



# Acoustic Properties of Subtypes of Creaky Voice

Patricia Keating<sup>1</sup>, Marc Garellek<sup>2</sup>, Jody Kreiman<sup>3</sup>, Yuan Chai<sup>1</sup>

<sup>1</sup>UCLA Linguistics, <sup>2</sup>UC San Diego Linguistics, <sup>3</sup>UCLA Head & Neck Surgery

<sup>1</sup>keating@humnet.ucla.edu

ASA  
Spring 2023  
4aSC42



## Introduction

- “Creaky voice” refers to a number of different kinds of voice production with similar percepts.
- Our goal is to explore the possible types of creaky voice, give a clear definition to each type, and analyze the acoustic properties of each type.

## Prototypical creak

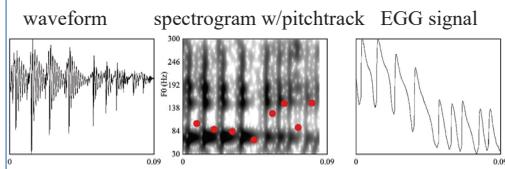
Prototypical creaky voice has three key properties:

- Low rate of vocal fold vibration (F0), w/ damped pulses
- Irregular F0, random or multiply pulsed
- Constricted glottis: vocal folds are close together, with a small peak glottal opening and a long closed phase (indexed here by EGG Contact Quotient  $\geq 0.60$ ), and thus with low airflow

From our existing audio recordings with EGG from several languages, speech tokens were found with 1, 2, or 3 of these properties. (Use the QR code at top right to listen.)

Prototypical creaky voice example: all 3 properties

(1) Low F0 + Irregular F0 + Constricted glottis  
(F0: 63 to 150 Hz; CQ: 0.6)



## Acoustics

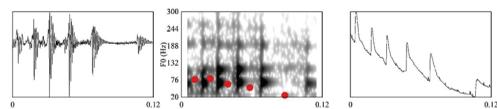
- Irregular F0 can be measured as pulse-to-pulse jitter; by the standard deviation of the F0; or, since it is perceived as spectral noise, as a low **Harmonic-to-Noise Ratio**.
- Strength of **period doubling** can be measured as relative strength of spectrum subharmonics (**Subharmonic to Harmonic Ratio**).
- H1-H2 correlates with the glottal Open Quotient and with Contact Quotient measures from electroglottography, thus indicating **glottal constriction/spreading**.

## Three key properties

- Low F0 and irregular F0 each suffice on their own for a creaky percept in informal listening by the authors;
- Constricted glottis alone is NOT sufficient for a creaky percept in informal listening by the authors.

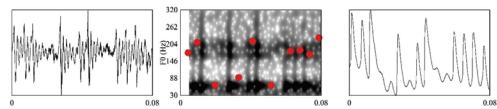
(2) Low F0 → Creaky

(F0: 26-81 Hz; CQ: 0.31, spread glottis)



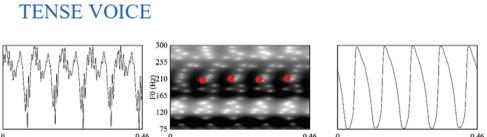
(3) Irregular F0 → Creaky

(F0: 62-246 Hz, mix of random and period-doubled; CQ: 0.38, spread glottis)



(4) Constricted glottis → Not creaky

(F0: 211 Hz; CQ: 0.61, constricted)

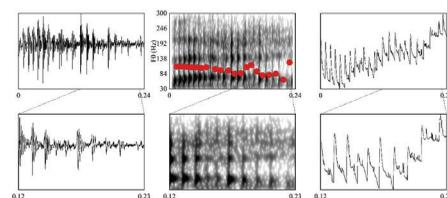


## Combinations of properties

(5) Low F0 + Irregular F0 → Creaky

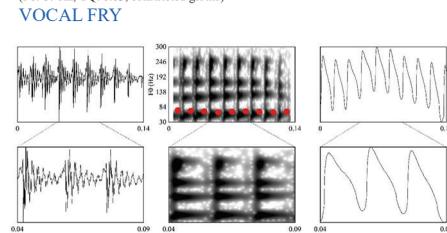
(F0: 62-124 Hz; CQ: 0.34, spread glottis)

SPREAD GLOTTIS CREAK (see also (2) + (3))



(6) Low F0 + Constricted glottis → Creaky

(F0: 67 Hz; CQ: 0.63, constricted glottis)



Irregular (high) F0 + Constricted glottis – Not found with creaky percept (yet)

