Multiple Realizations of Creaky Voice

Evidence for Phonetic and Sociolinguistic Change in Phonation

Patrick Callier
Lab41

Robert J. Podesva
Stanford University

Introduction

Sociolinguists and phoneticians have enriched our understanding of creaky voice's social distribution and phonetic properties.

Phonetics

- plurality of realizations (e.g., Keating, Garellek, and Kreiman 2015)
- detailed analysis of smaller, more socially uniform datasets

Sociolinguistics

- complex social patterning
- less detailed acoustic analyses in larger, socially stratified datasets (e.g., Stuart-Smith 1999, Podesva 2013, papers in this session)

Our Central Claims

1. The phonetic realization of creaky voice is constrained by phrase position.
2. Phrase position effects are socially constrained, as younger speakers expand the range of prosodic environments in which creak occurs.

Public Discourses About Creaky Voice

A Vocal 'Pandemic' Among Young Women — THE HUFFINGTON POST

Singers like Britney Spears slip into vocal fry when hitting low notes or for sultry effect, noted *Science* NOW's Marissa Fessenden, characterizing the creak as a "language fad." Kim Kardashian is guilty of it. So is Zooey Deschanel. — TIME

Listen to ***Slate***'s show about the much-reviled phenomenon, prominent among young American women, of "creaky voice." — Slate

the annoying young woman

Social Distribution of Creaky Voice

Variationist work complicates ideologies circulating in the media.

Social perception studies: creak not always evaluated negatively

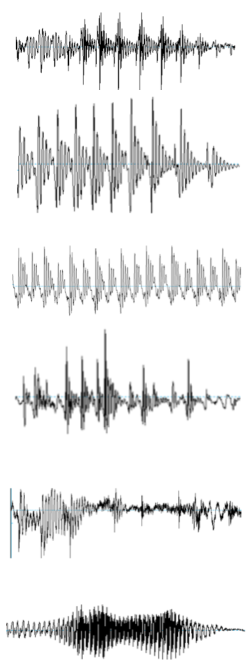
- Only older listeners evaluated creaky negatively (Eckert 2013)
- Creaky samples judged as sounding "professional" (Yuasa 2010)

Production studies: prevalence of creak outside the speech of young (white) women

- Men in UK (Esling 1978, Henton & Bladon 1988, Stuart-Smith 1999)
- Chicano character types (cholo, gangster) in media representations (Mendoza-Denton 2011)
- Women of all ages in DC, including African Americans (Podesva 2013)

Most sociolinguistic work has identified creak using auditory methods, which cannot differentiate different types of creak.

Different Kinds of Creak



phonetic property →	low F0	irregular F0	glottal constriction	damped pulses	sub-harmonics
main acoustic correlate →	(low F0)	(high noise)	(low H1-H2)	(low noise, narrow BW)	(high SHR)
1. prototypical	✓	✓	✓		
2. vocal fry	✓		✓	✓	
3. multiply pulsed		✓	✓		✓
4. aperiodic	NO	✓	✓		
5. nonconstricted	✓	✓	NO		
6. tense	NO		✓		

Keating, Garellek, and Kreiman (2015)

Three Approaches to Characterizing Creak

Single Acoustic Measure: H1*-H2*

- open quotient, inversely correlates with degree of glottal constriction
- interpretation: low values indicative of creakier phonation
- pro: nearly all types of creak characterized by glottal constriction
- con: does not correlate with nonconstricted creak (Slifka 2006)

Single Acoustic Measure: CPPS

- cepstral peak prominence (smoothed), correlates with degree of periodicity
- interpretation: low values indicative of creakier phonation
- pro: captures most types of creak, including nonconstricted
- con: also correlates with other less periodic (e.g., breathy) phonation types

Multiple Acoustic Measures: Creak Classification (Kane, Drugman & Gobl 2013)

- classification using neural network model of multiple acoustic measures
- interpretation: all intervals classified as ±creak
- pro: holistic, binary coding may approximate (some listeners') perception
- con: kinds of creak undifferentiated

Positional Constraints on Creaky Voice

Creaky voice generally favored in phrase-final position (e.g., Henton and Bladon 1988, Ogden 2011, Podesva 2013)

Stylistic use of non-final creak

- The most burned-out burnout uses more non-final creak than the most squeaky-clean jock (D'Onofrio, Hilton, and Pratt 2013).
- Chinese listeners evaluate non-final creak differently from final creak (Callier 2014).
- A Japanese adult video actress exhibits increased use of non-final creak (Kajino and Moon 2011) in “sexy” talk.

