

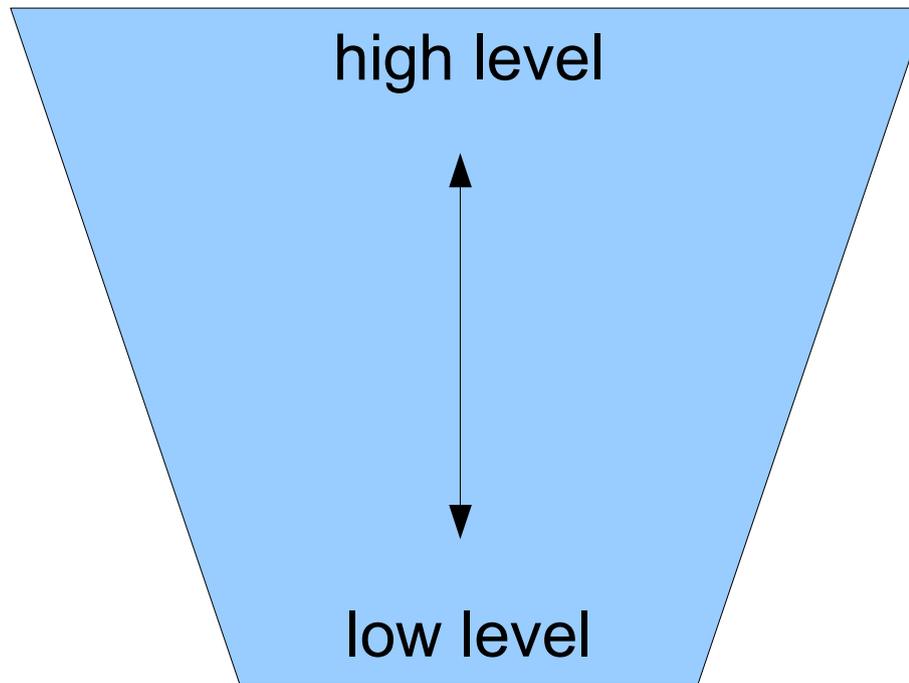
Voice Quality

R. L. Starr

Lecture for Language & Gender

May 26, 2009

Reminder: what aspects of human language can we study?



What can we study?

- Discourse-level:
 - how speakers interact in conversation
 - how speakers structure narratives, etc.

What can we study?

- Sentence-level:
 - How speakers use various types of sentence patterns.
 - Example: “I gave him the box” vs. “I gave the box to him.”

What can we study?

- Word-level:
 - What words speakers use in various situations.
 - Example: “he was hella cute” vs. “he was mad cute”

What can we study?

- Segmental-level:
 - Patterns of sound changes in the segments (or, sounds) that make up speech.
 - Segments: cat = /k/ /æ/ /t/ 3 segments

What can we study?

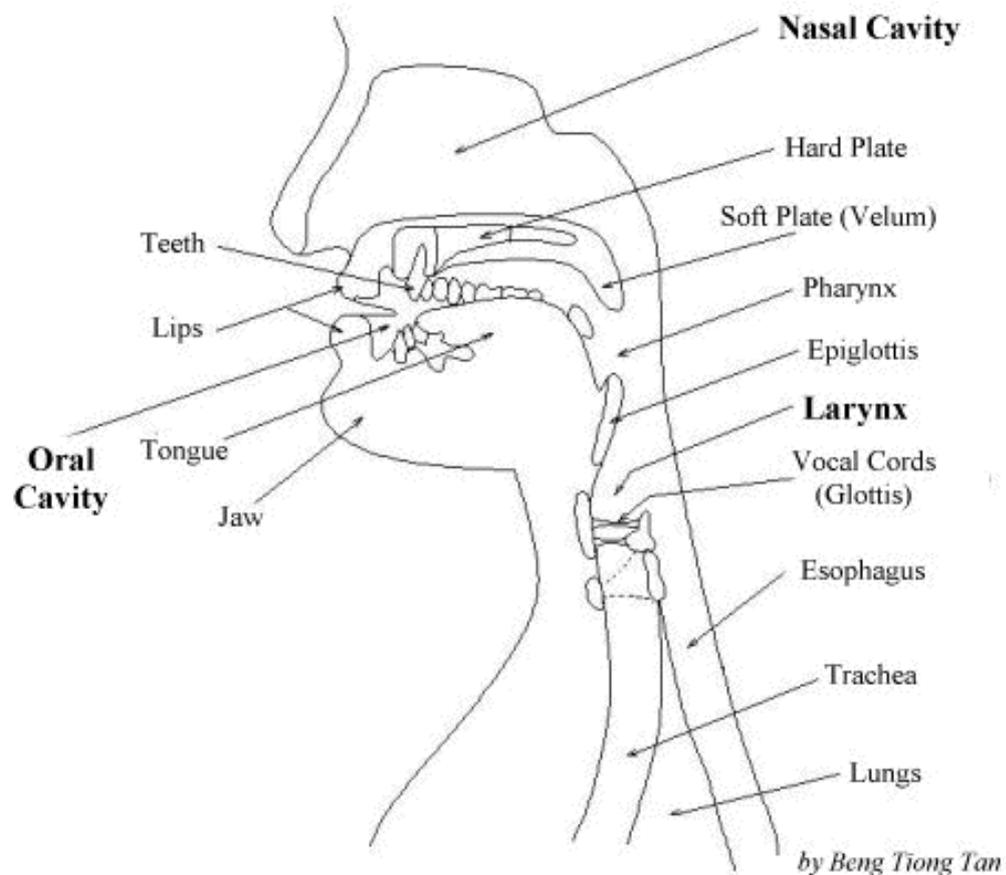
- Suprasegmental-level:
 - How speakers change their pitch and loudness over the course of their speech.
 - Example: “*Are legumes a good source of vitamins?*” vs. “*Are legumes a good source of vitamins?*”
 - How quickly or slowly speakers talk over the course of their speech.
 - How speakers change their **voice quality** over the course of their speech.

