GHOST SEGMENTS IN THE FLEMISH TUSSENTAAL

Mishko Bozhinoski & Paul Boersma

University of Amsterdam

ABSTRACT

The Flemish informal standard (*Tussentaal*) exhibits a word-final elision process whereby words such as ‘wat’ and ‘met’ are realised without their final plosive: [βɑ] and [mɛ]. This elided segment triggers devoicing in following fricatives. This process is not derivable from phonological context alone, and the current literature has not explained its distribution. This study therefore set off to (1) document this elision process, and (2) account for its distribution.

Using Python and the corpus engine OpenSonar, audio tokens of the fifty most common Dutch words containing final coronal plosives were scraped from the Corpus Gesproken Nederlands. Joint results of an algorithmic and a manual analysis reveal that this elision occurs in the present tense marker /-t/ and most function words. All exceptions found have historically undergone word-final schwa deletion, suggesting that final coronal plosive elision may be older. Evidence from 13th century texts supports this claim.

Keywords: Franconian dialectology, ghost segments, sound change, corpus linguistics, sandhi

1. INTRODUCTION

A peculiar and often ridiculed feature of colloquial Standard Flemish (henceforth *Tussentaal*) is the elision of expected final coronal plosives in words such as *wat*, *niet* and *met*, whereby they are regularly realized without their final plosive: [βɑ], [ni] and [mɛ]. The result of such elision frequently results in word-final lax vowels and monomoraic phonological words, which are impermissible in Standard Dutch [1, p.26]. It also results in a devoicing of following voiced fricatives, as if this dropped *-t* was still there [2]:

Ghost *-t*	Pronounced *-t*	No *-t*
<met zand> [mɛ sɑnt] ‘with sand’	<nat zand> [nɑt sɑnt] ‘wet sand’	<blauw zand> [blɑu zɑnt] ‘blue sand’

Table 1: Ghost *-t* and its effect on adjacent consonant

Additionally, in cases of hiatus across word boundaries, whereby the first word ends in an otherwise elided coronal plosive, a [d] surfaces in between the vowels [3]:

Ghost *-t*	Pronounced *-t*
<wat een> [βɑ**d**ə(n)] ‘what a...’	<vat een> [vɑ**t**ən] ‘catch a...’

Table 2: Liaison of ghost *-t*

The distribution of this ghost segment, henceforth transcribed as /^{t}/ or [^{t}][1], seems on the surface to be phonologically unpredictable from a synchronic perspective:

No *-t*	Ghost *-t*	Pronounced *-t*
[mu] ‘tired’	[mu^t] ‘must’	[mut] ‘courage’
[zɛː] ‘said’	[zɛː^t] ‘are’	[tɛːt] ‘time’
[ɣaː] ‘(I) go’	[ɣaː^t] ‘goes’	[laːt] ‘let’

Table 3: Unpredictable distribution of ghost *-t*

While the existence of this phenomenon has been documented in the literature [2, 3, 4], accounts for its distribution, either synchronically or diachronically, are cursory at best. The present research aims to fill this gap through an exploratory corpus study, with the following research questions:

1. What is the synchronic distribution of /^{t}/?
2. Is there a diachronic account for /^{t}/?

2. METHOD

2.1. Corpus

The present study is based on corpus data provided by the Corpus Gesproken Nederlands [5], henceforth CGN, consisting of speech collected in various contexts, ranging from spontaneous face-to-face conversations to TV and radio interviews to religious services and speeches. The data pool used in the present study is only that of face-to-face conversations in order to ensure an informal colloquial register as opposed to one in which Standard Dutch may be more appropriate.

The corpus is accessed using OpenSoNaR [6], a corpus search engine used alongside both CGN as well as a written corpus that is not relevant to the current research. CGN is annotated for phonetic and orthographic information in the *TextGrid* format, chiefly employed by *Praat* [7], in which each word is encapsulated within an interval associated with plain text. This facilitates the query of data based on orthographic (or phonetic) information.

2.2. Conditions and coding

Individual tokens examined by this study consisted of word pairs in which the first word ends (orthographically) in *-t* or *-d*, and the word right after begins with an underlyingly voiced fricative. The word pairs were selected on the basis of two conditions: the first word, which ends in *-t* or *-d*, and the fricative that the following word begins with.

2.2.1. Word conditions

The word condition is filled by the fifty most frequent words (orthographically) ending in *-t* or *-d*, as per *A Frequency Dictionary of Dutch* [8]. To limit the scope of the current study, nouns containing a maximum of one full (non-schwa) vowel are to be used. This restriction also minimizes potential stress-related effects. Thus, words such as /minyt/ ‘minute’, /rezʏltaːt/ ‘result’ and /ɪnhɑud/ ‘content’ are not allowed, while /ɣəbid/ ‘area’ and /ɣəlœyd/ ‘sound’ are. No such restriction was applied to non-nouns. The words were divided into four categories: verbs, adjectives/adverbs, nouns and other (prepositions, determiners and conjunctions).

