

# CODA GLOTTALIZATION IN AMERICAN ENGLISH

Scott Seyfarth and Marc Garellek

University of California, San Diego

sseyfarth@ucsd.edu, mgarellek@ucsd.edu

## ABSTRACT

Glottalization of coda /t, p/ is a common process in American English. This study uses acoustic measures to determine when coda glottalization occurs in the conversational speech of the Buckeye Corpus. Vowels preceding coda /t, p/ tokens for 40 speakers were analyzed using H1\*–H2\*, an acoustic correlate of glottal constriction. Results indicate that coda glottalization is more common before a sonorant, and this effect is still found phrase-finally, even when phrasal creak is taken into account. Nonetheless, the process occurs in other environments. While we conclude that coda glottalization may occur to enhance the voicelessness of coda /t/ before sonorants [20, 28], we argue that this cannot fully explain the phenomenon.

**Keywords:** glottalization, coda stops, voice quality

## 1. INTRODUCTION

In American English, /t/ and /p/ may undergo coda glottalization [19, 20, 7, 12, 29], which refers to either *glottal reinforcement* or *glottal replacement* in coda position. Glottal reinforcement occurs when the oral gestures for coda /t, p/ are produced with simultaneous glottal constriction: [ʔt, ʔp]. Glottal replacement occurs when the oral gestures for coda /t/ are replaced by glottal constriction: [ʔ] or constricted voicing. In American English, glottal reinforcement is found for both /t/ and /p/ in coda position, whereas glottal replacement is attested only for /t/. Coda glottalization occurs in other varieties of English [25, 5, 17, 1], as well as in other languages [7]. For example, glottal replacement is common in German [14] and glottal reinforcement is obligatory in many East and Southeast Asian languages [18, 6].

Previous work has shown that coda glottalization in American English is more common for /t/ than for /p/, and (phrase-medially) is more likely to occur when the coda precedes a sonorant, such as in ‘gate number’ [19, 20, 12]. It has been proposed that glottalization thus serves as an enhancement of coda voicelessness [28], particularly to prevent voicing from spreading from the following segment [20]. Glottalization weakens or prohibits voicing through

increased vocal fold constriction or sustained closure. Yet phrase-finally, coda glottalization rates generally increase [12]. This challenges the enhancement account, since there is often no coarticulatory voicing at a phrase boundary.

However, these findings are based on read speech from two to six speakers, using a relatively narrow selection of segment and word types. Prior identification of coda glottalization has also relied on visual inspection of voicing periodicity in waveforms. Yet there is no straightforward relationship between coda glottalization and irregular glottal pulses: coda glottalization by definition involves increased glottal constriction, but not all glottal constriction causes visibly-irregular voicing [3]. Further, irregular voicing is not always due to increased glottal constriction preceding coda stops. For example, in American English, vowels can be glottalized when word-initial and stressed [21, 4, 8], and creaky voice is also used to mark ends of phrases and index social identity [15, 9, 24, 23].

Thus, the goals of the present study are:

- (a) to use an acoustic measure that captures increased glottal constriction to test whether prior findings on coda glottalization in American English hold within a large corpus of spontaneous speech;
- (b) to evaluate phrase-final rates of coda glottal replacement, independently of irregular voicing;
- (c) to generate hypotheses for why coda glottalization occurs in American English.

## 2. CORPUS DATA

The data in this study come from the Buckeye Corpus of spontaneous American English [22], which consists of recorded speech of 40 adults (20 male, 20 female) from Ohio. The recordings took place in a quiet room and were digitized at 16 kHz with 16-bit resolution. The corpus contains both canonical phonemic and close phonetic transcriptions for each word. The canonical transcriptions were generated by automatic alignment software, which were then hand-corrected by corpus annotators to create the close phonetic transcriptions and segmentation.

