex mental lexicon.

Importantly, several authors assume that the absence of segments may not only result from gradient processes but also from categorical higher-level processes. One such categorical process is a phonological rule that deletes a segment completely (e.g., Chomsky and Halle, 1968; McCarthy and Prince, 1993). In addition, the mental lexicon may contain representations for multiple pronunciations of a word (e.g., Goldinger, 1998), and the absence of a segment may result from the selection of a variant with missing segments.

Support for the categorical nature of some reduction phenomena has been found in corpus-based and psycholinguistic studies. For instance, Torreira and Ernestus (2011) studied the processes underlying /e/-reduction in the highly frequent word combination *c'était* /setɛ/ "it was" using the Nijmegen Corpus of Casual French. The duration of /s(e)/ showed a bimodal distribution instead of a unimodal distribution, with one mode containing most tokens with /e/ and the other containing most tokens without /e/. Furthermore, the presence and duration of /e/ were conditioned by different variables: Whereas the presence of /e/ was conditioned by speech rate and the distance to the end of the Accentual Phrase, the duration of /e/ was conditioned by position in the Intonational Phrase and the duration of preceding /s/. Given these results, the authors concluded that the absence and shortening of /e/ can result from different processes, and that the absence of /e/ in French *c'était* sometimes results from a categorical process rather than from extreme vowel shortening.

Bürki and colleagues investigated the nature of word-internal schwa deletion in French. This deletion phenomenon is very frequent in many varieties of French and can also occur in formal speech and when words are uttered without context. Bürki *et al.* (2011b) showed that this schwa deletion phenomenon can result from a gradient shortening process, as the duration of schwa varies from very long to very short, like the duration of other vowels. In addition, it can result from categorical processes; Bürki *et al.* (2010) demonstrated in several production studies that speakers of French produce a pronunciation variant (e.g., *fenêtre*) faster the more frequent it is in comparison with the other variant (e.g., *fnêtre*), which suggests that both variants are stored. In a corpus study, Bürki *et al.* (2011a) found differences between the sets of variables that condition word-internal schwa duration and its absence. Schwa duration was influenced by voicing of the preceding and following consonants and by the word's predictability given the preceding word, whereas its absence was conditioned by word position in the utterance, the number of consonants in the sequence that results from the absence of schwa, and whether this sequence adheres to the sonority principle. The authors of these last two studies concluded that the absence of schwa is not always the extreme result of schwa shortening but can also result from a categorical process.

So far, evidence for the categorical absence of segments is restricted to the types of phenomena discussed above. They represent reduction in highly frequent (function) words or word combinations and reduction that also occurs in formal speech and in words uttered in isolation. This raises the question whether reduction processes that are restricted to informal speech and affect many different word types may also be categorical. If so, this has consequences for our view of the speech production process: It would indicate that the mental lexicon contains an immense number of pronunciation variants or that there are many more categorical (i.e., phonological) reduction processes than previously assumed, which are specified for the conditions in which they can be applied.

In this paper, we investigate this question by studying two reduction processes in Dutch that affect many words, mainly in informal connected speech. Although both processes have been well documented, little is known about their underlying nature.

The first process is the reduction of schwa, which has been studied before in several languages, such as English (e.g., Dalby, 1984; Patterson *et al.*, 2003; Davidson, 2006), French (e.g., Bürki *et al.*, 2011a), and Dutch (e.g., Van Bergem, 1994; Pluymaekers *et al.*, 2005). Schuppler *et al.* (2011) reported that in Dutch 44.7% of all schwas are absent. Generally, the studies on Dutch have shown that schwa tends to be more reduced in faster speech, in words that occur more frequently, after fricatives and before plosives, and in words in the middle of utterances.

The second process that we will examine is the reduction of /t/, which has been studied, among others, in Dutch (e.g., Booij, 1995; Cho and McQueen, 2005; Mitterer and Ernestus, 2006; Schuppler *et al.*, 2012) and English (e.g., Gregory *et al.*, 1999; Gahl and Garnsey, 2004). In careful Dutch, /t/ is usually produced with a closure and a burst. In informal Dutch, the closure and burst of /t/ may be weakened, shortened, and even be completely absent (Mitterer and Ernestus, 2006). For this speech register, /t/ has been reported to be absent in 11.9% of all word tokens and in one-third of the word tokens in which it is in final position (Schuppler *et al.*, 2011). Overall, previous research has shown that in Dutch /t/ is more likely to be reduced in more informal speech, in unstressed syllables, in non-final sentence position, in more predictable word pairs, and if preceded by /s/ or followed by a bilabial consonant than by vowels or other consonants. In some highly frequent words, such as *niet,* /t/ can also be absent in more formal speech registers.

By studying these two segments in the same words and with the same methods, we can reveal differences that can be directly ascribed to their phonetic characteristics, to their position in the word, or, more importantly, to the nature (categorical versus gradient) of the underlying processes. Previous studies have investigated the reduction of different segments within one word (e.g., Pluymaekers *et al.*, 2005) but only a few, if any, have addressed the question of whether these segments were reduced as a consequence of the same processes.

Our study focuses on the reduction of schwa and /t/ in Dutch past participles since both segments occur in regular past participles. Most Dutch regular past participles consist of a verb base and the circumfix /xə/ + /t/ (e.g., /mak/ *maak* "make": /xə+mak+t/ *gemaakt* "made"). Past participles of verbs that start with one of the prefixes *be*-, *er*-, *ge*-, *her*-, *ont*-, or *ver*- generally do not have an additional prefix *ge*- (e.g., /bətal/ *betaal* "pay": /bətal+t/ *betaald* "paid"). Irregular past participles typically deviate from regular past participles in that they end in /ə(n)/ instead of /t/ (e.g., /lop/ *loop* "walk": /xə+lop+ə(n)/ *gelopen* "walked"), or their base shows a vowel change (e.g., /krœyp/ *kruip* "crawl": /xə+krop+ə(n)/ *gekropen* "crawled").

This paper presents two corpus studies. Study 1 investigates the reduction of schwa, focusing on the initial syllables *ge*-, *be*-, and *ver*-, with the full forms /xə/, /bə/, and /vər/, respectively. These initial syllables all start with a consonant and given that the word is a past participle, the following schwa is almost completely predictable, and therefore probably prone to reduction. Importantly, the three syllables occur in many word types of various frequencies, which allow us to examine word frequency effects. In Study 2, we focus on reduction of word-final /t/.

For both schwa and /t/, we examined which variables condition their acoustic presence versus absence and which variables condition their durations. If a segment's absence and duration are conditioned by the same variables, we hypothesize that they result from the same shortening process, which in extreme cases results in the absence of a segment. If its absence is conditioned by different variables to its duration, we then hypothesize that its absence results from a categorical process. We studied variables that are known to play important roles in pronunciation variation, such as speech register (e.g., Van Son and Pols, 1999; Van Bael *et al.*, 2004), the position of a word in the utterance (e.g., Cambier-Langeveld, 2000; Bell *et al.*, 2003), speech rate (e.g., Dalby, 1984; Davidson, 2006), the characteristics of the surrounding consonants (e.g., Van Bergem, 1994; Mitterer and Ernestus, 2006), and word predictability (e.g., Pluymaekers *et al.*, 2005; Bell *et al.*, 2009). In order to be able to study register, we based our studies on CS from the Ernestus Corpus of Spontaneous Dutch (ECSD; Ernestus, 2000) and interviews (IN) and read speech (RS) from the Spoken Dutch Corpus (CGN; Oostdijk, 2002).

