

The Distribution of (Word Initial) Glottal Stop in Dutch

1. Introduction

Two kinds of vowel onset can be distinguished in Dutch, an abrupt and a more gradual one. The abrupt onset, also called "fast attack" or "glottal stop", is auditorily quite different from the vowel onset with "gradual attack" or "smooth onset".

A glottal stop can be defined as the loud and sudden start of a vowel. Immediately prior to the onset of a vowel, the vocal cords are adducted and kept in closed position for, say, 40 to 50 ms. During the closure phase, subglottal air pressure builds up rapidly. On releasing the closure, the vocal cords abruptly start vibrating for the production of the vowel, which results in a rapid increase of the vowel's intensity, especially in the second and third formants (cf. Malécot, 1975). Figure 1 shows an oscillogram of the utterance ... dat een [?]aantal ('... that a number'; [?] is our phonetic symbol for glottal stop), which was included in our speech database (see below).

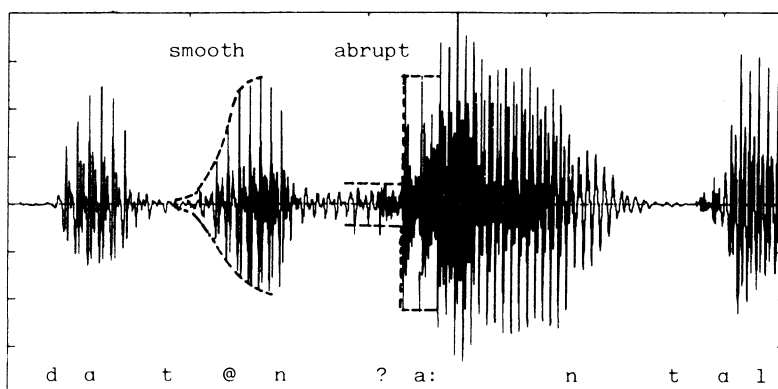


Figure 1: Oscillogram of the utterance dat een [?]aantal with a smooth vowel onset in een and an abrupt onset (glottal stop) in aantal. Notice the long and irregular glottal periods at the abrupt onset.

The word-initial schwa [@] (in een) is realized without a glottal stop; the amplitude increases relatively slowly and the glottal periods are regular from the first moment onwards. The word-initial vowel [a:] (aantal) is clearly realized with a glottal stop: the amplitude of the vowel increases suddenly and the first three glottal periods succeed at very irregular intervals.

Since in Dutch the glottal stop is not relevant on the phonological level (Booij, 1981:16), the phenomenon has received scant attention in the phonological literature. However, a number of phonologists have introduced the notion of "hiatus position", defined as any vowel at the beginning of a word or word-like morpheme, i.e. the sequence #V (Booij, 1981:103-110). In our view glottal stops can be inserted **only in hiatus position**, and nowhere else. When no glottal stop is inserted, a smooth transition from the pre-hiatus phoneme

into the hiatus vowel must be secured, either by resyllabifying a pre-hiatus consonant or by linking two vowels across the hiatus (cf. van Heuven & Hoos, 1991). Resyllabification/hiatus deletion and glottal stop insertion are therefore mutually exclusive choices.

Publications on other languages (e.g., Malécot, 1975), show that the distribution of (word-initial) glottal stop may be rule-governed. The present study was set up to shed more light on this issue for Dutch.

This study was motivated by our wish to improve the quality of the text-to-speech system for Dutch, which is being developed in the national research programme Analysis and Synthesis of Speech (ASSP). This programme is geared towards generating high quality speech synthesis while modeling the speaking behaviour of a single professional speaker PB (Philip Bloemendal, who is known as the former newsreader of the Dutch cinema news-bulletin). Our assumption is that intelligibility and naturalness of the synthetic speech can be improved by inserting glottal stops in the same positions where the human speaker produces them. We therefore analysed glottal stop distribution in a corpus of continuous prose read by PB in order to extract the optimal rule(s) for Dutch glottal stop insertion.

