Variable realization of the Arapaho glottal stop, despite its being distinctive and frequent

D. H. Whalen, Christian T. DiCanio and Wei-Rong Chen

City University of New York/Yale Child Study Center; University at Buffalo; Yale Child Study Center
douglas.whalen@yale.edu; cdicanio@buffalo.edu; wei-rong.chen@yale.edu

ABSTRACT

Complete glottal closure is typically treated as the canonical realization of a glottal stop, but it is quite unusual in running speech [1, 2]. Further, non-phonemic glottal stops are quite variable [3-5]. Even languages with contrastive glottal stops exhibit extensive variation [6-11]. Glottal stops are both common and contrastive in Arapaho (ISO 639 arp): Do they therefore have more canonical phonetic realizations? Glottal stops in an Arapaho corpus [12] were classified by phonetic realization and examined for duration and harmonics ratio (HR). Full glottal stops were rare (31.6%) and instead often realized as glottalization. Duration was longer for stops than for other realizations (but shorter than that of /t/). HR differed across all realizations. Arapaho reflects the pattern of other languages in using few stop realizations for the common, phonemic glottal stop, but may do so more often than reported for other languages.

Keywords: Glottal stop; Arapaho; variability; duration; harmonics ratio.

1. INTRODUCTION

Complete closure of the glottis is assumed to be the canonical realization of glottal stop [7]. Yet glottal stop realization has been found to be quite variable across many languages (e.g. [1, 10]). Glottal stops are found, at least allophonically, in 1131 of 3020 phonological inventories listed under PHOIBLE (37.5%) [13], but they are non-distinctive in other languages like English and Spanish. Does the presence of a phonemic contrast among oral stops and a glottal stop in a language predict the glottal stop's manner of realization? Some work assumes this to be true [14], but it has not been explicitly tested for many languages. The current study examines the variation in the realization of contrastive glottal stops in Arapaho as a test case.

While there is a substantial literature on variation in the production of non-modal phonation (for recent reviews, see [15, 16]), there is notably less instrumental phonetic work on the realization of glottal stops. In languages that lack contrastive glottal stops, such as English and German, glottal stops are often realized without complete glottal closure, but this variation is prosodically-mediated [3, 5, 17]. In languages with contrastive glottal stops, such as Itunyoso Triqui and Yucatec Maya, non-modal phonation is a frequent realization of the stop [6, 18]. A recent study on Hawaiian examined the realization of glottal stops across word positions within a corpus of radio program speech from eight speakers [2]. Unlike past studies, Davidson [2] categorized realizations of glottal stops into eight groupings related to the presence/absence of complete closure, the location of creaky phonation if complete closure was absent, and realizations involving modal phonation with or without local intensity perturbations (see [8, 21]). She found that glottal stops were realized with complete closure just 10-12% of the time in word-initial position and 2-6% of the time in word-medial position. Similar to studies on English and German, the realization of glottal stops in Hawaiian was prosodically-mediated. Therefore, overall, complete closure was rare in unscripted speech.

Glottal stop is sometimes treated as a glide in the phonology literature [19], but just as frequently as a stop [20]. We will be comparing it to /t/, and we can expect that the lack of an oral closure for glottal stop and/or its different phonological status may make it shorter in duration than /t/.

1.1. Arapaho

Arapaho (ISO 639-3: arp) is an Algonquian language currently spoken in Wyoming and Oklahoma. It has a relatively simple consonant inventory (see Figure 1), monophthongal vowels which contrast in length and quality /ɪ, iː, ɛ, æː, ʊ, uː, ɔ, ɔː/, and a few diphthongs /eɪ, ɪɫ, oʊ, aɪ/. For a phonetic description of the vowel system, see DiCanio and Whalen [21]. The phonetic properties of the stop inventory (excluding glottal stops) are investigated in Kakadelis [22].

	Bilabial	Interdental	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Stop	b		t			k	ʔ
Affricate				tʃ			
Nasal			n				
Fricative		θ	s			x	h
Approximant	w				j		

Figure 1: Consonant inventory of Arapaho

Historical changes (from Proto-Algonquian (PA) to Arapaho-Gros Ventre (AGV)) resulted in many, distinctive glottal stops [23]. Goddard (1974) states

*"PA *x, *m/*n, *q, and *ʔ as first members of clusters all fall together to *ʔ."* [24]. Indeed, a rough estimate from Conathan's online Arapaho-English dictionary shows it to be the third most common consonant, accounting for 13% of all consonants. Do both its greater frequency in the lexicon and its distinctive status affect its realization?

