THE ACOUSTICS OF PERCEIVED CREAKY VOICE IN AMERICAN ENGLISH

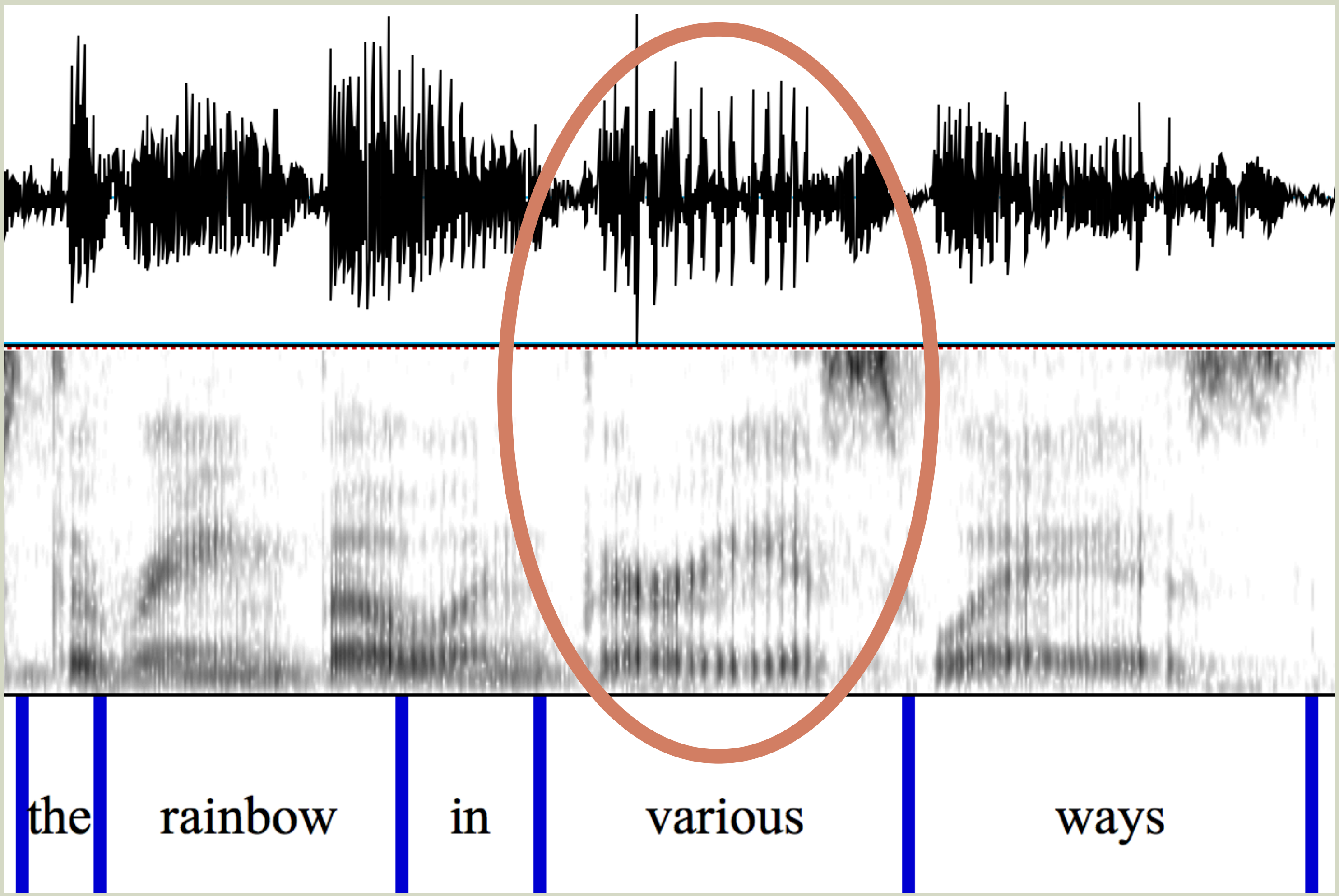
SAMEER UD DOWLA KHAN, REED COLLEGE
KARA BECKER, REED COLLEGE
LAL ZIMMAN, UC SANTA BARBARA

SLIFKA VOICE AND MEN'S SPEECH?

