The relationship between sonority and glottal vibration

Megan L. Risdal, Ann Aly, Adam J. Chong,
Pat Keating, Jesse Zymet

CUNY Phonology Forum — 2016 Sonority

The UCLA Phonetics Lab

Introduction

- Sonority and the sonority hierarchy have been an important basis underlying explanations for various phonological phenomena (e.g., assimilation, phonotactics of syllable structure).

Searching for a Physical Basis of Sonority

- How it manifests phonetically and how to quantify it is currently unresolved (Eduard, 1881; Heffner, 1950; Fletcher, 1972; Parker, 2002).
- Some have suggested that sonority may be related to:
  - degree of constriction (e.g., Hankamer and Aissen, 1974)
- While others have related sonority to:
  - voicing (e.g., Nathan, 1989)

Two Scales of Sonority

- Miller (2012) uses both of these parameters:
  - **Source scale**: No source < turbulence only < breathy voice (periodic source + glottal source) < modal voice (periodic source)
  - **Aperture scale**: explosive stops < voiced implosives < fricatives < nasals < liquids and glides < vowels
- The two parameters are considered to be completely separate.
- The two scales are combined additively to determine a segment's sonority.

Phonetic Correlates of Sonority

- Parker (2002) also took into account both voicing and degree of constriction in several of the phonetic correlates he measured (e.g. intensity, peak intraoral air pressure and peak total air flow) which depend *simultaneously* on both source and filter.
  - Other measures, F1 and duration, depend only on the filter.
- He reports that **intensity** exhibits the strongest (positive) correlation with the sonority hierarchy and duration exhibits the least (see also Gordon, this conference).

Glottal Source-Vocal Tract Interactions

- Voicing requires airflow through the glottis.
- With no oral constriction → air flows freely and voicing is easy.
- With a complete oral constriction → glottal airflow stops and voicing stops (Ohala, 1983; Westbury and Keating, 1986).
- In between these two extremes, voicing should vary dep. on the degree of the oral constriction.
- It has also been suggested that the way the vocal folds vibrate probably varies, too.

Previous evidence: effects of oral constriction on voicing

- Voicing amplitude varies (Solé, 2015).
- Glottal excitation (rate of glottal closing) varies (Fant, 1997; Stevens, 2000; Bickley and Stevens, 1987).
- Degree of glottal constriction (amount of glottal flow) varies (Fant, 1997; Stevens, 2000; Bickley and Stevens, 1987).
- Suggests that tighter **oral** constriction → **breathier voicing**.

Mittal et al. (2014)

- Six geminate voiced consonants: [r, z, ɣ, l, n, and ŋ].
- Measured **strength of glottal excitation** $\Delta\psi$:
  An acoustic measure of the relative amplitude of impulse-like excitation derived from the instant of significant excitation of the vocal-tract system during production of speech (Murty and Yegnanarayana, 2008).
- Result: [z] $<$ [r] $<$ [n,ŋ] $<$ [ɣ] $<$ [l].

Our Study

Our Study: Questions

- Replicate and extend Mittal et al. (2014) with more consonants and adding vowels, focusing more explicitly on:

1. Does voicing vary across segment types, and if so;
2. Do these distinctions depend on degree and type of constriction?
3. Do these differences among voiced segment types correlate with standard notions of sonority?

Segments (21 total)

- Including 5 of 6 segments from (Mittal et al., 2014). We eliminated [ŋ] because no difference between it and [n] was reported.

Consonants

- Approximants: [j, w, l, ɹ]
- Trill & Tap: [r, ɾ]
- Nasal: [n]
- Fricatives: [ð̞, z, ɣ, ʁ]
- Affricates & Stop: [ʤ, gɣ, d]

Vowels

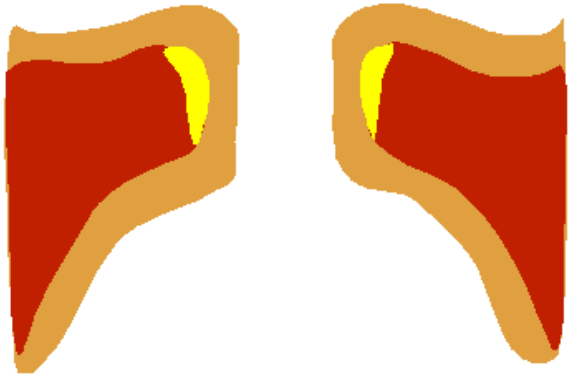
- Front unround: [i, e, a]
- Front round: [y, ø]
- Back round: [o, u]

Our Dependent Measures

- Acoustic signal: **Strength of glottal Excitation (SoE)**, the same dependent measure, $\Delta\psi$, reported in Mittal et al. (2014).
- Electroglottography (EGG) signal: the **Contact Quotient ($\mathbf{CQ}_H$)** as a measure associated with amount of vocal fold contact during phonation.