Words with syllable-final /t, p/ in both the canonical form and the close transcription were ex-

tracted from the corpus. Syllabification was adapted from [10]. Complex codas (‘kept’, ‘rant’) and phonetically-voiced stops (/t/ realized as [ɾ]) were excluded. For complex codas with two stops, glottalization cannot be attributed to a single stop; for complex codas with sonorant-stop sequences, the presence of a sonorant could affect the acoustic measures of voice quality (see §3.1). Further, because acoustic glottalization was measured on the vowel preceding /t, p/, we excluded tokens where the target vowel was < 50 ms, as short samples are problematic for voice analyses. We also excluded very long vowels (> 300 ms) and stops (> 150 ms), which are likely to be hesitations or disfluencies.

### 3. STUDY 1: GLOTTALIZATION ENVIRONMENTS

The first goal was to replicate previous findings, using acoustic measures, regarding the environments that trigger coda glottalization. The primary variable of interest in these analyses is the segment following the coda stop. Phrase-medially, glottalization is reported to be more prevalent before sonorants than obstruents and is predicted to be more prevalent before voiced obstruents than voiceless ones [20, 12]. This pattern should be observed if glottalization is being used as a strategy to prevent voiceless codas from undergoing coarticulatory voicing.

#### 3.1. Methods

Glottalization was measured using  $H1^*-H2^*$ , which is the difference in amplitude between the first and second harmonics, corrected for formant frequencies and bandwidths. Lower values of  $H1^*-H2^*$  are correlated with increased glottal constriction [11, 26, 16], meaning that this acoustic measure should reflect coda glottalization even in the absence of visually-irregular voicing.  $H1^*-H2^*$  was measured on the vowel preceding each coda stop in our dataset using VoiceSauce [27], and averaged over the target vowel’s entire duration.

For this analysis, we took out 6224 /t/ tokens transcribed as [ʔ], since the segment boundaries for [ʔ] were coextensive with regions of glottalization. These tokens were analyzed separately as described in §3.1.1, but we found that all results here are qualitatively the same with these tokens included.

Since the accuracy of the measurements depends on correct  $f_0$  and formant tracking, we took several steps to identify mistracked tokens. We excluded tokens if  $f_0$  increased by more than 100 Hz (for women) or 50 Hz (for men) during the vowel, or if the mean  $f_0$  was > 2.5 SD from each speaker’s

mean, as these suggest erroneous octave jumps. We also plotted F1–F2 distributions for each vowel type by sex, removed tokens > 2.5 SD from each category mean, and additionally hand-pruned a total of 104 vowels that were likely mistracked.

Finally, we removed tokens whose  $H1^*-H2^*$  measurements were > 2.5 SD from the global mean. Of the 5415 remaining tokens, 1549 were followed by a breath, silence, laugh, or other non-speech noise, and were removed from the present analysis, in order to exclude tokens that are likely phrase-final and where there is no immediately-following segment. In total, 3866 tokens and 274 word types are included in the analysis of  $H1^*-H2^*$ .

#### 3.1.1. Glottal stops

As a secondary measure, we examined separately the rate of glottal replacement for /t/ codas, which was identified using the glottal stop annotations in the close phonetic transcriptions. Hand inspection suggests that glottal stop annotations were largely accurate: we inspected 1824 tokens labeled as having a glottal stop, and identified glottal stops as having no [t] formant transitions or release burst and irregular voicing localized to the onset and offset of the target stop. Of these 1824 tokens, we revised only 62 annotations to glottal-reinforced [ʔ̥] because of the presence of a [t] release burst.

This dataset included only /t/ codas, since replacement is not attested in American English for /p/. Likely due to the segmentation strategy noted above,  $H1^*-H2^*$  was not correlated with replacement by [ʔ] ( $r_{pb} = -0.06$ ). Of the 11,594 /t/ tokens in the dataset, 3762 were followed by a silence or non-speech sound and removed from this analysis, leaving 7832 tokens and 312 word types.

