

ON THE USE OF THE GLOTTAL STOP IN FINNISH CONVERSATIONAL SPEECH

Mietta Lennes¹, Eija Aho², Minnaleena Toivola¹ and Leena Wahlberg¹

¹Department of Speech Sciences, University of Helsinki, Finland

²Department of General Linguistics, University of Helsinki, Finland

e-mail: mietta.lennes@helsinki.fi

Abstract

Glottal stops can be used for various tasks in speech. This study provides an overview of the different contexts where glottal stops tend to occur in casual Finnish conversation. Glottal stops in four phonetically annotated dialogues are investigated. Preliminary findings are reported on the durations of glottal stops in the different conditions. The possible interactional functions of glottal stops are discussed.

Keywords: glottal stop, Finnish, conversational speech

1 Introduction

A **glottal stop** is a plosive-like consonant sound whose closure is produced and released in the glottis. Some researchers refer to this sound as the **laryngeal plosive**. The glottal stop is not part of the Finnish phonemic system. However, nearly all Finns sometimes use glottal stops in their speech. One may say, e.g., *ota omena* [ʔotaʔ:omena] ‘take an apple!’ or *se oli ihan ihme juttu* [seoliʔihanʔihmej:utu] ‘it was quite a strange thing’.

The definition of the glottal stop may vary according to the research topic and the granularity of the phonetic analysis involved. In principle, during the production of a complete glottal stop, the vocal folds are held together tightly for a brief moment, preventing simultaneous phonation. Thus, a glottal stop should be unvoiced by definition. In the acoustic waveform of a complete glottal stop, a quiet closure phase may be distinguished, and a short burst of noise may often be observed just before the following sound begins.

In practice, however, it is often difficult to distinguish between a glottal stop, very low-pitch voice, and creaky voice (i.e., glottalization or laryngealization), since these are phonetically similar and may occur next to each other in speech. For instance, Ogden (2001) defines the glottal stop as one or more irregular glottal periods that occur within the same syllable. Thus, a glottal stop may sometimes be represented as a very short creaky phase in speech, and a complete stop closure is not required.

According to Ladefoged and Maddieson (1996:75), the complete closure will often be lacking from intervocalic glottal stops, and a complete glottal stop will consistently occur only in those languages and positions where the stop is geminated.

They also point out that in languages where the glottal stop has a phonemic status, this segment is frequently replaced with creak that tends to “spread” over the neighbouring sounds.

The weakly produced form of the glottal stop may be separately referred to as the **glottal approximant** (in Finnish, **puristussupistuma**). In this sound, the glottal closure is incomplete and consequently a clear release phase is lacking. A glottal approximant can be acoustically observed as the momentary lengthening and irregularity of glottal periods, and it is auditorily and acoustically similar to a short creaky phase in speech.

Ladefoged (1971) defined a phonetic continuum where unvoiced sounds are at one end, the glottal stop at the other, and modal phonation in the middle. Breathy voice quality resides between unvoiced sounds and modal phonation, whereas creaky voice is phonetically an intermediate form between modal voice and the full glottal stop. According to Esling and Harris (2005), glottal stops and creak are in fact produced with the same physiological mechanism. In addition to the glottal closure or approximation, both creaky voice and glottal stops are produced by adducting the arytenoid cartilages with the epiglottis above the larynx. This aryepiglottal constriction affects both the production and the perceived quality of the glottal periods. A glottal stop is perceived when a single glottal pulse is produced with the aforementioned laryngeal configuration, and creaky voice is perceived when several such pulses occur in a row (Esling & Harris 2005).

During a glottal stop, one is indeed “holding one’s breath” before the next speech sound or until there is another chance to speak. Tentatively, the glottal closure can help to maintain subglottal pressure in order to quickly produce a loud voice as soon as the next word comes to mind. Moreover, a similar, even tighter glottal and aryepiglottal closure is utilized by humans during heavy muscular strain. Thus, people may naturally interpret glottal stops as a signal of “hard work” or some kind of active processing.