Suprasegmental Features

- We call these features “suprasegmental” because they are overlaid on top of the segments of speech.
- They can affect more than one segment at a time.

What is voice quality?

- To produce speech, we move air through our vocal apparatus:

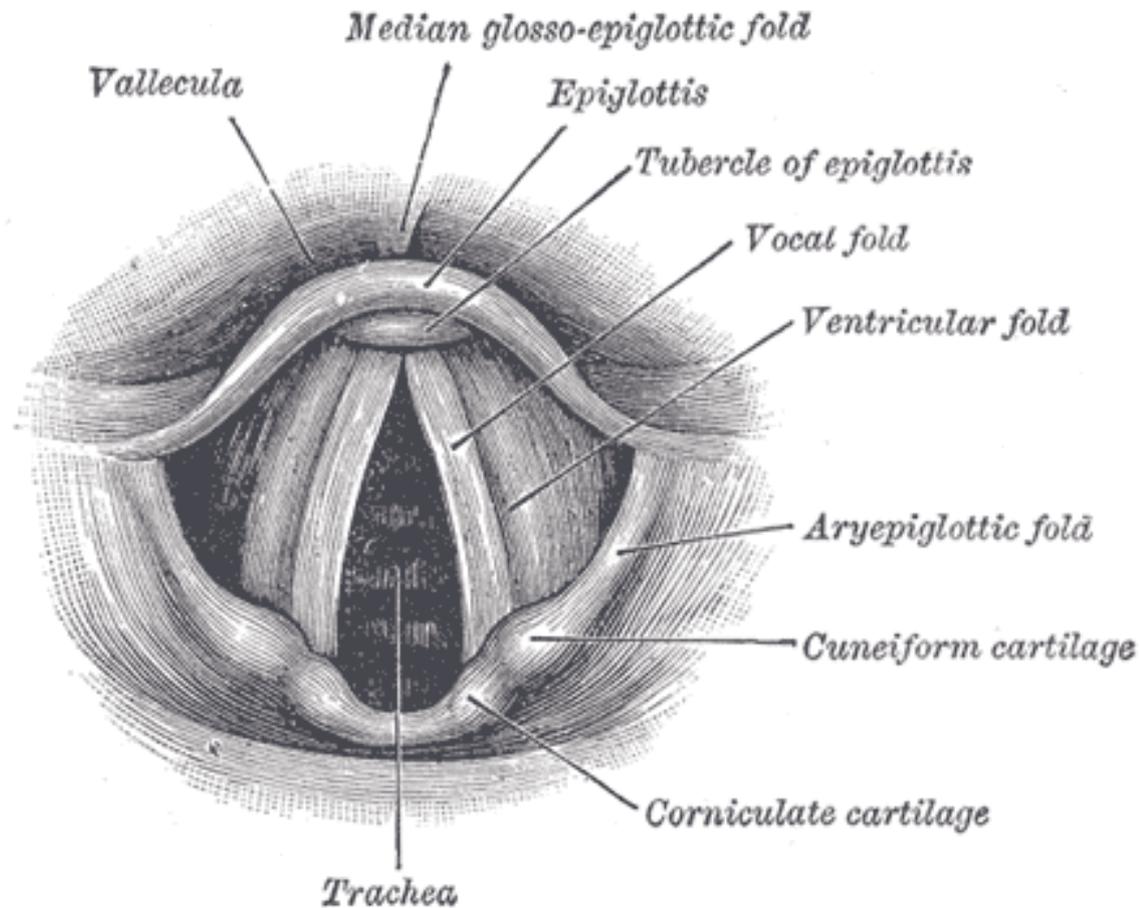


What is voice quality?

- Voice quality can refer to any of the suprasegmental properties of speech that result from how your vocal apparatus is configured.
 - Example: nasality
- Usually, though, we use “voice quality” to refer specifically to the properties of speech affected by stuff inside your larynx.

What's going on in the larynx? (NSFW)

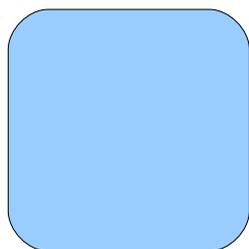
- Vocal folds (commonly called “vocal cords”):



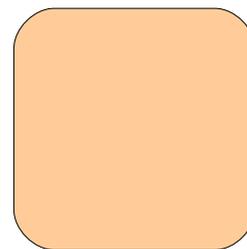
Voice Quality is complicated

- It's hard for us to talk about voice quality:
 - There are many complex things you can do with your vocal folds.
- We often mistake voice quality for pitch:
 - Pitch is easier to talk about, since it's just a scale.
 - Some voice quality features make things sound higher or lower to us, even when they're not.

How do listeners make use of voice quality information?



