VARIATION IN GLOTTALIZATION AT PROSODIC BOUNDARIES IN CLEAR AND PLAIN LAB SPEECH

Erin Luthern and Cynthia G. Clopper

Ohio State University
luthern.5@buckeyemail.osu.edu, clopper.1@osu.edu

ABSTRACT

Previous research on glottalization shows that this voice quality occurs more frequently at prosodic boundaries than in the middle of prosodic phrases. This study investigates ten speakers' use of glottalization at prosodic boundaries in five passages read in both clear and plain lab speech. We analyzed each syllable in every passage for its voice quality (glottalized or modal) and for its prosodic boundary strength using the ToBI system. We found that glottalization is used regularly in phrase-final syllables, and marginally in phrase-initial syllables. We also found that speaking style had no effect on overall glottalization.

Keywords: voice quality, glottalization, prosody, speaking style

1. INTRODUCTION

Glottalization is a voice quality defined by its distinctive acoustic and auditory expression. This quality is reflected in acoustic signals that include aperiodicity and low fundamental frequency [11], which can be seen by comparing the aperiodic, glottalized vowel in Figure 1 to the same vowel that is modally voiced in Figure 2. Glottalization also has a unique auditory impression that is often distinguishable from other voice qualities without visual inspection of the acoustic signal [11]. Glottalization defined thus is also known as creaky voice or vocal fry [6].

Glottalization as a broad category of voice quality is phonologically contrastive in some languages, such as Chong and Mazatec [2], although it is not similarly contrastive in English [12]. Early interpretations of the function of glottalization in English were based on an informal claim that glottalization could signal speaker boredom, and therefore contributed no significant meaning to our understanding of speech [9]. However, this interpretation of glottalization was not based on the analysis of actual speech data [6].

Figure 1: Example of a glottalized syllable "cheese".

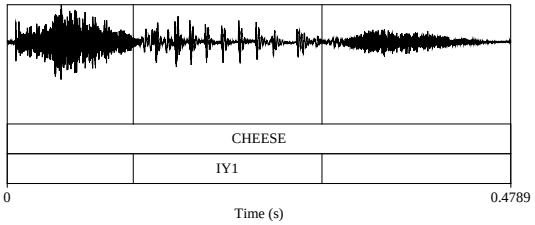
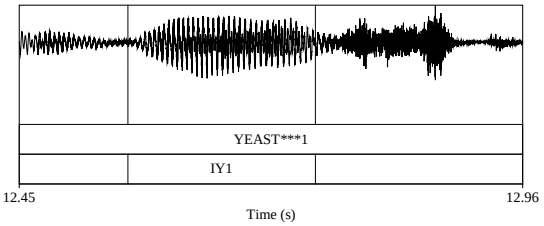


Figure 2: Example of a modal syllable "yeast," produced by the same speaker as in Figure 1.



More recent research suggests that glottalization is used in English to signal various prominent aspects of speech, including phrase edges and stressed vowels in word-initial position [5, 7]. In particular, glottalization is associated with prosodic phrasing. Both glottal stops and glottalization unrelated to a stop are more likely to be observed at phrase boundaries than in phrase-medial position [4, 6, 11]. Glottalization is also associated with phonological processes, such as word-final stops becoming glottalized, and the glottalization of vowel-initial words. The process of vowel-initial glottalization is also more likely to occur when the syllable occurs in phrase-initial position [10].

Phonetic prominence is also affected by speaking style. For example, clear lab speech, directed towards an imagined hearing-impaired or non-native speaker, has several defining qualities such as an exaggerated pitch range, decreased speaking rate, and exaggerated prominence of pitch accents on accented words in a phrase [13]. These enhancements of prominence in clear speech are interpreted as the talker's attempt to assist the listener in correctly parsing the

utterance. Given that glottalization is used to mark prominent syllables and prosodic boundaries, we expect that glottalization at prosodic boundaries may be produced more frequently in clear speech than in plain lab speech. The current study tests this prediction through an examination of the frequency of glottalization at prosodic boundaries in clear and plain lab speech.

2. METHODS

2.1. Talkers

The data for this study were collected from ten female speakers aged 18-25 years old. All 10 speakers are native speakers of the Midland dialect of American English, which is spoken in the southern half of the American Midwest, stretching from Ohio through Iowa, Missouri, Kansas, and Nebraska, bordered by the Ohio River to the south, and to its northern border including southern and central Ohio, Indiana, and Illinois [8].