1.1 The occurrence of glottal stops in different languages

In many languages where the glottal stop is not phonemic, both glottal stops and glottal approximants as well as their intermediate variants tend to occur at the beginning of vowel-initial words and/or at intervocalic morpheme boundaries. In Finnish, this sound is referred to as the “hard attack” (**luja aluke**), which is particularly typical for Savo dialects (cf. Kettunen 1940, Itkonen 1964). However, the hard attack is not obligatory and its occurrence may vary even in the speech of an individual speaker.

After specific Finnish words that end with a vowel, the consonant at the beginning of the next word or suffix may be “doubled”, i.e., it is produced like the corresponding long consonant. This phenomenon may be called **final** or **initial doubling**. When a complete glottal stop occurs at the end boundary of a word where final or initial doubling is applicable, the glottal stop may respectively be longer in duration, i.e., it may be “doubled” like any other consonant at a similar juncture (see, e.g., Itkonen 1975).

Although it is possible to list factors that increase the probability of a glottal stop in different languages, it is difficult to predict exactly when a glottal stop will occur. Prosodically, a glottal stop or a glottal approximant may in many languages be used for emphasizing the next word or a prosodic boundary. Word-initial vowels are more frequently glottalized (i.e., glottal stops or approximants are produced) at major prosodic boundaries (Pierrehumbert & Talkin 1992; Dilley, Shattuck-Hufnagel & Ostendorf 1996). In some forms of British English, the glottal stop is used as a segmental variant of /t/, or instead of word-final or intervocalic /p/ and /k/ (e.g.,

Trudgill, 1974:81). The glottal stop may even have sociolinguistic functions. Kirk (1967) showed that in Northern Ireland, Protestants employed these sounds more frequently than Catholics.

1.2 Hypotheses of the use of glottal stops in Finnish

Ogden (2001) has suggested that glottal stops and creaky voice have separate interactional functions in Finnish. According to his claim, speakers may use glottal stops for turn holding in association with, e.g., word search, self repair or incomplete Turn Constructional Units (TCUs). Creaky voice, on the other hand, would indicate possible turn transition points to the other participants.

Turn transitions are likely to occur at pauses. In case glottal stops were systematically used to hold turn, these sounds should rarely occur in a pre-pausal position, since it would be more efficient to entirely fill the pause with the glottal stop. Successful turn-holding with glottal stops would also imply that the other speakers frequently react to them either by quickly finishing their own speech or by not trying to speak at all. Therefore, it should be improbable to find overlapping speech during or after a glottal stop.

Nevertheless, all glottal stops in Finnish speech are probably not used for turn-holding. The aim of the present study is to give an overview of the different contexts of glottal stops in Finnish conversational speech. In case glottal stops have prosodic as well as interactional functions in Finnish, there may be phonetic differences between the glottal stops with different functional status. We will find out whether there are durational differences between glottal stops in, e.g., “word search” vs. “hard attack” contexts. We will also take a look at the proportion of overlapping speech in different contexts. Only glottal stops with complete closure will be analysed.

2 Material and methods

This study is based on the analysis of four unscripted dialogues, whose participants were young adults (age 21-28 years; 4 female and 4 male). The speakers had lived for all of their lives in the capital city region in Finland (in or near Helsinki and Espoo). Each pair of speakers knew each other well. The speakers will be referred to with the codes F1-F4 for female speakers and M1-M4 for males. Conversations were recorded among F1 and F2, F3 and F4, M1 and M2, and M3 and M4.

The recordings were performed in the anechoic room of the Laboratory of Acoustics and Audio Signal Processing at the Helsinki University of Technology. The speakers were sitting two to three meters apart, facing opposite directions. They heard both their own and the other speaker's voice through headphones. Thus, the situation somewhat resembled a telephone conversation. The speakers were allowed to chat freely and unmonitored for 45-60 minutes. Each speaker's voice was recorded with a high-quality headset microphone (AKG HSC-200 SR) to a separate channel of a DAT recorder (Tascam DA-P1, sample rate 44,1 kHz, sample size 16 bit). The digital stereo signal was then transferred from the DAT to a computer and downsampled to the rate of 22.05 kHz. The two channels of each stereo file, each channel representing one speaker's voice, were divided into two sound files of identical duration. This procedure allowed for the separate analysis of each speaker's phonetic output. Only minimal crosstalk was observed between the channels.

