Final /t/ reduction in Dutch past-participles: the role of word predictability and morphological decomposability

Iris Hanique[1,2], *Mirjam Ernestus*[1,2]

[1]CLS, Radboud University Nijmegen, the Netherlands
[2]Max Planck Institute for Psycholinguistics, the Netherlands

Iris.Hanique@mpi.nl, Mirjam.Ernestus@mpi.nl

Abstract

This corpus study demonstrates that the realization of word-final /t/ in Dutch past-participles in various speech styles is affected by a word's predictability and paradigmatic relative frequency. In particular, /t/s are shorter and more often absent if the two preceding words are more predictable. In addition, /t/s, especially in irregular verbs, are more reduced, the lower the verb's lemma frequency relative to the past-participle's frequency. Both effects are more pronounced in more spontaneous speech. These findings are expected if speech planning plays an important role in speech reduction.

Index Terms: pronunciation variation, acoustic reduction, corpus research, word predictability, morphological decomposability

1. Introduction

In everyday speech, words are often produced in reduced variants, in which segments are shorter or completely absent compared to their full forms [1]. For example, the Dutch word *gewenst* 'wished' has the full form /χəwɛnst/, but it may be pronounced as [χwɛns]. Research on when acoustic reduction tends to occur can aid the development of psycholinguistic speech production and comprehension models. Production models need information about which conditions favor reduced realizations and why, while comprehension models require information about which variation can be present in the acoustic input and which variables predict this variation. Many studies have investigated reduction of the phoneme /t/ (e.g., [2]). Nevertheless, many questions about the conditions favoring /t/-reduction are still open. The present corpus study examined to what extent the degree of reduction of word-final /t/ in Dutch past-participles is affected by the predictability of the word itself and of the preceding words and by the word's morphological decomposability.

The predictability of a word is reflected by its frequency of occurrence, which affects its degree of reduction: Words with higher frequencies tend to be more reduced than those with lower frequencies [3]. Another measure of a word's predictability is its likelihood given its neighboring words. Content words seem not to be influenced by their predictability given the directly preceding word, but by their predictability given the immediately following word. If a word is more predictable given the following word, it tends to be more reduced [2, 3].

The effect of a word's predictability given its following word has been explained as a speech planning effect [3]. Given two subsequent words (word₁ and word₂), word₂ is (at least partly) planned and processed during the articulation of word₁. If word₂ is highly predictable, it can be planned and pro-

cessed more easily. Consequently, word₁ can be produced more quickly, which may result in more reduction. This planning explanation also accounts for why there is no effect of the word's predictability given the preceding word: While articulating word₂, the articulation and planning of word₁ is already done, and can therefore not influence word₂.

If the explanation based on planning is correct, words that follow a highly predictable word pair may also be more reduced than words that follow less predictable pairs. As the planning of predictable word pairs is relatively easy, speakers can start planning the following words more quickly, and therefore do not have to slow at the word directly following the word pair to buy time for the (further) planning of the following words. Since the role of speech planning is probably greater for spontaneous than for read speech, we expect predictability of words to condition speech reduction more in more spontaneous speech.

A word's morphological decomposability is also assumed to play a role in the production of reduced speech. Decomposability indicates the ease with which a morphologically complex word can be parsed into its morphemes. Words that are less decomposable are more likely to be retrieved from the mental lexicon as single words, instead of computed from their parts. The affixes of these words may therefore be regarded as parts of their stems and expected to be more reduced than affixes in complex words that can be decomposed more easily. Morphologically complex words are claimed to be more easily decomposable the higher the frequency of the base relative to the frequency of the word itself [4]. For example, the English adverb *softly* is less frequent than its base *soft*, whereas *swiftly* is more frequent than *swift*, and accordingly, [4] found that /t/ is less likely to be reduced in *softly* than in *swiftly*.

The words of [4]'s study were English adverbs, which are the only inflectional variants of their adjective bases. For example, the only inflections of the base *swift* are *swift* itself and *swiftly*. In the case of Dutch verbs, the number of inflections is considerably higher. The verb *wensen* 'to wish', for instance, has the base *wens* 'wish' and the inflections {*wens, wenst, wensen, wenste, wensten, wensend, wensende, gewenst*} '{wish, wishes, wish, wished, wished, wishing, wishing, wished}'. The question arises which relative frequency predicts decomposability better: a relative frequency in which the frequency of the word is related to the frequency of the base in isolation, or to the frequency of all inflectional variants containing the base. [2] studied the effect of the first relative frequency and found, surprisingly, that greater decomposability increased reduction. In our study, we considered both relative frequency measures.

