Aperiodicity at Topic Structure Boundaries

Margaret Zellers & Brechtje Post

Research Centre for English & Applied Linguistics, University of Cambridge, UK
mkz21@cam.ac.uk, bmbp2@cam.ac.uk

Abstract

Topic structure in longer discourses has been shown to be marked in speech by prosodic variations, e.g. variations in fundamental frequency (F0) and speech rate. We investigated whether variations in voice quality, specifically aperiodicity as an aspect of glottalization, were also signals to topic structure by varying to indicate the strength of discourse boundaries. We found that variation in the presence of aperiodicity and length of aperiodic stretches were not good cues to topic structure; although there was some effect of topic structure on the presence of aperiodicity, the length of the aperiodic stretches did not correlate with topic structure at all. However, it still varied systematically in relation to F0 movements, and as such may be used as a cue to signal other linguistic structure.

Index Terms: voice quality, discourse, topic structure

1. Introduction

1.1 Prosodic cues to topic structure

Recent developments in the study of discourse structure have shown that it is possible, on the basis of variation in fundamental frequency (F0), to motivate the organization of long discourses around topics. Studies of the organization of topic structure have suggested that topics could be organized in a simple hierarchy (cf. Grosz & Sidner [1]) or in categories, such as those proposed by Nakajima & Allen [2] and Wichmann [3]. The existence of categories in discourse topic structure has been borne out to an extent by several studies (Zellers [4]; Zellers & Post [5]) which showed that variation in the size of F0 falls associated with sentence-initial target words in the studies, as well as variation in speech rate for some speakers, correlated with the topic structure categories defined as follows:

- (New) Topic: a shift in "aboutness" to something different than what precedes it
- Addition: new information on the same topic
- Elaboration: more detail on a previous utterance
- Continuation: completion of information begun in a previous utterance

The categories were shown to be ordered in terms of prominence; that is, a new Topic should be prosodically the most prominent, while a Continuation should be the least prominent. The F0 characteristics associated with the categories varied in prominence consistently with the predicted order of prominence of the categories, although it was unclear whether these parameters could signal a difference between Additions and Elaborations. The effects measured in these studies were all phrase-initial, varying with the onset of utterances in different topic structure categories. The question arises whether these categories also show phrase-final boundary cues. One possible acoustic candidate for signaling phrase-finality or boundary strength is a glottalized production.

Glottalization or creaky voice have been commonly identified as occurring at phrase boundaries in many languages, including English. For instance, Henton and Bladon [6] showed that glottalization was likely to occur, though not obligatory, at the ends of utterances; and numerous other studies have shown that glottalization of some form is much more likely at a variety of prosodic boundaries, as well as in other prominent locations such as pitch accents (cf. Redi & Shattuck-Hufnagel [7]). Glottalization may occur both before and after the location of the prosodic boundary. Dilley et al. [8] suggest that the glottalization of phrase-initial vowels in American English is an articulatory strengthening phenomenon, consistent with the consonantal strengthening found by Fougeron & Keating [9] for phrase-initial /n/. In this account, the role of strengthening would be to create a maximum differentiation between articulations in the two adjacent phrases. Huffman's [10] finding that syllable coda glottalization before obstruents is more likely in IP-final position is consistent with this account, if the higher incidence of glottalization is interpreted as evidence for an avoidance of coarticulation across these phrase boundaries. Similarly, Redi & Shattuck-Hufnagel [7] show that final glottalization is more likely in utterance-final position, as well as in utterance-medial position where there is a full IP boundary. They therefore propose that final glottalization may be a cue to the strength of phrase boundaries. However, the fact that glottalization is optional in final position, as well as the observation made by Fougeron & Keating [9], among others, that different speakers use different strengthening strategies at prosodic boundaries, raises the question of whether the situation is that simple in real production.

One complicating factor in studying the incidence of glottalization at phrase boundaries is that glottalization or creak are associated with a number of other segmental or prosodic characteristics that occur independently. For instance, syllable-final voiceless stops in English may lead to the production of creaky voice on preceding consonants if not to a complete allophonic substitution by a glottal stop (cf. Gordon & Ladefoged [11]). Creaky voice by definition has an F0 which is lower than that of modal speech, and this may be used by listeners as a perceptual cue to identify glottalization (Gerratt & Kreiman [12]). Some speakers may use a lowering of intensity instead of glottalization in comparable speech contexts (Dilley & Shattuck-Hufnagel [13]). Dilley et al. [8] also found that glottalization was more likely in contexts following pauses. Since pause length can be a cue to boundary strength (Swerts [14]), this may present another possible confound in investigating glottalization at prosodic boundaries.