The recorded signals were first annotated with the Praat program (Boersma & Weenink 2006) using multiple annotation tiers. Each speaker's utterances were

transcribed quasi-orthographically and utterance boundaries were marked in one tier. In this study, an utterance refers to a portion of the sound signal during which a speaker is continuously articulating. Word boundaries were also marked in a separate tier. For part of the material, phone segments were delineated and their phonetic transcriptions added in a phone tier. Glottal stops with a complete closure phase were included in the phone tier segmentation. Glottal approximants on the other hand were marked as creaky voice quality in yet another tier, but this information will not be utilized for the present purposes.

In order to obtain a general picture of the different uses of glottal stops, a preliminary auditory analysis was performed on a total of 80 glottal stops selected from different speakers. Impressionistic notes were taken on the function of each glottal stop occurrence, exploiting visual displays of the acoustic waveform and its fundamental frequency (F0) curve provided by Praat.

After the manual work, an automatic analysis was run on all the annotated glottal stops. This analysis was performed with a Praat script. As a result, a table was obtained with its each line corresponding to one glottal stop occurrence. Information was automatically collected on the duration of each glottal stop, whether it occurred directly before and/or after a vowel, which words preceded and followed the sound, whether the glottal stop occurred after a pause (i.e., utterance-initially), the index number of the following word within the utterance, the time elapsed from the beginning of the utterance, and whether there was overlapping speech by another speaker at the temporal midpoint of the glottal stop. A total of 323 glottal stops were analysed in this fashion.

3 Results

During the preliminary auditory analysis, a general impression was obtained about the different contexts and functions of glottal stops. The first observation was that glottal stops do indeed often occur in front of vowel-initial words or at an intervocalic boundary. However, there were only three glottal stops (approximately 1 % of all cases) that exhibited final/initial doubling in the whole dataset. Glottal stops were often used for the emphasis or accentuation of the vowel-initial word. An example of emotional emphasis is shown in figure 1, where the speaker F4 highlights the word *ikinä* ‘ever’ with a preceding glottal stop.

Many glottal stops were apparently associated with a “will to speak”: they may be used while searching for the next word (cf. the example in figure 2), or for obtaining (figure 3) and maintaining turn, i.e., a permission to continue speaking. However, not all glottal stops appeared to have such interactional use. Moreover, glottal stops did not always seem to prevent the other speaker from speaking, since many glottal stops were overlapped by the other participant in the dialogue.

Several or all of the aforementioned impressionistic functions or contexts of use may in principle occur simultaneously, and it was not possible to objectively determine the functional status of each case. However, it was possible to automatically separate several different contextual conditions where glottal stops tend to occur.

Utterance-initial as well as **utterance-final** glottal stops were treated as separate groups, since their delineation from the speech signal is often partly arbitrary and thus the duration of such cases cannot be reliably measured. The remaining data was further divided into three subsets. One group contains all glottal stops that occurred in front of vowel-initial words (but not utterance-initially) and were not associated with an immediately preceding or following incomplete word. We refer to this group as the

