UCLA Working Papers in Phonetics, No. 109, pp. 31-39

The benefits of vowel laryngealization on the perception of coda stops in English

Marc Garellek

(marcgarellek@ucla.edu)

Abstract

It has been previously noted that voiceless coda stops in English may undergo optional glottalization, i.e. they are produced with simultaneous glottal closure. The glottal closure usually produces laryngeal coarticulation on the previous vowel in the form of laryngealization (creaky voice). In this paper, the effects of vowel laryngealization on coda stop perception were investigated. Eighteen native speakers of English participated in a phoneme monitoring task where they were asked to monitor for /t/. The target stimuli were English monosyllabic words ending in a coda /t/, e.g. 'beat.' The stimuli differed according to two conditions: whether vowel was either modal or laryngealized, and whether the coda /t/ was either released or unreleased. The results show that presence of laryngealization resulted in faster and more accurate monitoring of /t/ in codas. Further analysis of the filler stimuli suggests that the perceptual advantage of laryngealization is a result of listeners' linguistic experience with glottalization and because laryngealization is useful for retrieving formant transition cues.

1 Introduction

Voiceless coda stops in English may undergo optional glottalization, i.e. produced with simultaneous glottal closure. This phenomenon is sometimes referred to as *glottal reinforcement* (Higginbottom 1964) or as preglottalization (Esling et al. 2005) and is known to occur in many other languages, especially East Asian languages with unreleased coda stops like Thai (Harris 2001). Previous studies on glottal reinforcement in English have shown that it is more likely to occur for /t/ and /p/ than for /k/, and that various segmental and prosodic factors influence its occurrence (Pierrehumbert 1995; Huffman 2005). For example, glottal reinforcement tends to be more common before a sonorant-initial word, and may be more likely to occur utterance-finally. The glottal closure for the glottalized stop usually produces laryngeal coarticulation in the form of laryngealization (creaky voice) on the previous vowel.

The reasons why glottal reinforcement occurs are still unclear. Pierrehumbert (1995) offered an enhancement-based explanation of why laryngealization derived from glottal reinforcement is more likely in some environments than in others. According to this view, the spectral characteristics of laryngealization would enhance the stop's difference from other sounds like nasals. This could account for why laryngealization tends to be more common before sonorants, which have their own spectral characteristics.

Huffman (2005) offered instead a coarticulatory account, whereby voicing in sonorants leads to a weakening of the glottal closure gesture associated with the coda stop, yielding laryngealization.

However, it is still unclear why a glottal closure gesture should ever come to be associated with coda stops in the first case. Glottal closure may be used to ensure a stop's voiceless characteristics, but many languages with glottalized codas like Thai do not have a voicing distinction in coda stops. Another hypothesis is that glottalization cooccurs with the loss of an audible release in coda stops. The languages with glottal reinforcement tend to have coda stops with no audible release. Glottal reinforcement in English seems to be more likely to occur with unreleased codas, though I am unaware of any literature showing the cooccurrence of these two phenomena for the language. The loss of stop release may harm perception, because the spectral properties of a stop's burst provide perceptual cues to the place of articulation of the stop (Winitz et al. 1972; Stevens & Blumstein 1978; Repp & Lin 1989; Alwan et al. 2010). Thus, perhaps glottalization occurs as a means of enhancing what cues are left to a stop's place of articulation when burst information is lost. Laryngealized phonation shows an increase in energy of higher frequencies compared to modal phonation, and laryngealized phonation consequently shows higher amplitudes of the first, second, and third formants than modal phonation, even when normalized to the amplitude of the first harmonic (Gordon & Ladefoged 2001). Formant transitions are a known perceptual cue to a stop's place of articulation in both onset and coda positions (Liberman et al. 1954; Delattre et al. 1955; Nguyen et al. 2009; Alwan et al. 2010), so the increased energy of those formants may be advantageous to the listener, especially when no burst cues are available. Further support for the potential beneficial nature of coda glottalization comes from production data on English, where more confusable words with voiceless coda stops were found to be produced with more laryngealization than less confusable ones (Garellek 2011). This suggests that laryngeal coarticulation from the glottalized stop increases in harder words. Thus, the current study will specifically test the hypothesis that laryngealization may aid the listener's perception of coda stops.

