Two Pragmatic Functions of Breathy Voice in American English Conversation

Nigel Ward, Ambika Kirkland, Marcin Włodarczak, Éva Székely
University of Texas at El Paso, KTH, Stockholm University

Research Question

What ***pragmatic*** functions does breathy voice serve?

Previous Work on Breathy Voice

Roles in emotional, identity-marking
Occurrence in laughter and sighs
Involvement in turn taking, questions
In Japanese, marks "talking to oneself," etc.

Discovery Methods

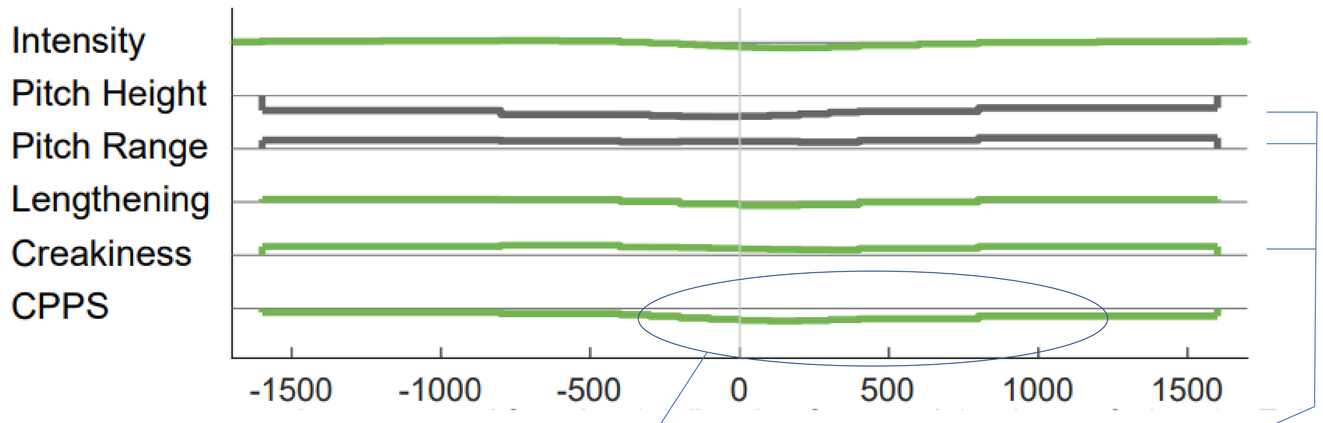
Data: Conversations among students, 80 minutes,
yielding 480K samples

Measure: CPPS "cepstral peak prominence smoothed",
the best single proxy for breathy voice
lower values imply lower harmonicity, more breathiness

Model: Temporal configurations of prosodic features
that may have functions.
found by applying PCA to 236 prosodic features,
including CPPS, computed over 12 windows per speaker

Inductive analysis of meanings of the resulting configurations
by examining exemplars of each

Pattern 1



~1.5 seconds of breathy voice

within a 3-second region of: low pitch,
rather wide pitch
slightly creaky voice

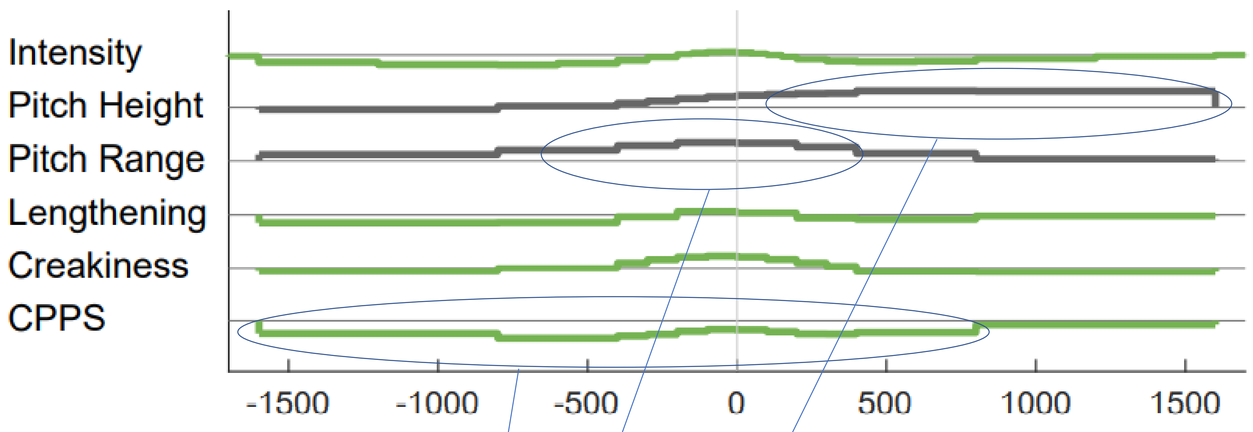
1. need to think about, so I can write it up in the document, to where we're gonna give to the taggers