Following Torreira and Ernestus (2011) and Bürki *et al.* (2011b), we also examined the distribution of the duration of the segment. If the absence of a segment only results from extreme gradient shortening, the duration is expected to show a unimodal distribution. In contrast, if the absence of that segment can result from both a gradient and a categorical process, this distribution is expected to be bimodal with one peak at 0 ms (since the absence of the segment can result from categorical processes and extreme gradient shortening) and one peak around the segment's average duration.

In summary, the present corpus study extends the research on pronunciation variation by investigating the nature of the processes underlying two highly frequent phenomena (i.e., reduction of schwa and /t/ in Dutch) that affect many word types but that hardly occur in formal speech or words uttered in isolation. By comparing the variables that affect a segment's presence and duration, and by inspecting the distribution of these segments' durations, we investigate whether the absence of a segment results only from extreme shortening or can also result from a categorical deletion process.

II. STUDY 1: SCHWA REDUCTION

A. Method

1. Materials

We extracted past participles from three subcorpora with different speech registers. The first corpus is the ECSD (Ernestus, 2000), which consists of 15 h of casual dialogues between 10 pairs of speakers. All 20 speakers were native speakers of standard Dutch, born and raised in the West of the Netherlands, aged between 21 and 55, and they all held academic degrees. The second and third subcorpora are components of the CGN (Oostdijk, 2002): We used those parts of the 43 h of IN with 681 speakers and the 64 h of lively read passages of the library of the blind produced by 324 speakers, that were phonetically transcribed (see below) at the moment of this study.

We extracted all past participles produced without background noise or overlapping speech from other speakers. Because *ge*- and *be*- consist of a consonant plus schwa, we only extracted tokens of *ver*- where /r/ was absent (which is the most common pronunciation of *ver*- and these tokens therefore form the majority of tokens). Furthermore, tokens directly preceded or followed by hesitations or with an extremely short (i.e., shorter than 15 ms) or long (i.e., longer than 478 ms) consonant preceding schwa were discarded. Since it is difficult to determine the presence and duration of schwa if it is followed by a vowel, we only used words for our study whose second syllable start with a consonant. All bi-syllabic and most longer past participles were stressed on the second syllable (e.g., /xə'dɑnst/ *gedanst* "danced," /bə'vɛstəxt/ *bevestigd* "confirmed," /xə'wɑndəlt/ *gewandeld* "walked"). Table I presents an overview of the final data set of schwa in word-initial syllables. The number of tokens per word type ranged from 1 to 151. Compared to the frequent words and word combinations discussed in the literature on

TABLE I. Number of tokens and types (between parentheses) in the data set for Study 1, divided by the three types of initial syllable and the three speech registers, namely CS, IN, and RS.

	Total	CS	IN	RS
Total	2150 (562)	738 (195)	539 (205)	873 (393)
ge-	1695 (387)	624 (145)	426 (145)	645 (268)
be-	242 (86)	57 (22)	67 (33)	118 (63)
ver-	213 (88)	57 (28)	46 (27)	110 (62)

the processes underlying reduction, the words in our data set are all of a low frequency. The most frequent word in our study was *geweest* "been," which occurs 54 times per 100 000 word tokens (based on counts from the data used for this study), whereas the highly frequent bigram *c'était* "it was" studied by Torreira and Ernestus (2011) occurs 280 times per 100 000 bigrams (based on the Nijmegen Corpus of Casual French). Interestingly, the number of word types in RS is twice as high as the number in CS and IN, indicating a difference in the type-token ratio.

2. Measurements

We determined the presence versus absence of schwa and its duration on the basis of automatically generated broad phonetic transcriptions. Automatic transcriptions are more consistent than human transcriptions and can be easily generated for large quantities of speech. For all materials used, we followed the forced alignment procedure described by Schuppler *et al.* (2011). For each word token in the orthographic transcriptions an automatic speech recognition (ASR) system retrieved the full phonetic form and possible pronunciation variants from a pronunciation lexicon. For instance, for the word *gemaakt* "made" it retrieved the full form /xəmakt/ and the variants /xmakt/, /xəmak/, and /xmak/. The possible pronunciation variants resulted from the application of 32 reduction rules to the full pronunciations of the words. These rules were based on earlier observations of reduced variants in casual Dutch (e.g., Ernestus, 2000), and deleted, among others, schwa in the syllables *ge*-, *be*-, or *ver*-, and word-final /t/. From all retrieved pronunciation variants, the ASR system HTK (Hidden Marcov Model Toolkit; Young *et al.*, 2002) selected the variant that best matched the speech signal, using 37 monophone models. These 32-Gaussian tri-state models (Hämäläinen *et al.*, 2009) had been trained on the RS component of the CGN. Because the acoustic models consisted of at least 3 emitting states and the frame shift was 5 ms, segments were assigned a minimum duration of 15 ms (even if the actual duration was shorter). On average, schwa was assigned a duration of 48 ms (range 15–225 ms) and /t/ a duration of 85 ms (range 15–355 ms).

In order to validate the automatic transcriptions, two of the authors manually transcribed 148 schwas in the ECSD data set. We chose ECSD for the validation because its speech style is the hardest to transcribe. The manual transcriptions were based on the audio and the waveform of the target word and a few surrounding words. The two transcribers first decided on transcription criteria and thereafter transcribed schwa and the preceding and following segments independently from each other using the speech analysis software package *Praat* (Boersma, 2001). They did so without being aware of the results of the ASR system.

As the schwas that were transcribed belonged to the syllables *ge*-, *be*-, or *ver*-, they were preceded by /x/, /b/, /p/, /v/, or /f/. The transcribers placed the boundary between a fricative and schwa at the offset of frication noise. The boundary in a /bə/ sequence was located directly after the burst of the plosive. Schwa could be followed by different kinds of consonants. If it was a fricative, the boundary was formed by the onset of frication noise. If it was a plosive, the boundary was located at the onset of the closure. Finally, if schwa was followed by a sonorant, the transcribers defined the transition as sudden changes in the audio or qualitative changes in the shape or complexity of the periodicity of the waveform.