2.2. Stimulus materials

Each speaker was prompted to read a series of five passages, which were read at a self-paced rate, with each passage displayed on the computer screen one at a time. All passages were first read in a plain lab speaking style, as though speaking to a friend; speakers were then prompted to read the same passages in a clear lab speaking style, as though speaking to a hearing-impaired or non-native listener. The data therefore comprise a total of 100 passages, with each of the 10 speakers reading the five passages in both clear and plain speaking styles. Speakers were recorded in a sound-attenuated booth with high-quality digital recording equipment. The read passages were 35-90 s in duration.

2.3. Voice quality coding

Every syllable in the recordings was annotated for voice quality using a series of automated and manual processes. The recordings were first analyzed with the Penn Phonetics Lab Forced Aligner [14], then manually checked for accuracy. Based on previous manual coding of voice quality in these data, we assumed an f0 cutoff of 150 Hz, with all f0 values above 150 Hz being identified as modal, and those below 150 Hz being considered glottalized. Because glottalization can be found at any point within a vowel, we sampled the f0 from each vowel in each passage at 10 equally spaced timepoints throughout the vowel's duration using Praat [3]. Undefined f0 values were manually corrected; f0 values that were above 350 Hz, and therefore outside the typical range for an adult female speaker, were examined and manually corrected as necessary. Following hand-correction, if six or more of the 10 samples within a given vowel were tagged as modal (with an f0 greater than 150 Hz), the vowel was considered modal. If fewer than six of the 10 samples within a given vowel had an f0 greater than 150 Hz, the vowel was considered glottalized. This automatic method with hand-correction was tested against previous hand-tagged vowels, and an agreement rate of 91% between the automatic and hand-tagged vowels was observed, suggesting that this automated process is highly reliable for identifying glottalization. The total data set included 17,243 tokens, of which 566 were excluded due to text misalignment in the forced-alignment phase or vowel deletion.

2.4. Prosodic annotation

For the analysis of the prosodic boundaries, we used the Tones and Break Indices (ToBI) system [1], which subjectively categorizes the perceived strength of the boundary separating the target word or syllable from those following it. The ToBI scale for break indices creates a 5-point scale from 0 to 4, with the highest values of 3 and 4 being associated with the prosodic constituents of intermediate phrases and intonational phrases, respectively. In the current study, a syllable with a break index of 3 or 4 immediately preceding it was classified as having a preceding prosodic boundary. Similarly, a syllable with a break index of 3 or 4 immediately following it was classified as having a following prosodic boundary.

3. RESULTS

For each speaker, the proportion of glottalized syllables relative to the total number of syllables was calculated separately for each prosodic position (the presence of a preceding boundary, following boundary, both a preceding and following boundary, or neither) and speaking style (plain vs. clear lab speech). Tokens with both a preceding and following boundary were excluded from the following analysis, as there were only 10 of them. Table 1 shows the mean proportion of glottalized syllables with a preceding boundary,

following boundary, or no boundary in both clear and plain speaking styles. A repeated-measures ANOVA with phrase position and speaking style as within-subject factors revealed that boundary presence was a significant factor in glottalization ($F(2,18)=9.35$, $p=.002$). This ANOVA also found that speaking style had no significant effect on glottalization ($F(1,18)=2.71$, $p=.134$), and the interaction between speaking style and boundary presence was also not significant ($F(2,18)=1.64$, $p=.221$). A post-hoc paired t-test comparing preceding boundary to no boundary, collapsed across speaking style, showed that preceding boundaries led to marginally more glottalization than no boundaries ($t(9)=1.99$, $p=.077$). Thus, syllables with a preceding boundary are more likely to be glottalized than those that are phrase-medial. A paired t-test also revealed that syllables are more likely to be glottalized when followed by a prosodic boundary than when preceded by a boundary ($t(9)=-2.29$, $p=.048$) or phrase medial ($t(9)=4.37$, $p=.002$).

Table 1: Proportions of glottalized syllables with a following prosodic boundary, preceding boundary, or no boundary in clear and plain speaking styles. Standard deviations are shown in parentheses.

