Aspects of Creaky Voice
Pavilion West
2:00 – 5:00 PM

Organizer: Irene Vogel (University of Delaware)

Participants: Angeliki Athanasopoulou (University of Delaware)
Patrick Callier (Stanford University)
Elizabeth Choi (Massachusetts Institute of Technology)
Jennifer Cole (University of Illinois)
Marc Garellek (University of California, San Diego)
Patricia Keating (University of California, Los Angeles)
Jianjing Kuang (University of Pennsylvania)
Taylor Miller (University of Delaware)
Nadya Pincus (University of Delaware)
Robert Podesva (Stanford University)
Stephanie Shattuck-Hufnagel (Massachusetts Institute of Technology)
Anita Szakay (Queen Mary, University of London)
Ela Thurgood (California State University, Chico)
Irene Vogel (University of Delaware)
Ikuko Patricia Yuasa (Illinois Wesleyan University)

Investigation of voice quality, or phonation, is a component of many types of linguistic research including articulatory and acoustic phonetics, speech perception, sociolinguistics, segmental and prosodic phonology and phonological typology. Despite common interests and challenges, there is relatively little dialogue among researchers in the various sub-disciplines. This Special Session (5 papers and 6 associated posters) hopes to develop such dialogue.

Core issues of what creaky voice is are addressed in the initial discussion of the phonetics of creaky voice and non-modal phonation in general: its articulation, acoustics, and distribution. Included are such questions as the effect of phonological properties on phrase-final creak and how the acoustic and articulatory properties of final creak compare to those in other positions (Garellek & Keating). The acoustic properties associated with creaky voice are known to play different roles in the phonological systems of different languages, however, it appears that they interact with tones in significantly similar ways across languages, as demonstrated for three unrelated languages: Iu-Mien, Hainan Cham, and Anong (Thurgood). Unlike most phonetic / phonological properties of language, creaky voice evokes strong reactions among many listeners, particularly as it is associated with the speech of (younger) females. The actual distribution of this phenomenon is examined with an acoustic analysis of a large corpus of conversational data from males and females of different ages, and contrary to common belief, it is shown that creaky voice has not suddenly emerged among females in the last few years (Podesva et al.). Nevertheless, people do express emotional reactions to speakers who use creaky voice, although the reactions differ in males and females, and in different regions of the US (Yuasa). Moreover, even if not novel, creaky voice seems to be quite prevalent among younger women, and this suggests that may be a new “norm” in American English. It would thus be acquired by children as part of their phonology, something that is in fact demonstrated in the analysis of the speech of children between 6 and 12 years old (Athanasopoulou et al.).

Posters follow up with a review of acoustic measures used to distinguish modal from laryngealized (including creaky) voice and information on how researchers can obtain relevant measurements semi-automatically using UCLA’s VoiceSauce program (Keating & Garellek). The challenge and reliability of different acoustic measures for creaky voice are further examined in several languages that exhibit the phenomenon in somewhat different ways (cf. Pincus et al.), and in a comparison of task-driven, as opposed to more controlled, speech (Shattuck–Hufnagel & Choi). The interaction of phonation and tone systems is also addressed in more detail for Mandarin Chinese (Kuang) and Vietnamese (Miller et al.). Finally, listener sensitivity to phonation is further examined in relation to perception and the extent to which phonation properties in English are recognized as part of the basic phonological composition of a sentence, or something that may exist as a distinct component that may or may not be adopted by an individual speaker (Shattuck–Hufnagel & Cole).
We address several issues regarding the articulation, acoustics, and distribution of phrase-final creak: (1) what phonological factors favor the occurrence of creak; (2) its acoustic characteristics; (3) whether phrase-final creak differs from non-final creak. Based on an analysis of recordings from four English speakers in the BU Radio News Corpus, results indicate that the presence of phrase-final creak is favored by several prosodic factors, including higher phrasal break indices and longer phrasal duration.

To determine what acoustic properties characterize phrase-final creak, creaky vs. non-creaky phrase-final syllables are compared on several acoustic measures. Results indicate that there exists substantial cross-speaker variability in the acoustic realization of phrase-final creak.