Electroglottography (EGG) & Contact Quotient

- Contact Quotient ($\mathbf{CQ}_H$): the proportion of a complete vibratory cycle from an inferred point of closure to an inferred point of opening.



Participants & Procedure

- 13 participants (6 males and 7 females) contributed data for the present study.
- All speakers were linguists trained in phonetics and fluent speakers of English (8 native speakers of North American English).
- Participants produced consonants (3 reps [aCa]) and vowels (3 reps [wV]) for a total possible 63 tokens per speaker or 819 overall.

Segmentation & Measurement

- Tokens were segmented using Praat TextGrids from onset through voiced duration.
  - Only fully-voiced portions of segments were analyzed.
  - The first voiced closure of the trill with at least 3 pulses was segmented.
  - Affricates were treated as a voiced closure (stop) and a voiced release (fricative) separately.
  - Stop releases were not included.
- Using VoiceSauce for SoE (Shue et al., 2011) and EggWorks for CQ_H (Tehrani, 2012) software.
- Missing values and outliers were excluded leaving 744 tokens for analysis. We report on speaker normalized values of CQ_H and SoE.

Results & Analysis

Missing Tokens: Evidence for Voicing Difficulty

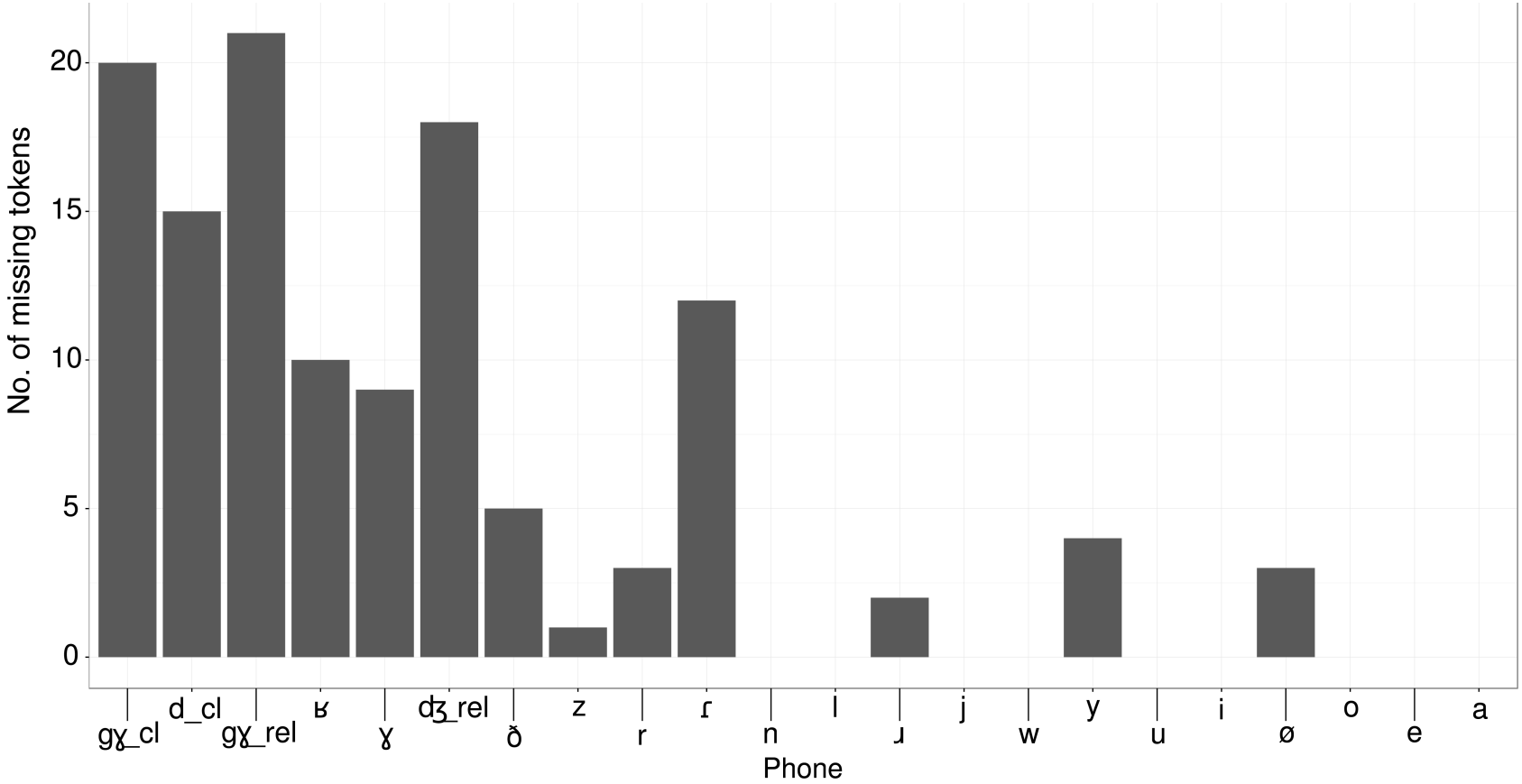


Figure: Segments eliminated due to lack of voicing. L $\longrightarrow$ R = decreasing constriction degree.

Analyses

- First we examine overall trends between CQ_H & SoE and voiced segments and discuss our findings as they relate to sonority.
- Following this, we take a closer look at several within-category distinctions.
- **Null hypothesis**: no difference in CQ_H or SoE across voiced segments, with the exception of voiced fricatives.