The structure of this paper is as follows. In §2 five factors that are potentially relevant to the glottal stop distribution are identified and discussed. §3 outlines the methodology of the research, and results are presented and discussed in §4. In §5, finally, a provisional rule schema will be given optimally covering the distribution of the glottal stop in Dutch.

2. Factors influencing the distribution of the glottal stop

We assume that the glottal stop distribution depends on two groups of factors. Firstly, physiological restrictions of the speech organs may lead to glottal stop insertion, such that the distribution of the glottal stop is partly based on considerations of speech comfort. Secondly, we propose that glottal stop insertion can be predicted in cases where the glottal stop may simplify word recognition for the listener, when the glottal stop can serve as an overt word boundary marker (Quené, 1989). In all, we shall consider five factors in this study that may influence glottal stop insertion; three of these serve the speaker's comfort, the remaining ones are motivated by potential ease of listening.

2.1. Factors motivated by ease of speaking

Prosodic pause preceding an initial vowel

Just before phonation resumes after a speech pause there is a speech initiation gesture: the soft palate is raised and the vocal cords are adducted. Because of the recent intake of air, subglottal air pressure will be much larger than oral air pressure. When no inhalation of air takes place during a speech pause, the vocal cords are closed tightly for a period of time between 200 and 500 ms, trapping what air remains in the lungs, so that considerable subglottal air pressure can build up. Given the significant difference in air pressure below and above the glottis after both types of pause, we anticipate a sudden and rather violent onset of vocal cord vibration at the beginning of a post-pausal vowel. If no pause precedes, the vowel onset will be smooth and well controlled.

The voicing feature of the preceding phoneme

Krech (1968) found a correlation between the voicing feature of the phoneme preceding an initial vowel and the occurrence of a glottal stop. Her data show that in German more glottal stops are found after voiceless than after voiced

speech sounds. Although we have difficulty understanding why this effect should apply, it seems a reasonable course of action to check whether the same regularity can be found in Dutch.

Prominence of syllables

Stressed syllables are articulated more precisely and energetically than unstressed syllables. It appears that in German (Krech, 1968) as well as in French (Malécot, 1975) glottal stops are more often inserted in a stressed syllable than in an unstressed syllable. Therefore, we included presence versus absence of stress on the syllable containing the initial vowel as a third factor in our study.

2.2. Factors motivated by ease of listening

Phonotactic restrictions on separating an onset

Phonotactic restrictions limit the possible combinations of speech sounds in a word or syllable. These restrictions potentially serve the listener in tracing (word) boundaries. When, for example, a Dutch listener hears the sequence [...rmdr...], he knows that this combination of phonemes is phonotactically not in order, and must contain a word boundary (#) in the middle: [...rm#dr...] (the segmentations ..r#mdr.. and ..rm#r.. do not lead to legal onset and offset consonant clusters).

Quené (1989) has shown that speakers tend to provide acoustic boundary markers only when other (for instance phonotactic) means facilitating boundary detection are relatively weak. Since the occurrence of the glottal stop in Dutch is restricted to the beginning of words (or at least morphemes), hearing a glottal stop is a sure sign that a new word (or morpheme) has just begun. In line with Quené's (1989) findings we predict that speakers will preferably realise a glottal stop in cases where word boundary ambiguity cannot be solved by other means (e.g., when phonotactic restrictions fail). For instance, the utterance bijt eer ('bite before' is normally pronounced as [beite:r]. The phonotactic restrictions of Dutch allow two segmentations: [beit#e:r] and [bei#te:r]. The speaker may assist the listener by inserting a glottal stop before the initial vowel [e:r] as a boundary marker, so that the listener will understand that bijt eer rather than bij teer ('near tar') is intended.

We shall differentiate between three word segmentation possibilities:

- (i) a sound sequence can be segmented only in one manner, as in, e.g., the sequence of er ('if there'): [ɔfer]. In Dutch a word cannot end in a short vowel [ɔ].
- (ii) a sound sequence can be segmented in more than one way, but no incorrect onset leads to an existing syllable, for instance: woorden uitspreken 'articulate words' is pronounced as [wo:rd@nytspre:k@n]; the [n] can be legally parsed as a syllable onset, but nuit is not an existing (word initial) syllable in Dutch.
- (iii) a sequence can be parsed in more than one way and the new onsets all lead to an existing initial syllable in Dutch, e.g., zijn eigen ('his own') [zein#eiɹ@n] can also be parsed as [zei#neiɹ@n] ('they tend').