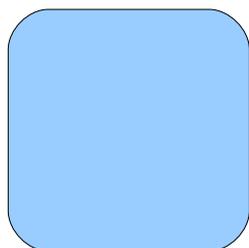
Ma #1



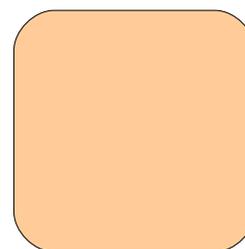
Ma #2

- Are these speakers male or female?

How do listeners make use of voice quality information?



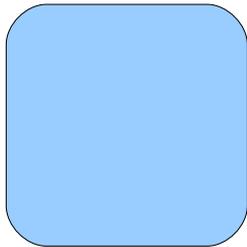
Ma #1



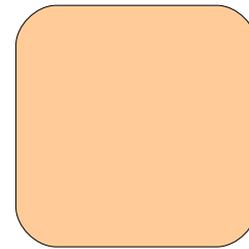
Ma #2

- Which speaker is younger?

How do listeners make use of voice quality information?



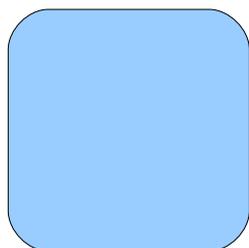
Ma #1



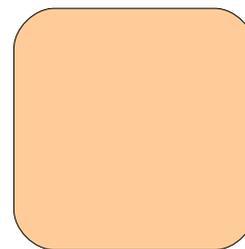
Ma #2

- Which one speaks with a higher voice?

How do listeners make use of voice quality information?



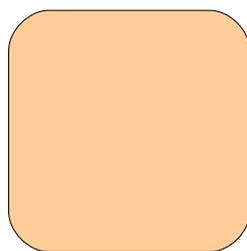
Ma #1



Ma #2

- Both these speakers are speaking the same word in Cantonese. Cantonese has high-pitched tones and low-pitched tones.
- Are they saying a high tone or a low tone?

How do listeners make use of voice quality information?



Keung

- What about this speaker, is she saying a high tone or a low tone?

Well, that was impressive

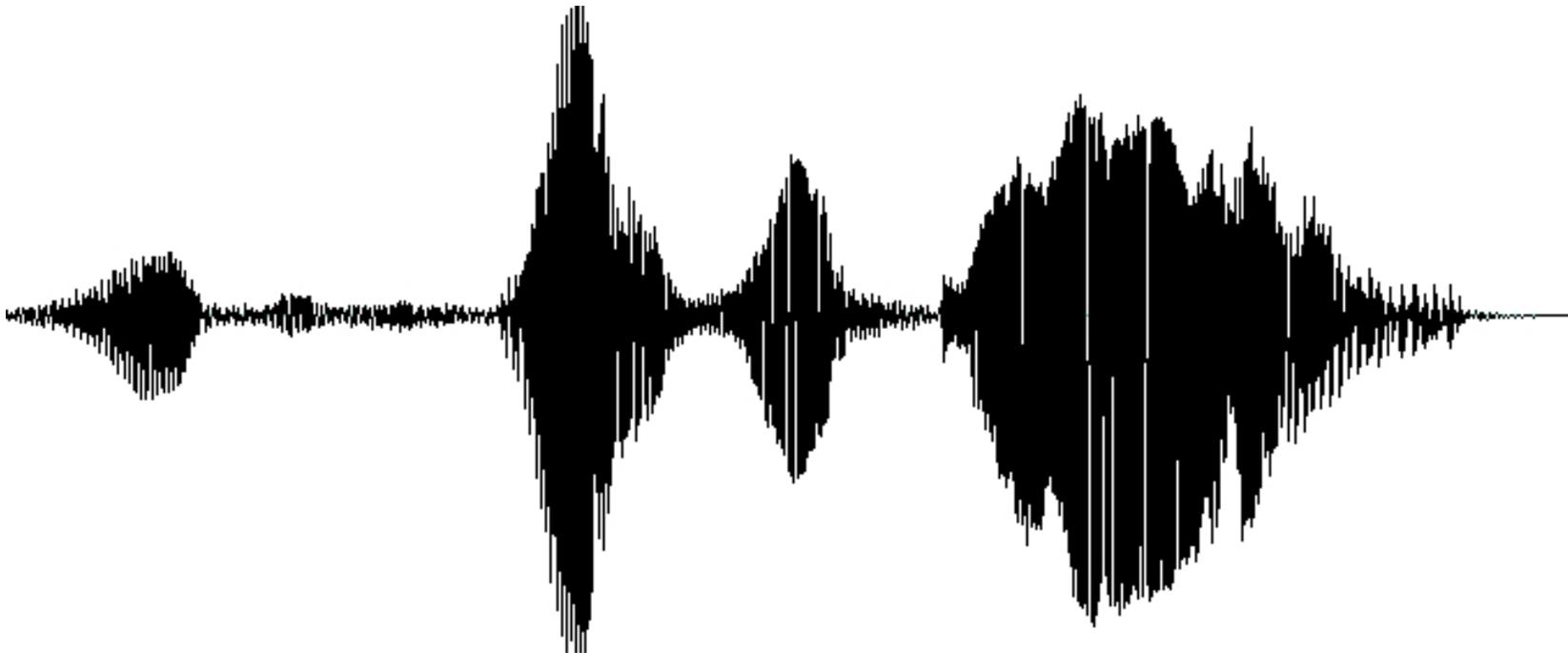
- Even though both “ma” speakers are producing the same exact absolute pitch, most listeners are able to figure out:
 - Which one is older
 - Which one has a higher voice
 - Whether they are saying a high or a low tone

Human listeners are really good at some tasks

- Listeners can reliably locate a pitch within a speaker's pitch range, without actually hearing any other speech from that speaker (Honorof & Whalen 2005).
- How on earth is that possible?

