

Importance of coding the Utterance domain in prosodic transcriptions

Marc Garellek

UCLA Linguistics Department

marcgarellek@ucla.edu

The Mainstream American English ToBI labeling convention codes break indices up to the level of the Intonation Phrase or IP (Silverman et al., 1992; Beckman and Ayers Elam, 1997). However, some models of prosodic structure have included an Utterance level above that of the IP (Nespor and Vogel, 1986). Although a break index above 4 (above the IP) has not been found to be useful for conditioning tonal phenomena, inclusion of the Utterance level can be useful for describing other prosodic phenomena. Indeed, previous work has shown some evidence that the Utterance level differs prosodically from the IP. For example, listeners seem to agree on perceived boundaries that are larger than the IP, and pauses at phrasal boundaries appear to be more frequent and longer in duration Utterance-finally than Utterance-medially (Wightman et al., 1992). For domain-initial strengthening, Fougeron and Keating (1997) and Keating et al. (2003) found that certain languages and speakers within a given language showed increased strengthening at onsets of Utterances compared to IP onsets.

We argue here that two voice quality phenomena in English, glottalization and phrase-initial voicing irregularity, are more prevalent and stronger at the onsets of Utterances than Utterance-medial IPs. Glottal stops and other cases of acoustic irregularity were analyzed for two female and two male radio newscasters in the Boston University radio news corpus (Ostendorf et al., 1995), previously analyzed for glottalization by Dilley et al. (1996) and Redi and Shattuck-Hufnagel (2001). This corpus includes break indices greater than 4, and in the present study we distinguished between 4 and 5 (the latter of which corresponds to the onset of an Utterance or breath group). All word-initial vowels were extracted, and we coded for presence of a complete glottal stop (defined as a period of silence followed by a clear burst and subsequent phonation onset (i.e., a full plosive [ʔ])). In total, 2086 tokens were included in the analysis. Using logistic mixed-effects regression, the two most important predictors of glottal stops were found to be vowel prominence (i.e., pitch-accented and/or unreduced) and the preceding break index. Therefore, glottal stops are more common before initial vowels that are both prominent and at the onsets of higher phrasal domains. The proportion of glottal stop occurrence increases as a function of the preceding break index, both for prominent and non-prominent vowels. Before prominent vowels however, glottal stops were more likely to occur Utterance-initially (after a break index of 5) than phrase-initially (after a break index of 4), as shown in Figure 1a. If we consider glottal stops to be optionally inserted before initial vowels, we can conclude that they are more likely to be inserted at the onset of an Utterance than at the onset of an IP. If instead we consider glottal stops to be obligatory before prominent initial vowels but can be reduced, then these data suggest that /ʔ/ is more likely to be realized as a full stop (i.e., greater strengthening) Utterance-initially than phrase-initially.

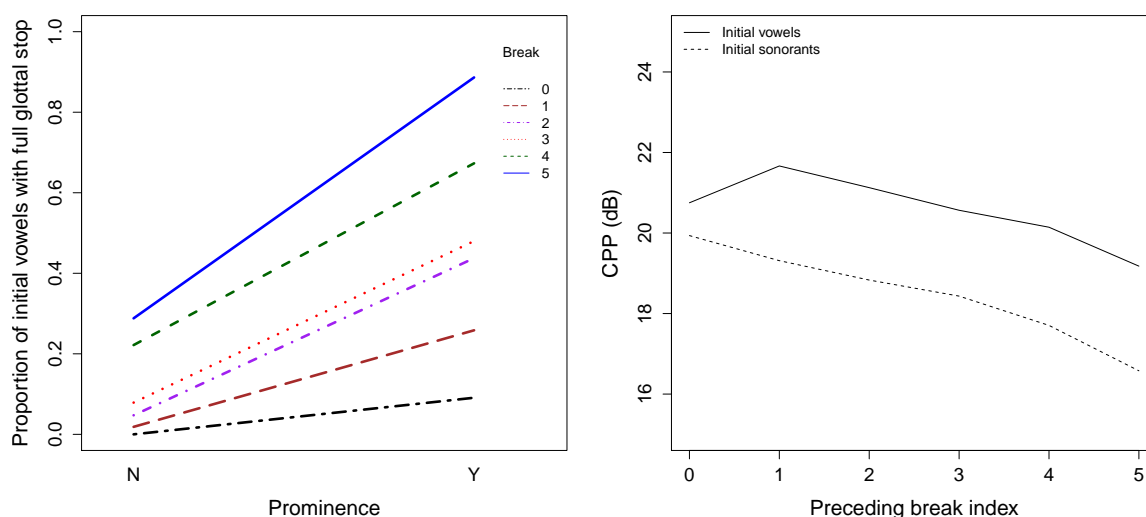
The Utterance-initial effects on glottal stop occurrence were also supported by quantitative acoustic measures of phonation for vowel- and sonorant-initial words. In total, 1288 word-initial sonorants were included in this analysis along with the vowel-initial words from the previous part. The quantitative measures, which were analyzed using linear mixed-effects regression, included both harmonic and noise measures. At the onsets of phrases, both vowel-initial and sonorant-initial words show a decrease in periodicity, as indicated by lower values of periodicity measures like cepstral peak prominence (CPP; see Figure 1b). Utterance-initial vowels and sonorants show even lower CPP values than Utterance-medial IP-initial ones, consistent with the idea that onsets of Utterances are more glottalized and/or less periodic than onsets of Utterance-medial IPs.

In sum, the qualitative and quantitative measures in this study show more frequent glottalization and lower periodicity as the phrasal domain increases, even beyond the Intonation Phrase in English. Coding for the Utterance level in prosodic corpora can therefore provide useful information about the extent of initial strengthening and variability in voice quality throughout the course of the utterance. Prosodic corpora should therefore be coded for an Utterance level, even if the tonal level does not require an additional level above the Intonation Phrase.

Figure 1: Differences between IP-initial (Utterance-medial) and Utterance-initial vowels and sonorants.

(a) Proportion [ʔ] as a function of prominence and preceding break index before the vowel-initial word. 'N' refers to non-prominent initial vowels; 'Y' to prominent ones.

(b) CPP (periodicity measure) as a function of preceding break index, when each prosodic unit begins with a vowel or sonorant.



References

- Beckman, M. E. and Ayers Elam, G. (1997). *Guidelines for ToBI labelling, Version 3*. The Ohio State University Research Foundation.
- 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.
- Fougeron, C. and Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, 101:3728–3740.
- Keating, P., Cho, T., Fougeron, C., and Hsu, C.-S. (2003). Domain-initial articulatory strengthening in four languages. In Local, J., Ogden, R., and Temple, R., editors, *Phonetic Interpretation (Papers in Laboratory Phonology 6)*, pages 143–161. Cambridge University Press, Cambridge.
- Nespor, M. and Vogel, I. (1986). *Prosodic Phonology*. Foris, Dordrecht.
- Ostendorf, M., Price, P. J., and Shattuck-Hufnagel, S. (1995). The Boston University radio news corpus. Technical Report ECS-95-001, Boston University.
- Redi, L. and Shattuck-Hufnagel, S. (2001). Variation in the realization of glottalization in normal speakers. *Journal of Phonetics*, 29:407–429.
- Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J., and Hirschberg, J. (1992). ToBI: A standard for labeling English prosody. In *Proceedings of the Second International Conference on Spoken Language Processing*, pages 867–870, Banff, Alberta.
- Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., and Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America*, 91:1707–1717.