The Affective Roots of Gender Patterns in the Use of Creaky Voice

ROBERT J. PODESVA
Stanford University
Introduction

Prevalence of creaky voice among young women
- Popular media
  (e.g., Fessenden 2011, Steinmetz 2011, Quenqua 2012, Arana 2013)
- Scholarly research

An exclusively gendered discourse about creak fails to explain why women might creak more than men.

Creaky voice serves a negative, disengaged affective function (which underlies gender differences in its use).
Two Questions

1. To what extent do (young) women produce more creaky voice than men (and older women)?

   - Quantitative variationist study of 93 speakers from California examining the incidence and acoustic character of creaky voice

2. Why do women (and men) produce creaky voice?

   - Quantitative study investigating correlations between non-vocal expressions of affect and the production of creaky voice, among 42 speakers from California (and the Western US)
A Vocal 'Pandemic' Among Young Women

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.

Listen to Slate’s show about the much-reviled phenomenon, prominent among young American women, of “creaky voice.”
vulgar

repulsive

I want the oil to stop frying!

mindless

annoying
Counter-Discourses About Creaky Voice

‘There are languages that use creak as part of the phonemic system,’ says [Linguist Patricia Keating of UCLA]. ‘The chances of it leading to vocal damage are very minimal.”’

Women have long tended to be the linguistic innovators...If you want to see where language is going...you find a young, urban woman.

In large part, the story of language is one of the dominant political group trying to fix the linguistic code in place, and those below them pushing and pulling it loose.
Quantitative Research on Gender Differences in Creak

Previous studies observing more creak among women:
• Young American English-speaking women creak a greater percentage of words than men (Yuasa 2010)
• College-aged women produced more vocal fry in read sentences than a matched sample of men (Wolk, Abdelli-Beruh, and Slavin 2012, Abdelli-Beruh, Wolk, and Slavin 2014)
• White and African American women in Washington, DC, creak a greater percentage of phrases than men (Podesva 2013)

Desired characteristics of quantitative study:
• Larger number of speakers  • Single regional dialect  • Longer samples of spontaneous speech  • Age-stratified speaker sample
Voices of California

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

Creak examined among speakers from 3 field sites

- Redding
- Merced
- Bakersfield
### Sample

<table>
<thead>
<tr>
<th>93 white speakers</th>
<th>32 from Bakersfield</th>
<th>31 from Merced</th>
<th>30 from Redding</th>
</tr>
</thead>
<tbody>
<tr>
<td>16 female</td>
<td>16 male</td>
<td>16 female</td>
<td>15 male</td>
</tr>
<tr>
<td>16 female</td>
<td>16 male</td>
<td>16 female</td>
<td>16 female</td>
</tr>
<tr>
<td>16 female</td>
<td>16 male</td>
<td>16 female</td>
<td>14 male</td>
</tr>
</tbody>
</table>

- 22-90 years old
- 24-81 years old
- 26-93 years old
- 18-90 years old
- 18-73 years old
- 18-63 years old
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, 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 of vowel midpoint in phrase from 0 to 1
• 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: ±creaky (logistic model), H1*-H2* (linear model), CPPS (linear model)

Random Effects: speaker, preceding sound, following sound

Fixed Effects:
- Linguistic: F1, F2, F0*, phone duration*, intensity, word frequency*, phrase position, IP duration*, A1*-P0
- Social: gender, age (linear and quadratic terms), field site, orientation to land

All continuous variables scaled and centered.
* log-transformed to ensure normal distribution
Summary of Results

Factors Influencing Incidence of Creak

Main Effects
- Phrase Position: creak increasingly likely as phrases progress
- Gender: women more likely to creak than men
- Age: younger speakers more likely to creak than older
- Land Orientation: speakers who earn living off land creak less

Interactions
- Phrase Position * Age: age effect varies as function of phrase position
- Gender * Age(q): youngest and oldest women more likely to creak than middle-aged women
1. Creak originates in phrase-final position.

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

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

Women are creakier than men.

Older men are breathier than younger men.

Women show a curvilinear pattern, with highest incidence of creak among the youngest and oldest women.
Interim Conclusion

Do young women creak the most?
• Yes, but...
• Young men and the oldest women creak a lot, too.
• A focus on gender can obscure other relevant social dimensions (e.g. orientation to land).