If these relative frequency measures affect degree of reduction because they reflect decomposability, we expect them to

play a larger role for regular than for irregular verbs. Irregular verbs are always retrieved as units from the mental lexicon and therefore the relative frequencies cannot reflect decomposability for these verbs.

The current study explores pronunciation variation in Dutch past-participles, which always occur with an auxiliary verb (e.g., *hebben* 'to have'). They typically consist of a stem, the prefix *ge*, and the suffix *en* (for irregular verbs; e.g., *ge+roep+en* 'called') or *t* (for both regular and irregular verbs; e.g., *ge+wens+t* 'wished'). The prefix *ge*, however, is not always part of a Dutch past-participle: If a verb starts with an unstressed prefix, the prefix *ge* is not present (e.g., *ver+maak+t* 'entertained').

This corpus study investigates whether the presence of word-final /t/ in Dutch past-participles and its duration is affected by predictability of the past-participle itself and of the preceding words, and by two measures of morphological decomposability. We investigate possible interactions of these predictors with speech register and the verb's regularity.

2. Method

2.1. Materials

This study focuses on the final /t/ of Dutch past-participles starting with *ge*, *be*, or *ver*. We studied all tokens of these participles occurring in two corpora containing three different speech registers: Spontaneous conversational speech was obtained from the Ernestus Corpus of Spontaneous Dutch (ECSD) [1], while semi-formal interview and formal read speech were extracted from the *Corpus Gesproken Nederlands* (CGN; Spoken Dutch Corpus) [5]. We removed all tokens from the dataset that contained any background noise, including other speakers' speech. The final dataset consisted of 1369 word tokens, representing 380 word types (see Table 1).

	Total	Regular	Irregular
Total	1369 (380)	967 (367)	402 (13)
ECSD$_{spontaneous}$	498 (148)	320 (137)	178 (11)
CGN$_{interviews}$	363 (141)	245 (132)	118 (9)
CGN$_{read\ speech}$	508 (238)	402 (225)	106 (13)

Table 1: *The number of word tokens and word types (in parentheses) in the final dataset, split for the three speech registers and the verb's regularity.*

We determined the presence versus absence of /t/ and its duration on the basis of automatically generated broad phonetic transcriptions [6]. The automatic speech recognition (ASR) system HTK [7] retrieved for each word in the orthographic transcriptions its full form and possible pronunciation variants, and selected the variant that best matched the speech signal on the basis of 37 32-Gaussian tri-state monophone acoustic models [8]. These acoustic phone models had been trained on the read speech component of the CGN. Because the models consisted of at least three emitting states and the frame shift was 5 ms, phones were assigned durations of at least 15 ms (even if they were actually shorter).

To test the reliability of these transcriptions, two trained transcribers decided whether or not schwa is present in 230 target tokens from the ECSD's spontaneous speech. The validation was based on schwa, since agreements for schwa are generally lower than for other segments [6]. Results showed that

the agreement between the two transcribers on the presence of schwa (81.7%) was very similar to those between each transcriber and the ASR system (78.3% and 74.8%). Furthermore, the mean differences in the duration of schwa were small, that is, 6.8, 12.4, and 6.3 ms.

2.2. Variables

We tested three measures of word predictability. The first is the word's frequency of occurrence, which was log-transformed. The second and third measures are contextual predictabilities: We assessed the mutual information (MI) between the two words preceding the target word and between the target word and the following word. We calculated these measures with Formula 1, where X and Y denote the two words of a word pair. All frequencies used in this study were taken from the CGN.

$$MI\,(X,Y) = log_2(\frac{Freq(XY)}{Freq(X) \cdot Freq(Y)}) \qquad (1)$$

The second type of independent variable reflects the word's morphological decomposability. We defined base-word relative frequency as the frequency of the base in isolation divided by the frequency of the word. Additionally, we computed lemma-word relative frequency as the sum of the frequencies of all verb forms with the base divided by the frequency of the word. Both relative frequencies were log-transformed. For example, for the base-word relative frequency of the past-participle *gewenst* 'wished', we divided the frequency of its base *wens* 'wish' by the frequency of *gewenst*, and for the lemma-word relative frequency, we divided the sum of the frequencies of *wens* 'wish', *wenst* 'wishes', *wensen* 'wish', *wenste* 'wished', *wensten* 'wished', *wensend* 'wishing', *wensende* 'wishing', and *gewenst* 'wished' by the frequency of *gewenst*.