CQ_H: Consonants

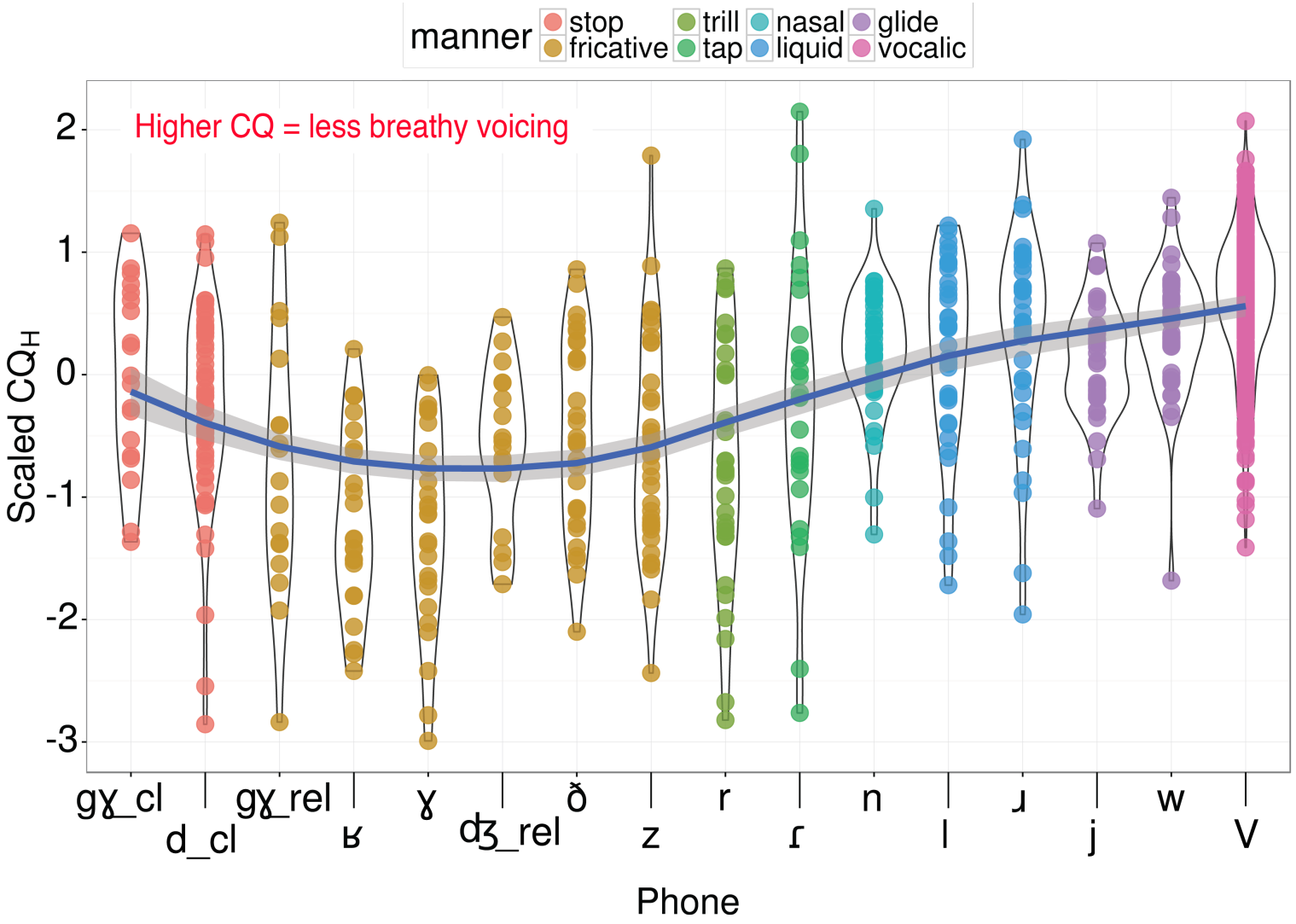


Figure: CQ_H across all consonants, collapsing vowels into a single category. L $\longrightarrow$ R = decreasing constriction degree.