How are we able to do this?

- The sound of someone's voice reaches us after traveling through the speaker's vocal tract.
- Therefore, the soundwave has certain characteristics, depending on the size and shape of the vocal tract.



Yay, Voice Quality

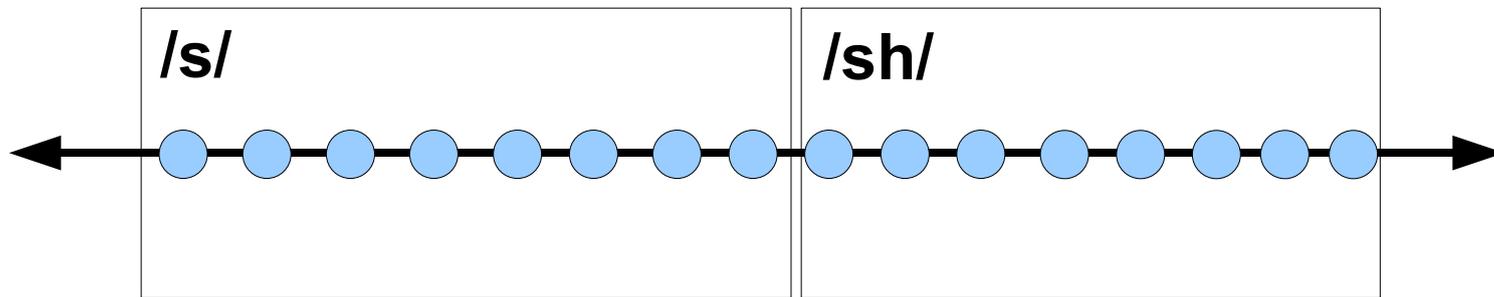
- These acoustic characteristics are part of voice quality.
- Voice quality can also trick us into being bad at other tasks:
 - We are not great at identifying absolute pitch.
 - Two speakers producing the same absolute pitch can sound like they are producing different pitches, due to voice quality differences.

Liz Strand 1999: Gender Stereotypes in Speech Perception



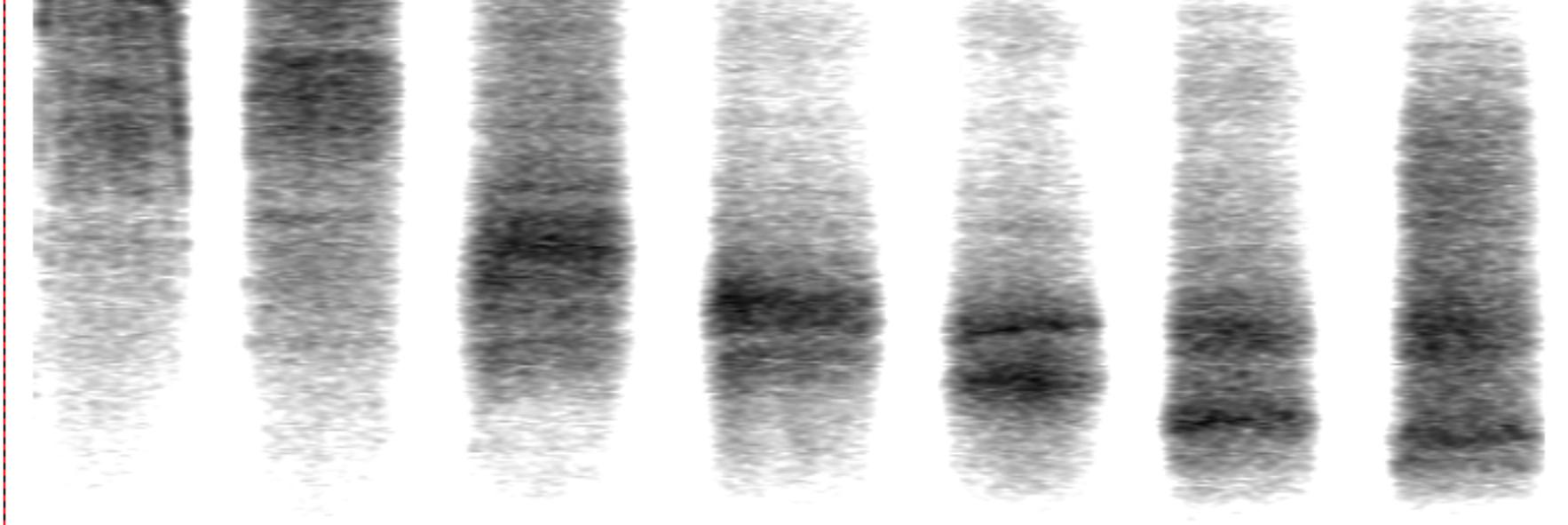
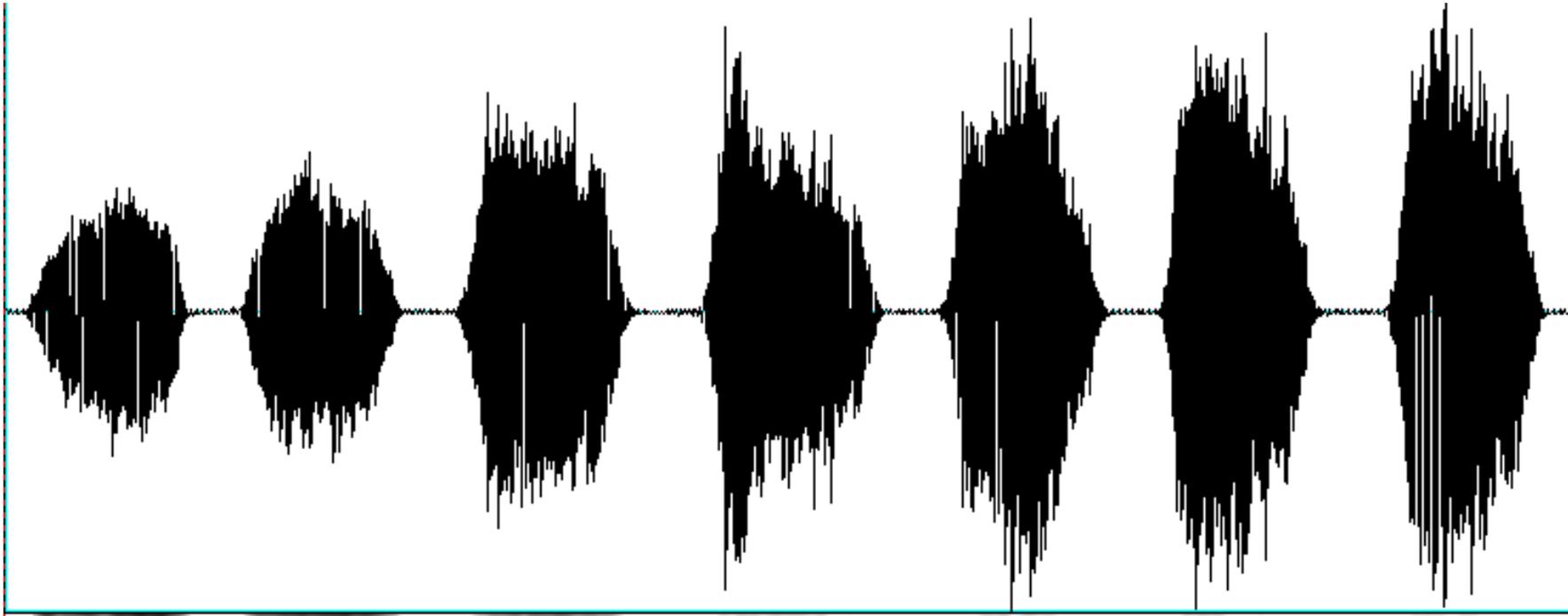
How listeners perceive sounds