## Acoustic correlates (signature) of each type

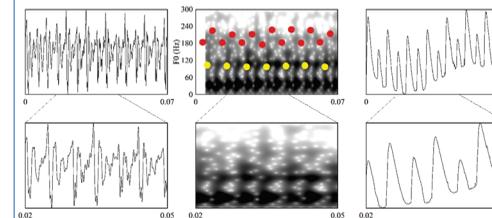
Type – sounds creaky	Low F0	Low HNR	High SHR	Low H1-H2	High CQ
Low F0	✓		(optional)		
Irregular F0		✓		(not defined)	
Prototypical creak	✓	✓		✓	✓
Vocal fry	✓			✓	✓
Spread glottis creak	✓ (and/or)	✓		NO, high	
Multiple-pulsed spread glottis creak	✓ (and/or)	✓	✓	NO, high	
Type – does not sound creaky					
Tense voice				✓	✓
Multiple-pulsed			✓		

## Multiple-pulsing

- Multiple-pulsing (e.g. period doubling) is a special case of irregular F0, as it is not random;
- By itself is NOT sufficient for creaky percept, instead sounds rough.
- Often the fundamental period includes the doubled pulses, not just a single pulse, so a lower F0 is possible (yellow dots in pitchtracks below) – if this is low enough, then percept is creaky

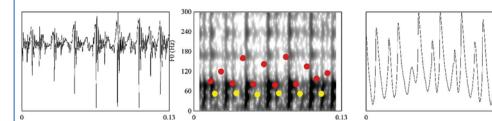
(7) Multiple-pulsing alone → Not creaky

(lower F0 (yellow dots): 100 Hz; CQ: strong pulse: 0.46; weak pulse: 0.47; mean: 0.47)



(8) Multiple-pulsing + Low F0 → Creaky

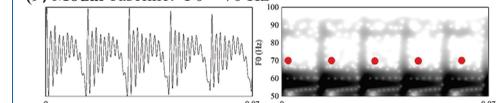
(lower F0 (yellow dots): 50 Hz; CQ: weak pulse: 0.36; strong pulse: 0.43)



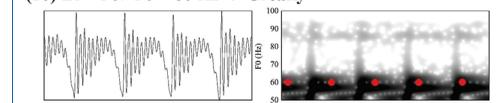
We test our observations from natural speech through informal listening to Klatt synthesis in Praat. →

## Klatt Synthesis in Praat

(9) Modal baseline: F0 = 70 Hz

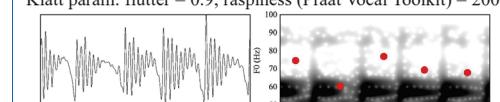


(10) Low F0: F0 = 60 Hz → Creaky



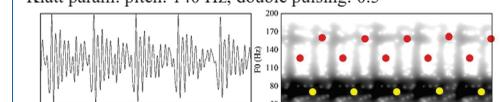
(11) Irregular F0: F0 = 70 Hz → Creaky

Klatt param: flutter = 0.9; raspiness (Praat Vocal Toolkit) = 200%



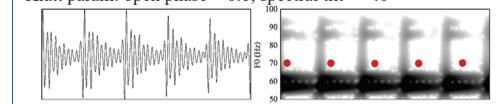
(12) Period-doubling: lower F0 (yellow dots) = 70 Hz → Not Creaky

Klatt param: pitch: 140 Hz; double pulsing: 0.5



(13) Constricted glottis: F0 = 70 Hz → Not Creaky

Klatt param: open phase = 0.1; spectral tilt = -40



## Conclusion

- Each type of creaky voice has a different (sub)set of the three key properties **low F0, irregular F0, constricted glottis**. None are necessary, and only the first two are sufficient by themselves for a creaky percept.
- Each has its own acoustic correlates, thus each type of creaky voice has a different acoustic signature.
- Thus acoustic analysis of creaky voice will give different results depending on which kind of creak is at issue.
- No single acoustic measure is criterial for all types of creaky voice. Most notably:
- While H1-H2 is the most common measure, glottal constriction is neither necessary nor sufficient for a creaky voice percept! – By itself it does not give a creaky percept, and creaky voice can have spread glottis.
- Low/irregular F0 are good correlates for phonemic creaky voice.