2.2.2. Fricative conditions

This study concerns itself with the phonetic realization of the ghost segment /^{t}/, and its realization when adjacent to a fricative. Namely, what is relevant is (a) whether the segment is realized and (b) whether the following underlyingly voiced fricative in question is voiceless. As Dutch contains three native voiced fricatives, namely /v/, /z/ and /ɣ/, each word undergoes three batches of data collection: one in which it is followed by /v/, one by /z/ and one by /ɣ/.

2.3. Data collection

To limit the time required to manually process the data, each word–fricative pair was limited to ten tokens. This means that the ideal number of tokens collected is 1500 (3 fricatives $\cdot$ 50 words $\cdot$ 10 tokens). The collection of the relevant data was carried out using a process by which web-hosted data are extracted and saved locally, a process known as web-scraping. This was carried out using a script created for this study written in the scripting language *Python*, and chiefly using the web-automation library *splinter* [9] and the audio processing library *pydub* [10]. The program’s workflow can be summarized as follows:

1. Send a corpus query to OpenSoNaR per element of the cartesian product $W \times F$, where W consists of all target words and F consists of the three voiced fricatives.
2. Discard all but one result per audio file. This is to guarantee that each audio clip comes from a different speaker.
3. Retrieve the first 10 results (where available) of each query, as well as the TextGrid boundaries of the beginning and end of the word pair.
4. Download the audio file of each query and cut it in accordance with the boundaries acquired in step 2.

2.4. Analysis

Using *Praat* [7], an acoustic analysis was performed per word pair with respect to two variables: the presence of an alveolar plosive in the first word, and the voicing of the initial fricative in the second word. Due to the unreliable quality of the audio, plosive presence was analyzed using a hybrid automated–manual approach.

2.4.1. Plosive presence

The detection of the presence of a plosive was conducted using *Praat*. Firstly, the data were manually annotated so that the whole span of the fricative was tightly encapsulated within a TextGrid interval. These intervals were additionally marked with *1* or *0*, representing respectively the perceived presence or absence of a plosive before them. This binary marking is henceforth referred to as T. Then, a script was written to execute the following tasks:

1. For each fricative interval, extract a clip running from 100 ms to 0 ms before the beginning of the fricative.
2. Cut off all frequencies below 500 Hz. This was done due to the presence of unwanted noise, presumed to be a result of the relatively low recording quality.
3. Take the difference between the highest and lowest points of intensity ΔI within the clip.

It was expected that the presence of a plosive would result in a high ΔI, and inversely that the lack of a plosive would show a low ΔI.

2.4.2. *Fricative voicing*

Cues for fricative voicing in Dutch have been documented to include periodicity, duration and to a lesser degree intensity [11]. The cue used to identify fricative voicing within this study was that of duration, as it is least dependent on the quality of the data. First, a control study was conducted to establish whether a duration cue for voicing is reliably found in the data. Data were scraped using the same script as in §2.3, with the word conditions all ending in a vowel or sonorant, and the fricative conditions containing both underlyingly voiced and voiceless fricatives. The extracted audio files were then annotated in the TextGrid format, such that the span of each target fricative corresponds to the span of an interval. The annotation was carried out manually. There was no way to guarantee unbiased annotation within the temporal constraints of the study.

Using a Praat script, the duration of each interval and the consonant it corresponds to was extracted and saved onto a plain text file, such that each token interval is represented by its duration d and a binary voicing parameter v, such that <z>, <v> and <g> are represented by '+v', and <s>, <f> and <ch> are represented by '-v'. A total of 66 token fricatives were scraped and annotated. Finally, with the help of R [12], the data were fit onto a linear model of $\ln d$ as a function of the binary voicing parameter v. This yielded that the average duration of voiceless fricatives was longer than that of voiced fricatives (v-coef. $= 0.752$; conf. interval $= 0.592 \ldots 0.911$; $t(64) = 9.452$; $p = 9.4 \cdot 10^{-17}$).

2.5. Data processing

The majority of the data was formatted and partially processed using *Python* and *NumPy* [13]. Per word, the following values were calculated:

- RT = the ratio of tokens with deleted [t] over all tokens of a given word
- n = the number of tokens per word
- ΔI_m = the mean of ΔI over n

The values of T, ΔI, and d per token were also saved to a text file to be fit onto several linear models using *R*. This is the per-token analysis.