Table II compares the presence of schwa in the word-initial syllables as annotated by the ASR system and the two transcribers. We found an agreement of 82.4% and a Cohen's kappa of 0.64 between the two human transcribers, and agreements of 75.7% and 77.0% and kappas of 0.51 and 0.53 between the ASR system and each of the human transcribers. A regression model with the dependent variable *agreement on the presence of schwa*, the independent variable *transcriber pair* and the random variable *word type* showed no significant effects (A-H1 versus A-H2: $\beta = 0.01$, $t = 0.30$, *n.s.*; A-H1 versus H1-H2: $\beta = 0.07$, $t = 1.48$, *n.s.*; A-H2 versus H1-H2: $\beta = 0.05$, $t = 1.19$, *n.s.*). This indicates that the agreements between the ASR and each human transcriber do not differ substantially from the agreement between the two human transcribers. These agreements are also similar to other transcriber agreements obtained for informal speech. For example, Kipp *et al.* (1997) reported agreements between three human-made transcriptions of spontaneous German of at least 78.8%, and Pitt *et al.* (2005) presented an overall agreement of 80.3% between human-made transcriptions of conversational American English.

We also compared the durations of those schwas that were transcribed as present by the human transcribers or the ASR system. For this analysis, three outliers with durations longer than 210 ms were excluded (as they were in the duration analysis of Sec. II B). Table II presents the results (including mean duration differences) without these outliers. A regression model with *transcriber* as the independent factor and *word type* as the random factor indicated that the durations assigned by the different transcribers differed significantly (1-A: $\beta = 0.02$, $t = 6.23$, $p < 0.0001$; 1-2: $\beta = 0.01$, $t = 2.04$, $p < 0.05$; 2-A: $\beta = 0.02$, $t = 4.22$, $p < 0.0001$). Those of the ASR system are the longest, which is unsurprising as these are minimally 15 ms. Those of Transcriber 2 are longer than those of Transcriber 1. Further comparison of the durations assigned by the ASR system and the human transcribers showed that the differences were mainly caused by three outliers (5% of the data points). Removal of these data points resulted in mean differences of 12.8 and 3.7 ms between the ASR system and Transcribers 1 and 2, respectively. These average differences are similar to those obtained in other studies. For example, Wesenick and Kipp (1996) reported that 96% of the segment boundaries determined by three

TABLE II. Comparison of the automatically generated (A) and the human-made (H1 and H2) transcriptions.

	Presence of schwa			Duration of schwa		
Comparison	N	Agreement	κ	N	Mean difference	95% confidence interval
A-H1	148	75.7%	0.51	62	18.0 ms	10.6–25.4 ms
A-H2	148	77.0%	0.53	63	8.5 ms	1.5–15.5 ms
H1-H2	148	82.4%	0.64	69	8.2 ms	5.3–11.1 ms

human transcribers for 64 read sentences of German differ by less than 20 ms. Similarly, Raymond *et al.* (2002) presented an average alignment difference of 16.4 ms between human transcribers of the Buckeye corpus.

This evaluation shows that our automatically obtained transcriptions are comparable to human-made transcriptions. As it is currently not possible to easily obtain significantly better transcriptions for a large number of tokens, we accept our automatic transcriptions to be valid. Given the average duration of schwa (48 ms) and the mean differences in duration assigned by the different transcribers and the ASR system (3.7 and 12.8 ms), our data will, however, have to be interpreted with care.

The validation described above did not include tokens based on /h/-initial verbal bases (e.g., *gehaast* "rushed") because manual inspection showed that the automatic transcriptions of these tokens were not reliable. We therefore decided not to rely on the automatic transcriptions but have the first syllable and /h/ of 210 past participles transcribed by the same two transcribers. They used the same acoustic criteria as described above and identified boundaries between schwa and /h/ as transitions in the waveform and audio. The two transcribers agreed on the presence of schwa for approximately 73.3% of the tokens (κ: 0.46), and the differences in schwa duration were only 0.7 ms on average (with a 95% confidence interval from -3.3 to 1.9 ms) and non-significant. We included only those tokens in the data set for which the transcribers agreed on the presence of schwa, and we used the average durations from their transcriptions.

3. Predictors

We tested the role of several predictors that have previously been shown to condition segment reduction. A first predictor was *speech register* since reduction has been shown to be more prominent in more informal speech (e.g., Van Son and Pols, 1999; Van Bael *et al.*, 2004). This factor differentiated between the CS of the ECSD, the IN of the CGN, and the RS of the CGN. A second variable was *speech rate* since words tend to be more reduced in faster speech (e.g., Dalby, 1984; Davidson, 2006). We defined *speech rate* as the number of syllables/s in the continuous uninterrupted stretch of speech that contained the target word assuming that all words had been produced in their full forms. This variable ranged from 2.36 to 10.67 syllables/s.

In addition, the segmental context of schwa may affect its pronunciation (Davidson, 2006). We therefore tested the role of the identity of the first *syllable*. This predictor contrasted the syllable *ge*- with the syllables *be*- and *ver*-. The syllables *be*- and *ver*- were pooled since there appeared to be fewer tokens of *be*- ($n=242$) and *ver*- ($n=213$) than of *ge*- ($n=1695$), and because preliminary results indicated that *be*- and *ver*- show similar behavior. Note that this predictor also distinguishes between the different preceding consonants (i.e., /x/ versus /b/ and /v/). Since schwa may be co-articulated with the preceding consonant, which then may result in a longer preceding consonant and a shortened or absent schwa, another potential conditioning variable is the duration of the preceding segment (*duration preceding segment*). We used the log-transformed duration (see below), which ranged from 4.00 to 8.83. Another predictor indicated whether the places of articulation of the consonant preceding and the consonant following schwa are identical (*identical place*; identical: $n=530$, different: $n=1620$). If the surrounding consonants share their place of articulation, schwa tends to be longer due to the physiological difficulty of gesture repetition (e.g., Walter, 2007).

Several prosodic variables have also been shown to affect speech reduction. We tested the role of *word length* since segments tend to be shorter if they are followed by more syllables (Lindblom, 1963; Nooteboom, 1972). This binary factor distinguished bi-syllabic words ($n=1142$) from words with more than two syllables ($n=1008$). Another prosodic variable is *chunk finality*, which indicates whether the word was the last word of a chunk of speech (final: $n=872$; non-final: $n=1278$). A chunk is surrounded by pauses and may contain a part of a sentence, or one or multiple complete sentences. Especially in RS, a chunk corresponds to a prosodic unit, and words at the end of a chunk may therefore undergo prosodic lengthening (e.g., Cambier-Langeveld, 2000).

Finally, we added three predictability variables because words that occur more often or are more likely to occur given their context have been found to be more reduced (e.g., Pluymaekers *et al.*, 2005; Bell *et al.*, 2009). We tested: (1) The log-transformed *word frequency*; (2) the conditional probability of the target word (w_{target}) given the preceding word ($w_{\text{preceding}}$), which was estimated by

$$\log_2\left(\frac{\text{frequency}\left(w_{\text{preceding}}, w_{\text{target}}\right)}{\text{frequency}\left(w_{\text{preceding}}\right)}\right),$$

and (3) the conditional probability of the target word given the following word ($w_{\text{following}}$), calculated with

$$\log_2\left(\frac{\text{frequency}\left(w_{\text{target}}, w_{\text{following}}\right)}{\text{frequency}\left(w_{\text{following}}\right)}\right).$$

All frequency measures used for the calculation of these predictors are based on counts from all CGN components.