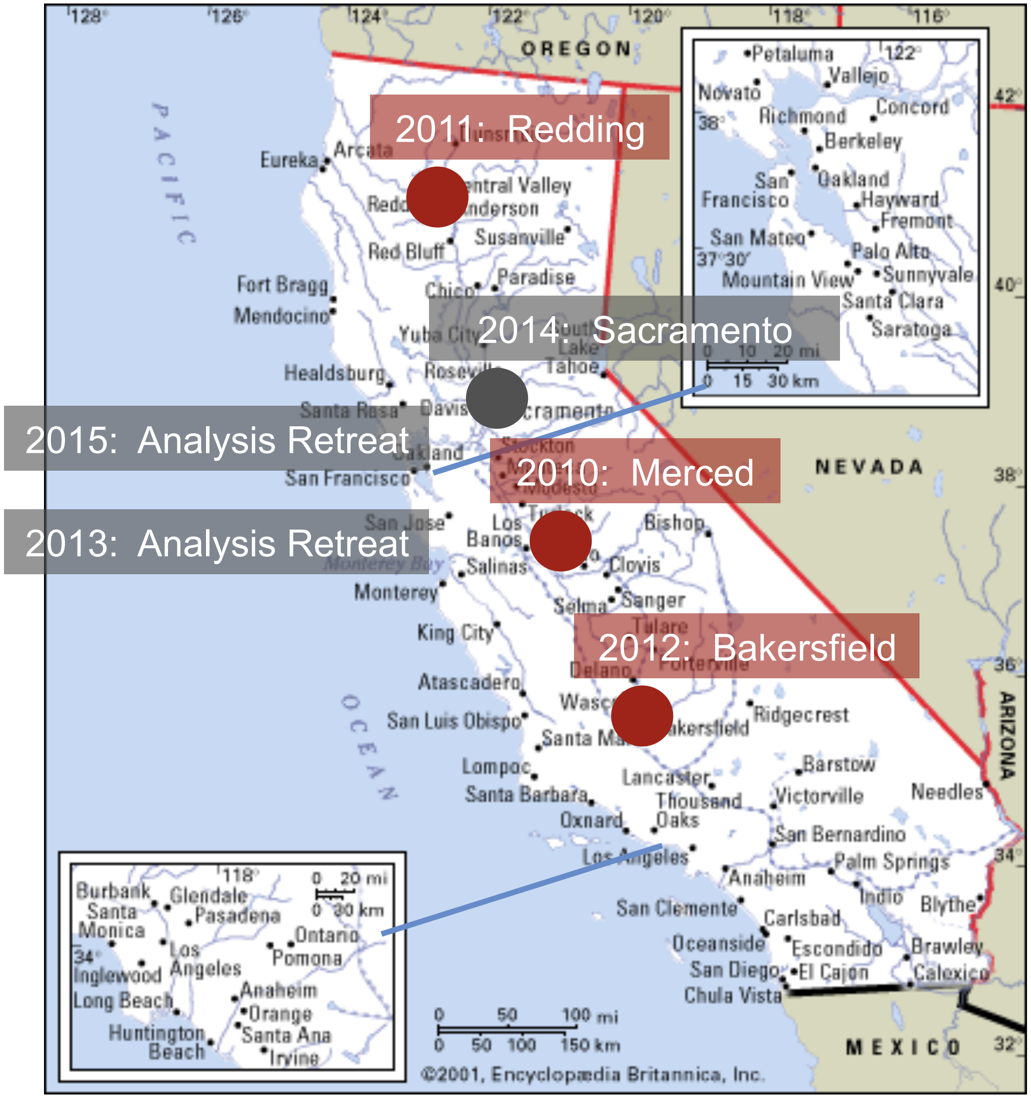
Extensive Creaky Voice

Interviewer:	When- When did they- When did your parents *get* a divorce?
Jessica:	Uh- Shortly after (.) *we had moved there*,
Jessica:	They were in the *process of getting a divorce*
Interviewer:	Oh I see.
Jessica:	So we moved up there,
Jessica:	And then,
Jessica:	*They decided to get a divorce so we moved back*.

Our Central Claims

1. The phonetic realization of creaky voice is constrained by phrase position.
2. Phrase position effects are socially constrained, as younger speakers expand the range of prosodic environments in which creak occurs.

Data



Roughly hour-long sociolinguistic interviews by student and faculty fieldworkers for Voices of California Project

3 field sites

- Redding
- Merced
- Bakersfield

Sample

93 white speakers					
32 from Bakersfield		31 from Merced		30 from Redding	
16 female	16 male	16 female	15 male	16 female	14 male
22-90 years old	24-81 years old	26-93 years old	18-90 years old	18-73 years old	18-63 years old

1/3 of the speakers (represented in all cells) earn their livelihood off the land (e.g., agriculture, ranching, logging, oil).

Methods

Annotation

- Orthographic transcriptions in ELAN (Lausberg & Sloetjes 2009) or Transcriber (Barras et al. 1998)
- Forced alignments generated with FAVE (Rosenfelder et al. 2011)

Extraction of Acoustic Measurements

- Measurements taken for all vowel intervals every 10 ms in Praat (Boersma and Weenink 2015), based on methods in Vicenik (nd), Iseli et al. (2007), Shue (2009)
- Spectral tilt: **H1*-H2***, H1*-A1*, H1*-A2*, H1*-A3*, A1*-A3*, 2k-5k
- Periodicity: cepstral peak prominence (CPP), **smoothed CPP (CPPS)**, harmonics-to-noise ratio (HNR), HNR on low-pass filtered spectrum (500Hz: HNR05, 1500 Hz: HNR15, 2500 Hz: HNR25)
- Nasality: A1*-P0
- F0, F1, F2, intensity

Methods

Post-Processing

- Data reduced to one record (median) per vowel segment
- Exclusions
  - phone duration ≤ 50ms or ≥ 283 ms (median of log duration + 2 s.d.)
  - outliers (± 2 s.d) in F1, F2, intensity, log F0, A1*-P0
- Phrase segmentation from pauses; position in phrase from 0 to 1, based on vowel midpoint
- Preceding and following segments from aligned TextGrids
- Intensity normalization by speaker mean intensity
- Word frequency from in-corpus token count

Creak Detection

- All vowels coded as **±creaky** by a neural network classifier
- MATLAB implementation of Kane, Drugman, and Gobl (2013) algorithm, which factors in a number of acoustic parameters

Methods

Statistical Analysis

Response	H1*-H2* (linear model), CPPS (linear model), ±creaky (logistic model)	
Random Effects (intercepts)	speaker, preceding sound, following sound	
Fixed Effects	Linguistic	Social
	F1, F2	sex
	F0*	age (linear and quadratic terms)
	phone duration*	field site
	intensity	land orientation
	word frequency*	
	phrase position	
	IP duration*	
	A1*-P0	