2 Method

A phoneme monitoring task (Connine & Titone 1996) was created, whereby participants were asked to press a computer key as soon as they heard /t/. The experiment was implemented in Matlab, and was carried out in a soundbooth at UCLA. After presented with instructions, participants had a practice round where they were asked to press a key as soon as they heard /t/. The instruction phase provided feedback, such that participants were informed whether they were correct in pressing or not pressing the key. This was designed so that participants would know to press the key when they heard even an unreleased /t/. After the instruction round came the actual experiment. As in the instruction round, participants were asked to press the key as soon as they heard a /t/. The stimuli played automatically with an inter-stimulus interval of one second. At the midpoint of the experiment, participants were able to take a break. The testing period lasted 20 to 30 minutes.

The stimuli, which were produced by a phonetically-trained native English speaker, consisted of 240 tokens of monosyllabic English words. The 240 tokens were repeated in a second block, for a total of 480 tokens. The target stimuli consisted of English words with long nuclei /i, eɪ, aʊ, aɪ, oʊ, u/ and a coda /t/, e.g. /bit/. Eighteen target stimuli were repeated twice, for a total of 36 targets per block. Additionally, fillers beginning with a /t/ in onset position were included so that listeners would be exposed to /t/ in a variety of syllable positions. The other filler types included words ending in a coda /p/ or /k/, words ending in an alveolar, e.g. /bin, biz/, and words without codas, e.g. /bi, peɪ/. Every word appeared minimally twice per block: once with a modal vowel, and once more with a laryngealized vowel. For each phonation type (modal vs. largyngealized), the words that ended in a coda stop /p, t, k/ also appeared with both released and unreleased codas. The unreleased tokens were created by cropping the sound file right before the release of the coda stop. Therefore, unreleased and released tokens were identical except for the presence of a stop release at the end (and thus also differed in duration). The phonetic outcomes for the sample target 'beat' are shown in IPA in Table 1, and the complete list of stimuli can be found in the Appendix.

Table 1. IPA transcriptions for a sample target word 'beat,' for the four levels of interest.

	Modal	Laryngealized
Unreleased	bit̚	bḭt̚
Released	bit	bḭt

The stimuli were analyzed acoustically using VoiceSauce (Shue et al. 2011). Generally, laryngealized tokens were longer than modal ones**.** Crucially, they did not differ statistically in F0 from modal tokens, but the amplitudes of F1, F2, and F3 were higher (as evidenced by lower values of H1*-A1*, H1*-A2*, and H1*-A3*) for laryngealized tokens than for modal ones.

In total, 18 native English speakers (nine female and nine male) participated in the experiment. All were American undergraduate students enrolled at UCLA, two of whom had taken an introductory course in linguistics. Their mean age was 20, and they received extra credit for their participation.

3 Results

3.1. Target stimuli ending in coda /t/

Response time (RT) was measured from the onset of the stimulus, and accuracy was assessed by whether the participants responded within 1000 ms of the onset of the stimulus. The results are separated according to the release of the coda /t/ and voice quality of the preceding vowel. The RT results were analyzed using a linear mixed-effects model, with RT as the dependent variable, coda release and voice quality as fixed effects, and subject and item as random effects. No other factors like word frequency, vowel quality, and stimulus duration were found to improve the model's fit significantly. Thus, although laryngealized and unreleased words were shorter than modal and released ones, respectively, the inclusion of duration as a variable in the statistical model did not yield a better fit to the results. Accuracy was assessed using a logistic mixed-effects model, with accuracy as dependent variable. The models were run in R using the *lme4* package and p-values were calculated using the *pvals.fnc* function with 10,000 simulations **(**Baayen et al. 2008).