1.2 Hypothesis

Given that Arapaho treats glottal stop as a stop and uses it with great frequency, we hypothesized that this would lead to greater use of the canonical, full glottal stop realization than in studies of other languages.

2. DATA

This question was studied by examining glottal stops in a portion of the corpus of Arapaho collected in Wyoming by Lisa Conathan (2004-2007) [12]. We selected two male and four female speakers for our analysis. Both narrative and elicited speech were analyzed. Overall, there were 493 tokens of glottal stop in elicited speech (98.6 per speaker) and 2117 in narratives (423.4 per speaker). Phonemic /t/ was measured for comparison, with 263 tokens in elicited speech and 1029 in narratives. A total of 284 minutes (4 hours, 44 minutes) of Arapaho speech was analyzed.

2.1. Alignment

Transcriptions were modified to use only ARPABET characters [25] so that an English-based aligner, P2FA [26], could be used. This required treating Arapaho glottal stop as a <T> since English lacks this phoneme. Those boundaries were hand corrected.

2.2 Delimiting glottal stops

We classified the realization of glottal stops into five groups (below):

Full glottal stop.
Creaky phonation.
Creaky phonation with local glottalization.
No glottalization.
Problematic stop.

These groupings were chosen organically by several raters of the Arapaho speech. The categories of *creaky phonation* and *creaky phonation with local glottalization* differed mainly in terms of the temporal extent of creaky voice. In the former, creaky phonation was found uniformly across the adjacent vowels, but in the latter, it was adjacent to a notably more aperiodic region corresponding to the lexically expected location for the glottal stop. The last category was used for those glottal stops whose realizations we could not categorize. Only 6.1% of all observed glottal stops fell into this category.

Each category was also measured for its duration. For full stops, this included the silence and adjacent noise (aspiration). The two creaky phonation types were challenging, but the portion that had the least clear formant structure was taken. Duration for tokens that lacked glottalization was taken as the stretch of low amplitude and/or low F0 associated with the glottal gesture; see [8, 27] for a discussion of glottalization-induced effects on amplitude and F0.

The underlying /t/s were analyzed for comparison with /ʔ/. As with full glottal stops, silence and associated aspiration were included as part of the durational measurement.

2.3. Calculating harmonics-to-noise ratio

One useful measure of voice quality is the harmonics-to-noise ratio (HNR). The less noise, the higher the ratio, and the more "modal" the phonation [28]. Though there are various implementations of HNR, it always relies on the presence of periodicity in the signal. Creaky phonation or glottalization, which we often find in the context of different realizations of a glottal stop, can strongly affect periodicity. HNR values may be indeterminate if there is substantial aperiodicity.

Instead of a typical measure of HNR, we used the "harmonic ratio" (HR) of Kim et al. [29] implemented in MATLAB, bounded between 0 (white noise) and 1 (sine wave). HR values were calculated continuously with 30 ms sliding windows stepped by 1 ms. We measured the median HRs during the glottal stop itself, and during /t/ as a comparison, weighted by the number of tokens per speaker. The advantage of this method is the ability to track HR changes in the speech signal regardless of the observed periodicity.

3. RESULTS

3.1 Category frequencies

The percentages of realizations for the glottal stops are shown in Table 1. These are averaged across talkers and speaking style (narrative or elicited speech). More glottal stops were realized without full closure than with full closure in the data (65.9% vs. 34.1%, respectively). This matches findings in other languages, such as Hawaiian [2]. However, the proportion is higher in Arapaho by a factor of 2 or 3, depending on the study. See Table 2 for realization percentages by speaker.

Realization	Average %	N
Full glottal stop	34.1	890
Creaky phonation	28.1	734
Creaky phonation with local glottalization	19.5	508
No glottalization	12.2	319
Problematic stop	6.1	159

Table 1: Glottal stop realization across speakers and styles.

Speaker	% stop	N
ZM	59.8	417
RMG	14.2	54
JA	33.3	23
MKU	24.6	170
ASE	25.1	87
LD	32.5	139
Mean/total	31.6	890

Table 2: Glottal stop realized as full stop, by speaker, across styles.