Why do people use creaky voice?
Counter-Discourses About Creaky Voice

The New York Times

‘If women do something like uptalk or vocal fry, it’s immediately interpreted as insecure, emotional or even stupid,’ said Carmen Fought, a professor of linguistics at Pitzer College in Claremont, Calif. ‘The truth is this: Young women take linguistic features and use them as power tools for building relationships.’

‘A lot of these really flamboyant things you hear are cute, and girls are supposed to be cute,’ said Penny Eckert, a professor of linguistics at Stanford University. ‘But they’re not just using them because they’re girls. They’re using them to achieve some kind of interactional and stylistic end.’
The Social Meaning of 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*
Discourse Analytic Work on Affective Function of Creak

Previous studies suggest that creak conveys negative affect/disengagement:
• Creaky voice distances parenthetical speech from main narrative thread; also to distance speakers from issue under discussion (Lee 2015)
• Creaky *yeah* expresses either disalignment between interlocutors or dispreference to continue on current topic (Grivičić and Nilep 2004)
• Transmasculine speaker creaks to index “a stance of disaffectation, an aloof persona, or a kind of emotional stoicism” (Zimman 2014)

Can this affective function of creak be generalized?
If creaky voice conveys negative, disengaged affect...

<table>
<thead>
<tr>
<th>Modality of Affect Expression</th>
<th>Form of Affect Expression</th>
<th>Predicted Correlation with Creaky Voice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-Report</td>
<td>Comfortable</td>
<td>Less</td>
</tr>
<tr>
<td>Self-Report</td>
<td>Enjoyable</td>
<td>Less</td>
</tr>
</tbody>
</table>
Interactional Sociophonetics Laboratory

Desired characteristics of quantitative study:

- Larger number of speakers
- Audio-visual data
- Comfortable social environment
- Assessments of affect

Acoustical specifications of sound booth, staged as living room
Data Collection

Dyadic interactions, video and audio (wireless microphones) recorded 
~30 min of conversation, with aid of prompts  
Post-recording survey (demographic information, assessments)
Example Recording
Sample: 42 speakers from the Western United States

<table>
<thead>
<tr>
<th>Category</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>26 women&lt;br&gt;16 men</td>
</tr>
<tr>
<td>Age</td>
<td>25 undergraduates (18-22 years old)&lt;br&gt;17 older adults (23 years old and up)</td>
</tr>
<tr>
<td>Race</td>
<td>21 white&lt;br&gt;6 African American/white&lt;br&gt;5 Asian/Pacific Islander&lt;br&gt;3 Asian/white&lt;br&gt;2 Native American&lt;br&gt;2 other multiracial&lt;br&gt;1 each of African American, Latinx, Middle Eastern</td>
</tr>
<tr>
<td>Sexual Orientation</td>
<td>32 straight&lt;br&gt;7 LGBTQ&lt;br&gt;3 unspecified</td>
</tr>
</tbody>
</table>
## Valence, Arousal, and Dominance Lexicons

<table>
<thead>
<tr>
<th>word</th>
<th>valence</th>
<th>arousal</th>
<th>dominance</th>
</tr>
</thead>
<tbody>
<tr>
<td>emptiness</td>
<td>0.177</td>
<td>0.173</td>
<td>0.184</td>
</tr>
<tr>
<td>menace</td>
<td>0.125</td>
<td>0.894</td>
<td>0.759</td>
</tr>
<tr>
<td>floral</td>
<td>0.833</td>
<td>0.235</td>
<td>0.264</td>
</tr>
<tr>
<td>champion</td>
<td>0.89</td>
<td>0.88</td>
<td>0.945</td>
</tr>
</tbody>
</table>

(Mohammad 2018)
Sentiment Analysis

“but no- nothing long term -- I mean, I haven't had that many long term”

(Socher, Perelygin, Wu, Chuang, Manning, Ng, and Potts 2013)
Smiling Annotation

Haar cascade classifier trained on open source corpus of photographs hand-annotated for ±smiling (http://github.com/hromi/SMILE_EsmileD).

Each frame of video run through classifier.
Quantifying Movement

Movement Amplitude
Voigt, Podesva & Jurafsky (2014)
If creaky voice conveys negative, disengaged affect...