BACKGROUND

- **Creaky voice quality** is typically described as having:
  - Low pitch (f0)
  - Irregularity in pitch track
  - Multiple pulsing
- How can we identify it?
  - **Auditory/visual** impression
  - **Acoustic** correlates

AUDITORY/VISUAL IMPRESSION



ACOUSTIC CORRELATES

- **Pitch**
  - **Low f0**: **English** (Garellek & Keating 2015), Black Miao (Kuang 2013)
  - High f0 (tense voice): Takhian Thong Chong (DiCanio 2009), Black Miao
- **Spectral amplitudes**
  - Low H1c: White Hmong (Esposito 2012)
  - High H2c: Chanthaburi Khmer (Wayland & Jongman 2003)
  - Low H1c-A1c: Jalapa Mazatec (Garellek & Keating 2010), Yi (Kuang 2011), Khmer women
  - Low H1c-A2c: Mazatec
  - Low H1c-A3c: Santa Ana del Valle Zapotec men (Esposito 2010)
  - **Low H1c-H2c**: Mazatec, Chong (DiCanio 2009), Yi, **English**, Zapotec women
- **Regularity/periodicity**
  - **Low HNR** (harmonics to noise ratio): **English** (Garellek & Keating 2015)
  - **Low CPP** (cepstral peak prom.): **English** (Callier & Podesva 2015), Mazatec, Yi
- **Multiple pulsing**
  - **High SHR** (subharmonics to harmonics ratio): **English** (Garellek & Keating 2015)

BACKGROUND

- Some languages **contrast** words based on the use of creak
  - Burmese [lɛ́] ‘falling’ vs. [lɛ̰] ‘crystal’ (Gruber 2011)
  - Santa Ana del Valle Zapotec [bǎd] ‘duck’ vs. [bâ̰d] ‘scabies’ (Esposito 2010)
  - Danish [lɛːsɐ] ‘reader’ vs. [lɛ̰ːsɐ] ‘reads’ (Grønnum 2014)

- In **English**, however, creak is not a contrastive or grammatical feature, rather it is **variable**, influenced by:
  - Phonological factors
  - Social factors

FACTORS FAVORING CREAK

- **Phonological** factors that favor creak in American English:
  - **Vowel onsets** (Pierrehumbert & Talkin 1992; Pierrehumbert 1995; Dilley et al. 1996; Garellek 2013, 2014)
  - **Voiceless stops**, esp. /t/ → [ʔ] (Milroy et al. 1994; Pierrehumbert 1995; Huffman 2005; Eddington & Channer 2010; Eddington & Savage 2012)
  - **IP-final position** (Kreiman 1982; Henton & Bladon 1998; Pierrehumbert & Talkin 1992; Epstein 2002)

- **Gendered** identities associated with creak in the US:
  - **Women in California** (Podesva 2013; Podesva et al. 2015)
  - **Young white women** (Yuasa 2010)
  - **Men** who are (perceived as) **gay/queer** (Podesva 2007; Zimman 2013)
  - **Transgender men** (Zimman 2012, 2013)

QUESTIONS

- How best to identify creak in American English?
- **Auditorily**: are listeners **reliable raters** of creak?
- **Acoustically**: what acoustic features **correlate** with relative ratings of creak?