Lastly, creaky phrase-final syllables are compared to creaky non-final syllables in another corpus that includes recordings from 12 speakers of English and Spanish. Differences between phrase-final and non-final types of creak are described acoustically, as well as articulatorily via electroglottographic measures.

The paper focuses on the interaction of pitch and phonation in three languages, each from a different language family: Iu-Mien (Hmong-Mien), Hainan Cham (Austronesian), and Anong (Tibeto-Burman). In each case, the phonation could be ignored allowing the tones to be described as phonemicized pitch patterns, but a closer examination of each language shows that this is not the case. The study shows that when a tightly clustered pitch patterns occur in multiple tone languages, phonation differences not only provide a way to keep the tones separated but also a way by which tones merge through loss of pitch differences, and a way by which tones realign in acoustic space.

This paper draws on an acoustic investigation of a large corpus of conversational data from speakers of English in California to illustrate that women exhibit robustly creakier phonation than men. Measures of spectral tilt (H1-H2, H1-A1, H1-A2, H1-A3) and cepstral peak prominence were taken at the midpoint of all vowels in the corpus, and data for each measure were fitted to mixed effects linear regression models, which reveal that several linguistic factors influence measures of phonation, and women are creakier than men. Even though the rise of creaky voice among young women is characterized as a recent trend in media coverage of “vocal fry,” our study offers no evidence that young women exhibit acoustically distinct patterns from women of other ages.

The present investigation compares and contrasts how young male and female informants (99 males; 77 females) across two US regions of the West Coast and Midwest perceive female American creaky voice. The study generated complex results, showing the most negative comments (particular by females) in the Midwest, however, there were generally mostly positive or neutral comments made by the informants in both regions. There was a tendency for male listeners in the Midwest to perceive older-sounding female creaky voice as more attractive, but for male listeners on the West Coast to perceive younger-sounding female creaky voice as more attractive. The results of this study indicate that young American listeners may be in the process of reinterpreting females creaky voice as a new type of feminine voice.
Acquisition of creaky voice in English

It has been observed that creaky voice has been increasing in American English, particularly among young women. The question we address is whether children who are exposed to this variety of speech develop creaky voice as an integral part of their language acquisition. To this end, we analyze the speech of children 5, 7 and 10 years old, and adults, specifically the first and final words in sentences of the structure: “Tom’s/Emma’s X is at the top/bottom.” As expected, the adults use creaky voice extensively at the end of the sentence, but less at the beginning, and a similar pattern is observed in the oldest group of children. The younger children use more modal phonation, however, it appears that they are indeed acquiring creaky voice as they develop other aspects of their phonology and creaky voice is thus becoming a standard, as opposed to a stylistic, aspect of English prosody.

Abstracts: Posters

Patricia Keating (University of California, Los Angeles)
Marc Garellek (University of California, San Diego)

Acoustic analysis of creaky voice

In this presentation we will review acoustic measures that have been shown to distinguish modal voice from laryngealized voice, including creaky voice. These measures of the speech spectrum include various differences in harmonic amplitudes (e.g., H1*-H2*, H1*-A2*), various relations between energy in harmonics vs. in spectral noise (e.g., HNR and CPP), and a measure of energy in harmonics vs. subharmonics (SHR), as described by previous researchers. We will then show how these measurements can be obtained semi-automatically and for large corpora using VoiceSauce, a free analysis program from UCLA. We will discuss correction of harmonic amplitudes for vowel formants, as well as the challenge of estimating fundamental frequency (F0) in creaky voice, which may arise from unpredictable changes in F0 and the amplitude of voicing, as well as from prominent subharmonics.

Jianjing Kuang (University of Pennsylvania)

The covariation between pitch and phonation: creaky voice in Mandarin tones

This study aims to provide several pieces of experimental evidence to answer the question whether allophonic creaky voice in Mandarin is tied to tonal categories or is driven by phonetic pitch ranges. We showed that: the presence of creak is not exclusively limited to Tone 3, but can occur with any of the low targets in the Mandarin tones; Tone 3 is less creaky when the overall pitch range is raised, but more creaky when the overall pitch range is lowered; Tone 1 is also subject to similar variations; overall, voice quality co-varied with pitch height in a wedge-shaped function. In sum, voice quality is quite systematically tied to F0 in Mandarin. Voice quality thus has the potential to enhance the perceptual distinctiveness of extreme pitch targets.