<table>
<thead>
<tr>
<th>Modality of Affect Expression</th>
<th>Form of Affect Expression</th>
<th>Predicted Correlation with Creaky Voice</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-Report</td>
<td>Comfortable</td>
<td>Less</td>
<td></td>
</tr>
<tr>
<td>Self-Report</td>
<td>Enjoyable</td>
<td>Less</td>
<td>YES</td>
</tr>
<tr>
<td>Lexical/Semantic</td>
<td>Valence</td>
<td>Lower</td>
<td></td>
</tr>
<tr>
<td>Lexical/Semantic</td>
<td>Arousal</td>
<td>Lower</td>
<td></td>
</tr>
<tr>
<td>Lexical/Semantic</td>
<td>Dominance</td>
<td>Lower</td>
<td>YES</td>
</tr>
<tr>
<td>Lexical/Semantic</td>
<td>Sentiment</td>
<td>More Negative</td>
<td></td>
</tr>
<tr>
<td>Embodied</td>
<td>Smiling</td>
<td>Less</td>
<td>YES*</td>
</tr>
<tr>
<td>Embodied</td>
<td>Body Movement</td>
<td>Less</td>
<td>YES</td>
</tr>
</tbody>
</table>

* female speakers only
Discussion

Data largely support hypothesis that creaky voice conveys negative, disengaged affect (all four significant results, spanning all three modalities, in predicted direction)

Potential explanations for no observed correlation:
  • There is no connection
  • Not enough statistical power (single comfort level rating per interaction) Crowdsourcing to obtain dynamic assessments as interactions unfold?
  • Methods insufficiently robust (sentiment analysis developed for movie review text, not spontaneous speech)
Why do young women creak the most?

- (Young) female speakers creak more than others.
- Regardless of gender, people creak more when moving less, on words that convey low dominance, in less enjoyable interactions.
- Women do not move less, use less dominant lexical items, or enjoy interactions less than men overall.

Women creak more because they appear to prefer creak as a means of enacting a negative, disengaged stance.
Why convey negative, disengaged stance with creak?

The social meaning of creak, though interactionally relevant, does not contribute to the referential meaning of the utterance that carries it.

Because this meaning is not at issue, speakers can express affect without doing so directly (Besnier 1990).

A:  Yeah
B:  # Why are you disengaging?

Plausible deniability of disengagement is useful!
### Meaning of Creak

- **Yeah**
- ‘negative, disengaged affect’
- better than ‘young female’
- but we can probably do better
  - ‘negative’ = evaluative stance
  - ‘disengaged’ = (lack of) alignment between interactants

---

### Meaning of Uptalk

- Assertive rising declaratives (Jeong 2018)
- *My name is Anna? I’ll be your waitress?*

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Speaker committed to proposition</td>
</tr>
<tr>
<td>2.</td>
<td>Speaker introduces metalinguistic issue (MLI) to the Table</td>
</tr>
</tbody>
</table>

- ARDs perceived as more polite, follows from (2)
- ARDs perceived as feminine for similar reasons

---

**Gendered meanings follow from core meaning.**
Conclusion

Gender can’t be the whole story.

The prevalence of creak among women is less about gender than affect expression, which can neither be separated from nor reduced to gender.

Creaky voice is interactionally useful.
• This point combats sexism underlying many linguistic ideologies.
• This point is a descriptive fact.
Thank You!

Thanks to the ETAP4 organizers! The collection of audiovisual data was supported by a grant from the Roberta Bowman Denning Initiative in the Digital Humanities. Data collection for the Voices of California Project was supported by Stanford University and the Richard A. Karp Foundation. Many thanks to audiences at Cornell University, Sheffield University, University of Nevada, Reno, University of California, Merced, Brown University, the Ohio State University, Michigan State University, and the University of California, Berkeley for feedback on segments of the work presented here. Thanks also to participants in a seminar on Language and Embodiment, and Annette D’Onofrio, Penny Eckert, Lauren Hall-Lew, Katherine Hilton, Sunwoo Jeong, Emma Moore, Teresa Pratt, Janneke Van Hofwegen, Rob Voigt, and Rob Xu for discussions of this material.

Questions?
podesva@stanford.edu