- To help control for some of the phonological and social variation, we restricted our sample to:
  - Words in **IP-final** position
  - Speech of **transgender men**

RECORDINGS

- Talker group: **5 speakers** of American English
- **Transgender men**
  - Participants in /s/-articulation study (Zimman 2015)
  - Demographic associated with higher use of creak (Zimman 2012, 2013)
- **5 tokens per speaker of IP-final word "bows"**
  - Extracted from recordings of the Rainbow Passage
  - Chosen to capture a range of degrees of creak
  - Pitch range 40-220Hz
  - Scaled to 70dB intensity

LISTENERS

- **14 listeners**, all linguistics majors (9F, 5M)
- **Pre-test**
  - A 6th talker's most unambiguously creaky and modal tokens were used as initial tokens to test for familiarity with phonetic terms
  - A 15th listener was disqualified at this stage
- **Setup**
  - Lab of Linguistics (LoL) at Reed College
  - Run in SuperLab (Cedrus) wearing headphones

TASK

You may repeat a sound as many times as you like.
To repeat: press the spacebar and follow the instructions on the next screen.

When you're ready, please indicate how creaky the sound was by pressing the number key that corresponds to your rating on the scale below. (0, 1, 2, 3, 4, 5)

not creaky	creaky				
0	1	2	3	4	5
not creaky	not very				extremely

RESULTS: RATINGS

- Listeners generally used the full scale of ratings: 0-5
  - 3 listeners used only 1-5, i.e. nothing was devoid of creak
- Mean rating = 1.44, SD = 1.83
- **Ratings across listeners are reliable**
- **Intra-class correlation = 0.615**

Listener	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Max	5	5	5	5	5	5	5	5	5	5	5	5	5	5
Mean	2.12	2.68	1.64	1.56	1.64	2.04	1.88	2.20	3.56	2.72	2.92	2.64	3.04	2.32
Min	0	0	0	0	0	0	0	0	1	1	0	0	1	0
SD	1.96	1.95	2.03	1.58	1.50	2.11	1.83	1.83	1.19	1.31	1.66	2.00	1.72	1.89

ACOUSTIC ANALYSIS

- Vowel portion of each token was subjected to **automated acoustic measurements in VoiceSauce** (Shue 2010)

  f0: frequency of fundamental, calculated by Snack method (Sjölander 2004)
  H1c: amplitude of fundamental, c = corrected (Hanson 1995; Iseli et al. 2007)
  H2c: amplitude of second harmonic
  H1c-H2c: amplitude difference between first and second harmonics
  H1c-A1c: amplitude difference between H1 and harmonics around F1
  H1c-A2c: amplitude difference between H1 and harmonics around F2
  H1c-A3c: amplitude difference between H1 and harmonics around F3
  HNR-05: amplitude ratio of harmonics to noise 0-500Hz (Hillenbrand et al. 1994)
  HNR-15: amplitude ratio of harmonics to noise 0-1500Hz
  HNR-25: amplitude ratio of harmonics to noise 0-2500Hz
  CPP: cepstral peak prominence (de Krom 1993)
  SHR: amplitude ratio of subharmonics to harmonics (Sun 2002)

RESULTS: PITCH

- **Correlations** between mean ratings and acoustic measures
  - Bonferroni-corrected alpha for multiple correlations = .003
- As expected, higher creak ratings were **strongly correlated with lower f0**

Measure	Expected	Observed	Pearson	p-value
f0	lower	lower	-.91	< .001*
H1c	lower	lower	-.02	.914
H2c	higher	lower	-.20	.376
H1c-H2c	lower	higher	.21	.342
H1c-A1c	lower	higher	.50	.018
H1c-A2c	lower	higher	.51	.016
H1c-A3c	lower	higher	.52	.014

RESULTS: SPECTRAL AMPLITUDES

- However, creakier ratings were **not significantly correlated with spectral amplitude measures**
- Non-significant trend in the **unexpected direction** for H1c-A1, H1c-A2c, H1c-A3c

Measure	Expected	Observed	Pearson	p-value
f0	lower	lower	-.91	< .001*
H1c	lower	lower	-.02	.914
H2c	higher	lower	-.20	.376
H1c-H2c	lower	higher	.21	.342
H1c-A1c	lower	higher	.50	.018
H1c-A2c	lower	higher	.51	.016
H1c-A3c	lower	higher	.52	.014

