IRREGULAR PHONATION AND ITS PREFERRED ROLE AS CUE TO SILENCE IN PHONOLOGICAL SYSTEMS

Janet Slifka

Speech Communication Group, MIT and Eliza Corporation, MA, USA
slifka@gmail.com

ABSTRACT

Regions of silence in the speech stream are commonly produced by pressing the vocal folds together and by spreading the vocal folds apart. Given that irregular phonation normally arises from both of these actions, it is proposed that there is a preferred role for irregular phonation in phonological systems – that of a cue to silence whether that silence is related to a segmental contrast or a prosodic structure and whether or not that silence is actually fully achieved. This preferred role for irregular phonation as a meaningful sound is grounded in the physical reality of managing vocal fold vibration.

Keywords: Prosody, pauses, stop consonants, glottalization, irregular phonation.

1. INTRODUCTION

Phonology characterizes the meaningful sounds of a language and the ways in which those sounds function as a structured system. The question of the extent to which properties of the speech production and perception system shape phonological systems in the world's languages can be posed in a variety of ways. One approach is to ask if there are aspects of either the perception or production system that bias the manner in which a sound is associated with a meaning, and is that bias evident across languages. In this paper, a particular sound, irregular phonation, is examined along with the mechanisms that generate irregular phonation and the ways in which this sound generally functions in phonological systems.

Irregular phonation is used as an umbrella term for phonation which differs from the normal range of quasi-periodic vocal fold vibration. Irregular phonation is characterized by an unusual difference in time or amplitude over adjacent pitch periods (exceeding normal ranges for jitter and shimmer), or an unusually wide-spacing between pitch periods for a given speaker. Breathy phonation is not included in this definition. Terms commonly used to describe irregular phonation include glottalization, creaky voice, and vocal fry.

A region of silence is created in the speech stream by suspending the sound sources, and one type of sound source arises from vocal fold vibration. Vocal fold vibration is most commonly suspended with one of two basic mechanisms: (1) press the vocal folds tightly together or (2) move the vocal folds wide apart.

This paper proposes that: given that irregular vocal fold vibration normally arises from both of the above sound-suspension mechanisms, irregular phonation has a preferred role in phonological systems – that of a cue to silence whether that silence is related to a segmental contrast or a prosodic structure and whether or not that silence is actually fully achieved. This preferred role for irregular phonation as a meaningful sound is grounded in the physical reality of managing vocal fold vibration.

2. PROSODIC BOUNDARIES

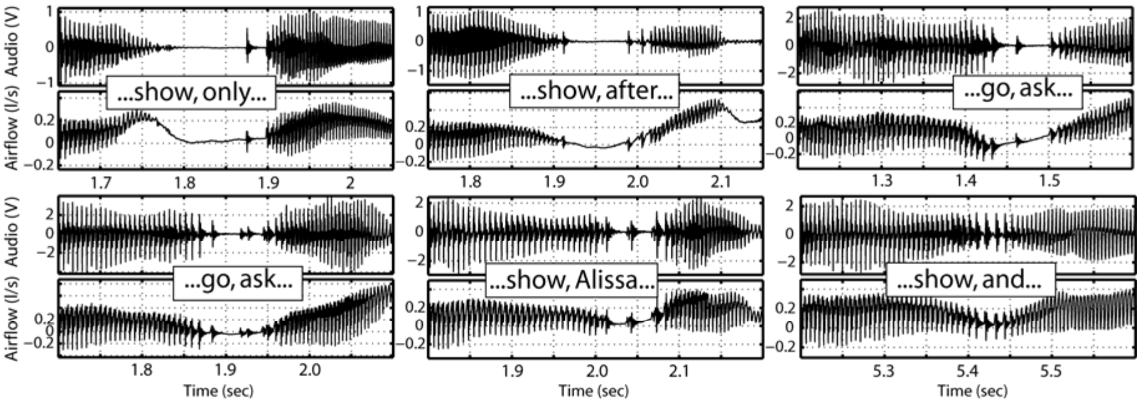
Prosodic boundaries are generally signaled with a range of acoustic cues including changes in fundamental frequency, irregular phonation, segmental duration, and pauses. Speakers generally place pauses at structural boundaries in both read speech (e.g. [2], [9]), and spontaneous speech (e.g.[2], [10]). This section discusses (1) irregular phonation at pauses produced with tightly adducted vocal folds, and (2) irregular phonation produced with spreading vocal folds.