#### 3.2. Models

Using *lme4* [2], we fit a linear mixed-effects model to the  $H1^*-H2^*$  data described in §3.1, and a logistic model to the glottal stop data in §3.1.1. The variable of interest—the segment following the stop—was Helmert-coded to test two contrasts: voiced vs. voiceless obstruents, and obstruents vs. sonorants.

As control variables, we included three factors that could influence voice quality: absence of a syllable onset (favoring word-initial glottalization), position of the syllable in the word, and phrasal creak, which was identified using the ‘creaky voice’ labels provided by corpus annotators. However, when this label was applied only to a local region around a target coda (less than twice the nucleus vowel duration), we treated the label as referring to coda glott-

talization rather than phrasal creak. For the model of  $H1^*-H2^*$ ,  $f_0$  was also included as a control. Both models included maximal converging per-speaker random effects. Also included were per-word intercepts and slopes for the two Helmert contrasts, and intercepts for the identity of the following segment.

### 3.3. Results

Figure 1 shows a summary of the data; fixed-effects estimates for the two models are shown in Table 1. In the model of  $H1^*-H2^*$ , there is a significant decrease in the measure (i.e., more glottal constriction) when a sonorant follows the coda stop, replicating the findings in [20, 12].  $H1^*-H2^*$  is not significantly lower when the coda stop is followed by a voiced obstruent, relative to a voiceless one.

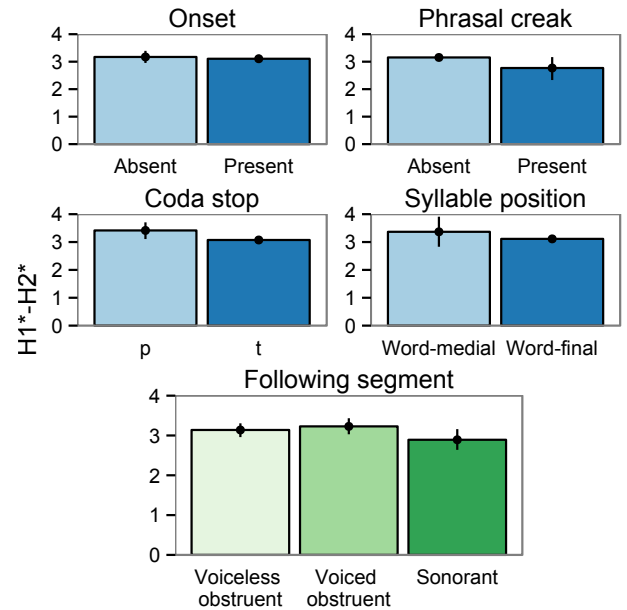
The model of glottal stop replacement showed the same pattern: glottal replacement of coda /t/ is significantly more likely when there is a following sonorant, but there was not a significant difference between voiced and voiceless obstruents.

For the most part, the control variables patterned in expected directions, as shown in Figure 1, although in the model they were non-significant (stop place was marginal,  $p < 0.07$ ). This may be because there was insufficient variation in the controls: for example, less than 4% of tokens were word-medial (about 75% of the data are monosyllabic function words).

	$H1^*-H2^*$ (dB z)	/t/ → [ʔ]
Intercept	0.072 (0.096)	0.465 (0.230)
Voiced vs. voiceless obstruent (Helmert comparison)	0.023 (0.019)	-0.047 (0.168)
Sonorant vs. obstruent (Helmert comparison)	-0.041** (0.014)	0.821*** (0.112)
Onset (present: -1, absent: 1)	0.007 (0.021)	0.006 (0.112)
Phrasal creak (present: -1, absent: 1)	0.008 (0.037)	-0.352*** (0.069)
Coda syllable position (word-final: -1, medial: 1)	0.020 (0.031)	-0.098 (0.110)
Coronal vs. labial coda stop (p: -1, t: 1)	-0.050 (0.026)	—
$f_0$ (Hz, z-score)	0.351*** (0.048)	—