- Even though we think of different sounds in a language as being distinct, in fact they are categories imposed on a continuum of sounds.



- At some point along the continuum, we draw the line between what is categorized as one sound, and what is categorized as another.

/s/ vs. /sh/



Recognizing sounds between two speakers

- Different speakers produce sounds slightly differently, depending on the size of their vocal tract, etc.
- This varies particularly by gender.
- Even when two speakers produce the same segment, like /s/, quite differently, we are able to interpret it as the same.
- We “normalize” our perception between speakers.

How do we normalize?

- Based on acoustic information present in voice quality, which gives us clues as to the size of the speaker's vocal tract, etc.
- Also, other information (e.g., visual cues).
- For example: if you believe a speaker has a large vocal tract, you will assume that the frequency of their /s/ will be lower than for a speaker with a small vocal tract.

Strand's studies

- Focus is on where listeners draw the line between /s/ and /sh/, and how that is affected by visual and audio gender information.

Gender as gradient

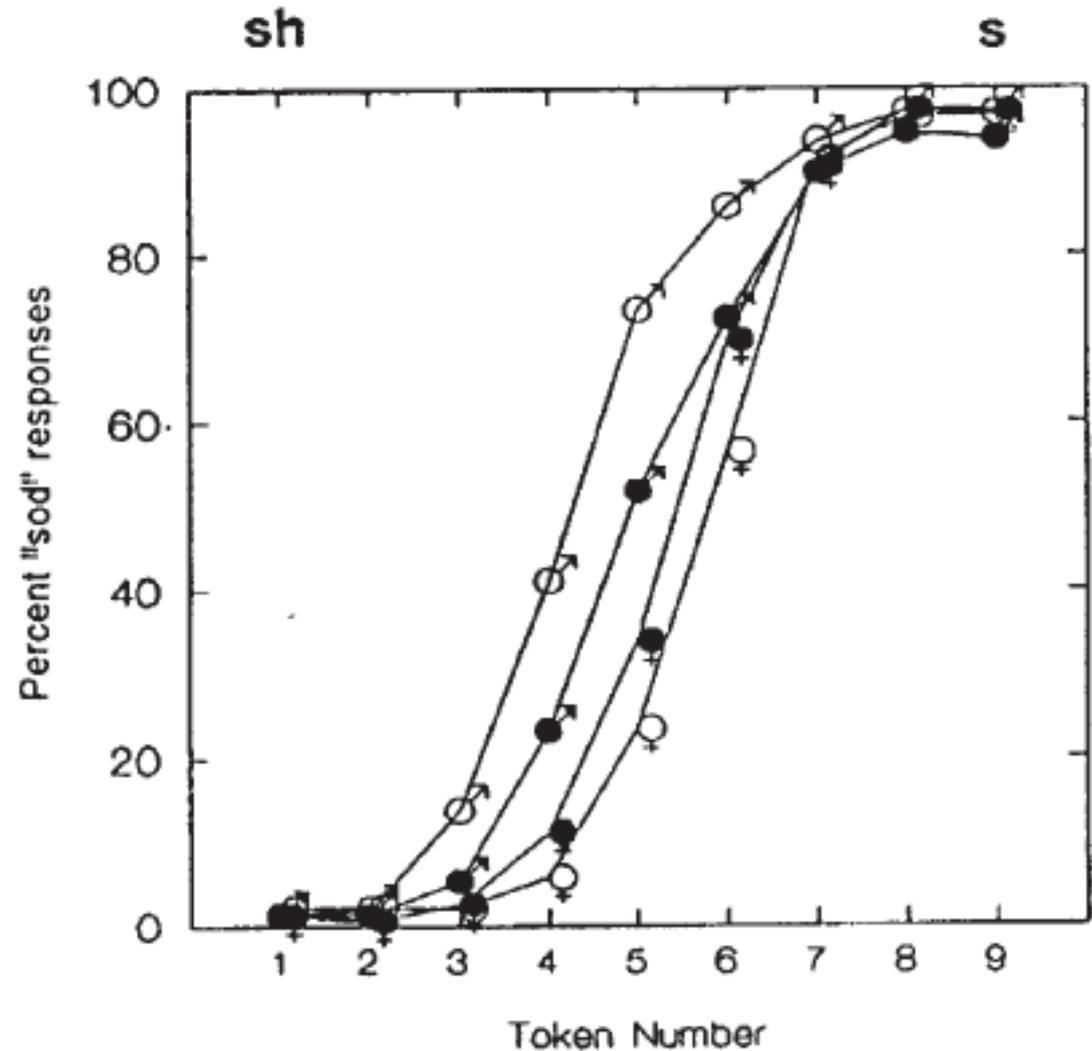
- Previous studies found that speakers draw the line between /s/ and /sh/ differently for women than for men (May 1976).
- But Strand goes further, looking at gender as more gradient:
 - some voices sound more prototypically “male,” some more prototypically “female.”

Strand Study #1

- Four voices: prototypical male, non-prototypical male, prototypical female, non-prototypical female.
 - None of the voice are so weird that people confuse the sex of the speaker.
- Strand synthesized a 9-step continuum of sounds that go from “shod” to “sod”, with a bunch of steps in between.

Strand Study #1

- Listeners presented with examples in the continuum, asked to identify the word as “shod” or “sod.”



Strand Study #1 results

- Speakers identified tokens spoken by prototypical male voices as transitioning to “sod” earlier than for other voices.
- In other words, the same exact token was perceived as “sod” when spoken by the prototypical male voice, and as “shod” when spoken by other voices.
- The four voices each patterned differently, as predicted.

Strand Study #2: The Face Gender Effect

- Audio tracks from before now paired with videos of male and female faces:



Strand Study #2: Results

- The gender of the face affects perception.
- Female faces shift the boundary between /sh/ and /s/ up in frequency, male faces shift it down.
- Consistent with the direction we expected.
- Conclusion: listeners are able to integrate visual and audio information when they perceive speech.

The McGurk Effect

- Let's watch a video about the McGurk effect!

