

Co-occurrence of reduced word forms in natural speech

Malte C. Viebahn¹, Mirjam Ernestus^{2,1}, James M. McQueen^{3,4,1}

¹Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

²Center for Language Studies, Radboud University Nijmegen, The Netherlands

³Behavioural Science Institute, Radboud University Nijmegen, The Netherlands

⁴Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands

malte.viebahn@mpi.nl, m.ernestus@let.ru.nl, j.mcqueen@pwo.ru.nl

Abstract

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

Index Terms: speech production, spontaneous speech, speech reduction

1. Introduction

In spontaneous conversations, the way speakers produce words often deviates from dictionary pronunciations. For example, the schwa in the English word *separate* may be shortened or omitted resulting in the reduced form *sep'rate*. In Dutch, schwas are frequently reduced when they occur in the prefix of past participles [1]. For instance, the Dutch word *geweest* may be pronounced as *g'weest*. In the present study, we investigated whether reduced word forms are likely to co-occur in natural speech. The main question we asked was whether the degree of reduction of a given word is influenced by the degree of reduction of a previous word. In addition, we investigated if this effect is influenced by the time lag between the succeeding words and the degree to which words overlap phonologically.

Data on co-occurrences among word forms can inform us about the nature of lexical representation. Theories of the mental lexicon differ with respect to the degree to which phonetic variation is represented in lexical memory. On the one hand, there are abstractionist theories that claim that words are stored as single abstract representations [2]. On the other hand, there are episodic theories that assume that fine-grained phonetic detail of each occurrence of a word is stored [3]. Between these two extreme viewpoints there are hybrid theories that vary

with respect to the number of phonetic variants that are stored and the degree to which phonetic detail is retained in memory [4]. Whereas extreme abstractionist accounts predict no increase in acoustic similarity between word forms that occur together, extreme episodic theories predict that co-occurring words have similar acoustic properties, especially if they occur closely together and show greater phonological overlap.

2. Method

2.1. Materials

The data were extracted from the Ernestus Corpus of Spontaneous Dutch (ECSD) [5] and two components of the Corpus of Spoken Dutch (Corpus Gesproken Nederlands, CGN) [6]. We collected a total of 3,241 Dutch past participles from three speech registers: spontaneous speech from the ECSD, interviews from the CGN, and read speech from the CGN. We focused on past participles because they are frequent and often subject to reductions in Dutch [1]. Furthermore, they provide two different kinds of reductions: schwa reduction in the prefixes and /t/ reduction in the suffix. Following Hanique et al. [7], our analyses were restricted to tokens that fulfilled the following criteria: words were only included if they were not followed by a hesitation, participles with the prefix /vər/ were only included if the /r/ in the prefix was not realized, the duration of the following consonant (after the schwa) had to fall between 8 and 478 milliseconds, and, for the /t/ reduction data, words were only included if they were not followed by a word starting with a stop consonant (i.e. /tɔpɪk/). Each pair of past participles was spoken by the same speaker. The time lag between two succeeding past participles was defined as the interval between the offset of the first and the onset of the second past participle. Our analyses were based on the output of the automatic speech recognition system used by Schuppler et al. [8]. Using a very similar dataset, Hanique et al. [7] have shown that the agreement between human transcribers and the automatically generated transcriptions is approximately as high as the agree-

ment among the human transcribers.

2.2. Variables

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

2.3. Statistical analyses

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

3. Results

3.1. Schwa reduction

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

In the second model, the critical predictor variable

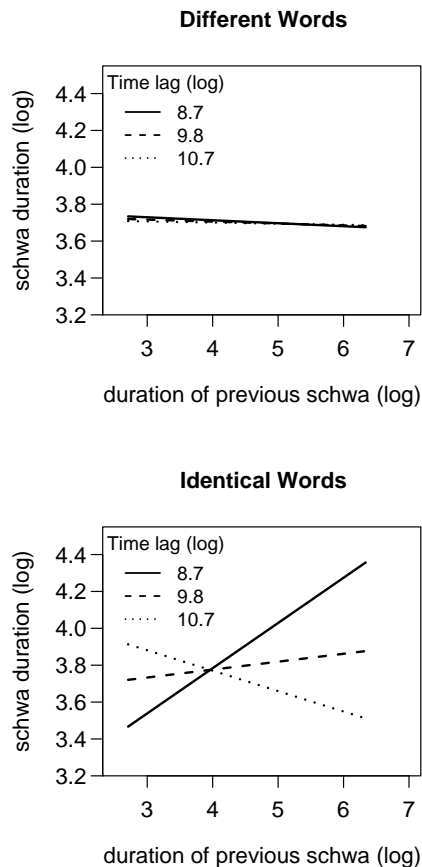


Figure 1: Interaction between relatedness, time lag, and duration of previous schwa for schwa duration predicted by previous schwa duration. The different lines for time lag represent the 25th, 50th, and 75th percentiles.

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

In the third model, the dependent variable was schwa duration and the critical predictor variable was the duration of the schwa in the preceding participle. All tokens in which the schwa was absent in either the first or the second past participle were removed, leaving 918 tokens. This model indicates a significant three-way interaction between the duration of the previous schwa, time lag, and relatedness (see Table 1 and Figure 1).

Separate analyses were then conducted for different and identical words. For different words ($N = 864$), we found no effects of the duration of the previous schwa,

Table 1: Results for schwa duration predicted by previous schwa duration (prev. schwa).