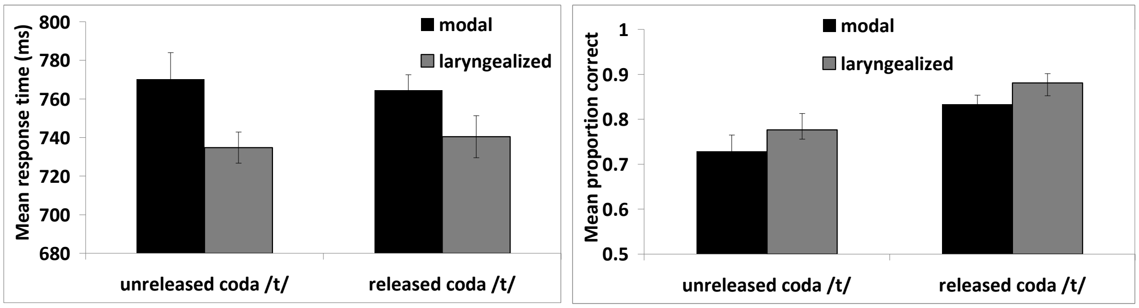


Figure 1. Mean response time and proportion correct of /t/ identification in targets. Error bars show standard error values.

The response time (RT) and accuracy results (in terms of proportion correct) for the target stimuli ending in coda /t/ are shown in Figure 1. The release of the coda /t/ did not result in faster results even though unreleased tokens were longer. However, laryngealized tokens showed faster response times for both unreleased and released tokens ($p < 0.0001$).

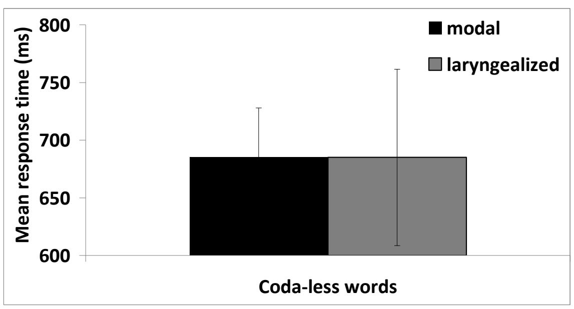
The accuracy data were analyzed in terms of the proportion correct. Listeners were more accurate for released words than for unreleased words ($p < 0.0001$). Moreover, for both unreleased and released tokens, laryngealization resulted in more accurate responses ($p = 0.0363$). In sum, both the RT and accuracy results suggest that laryngealized voice quality does indeed help cue the listener to the presence of a /t/. The benefit of laryngealization is found not only for unreleased tokens but also for fully released tokens. Subsequent analyses will therefore average over all modal vs. laryngealized tokens, regardless of the coda's release.

3.2. Words ending in coda /p/ and /k/

If laryngealization helps in the detection of codas by increasing the energy of the formant transitions into the coda stop, then I would expect the accuracy for words ending in coda /p/ and /k/ to improve under laryngealization as well. The results however show a decrease in accuracy for words ending in /p/ or /k/ when the previous vowel is laryngealized ($p < 0.001$ for both /p/ and /k/). The reasons for this are still unclear. All words in the experiment had modal and laryngealized versions, so listeners were not expected to associate voice quality changes with a given set of words. Nevertheless, the effect might be lexical: preglottalized /k/ is known to be less common in English (Pierrehumbert 1995), and although preglottalized /p/ is attested, it is undoubtedly less common than preglottalized /t/, given that the latter phoneme is more common than the former. Moreover, Huffman (2005) generally found higher rates of glottalization for coda /t/ than /p/. Thus, it is possible that English listeners have learnt to associate preglottalization (and through coarticulation, also vowel laryngealization) as a phenomenon occurring before instances of coda /t/ only. Thus, if laryngealization does help in coda /t/ perception, it is probably due to both phonetic and lexical reasons.