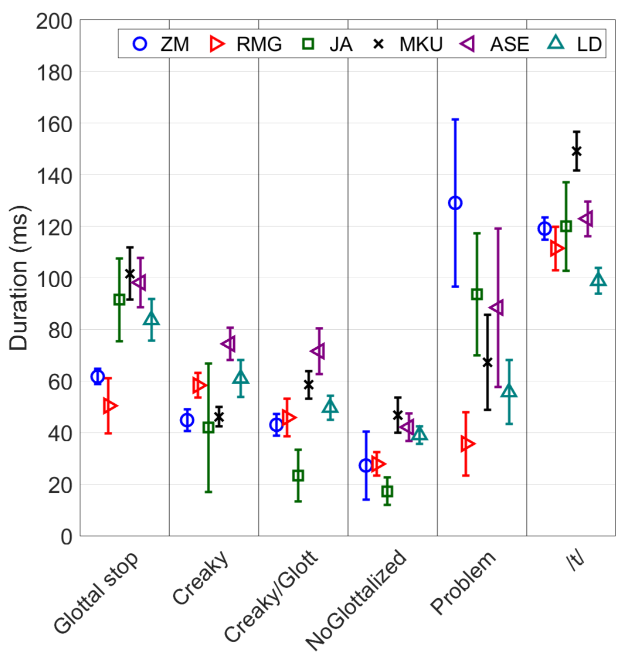

Figure 2: Duration of realizations of glottal stop and of /t/. Individual speakers' values are given, collapsed across speaking style. Error bars show 95% confidence intervals (CIs) of the mean.

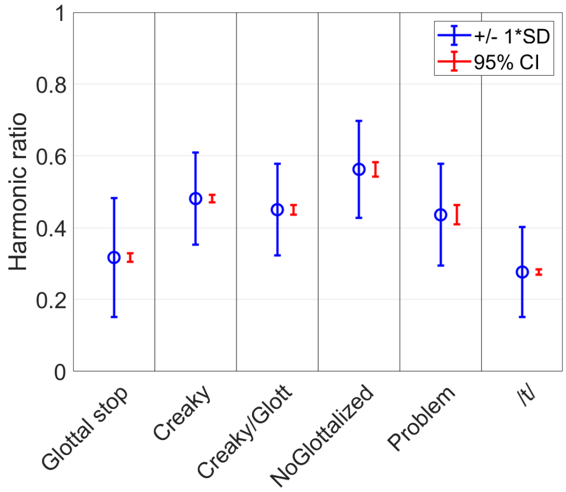

Figure 3: HR for realizations of glottal stop and of /t/. Averaged across speaker and style. Error bars show 95% Cis in red and +/- 1*SD in blue.

3.2 Duration

The average durations of the five realizations, and that of /t/, are shown in Figure 2. Complete closures are about 50% longer than the creaky phonation realizations (82.0 vs. 52.3 ms). Cases where no glottalization (i.e., aperiodicity) was observed had the shortest average duration (31.8 ms). These patterns suggest that there is a relationship between the duration of a glottal stop and its degree of lenition: full closure > creaky phonation > only intensity/f0 perturbations. These span the range from longer duration to shorter duration.

We also compared durations across the two speech styles, narrative and elicitation. Though we have recordings of both styles for most speakers, for speaker ZM, we had only narrative speech data, and for speaker JA, we had only elicited speech data. For some speakers, there were no instances of a glottal stop category for a particular style. The durations of both glottal stop and /t/ were quite similar in elicited speech and narratives. Because of these gaps, only about half of the comparisons were available for comparing realizations across styles. Either the speaker lacked any data for the style (two speakers), or the speaker had both styles but one of the styles did not have any instances of the realization. The average difference in duration across speech styles was 15.7 ms. The range was 77.3 ms. The slightly positive difference may simply be due to a slower speaking rate in elicited speech.

3.3 HR

Results for HR are shown in Figure 3. The non-glottalized realizations had the largest HR. The other glottalized versions were lower than that and similar to each other. As can be expected, both the stops (glottal and underlying /t/) had the lowest values. Although not a full statistical test, CI overlap showed /t/ to have the lowest value, closely followed by glottal stop; the realization with no glottalization has the highest value, and the two realizations with creaky phonation had similar values.

4. DISCUSSION

Glottal stops were highly variable in the production of Arapaho speech, though as observed in past studies, complete glottal closure was significantly less common than realizations consisting of non-modal phonation. Complete glottal closure was rather more common in Arapaho (34.1%) than in Hawaiian (2-12% by context) [2]. One explanation for this may relate to differences in speech rate across languages. Though we do not have data on Hawaiian speech rate, previously reported data indicate that Arapaho speech is notably slower than either Sierra Norte de Puebla Nahuatl or Bardi [16]. Since the presence of complete glottal closure varies inversely with produced duration (both here and in Hawaiian), one possibility is that the longer realizations of underlying glottal stop in Arapaho result in a reduced potential for target undershoot [30] to occur with glottal closure.