2.1. Tightly adducted vocal folds

Glottalization has been defined as “a voice that contains frequent transient sounds (clicks) that result from relatively forceful adduction or abduction during phonation.” [8] Glottalization is produced with a long closed-phase and a short open-phase resulting in an overall decrease in airflow [7].

Figure 1 shows data from two female speakers for recordings of the audio signal and the

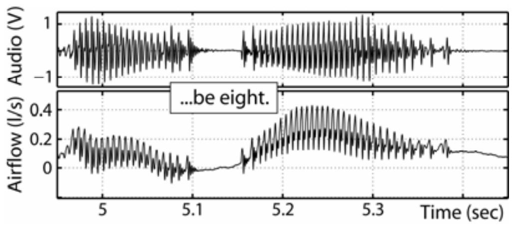
Figure 1: From upper-left to lower-right, a possible continuum of irregular phonation at a pause.



airflow signal as estimated using the Rothenberg mask [3]. These data are typical examples drawn from a larger database where the details of the acquisition are given in [5]. (Statistical analyses are in progress.) Tokens are from read speech where the printed prompts contained a comma at a major phrase boundary. (While the syntactic structure does not predict the prosodic structure, a panel of listeners judged a pause to be present at the comma in each of these simple sentences). The region of speech in the vicinity of that phrase boundary is plotted in the figure. In each case, the silence, if it exists, is a result of glottalization. The airflow signal shows an overall decrease in flow in the region of glottalization and a complete stop to airflow in those cases with a silent pause.

The upper left corner shows a silent pause where a single glottal pulse appears at the initial vowel of the phrase-initial word “only.” This series of tokens is proposed to represent samples of a continuum in irregular phonation associated with a pause which, through an increased number of glottal pulses and/or decreased spacing between the pulses, ends at the lower right token with a ‘filled’ pause, or continuous voicing through the phrasal boundary. Glottalization can arise as the vocal folds are pressed together (at the ‘start’ of a pause) or can occur as the vocal folds separate (at the ‘end’ of a pause). At the end of the continuum, there is not an actual silent pause, but the irregular phonation persists. A speaker does not need to precisely control for the exact number or spacing of irregular pulses, and acoustic cues to an underlying silent pause are maintained in the filled pause realization. In a similar manner, speakers could insert separation in the form of silence or underlying silence between smaller units (such as words (Figure 2)) or initiate from silence preceding the start of an utterance (such as a “hard attack” in a glottalized onset).

Figure 2: Irregular phonation at a word boundary.

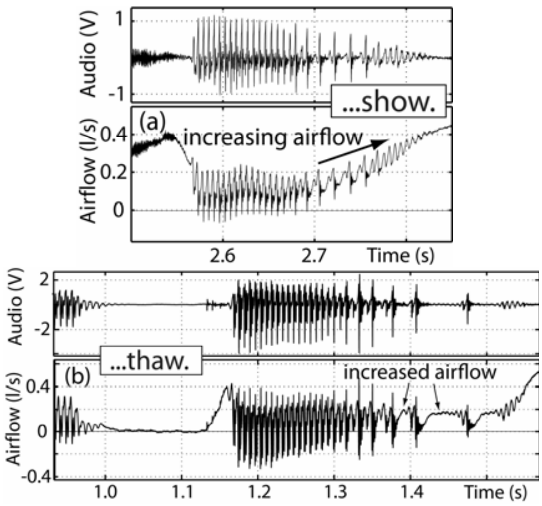


2.2. Vocal folds in the process of abducting

Utterances where the end of the utterance coincides with the end of an exhalation have been shown to associated with a particular type of irregular phonation characterized by a short closed-phase and a long-open phase (in contrast to glottalization) [6]. Figure 3a is an example in which there is an increasing region along the folds where the folds do not make contact, resulting in a rising airflow contour as the utterance ends. Figure 3b is an example where the vocal folds separate quickly resulting in a relatively long open-phase.