Phrase position	Clear speech	Plain speech
Following boundary	.48 (.29)	.51 (.25)
Preceding boundary	.22 (.17)	.34 (.24)
No boundary	.16 (.12)	.18 (.11)

Given the marginal effect of a preceding boundary on glottalization, we wanted to explore how much of that glottalization came from vowel-initial versus consonant-initial words, given the proposed interaction between vowel-initial words and phrase-initial contexts [10]. Table 2 shows the proportions of glottalization in syllables in phrase-initial position when the syllable is consonant-initial or vowel-initial in both speaking styles. A repeated-measures ANOVA revealed that vowel-initial words are more likely to be glottalized when preceded by a prosodic boundary than consonant-initial words ($F(1,9)=20.03$, $p=.002$). The effect of speaking style was not significant ($F(1,9)=2.93$, $p=.12$), nor was the interaction between speaking style and the initial segment ($F(1,9)=2.19$, $p=.17$).

Table 2: Proportions of glottalized syllables preceded by a prosodic boundary when the syllable is consonant-initial vs. vowel-initial. Standard deviations are shown in parentheses.

Phrase-initial onset	Clear speech	Plain speech
Vowel-initial	.32 (.20)	.49 (.26)
Consonant-initial	.17 (.17)	.28 (.27)

One striking aspect of the data that is masked by the statistical analysis is the considerable inter-speaker variation in glottalization between speaking styles with respect to following boundaries, as shown in Table 3. Whereas some speakers showed a large difference in the proportion of glottalization relative to a following intonational phrase boundary in both speaking styles, other speakers did not. For example, Speaker 72 consistently produced glottalization before prosodic boundaries, regardless of speaking style, and is an example of the overall trend towards the importance of following prosodic boundaries in the data. By contrast, Speaker 153, who showed a similar overall degree of glottalization to Speaker 72, exhibited a larger effect of speaking style than prosodic position, using glottalization more frequently in plain speech than clear speech, and is therefore less representative of the sample as a whole.

Table 3: Proportion of individual speakers' use of glottalization with (F) and without (No F) a following prosodic boundary in both clear and plain speech. Total proportion of glottalization for each speaker is shown in the last column.

Speaker	Clear speech		Plain speech		Total
	No F	F	No F	F	
70	0.06	0.68	0.12	0.78	0.15
72	0.24	0.77	0.22	0.67	0.28
73	0.08	0.65	0.07	0.57	0.12
76	0.05	0.41	0.06	0.30	0.09
81	0.14	0.56	0.19	0.56	0.21
86	0.34	0.73	0.31	0.80	0.38
89	0.36	0.90	0.41	0.66	0.43
136	0.16	0.13	0.16	0.11	0.15
150	0.14	0.10	0.14	0.18	0.14
153	0.19	0.14	0.32	0.35	0.25

Although our analysis did not uncover a significant effect of speaking style on glottalization, some speakers showed large stylistic differences that were not captured by the statistical analysis. For example, Speaker 89

glottalized nearly all (90%) of her syllables with a following boundary in the clear speaking style, but only 66% of her syllables with a following boundary in the plain speaking style. This pattern is consistent with our prediction that glottalization in phrase-final position would be enhanced in clear speech relative to plain speech. However, as noted above, Speaker 153 showed an increased use of glottalization in plain speech relative to clear speech, regardless of the presence of a following boundary, contrary to our prediction. Thus, the lack of a speaking style effect in our analysis may reflect individual differences in the use of glottalization across styles.

4. DISCUSSION

In this research, we found greater proportions of glottalization in phrase-final syllables than in non-phrase-final syllables, as well as marginally more glottalization in phrase-initial syllables relative to phrase-medial syllables. The results support the role of glottalization as a marker of prosodic boundaries, even when there is no stop to be glottalized at these boundaries [11]. Previous research suggested that glottalization occurs for various reasons, whether prosodic or phonological, and that there is some interaction between these functions [4, 6, 10, 11]. Our results show that glottalization is used in both phrase-initial and phrase-final positions, which suggests that glottalization is closely linked to suprasegmental prosodic structures.

Our finding of marginally more glottalization in phrase-initial syllables than in non-initial syllables is in line with previous research [4]. We considered the possibility that the proportion of glottalization in phrase-initial position might be strongly affected by the frequency of vowel-initial words at phrase onsets. To explore this possibility, we analyzed phrase-initial syllables, separating them into vowel-initial and consonant-initial categories. The results showed that vowel-initial syllables in phrase-initial position are more likely to be glottalized than their consonant-initial counterparts. These data are in accordance with previous work, which describes vowel-initial glottalization as the result of simple physiological difficulty in producing a vowel after a pause in speech [4].