3. RESULTS

3.1. Per token

In total, 848 tokens from 39 words were collected. Data were not collected for any pair for which fewer than three tokens were available.

3.1.1. *Computation of ΔI versus manual marking*

To check whether the manual annotation T – the manual binary annotation of 1 (final plosive present) and 0 (final plosive absent), agrees with the automated measurement ΔI – the difference between the highest and lowest point of intensity of a token, they were fitted onto a linear model of ΔI as a function of T. The value of ΔI was an estimated 5.21 dB higher (as expected in the case of a present plosive) when $T = 1$ than when $T = 0$. This effect was statistically significant (95% conf. interval $=$ 4.01 dB ... 6.41 dB; $t(363) = 8.66$; $p = 1.6 \cdot 10^{-16}$).

For §3.1.2, in which a binary decision of voicing is needed, we use the manual voiced/voiceless marking rather than the continuous ΔI. Using ΔI would require a conversion to a binary value that the reliability of ΔI cannot afford.

3.1.2. *Voicing of fricative after ghost t*

To determine the ghost *-t*'s effect on fricative voicing, the durations of fricatives that follow a phonetically absent *-t* was compared to those of known voiced and voiceless fricatives, as established in §2.4.2. A data set was constructed consisting of all known voiced (marked 'v') and voiceless (marked 'u') tokens alongside all tokens of fricatives directly following a ghost *-t* (marked 'a'). The values of d (fricative duration) and V (ternary marking of 'a', 'u' and 'v') were fit onto a linear model of $\ln d$ as a function of V, with an orthogonal contrast measuring the distance of $\ln d$ when $V =$ 'a' from the average of $\ln d$ of fricatives following known voiceless and voiced plosives. The value of $\ln d$ when $V =$ 'a' was shown to be significantly greater than the average $\ln d$ of voiced vs voiceless fricatives (est. distance from mean $= 0.267$; 95% conf. interval 0.171 ... 0.363; $t(187) = 5.46$; $p = 1.5 \cdot 10^{-7}$). We conclude that fricatives that follow a ghost *-t* are on average closer to postvocalic voiceless than to postvocalic voiced fricatives (any remaining difference with the postvocalic voiced fricatives can be due to lexicalized cases of variation, which there is no room to discuss here).

3.2. Per word

The per-word analysis is taken as relevant only if the word in question has ten or more tokens representing it. This adds up to thirty words. Below is a summary of those thirty words depending on their value of *RT* – the ratio of tokens which the manual data marks as having a deleted final [t]. In bold are function words as well as conjugational affixes.

- $RT \geq 0.5$: ***dat***, ***doet***, ***niet***, ***omdat***
- $0.2 \leq RT < 0.5$: ***gaat***, *goed*, ***met***, *staat*, ***wat***
- $RT < 0.2$: ***moet***, *weet*, ***had***, *zat*, *altijd*, *bed*, *groot*, *kwaad*, *kwijt*, *laat*, ***net***, *nooit*, *ooit*, *stad*, *straat*, *tijd*, ***tot***, ***uit***, *zet*, *zit*

4. DISCUSSION

4.1. Effect of ghost -*t* on following fricatives

The results of §3.1.2. show us that the ghost -*t* have a duration (which we take to be a voicing cue) that is significantly more similar to voiceless fricatives than to voiced fricatives, suggesting that the ghost -*t* has a devoicing effect on following fricatives. This is in-line with the observations of Camerman [3] and De Schutter [2].

4.2. Distribution of ghost -*t*

The distribution of *t*-deletion seems to be limited to function words, with the potential exception of *goed* 'good, well' – which is a high frequency content word, although its vague meaning and modal usage may qualify it as a function word. While *t*-deletion seems to only happen in function words, being a function word alone is not enough to predict *t*-deletion. Namely, the words, *had* 'had', *tot* 'to, until', *uit* 'from, out' and *net* 'just' do not show any *t*-deletion in the dataset.

The lack of deletion in *net* can be explained by considering that it may be a loan from the standard language. This is supported by its lack in dialect dictionaries such as the Antwerps Woordenboek [14] (in Antwerp Brabantian, for instance, we see the word *zjust*, from French *juste*, covering the semantic range of *net*).

A possible explanation for the lack of *t*-deletion in *had*, *tot* and *uit* is the presence of a schwa following the final obstruent in their older forms. Namely, *had*, *tot* and *uit* appear as *had(d)e*, *tote* and *uite* in their Middle Dutch forms. This suggests that this *t*-deletion may have occurred before word-final schwa was deleted, and indeed there is evidence of this in the corpora. See the below examples sourced from Corpus Gysseling Vol I [15]:

(1) (0295, Oudenbiezen, 1280 april 22–28)
Des lants hef si geloft te mestene no hore magt, en **dassi** (< dat zij) mest do salsi VI joer derno 15ane hebben. Die selue Mente heft XV ruden bamt ombe VII vaet roghen Tungers ter seluer getonst **dassi** (< dat zij) dlant heft.