Despite these difficulties, the anecdotal evidence for glottalization as a cue to phrase boundary strength is suggestive. Since even untrained listeners are able to identify creaky voice (based on a variety of acoustic parameters) with 95% accuracy (Blomgren et al. [15]), it would appear that glottalization is a readily available acoustic parameter for use as a cue to prosodic phrasing.

1.2 Current study and hypotheses

The current study investigates the occurrence of creaky phonation, specifically aperiodicity, at discourse structure boundaries in Standard Southern British English (SSBE). If this is the case, it should be possible to observe variation in aperiodicity running parallel to the topic structure categories noted above. On the basis of previous studies and to a certain extent based on impressionistic observations by the first author, we hypothesize that the presence or absence of glottalization could be a cue to distinguishing discourse structure boundaries of different strengths. In particular, we hypothesize that:

- Aperiodicity will be more likely to occur at boundaries of higher strength. That is, the ends of utterances preceding a New Topic utterance will be most likely to show aperiodicity, and ends of utterances preceding Continuations will be least likely.
- The length of aperiodicity will be greater (either absolutely or in proportion to the word or phrase length) at boundaries of higher strength. That is, stretches of aperiodicity will be longest in utterances preceding a New Topic, and shortest in utterances preceding a Continuation.

These hypotheses may be valid either simultaneously or independently of one another.

2. Methodology

2.1 Materials

The recordings used in this study were part of a larger project investigating prosodic correlates of topic structure in SSBE. Eighteen monolingual native speakers of SSBE read aloud a written text that had been designed to control for topic structure as well as for a number of potential segmental influences on the realization of prosody, including the segmental structure of the target words, the presence or absence of an anacrusis, and the position of the utterance in a group of utterances (to account for declination across a long stretch of speech).

Since different speakers use different strategies for glottalization, a subset of five speakers (4 female and 1 male) from the above study were used for the current investigation. These speakers were those who fairly consistently produced identifiable aperiodicity in utterance-final position as an aspect of glottalized productions. While most speakers produced glottalization in utterance-final contexts, not all produced aperiodicity as an aspect of that glottalization.

2.2 Identification of glottalization

Following Dilley et al. [8], two criteria had to be met for a stretch of speech to be marked as glottalized (or specifically, aperiodic). First, there had to be an auditory perception of glottalization, identified as creak, roughness or unevenness. Second, there had to be identifiable aperiodicity, or unevenness, in the waveform.

Instances of aperiodicity were excluded from the study if final glottalization could be related to the presence of (1) a voiceless stop consonant at the end of the utterance, or (2) an initial vowel in the following utterance, since in these cases, it would be difficult to determine whether the glottalization was related to the phrase boundary or to the segmental make-up of the tokens.

One common measure of aperiodicity in sound signals is jitter. For all speakers investigated, the jitter of stretches labeled "aperiodic" was approximately double the jitter of the modal stretches preceding them, as calculated by Praat (Boersma & Weenink [16]). The jitter of modal stretches was approximately 0.0123, while the jitter of aperiodic stretches was 0.0237 (t-test, t=-6.18, df=208.36, $p<0.001$). All other acoustic measurements were also taken in Praat.

3. Results

3.1 Occurrence of aperiodic stretches at boundaries

Aperiodicity as a cue to glottalization was present in 80.1% of the utterances studied (after discarding items as described in 2.2). For the male speaker in the study, 92.9% of all the utterances, regardless of other glottalization, were produced with final aperiodicity. For the four female speakers, aperiodicity was produced in 74.8% to 82.4% of the utterances.