RESULTS: REGULARITY

- In terms of f0 regularity, higher creak ratings were **strongly correlated with lower HNR**, as expected

Measure	Expected	Observed	Pearson	p-value
HNR-05	lower	lower	-.75	< .001*
HNR-15	lower	lower	-.71	< .001*
HNR-25	lower	lower	-.77	< .001*
SHR	higher	lower	-.54	.009
CPP	lower	lower	-.54	.009

RESULTS: MULTIPLE PULSING

- Unexpectedly, higher creak ratings had a non-significant correlation with **weaker cues to multiple pulsing**

Measure	Expected	Observed	Pearson	p-value
HNR-05	lower	lower	-.75	< .001*
HNR-15	lower	lower	-.71	< .001*
HNR-25	lower	lower	-.77	< .001*
SHR	higher	lower	-.54	.009
CPP	lower	lower	-.54	.009

SUMMARY OF RESULTS

- Listeners were generally **reliable** in rating relative creak
- Creakier ratings are significantly correlated with:
  - **Lower pitch** (lower f0)
  - **More irregularity** (lower HNR-05, HNR-15, HNR-25)
- Non-significant correlations in unexpected directions:
  - **Steeper spectral slope** (higher H1c-A1c, H1c-A2c, H1c-A3c)
  - **Weaker cues for multiple pulsing** (lower SHR)

DISCUSSION

- The features we've found resemble **both creaky voice and breathy voice**, which involves a **spreading glottis**
  - Gujarati (Khan 2012)
  - Chanthaburi Khmer (Wayland & Jongman 2003)
  - Santa Ana del Valle Zapotec (Esposito 2010), etc.

- We need to account for the breathy-like qualities seen in the steeper spectral slope of this kind of creak

DISCUSSION

- Recent typology (Keating et al. 2015) distinguishes **6 kinds of creaky voice** on acoustic criteria:
  - **Prototypical creak**: low, irregular f0 + shallower spectral slope
  - **Vocal fry**: regular f0
  - **Multiply pulsed voice**: multiple f0's
  - **Slifka voice**: steeper spectral slope
  - **Tense/pressed voice**: mid/high, regular f0
  - **Aperiodic voice**: no perceived pitch

DISCUSSION

- **"Slifka voice"** (Keating & Garellek 2015)

  a.k.a. "Nonconstricted creak" (Keating et al. 2015)

  a.k.a. "Non-constricted voice" (Keating & Garellek 2015)

  a.k.a. "Less constricted creaky voice" (Garellek & Keating 2015)

- Identified in **acoustic+aerdodynamic** study (Slifka 2000, 2007) as 1 of 3 "voicing termination types"
  - Acoustic: Low, irregular f0
  - Aerodynamic: **Airflow increases** as subglottal pressure falls
  - Interpretation: Slow, spreading glottis, **partially abducted VFs**

FURTHER QUESTIONS

- How common is Slifka voice?
  - Universal in English but **underreported**?
  - More common among **men**?
  - Or specifically **trans men**?
- Very recent studies provide some indication that this could be a feature of **men's speech**

FURTHER QUESTIONS

- IP-final creak in BU news corpus (4 talkers) and lab speech was found to have (Garellek & Keating 2015):
  - Low f0, low HNR, high SHR
  - **Variable spectral slope: steeper for men?**
- IP-final creak in interviews of Central Californians found **gender-based variation** (Callier & Podesva 2015):
  - Women produce shallower spectral slope towards creak
  - **Men produce steeper spectral slope** towards creak
- Unfortunately, neither study discusses cis/trans distinction, or looks at spectral slope measures other than H1c-H2c

FURTHER QUESTIONS

- We're investigating this in **ongoing work** (Becker et al. 2015) looking at **creak across 8 gender+sex identities**:

	Women	Men	Non-binary
Assigned female at birth, not taking testosterone	Cis women	Trans men not taking T	AFAB NB not taking T
Assigned female at birth, taking testosterone	(N/A)	Trans men taking T	AFAB NB taking T
Assigned male at birth	Trans women	Cis men	AMAB NB

- We look forward to reporting our results soon

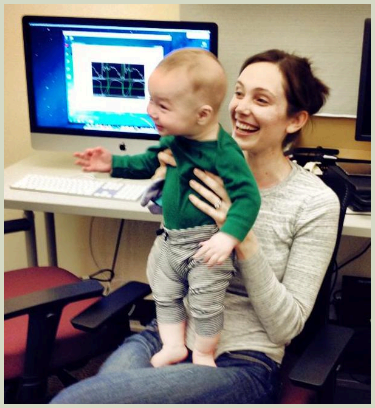
CONCLUSIONS

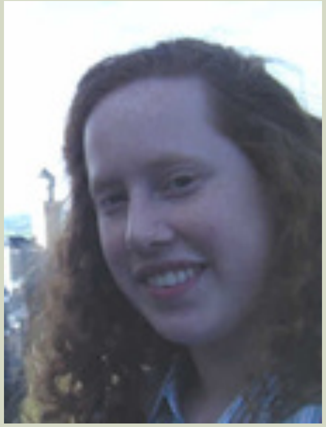
- Creak was **reliably rated by listeners**
- Ratings were correlated with:
  - **Low pitch*** (low f0)
  - **More irregularity*** (low HNR-05, HNR-15, HNR-25)
  - **Steeper spectral slope** (high H1c-A1c, H1c-A2c, H1c-A3c)
  - **Less audible multiple pulsing** (low SHR)

- The subtype of creak we found is **Slifka voice**
- Aligns with very recent work suggesting that this **may be a more common voice quality among men**

ACKNOWLEDGMENTS

- This study was funded by Reed College's Stillman Drake Fund and Summer Scholarship Fund.
- **Many thanks** to my coauthors Kara and Lal (pictured below), our listeners, our talkers, our RAs Dean and Molly (also pictured below), and the audience here at ASA!



REFERENCES

Becker, Kara; Khan, Sameer ud Dowla; Zimman, Lal. 2015. Creaky voice in a diverse gender sample: Challenging ideologies about sex, gender, and creak in American English. Presented at NWAV 44, Toronto.

Callier, Patrick; Podesva, Robert. 2015. Multiple realizations of creaky voice: Evidence for phonetic and sociolinguistic change in phonation. Presented at NWAV44, Toronto.

DiCanio, Christian. 2009. The phonetics of register in Takhian Thong Chong. *JIPA* 39(2), 162-188.

Dilley, Laura; Shattuck-Hufnagel, Stefanie; Ostendorf, Mari. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *JPhon* 24, 423–444.

Eddington, D., & Channer, C. 2010. American English has goʔ a loʔ of glottal stops: Social diffusion and linguistic motivation. *American Speech* 85, 338–351.

Eddington, D., & Savage, M. 2012. Where are the moun[ʔə]ns in Utah? *American Speech* 87, 336–349.

Epstein, Melissa A. 2002. Voice quality and prosody in English (Ph.D. thesis). UCLA.

Esposito, Christina M. 2010. Variation in contrastive phonation in Santa Ana del Valle Zapotec. *JIPA* 40(2), 181-198.

Esposito, Christina M. 2012. An acoustic and electroglottographic study of White Hmong phonation. *JPhon* 40, 466-476.

Garellek, Marc. 2013. Production and perception of glottal stops (Ph.D. thesis). UCLA.

Garellek, Marc. 2014. Voice quality strengthening and glottalization. *JPhon* 45, 106-113.

Garellek, Marc & Keating, Patricia. 2015. Phrase-final creak: Articulation, acoustics, and distribution. Presented at LSA Portland.

Gruber, James. 2011. An articulatory, acoustic, and auditory study of Burmese tone. (Ph.D. thesis). Georgetown University.