3.3. Coda-less words

Another set of interesting filler words were those ending in open syllables, for example 'bee' /bi/. In this experiment, all such words had minimal pairs with coda /t/ (and usually also with coda /p/ and /k/) which were also presented to the listeners. If listeners associate laryngealization with coda /t/, as the results from Section 3.2 suggest, then they should perform less accurately on open syllables when they are laryngealized. For example, listeners would be expected to false alarm more frequently on a laryngealized token of /bi/, thinking it was a token of /bit/. The results, however, show that while in general accuracy is quite high for these open syllables, with an overall proportion correct of over 0.95 for modal tokens, this number increases to over 0.98 for laryngealized tokens. Using a mixed-effects model to the previous ones except without the fixed effect of stop release, the difference in means was found to be significant at $p = 0.0193$. No difference in response time between modal and laryngealized tokens was found for coda-less words. The results are shown in Figure 2.



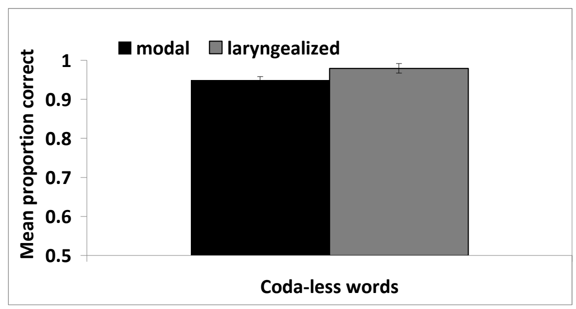


Figure 2. Response time and accuracy for modal vs. laryngealized tokens without codas. Error bars show standard error values.

Therefore, although laryngealization seems to trick listeners for words ending in coda /p/ and /k/, it helps them for coda-less words. Note that the generally high accuracy for coda-less words is likely due to the fact that these words had longer vowels than the closed-syllable words, as is common in English. Nevertheless, the beneficial effect of laryngealization in these cases is well explained by the fact that laryngealization amplifies the steady formants of such words.

3.4 Words with onset /t/

Another way to test whether the effect of laryngealization on coda /t/ monitoring is due to lexical effects alone is by analyzing the fillers that began with /t/. Recall that none of these words ended in a coda /t/, so any effect of laryngealized vowels in these words could only be due to the /t/ in onsets. The statistics for this subset of the data were obtained using a mixed-effects regression model identical to that from Section 3.3, except for the different data sets. The results, schematized in Figure 3, indicate that listeners were more accurate at monitoring for a /t/ when the following vowel is laryngealized (p =

0.0352). Listeners were also faster at perceiving onset /t/ when the following vowel was laryngealized ($p < 0.002$).

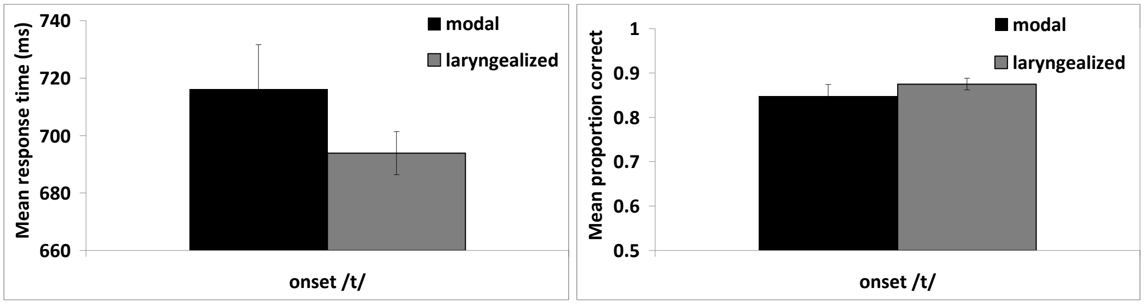


Figure 3. Response time and accuracy for modal vs. laryngealized tokens beginning with onset /t/. Error bars show standard error values.

This finding cannot be due to the lexical distribution of vowel laryngealization, which (when derived segmentally) only occurs *before* /t/, not *after*. Thus, this appears to be a phonetic rather than lexical effect.