Although both Arapaho and Hawaiian have distinctive glottal stops, their functional load is likely to be higher in Hawaiian. Most word-initial glottal stops in Arapaho occur within a small set of particles or interjections, whereas Hawaiian makes extensive use of it word-initially. Many final glottal stops in Arapaho are also epenthetic, while Hawaiian lacks coda consonants altogether.

Examining prosodic effects on the realization would be of extreme interest given the observations that the realization of glottal stops varies by prosodic position. However, we did not have the ability to do so here. Using unscripted speech corpora for phonetic data analysis increases the ecological validity of the observed phonetic patterns. However, a disadvantage is the difficulty of annotating prosodic detail given the great variability of the prosodic contexts. We could only reliably code for word position in the current corpus, and the limitations on initial glottal stop (only a few interjections and particles) would have made the comparison a weak one.

5. CONCLUSION

We have shown that Arapaho glottal stops are highly variable in their realization. This variation is strongly correlated with the duration of the glottal stop itself. Full glottal closure occurs in contexts of longer duration and is produced with low HR values; non-modal phonation is more common with shorter duration glottal stops and is produced with higher HR values; further, a lack of glottalization is more common with the shortest overall duration and is produced with the highest HR values. Full glottal closure is uncommon in the Arapaho corpus and these findings demonstrate a strong pattern of duration-induced reduction in spontaneous speech production [11, 30].

The Arapaho realization of distinctive glottal stop is similar to what has been observed in other languages, but the frequency of the "canonical" realization as a true stop is, in this corpus, approximately twice as high as that found for other languages. We theorize that this particular finding may relate to differences in speech rate across languages, though research on additional corpora of languages with contrastive glottal stops is required to test this hypothesis.

The lack of an effect of phonemic status both in Arapaho and other languages on the proportion of full glottal stop realizations would appear to relate to the ease with which listeners accommodate to the various realizations. The great variability in the realization of tones or nasalization suggests that canonical forms need not be the most frequent for a wide variety of linguistic signals.

Phonetic descriptions of endangered languages seldom include accounts of speech variation [31], even though such variation is important for language learners. The work here serves the additional goal of providing a description of this variation for Arapaho. The current results show that Arapaho is typical in its realization of distinctive glottal stops. However, it is atypical relative to Hawaiian and other languages in the frequency with which full glottal closure is produced by speakers. We hope that further research will reveal the precise mechanisms explaining these patterns.

6. ACKNOWLEDGMENTS

This work was supported by NSF Grant No. 0966411 and NIH grant DC-002717 to Haskins Laboratories.

7. REFERENCES

[1] P. Ladefoged and I. Maddieson, *The sounds of the world's languages*. Oxford, UK; Cambridge, MA: Blackwell, 1996.

[2] L. Davidson, "Effects of word position and flanking vowel on the implementation of glottal stop: Evidence from Hawaiian," *J. Phon.,* vol. 88, p. 101075, 2021/09/01/ 2021.

[3] L. C. Dilley, S. Shattuck-Hufnagel, and M. Ostendorf, "Glottalization of word-initial vowels

as a function of prosodic structure," *J. Phon.,* vol. 24, pp. 423-444, 10// 1996.

[4] J. B. Pierrehumbert and D. Talkin, "Lenition of /h/ and glottal stop," in *Papers in laboratory phonology II: Gesture, segment, prosody*, G. J. Docherty and D. R. Ladd Eds. Cambridge: Cambridge University Press, 1992, pp. 90-117.

[5] M. Garellek, "Voice quality strengthening and glottalization," *J. Phon.,* vol. 45, pp. 106-113, 7// 2014.

[6] C. T. DiCanio, "Coarticulation between tone and glottal consonants in Itunyoso Trique," *J. Phon.,* vol. 40, pp. 162-176, 2012.

[7] J. H. Esling, K. E. Fraser, and J. G. Harris, "Glottal stop, glottalized resonants, and pharyngeals: A reinterpretation with evidence from a laryngoscopic study of Nuuchahnulth (Nootka)," *J. Phon.,* vol. 33, pp. 383-410, 10// 2005.

[8] C. Gerfen and K. Baker, "The production and perception of laryngealized vowels in Coatzospan Mixtec," *J. Phon.,* vol. 33, pp. 311-334, 7// 2005,.

[9] C. Gerfen, "Topics in the phonology and phonetics of Coatzospan Mixtec," Ph.D. dissertation, University of Arizona, 1996.