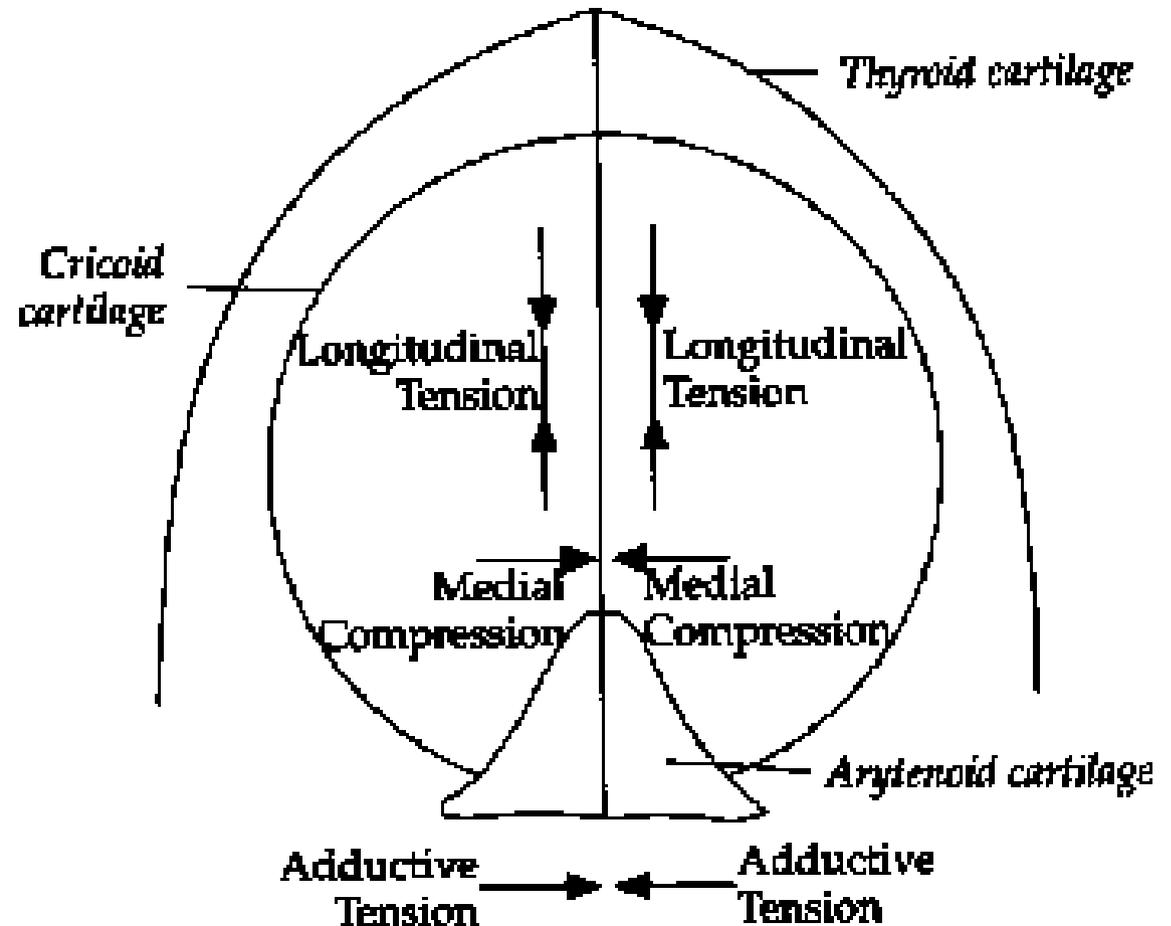


The McGurk Effect

- The video shows a guy saying “ga”
- The audio is of a guy saying “ba”
- Result: most people hear “da,” which is phonetically kinda in between “ga” and “ba.”
- The effect doesn't work on everyone:
 - If it doesn't work for you, consult your physician.
 - No, you'll probably be fine. Probably.

Voice Quality: Phonation

- The vocal folds are complex: there are a number of things you can do with them.



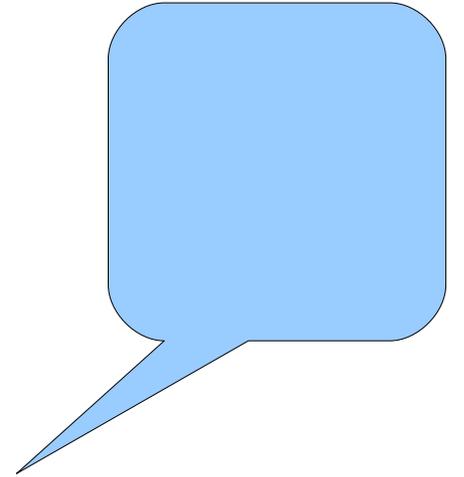
Phonation scale

- Phonation refers to how air comes through the vocal folds.
- Three of the most common phonation types are often presented as a phonation scale:

creaky voice ---- modal voice ---- breathy voice

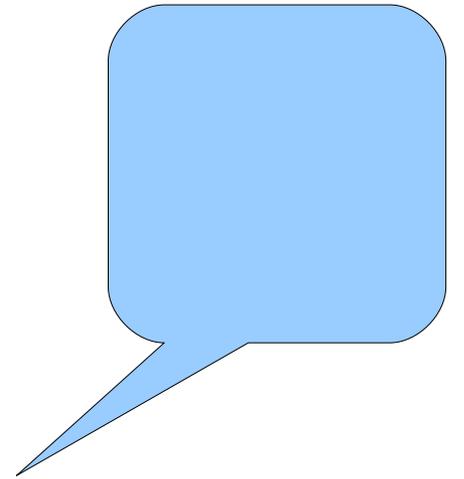
Creaky voice

- Vocal folds are pressed tightly together
- Not a lot of tension lengthwise
- The vocal folds get bunched up
- Vibration is slow and irregular
- Associated with lower pitch



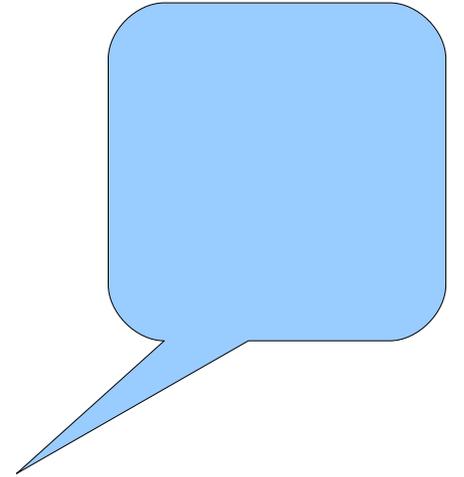
Modal voice

- This is the “normal” way of talking
- Medium amount of tension in all parts of the vocal folds



Breathy voice

- Moderate tension lengthwise
 - Low tension pushing folds together
 - Results in frication as a lot of air escapes through the opening
-
- Now YOU try it!



How do we measure phonation?

- Articulatory methods:
 - Attach devices onto parts of speakers' bodies, or scan them using fancy medical scanners
 - Measures what they are doing with different parts of their vocal apparatus
- Acoustic methods:
 - Analyze and measure recordings with computer software.
- Perceptual methods:
 - Categorize speech through our own perceptual intuitions.