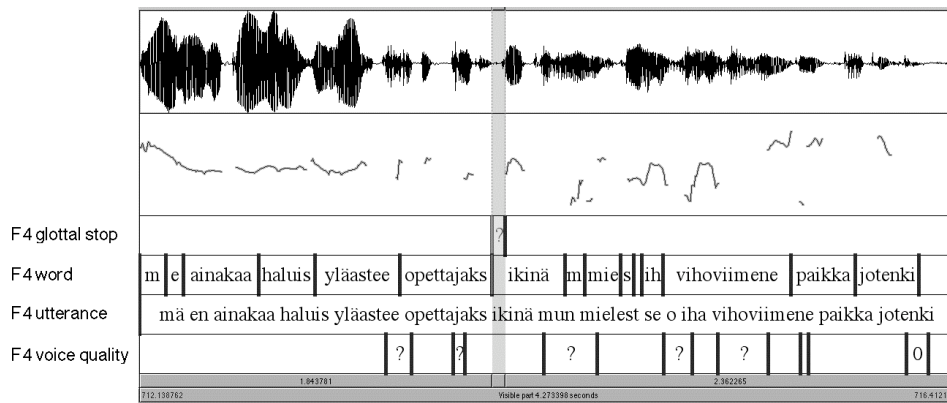


Figure 1. Speaker F4 uses a glottal stop in order to emphasize the vowel-initial word *ikinä* ‘ever’. Glottal stops and creaky voice are indicated in their separate annotation tiers with question marks. The area of the glottal stop is shadowed. The voice quality mark 0 indicates the voiceless or whispery end of the utterance. The acoustic waveform is shown at the top, and pitch is displayed in the middle. Note that the pitch curve at the final part of the utterance is not reliable due to “octave jumps” caused by voice quality changes.

“**emphatic**” glottal stops, although some less emphatic cases of hard attack are also covered. The glottal stops associated with **incomplete words** were separated into another group. Such words had been marked in the orthographic transcript with a dash, e.g., *mit-*, and thus they were easy to pick out from the data. The remaining glottal stops were, on closer inspection, most often associated with **word search** of some kind: there were many false starts or repeated words, and the speaker appeared to be “taking some time to think” in practically all of the cases in the fourth group. Next, we will take a closer look at these different contexts.

3.1 Position within utterance

Glottal stops tend to occur mostly around the beginning and the middle portion of utterances (see figure 4). This is hardly surprising, as glottal stops are most frequently used in word-initial positions. Approximately 27 % of all glottal stops occurred utterance-initially. The female speaker F3 used more than half of her glottal stops at utterance-initial positions (16 cases out of 27). In only three cases (1 %) did a glottal stop occur pre-pausally.

3.2 Overlaps

In case the glottal stop is a turn-holding signal by Finnish speakers, it should be unlikely for glottal stops to be overlapped by another speaker’s speech. Only 10 % of the “emphatic” glottal stops were overlapped. This may be either a pure coincidence or due to the perceptually prominent nature of these vowel-initial words, which may prevent attempts to overlap.

None of the three pre-pausal glottal stops were overlapped by another speaker, which might indicate that they had been interpreted as a continuation signal. On the

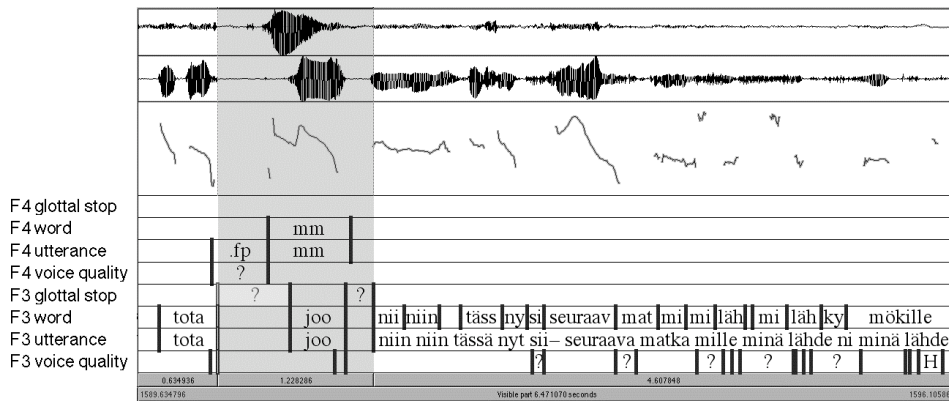


Figure 2. Speaker F3 (the four annotation tiers at the bottom of the figure) uses two glottal stops for holding her turn while thinking about something to say about a newly introduced topic. Speaker F4 acknowledges this by uttering *mm* with a creaky-voiced start. Glottal stops and creaky voice are indicated with question marks in separate annotation tiers. The area of the two glottal stops is shadowed. Note also the gradual increase in creaky voicing towards the end of speaker F3’s turn. The acoustic waveforms of each speaker’s speech are shown at the top of the figure: F4 is topmost and F3 below F4.

other hand, there was overlapping speech during 58 % of those glottal stops that occurred utterance-initially (cf. figure 6). This is not surprising, however, since the start of an utterance would often represent a turn transition point. Speaker F3 produced the largest proportion of overlapped glottal stops (13 out of 27). Nine of her overlapped glottal stops occurred utterance-initially, which means that she caused these overlaps herself.