[10] L. Storto and D. Demolin, "The phonetics and phonology of South American languages," in *The Indigenous languages of South America*, L. Campbell and V. Grondona Eds. Berlin: Mouton de Gruyter, 2012, pp. 331-390.

[11] C. T. DiCanio, W.-R. Chen, J. Benn, J. Amith, and R. C. García, "Extreme stop allophony in Mixtec spontaneous speech: data, prosody, and modelling," *J. Phon.,* vol. 92, no. 101147, pp. 1-18, 2022.

[12] L. Conathan, *Northern Arapaho linguistic corpus*: Endangered Language Archive, School of Oriental and African Studies, U. of London, 2004-2007.

[13] *PHOIBLE 2.0.* (2019). Max Planck Institute for the Science of Human History, Jena. Accessed: 2021-12-15. [Online]. Available: http://phoible.org

[14] D. Silverman, *Phasing and recoverability*. New York: Garland Publishing, 1997.

[15] L. Davidson, "The versatility of creaky phonation: Segmental, prosodic, and sociolinguistic uses in the world's languages," *WIREs Cognitive Science,* vol. 12, no. 3, p. e1547, 2021/05/01 2021, doi: 10.1002/wcs.1547.

[16] M. Garellek, "The phonetics of voice," in *The Routledge handbook of phonetics*, W. F. Katz and P. F. Assmann Eds. London: Routledge, 2019, pp. 75-106.

[17] M. K. Huffman, "Segmental and prosodic effects on coda glottalization," *J. Phon.,* vol. 33, pp. 335-362, 7// 2005.

[18] M. Frazier, "The production and perception of pitch and glottalization in Yucatec Maya," Ph.D. dissertation, University of North Carolina, Chapel Hill, 2009.

[19] R. Jakobson, G. Fant, and M. Halle, *Preliminaries to speech analysis: The distinctive features and their correlates*. Cambridge, MA: MIT, 1951/1963.

[20] R. Jakobson, *Child language, aphasia and phonological universals*. The Hague: Mouton, 1968.

[21] C. T. DiCanio and D. H. Whalen, "The interaction of vowel length and speech style in an Arapaho speech corpus," in *Proceedings of the 18th International Congress of Phonetic Sciences*, vol. Paper number 513, The Scottish Consortium for ICPhS 2015 Ed. Glasgow: University of Glasgow, 2015, pp. 1-5.

[22] S. Kakadelis, "Phonetic properties of oral stops in three languages with no voicing distinction," Ph.D. dissertation, City University of New York, 2018.

[23] I. Goddard, "Recovering Arapaho etymologies by reconstructing forwards," in *Mír curad: Studies in honor of Calvert Watkins*, J. Jasanoff, H. C. Melchert, and L. Olivier Eds. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck, 1998, pp. 183-200.

[24] I. Goddard, "An outline of the historical phonology of Arapaho and Atsina," *International Journal of American Linguistics,* vol. 40, pp. 102-116, 1974..

[25] M. A. Mines, B. F. Hanson, and J. E. Shoup, "Frequency of occurrence of phonemes in conversational English," *Language and Speech,* vol. 21, pp. 221-241, July 1, 1978 1978, doi: 10.1177/002383097802100302.

[26] J. Yuan and M. Y. Liberman, "Speaker identification on the SCOTUS corpus," *The Journal of the Acoustical Society of America,* vol. 123, p. 3878, 2008.

[27] J. M. Hillenbrand and R. A. Houde, "Role of F0 and amplitude in the perception of intervocalic glottal stops," *J. Speech, Lg., Hear. Res.,* vol. 39, pp. 1182-1190, 1996.

[28] E. Yumoto, W. J. Gould, and T. Baer, "Harmonics-to-noise ratio as an index of the degree of hoarseness," *Journal of the Acoustical Society of America,* vol. 71, pp. 1544–1550, 1982.

[29] H.-G. Kim, N. Moreau, and T. Sikora, *MPEG-7 audio and beyond: Audio content indexing and retrieval*. Hoboken, NJ: Wiley Online Library, 2005.

[30] B. Parrell and S. Narayanan, "Explaining coronal reduction: Prosodic structure and articulatory posture," *Phonetica,* vol. 75, no. 2, pp. 151-181, 2018, doi: doi:10.1159/000481099.

[31] D. H. Whalen, C. T. DiCanio, and R. Dockum, "Phonetic documentation in three collections: Topics and evolution," *Journal of the International Phonetic Association,* vol. 52, no. 1, pp. 95-121, 2022.