The lack of an effect of speaking style on glottalization suggests that glottalization as a phrase-final marker is used similarly across speaking styles. Thus, although glottalization is used for marking the edges of phrases, this prosodic marking is not exaggerated in clear speech, unlike other dimensions of speech such as pitch range, duration, and vowel quality.

The marked variation between speakers' use of glottalization within and across speaking styles suggests that other factors may drive the use of glottalization within this group of speakers. Although an overall effect of speaking style was not observed, some speakers produced the predicted pattern of more glottalization in clear speech than in plain speech, especially before prosodic boundaries (e.g., Speakers 72, 76, and 89). We attempted to control speaker homogeneity by using participants who were all young (18-25 years old), female speakers of the Midland dialect of American English. Because of this control over some basic social factors, we expected to see more uniformity in the overall rates of glottalization. However, our data cannot clearly explain the varied use of glottalization between speaking styles for individual speakers. For example, Speakers 86 and 89 have similar overall glottalization proportions, but exhibit opposite effects of speaking style. Whereas Speaker 86 shows an increased use of glottalization in phrase-final syllables in plain speech relative to clear speech, Speaker 89 uses much more glottalization in phrase-final syllables in clear speech than in plain speech. The variation in these data suggests that further studies of glottalization and speaking style should explore other potential factors that influence its use.

Further, these data were analyzed from read speech, which may differ from spontaneous speech in terms of glottalization and its prosodic functions. Using read speech allowed us to directly compare identical passages from the same speakers, which allowed for more direct comparisons and provides stronger evidence for individual differences. However, exploration of variation in glottalization across prosodic positions and speaking styles in spontaneous speech may uncover additional effects that were not observed with the read speech materials in this study.

5. ACKNOWLEDGMENTS

This work was partially supported by an Undergraduate Research Scholarship from the College of Arts and Sciences at Ohio State University and the National Science Foundation (BCS-1056409).

6. REFERENCES

[1] Beckman, M. E., Ayers, G. 1997. Guidelines for ToBI labelling. The OSU Research Foundation, 3.

[2] Blankenship, B. 1997. *The time course of breathiness and laryngealization in vowels*. M.A. Thesis, UCLA.

[3] Boersma, P. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5, 341-345.

[4] Dilley, L., Shattuck-Hufnagel, S., Ostendorf, M. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24, 423-444.

[5] Garellek, M. 2014. Voice quality strengthening and glottalization. *Journal of Phonetics* 45, 106-113.

[6] Henton, C., Bladon, A. 1988, Creak as a sociophonetic marker. In: Hyman, L.M. (ed.), *Language, Speech and Mind: Studies in Honour of Victoria A. Fromkin*. London: Routledge, 3-29.

[7] Kreiman, J., Sidtis, D. 2011. *Foundations of Voice Studies*. Malden, MA: Wiley-Blackwell.

[8] Ash, S. 2006. What is dialect? Revisiting the Midland. In: Murray, T.E., Beth, E.L. (eds.), *The North American Midland as a Dialect Area.* Philadelphia: John Benjamins, 33-56.

[9] Laver, J. 1980. *The Phonetic Description of Voice Quality*. Cambridge Cambridge University Press.

[10] Pierrehumbert, J., & Talkin, D. 1992. Lenition of /h/ and glottal stop. In: Doherty, G, Ladd, D.R. (eds.) *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, Cambridge : Cambridge University Press, 90-117.

[11] Redi, L., Shattuck-Hufnagel, S. 2001 Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29, 407-429.

[12] Shockey, L. 2014. A personal note on the larynx as articulator in English. In: Waniek-Klimczak, E., Pawlak, M. (eds.), *Teaching and Researching the Pronunciation of English.* Heidelberg: Springer, 219-222.

[13] Smiljanic, R., Bradlow, A.R. 2009. Speaking and hearing clearly: Talker and listener factors in speaking style changes. *Language and Linguistics Compasss* 3, 236-264.

[14] Yuan, J., Liberman, M. 2008. Speaker identification on the SCOTUS corpus, *Proceedings of Meetings on Acoustics 2008*, 5687-5690.