(2) (0729, Brugge, 12 juli 1287)
en*de* in ke*n*nesse van desen sticke*n* dat sie vast bliue*n* en*de* ghestade hebbe*n* vorseid*e* redenars dese lett*re*n gheseghelt **me** (< met) onsen zeghell*en* hutha*n*ghe*n*de dit was ghedaen sat*er*daghes vore alf hoimaent.

5. CONCLUSION

Pertaining to the topic of the current study, one may take away the following.

Firstly, the ghost segment $/^t/$ is limited to function words that historically end in /t/ or /d/, represented orthographically by <t> and <d> respectively. Function words that synchronically end in /t/ or /d/ robustly realized as [t] seem to comprise words that historically end in a schwa instead of a plosive, suggesting that this phonetic weakening of final coronal plosives chronologically precedes word-final schwa-deletion. This is supported by historical corpus data dating from the 13th century.

Secondly, the ghost -*t*, despite its phonetic absence, causes devoicing in fricatives that directly follow it.

Finally, there are many gaps in our knowledge pertaining the ghost segment $/^t/$, gaps which hide critical insight into both synchronic and diachronic Franconian linguistics. These gaps may viably be explored using computational methods, but our current spoken language corpora are thusfar insufficient for these purposes. We plan to address this in a future project, which will recreate this study with a corpus at least three orders of magnitude the size of the one of the present study.

6. REFERENCES

[1] G. E. Booij, *The phonology of Dutch*, 1st ed., ser. The phonology of the world's languages. Amsterdam: Oxford Univ. Press, 1999.

[2] G. De Schutter, "Het Antwerps. een schets van zijn evolutie tussen 1898 en 1998," *Honderd jaar stadstaal.*, pp. 301–315, 1999.

[3] F. Camerman, *Antwerps schrijven : spelling en grammatica van het 21e-eeuwse Antwerps*. Antwerpen: de Vries-Brouwers, 2007.

[4] K. R. Haeseryn, K. Romijn, G. Geerts, J. de Rooij, and M. van den Toorn, "1.5.2.6 Deletie van t," 2020, retreived January 03, 2023. [Online]. Available: https://e-ans.ivdnt.org/topics/pid/topic-16007638353447182

[5] L. van Eerten, "Corpus gesproken Nederlands," *Nederlandse taalkunde (Groningen)*, vol. 12, no. 3, pp. 193–214, 2007.

[6] N. Oostdijk, M. Reynaert, V. Hoste, and I. Schuurman, "The construction of a 500-million-word reference corpus of contemporary written Dutch," in *Essential speech and language technology for Dutch*. Springer, Berlin, Heidelberg, 2013, pp. 219–247.

[7] Boersma, Paul and Weenink, David, "Praat: doing phonetics by computer," 2022. [Online]. Available: http:\\praat.org

[8] C. Tiberius and T. Schoonheim, *A frequency dictionary of Dutch: Core vocabulary for learners*. Routledge, 2013.

[9] A. Medina, B. Barreto, D. Fernandes, D. S. de Andrade, D. Camata, R. Goncalves, E. Flores, F. M. Costa, F. Missi, F. Ribeiro, L. Geron, B. Heynemann, H. L. Tavares, H. S. Antunes, J. Bochi, L. Lemos, M. de Oliveira, G. Rezende, R. Caricio, E. B. d. Luz, T. Azevedo, T. Al-Chueyr, and T. Santos, "splinter (version 0.17)," 2021. [Online]. Available: https://github.com/cobrateam/splinter

[10] J. Robert, M. Webbie *et al.*, "Pydub," 2018. [Online]. Available: http://pydub.com/

[11] de Jonge, Mirjam and Aligbeh, Isaura, "Duration as a voicing cue in Dutch," in *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne, 2019.

[12] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2021.

[13] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant, "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357–362, Sep. 2020.

[14] F. Camerman, "Antwerps.be," 2009. [Online]. Available: https://www.antwerps.be/

[15] M. Gysseling, "Corpus van Middelnederlandse teksten (tot en met het jaar 1300), mmv Willy Pijnenburg," *Reeks I: Ambtelijke bescheiden. 's-Gravenhage, Nijhoff*, 1977.

[1] The transcription $/^t/$ is not meant to imply that the ghost -*t* is underlyingly voiceless. It is a largely arbitrary choice motivated by its orthographical representation and its correspondence to historical /t/ in most cases.