

# Defining and Measuring Voice Quality

Jody Kreiman\*, Diana Vanlancker-  
Sidtis\*\*, and Bruce R. Gerratt\*

\*UCLA

and

\*\*New York University



# Voice Quality is Hard to Measure

- Complex
- Unstable



# The Definitional Dilemma

- Voice is hard to define
- Voice quality is also hard to define

# The ANSI Definition

“...that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar.”

# The ANSI Definition

- Often maligned
  - A negative definition
  - Defines quality in the context of one specific task; hard to operationalize or generalize to other tasks
  - Implies quality is independent of frequency and amplitude

# Virtues of the ANSI Definition

- Treats sound quality as the result of a perceptual process
- Highlights importance of signals and listeners in determining quality

# Why Include the Listener?

- Just as loudness and pitch do not exist without the listener, vocal quality is an acoustic-PERCEPTUAL phenomenon.

# How Listeners Introduce Variability

- Listeners may pay attention to different acoustic aspects of signals, even in the same task
- Importance of a given cue may depend on context or task demands
- Different listeners may use different cues
- Definitions of quality that focus on production or acoustics cannot account for such effects



# So: How Should We Measure Quality?

- Create lists of terms to describe listeners' auditory impressions

# Venerable and Modern Labels for Voice Quality

## Julius Pollux

Clear

Deep

Brilliant

Small, feeble,  
faint

Thin

Hollow, indistinct

## Moore, 1964

Clear, light, white

Deep

Bright, brilliant

Breathy, whispery

Thin, pinched,  
shallow,

Hollow, covered

## Gelfer, 1988

Clear

Resonant, low

Bright, vibrant

Breathy, soft,  
babyish, weak

Thin

Muffled

# More Labels for Voice Quality

## Julius Pollux

Brassy

Harsh

Shrill, sharp

Smooth

Dull

## Moore, 1964

Buzzy, clangy,  
metallic

Harsh, strident,

Shrill, sharp,  
piercing, cutting,  
pointed

Smooth, velvety

Dull, heavy, dead

## Gelfer, 1988

Metallic

Harsh, gravelly

Shrill, sharp

Smooth

Dull, heavy, thick

# Well-known Problems with Rating Scale Approaches

- Atheoretical approach
- Which scales to include?
- Poor reliability and questionable validity
- Redundancies and ambiguities
  - MDS and factor analytic studies have not resolved this problem

# Vagaries of Scale Definition

Breathiness = dry, hard, excited, pointed, cold, choked, rough, cloudy, sharp, poor, bad? (Isshiki et al.)

Or:

Breathiness = breathy, wheezing, lack of timbre, moments of aphonia, husky, not creaky? (Hammarberg et al.)

# What to Do?

- Voice Profile Analysis
  - Consistent from phonetic theory
  - Specifies how scales are related to each other
  - Specifies where information about quality *might* be, but does not model listeners' behavior

# What to Do?

- Acoustic assessment protocols
  - e.g., Dysphonia severity index, Hoarseness diagram
  - Depend on inconsistent correlations with perceptual measures for validity as measures of quality

# What to Do?

- Method-of-adjustment task using speech synthesis
  - Does not depend on selection/definition of labels for quality dimensions
  - Helps listeners focus attention and avoids reliance on internal standards
  - Demonstrates causation between acoustic attributes and perceived quality
  - Follows directly from ANSI definition of quality



# Strengths/Limitations

- Reliability
- Directly links perception to acoustics
- Technically difficult at present

# Conclusion

- When we cannot measure, our knowledge is meager and unsatisfactory.
  - Attributed to Lord Kelvin
- If it exists, it exists in amounts, and if it exists, it can be measured.
  - Lord Thorndyke

# Acknowledgment

This research was supported by grant DC01797 from the NIH/NIDCD.