\*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$  (adj. for multiple tests)

**Table 1:** Model estimates for Study 1 (SEs below in parentheses). Italics show units or coding scheme.

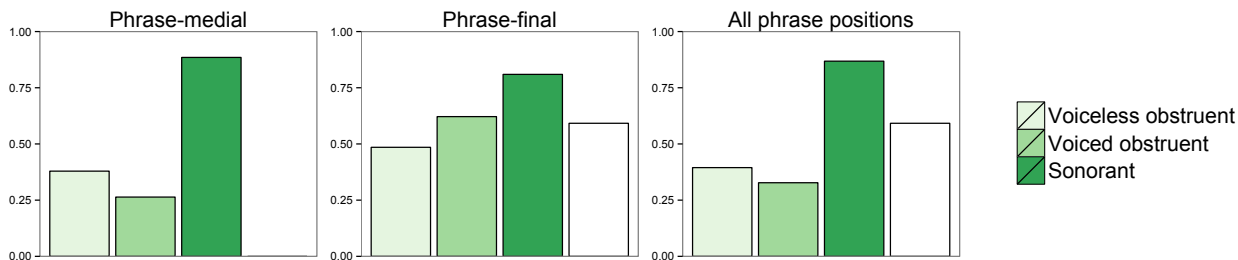


**Figure 1:** Means of  $H1^*-H2^*$  measurements by variable. Error bars show bootstrapped 95% CIs.

## 4. STUDY 2: GLOTTAL REPLACEMENT BY PHRASE POSITION

Study 1 supported prior work showing that coda glottalization (as both glottal constriction and glottal replacement) is more prevalent preceding a sonorant. However, prior work has also reported high rates of coda glottalization phrase-finally. Why might coda glottalization be more common both when the following sound is a sonorant and at the ends of phrases? If coda glottalization occurs primarily to enhance voicelessness of the coda stop (by preventing sonorant voicing from spreading [20]), then it is unclear why coda glottalization rates would also increase phrase-finally, especially for utterance-final phrases that are not followed by a speech sound.

On the other hand, it is also plausible that phrase-final coda glottalization is mainly an expression of phrasal creak, which is common at the ends of prosodic phrases in American English [15, 24]. Study 2 thus examines the effect of phrase position only on glottal replacement, where the presence of glottalization is probably not a result of mistaking phrasal creak for coda glottalization. Unlike glottal reinforcement, glottal replacement involves a glottal stop with no coronal formant transitions during the preceding vowel and no [t] release. Thus, even if a vowel before coda /t/ is creaky because of phrasal creak rather than coda glottalization, we would still expect to see both formant transitions and a [t] release. In this study, we also attempt to control for the presence of phrasal creak, as defined in §3.2.



**Figure 2:** Proportion of /t/ codas realized as [ʔ], by type of following segment. White bars show the proportion realized as [ʔ] when the /t/ was not followed by a speech sound; all of these codas were phrase-final.

#### 4.1. Coding of phrasal position

Five coders noted whether the target word with a coda stop was phrase-final (at the end of an intermediate or full intonational phrase). The end of a phrase was identified by lengthening of the phrase-final vowel and/or by the presence of a following phrase accent, pause, breath, silence, or disfluency.

This study reports a preliminary analysis of 6347 /t/ words that have been hand-coded for phrasal position; annotation of the remaining words is currently underway. Figure 2 shows the proportion of /t/ codas that were replaced with glottal stop, by phrase position and the following segment type.

#### 4.2. Model and Results

A logistic mixed-effects model was fit to the annotated glottal replacement data using the procedure, variables, and coding in §3.2, with two differences. First, phrasal position was added as a variable, including interactions with the type of segment following the coda. Second, the following segment variable was re-coded in the model to test three contrasts, based on visualizing the data (Figure 2): obstruents vs. sonorants, obstruents vs. utterance-final tokens (those not followed by a speech sound), and voiced vs. voiceless obstruents.