Strength of Excitation: Consonants

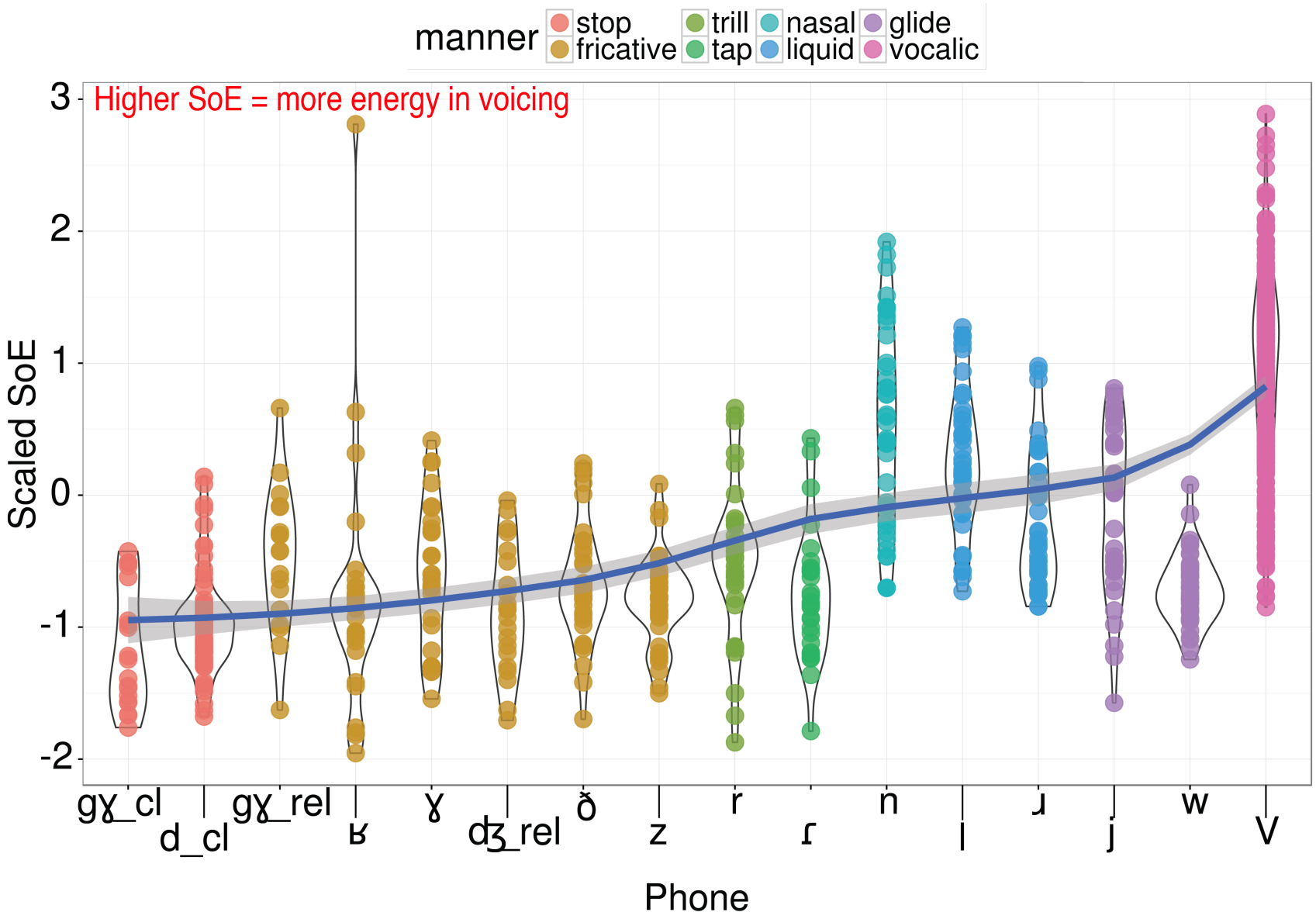


Figure: SoE across all consonants, collapsing vowels into a single category. L $\longrightarrow$ R = decreasing constriction degree