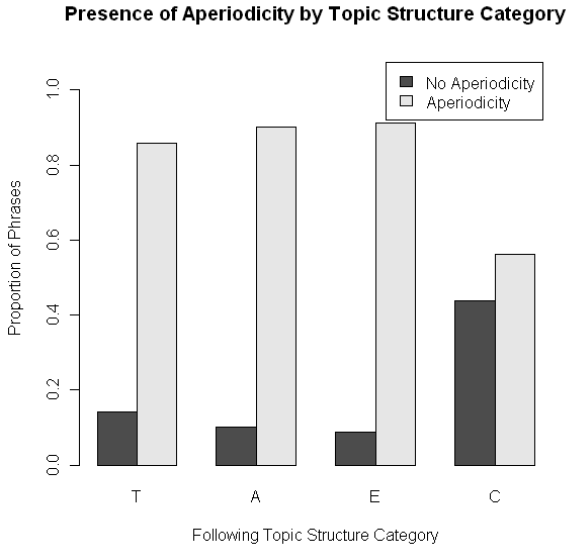

Figure 1: Presence of aperiodicity by following topic structure category. T=New Topic, A=Addition, E=Elaboration, C=Continuation.

Although the absolute number of items with aperiodicity in each topic structure category varied for each speaker, the overall trend was similar, and therefore all speakers are considered together (see figure 1). Stretches of aperiodicity were possible in all boundary conditions (that is, preceding all four topic structure categories) and in any position in the group of utterances (that is, initial, medial or final). A lack of aperiodicity was very uncommon in positions preceding new Topics (15%), Additions (10%) and Elaborations (9%), which did not vary significantly. However, in Continuations, 44% of utterances had no aperiodicity; this difference from the other three categories is statistically significant (χ^2=72.93, df=3, $p<0.001$). There was no main effect of the position of the utterance in the group of utterances, and no interaction with this factor; in particular, utterances in final position in the group of utterances were no more likely to contain aperiodic stretches than utterances in other positions, except that utterances in final position were never followed by Continuations.

3.2 Length of aperiodic stretches at boundaries

To compare the length of aperiodic stretches in different topic structure conditions, a subset of directly comparable target words was analyzed further (N=120). These words were all three syllables long, with lexical stress falling on the first syllable, and they were composed only of voiced sonorant segments. Differences between the target words were not found to be significant, so they are treated together in the following.

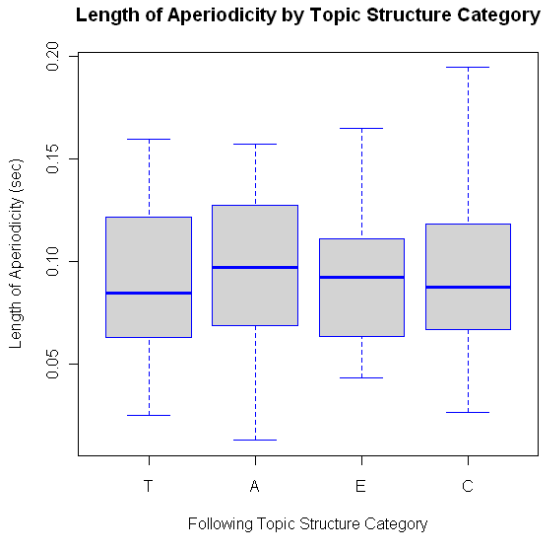

Figure 2: Length of aperiodic stretches by topic structure category.

There was no effect of topic structure category on the length of aperiodic stretches in the utterances. In the overall results, there were no significant differences between groups (see figure 2). For some speakers, there were significant differences between two or three categories, but these did not follow any consistent pattern, and may therefore be assumed to be incidental, unsystematic variation (at least with regards to the topic structure categories). There was also no effect of the position of the utterance in the utterance group; utterances in final position in the utterance group did not show longer or shorter stretches of aperiodicity than utterances in initial position in the group. This was true whether the absolute length of aperiodicity, the length of aperiodicity in proportion to the word, or the length of aperiodicity in proportion to the phrase were measured. There was a significant effect of speaker, with the speakers falling into three groups: the male speaker produced the longest aperiodic stretches, and the female speakers clustered into two pairs producing longer or shorter aperiodic stretches (ANOVA, $F(4, 119)=5.93$; $p<0.01$).

4. Discussion

Neither of our initial hypotheses turned out to be sufficient to explain the aperiodicity data presented here. The first hypothesis was shown to be only partly true: although utterances preceding a Continuation showed a different pattern with regards to the presence or absence of an aperiodic stretch, the other topic structure categories were not distinguishable using this variation. The second hypothesis was shown to be completely incorrect; there was no apparent relationship between the topic structure categories and the length of the aperiodic stretches.