4. Analyses

We investigated the influences of the predictors on the presence of schwa and its duration by means of mixed effects regression models (Baayen *et al.*, 2008). In all models *speaker* and *word type* were considered as crossed random factors. Further, we used contrast coding, where for each qualitative variable, one level is treated as the standard level to which all other levels are compared. In order to create a normal distribution for the duration of schwa, we applied a log-transformation to this duration (milliseconds) and also transformed the duration of the preceding segment. In the duration analyses, segments differing more than 2.5 times the residual standard errors from the values predicted by the statistical models were considered outliers, thus removed, and the models were refitted.

Most studies assume that a predictor influences reduction if it is statistically significant in a statistical model. As a consequence, these studies also assume predictors to be relevant if they hardly improve or even worsen the model's relative goodness of fit. We initially applied this procedure but then decided to abandon it as our corpus study is based on many data points (i.e., 2150) and we found many significant effects, including high-order interactions. Most of these effects had very small effect sizes and several were not interpretable. We therefore adopted a more conservative approach based on the Akaike Information Criterion (AIC; Akaike, 1973). The AIC measure is defined as $-2\log\mathcal{L} + 2p$, where $\mathcal{L}$ is the maximum of the likelihood for the model, and p is the number of parameters in the model. We added predictors one by one to a regression model. A predictor or interaction was only retained if it showed a significant effect and if the model with this predictor or interaction had a lower absolute AIC value than the same model without this predictor or interaction. Only predictors that passed these criteria are presented in Sec. II B.

Before we added a predictor to a model, we checked for possible correlations with the predictors already present in that model. When predictor A correlated with predictor B, we first orthogonalized them. That is, we replaced predictor A by the residuals of a linear regression model that predicted this predictor A as a function of predictor B, or vice versa.

Following Torreira and Ernestus (2011) and Bürki *et al.* (2011b), we also examined the distribution of the duration of schwa. We restricted the analysis to those tokens of schwa that are not highly likely to be absent due to co-articulation with the surrounding segments. If we find a peak at 0 ms, this peak can therefore be interpreted as resulting from categorical reduction processes.

B. Results

The statistical results are presented in Table III. The analysis of the presence of schwa was based on 2150 tokens of past participles, 68.4% of which were produced with schwa. The statistical model showed significant effects of *speech register* and *syllable*, indicating that schwa was more often absent in CS (52.0%) than in IN (35.3%), and least often in RS (12.1%), and that it was more often absent in the syllable *ge*- (35.5%) compared to *be*- and *ver*- (17.4%). In addition, there were main effects of *word length* and *speech rate*, which were modulated by an interaction of these two variables. Schwa was more likely to be absent in bi-syllabic past participles, especially at higher speech rates (see Fig. 1). Finally, the *duration preceding segment* was significant, showing that schwa was more likely to be absent if preceded by a longer segment. The model estimated the standard deviation of the random variable *word type* at 0.82 and of *speaker* at 0.42.

The analysis of the duration of schwa was based on the 1470 schwas that were present. We found significant effects of *speech rate* and *identical place*: Schwa was shorter at higher speech rates and if the preceding and following consonant had different places of articulation (mean duration for different place: 44 ms; for identical place: 54 ms). Also, the effects of *syllable* and *speech register* were significant, and they were modulated by their interaction: In RS, schwa

TABLE III. Results for the presence and duration of schwa. Only those comparisons between levels of a factor that differ significantly are presented.

Presence of schwa	β	z	p
Speech rate	−0.43	−6.00	<0.0001
Syllable	−0.71	−3.61	<0.001
Speech register			
RS–CS	−2.21	−12.06	<0.0001
RS–IN	−1.29	−7.39	<0.0001
IN–CS	−0.92	−5.11	<0.0001
Word length	−1.66	−2.86	<0.01
Word length × speech rate	0.27	2.74	<0.01
Duration preceding segment	−0.16	−2.09	<0.05
Duration of schwa	β	t	p
Speech rate	−0.10	−6.68	<0.0001
Syllable	−0.26	−4.94	<0.0001
Speech register			
RS–CS	−0.36	−3.89	<0.001
RS–IN	−0.23	−2.58	<0.01
Syllable × speech register			
RS–CS	0.31	2.89	<0.01
RS–IN	0.32	2.99	<0.01
Identical place	0.29	5.72	<0.0001

tended to be shorter in *ge*- (mean: 43 ms) than in *be*- and *ver*- (mean: 52 ms). Finally, the model contained random effects involving *word type*: The standard deviation (in log ms) for the intercept was 0.22, the estimated standard deviations for IN was 0.36, and for CS it was 0.31, with estimated correlations of −0.20 (IN) and −0.11 (CS). The residual standard deviation was 0.52.

Figure 2 shows the kernel density plot for the duration of schwa. This plot is an estimation of the duration density and can be viewed as a smoothed version of a histogram. The plot is based on all data points that are not outliers (i.e., schwas shorter than 210 ms) and for which we expect no strong co-articulation (i.e., schwas not followed by a voiceless obstruent or a voiced fricative; Ernestus, 2000). Importantly, this plot clearly shows two peaks. This is as expected assuming that the absence of schwa can result from a categorical process and from a gradient shortening process.

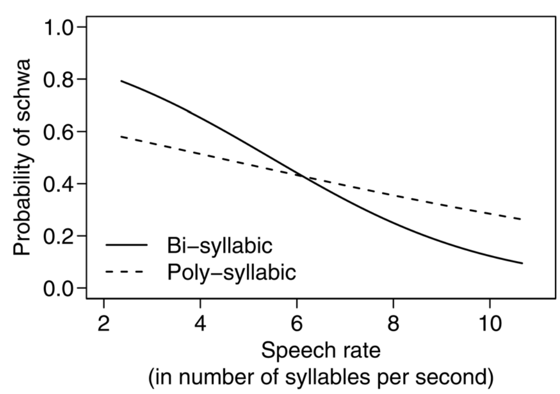

FIG. 1. The effects of speech rate and word length on the presence of schwa. The curves are adjusted for the means of all other predictors, and depict past participles from the ECSD carrying the prefix *ge*-.

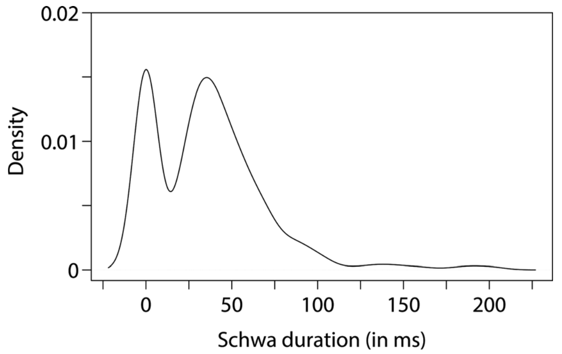

FIG. 2. The distribution of the duration of schwa.