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

time lag, and no interaction. The only significant predictor was speech rate (shorter schwas at higher rates: $\beta = -0.07$, $t = -5.18$). In contrast, for the tokens that were preceded by identical words ($N = 54$), we found effects of the duration of the previous schwa ($\beta = 2.07$, $t = 3.83$) and lag ($\beta = 0.79$, $t = 3.33$) and these predictors interacted ($\beta = -0.21$, $t = -3.43$). We then performed a median split on time lag and ran separate models for tokens with lags shorter or longer than the median. For the short time lag, we found a significant effect of previous schwa duration ($\beta = 0.42$, $t = 3.07$): schwa durations of succeeding tokens are similar when they occur closely together in time. In contrast, we found no significant effect for the long time lag ($\beta = -0.03$, $t = -0.22$).

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

3.2. /t/ reduction

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

In the second model, /t/ presence was the dependent variable and the duration of the previous /t/ was the critical predictor variable. We excluded all tokens in which the /t/ was not present in the previous participle, leaving 635 tokens. Significant effects were present only for speech rate (fewer /t/s in faster speech: $\beta = -0.32$, $t = -3.28$) and speech register (fewer /t/s in spontaneous

than in read speech: $\beta = 1.47$, $t = 4.29$).

In the third model, /t/ duration was the dependent variable and the duration of the preceding /t/ the critical predictor variable. Tokens in which the /t/ was not present either in the first or the second participle were removed, leaving 496 data points. This model showed significant main effects of previous /t/ duration, relatedness, speech rate, and speech register (see Table 2). Shorter /t/s were associated with longer /t/s in the previous participle. Furthermore, /t/s were shorter when the preceding word was identical than when it was different.

Table 2: Results for /t/ duration predicted by previous /t/ duration (prev. /t/).

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

In the fourth model, /t/ duration was the dependent variable and the presence of the preceding /t/ was included as a predictor variable. We excluded tokens in which the /t/ was not present in the second participle, leaving 642 data points. We found significant effects of speech rate and speech register and an interaction of time lag and relatedness (see Table 3 and Figure 2). An analysis of the simple effects showed a significant negative slope for different words (shorter /t/s at longer lags: $\beta = -0.03$, $t = -2.02$; $N = 606$) and a significant positive slope for identical words (longer /t/s at longer lags: $\beta = 0.12$, $t = 2.36$; $N = 36$).

Table 3: Results for /t/ duration predicted by previous /t/ presence.

Predictor	β	t value
lag	-0.03	-1.94
identical words	-1.50	-3.39
speech rate	-0.08	-5.18
read speech (CGN)	0.06	1.07
interview speech (CGN)	-0.24	-3.898
lag \times identical words	0.14	3.07

4. Discussion and conclusions

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

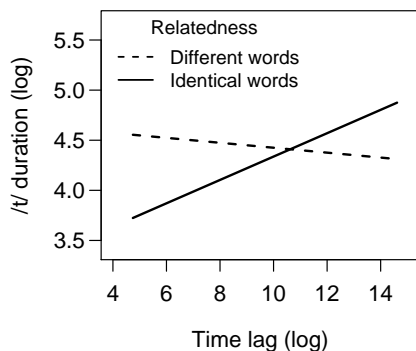


Figure 2: Interaction between relatedness and time lag for /t/ duration predicted by previous /t/ presence for spontaneous speech.

positive relationship between segment duration and time lag for identical words and a negative relationship for different words.

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

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

in the former model further supports the notion that it had insufficient power.

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

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

5. Acknowledgements

The authors would like to thank Tijn Grootswagers and Iris Hanique for help with the corpus analysis. This work was partly funded by an ERC starting grant (284108) to the second author.

6. References

- [1] M. Pluymaekers, M. Ernestus, and R. Baayen, "Lexical frequency and acoustic reduction in spoken Dutch," *The Journal of the Acoustical Society of America*, vol. 118(4), pp. 2561–2569, 2005.
- [2] W. Levelt, A. Roelofs, and A. Meyer, "A theory of lexical access in speech production," *Behavioral and Brain Sciences*, vol. 22, pp. 1–38, 1999.
- [3] S. D. Goldinger, "Echoes of echoes? An episodic theory of lexical access," *Psychological Review*, vol. 105, pp. 251–279, 1998.
- [4] A. Bürki, M. Ernestus, and U. Frauenfelder, "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*, vol. 62, pp. 421–437, 2010.
- [5] M. Ernestus, "Voice assimilation and segment reduction in casual Dutch. A corpus-based study of the phonology-phonetics interface," Ph.D. dissertation, Utrecht: LOT, 2000.
- [6] N. Oostdijk, "The design of the Spoken Dutch Corpus," in *New Frontiers of Corpus Research*, P. Peters, P. Collins, and A. Smith, Eds. Amsterdam: Rodopi, 2002, pp. 105–112.
- [7] I. Hanique, M. Ernestus, and B. Schuppler, "Can productive casual speech phenomena be categorical in nature?" submitted.
- [8] B. Schuppler, M. Ernestus, O. Scharenborg, and L. Boves, "Acoustic reduction in conversational Dutch: A quantitative analysis based on automatically generated segmental transcriptions," *Journal of Phonetics*, vol. 39, pp. 96–109, 2011.