Taylor Miller (University of Delaware)
Angeliki Athanasopoulou (University of Delaware)
Nadya Pincus (University of Delaware)
Irene Vogel (University of Delaware)

The effect of focus on phonation in Northern Vietnamese tones

The Northern Vietnamese tonal system involves pitch movement and phonation differences. Given that focus typically enhances the acoustic manifestation of words, we ask to what extent focus affects specific properties in Vietnamese. Specifically, we compare two rising tones that differ primarily in phonation: (i) modal sác and (ii) creaky ngã. Ten Hanoi speakers produced a total of 1920 target vowels (/a, i, u/) in real three-syllable compounds appearing in focus and non-focus contexts. Our findings show focus enhancement of duration and intensity for both tones, but focus did not enhance phonation or pitch properties. These findings are consistent with the broader claim that the acoustic properties used to express prominence in a language are selected such that they do not interfere with the manifestations of crucial contrasts of the language, specifically tone and creaky voice (and other phonation properties) in Vietnamese.
**Saturday, January 10**

**Tutorial**

**Nadya Pincus** (University of Delaware)
**Angeliki Athanasopoulou** (University of Delaware)
**Taylor Miller** (University of Delaware)
**Irene Vogel** (University of Delaware)

*The reliability of various phonation measurements*

Many phenomena are referred to as “creaky voice”, but both visual inspection of spectrograms and waveforms, and auditory assessment of speech samples suggest that several distinct phonation patterns are involved. Various measurement procedures have been proposed for phonation phenomena, raising the question of which is the most reliable for creaky voice. A crucial problem in analyzing phonation is that measurements typically depend on formants and harmonics, but if F0 is disrupted or missing, such measurements are questionable. We present conflicting findings from previous studies, and examine different options for our data from English, French and Vietnamese, languages that are claimed to have creaky voice, but where the appearance and auditory properties are quite diverse. We propose that measurement of specific aspects the disruption of F0 in creaky voice might also be needed in order to understand phonation, and shed light on different types of phonation that are termed “creaky voice.”

**Stefanie Shattuck-Hufnagel** (Massachusetts Institute of Technology)
**Elizabeth Choi** (Massachusetts Institute of Technology)

*Distribution of glottalized onsets in task-directed American English speech*

The systematic appearance of irregular pitch periods at prosodically significant locations, such as the onsets of intonational-phrase-initial and pitch-accented vowel-initial words in American English, has been reported in read laboratory speech (Pierrehumbert and Talkin 1992) and professionally-read radio news speech (Dilley et al. 1996). We extend these findings in a survey of irregular pitch period episodes in a corpus of task-driven speech (Kenney et al. 2013), using the map-task method (Anderson et al. 1991) to elicit 16 short dialogues from 8 pairs of adult female speakers of an East Coast dialect of American English. These dialogues were hand-labelled for episodes of irregular pitch periods, as part of a larger effort to mark the acoustic cues to prosodic structure and phonological feature contrasts. This study compares the consistency of laryngeal marking of prosodic structure across several speech samples from the same speaker, as well as the differences among the 8 speakers.

**Stefanie Shattuck-Hufnagel** (Massachusetts Institute of Technology)
**Jennifer Cole** (University of Illinois)

*Imitation as a tool for investigating cues to prosodic structure*

Recent work has revealed systematic variation in the phonetic realization of word forms, including voice quality variation across different prosodic contexts. Phenomena related to glottal configuration (e.g. breathiness and irregular pitch periods) occur at the onsets of prosodic constituents, and at prominences marked by pitch accents (Pierrehumbert & Talkin 1992, Dilley et al. 1996, inter alia); phrase-finally (Henton and Bladon 1987, Docherty et al. 1997, Redi and Shattuck-Hufnagel 2001); and as glottalized variants of word-internal -final /t/. We asked whether speakers imitating utterances reproduce voice quality variations in the target, or use their own typical voice quality patterns. Results suggest that listeners do not reproduce target voice quality variation, consistent with the hypothesis that, although patterns of laryngeal marking of prosodic structure are systematic, and contain information perceptible to listeners, they do not form part of the representation formed by a listener to govern an imitation of a heard utterance.