The relationship between CQ_H and SoE

- CQ_H makes distinctions among obstruents and SoE among sonorants.

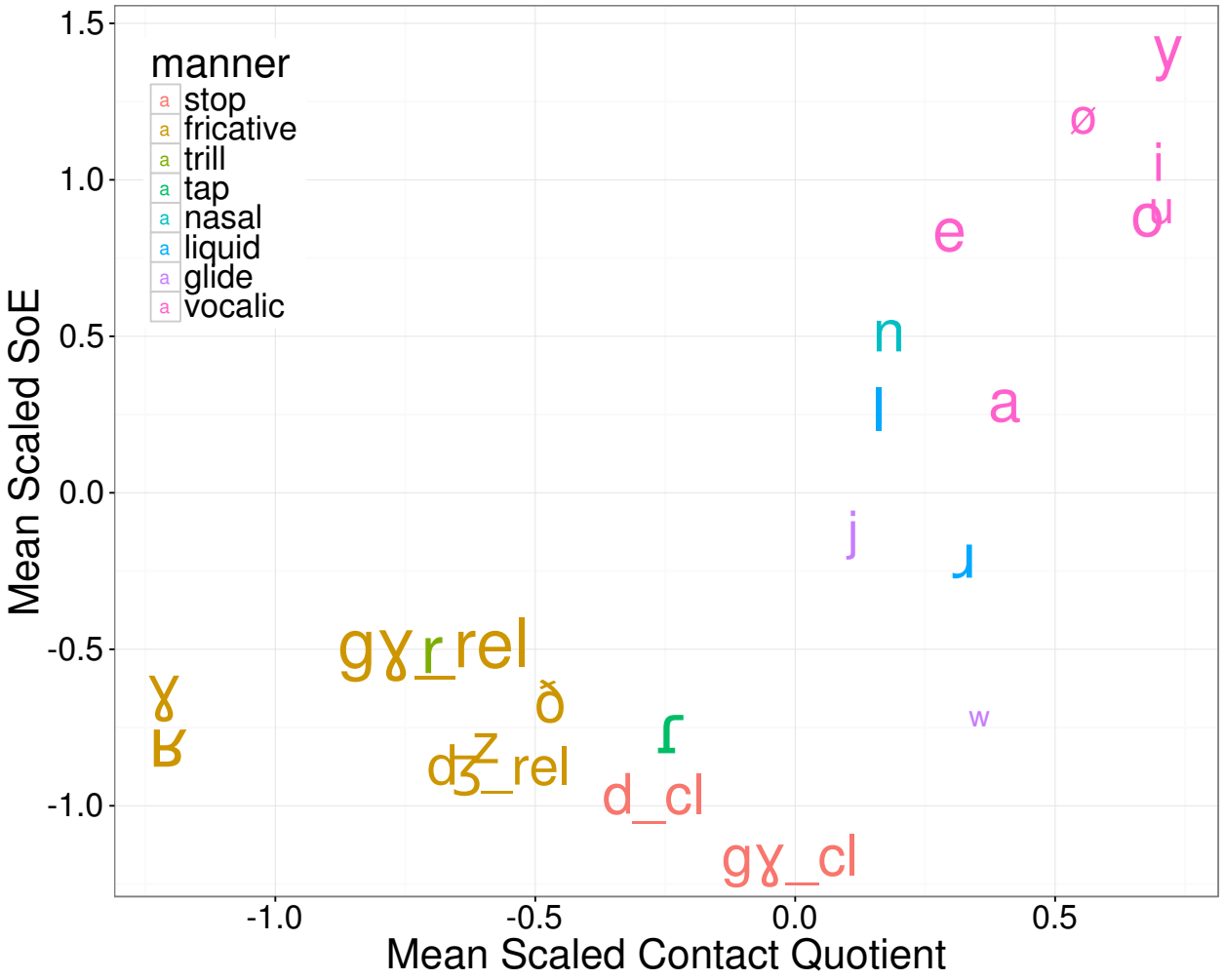


Figure: Scaled CQ_H by Scaled SoE by segment. Size of symbol indicates standard deviation.