In Study 2, there was a marginal overall effect of phrase-final position ( $\beta = 0.20$ ,  $p < 0.07$ ). As in Study 1, there was more replacement before sonorants than obstruents, both phrase-medially ( $\beta = 1.51$ ,  $p < 0.001$ ) and phrase-finally ( $\beta = 0.83$ ,  $p < 0.001$ ). The sonorant effect was significantly smaller phrase-finally ( $p < 0.001$ ).

There was no significant difference between voiced and voiceless obstruents in either phrase position. However, phrase-finally, the replacement rate increased more before voiced obstruents than voiceless ones ( $\beta = 0.29$ ,  $p < 0.01$ ). Utterance-final tokens (those not followed by a speech sound) did not undergo more replacement than phrase-final tokens preceding obstruents ( $p > 0.5$ ).

## 5. GENERAL DISCUSSION

This study tested whether prior findings on coda glottalization in American English hold across a diverse selection of phonological environments and a larger number of speakers, using an acoustic correlate of glottal constriction,  $H1^* - H2^*$ , as well as codings for [ʔ]. We find that coda glottalization is more common predominantly before sonorants, confirming previous findings [20, 12]. On an annotated subset of the corpus, we also find that the sonorant effect still exists phrase-finally, contra [12]. These results support a glottalized allophone of coda /t/ before sonorants (regardless of phrasal position), whose precise phonetic articulation ranges from glottal-reinforced [ʔt̚] to a glottal stop [ʔ]. This stems from our finding that both glottal reinforcement and glottal replacement are more common before sonorants.

Our results support the claim that coda glottalization occurs when it helps to prevent coarticulatory anticipatory voicing from a following sonorant [20]. Although coda glottalization rates were not found to be different preceding voiced and voiceless obstruents, voicing is relatively weak for English obstruents. Thus, it may be that there is less need to prevent anticipatory voicing in codas before voiced obstruents relative to those before sonorants.

Utterance-finally, when there is no following sound, glottal replacement nonetheless occurs over 50% of the time. Additionally, replacement rates before obstruents are somewhat higher phrase-finally than medially. This suggests that there are other considerations that trigger coda glottalization beyond a need to prevent coarticulatory voicing, which would not be present in final position. For example, it is possible that glottalization serves to enhance the voiceless/voiced distinction more generally. Thus, it may enhance the relative percept of voicelessness in final position, where cues to voicedness are weakened and less reliable [13].

Nonetheless, even phrase-medially before voiceless sounds, glottalization is hardly rare (Figure 2). Further work is needed to better understand why it occurs across such a wide variety of environments.