III. STUDY 2: /t/ REDUCTION

A. Method

1. Materials

From the data set used in Study 1, we selected all past participles that ended in /t/. We excluded those that were directly followed by a plosive, as /t/ is then often unreleased (55.3% according to Schuppler *et al.*, 2011), which makes it difficult to determine the boundary between /t/ and the following plosive. The data set for Study 2 consisted of 1166 word tokens, representing 355 word types (404 word tokens representing 133 types from CS, 292 tokens representing 132 types from IN, and 470 tokens representing 227 types from RS).

2. Measurements, predictors, and analyses

To determine the presence versus absence of /t/ and its duration, we used the method described in Study 1. In connected speech, /t/ may have no oral closure and thus exist of weak frication only. Since the /t/ phone model of the ASR system was trained on both full and reduced /t/s; it is also able to recognize reduced /t/s (see, for instance, Fig. 3).

The reduction of /t/ was analyzed using the same method as in Study 1, and the predictors used in these analyses were also the same with a few exceptions: We did not include the variables *syllable* and *identical place* as these are specific for the initial syllable. Instead, we added a factor that identified the *type of segment preceding /t/*. Since /t/ is often reduced after obstruents, especially /s/ (e.g., Ernestus, 2000), this factor distinguished between /s/ ($n = 188$), /x,f/ ($n = 191$), /p,k/ ($n = 191$), and other segments ($n = 787$). Preliminary analyses showed that these other segments (vowels and sonorants) had the same effect.

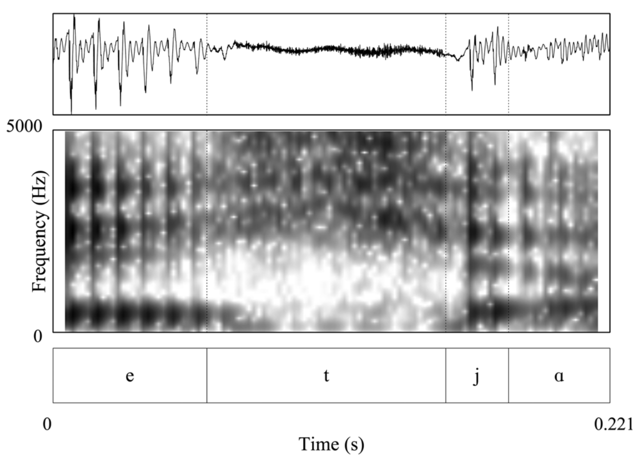

FIG. 3. An example of a reduced /t/, transcribed by the automatic system.

B. Results

Table IV presents the results of Study 2. The analysis of the presence of word-final /t/ was based on 1166 past participles, of which 77.9% were produced with /t/. The statistical model yielded significant main effects of *speech rate*, *speech register*, and *type of preceding segment*: Word-final /t/ was more likely to be absent at higher speech rates and in more informal speech (CS: 29.0%; IN: 29.1%; RS: 11.9%). In addition, it was most often absent if preceded by a fricative (/s/: 44.1%; /x,f/: 40.3%), than if preceded by a plosive (25.1%), and least often in other cases (8.4%). The standard deviation of the random effect *word type* was estimated at 0.49.

The statistical model for the durations of the 908 /t/s that were present showed significant effects of *speech rate* and *type of preceding segment*: Word-final /t/ tended to be shorter at higher speech rates, and if the preceding segment was a fricative, especially /s/ (mean durations: /s/ 69 ms; /x,f/ 78 ms; /p,k/ 85 ms; other 90 ms). Furthermore, *chunk finality* and *speech register* showed significant main effects and an interaction: Past participles tended to end in longer /t/s in chunk final than in chunk initial or medial position,

TABLE IV. Results for the presence and duration of /t/. Only those comparisons between levels of a factor that differ significantly are presented.

Presence of /t/	β	z	p
Speech rate	−0.20	−2.66	<0.05
Speech register			
RS–CS	−1.28	−6.31	<0.0001
RS–IN	−1.04	−4.73	<0.0001
Type of preceding segment			
other–/x,f/	−2.03	−7.64	<0.0001
other–/s/	−2.14	−6.62	<0.0001
other–/p,k/	−1.18	−4.18	<0.0001
/x,f/–/p,k/	0.85	2.78	<0.01
/s/–/p,k/	0.96	2.69	<0.01

Duration of /t/	β	t	p
Speech rate	−0.08	−4.32	<0.0001
Speech register			
RS–IN	−0.18	−2.33	<0.05
IN–CS	0.22	2.75	<0.01
Type of preceding segment			
other–/x,f/	−0.22	−3.60	<0.01
other–/s/	−0.46	−7.54	<0.0001
/x,f/–/s/	−0.25	−3.14	<0.01
/s/–/p,k/	0.35	4.79	<0.0001
Chunk finality	0.75	12.53	<0.0001
Chunk finality × speech register			
RS–CS	−0.53	−5.52	<0.0001
RS–IN	−0.34	−3.34	<0.01

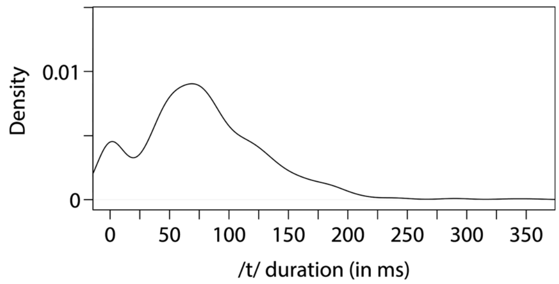

FIG. 4. The distribution of the duration of /t/.

especially in less informal speech (mean durations: CS final: 88 ms, CS non-final: 74 ms, IN final: 80 ms, IN non-final: 58 ms, RS final: 114 ms, RS non-final: 70 ms). The estimated standard deviation of *speaker* was 0.17 log ms, and the residual standard deviation was 0.55 log ms.

Figure 4 shows the kernel density plot of the distribution of /t/ duration. It is based on those tokens in which strong co-articulation is not very likely (i.e., if /t/ is not preceded or followed by a bilabial or alveolar segment nor by a fricative). This plot shows two peaks but the peak at zero is small and not very distinctive, suggesting that if a categorical process plays a role in /t/ duration, it is a very minor one.

IV. GENERAL DISCUSSION

This paper investigated whether highly frequent reduction processes that affect many word types but only in informal speech can be categorical in nature. In particular, we investigated the presence versus absence and duration of schwa and /t/ in Dutch past participles. If a segment's absence and duration are conditioned by the same variables, they probably result from the same gradient shortening process. In contrast, if the absence of a segment is conditioned by different variables than its shortening, we hypothesize that the absence may result from categorical processes; for instance, a phonological rule or the selection of a reduced pronunciation variant from the mental lexicon.

We extracted past participles from the ECSD (Ernestus, 2000) and the CGN (Oostdijk, 2002). Study 1 analyzed 2150 schwas in the word-initial syllables *ge*-, *be*-, and *ver*-, and Study 2 focused on the subset of 1166 past participles ending in /t/. We determined a segment's presence and duration on the basis of automatically generated broad phonetic transcriptions and found that schwa and /t/ were present in 68.4% and 77.9% of the tokens, respectively. An evaluation of the automatic transcriptions revealed that they were comparable to human-made transcriptions.