We assume that sound sequences that can be segmented plausibly in more than one way, will cause more confusion for the listener. If the speaker inserts glottal stops on purpose, so as to facilitate word segmentation and recognition as supposed by Quené (1989), more glottal stops are expected for boundaries of type (iii) than for (ii) than for (i).

Word length

The length of a word, too, potentially influences glottal stop insertion. Polysyllabic words can typically be recognized long before the listener has heard the whole word. In order to recognize the word elephant the listener need only hear the initial sound sequence eleph, since there are simply no other words in the (English) lexicon that begin with this sequence. Short (especially monosyllabic) words such as man should, in order to be recognized, be heard in their entirety. Moreover, the listener has to make certain that the speaker was not actually pronouncing the first syllable of a longer word, e.g., manual. As a result, establishing word boundaries in a sequence of short words is more difficult than in long words (Nootboom, 1985; Scharpf, 1987; Scharpf & van Heuven, 1988). Assuming that the speaker aims to assist his listener, a glottal stop (marking the onset of a new word) will be inserted sooner after a monosyllabic word (or in between two monosyllabic words) than after a longer word (or in between long words).

3. Method

The corpus we examined contained approximately 1,500 words of text, divided into a number of short coherent pieces of prose, typically taken from newspaper editorials or magazine columns. These texts are part of a larger speech database that is currently under construction. All the positions in the text where a glottal stop could potentially be realised (cf. §1), i.e., all the hiatus positions, were automatically marked by extending and executing a rule-based letter-to-sound conversion routine (Berendsen, Langeweg & van Leeuwen, 1986) using the phonological rule compiler Toolip (van Leeuwen, 1989). In this way 424 hiatus positions were identified.

The first author then listened to the corresponding speech materials, and indicated for each hiatus position whether or not a glottal stop had actually been realised (dependent variable). Only in the (few) cases where the first author felt uncertain, a joint decision was taken by both authors together, after visual inspection of the waveform.

Next, each hiatus position in the corpus was scored in terms of the five factors (independent variables) identified in §2. We shall now briefly explain how the relevant information was collected.

1. In order to establish whether a hiatus position was preceded by a **speech pause** we examined the relevant portions of the waveform. An interruption of a fluent utterance had to be longer in duration than 200 ms in order to be scored as a speech pause.
2. Whether the pre-hiatus sound was **voiced** or **voiceless** was determined by referring to the phonemic transcription of the corpus. Obstruents preceding hiatus position are voiceless, whereas all other sounds were considered voiced. Vowels and other sonorants were scored separately.
3. The initial syllable of a polysyllabic word was scored as **prominent** if it has lexical (main) stress. If the main stress was elsewhere in a polysyllabic word the initial syllable was considered non-prominent. However, lexical stress is ill-defined for monosyllabic words, since there is no strong/weak-opposition. We used the computer implemented algorithm PROS (Quené & Kager, 1990) to determine for each monosyllabic word containing a hiatus position whether it would be accented or not. Typically, (monosyllabic) content words are assigned accent by this algorithm, whereas function words are not. Although this procedure may have led to an occasional infelicitous choice, it has the advantage of being explicit and automatic.
4. The context surrounding each hiatus position was examined in terms of phonotactic ambiguity. Three categories were allowed as defined above.

5. Word length of the target word (i.e., containing the hiatus position) and its left neighbour were scored separately as either monosyllabic or polysyllabic; four categories resulted.

4. Results and preliminary conclusions

The corpus contained 424 hiatus positions. Twenty-seven of these occurred word-internally, in words such as re[?]ageren ('react'), spreek[?]onderwijs ('speech training'). The distribution of glottal stop is probably different within words than across (orthographic) word boundaries, so that these 27 cases would have to be studied separately. This number is so small that no useful conclusions could ever be drawn; we therefore decided not to analyse these cases, and to concentrate on the remaining 397 hiatus positions at the beginning of orthographic word forms.