## 6. REFERENCES

- [1] Ashby, M., Przedlacka, J. 2014. Measuring incompleteness: Acoustic correlates of glottal articulations. *Journal of the International Phonetic Association* 44, 283–296.
- [2] Bates, D., Maechler, M., Bolker, B., Walker, S. 2014. *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-7.
- [3] Blankenship, B. 2002. The timing of nonmodal phonation in vowels. *Journal of Phonetics* 30, 163–191.
- [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] Docherty, G., Foulkes, P. 1999. Sociophonetic variation in glottals in Newcastle English. *Proceedings of the International Congress of Phonetic Sciences* San Francisco. 1037–1040.
- [6] Edmondson, J. A., Chang, Y., Hsieh, F., Huang, H. J. 2011. Reinforcing voiceless finals in Taiwanese and Hakka: Laryngoscopic case studies. *Proceedings of the 17th International Congress of Phonetic Sciences* Hong Kong.
- [7] Esling, J. H., Fraser, K. E., Harris, J. G. 2005. Glottal stop, glottalized resonants, and pharyngeals: A reinterpretation with evidence from a laryngoscopic study of Nuuchahnulth (Nootka). *Journal of Phonetics* 33, 383–410.
- [8] Garellek, M. 2014. Voice quality strengthening and glottalization. *Journal of Phonetics* 45, 106–113.
- [9] Garellek, M. 2015. Perception of glottalization and phrase-final creak. *Journal of the Acoustical Society of America* 137, 822–831.
- [10] Gorman, K. 2013. *Generative phonotactics*. PhD thesis University of Pennsylvania.
- [11] Holmberg, E. B., Hillman, R. E., Perkell, J. S., Guiod, P., Goldman, S. L. 1995. Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech and Hearing Research* 38, 1212–1223.
- [12] Huffman, M. K. 2005. Segmental and prosodic effects on coda glottalization. *Journal of Phonetics* 33, 335–362.
- [13] Keyser, S. J., Stevens, K. N. 2006. Enhancement and Overlap in the Speech Chain. *Language* 82(1), 33–63.
- [14] Kohler, K. J. 1994. Glottal stops and glottalization in German. Data and theory of connected speech processes. *Phonetica* 51, 38–51.
- [15] Kreiman, J. 1982. Perception of sentence and paragraph boundaries in natural conversation. *Journal of Phonetics* 10, 163–175.
- [16] Kreiman, J., Shue, Y.-L., Chen, G., Iseli, M., Gerratt, B. R., Neubauer, J., Alwan, A. 2012. Variability in the relationships among voice quality, harmonic amplitudes, open quotient, and glottal area waveform shape in sustained phonation. *Journal of the Acoustical Society of America* 132, 2625–2632.
- [17] Mees, I., Collins, B. 1999. Cardiff: a real-time study of glottalisation. In: Foulkes, P., Docherty, G., (eds), *Urban Voices: Accent Studies in the British Isles*. London: Arnold 85–202.
- [18] Michaud, A. 2004. Final consonants and glottalization: New perspectives from Hanoi Vietnamese. *Phonetica* 61, 119–146.
- [19] Pierrehumbert, J. 1994. Knowledge of variation. *Papers from the parasession on variation, 30th meeting of the Chicago Linguistic Society* Chicago. Chicago Linguistic Society 232–256.
- [20] Pierrehumbert, J. 1995. Prosodic effects on glottal allophones. In: Fujimura, O., Hirano, M., (eds), *Vocal fold physiology: voice quality control*. San Diego: Singular Publishing Group 39–60.
- [21] Pierrehumbert, J., Talkin, D. 1992. Lenition of /h/ and glottal stop. In: Docherty, G. J., Ladd, D. R., (eds), *Papers in Laboratory Phonology II*. Cambridge: Cambridge University Press 90–117.
- [22] Pitt, M. A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E., Fosler-Lussier, E. 2007. *Buckeye Corpus of Conversational Speech (2nd release)*. Department of Psychology, Ohio State University Columbus, OH.
- [23] Podesva, R. J., Callier, P. 2015. Voice quality and identity. *Annual Review of Applied Linguistics* 35, 173–194.
- [24] Redi, L., Shattuck-Hufnagel, S. 2001. Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29, 407–429.
- [25] Roach, P. J. 1973. Glottalization of English /p/, /t/, /k/ and /tʃ/ – a re-examination. *Journal of the International Phonetic Association* 3, 10–21.
- [26] Samlan, R. A., Story, B. H. 2011. Relation of structural and vibratory kinematics of the vocal folds to two acoustic measures of breathy voice based on computational modeling. *Journal of Speech, Language, and Hearing Research* 54, 1267–1283.
- [27] Shue, Y.-L., Keating, P. A., Vicenik, C., Yu, K. 2011. VoiceSauce: A program for voice analysis. *Proceedings of the International Congress of Phonetic Sciences* Hong Kong. 1846–1849.
- [28] Stevens, K., Keyser, S. J. 1989. Primary features and their enhancement in consonants. *Language* 65(1), 86–106.
- [29] Sumner, M., Samuel, A. G. 2005. Perception and representation of regular variation: The case of final /t/. *Journal of Memory and Language* 52, 322–338.