The reduction of schwa and /t/ have been hypothesized to occur more often in more informal speech (e.g., Van Son and Pols, 1999; Van Bael *et al.*, 2004). Our data confirm this hypothesis. In our CS, we found that schwa and /t/ were absent in 44.9% and 29.0% of the tokens, respectively, whereas in the RS they were absent in only 12.1% and 11.9% of the tokens, although the RS that we investigated consists of lively read passages from the library of the blind, and is not as formal as, for instance, speeches or lectures.

Our data present three types of evidence that the absence of /t/ is mostly due to extreme gradient shortening. The first evidence comes from Fig. 4, which shows the distribution of the duration of /t/ for those tokens that are surrounded by segments with which it is not easily co-articulated. This distribution shows only one clear peak, which suggests the absence of /t/ results from gradient shortening. The small peak at 0 ms, however, leaves open the possibility that /t/ may occasionally also be absent due to categorical processes. Possibly /t/ may be categorically absent in a small number of (high frequency) words.

The second type of evidence comes from the types of variables that affect /t/ reduction (see Table V) since the effects of all these variables may result from co-articulation. Like all co-articulation processes, /t/ was more often reduced in more informal speech and at higher speech rates. Moreover, it was also more reduced if preceded by a voiceless obstruent, particularly /s/. If articulatory gestures are decreased in magnitude (Browman and Goldstein, 1990), a /t/ may be realized with an incomplete closure and consequently be difficult to distinguish from a preceding obstruent. This is especially the case if this preceding obstruent is /s/, which shares its place of articulation with /t/. Finally, we found that /t/s were longer in chunk final position, especially in less informal speech. The chunks in RS more often coincided with the ends of prosodic units (e.g., intonational phrases) than the chunks in CS. As a consequence, the end of a chunk in RS is more likely to show final lengthening (see Cambier-Langeveld, 2000).

Third, the duration and presence of /t/ are influenced by exactly the same variables. The only exception is the position of the word in the chunk, which affected only /t/ duration. We probably did not find an effect of this variable on the presence of /t/ because analyses of binomial factors usually have less statistical power than analyses of continuous measures. Also, this similarity in variables suggests that the absence of /t/ mainly results from gradient shortening.

The results for schwa reduction are similar to those of /t/ reduction in that most effects on the presence and duration of schwa (see Table VI) can be interpreted as resulting from co-articulation. First, like other co-articulation processes, schwa reduction was more frequent in more informal speech and at higher speech rates. Second, schwa was more often absent after longer consonants, which strongly suggests that its absence often resulted from the articulatory overlap of the preceding consonant and schwa (Browman and Goldstein, 1990). Third, schwa tended to be longer if the places of articulation of the preceding and following consonants were identical. This finding is in line with findings by Walter (2007), who argues that the production of two subsequent consonants with the same place of articulation requires more

TABLE V. Significant predictors for the presence and duration of /t/.

Presence of /t/	Duration of /t/
Speech rate	Speech rate
Speech register	Speech register
Type preceding segment	Type preceding segment
	Chunk finality
	Chunk finality × speech register

TABLE VI. Significant predictors for the presence and duration of schwa.

Presence of schwa	Duration of schwa
Speech register	Speech register
Speech rate	Speech rate
Syllable	Syllable
	Syllable × speech register
	Identical place
Word length	
Word length × speech rate	
Duration preceding C	

effort than the production of a non-repetition sequence, which leads to longer vowels between two consonants with the same place of articulation.

Fourth, we found a clear effect of the different word-initial syllables. In general, schwa tended to be more often present in *be*- and *ver*- than in *ge*-. We found the same effect for the duration of schwa but only in RS. The difference between the syllables probably results from their morphological functions. The syllables *be*- and *ver*- are derivational pre-fixes and therefore contain crucial semantic information about the identity of the word, as illustrated by the minimal word pair /bətalt/ *betaald* "paid" versus /vərtalt/ *vertaald* "translated." In contrast, the syllable *ge*- is an inflectional prefix in 375 of the 387 examined word types, indicating that the word is a past participle (as in /xəfitst/ *gefietst* "cycled," whose stem does not occur with other prefixes), which is usually expected given the context. As a consequence, the syllables *be*- and *ver*- tend to be less redundant than *ge*-, and this may result in differences in reduction (see Hanique and Ernestus, 2012b, for an overview of the role of predictability in reduction). Probably, the effect on duration is only demonstrated in RS since, in this register, speakers are more aware of the listener and their speech tends to be more listener-driven than in less formal registers (Lindblom, 1990).

Finally, we found an interaction of word length with speech rate, with schwa being more often absent in bi-syllabic than in longer words, especially at higher speech rates. The direction of this effect is unexpected given previous studies which have shown that vowels tend to be shorter if they are followed by more syllables in the word (Lindblom, 1963; Nooteboom, 1972). We therefore conducted further analyses and found that bi-syllabic and longer words did not only differ in terms of the number of syllables but that they also differed in several measures of contextual predictability: For instance, the bi-syllabic words are significantly more frequent (mean log frequency: 9.06) than the longer words (mean log frequency: 7.68; $t = 10.94$, $p < 0.0001$). Since more predictable words are more likely to be reduced, especially at higher speech rates (e.g., Pluymaekers *et al.*, 2005), the effect of word length may actually be an effect of contextual predictability.

The variables affecting schwa reduction thus indicate that the absence of schwa also results from gradient shortening. In addition, the absence of schwa may result from categorical processes, as is indicated by two other types of results. First, the density plot of the duration of schwa (Fig. 2) clearly shows two modes: One mode representing the absent schwas and the other mode containing the present schwas. Second, if the absence of a segment only results from gradient shortening, we expect that its presence and duration would be affected by the same variables, as we found for /t/ above, and as has been found in several other studies (e.g., Torreira and Ernestus, 2010; Hanique and Ernestus, 2012a). In contrast to this expectation, we found that the presence of schwa but not its duration is affected by word length, an interaction of word length with speech rate, and the duration of the preceding consonant, whereas its duration but not its presence is affected by an interaction of the identity of the first syllable with speech register and by whether the surrounding consonants share their place of articulation (see Table VI). Moreover, we found that the presence of schwa was influenced by more variables than its duration. Given that analyses of continuous variables generally have greater statistical power than analyses of binomial factors, this difference is unexpected if the presence and duration of schwa result from the same articulatory process.

There is, however, one possible explanation for the greater number of predictors for the absence of schwa versus its duration. The duration data set contained 31.6% fewer data points than the data set on the presence of schwa, and the duration analysis may therefore have had less statistical power. However, for /t/, the duration data set also contained 22.1% fewer data points than the data set on its presence, and for /t/ we did not find that its duration was conditioned by fewer variables than its presence. It seems therefore unlikely that the difference in number of data points can explain the difference in the number of significant effects in the analyses of the presence and duration of schwa.