A first breakdown of the counts reveals that glottal stops are realised by PB in 56% of the hiatus positions, which is more or less a random distribution. Let us now examine the effects of each of the factors in the design separately.

Effect of speech pause preceding hiatus

The effect of presence versus absence of a speech pause immediately preceding the hiatus position is apparent from table 1.

Table 1: Effect of a preceding speech pause (presence versus absence) on distribution of glottal stop (glottal stop inserted or not). Both absolute and relative frequencies are given (re. row totals). The significance of bias in the row distributions is specified (binomial test, two-tailed).

		glottal stop insertion		
		applied	not applied	Row Total
after pause		81	0	81
		100.0	0	20.4
not after pause		142	174	316
		44.9	55.1	79.6
Column		223	174	397
Total		56.2	43.8	100.0

$p < .01$

$p = .08$

It appears from Table 1 that a glottal stop is invariably inserted after a speech pause ($p < .01$). When no pause precedes the initial vowel, i.e., when the hiatus position occurs somewhere in the middle of a phonological phrase, we observe no regularity. For this category of hiatus positions, glottal stops are realised more or less at random. However, one or more of the remaining factors in the design may narrow down the choice further.

Effect of pre-hiatus phoneme

A crosstabulation of realisation of glottal stop by type of pre-hiatus phoneme is provided in table 2 (note that the 81 cases of hiatus after a speech pause have been left out of this table).

If all sonorants (voiced consonants and vowels alike) are lumped together, the distribution of the glottal stop after this class of sounds is more or less random (97 cases realised versus 82 cases not realised). Voiceless phonemes

preceding the hiatus position, however, seem to block glottal stop insertion (only 33% realised). The predictive power of this factor may be enhanced by separating off vowels from the general class of sonorants. The hiatus position following a vowel tends to be filled by a glottal stop (74% realised). Glottal stops are then distributed at random for the third category: the sonorant consonants (45% realised).

Table 2: Effect of sonorance (vowel, sonorant consonant, voiceless obstruent) of pre-hiatus phoneme on distribution of glottal stop (excluding post-pausal hiatus positions); further see table 1.

glottal stop insertion				
	applied	not applied	Row Total	
after (voiceless) obstruent	45 32.8	92 67.2	137 43.4	p < .01
after (voiced) sonorant	54 44.6	67 55.4	121 38.2	p = .28
after (voiced) vowel	43 74.1	15 25.9	58 18.4	p < .01
Column Total	142 44.9	174 55.1	316 100.0	

It would appear from these data that glottal stops are more likely to be realised after vowels than after sonorant consonants than after obstruents. This effect runs counter to the German data reported by Krech (1968), whose hypothesis seemed unmotivated to us all along (see above).

Effect of prominence of syllable containing hiatus

Table 3 presents a crosstabulation of glottal stop realisation by the prominence of the syllable that contains the hiatus position. The 81 cases of hiatus after a speech pause have been left out of this table.

Table 3: Effect of prominence versus non-prominence of syllable containing hiatus position on distribution of glottal stop; further see table 2.

glottal stop insertion				
	applied	not applied	Row Total	
in prominent syllable	78 69.6	34 30.4	112 35.4	p < .01
in non-prominent syllable	64 31.4	140 68.6	204 64.6	p < .01
Column Total	142 44.9	174 55.1	316 100.0	

These data reveal a significant, yet moderate effect of the prominence of the syllable containing the initial vowel. Generally, the hiatus is filled by a glottal stop in 70% of the prominent syllables, and, quite symmetrically, in only 30% of the non-prominent syllables.

Effect of word boundary ambiguity due to phonotactic constraints

Table 4 presents the distribution of glottal stops broken down by the three types of word boundary ambiguity identified in §3 sub 4. Note that in this table we have left out the 58 cases after a vowel. The distribution of glottal stops in VV-sequences, given above in table 2, runs counter to that in the CV-sequences, and would obscure any effects of word boundary ambiguity in the latter set. The cases after a pause (N=81) have also been left out.