4 Discussion

The present study provides evidence that laryngealization helps cue listeners to the presence of a /t/ in English. Listeners monitor for /t/ faster and more accurately when the target word contains a laryngealized vowel than when the vowel is modal. This effect is true even when the coda stop is fully released. As mentioned earlier, this effect could be due to two factors. The first is that in English, glottal reinforcement (and thus vowel laryngealization) is common with /t/, and more common than for /p/ and /k/. The second factor is acoustic: laryngealization results in more intense formants than modal phonation, thus potentially amplifying the vowel's formant transitions, which are a major cue to the place of articulation of the following stop. Although the results of this study cannot determine with certainty whether one factor is more important than the other, they suggest that both factors are at play. The role of the lexicon can explain why accuracy is lower for words ending in coda /p/ and /k/, because native listeners might be aware that laryngealization is most common before coda /t/ in English. Therefore, despite the fact that the preceding vowel's formants are amplified, giving stronger cues to the presence of a following /p/ or /k/, English listeners associate the laryngealization with a following /t/. However, the lexicon cannot account for the finding that open syllables were more accurately detected as *not* ending in /t/ when the vowels were laryngealized than when they were modal. Laryngealized tokens were shorter, so the shortened vowel could have confused listeners into thinking the word ended in a coda. But the fact that laryngealized vowels helped instead of confused the listeners suggests that the amplified energy in the higher frequencies cued them to the absence of formant transitions for /t/.

Lexical properties also cannot easily account for the fact that onset /t/ was detected faster and more accurately when the following vowel was laryngealized, because in English only coda /t/ is associated with laryngealization. Therefore, the results of this study suggest that laryngealization aids listeners due to both its lexical distribution and cross-linguistic phonetic properties. But whether one of these factors is more important than the other is left to further research.

This study provides support for the hypothesis that glottal reinforcement might be perceptually useful. The vowel laryngealization that is derived from coarticulation with the glottal closure is used by listeners to monitor for /t/. Thus, it is possible that glottal reinforcement cooccurs with the loss of audible release of coda stops in many languages in order to facilitate perception of the sounds' place of articulation. A corpus study investigating the cooccurence of preglottalization and lack of audible release would be particularly enlightening for this claim.

Lastly, this study provides further evidence that laryngealization should be viewed as primarily beneficial to the listener, in line with studies of tone recognition (Yu 2010; Yu & Lam 2011; Brunelle & Finkeldey 2011). It also helps explain why greater laryngealization was found before coda stops in English for words with lower relative frequencies (Garellek 2011). Given that laryngealization was found in the current study to be utilized by listeners to perceive /t/, greater laryngealization in harder words would amplify this perceptual cue.

5 Conclusion

In this study, the effects of laryngealization on coda stop perception were investigated using a phoneme monitoring task. Eighteen native speakers of English were asked to monitor for the sound /t/. The results showed that presence of laryngealization resulted in faster and more accurate monitoring of /t/ in codas, showing that laryngealization is a useful cue to /t/ detection. Investigation of filler words ending in coda /p/ and /k/ suggests that laryngealization is a cue to coda /t/ due to lexical properties of English, given that glottalized /t/ is more common than other glottalized stops.

However, the analysis of coda-less fillers and /t/-initial words suggests that laryngealization also helps cue listeners to /t/ by amplifying the energy in the formant transitions, because English listeners have no lexical experience with laryngealization in such words. Further research, possibly by presenting foreign words to English listeners, would be useful for determining whether laryngealization is beneficial for /t/ detection even for non-English words. The perceptual benefit of laryngealization might also explain why unreleased stops tend to occur with glottal reinforcement across languages, though this relationship merits further study.

Acknowledgements

I would like to thank Pat Keating and Megha Sundara for help with the early stages of this project, Henry Tehrani for help with the Matlab scripting, Robyn Orfitelli for help with the stimuli, and Chad Vicenik, Kaeli Ward, and Lauren Winans for useful comments during the piloting stage.