All continuous variables were scaled and centered.
* log-transformed to ensure normal distribution

Stanford University

H1*-H2*: Linguistic Factors

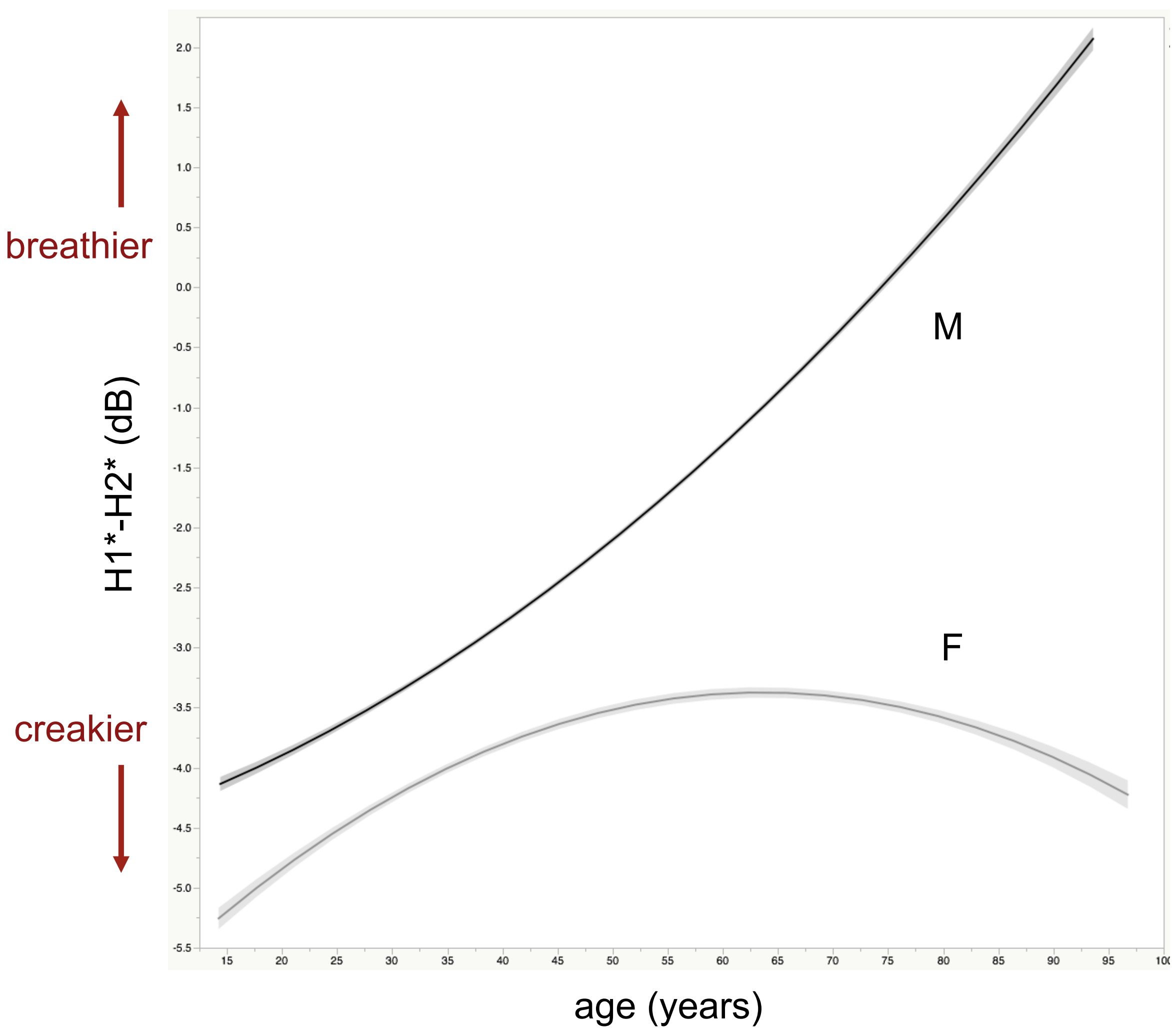
Previous Results (Podesva, Callier, and Szakay 2015)

Creaky voice stronger

- at lower F0 (F0 effect decreases with intensity)
- for vowels exhibiting longer duration
- at later phrase positions (for female speakers only)
- with greater nasality for men, lesser for women
- for words with higher frequency