Table 4: Effect of status of syllable onset ambiguity (as defined by phonotactic restrictions, see text) on distribution of glottal stop (excluding VV-sequences); further see table 2.

		glottal stop insertion		Row Total	
		applied	not applied		
no onset can be separated		52	80	132	p = .02
		39.4	60.6	51.2	
onset can be separated no existing syllable		15	24	39	p = .20
		38.5	61.5	15.1	
onset can be separated existing syllable		32	55	87	p = .02
		36.8	63.2	33.7	
Column		99	159	258	
Total		38.4	61.6	100.0	

Disappointingly, there is simply no effect at all due to the different types of word boundary ambiguity. In this subset of the data, hiatus positions are filled by glottal stops in 37 to 39 percent of the cases. Presumably, Bloemendal (or any other Dutch speaker) does not use his implicit knowledge of phonotactic restrictions on word segmentation: he does not insert glottal stops in order to prevent confusion for the listener.

Effect of word length

In Table 5 we present the results for four possible word length combinations; note that we left out the 81 cases after a pause.

It is quite clear from table 5 that the length of the words surrounding the hiatus position exerts an effect on the distribution of the glottal stop. If the word containing the hiatus vowel is a monosyllable, chances of a glottal stop being realised are slender: 31%. However, when the hiatus vowel occurs at the onset of a longer word, the incidence of glottal stops rises remarkably: 76 (72%) realised versus 29 (18%) not realised. Only for hiatus vowels in long words does the length of the preceding word make an independent contribution. The chance of a glottal stop is diminished by about 10% when the preceding word is monosyllabic, but increased if the preceding word is longer.

Although an effect was predicted in §2.2 for the preceding word, we could not foresee that it would be weaker than the effect of the length of the word containing the hiatus vowel, nor that preceding word length is irrelevant when the hiatus word itself is a monosyllable.

Table 5: Effects of length (monosyllabic vs. polysyllabic) of pre-hiatus and hiatus word on distribution of glottal stop; further see table 2.

		glottal stop insertion		
		applied	not applied	Row Total
mono-monosyll.		45	103	148
		30.4	69.6	46.8
poly-monosyll.		21	42	63
		33.3	66.7	19.9
mono-polysyll.		45	24	69
		65.2	34.8	21.8
poly-polysyll.		31	5	36
		86.1	13.9	11.4
Column		142	174	316
Total		44.9	55.1	100.0

5. Towards an integrated model of glottal stop distribution

In our final section we shall attempt to formulate a simplified decision algorithm that will allow us to optimally determine whether a hiatus position will or will not be filled by a glottal stop, depending on the combination of context features. In §4 we have examined the effects of five potentially relevant factors separately. Clearly, one factor proved totally worthless, viz., the possibility of reducing word boundary ambiguity. A comprehensive model of the glottal stop distribution need not take this factor into account. Second, the effect of a prededing speech pause was clear cut: whenever a hiatus occurs after a speech pause it will be filled by a glottal stop; any other considerations are irrelevant here. A similar hierarchical ordering of decision criteria was established on a lower level, viz. for the effects of the length of the word preceding and following the hiatus position. The optimal model (concise yet efficient) would therefore contain the effects of and interactions between (i) speech pause, (ii) sonority of the pre-hiatus phoneme, (iii) prominence of the hiatus vowel and (iv) length of the words preceding and following the hiatus position. Table 6 contains the optimal model.

Notice, first of all, that - in our corpus - hiatus positions never occur at the onset of unstressed syllables in polysyllabic words. Therefore the word-length factor need not be specified for unstressed hiatus positions. Although this table is normally taken as the input to probabilistic rules (so called "variable rules"), we are only interested in generating deterministic rules from it. Variable rules have no application in text-to-speech systems. Therefore the criterion for glottal stop insertion was simply set at 50%: if for a particular combination of factor levels the number of glottal stops exceeded 50%, we assume that glottal stop insertion is the rule for this category ("rule on"), if the number remains below threshold, glottal stop insertion is taken to be inapplicable ("rule off").