References

Alwan, A., J. Jiang, & W. Chen. (2010). Perception of place of articulation for plosives and fricatives in noise. *Speech Communication 53*, 195–209.

Baayen, R. H., D. J. Davidson, & D. M. Bates. (2008). Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language 59, 390-412.

Brunelle, M. & J. Finkeldey. (2011). Tone perception in Sgaw Karen. In *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS 17)*.

Connine, C. M. & D. Titone. (1996). Phoneme monitoring. *Language and Cognitive Processes 11*, 635–645.

Delattre, P. C., A. M. Liberman, & F. S. Cooper. (1955). Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America 27*, 769–773.

Esling, J. H., K. E. Fraser, & J. G. Harris. (2005). Glottal stop, glottalized resonants, and pharyngeals: A reinterpretation with evidence from a laryngoscopic study of Nuuchahnulth (Nootka). *Journal of Phonetics 33*, 383–410.

Garellek, M. (2011). Lexical effects on English vowel laryngealization. In *Proceedings of the International Congress of Phonetic Sciences (ICPhS 17)*, Hong Kong.

Gordon, M. & P. Ladefoged (2001). Phonation types: a cross-linguistic overview. *Journal of Phonetics 29*, 383–406.

Harris, J. G. (2001). States of the glottis of Thai voiceless stops and affricates. In K. Tingsabadh & A. S. Abramson (Eds.), *Essays in Tai linguistics*, pp. 3–11. Bangkok: Chulalongkorn University Press.

Higginbottom, E. (1964). Glottal reinforcement in English. *Transactions of the Philological Society 63*, 129–142.

Huffman, M. K. (2005). Segmental and prosodic effects on coda glottalization. *Journal of Phonetics 33*, 335–362.

Liberman, A. M., P. C. Delattre, F. S. Cooper, & L. J. Gerstman. (1954). The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs: General and Applied 64*, 1–13.

Nguyen, V. S., E. Castelli, & R. Carré. (2009). Vietnamese final stop consonants /p, t, k/ described in terms of formant transition slopes. In *2009 International Conference on Asian Language Processing*, pp. 86–90. IEEE.

Pierrehumbert, J. (1995). Prosodic effects on glottal allophones. In O. Fujimura and M. Hirano (Eds.), *Vocal fold physiology: voice quality control*, pp. 39–60. San Diego: Singular Publishing Group.

R Development Core Team. 2007. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. http://www.R-project.org (14 January 2009)

Repp, B. H. & H.-B. Lin. (1989). Acoustic properties and perception of stop consonant release transients. *Journal of the Acoustical Society of America 85*, 379–396.

Shue, Y.-L., P. A. Keating, C. Vicenik, & K. Yu. (2011). VoiceSauce: A program for voice analysis. In *Proceedings of the International Congress of Phonetic Sciences (ICPhS 17)*, Hong Kong.

Stevens, K. N. & S. E. Blumstein. (1978). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America 64*, 1358–1368.

Winitz, H., M. E. Scheib, & J. A. Reeds. (1972). Identification of stops and vowels for the burst portion of /p, t, k/ isolated from conversational speech. *Journal of the Acoustical Society of America 51*, 1309–1317.
Yu, K. M. (2010). Laryngealization and features for Chinese tonal recognition. In *Proceedings of Interspeech 2010*.
Yu, K. M. & H. W. Lam. (2011). The role of creaky voice in Cantonese tonal perception. In *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS 17)*.

Appendix: Wordlist

bait
bay
be
beak
bean
beat
beep
bite
boat
boo
boot
bout
bow
cane
cape
coat
date
day
dean
deep
dote
doubt
dough
gain
gape
gate
gay
go
goat
gout
Kate
kay
keep
key
kite
pain
pate
pay
peat
pee
peek
pout
take
tale
tame
tape
taupe
teak
tease
tide
ties
tile
toes
toil
toll
tome
tool
town
tyke
tine
type