Furthermore, we examined whether a difference in the sizes of the data sets explains why the presence of schwa was affected by more variables than its duration, whereas this is not the case for /t/. We restricted the schwa data set to those 1162 tokens that also occur in the /t/ data set. Due to the smaller sample size, this data set showed fewer effects on the presence of schwa than the original data set (see the Appendix). Importantly, the variables that condition the presence and duration of schwa are still different. For /t/, we expanded the data set in order to obtain a data set that was as large as the original schwa data set. Since the corpus did not contain past participles that met all restrictions and that were not yet part of the data set, we expanded the data set with other word types (e.g., adjectives and other verb forms). This new data set consisted of 1993 tokens in total. The presence and duration of these new tokens of /t/ were determined with the same procedures described in Study 2. This analysis yielded three additional effects influencing *both* the presence and duration of /t/ (see the Appendix). Overall, these re-analyses show the same patterns as the original analyses. Apparently, statistical power cannot explain the differences in conditioning variables.

On the basis of these findings, we conclude that the absence of schwa in Dutch can result both from gradient and categorical reduction processes. In future research, the evidence for categorical reduction may be extended with more detailed acoustic analyses of our data, like the analyses reported on French *c'était* "it was" by Torreira and Ernestus (2011).

Thus far, categorical reduction processes have only been identified for highly frequent words or word combinations (e.g., /e/-deletion in French *c'était,* Torreira and Ernestus, 2011) and for phenomena that do not only occur in informal conversations but also in formal speech or in words uttered in isolation (e.g., word-internal schwa deletion in French, Bürki *et al.*, 2011a). Gradient reduction has been observed for reduction that affects many words but only if uttered in connected informal speech (e.g., vowel devoicing in French, Torreira and Ernestus, 2010). In line with this pattern, we found that reduction of word-final /t/ in Dutch, which affects many words in connected informal speech, is mainly gradient in nature. In contrast, the results for schwa point to a more complex picture. The comparison of the variables conditioning the absence and duration of schwa as well as the distribution of schwa duration suggest that extreme vowel shortening is not the only process that can result in its absence. Hence, categorical reduction is not restricted to highly frequent word combinations or phenomena that also occur in formal speech registers.

The categorical process that results in the absence of schwa may consist of either the application of phonological rules or the selection of a reduced pronunciation variant from the mental lexicon during speech production. If the absence of schwa results from phonological rules, we expect these rules to apply to all words equally often (word-specific rules would be similar to the storage of the reduced pronunciation variants together with their frequencies). Our analyses showed that the random factor *word type* has a standard deviation of 0.81, which indicates that words differ substantially in their probability that schwa is absent (approximately 20%). This suggests that words are not reduced equally often. The categorical absence of schwa is therefore not likely to result from phonological rules but from variant selection. This implies that for many words, the mental lexicon contains at least two pronunciation variants.

Our findings raise the question why the absence of /t/ results from a gradient process, whereas the absence of schwa can result from both a gradient and a categorical process. One possible explanation is that the schwa and /t/ in past participles differ in their perceptual prominence. Whereas schwa is part of the initial syllable, the /t/ is in word-final position. In addition, schwa is part of a prefix that contributes to the meaning of the past participle or signals that the word is a past participle, and therefore carries information. The /t/, in contrast, only indicates the past participle function, which in most cases is also signaled by the prefix (if it is *ge*-) or by the syntactic structure of the sentence. Finally, schwa is probably more prominent because of its acoustic characteristics: A reduced schwa is most likely to be clearly audible, whereas a /t/ with a weakened burst may be difficult to perceive. As a consequence of this difference in perceptual prominence, language users may more easily detect the absence of schwa in the prefix of past participles and build lexical representations of pronunciation variants without schwa. They may less often notice the absence of word-final /t/ and may interpret its absence as resulting from its low prominence. Accordingly, they may be less inclined to create lexical representations for past participles without /t/.

Unexpectedly, our data showed no effects of the word's (contextual) predictability on schwa and /t/ reduction. We believe that this is due to our conservative analysis method: We only retained those variables in our statistical models that were significant *and* improved the model's AIC value. If we do not apply this AIC criterium, effects of a word's frequency of occurrence and its predictability given the following word emerge. Unfortunately, it is currently not possible to draw general conclusions about the role of predictability in reduction. On the one hand, predictability effects reported in earlier studies may disappear if our conservative analysis method is applied to these studies. On the other hand since past participles tend to be more predictable than other types of content words, predictability may play no substantial role in the production of these particular verb forms.

An advantage of using corpus data over experimental data is that the examined speech is more natural. Our results suggest that corpus data can indeed be used to infer insights on the processes that drive pronunciation variation. More reliable phonetic transcriptions will even further increase the usefulness of corpus data. The findings based on corpus data should nevertheless be confirmed by data obtained from controlled experiments. It is currently not clear which experimental task would allow us to investigate the production of casual speech. Hanique and Ernestus (2012a) demonstrated, for instance, that one experimental paradigm that may seem promising, the shadowing task, elicits fast speech that cannot be compared to casual speech.

In summary, our results show that the absence of word-final /t/ results from extreme shortening. The absence of schwa, in contrast, can result from gradient vowel shortening as well as from a categorical process. We therefore conclude that reduction phenomena that mostly occur in informal conversations and that affect many different words can also result from categorical processes. This categorical process is probably the selection of a variant without schwa from the mental lexicon, which implies that for many words, the mental lexicon contains at least two pronunciation variants.

ACKNOWLEDGMENTS

The authors would like to thank Lou Boves and two anonymous reviewers for their useful comments. This work was partly funded by a European Young Investigator Award to M.E. and by the Marie Curie Research Training Network *Sound To Sense* and by the Austrian Science Foundation (FWF).

APPENDIX: RE-ANALYSIS

To examine whether the difference in power explains why the presence of schwa was affected by more variables than its duration, we re-analyzed schwa and /t/ with better matched data sets. The statistical models for schwa showed the same significant effects as Study 1 but the main effects of syllable and word length and the interaction between word length and speech rate on the presence of schwa were absent.

TABLE VII. Additional effects found in the re-analyses for the presence and duration of /t/. Only those comparisons between levels of a factor that differ significantly are presented.

Presence of /t/	β	z	p
Past participle	0.53	2.97	<0.05
Chunk finality	1.07	3.35	<0.01
Chunk finality × past participle	−0.65	−1.98	<0.05
Chunk finality × type preceding segment			
other–/x,f/	−1.26	−3.49	<0.001
Duration of /t/	β	t	p
Past participle	0.20	4.68	<0.0001
Chunk finality	0.94	14.52	<0.0001
Chunk finality × past participle	−0.79	−2.52	0.05
Chunk finality × type preceding segment			
other–/s/	−0.24	−2.16	<0.05
other–/p,k/	−0.22	−2.28	<0.05

For /t/, the statistical models showed the same effects as in Study 2 plus some additional effects (see Table VII for these additional effects): Word-final /t/ tended to be more often absent and shorter in word types other than past participles, and in chunk non-final position, particularly in words other than past participles. Furthermore, /t/ was more likely to be reduced in a chunk non-final position if not preceded by fricatives.