It would seem, then, that aperiodicity may not be the best choice to signal subtle differences between boundary strengths, at least not on the level of discourse structure. However, the varying presence of aperiodicity in different contexts, as well as the perceptibility of the modal-creaky contrast by untrained listeners, suggests that aperiodicity may still be varied meaningfully, perhaps to aid in some other contrast. Since the second analysis compared different tokens of the same target word in similar contexts, we may discard lexical effects for the moment. However, it has been noted that phonation quality changes are often tied to F0 changes (cf. Gerratt & Kreiman [12]). It is possible that variation in phonation quality could either result from or possibly contribute to an F0 or tonal contrast. Dilley et al. [8], for example, have suggested that glottalization could be associated with an L* in the tonal structure.

The aperiodic stretches in utterances did not normally extend all the way until the end of the utterance; in most cases, there were several regular pitch periods at the very end, within a similar pitch range to that of the voiced stretch preceding the aperiodic region. The pitch movements before and after the aperiodic stretches were identified as either Rising or Falling on the basis of F0 measurements and the auditory judgments of the first author. (Note that the label "Rise" in the post-aperiodic stretch could also be applied to high pitch, and "Fall" in that context could be applied to non-high or non-rising pitch.) If we now examine the length of aperiodic stretches in these different contexts, we find an interacting effect of the F0 movements preceding and following the aperiodic stretches (see figure 3). Rising movements followed by low or falling movements show the shortest aperiodic stretches. This is fairly unsurprising, given that (at least in English) glottalization is not associated with high pitch, and the change from a rising to a falling movement would by definition include a pitch peak. However, the change from a falling movement to a high or rising movement, where we would expect an F0 valley, does not show the longest stretches of aperiodicity. Instead, the longest stretches of aperiodicity are found when the directionality of the pitch movement is

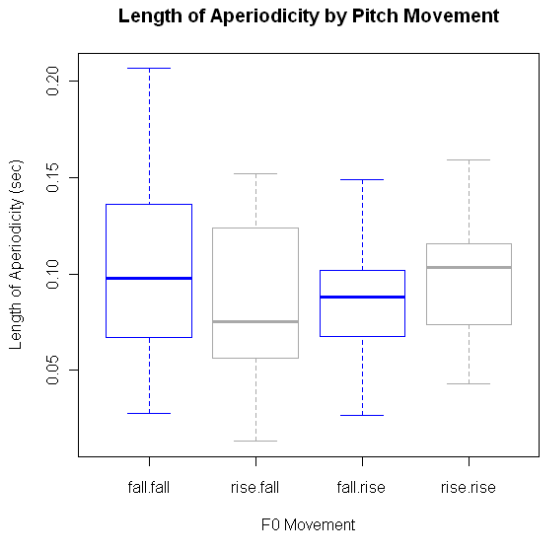

Figure 3: Length of aperiodic stretches by surrounding pitch movements

consistent: a fall to a fall, or a rise to a rise. Statistically, there is no main effect of either the preceding or the following pitch movement; only the interaction is significant (ANOVA $F(1, 120)=4.01$; $p<0.05$).

The correlation of longer aperiodic stretches to stretches where the directionality of the pitch movements is consistent suggests that variation in the length of aperiodicity may be a phonetic effect which is modulated to help create acoustic space for more complex pitch movements (and by extension, more complex tonal configurations). Perhaps the length of an aperiodic stretch aids listeners in identifying what kind of F0 movements are occurring at the ends of phrases. This could be especially relevant given that the falling off of amplitude utterance-finally and the relatively short length of the post-aperiodic stretches could create adverse listening conditions for perception of the tonal configuration. A more detailed study of the interaction between F0 and aperiodicity (and other correlates of glottalization), as well as a perceptual study of this phenomenon, could test this hypothesis. Alternatively, the length of creaky periods could be related to the amount of attention placed on the vocal folds by the speaker. Creaky phonation could be seen as a correlate of there being less attention on the vocal folds; more complex pitch movements would require the speaker to attend more strongly to the phonation mechanism, thus decreasing the length of stretches of aperiodicity and/or other phonetic cues to creakiness (Francis Nolan, personal communication).