To investigate possible interactions with speech register and regularity, one predictor distinguised between the spontaneous ECSD, the CGN$_{interviews}$, and the CGN$_{read\ speech}$, while another predictor indicated the regularity of the verb (regular versus irregular). Finally, to reduce the variance in the data, we added a control variable indicating whether the word was in utterance-final position, and a control variable indicating the speech rate of the continuous chunk of speech containing the target word, which was defined as the number of syllables in the full word forms of that chunk produced per second.

3. Results

We investigated the presence versus absence and duration of /t/ by means of mixed effects regression models with contrast coding. All models contained the crossed random effects word type and speaker. For both dependent variables, we first fitted a model with utterance finality, speech rate, regularity, and speech register. Subsequently, we added each probabilistic variable to the model using that subset of the data for which the variable could be calculated. In all models, we only kept those predictors and interactions that were significant.

3.1. /t/ presence

The model ($N = 1369$) for the presence of /t/ showed significant effects of speech register ($F = 16.72, p < .0001$), speech rate ($\beta = -0.17, F = 12.89, p < .0001$), and regularity ($\beta = 0.75, F = 4.59, p < .05$). Participle-final /t/ is more often present in CGN$_{read\ speech}$ compared to CGN$_{interviews}$ and ECSD, if it is produced at a lower speechrate, or if it belongs to a regular verb.
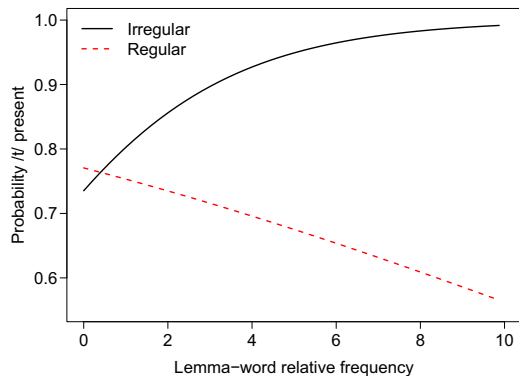
Figure 1: *The effect of lemma-word relative frequency on the presence of /t/ split for regular and irregular past-participles, as estimated by the regression model.*
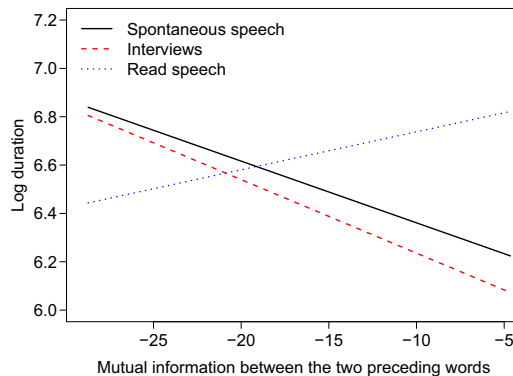

Figure 2: *The effect of mutual information between the two preceding words on the duration of /t/ split for the three corpora, as estimated by the regression model.*

Next we added the predictability variables to the model. Because MI between the two preceding words correlated with speech register, we orthogonalized these two predictors by replacing MI by the residuals of a linear regression model predicting MI as a function of speech register. The residuals of MI showed a significant main effect on the presence of /t/ ($N = 1085; \beta = -0.08, F = 6.63, p < .05$), suggesting that /t/ is more likely to be absent if the MI between the two preceding words is higher. In this statistical model, speech register and regularity were still significant, but speech rate no longer was. We found no effects of word frequency and MI between the target and following word.

We then added the two morphological frequency measures to the model. As both were highly correlated with regularity, we replaced these variables by the residuals of the models predicting these frequencies as a function of regularity. Only the residuals of lemma-word relative frequency appeared to predict the presence of /t/, in interaction with regularity ($N = 1078$; regularity: $\beta = 0.19, F = 4.30, p < .05$; relative frequency: $\beta = 0.38, F = 0.31, p < .5$; interaction: $\beta = -0.48, F = 6.71, p < .05$). Figure 1 shows how the effect of this relative frequency differs for regular and irregular verbs. Further analysis suggests that the effect is only present for irregular verbs.

3.2. /t/ duration

We then analyzed the durations of the 1034 present /t/s. In order to create a more normal distribution, we log-transformed the durations. Moreover, segments longer than 2.5 times the residual standard errors of the values predicted by the statistical models were considered outliers and removed.