While this utterance-final irregularity could be planned for and executed in the form of a deliberate action – such as strongly adducted vocal folds and a short open-phase/long closed-phase, evidence for such adduction is not present for most utterances where the end of the utterance coincides with the end of an exhalation [6]. Since irregular phonation is not lexically contrastive for vowels in American English, the speaker does not have to prevent irregular phonation at the end of the utterance. In fact, it is an acceptable cue to utterance termination. The character of the phonation arises from the particular combination of phonation control parameters as an utterance ends – sometimes giving rise to regular phonation and at other times, giving rise to irregular phonation. Additionally, parameters that influence the mode of vocal fold vibration change in similar patterns for utterances that end with irregular phonation and those that end with regular phonation [6]. In many physical systems, a slight change in a single parameter can give rise to a large acoustic change. Some instances of irregular phonation may occur without a plan for deliberate insertion of irregularity, but result from a slight change in one or more control parameters – a change made for other reasons. At the ends of utterances that end with a breath, this condition may be more likely than at other parts of an utterance, because many of the parameters that control phonation are changing in the same span.

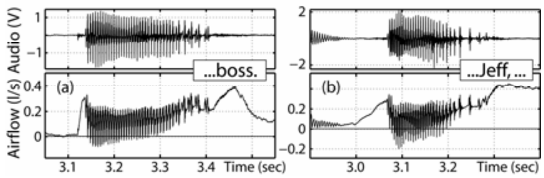
Figure 3: Irregular phonation produced with spreading vocal folds and increased airflow [6].



Instances of irregular phonation arising from spreading vocal folds are not limited to words ending with a vowel. In Figure 4a, the utterance ends with the word "boss" and coincides with the end of an exhalation. The vocal folds abduct in support of the voiceless fricative /s/, but again multiple parameters related to the mode of vocal fold vibration are changing in this region, and irregular phonation results. Figure 4b is an utterance-medial example at a phrase-final position where the word, "Jeff" ends with a voiceless fricative. Again, the vocal folds abduct and transition into a region of irregular phonation. In such cases, the speaker does not have to control conditions so as to prevent a transition into irregular phonation, because that phonation is an accepted cue to the prosodic structure and does not inappropriately affect the identity of the segments.

In general, there appears to be a relatively controllable mechanism to generate irregular phonation, i.e. speakers can produce it upon demand, and that mechanism involves adducted vocal folds and a decrease in airflow. There is also what might be called a less controllable mechanism for generating irregular phonation that involves partially-abducted or abducting vocal folds. This latter case is associated with an increased airflow and generally arises when several parameters that control the mode of phonation are changing in the same region (degree of adduction, transglottal pressure, vocal fold tension, etc.) If the particular set of conditions permit, then a transition into irregular phonation is observed, but reaching that set of conditions may or may not be planned by the speaker.

Figure 4: Fricatives with irregular phonation from spreading vocal folds (data from [4])



3. SEGMENTAL CONTRAST

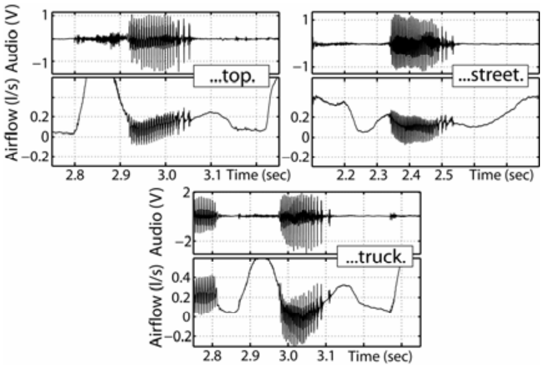
Stop consonant production involves complete obstruction of the airway and a change in air pressure behind the obstruction. For voiceless stops (and frequently for voiced stops as well) an interval of silence is created during the obstruction.

Some stop consonants obstruct the airflow at the place of articulation such as the lips, tongue tip, or tongue body. Other stop consonants simultaneously obstruct the airflow at the glottis and at the place of articulation. The ejective form of these stop consonants, or glottalic stops, is associated with irregular phonation at release.

In some languages, such as American English, irregular phonation is also an accepted cue to voiceless stop consonants. It is most commonly discussed in relation to syllable-final /t/ with terms such as "glottal stop". A glottal stop is produced by pressing the vocal folds tightly together to briefly obstruct air flow. In the case of a glottal stop, the irregular phonation serves as a cue to the silence normally associated with the stop consonant production.