Conditional Inference Tree

- Do CQ_H values discriminate among voiced sounds along the sonority hierarchy?
- Conditional inference trees work by recursively partitioning observations into subsets of the data which are significantly different with respect to an outcome variable (Hothorn et al., 2006).

CQ_H: All Segments

- V > liquids/glides + [gɣ] (cl) > non-dorsal fricatives/tap/trill + [d] (cl) > dorsal fricatives

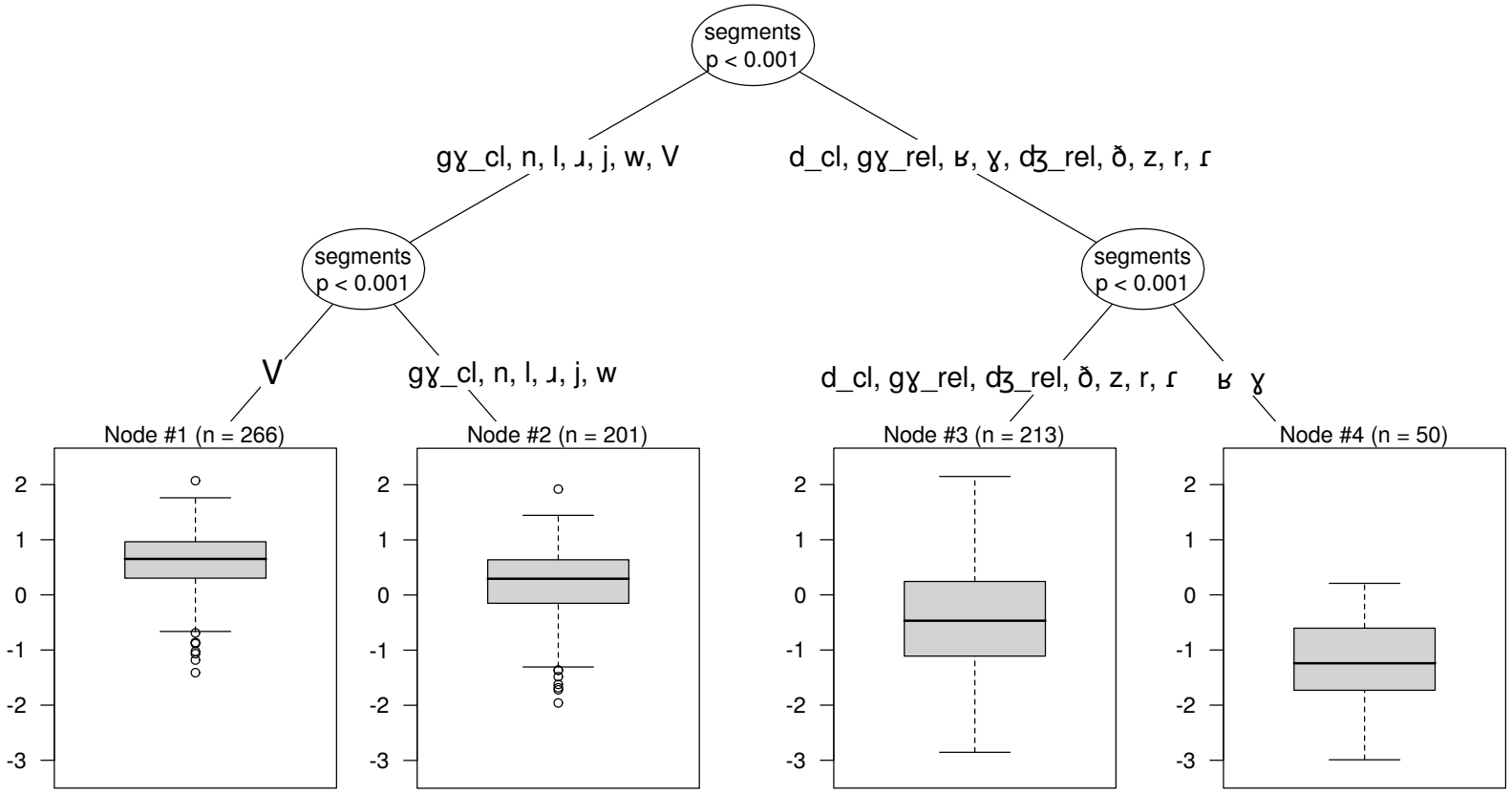


Figure: CQ_H across all segments.