>> QR code above for sound files and references

- Ashby, M., Przedlacka, J. 2014. Measuring incompleteness: Acoustic correlates of glottal articulation. *J. Intern. Phon. Assoc.* 44, 283-296.
- Batliner, A., Berger, S., John, B., Kießling, A. 1993. MÜSLI: A classification scheme for laryngealizations. *Proc. ESCA Workshop on Prosody*, Lund, 176-179.
- Blomgren, M., Chen, Y., Ng, M.L., Gilbert, H.R. 1998. Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers. *J. Acoust. Soc. Am.* 103, 2649-2658.
- Chernyak, B.R., Simon, T.B., Segal, Y., Steffman, J., Chodroff, E., Cole, J.S., Keshet, J. 2022. DeepFry: Identifying vocal fry using deep neural networks. arXiv preprint arXiv:2203.17019.
- Childers, D.G., Lee, C.K. 1991. Vocal quality factors: Analysis, synthesis, and perception. *J. Acoust. Soc. Am.* 90, 2394-2410.
- Coleman, R.F. 1963. Decay characteristics of vocal fry. *Folia Phoniatrica et Logopaedica* 15, 256-263.
- de Krom, G. 1993. A cepstrum-based technique for determining a harmonic-to-noise ratio in speech signals. *J. Sp. Hear. Res.* 36, 254-66.
- Davidson, L. 2019. Perceptual coherence of creaky voice qualities. *Proc. 19<sup>th</sup> ICPhS Melbourne*, 147-151.
- Devaraj, V., Roesner, I., Wendt, F., Schoentgen, J., Aichinger, P. 2023. Auditory perception of impulsiveness and tonality in vocal fry. *Appl. Sci.* 13, 4186.
- Drugman, T., Kane, J., Gobl, C. 2014. Data-driven detection and analysis of the patterns of creaky voice. *Computer Speech & Language* 28, 1233-1253.
- Garellek, M., Keating, P. 2015. Phrase-final creak: Articulation, acoustics, and distribution. Annual Meeting of the Linguistic Society of America, Portland, OR.  
[https://linguistics.ucla.edu/people/keating/Garellek-Keating\\_2015\\_LSAtalk\\_creakyvoice.pdf](https://linguistics.ucla.edu/people/keating/Garellek-Keating_2015_LSAtalk_creakyvoice.pdf)
- Gerratt, B.R., Kreiman, J. 2001. Toward a taxonomy of nonmodal phonation. *J. Phonetics* 29, 365-381.
- Gobl, C., Ní Chasaide, A. 2010. Voice source variation and its communicative functions. In: Hardcastle, W., Laver, J., Gibbon, F. (eds), *The Handbook of Phonetic Sciences (Second Edition)*. Oxford: Blackwell, 378-423.
- Hedelin, P., Huber, D. 1990. Pitch period determination of aperiodic speech signals. *Proc. ICASSP Albuquerque*, 361-364.
- Huang, Y. 2023. *Phonetics of period doubling*. Unpublished dissertation, UC San Diego.  
<https://escholarship.org/content/qt33n789f2/qt33n789f2.pdf?t=rt5k6i>
- Kane, J., Drugman, T., Gobl, C. 2013. Improved automatic detection of creak. *Computer Speech & Language* 27, 1028-1047.
- Klatt, D., Klatt, L. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J. Acoust. Soc. Am.* 87, 820-857.
- Kreiman, J., Shue, Y.-L., Chen, G., Iseli, M., Gerratt, B. R., Neubauer, J., Alwan, A. 2012. Variability in the relationships among voice quality, harmonic amplitudes, open quotient, and glottal area waveform shape in sustained phonation. *J. Acoust. Soc. Am.* 132, 2625-2632.
- Kreiman, J., Antoñanzas-Barroso, N., Gerratt, B. R. 2010. Integrated software for analysis and synthesis of voice quality. *Behavior Research Methods* 42, 1030-1041.
- Kreiman, J., Gerratt, B.R. 2005. Perception of aperiodicity in pathological voice. *J. Acoust. Soc. Am.* 117, 2201-2211.
- Laver, J., 1980. *The phonetic description of voice quality*. *Cambridge Studies in Linguistics* 31, 1-186.
- Redi, L., Shattuck-Hufnagel, S. 2001. Variation in the realization of glottalization in normal speakers. *J. Phonetics* 29, 407-429.
- Shue, Y.-L., Keating, P., Vicenik, C., Yu, K. 2011. VoiceSauce: A program for voice analysis. *Proc. 17<sup>th</sup> ICPhS Hong Kong*, 1846-1849.
- Slifka, J. 2006. Some physiological correlates to regular and irregular phonation at the end of an utterance. *J. Voice* 20, 171-186.
- Sun, X. 2002. Pitch determination and voice quality analysis using Subharmonic-to-Harmonic Ratio. *Proc. ICASSP Orlando*, 333-336.
- Tehrani, H. 2022. EGGworks. <http://www.linguistics.ucla.edu/faciliti/sales/software.htm> (accessed April 24, 2022; no longer available).
- Weenink, D. 2009. The klattgrid speech synthesizer. *Proc. Interspeech 2009*, 2059-2062.
- Whitehead, R. L., Metz, D., Whitehead, B.H. 1984. Vibratory patterns of the vocal folds during pulse register phonation. *J. Acoust. Soc. Am.* 75, 1293-1297
- Yanushevskaya, I., Gobl, C., Ní Chasaide, A. 2018. Cross-language differences in how voice quality and F0 contours map to affect. *J. Acoust. Soc. Am.* 144, 2730-50.
- Yu, K. 2019. Probing voice quality's contribution to tone perception: Challenges for synthesis software. *Proc. 19<sup>th</sup> ICPhS Melbourne*, 2104-2108