Overall, the common thread is that the vocal folds are pressed together to obstruct airflow and create a small region of silence. The examples in Figure 5 open the question of whether or not this is true in all cases of irregular phonation associated with stop consonants, at least in American English. Tokens are utterance-final where the utterance again coincides with the end of an exhalation. A general trend for a rising airflow realization of irregular phonation is observed (as opposed to reduced-airflow glottalization). Again, the speaker does not have to deliberately prevent a transition into irregular phonation because not only is it an accepted cue to utterance termination but it is also a cue to voiceless stop consonants.

Figure 5: Stop consonants with irregular phonation from spreading vocal folds (data from [4])



4. DISCUSSION AND SUMMARY

This discussion is based upon observations that irregular phonation is generated by both a very-controllable mechanism (arising from vocal folds pressed tightly together) and a less controlled mechanism (arising in the course of vocal fold spreading). If listeners do not readily discriminate the types of irregular phonation arising from these two mechanisms, then a robust structure for meaningful sounds would assign these two types of irregular phonation the same role. This proposed role is to signal a region of silence or underlying silence. This representation is consistent for the listener whether that silence is associated with a prosodic structure or whether that silence is associated with a segmental contrast such as in the case of stop consonants.

Rather than considering the prosodic role and the segmental role as different classes of phenomena altogether, these roles within phonological systems for irregular phonation are linked by a common grounding in speech production physics and aerodynamics — arising as a by-product of mechanisms for creating silence and codified as a substitute for actual silence. This grounding in the physical production system is proposed to lead to a bias for the role of irregular phonation. It is not to say that sometimes this bias may be overridden; there are languages which are reported to have glottalized sonorant consonants and/or glottalized vowels (For a review, see [1]).

Listeners can distinguish regular phonation from irregular phonation, leading to the inclusion of irregular phonation as a possible entry in the set of meaningful sounds for a given phonological structure. Given that irregular phonation is distinct from other sounds in the human speech perception system, then the meaning associated with that sound could have been deliberately chosen to be anything. However, the widespread role for irregular phonation as a cue to silence suggests a grounding for that role is in the physical realities of the production system. Essentially, there is a preferred phonological role for irregular phonation based on the behavior of the speech production system, and that role is to signal silence.

5. REFERENCES

[1] Gordon, M., Ladefoged, P. 2001. Phonation types: a cross-linguistic overview. *J. Phonetics*, 29, 383-406.

[2] Henderson, A., Goldman-Eisler, F., Skarbek, A. 1965. Temporal patterns of cognitive activity and breath control in speech, *Language and Speech*, 8, 236-242.

[3] Rothenberg, M. 1973. A new inverse-filtering technique for deriving the glottal airflow waveform during voicing, *J. Acoust. Soc. Am*, 89, 1777-1781.

[4] Slifka, J. 2000. Respiratory constraints on speech production at prosodic boundaries, Ph.D. Thesis, Massachusetts Institute of Technology.

[5] Slifka, J. 2003. Respiratory constraints on speech production: Starting an utterance, *J. Acoust. Soc. Am*, 114, 3343-3353.

[6] Slifka, J. 2006. "Some physiological correlates to the end of phonation at the end of an utterance," *J. Voice*, 20, 171-186.

[7] Stevens, K. 1988. Modes of vocal fold vibration based on a two-section model. In: Fujimura, O. (ed.), *Vocal physiology : voice production mechanisms and functions.* New York: Raven Press, 357-371.

[8] Titze I. 1995. Definitions and nomenclature related to voice quality. In: Fujimura O, Hirano M (eds). *Vocal Fold Physiology, Voice Quality Control*. San Diego: Singular Publishing Group; 335-342.

[9] Winkworth, AL, Davis, PJ, Ellis, E, Adams, RD. 1994. Variability and consistency in speech breathing during reading: Lung volumes, speech intensity, and linguistic factors, *J. Speech and Hearing Research*, 37, 535-556.

[10] Winkworth, AL, Davis, PJ, Adams, RD, Ellis, E. 1995. Breathing patterns during spontaneous speech, *J. Speech and Hearing Research*, 38, 124-144.