Interim Summary: The relationship between CQ_H and SoE

- A range of values of both measures is observed across segments of different constriction degrees.
- In broad terms, the trends reflect differences along the sonority hierarchy.
  - Vowels, on the whole, have the highest CQ_H and SoE values, indicating strongest voicing energy and least breathy voicing.
  - Compared to nasals, liquids and glides, voiced obstruents also have lower CQ_H and SoE values - less energy in voicing and breathier voicing.

Analyses: Within Segment Types

- Cross-linguistically, the sonority hierarchy is stable at each level; however, there is some variation within levels (Jany et al., 2007).
  - In particular, liquids and rhotics differed across languages (e.g., Egyptian Arabic, Hindi, Mongolian, Malayalam).
  - Other contrasts may pattern differently cross-linguistically, for example sibilants versus other fricatives.
- This motivates us to examine some within-category differences.
- Linear mixed effects models[1] in `R` predicting CQ_H/SoE values.

[1] Speaker as random intercept. P-values obtained using likelihood ratio tests.

Trill, tap, and [d]

- We were interested in examining segments that involved full closures, but are nonetheless in different parts of the sonority hierarchy.
- The trill has a significantly **lower $\mathbf{CQ}_H$** value (**breathier**) compared to [d] and [r] which aren't significantly different.
- That is: **[r] $<$ [d] = [r] on $\mathbf{CQ}_H$ values.**
- Trills pattern like voiced fricatives which is not surprising as their contact intervals are often more like voiced fricatives than like stops.

Place (dorsal/coronal) versus Sibilance

- Mittal et al. (2014) attributed a difference between **dorsal, non-sibilant** [ɣ] and **coronal, sibilant** [z] in SoE to a place of articulation effect rather than an effect of sibilance.
- We compared the dorsal, non-sibilant fricatives [ʁ] and [ɣ] ($n = 59$) to the coronal, non-sibilant [ð̞] ($n = 34$) and the coronal, sibilant [z] ($n = 38$) to tease this apart.
- Place: Dorsal fricatives (non-sibilants) have **lower $\mathbf{CQ}_H$** values than coronal [ð̞] (non-sibilant) and [z] (sibilant) $\rightarrow$ **breathier voicing**.
- Sibilance: **no significant effect** on model fit.

Strength of Excitation: Vowels

- We tested the relationship between voicing across vowels and intrinsic aperture using **scaled SoE** and **scaled F1** as an indicator of vowel height/aperture.