Grønnum, Nina. 2013. Danish stød: laryngealization or tone. *Phonetica* 70(1,2), 66-92.

Hanson, Helen M. 1995. *Glottal characteristics of female speakers* (Ph.D. dissertation), Harvard University.

Henton, Caroline; Bladon, Anthony. 1988. Creak as a sociophonetic marker. In Hyman, Larry; Li, Charles N. (eds.) *Language, Speech, and Mind*. Routledge 3-29.

Hillenbrand, James M., Cleveland, Ronald A.; Erickson, Robert L. 1994. Acoustic correlates of breathy vocal quality. *JSHR* 37, 769–778.

Huffman, M. K. 2005. Segmental and prosodic effects on coda glottalization. *JPhon* 33, 335–362.

Iseli, Markus; Shue, Yen-Liang Shue; Alwan, Abeer. 2007. Age, sex, and vowel dependencies of acoustic measures related to the voice source. *JASA* 121, 2283–2295.

REFERENCES, CONTINUED

Keating, Patricia; Garellek, Marc; Kreiman, Jody. 2015. Acoustic properties of different kinds of creaky voice. *Proceedings of ICPhS*.

Keating, Patricia & Garellek, Marc. 2015. Acoustic analysis of creaky voice. Presented at LSA Portland.

Kreiman, Jody. 1982. Perception of sentence and paragraph boundaries in natural conversation. *JPhon* 10, 163-175.

de Krom, G. 1993. A cepstrum-based technique for determining a harmonics-to- noise ratio in speech signals. *JSHR* 36, 224–266.

Kuang, Jianjing. 2011. Production and perception of the phonation contrast in Yi (Master's thesis). UCLA.

Kuang, Jianjing. 2013. The tonal space of contrastive five level tones. *Phonetica* 70, 1-23.

Pierrehumbert, Janet. 1995. Prosodic effects on glottal allophones. In Fujimura O; Hirano, M. (eds.), *Vocal fold physiology*, pp. 39–60. Singular.

Pierrehumbert, Janet; Talkin, David. 1992. Lenition of /h/ and glottal stop. In Docherty, G.J. & Ladd, D.R. (eds.), *Papers in LabPhon II*, 90–117. Cambridge University Press.

Podesva, Robert J. 2007. Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Sociolinguistics* 11: 478-504.

Podesva, Robert J. 2013. Gender and the social meaning of non-modal phonation types. *Proceedings of BLS37*, 427-448.

Shue, Yen-Liang. 2010. The voice source in speech production: Data, analysis and models (Ph.D. thesis). UCLA.

Sjölander, Kåre. 2004. Snack Sound Toolkit. KTH. Available online: http://www.speech.kth.se/snack/

Slifka, Janet. 2000. Respiratory constraints on speech production at prosodic boundaries (Ph.D. thesis). MIT.

Slifka, Janet. 2006. Some physiological correlates to regular and irregular phonation at the end of an utterance. *Voice* 20: 171-186.

Sun, X. 2002. Pitch determination and voice quality analysis using Subharmonic-to-Harmonic Ratio. *Proc. ICASSP-2002*, Orlando, 333-336.

Wayland, Ratree; Jongman, Allard. 2003. Acoustic correlates of breathy and clear vowels in the case of Khmer. *JPhon* 31, 181-201.

Yuasa, Ikuko Patricia. 2010. Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech* 85: 315-337.

Zimman, Lal. 2012. Voices in transition: Testosterone, transmasculinity, and the gendered voice among female-to-male transgender people. (Ph.D. thesis), University of Colorado Boulder.

Zimman, Lal. 2013. Hegemonic masculinity and the variability of gay-sounding speech: The perceived sexuality of transgender men. *Journal of Language & Sexuality* 2(1): 5-43.

Zimman, Lal. 2015. Transmasculinity and the voice: Gender assignment, identity, and presentation.