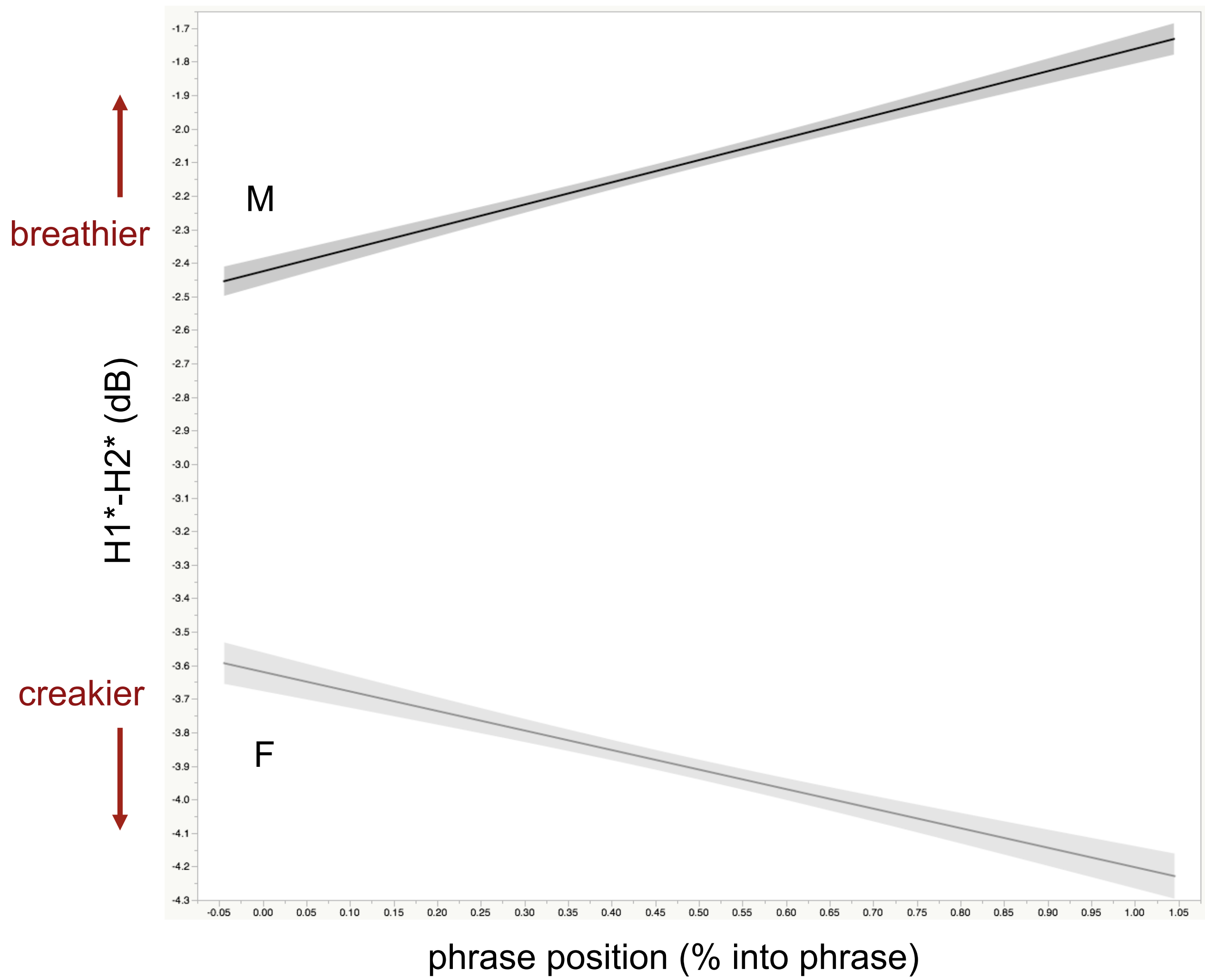
All linguistic factors and relevant interactions were included in statistical models that incorporate social factors.

H1*-H2*: Sex * Age Interaction



- Women are creakier than men (in spite of Simpson's 2012 finding that H1-H2 inflates breathiness values for females).
- Older men are breathier than younger men (linear term for age interacts with gender).
- Women show a curvilinear pattern, with highest incidence of creak among the youngest and oldest women (quadratic term for age interacts with gender).

H1*-H2*: Sex * Phrase Position Interaction



- Men appear to exhibit the opposite pattern from women… but are they really becoming less creaky at the ends of phrases?

- Women exhibit the canonical pattern, with the degree of creak increasing at later phrase positions.