22 % of the glottal stops in incomplete words and 21 % of other possible word search cases were overlapped by the other speaker. A general impression of these overlaps was that the other participant wanted to help or encourage the speaker who was having trouble with finding the right word. Overlapping may thus signal cooperation and mutual understanding instead of competition. Turn-holding and turn-taking are not necessarily “selfish” activities but they often reflect the social conventions and means of negotiation amongst the speakers.

3.2 Duration

As seen in figure 5, the distribution of segmental duration values is slightly skewed and not a Gaussian one, which is typical for segmental duration measurements in general. Thus, the mean value is not a good predictor of the expected duration and median values are a better alternative for making comparisons. For all 323 glottal stops that were analyzed, the mean duration was 120 ms and median 85 ms. When utterance-initial glottal stops are excluded, the total median was 96 ms. For “emphatic” glottal stops in vowel-initial words (group 2) the median was 67 ms and there was much less variation than for the other groups. In the three cases where an “emphatic” glottal stop was lengthened due to final/initial doubling, its duration was 100-150 ms, which is approximately twice the expected duration of the “emphatic” glottal stops. These values correspond to the expected duration ratio of long and short consonants in Finnish (cf.,

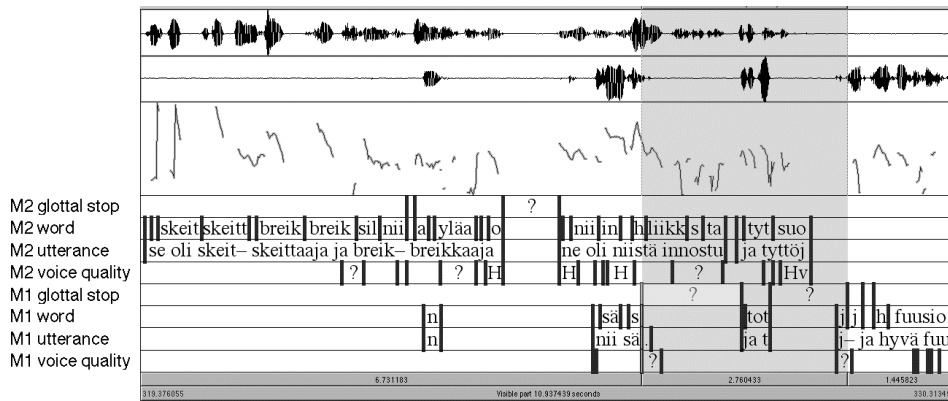


Figure 3. The male speaker M1 is trying to take a turn in order to give a comment. M1 uses several glottal stops for signaling his wish to speak while speaker M2 is still continuing. Glottal stops and creaky voice are indicated in separate annotation tiers with question marks. The acoustic waveforms of each speaker’s speech are shown at the top of the figure: M2 topmost and M1 below M2.

e.g., Lehtonen, 1970).

Glottal stops tend to be somewhat longer in duration in cases where they do not occur utterance-initially or in front of “emphatic” vowel-initial words. The longest glottal stops tend to occur in incomplete words and during word search.

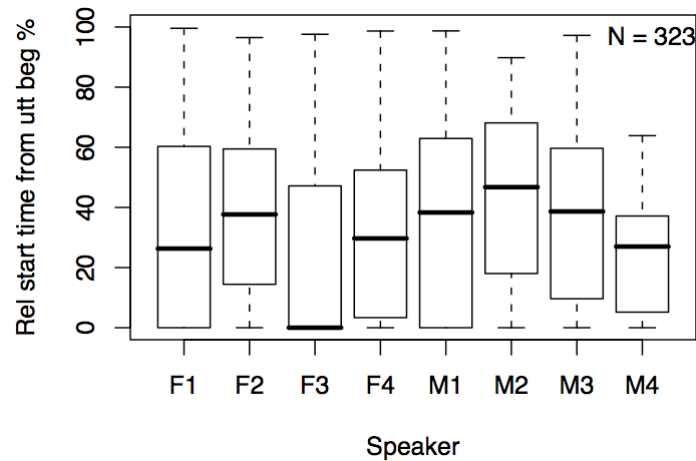


Figure 4. The position of glottal stops within utterances of eight speakers. The starting point of glottal stops within the utterance is shown in the vertical axis in proportion to total utterance duration. The majority of glottal stops occurred within the first halves of utterances.