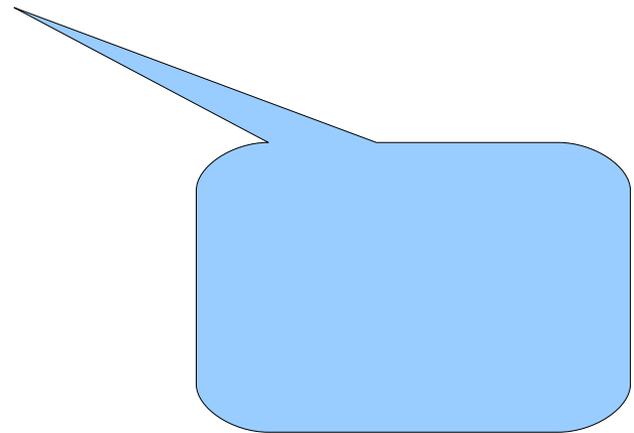
What do languages use phonation for?

- Some languages (e.g., Gujarati) use phonation types as part of their sound system (Keating & Esposito 2007)
 - For example, sounds produced with creaky voice would mean something different from sounds produced with modal voice.
- Most languages don't have phonation type as part of their sound system.
- But we can all use phonation for stylistic purposes.

Rob Podesva 2007:
Phonation type as a stylistic
variable: the use of falsetto in
constructing a persona

Falsetto

- Falsetto is another phonation type.
- Vocal folds are strongly stretched lengthwise, causing them to become thin and vibrate at a higher frequency.
- Correlates with high pitch (high f_0) due to the way it's produced.



Heath

- Heath is a gay med student.
- Podesva looks at Heath's speech in various contexts:
 - bbq with friends
 - phone call with family
 - meeting with a patient

Heath's use of falsetto

- Uses falsetto most frequently at bbq with friends
- Duration of his falsetto longer at bbq
- f_0 range wider, meaning he varies up and down more in pitch.
- Heath also uses creaky voice, possibly to widen his pitch range.

What is the significance of falsetto?

- Podesva: falsetto carries a core meaning of “expressiveness.”
- Functions:
 - yelling
 - expressing surprise or excitement
 - offering evaluative commentary
 - enlivening a direct quotation
 - engaging audience when telling narrative
- Heath uses falsetto to construct a diva persona.

Let's check out some falsetto

- Video of Ross the Intern (from The Tonight Show)

Where does Ross use falsetto?

- Examples:

- What does it do? [yelling] 

- She has *huge* lips! [evaluative commentary] 

Voice Quality in Cartoons



Why are cartoon voices interesting?

- Voices, sounds and images are often exaggerated in cartoons, giving us the essences of characters and contexts.
- Because cartoons exaggerate voice quality, they provide us with an interesting opportunity to examine the social significance of voice quality features.

What's the deal with Russian Sherlock Holmes?

- The Russian image of Sherlock Holmes was primarily formed by a very popular Russian live-action TV show in the late 70's / early 80's.
- Holmes played by Vasiliy Livanov:



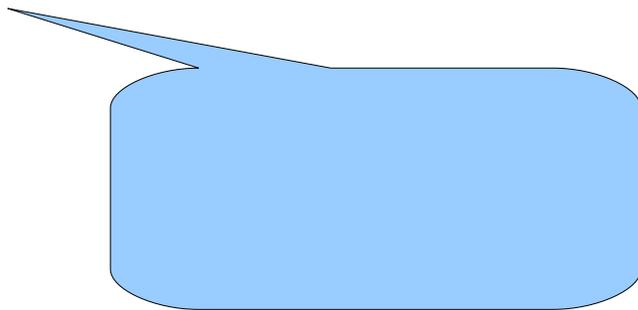
From live-action to cartoon

- Takes some of Livanov's voice quality features and exaggerates them.
- Personality features that may be associated with this voice:
 - eccentric
 - antisocial
 - authoritative
 - serious
 - smoker



Japanese sweet voice: voice of the perfect woman (Starr 2006)

- Sweet voice is a popular professional voice-acting style in Japan.
- Appears in voice-overs for commercials, train station announcements, cartoons.



Characteristics of Sweet Voice

- Acoustic characteristics:
 - dramatic swings from modal to breathy
 - relatively low pitch
 - produced with “head voice” phonation
- Linguistic correlates:
 - use of Japanese Women's Language features

Characteristics of Sweet Voice

- Social correlates:
 - motherly
 - kind
 - mature
 - passive
 - conservative
 - traditionally beautiful
- Sweet voice characters tend to be supporting characters, not heroines.



Sweet Voice is not cute

- In contrast to sweet voice characters, cute characters are relatively:
 - young
 - non-traditional
 - not as beautiful (but cute!)
 - energetic
 - assertive
 - high-pitched
 - can be main characters



What does sweet voice tell us?

- There are multiple ways of being feminine in Japanese popular culture.
- There are strong perceived links between voice quality, language use, and personality characteristics.
- The notion of the perfect woman who is a devoted wife and mother, which has old roots in Japan's history, is still alive and well.

But how does voice quality affect ME?

- Creaky voice:
 - Young people today use a LOT of creak.
 - Particularly young women.



Creak

- Clip: Molly McAleer
 - Where does Molly use the most creak?
 - What do you think creak means? What social message is she trying to send with it?
 - Have you noticed students at Stanford using a lot of creak?