CPPS

cepstral peak prominence (smoothed), correlates with degree of periodicity

interpretation: low values indicative of creakier phonation

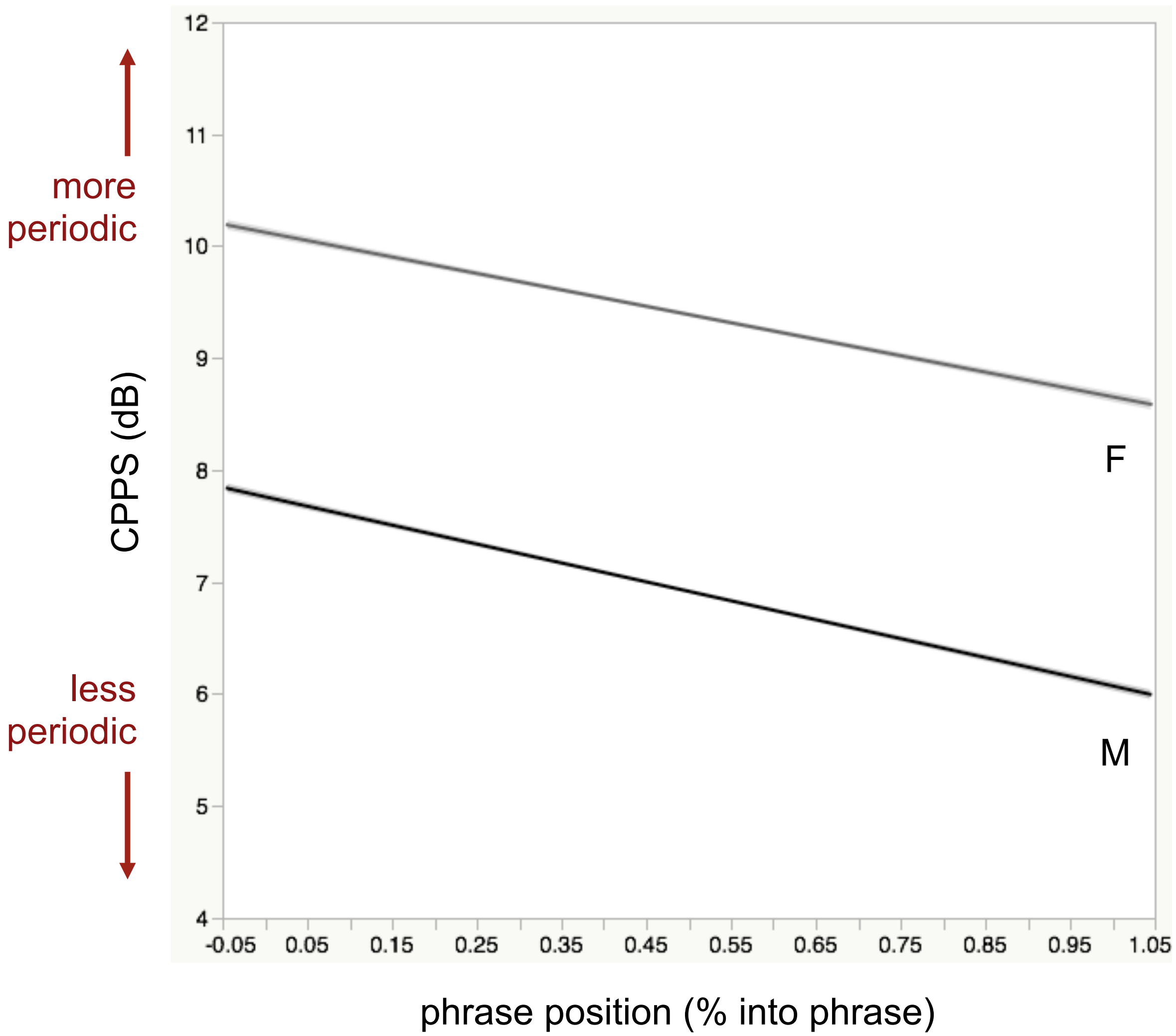
Linguistic Factors

- generally the same as H1*-H2*, except
- longer vowels more periodic than short vowels

Social Factors

- women more periodic than men
- no effects of age

CPPS: Phrase Position (by Sex)



- Both women and men become less periodic as the phrase progresses.
- Women are creakier at the ends of phrases (H1*-H2* and CPPS patterns converge).
- Men might be creakier at the ends of phrases (CPPS patterns could indicate increased breathiness).

Creak Detection

All vowels classified as ±creaky using Kane, Drugman, and Gobl's (2013) neural network model, which factors takes several acoustic measures into account.

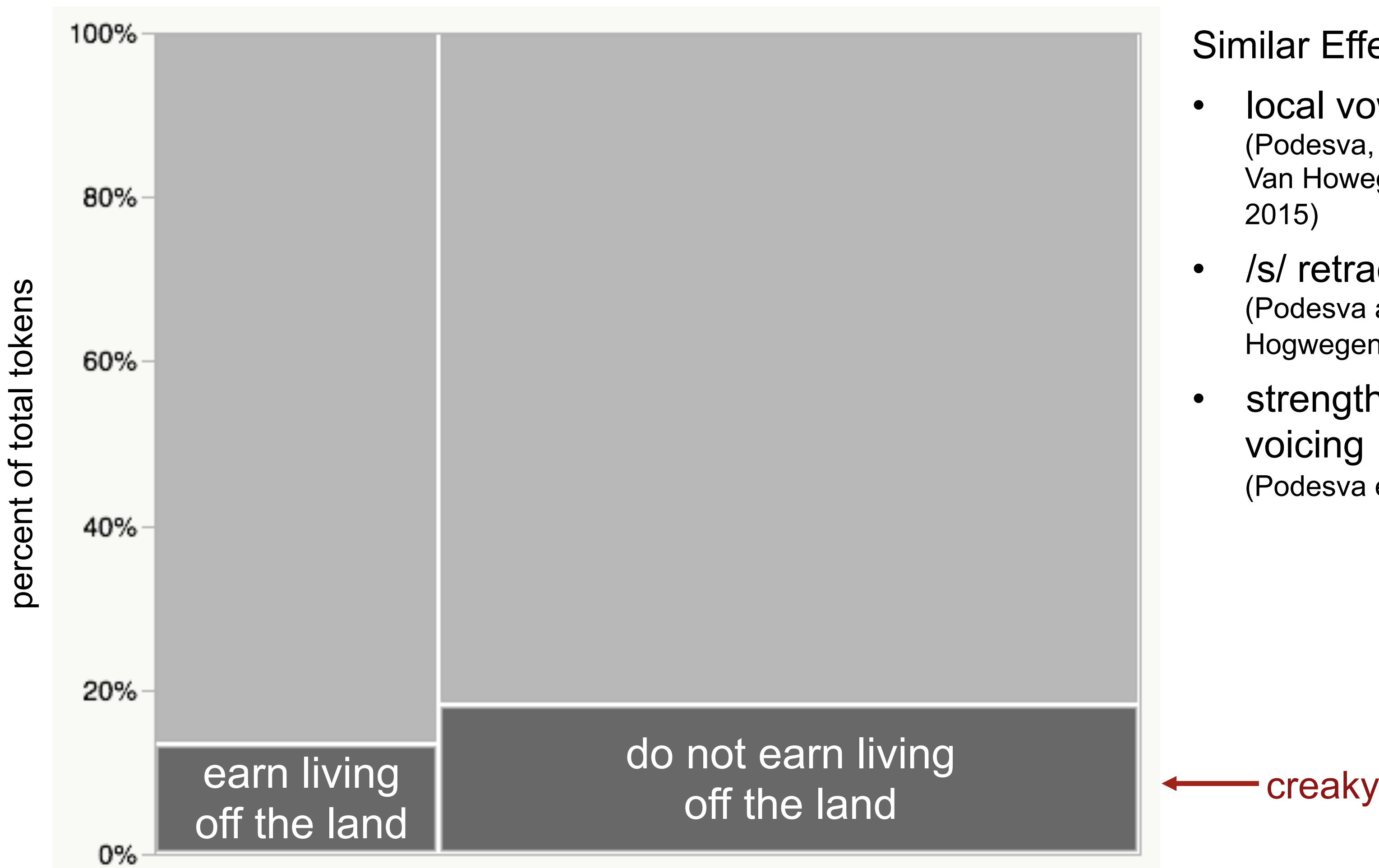
Linguistic Factors

- generally the same as H1*-H2*
- exception: higher incidence of creak for vowels in shorter phrases