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

Baayen, R., Davidson, D., and Bates, D. (**2008**). "Mixed-effects modeling with crossed random effects for subjects and items," J. Mem. Lang. **59**, 390–412.

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

Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., and Gildea, D. (**2003**). "Effects of disfluencies, predictability, and utterance position on word form variation in English conversation," J. Acoust. Soc. Am. **113**, 1001–1024.

Boersma, P. (**2001**). "Praat, a system for doing phonetics by computer," Glot Int. **5**, 341–355.

Booij, G. E. (**1995**). *The Phonology of Dutch* (Clarendon Press, Oxford), Chap. 7, pp. 143–164.

Browman, C., and Goldstein, L. (**1990**). "Tiers in articulatory phonology with some implications for casual speech," in *Between the Grammar and Physics of Speech [Papers on Laboratory of Phonology 1]*, edited by J. Kingston and M. Beckman (Cambridge University Press, Cambridge), pp. 341–376.

Browman, C., and Goldstein, L. (**1992**). "Articulatory phonology: An overview," Phonetica **49**, 155–180.

Bürki, A., Ernestus, M., and Frauenfelder, U. (**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," J. Mem. Lang. **62**, 421–437.

Bürki, A., Ernestus, M., Gendrot, C., Fougeron, C., and Frauenfelder, U. (**2011a**). "What affects the presence versus absence of schwa and its duration: A corpus analysis of French connected speech," J. Acoust. Soc. Am. **130**, 3980–3991.

Bürki, A., Fougeron, C., Gendrot, C., and Frauenfelder, U. (**2011b**). "Phonetic reduction versus phonological deletion of French schwa: Some methodological issues," J. Phonetics **39**, 279–288.

Cambier-Langeveld, T. (**2000**). "Temporal marking of accents and boundaries," Ph.D. thesis, LOT, Utrecht, pp. 171.

Cho, T., and McQueen, J. (**2005**). "Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress," J. Phonetics **33**, 121–157.

Chomsky, N., and Halle, M. (**1968**). *The Sound Pattern of English* (Harper and Row, New York), pp. 470.

Dalby, J. (**1984**). "Phonetic structure of fast speech in American English." Ph.D. thesis, Indiana University, pp. 102.

Davidson, L. (**2006**). "Schwa elision in fast speech: Segmental deletion or gestural overlap?," Phonetica **63**, 79–112.

Ernestus, M. (**2000**). "Voice assimilation and segment reduction in casual Dutch. A corpus-based study of the phonology-phonetics interface," Ph.D. thesis, LOT, Utrecht, pp. 369.

Gahl, S., and Garnsey, S. (**2004**). "Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation," Language **80**, 748–775.

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

Gregory, M., Raymond, W., Bell, A., Fosler-Lussier, E., and Jurafsky, D. (**1999**). "The effects of collocational strength and contextual predictability in lexical production," Chicago Linguistics Soc. **35**, 151–166.

Hämäläinen, A., Gubian, M., ten Bosch, L., and Boves, L. (**2009**). "Analysis of acoustic reduction using spectral similarity measures," J. Acoust. Soc. Am. **126**, 3227–3235.

Hanique, I., and Ernestus, M. (**2012a**). "The processes underlying two frequent casual speech phenomena in Dutch: A production experiment," in *Proceedings of Interspeech-2012* (Portland, Oregon), pp. 2053–2056.

Hanique, I., and Ernestus, M. (**2012b**). "The role of morphology in acoustic reduction," Lingue e Linguaggio **11**, 147–164.

Johnson, K. (**2004**). "Massive reduction in conversational American English," in *Proceedings of the Workshop on Spontaneous Speech: Data and Analysis* (Tokyo, Japan), pp. 29–54.

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

Levelt, W., Roelofs, A., and Meyer, A. (**1999**). "A theory of lexical access in speech production," Behav. Brain Sci. **22**, 1–38.

Lindblom, B. (**1963**). "Spectrographic study of vowel reduction," J. Acoust. Soc. Am. **35**, 1773–1781.

Lindblom, B. (**1990**). "Explaining phonetic variation: A sketch of the H&H theory," in *Speech Production and Speech Modelling*, edited by W. Hardcastle and A. Marchal (Kluwer Academic Publishers, Dordrecht), pp. 403–439.

McCarthy, J. J., and Prince, A. (**1993**). "Generalized alignment," in *Yearbook of Morphology*, edited by G. E. Booij and J. Van Marle (Kluwer Academic Publishers, Dordrecht), pp. 79–154.

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

Nooteboom, S. (**1972**). "Production and perception of vowel duration: A study of durational properties of vowels in Dutch," Ph.D. thesis, University of Utrecht, pp. 166.

Oostdijk, N. (**2002**). "The design of the Spoken Dutch Corpus," in *New Frontiers of Corpus Research*, edited by P. Peters, P. Collins, and A. Smith (Rodopi, Amsterdam), pp. 105–112.

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

Pitt, M., Johnson, K., Hume, E., Kiesling, S., and Raymond, W. (**2005**). "The Buckeye corpus of conversational speech: Labeling conventions and a test of transcriber reliability," Speech Commun. **45**, 89–95.

Pluymaekers, M., Ernestus, M., and Baayen, R. (**2005**). "Lexical frequency and acoustic reduction in spoken Dutch," J. Acoust. Soc. Am. **118**, 2561–2569.

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

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

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

Torreira, F., and Ernestus, M. (**2010**). "Phrase-medial vowel devoicing in spontaneous French," in *Proceedings of Interspeech-2010* (Makuhari, Japan), pp. 2006–2009.

Torreira, F., and Ernestus, M. (**2011**). “Vowel elision in casual French: The case of vowel /e/ in the word c’était,” J. Phonetics **39**, 50–58.
Van Bael, C., Van Den Heuvel, H., and Strik, H. (**2004**). “Investigating speech style specific pronunciation variation in large spoken language corpora,” in *Proceedings of Interspeech-2004* (Jeju, Korea), pp. 2793–2796.
Van Bergem, D. (**1994**). “A model of coarticulatory effects on the schwa” Speech Commun. **14**, 143–162.
Van Son, R., and Pols, L. (**1999**). “An acoustic description of consonant reduction,” Speech Commun. **28**, 125–140.
Walter, M. (**2007**). “Repetition avoidance in human language,” Ph.D. thesis, MIT, pp. 225, online available at http://faculty.wcas.northwestern.edu/∼maw962/docs/walter-dissertation. pdf (Last viewed: 11/7/2012).
Wesenick, M., and Kipp, A. (**1996**). “Estimating the quality of phonetic transcriptions and segmentations of speech signals,” in *Proceedings of ICSLP-1996* (Philadelphia, PA), pp. 129–132.
Young, S., Evermann, G., Hain, T., Kershaw, D., Moore, G., Odell, J., Ollason, D., Povey, D., Valtchev, V., and Woodland, P. (**2002**). *The HTK Book 3.2* (Entropic, Cambridge).