---------- = breathy

2. and like, I could just imagine being part of something like that

In general this pattern serves to marks the utterance as **self-directed**.

Pattern 2



~2 seconds of breathy voice

with a moment of wide pitch
rising to high pitch

3. okay, so it's been a little while since you've been programming⇑, or // no↑, it, um, it's actually been really quick; I've only had a summer off.

↑⇑ = rising pitch

4. They're searching for content? they're searching for↑ …// yeah⇑, content, um↑ // a feeling↑

In general this pattern serves to marks an attempt at **grounding**.

Perception Study

Hypothesis 1: People will frequently judge stimuli that exemplify Pattern 1 to be "self-directed".

Hypothesis 2: People will frequently judge stimuli that exemplify Pattern 2 to be "working towards common ground".

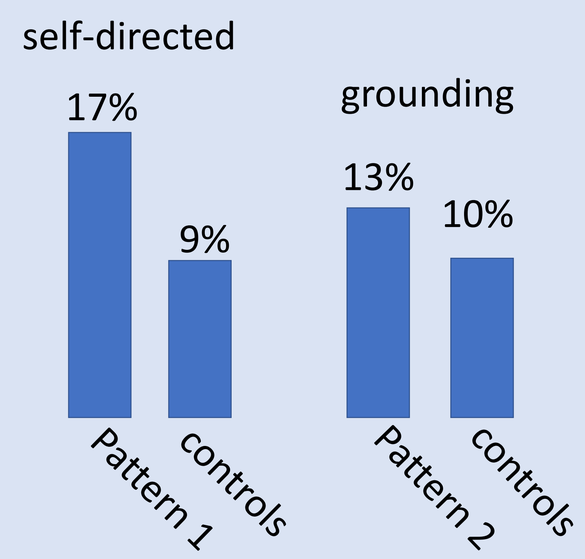
Stimuli: for each, 2 exemplars from each of 6 conversations
with 35 controls

Procedure: "Listen to the audio, and, focusing on the last few seconds, select one or more descriptors that best describe what is happening in the conversation".

Descriptors: {self-directed, other-directed, working towards common ground, explanation, low involvement, high involvement, agreement, elaboration, disagreement, confusion, anger, amusement, hesitation}

Participants: 99 native speakers of English (4% American)

Results:



Differences significant ($p < .002$, χ^2)

Implications

Breathy voice is not just paralinguistic.

Speech synthesizers should be able to control breathiness.

Key References

Heldner, M., Wlodarczak, M., et al. (2019). Voice quality as a turn-taking cue. In *Interspeech 2019.*

Heman-Ackah, Yolanda D., et al. "Cepstral peak prominence: a more reliable measure of dysphonia." *Annals of Otology, Rhinology & Laryngology* 112.4 (2003): 324-333.

Ishi, Carlos T., et al. "Analysis of the roles and the dynamics of breathy and whispery voice qualities in dialogue speech." *EURASIP Journal on Audio, Speech, and Music Processing* 2010 (2010): 1-12.

Panfili, Laura. *Cross-linguistic acoustic characteristics of phonation.* U. Washington Dissertation, 2018.

Ward, Nigel G. *Prosodic patterns in English conversation*. Cambridge University Press, 2019.