Social Factors

- generally the same as H1*-H2*, including sex, age, and interaction
- exception: less creak among land-oriented speakers
- exception: interaction between phrase position and age

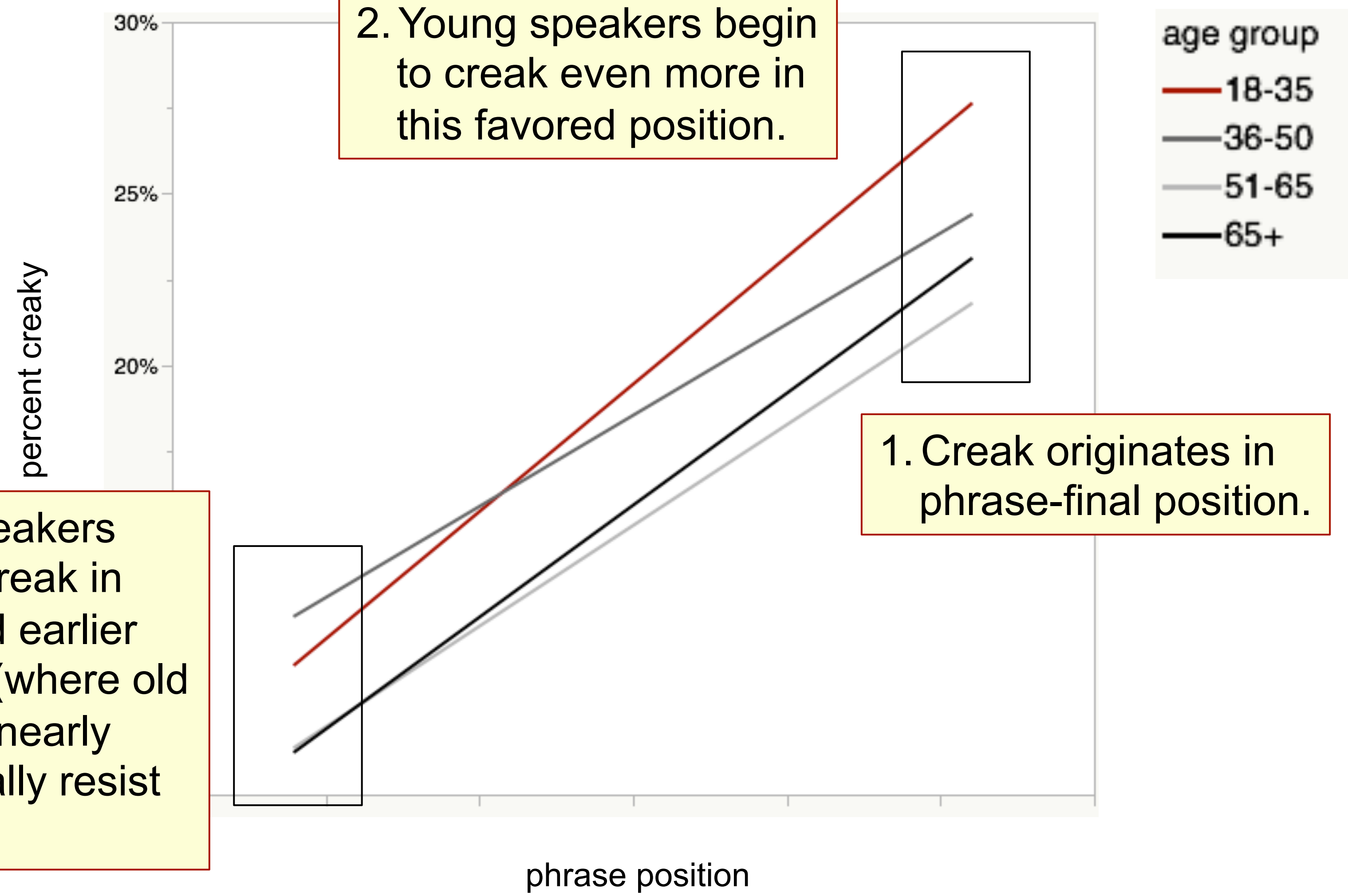
Creak Detection: Relationship to Land



Similar Effects

- local vowel shift (Podesva, D'Onofrio, Van Howegen, and Kim 2015)
- /s/ retraction (Podesva and Van Hogwegen 2014)
- strength of stop voicing (Podesva et al. 2015)

Creak Detection: Age * Phrase Position Interaction



2. Young speakers begin to creak even more in this favored position.

1. Creak originates in phrase-final position.

3. Young speakers begin to creak in disfavored earlier positions (where old speakers nearly categorically resist creaking).