The model yielded significant effects of speech rate ($\beta = -0.11, F = 20.14, p < .0001$) and regularity ($\beta = 0.16, F = 7.45, p < .05$). Word-final /t/ tends to be longer when produced at lower speech rates or if it belongs to a regular verb. Besides, this model contained significant interactions between utterance finality and speech rate (utterance finality: $\beta = -0.26, F = 196.87, p < .0001$; interaction: $\beta = 0.08, F = 5.36, p < .05$) and between utterance finality and speech register (speech register: $F = 27.99, p < .0001$; interaction: $F = 17.36, p < .0001$), indicating that /t/ tends to be longer in utterance final position, but only at lower speech rates or in the $\text{CGN}_{read\ speech}$.

Subsequently, each predictability variable was separately

added to the model. As word frequency correlated with regularity, this variable was replaced by the residuals from a model predicting word frequency as a function of regularity. The model for t-duration showed a significant main effect of these residuals ($\beta = -0.02, F = 12.01, p < .0001$), suggesting that /t/s tend to be shorter if word frequency is higher. Furthermore, we found an interaction between speech register and MI between the two preceding words ($N = 811$; speech register: $F = 53.35, p < .0001$; MI: $\beta = -0.03, F = 2.45, p > .1$; interaction: $F = 6.38, p < .0001$). As shown in Figure 2, word-final /t/ tends to be shorter if the preceding two words are more predictable, but only in $\text{CGN}_{interviews}$ and ECSD. MI between the target and following word was not significant.

Finally, due to correlations with regularity, both morphological frequency measures had to be orthogonalized before they could be added to the model. The residuals of lemma-word relative frequency interacted with MI between the two preceding words and regularity (see Table 2). Because the previous analyses showed an interaction between MI and speech register (see Figure 2), we also tested this three-way interaction for $\text{CGN}_{interviews}$ and ECSD only, which resulted in stronger effects than in the previous model (see Table 2). Final /t/ tends to be longer above all in irregular verbs, the higher the lemma frequency relative to the word frequency, especially if the MI between the two preceding words is lower (see Figure 3).

Factor	All speech registers			Spontaneous speech		
	β	F	$p <$	β	F	$p <$
Speech (Sp.) rate	-0.14	19.16	.0001	-0.12	7.42	.05
Utterance Finality	-0.50	107.27	.0001	-0.58	10.23	.0005
Speech register	—	52.06	.0001	—	—	—
Regularity (Regu.)	-0.55	6.87	.05	-0.84	3.90	.05
Word frequency	-0.02	5.12	.05	-0.02	3.12	.1
MI	0.007	3.27	.1	0.01	2.93	.1
Rel.freq.	-0.69	2.83	.1	-0.95	2.63	.2
Utt.fin. × Sp. rate	0.12	9.23	.0001	0.14	6.66	.05
Utt.fin. × Sp. register	—	14.39	.0001	—	—	—
MI × Sp. register	—	7.13	.0001	—	—	—
MI × Regu.	-0.03	1.50	.5	-0.05	3.37	.1
Rel.freq. × Regu.	0.70	1.36	.5	0.71	1.70	.2
Rel.freq. × MI	-0.04	2.84	.1	-0.05	9.66	.005
Rel.freq. × MI × Regu.	0.04	7.21	.05	-0.04	3.15	.1

Table 2: *Results of the /t/-duration model with MI between the two preceding words and lemma-word relative frequency.*

Mutual information	
—	−28.7
---	−22.6
···	−21.2
-·-	−19.4
--	−4.9

(a) Regular verbs

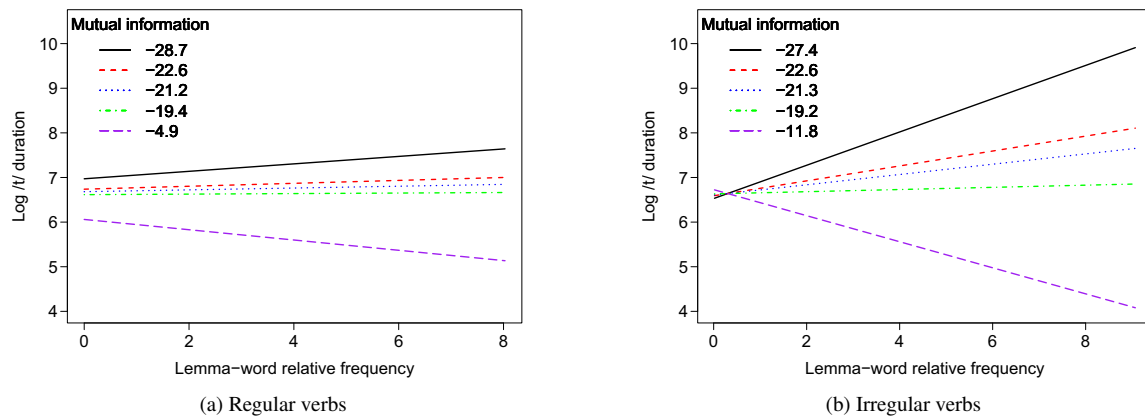| Mutual information	
—	−27.4
---	−22.6
···	−21.3
-·-	−19.2
--	−11.8