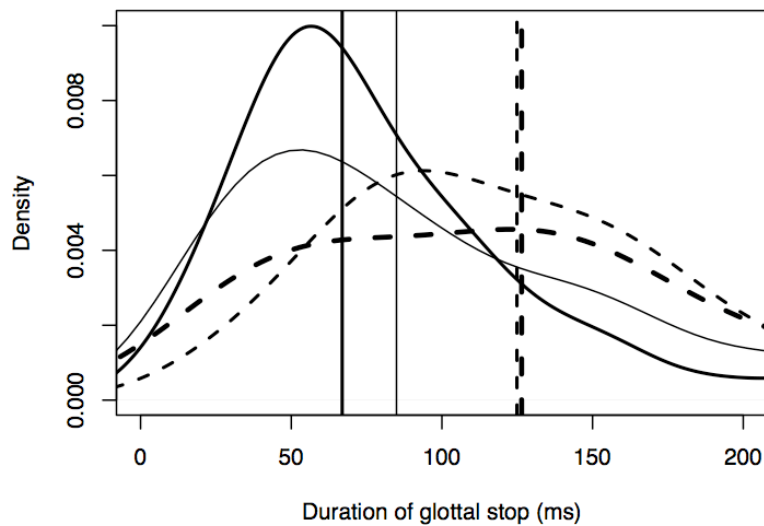


Figure 5. The distributions of glottal stop duration for the different subsets of the data. Vertical lines denote the median duration for each subset. Thin solid line = the overall distribution of all 323 glottal stops. Thick solid line = “emphatic” glottal stops, i.e., non-utterance initial, immediately preceding a vowel-initial word, not associated with an incomplete word (N=76). Thin dashed line = immediately following an incomplete word (N=64). Thick dashed line = “word search”, i.e., between two words, not associated with a vowel-initial or incomplete word or a pause (N=78).

3.4 Speaker-dependent differences in the use of glottal stops

All of the speakers did sometimes use glottal stops in utterance-initial positions, for emphasizing a vowel-initial word or the word boundary preceding it, at the ends of incomplete words or “false starts”, or between words during word search. However, as seen in figure 6, there were differences in the distribution of the different types of glottal stops among the speakers and dialogues. Some speakers apparently favoured glottal stops for emphasis or “hard attack” (especially F4), whereas others (e.g., M1) tended to use these sounds more often for word search.

The number of incomplete word cases tended to be rather similar amongst the speakers within each dialogue. The speakers F4 and M4 did not produce any glottal stops that would have been overlapped by their fellow speakers, although the other six speakers did. This indicates that speakers may have had different strategies for turn-taking and turn-holding, i.e., for balancing the speaking time between speakers during the dialogue.

4 Conclusions

The term ‘glottal stop’ may have various definitions and interpretations. In Finnish, glottal stops can be used for different linguistic, phonetic and interactional purposes. In this study, we investigated a Finnish dialogue corpus, focusing on glottal stops that are produced with a complete glottal closure. Due to the method of automatic analysis that was selected in the present study, it was impossible to make comparisons with positions