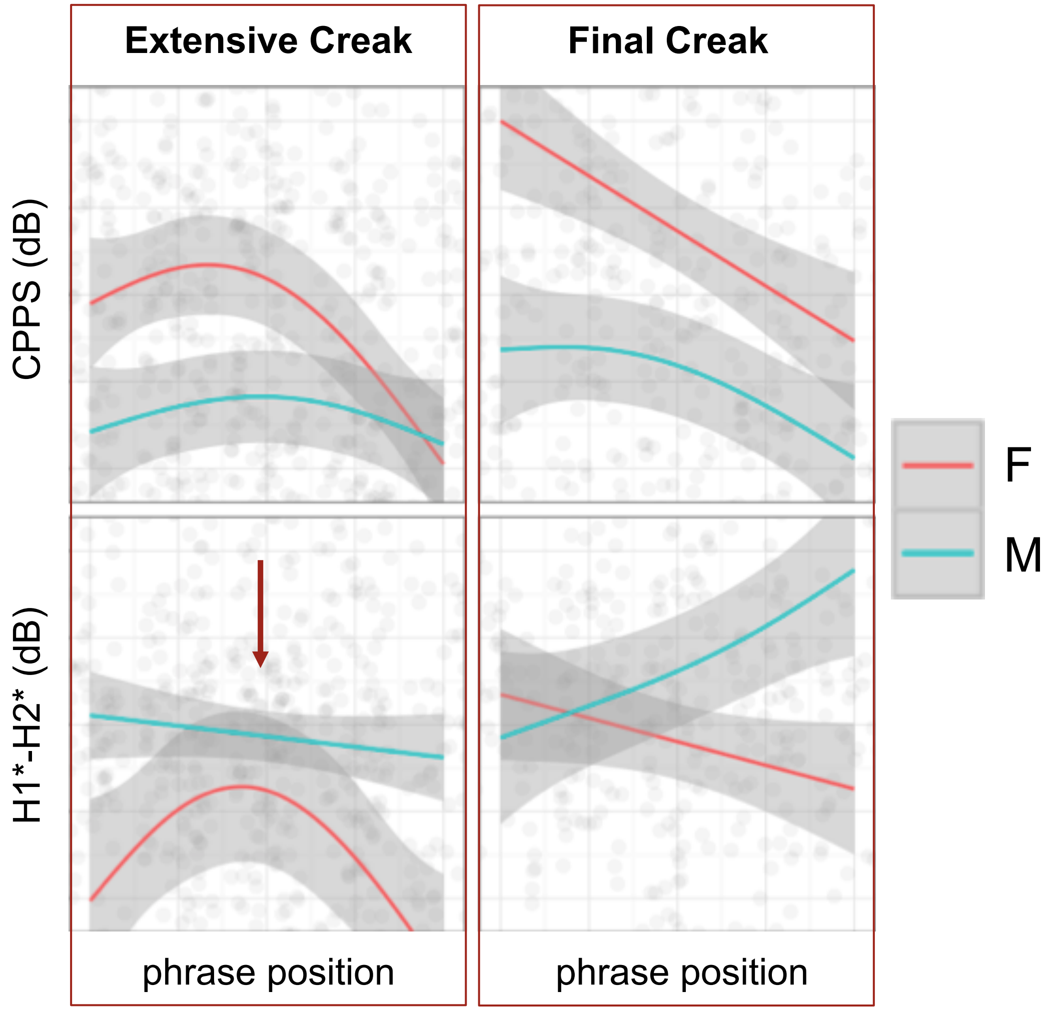
Acoustic Character of Extensive Creak

Hand-coded subset (500 random phrases)

- Final creak (on or after nuclear phrase accent)
- Extensive creak (one or more syllables before nuclear phrase accent)

The final creak pattern is the same as that in the full corpus, suggesting sex class differences in implementation of creak.

The sex class difference does not hold for extensive non-final creak, where women and men both show increased glottal constriction (low H1*-H2*).



Conclusion

Summary of Social Distribution

Sex	Women exhibit stronger, more common creak than men. But strong creak among young men and older women, too!
Age	Change in apparent time, achieved by 1. more phrase-final creak 2. expansion of domain to earlier in the phrase
Land	Speakers who earn their living off the land creak less (cf. Yuasa's 2010 claim about urbanity)

Conclusion

Summary of Phonetic Variation in Realization of Creaky Voice

1. Sex differences in the realization of final creak

   Men exhibit higher H1*-H2* (decreased glottal constriction)
   (cf. Slifka 2006)

2. Positional differences in the realization of creak

   Non-final, extensive creak characterized by uniformly low H1*-H2* (increased glottal constriction), exhibiting no sex differences

Conclusion

Implications for Public Discourses About Creak

Creak's appearance in recent public discourses may be due to

- increased use
- distinctive acoustic character in non-final position, where it is gaining ground

Creak is prevalent among women *of a variety of ages*, as well as young *men*.

Conclusion

Value of Taking Multiple Approaches

A single acoustic measurement may not always be available as a proxy for a phenomenon of interest.

Example: creaky voice and phrase position among men

1. CPPS lowers as the phrase progresses.
   But is it breathier or creakier?
2. Creak detection shows increased incidence of creak.
   But what kind of creak is it?
3. H1*-H2* increases as the phrase progresses.

   So men exhibit a decreased degree of glottal constriction (i.e., a nonprototypical type of creaky voice) for final creak.

Conclusion

Future Work

Trans men and the biological vs. learned basis of sex differences (collaboration with Zimman)

Discursive, interactional, and embodied contexts in which speakers creak

> Podesva, Callier, Voigt, and Hilton (this conference)
> Creaky voice more common when speakers move less, aren't smiling, and report feeling less comfortable.

Understanding the range of social meanings that creaky voice conveys is essential for understanding its trajectory of change.

Thank You!