(b) Irregular verbs

Figure 3: *The interaction of lemma-word relative frequency and mutual information between the two preceding words on the duration of /t/ in the two more spontaneous corpora split for (a) regular and (b) irregular verbs, as estimated by the regression model.*

4. Discussion and conclusions

The present corpus study investigated which variables affect the presence and duration of the Dutch participle-final segment /t/. We analysed /t/ in 1369 past-participles, which were extracted from corpora containing three different speech styles.

The first predictor that we focused on is the predictability of the word itself. /t/ appeared shorter in words of a higher frequency. This result confirms the general finding that degree of reduction increases with word predictability (e.g., [3]).

Second, we investigated the role of contextual predictability. Although earlier research showed effects on degree of reduction of the predictability of the word given the following word [2, 3], we found no such effect. In contrast, we found an effect of the predictability of the two preceding words. The number of tokens on which these analyses were based perhaps explains the absence of an effect of the following word, since the analyses of the MI with the following word were based on fewer tokens (i.e., 710 for presence and 521 for duration for MI with the following word versus 1085 for presence and 811 for duration for MI of the two preceding words).

A higher predictability of the two preceding words leads to final /t/s that are shorter and more often absent in spontaneous speech. This predictability effect is probably due to the mechanisms of speech planning. If two words are highly predictable and therefore easy to plan, planning of the subsequent words can start early. If so, speakers do not have to slow down at the following word in order to buy time for the planning of the further following words, and can therefore produce the word more quickly and with a higher degree of reduction. Our finding that the effect of predictability is greater for spontaneous speech supports an explanation based on planning.

For those tokens in the spontaneous speech corpora that followed low predictability word pairs, we found an effect of lemma-word relative frequency: /t/ tends to be more reduced the lower the relative frequency, especially in irregular words. Since we defined this relative frequency to reflect morphological decomposability, we expected it to affect primarily regular verbs, contrary to our findings. We note that the lemma-word relative frequency also represents a past-participle's likelihood within the verb's paradigm (e.g., [9]). If the lemma frequency is higher relative to the word frequency, the word has to be chosen among highly likely alternatives within the paradigm and thus has a lower likelihood. Consequently, a speaker needs more

time for planning, and produces a less reduced variant.

Finally, we also found a main effect of the verb's regularity: /t/ is more likely to be present and longer if the past-participle belongs to a regular verb. Probably, /t/ is more reduced in irregular past-participles because these forms are more likely to be retrieved from the lexicon, whereas the regular forms are more likely to be composed from their parts.

In conclusion, we showed that reduction of Dutch participle-final /t/ is conditioned by the predictability of the two preceding words and by the frequency of a word relative to its lemma frequency. Both effects may result from speech planning.

5. References

[1] 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.

[2] B. Schuppler, W. van Dommelen, J. Koreman, and M. Ernestus, "How linguistic and probabilistic properties of a word affect the realization of its final /t/: Studies at the segmental and sub-segmental level," submitted.

[3] A. Bell, J. Brenier, M. Gregory, C. Girand, and D. Jurafsky, "Predictability effects on durations of content and function words in conversational English," *JML*, vol. 60, pp. 92–111, 2009.

[4] J. Hay, "Causes and consequences of word structure," Ph.D. dissertation, Northwestern University, 2000.

[5] N. Oostdijk, "The design of the Spoken Dutch Corpus," in *New Frontiers of Corpus Research*, P. Peters, P. Collins, and A. Smith, Eds., 2002, pp. 105–112.

[6] 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.

[7] S. Young, G. Evermann, T. Hain, D. Kershaw, G. Moore, J. Odell, D. Ollason, D. Povey, V. Valtchev, and P. Woodland, *The HTK Book 3.2*. Cambridge: Entropic, 2002.

[8] A. Hämäläinen, M. Gubian, L. ten Bosch, and L. Boves, "Analysis of acoustic reduction using spectral similarity measures." *JASA*, vol. 126, no. 6, pp. 3227–3235, 2009.

[9] F. Moscoso del Prado Martín, A. Kostić, and R. Baayen, "Putting the bits together: An information theoretical perspective on morphological processing," *Cognition*, vol. 94, pp. 1–18, 2004.