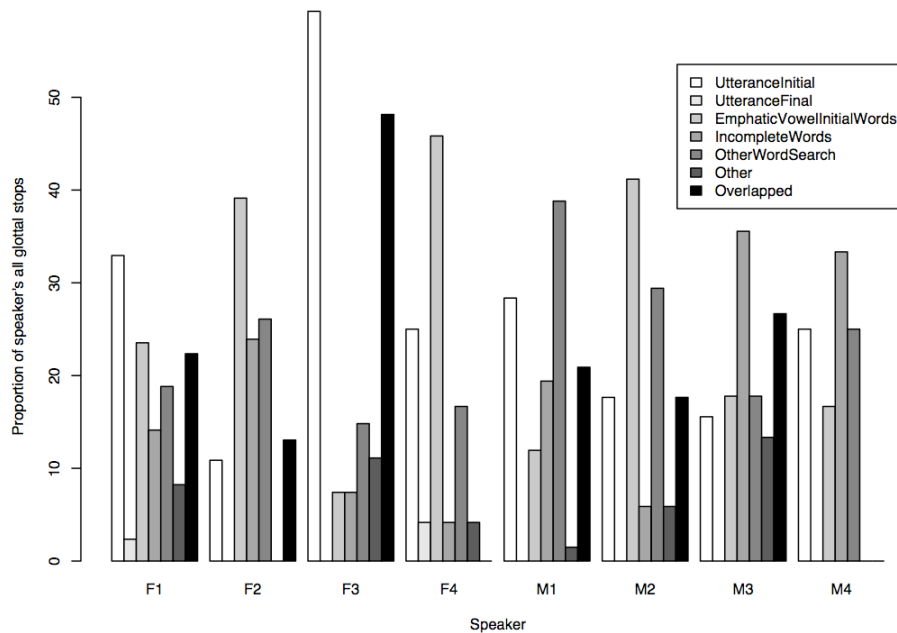


Figure 6. The proportions of different glottal stop contexts for eight speakers. Female speakers are shown on the left, male speakers on the right side. Each adjacent pair of speakers participated in one dialogue together (F1 with F2, F3 with F4, M1 with M2, and M3 with M4).

where glottal stops did *not* occur (contrary to a given prediction). For this reason, it is not yet possible to state exactly when glottal stops will be used in spoken Finnish.

However, we were able to determine some contexts where glottal stops generally occur. Glottal stops were primarily used as word boundary signals before vowel-initial words (i.e., as either “hard attack” sounds or for emphasizing words) or during word search and incompletely produced words. In front of vowel-initial words, glottal stops tended to have roughly the same duration as other consonants in similar position. However, in case a glottal stop was associated with word search or a false start, it was often much longer in duration. Glottal stops were rarely found immediately preceding a pause, but they were rather common in utterance-initial positions. These findings indicate that glottal stops with complete closure can be used for signaling one’s intention to continue speaking, i.e., for holding turn.

Acknowledgments

The work of Mietta Lennes and Leena Wahlberg has been partly supported by the Academy of Finland (Grant No. 107606).

References

- Boersma, P. and Weenink, D. (2006). Praat: doing phonetics by computer (Version 4.4.24) [Computer program].
- Dilley, L., Shattuck-Hufnagel, S., and Ostendorf, M. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics*, 24:423-444.

- Esling, J. H. and Harris, J. G. (2005). States of the glottis: An articulatory phonetic model based on laryngoscopic observations. In Hardcastle, W. J. and Mackenzie Beck, J., editors, *A Figure of Speech: a Festschrift for John Laver*. Mahwah: Lawrence Erlbaum.
- Itkonen, T. (1964). Proto-Finnic final consonants: their history in the Finnic languages with particular reference to the Finnish dialects. Helsinki: SKS.
- Itkonen, T. (1975). Suomen ja eteläviron sandhiongelmia. *Virittäjä*, 371–386.
- Kettunen, L. (1940). Suomen murteet III A. Murrekartasto. Helsinki.
- Kirk, T. (1967) *The Religious Distribution in Lurgan*. Queen's University Belfast.
- Ladefoged, P. (1971). *Preliminaries to Linguistic Phonetics*. Chicago: The University of Chicago Press.
- Ladefoged, P. and Maddieson, I. (1996). *The Sounds of the World's Languages*. Cambridge, MA: Blackwells.
- Lehtonen, J. (1970). *Aspects of Quantity in Standard Finnish*. Studia Philologica Jyväskyläensia VI. Jyväskylä: University of Jyväskylä.
- Ogden, R. (2001). Turn transition, creak and glottal stop in Finnish talk-in-interaction. *Journal of the International Phonetic Association*, 31:139-152.
- Pierrehumbert, J. and Talkin, D. (1992). Lenition of /h/ and glottal stop. In Docherty, G. and Ladd, D. R., editors, *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, pp. 90-127. Cambridge: Cambridge University Press.
- Trudgill, P. J. (1974). *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.