Many thanks to the Richard A. Karp Foundation and Stanford University for funding data collection.

Questions?

patrick.r.callier@gmail.com
podesva@stanford.edu

Summary of Model for H1*-H2*

Term	Estimate	Std Error	DFDen	t Ratio	Prob > |t|
Intercept	-0.13788	0.065776	92.65	-2.1	0.0388*
F1	0.1817783	0.001491	5.00E+05	121.93	<.0001*
F2	-0.020828	0.001306	5.00E+05	-15.95	<.0001*
log_phone_duration	-0.013746	0.001249	5.00E+05	-11	<.0001*
log_f0	0.3393264	0.001652	5.00E+05	205.39	<.0001*
log_phone_duration*log_f0	0.0070103	0.001166	5.00E+05	6.01	<.0001*
intensity	-0.239248	0.001319	5.00E+05	-181.3	<.0001*
log_phone_duration*intensity	0.0176932	0.001159	5.00E+05	15.27	<.0001*
log_f0*intensity	-0.038822	0.001107	5.00E+05	-35.06	<.0001*
log_phone_duration*log_f0*intensity	0.015482	0.001055	5.00E+05	14.67	<.0001*
log_word_frequency	-0.00194	0.001232	4.00E+05	-1.57	0.1154
position_in_ip	0.0057479	0.001346	4.00E+05	4.27	<.0001*
log_ip_duration	-0.006916	0.001187	5.00E+05	-5.83	<.0001*
log_ip_duration*position_in_ip	0.0112514	0.001178	5.00E+05	9.55	<.0001*
sex[female]	-0.360134	0.04706	87.07	-7.65	<.0001*
sex[female]*log_f0	0.2848269	0.001504	5.00E+05	189.43	<.0001*
sex[female]*position_in_ip	0.0160396	0.001116	5.00E+05	14.37	<.0001*
sex[female]*log_ip_duration	0.0010739	0.001152	5.00E+05	0.93	0.3513
sex[female]*log_ip_duration*position_in_ip	0.0049335	0.00111	5.00E+05	4.44	<.0001*
sex[female]*A1P0	-0.044968	0.001233	5.00E+05	-36.46	<.0001*
A1P0	0.0538803	0.001422	5.00E+05	37.9	<.0001*
age	0.2167578	0.047353	87.06	4.58	<.0001*
sex[female]*age	-0.092431	0.047353	87.06	-1.95	0.0542
age*log_f0	0.0931864	0.001584	5.00E+05	58.84	<.0001*
age.q	-0.02714	0.042568	87.05	-0.64	0.5254
sex[female]*(age.q-1)	-0.127123	0.042568	87.05	-2.99	0.0037*
age*A1P0	0.0130678	0.001256	5.00E+05	10.41	<.0001*
sex[female]*age*A1p0	-0.018682	0.001254	5.00E+05	-14.9	<.0001*
sex[female]*age*log_f0	0.0008053	0.001585	5.00E+05	0.51	0.6115
log_f0*(age.q-1)	0.05294	0.001416	5.00E+05	37.39	<.0001*
sex[female]*log_f0*(age.q-1)	-0.027165	0.001414	5.00E+05	-19.21	<.0001*

Summary of Model for CPPS

Term	Estimate	Std Error	DFDen	t Ratio	Prob>|t|
Intercept	4.0220218	0.253624	164.6	15.86	<.0001*
F1	0.0054344	6.15E-05	3.00E+05	88.4	<.0001*
F2	0.0011726	2.04E-05	3.00E+05	57.45	<.0001*
log_phone_duration	0.0532018	0.015269	3.00E+05	3.48	0.0005*
log_word_freq	-0.059068	0.002775	3.00E+05	-21.29	<.0001*
position_in_ip	-2.024564	0.023628	3.00E+05	-85.68	<.0001*
sex[female]	0.9148948	0.205943	87.05	4.44	<.0001*
sex[female]*(position_in_ip)	0.1212415	0.021024	3.00E+05	5.77	<.0001*
log_ip_duration	0.1286475	0.016551	3.00E+05	7.77	<.0001*
log_ip_duration*position_in_ip	-0.25419	0.054274	3.00E+05	-4.68	<.0001*
sex[female]*log_ip_duration	0.1373617	0.016412	3.00E+05	8.37	<.0001*
sex[female]*log_ip_duration*position.in.ip	0.3287143	0.053547	3.00E+05	6.14	<.0001*

Summary of Model for ±Creak

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-1.0007947	0.0713177	196.92	<.0001*
F1	0.00058102	4.83E-05	144.85	<.0001*
F2	-7.45E-05	1.69E-05	19.5	<.0001*
log_phone_duration	0.80940311	0.0128238	3983.8	<.0001*
log_word_freq	0.04823304	0.002319	432.61	<.0001*
position_in_ip	1.09860201	0.0201423	2974.8	<.0001*
log_ip_duration	-0.373264	0.0139903	711.83	<.0001*
sex[female]	0.50244232	0.0065753	5839.1	<.0001*
orientation[land]	-0.1099663	0.0065974	277.83	<.0001*
age	0.02774937	0.0014943	344.84	<.0001*
age_sqrd	-0.0003621	1.56E-05	535.68	<.0001*
sex[female]*(age)	-0.0032192	0.0014958	4.63	0.0314*
sex[female]*(age.q)	0.00012023	0.0000157	58.65	<.0001*
sex[female]*(age)*(position_in_ip)	-0.0032808	0.0010473	9.81	0.0017*
(age)*(position_in_ip)	0.00430619	0.0010475	16.9	<.0001*
sex[female]*(position_in_ip)	0.03696844	0.0198191	3.48	0.0621