5. Conclusion

We found that aperiodic stretches were less common in utterances followed by a Continuation (i.e., at the weakest topic structure boundaries in our data), but that none of the other topic structure categories could be distinguished in this way, in contrast with results obtained for other prosodic cues such as F0 variation and speech rate (cf. Zellers & Post [5]). Furthermore, the length of aperiodic stretches did not correlate with topic structure at all. Therefore, despite previous proposals that glottalization could be a cue to the strength of prosodic boundaries, there appears to be little evidence to support this idea in the realm of topic structure, other than in the distributional patterns of aperiodicity.

Topic structure is organization at the level of longer discourses, and the domains involved may have been too large for aperiodicity to be a relevant cue in the context. It may well be relevant in distinguishing between levels of prosodic structure and phrasing on lower levels, for example, between Intonational Phrases and smaller units. It is also possible that correlates of glottalization other than aperiodicity are more relevant to signaling the contrasts tested here. A future study investigating a wide variety of acoustic parameters associated with glottalization will provide an avenue for testing this possibility.

Interestingly, this study did find evidence to support the hypothesis that aperiodicity in final contexts varies systematically; but instead of the variation being tied to topic structure, it appears to be related to F0 movements in the vicinity of the aperiodic region. Therefore we may maintain the hypothesis that variation in the production of glottalization can be meaningful, but a wider range of sources of this variation will need to be examined to establish how these cues can be exploited to signal structural features of spoken discourses.

6. Acknowledgements

Many thanks to Sarah Hawkins, Francis Nolan, and Radek Skarnitzl for providing helpful advice and resources, and to the Cambridge prosody reading group and the anonymous reviewers for their comments. This research was supported by the EC Marie Curie Training Network/*Sound to Sense*/(MRTN-CT-2006-035561), and by the ESRC First Grant *Categories and gradience in intonation* (RES-061-25-0347).

7. References

[1] Grosz, B.J. & Sidner, C. Attention, intentions, and the structure of discourse. *Computational Linguistics* 12(3):175-204, 1986.

[2] Nakajima, S. & Allen, J.F. A study on prosody and discourse structure in cooperative dialogues. *Phonetica* 50:197-210, 1993.

[3] Wichmann, A. *Intonation in Text and Discourse*. Harlow: Longman, 2000.

[4] Zellers, M. Fundamental frequency and discourse meaning in SSBE. Presentation given at Phonetics and Phonology in Iberia, Las Palmas de Gran Canaria, 17-18 June 2009.

[5] Zellers, M. & Post, B. Fundamental frequency and other prosodic cues to topic structure. *Proceedings of IDP 2009*, Paris, to appear.

[6] Henton, C. & Bladon, A. Developing computerized transcription exercises for American English. *Journal of the International Phonetic Association*, 17(2): 72-82, 1987.

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

[8] Dilley, L., Shattuck-Hufnagel, S., and Ostendorf, M. Glottalization of vowel-initial syllables as a function of prosodic structure. *Journal of Phonetics* 24:423-444, 1996.

[9] Fougeron, C. & Keating, P.A. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 101:3728-3740, 1997.

[10] Huffman, M. (2005). Segmental and prosodic effects on coda glottalization. *Journal of Phonetics* 33(3): 335-362.

[11] Gordon, M. & Ladefoged, P. Phonation types: a cross-linguistic overview. *Journal of Phonetics* 29:383-406, 2001.

[12] Gerratt, B. & Kreiman, J. Toward a taxonomy of nonmodal phonation, *Journal of Phonetics* 29:365-381, 2001.

[13] Dilley, L. & Shattuck-Hufnagel, S. Variability in glottalization of word onset vowels in American English. *Proceedings of the XIII International Congress of the Phonetic Sciences*, Stockholm, Vol. 4, 586-589, 1995.

[14] Swerts, M. Prosodic features at discourse boundaries of different strength. *Journal of the Acoustical Society of America* 101:514-521, 1997.

[15] Blomgren, M., Chen, Y., Ng, M.L. & Gilbert, H.R. Acoustic, aerodynamic, physiologic and perceptual properties of modal and vocal fry registers. *Journal of the Acoustical Society of America* 103:2649-2658, 1998.

[16] Boersma, P. & Weenink, D. *Praat: doing phonetics by computer* (Version 5.1.20). [Computer program]. http://www.praat.org/ 31 October 2009.