The performance of this model is reasonable: the model predicts the insertion of a glottal stop correctly in 180 of the 223 times that speaker PB actually used a glottal stop (81% correct), and it correctly predicts that a glottal stop should not be inserted in 130 of the 174 hiatus positions that were not filled by PB (75% correct). Overall performance of the model is at 78% correct. This performance is much better than chance, even if we allow for the fact that the chance distribution is .56/.44 (cf. table 1) rather than .50/.50.

With marginal loss of performance the model can be expressed in terms of a single linguistic rule:

insert a glottal stop in hiatus position except at the onset of a non-prominent syllable not preceded by a speech pause (I-domain boundary).

This deterministic rule predicts the distribution of the glottal stop correctly in 308 of the 397 hiatus positions in the PB-speech corpus (78% correct). This is a lot better than the phonological rule that inserts a glottal stop in any hiatus position (this would perform at 56% correct), but a lot of additional work is required if we want to come up with rules that score close to 100%.

Table 6: Effects of pause (speech pause preceding hiatus), degree of sonority of phoneme preceding hiatus (obstruent, sonorant, vowel), stress on hiatus vowel, and word length (1: hiatus word short; 2: hiatus word long, pre-hiatus word short; 3: both words long) on distribution of glottal stop (insertion yes/no). If glottal stop is realised in more than 50% of the cases (%ins) the insertion rule is considered to be on, else off. The number of erroneous applications is specified (errors).

pause	sonor	stress	len	yes	no	%ins	rule	errors
+				81	0	100	on	0
-	obs	+	1	5	13	28	off	5
			2	10	8	56	on	8
			3	8	0	100	on	0
		-		22	71	24	off	22
	son	+	1	8	4	67	on	4
			2	21	13	69	on	13
			3	9	4	62	on	4
		-		16	46	26	off	16
	vow	+	1	4	2	67	on	2
			2	14	3	82	on	3
3			14	1	93	on	1	
	-		11	9	55	on?	9	
Σ				223	174			87

Acknowledgement

This work was supported by the Speech Technology Foundation, which is funded by the Netherlands Stimulation Project for the Information Sciences, SPIN. We thank Hugo Quené for ideas and discussion. Automatic generation of hiatus positions was done in collaboration with Anneke Nunn.

References

- BERENDSEN, E., LANGEWEG, S. & LEEUWEN, H.C. van
 1986 Computational Phonology, merged not mixed. Proceedings International Conference on Computational Linguistics 1986, 612-614.
- BOOIJ, G.E.
 1981 Generatieve fonologie van het Nederlands [Generative phonology of Dutch], Spectrum, Utrecht.
- HEUVEN, V.J. van, HOOS, A.
 1991 Hiatus deletion, phonological rule or phonetic coarticulation? This issue.
- KRECH, E.A.
 1968 Sprechwissenschaftlich-phonetisch Untersuchungen zum gebrauch des Glottisschlageinsatzes in der allgemeinen deutschen Hochlautung, Basel.
- LEEUWEN, H.C. van
 1989 TooLip: A Development Tool for Linguistic Rules, doct. diss. Technical University Eindhoven.
- MALECOT, A.
 1975 The Glottal Stop in French, Phonetica, 31, 51-63.
- NOOTEBOOM, S.G.
 1985 A functional view of prosodic timing in speech, in J.A. Michon, J.L. Jackson (eds.): Time, mind and behavior, Berlin: Springer, 242-252.
- QUENE, H.
 1989 The influence of acoustic-phonetic word boundary markers on perceived word segmentation in dutch, doct. diss, Utrecht University.
- QUENE, H., KAGER, R.
 1990 Pros: automatic prosodic sentence analysis, accentuation and phrasing for dutch text-to-speech conversion, SPIN/ASSP Report nr.17, Stichting Spraaktechnologie, Utrecht.
- SCHARPFF, P.J.
 1987 Effect of context and lexical redundancy on continuous word recognition, Proc. 11th Int. Congress Phonetic Sciences, vol. 5, 43-47.
- SCHARPFF, P.J., HEUVEN, V.J. VAN
 1988 Effects of pause insertion on the intelligibility of low quality speech, Proc. 7th FASE/SPEECH-88, 261-268.