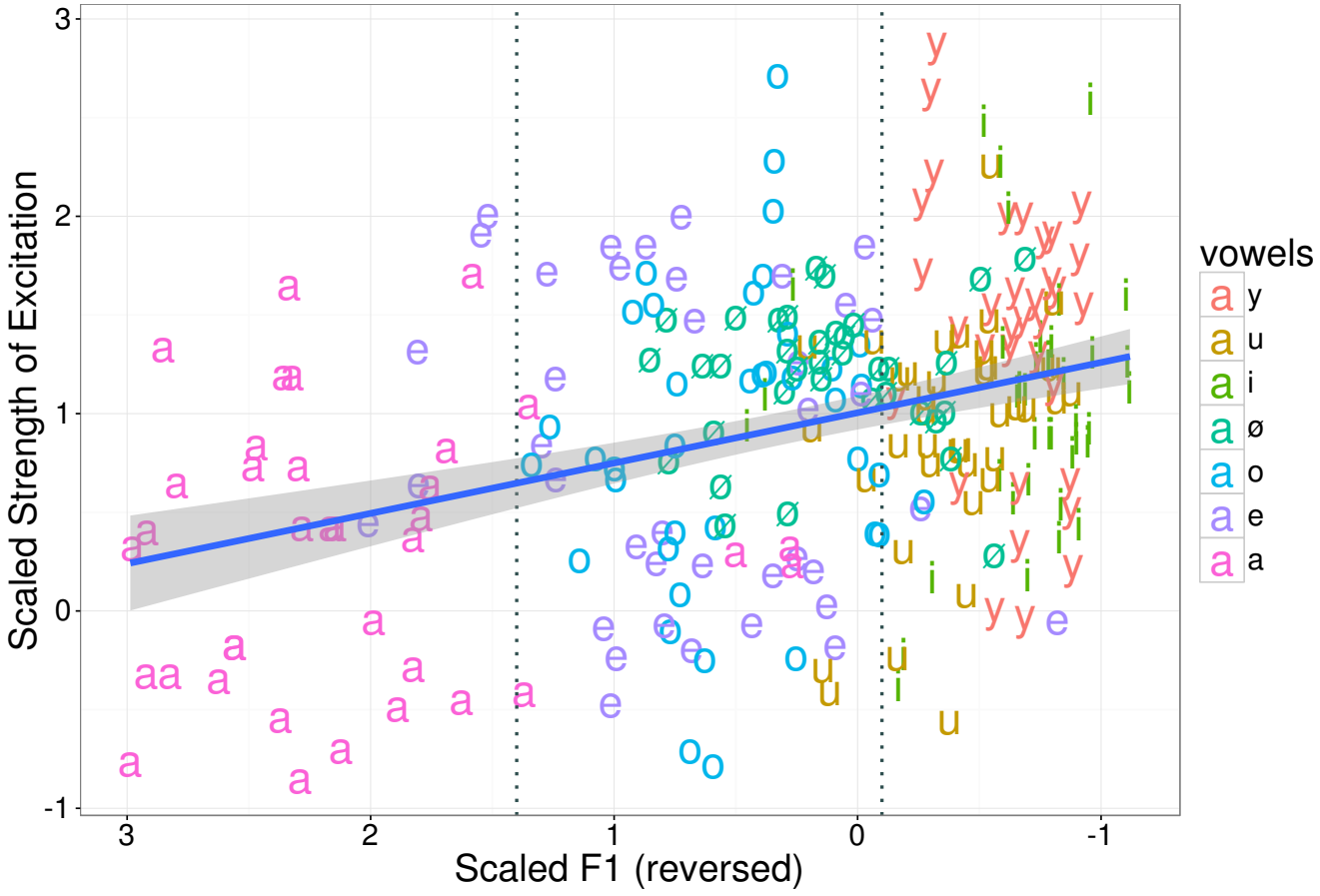


Figure: Strength of Excitation versus scaled F1 (reversed).

Vowels: [i, e, a, y, ø, o, u]

What were the effects of the following factors on **SoE**?

- **Height** (F1): High vowels > Mid vowels > Low vowels.
- **Backness** (F2): No significant effect.
- **Roundness**: Unrounded vowels have **lower** SoE values (less strong voicing) than round vowels.

- These effects are in the **opposite direction** from the prediction based on the idea that rounding/height decrease vowel aperture.

Summary: Within Categories

- Compared to tap [ɾ] and stop [d], the trill [r] has the lowest CQ_H values meaning it patterns more like a fricative.
- Place vs. Sibilance:
  - Found a significant difference in voicing dep. on place of articulation but not sibilance.
- Vowels: The expected relationship between vowel height (and roundness) and SoE is not borne out; however —
  - The reversal in expected SoE values among vowels is also observed among sonorants (nasal > liquids > approx.).

Conclusions

1. Does voicing vary across segment types, and if so;
   - **Yes! Null hypothesis can be rejected.**
2. Do these distinctions depend on degree and type of constriction?
   - **Generally speaking, the tighter the constriction $\rightarrow$ the breathier the voicing**
3. Do these differences reflect distinctions made by (phonological) sonority?
   - **Only at the broadest level: Vowels $>$ approximants $>$ obstruents**
   - **A number of reversals within each broad manner class.**

Acknowledgments

Thank you to Yen Shue for VoiceSauce and Henry Tehrani for EggWorks.

We would also like to thank the students in the UCLA Winter 2015 Speech Production graduate course for participating in the design and execution of this study, as well as Soo Jin Park for adding the SoE measure to VoiceSauce.

Thank you also to members of the UCLA Phonetics Lab for helpful comments and discussion.

Bickley, C. A. and Stevens, K. (1987). Effects of a vocal tract constriction on the glottal source: Data from voiced consonants. In Baer, T., Sasaki, C., and Harris, K., editors, *Laryngeal Function in Phonation and Respiration*, pages 239–253. College-Hill Press, Boston, MA.

Eduard, S. (1881). *Grundzüge der Phonetik*. Breitkopf und Hartel, Leipzig.

Fant, G. (1997). The voice source in connected speech. *Speech Communication*, 22:125–139.

Fletcher, M. (1972). *Speech and Hearing in Communication*. Krieger, Robert E., Huntington.

Hankamer, J. and Aissen, J. (1974). The sonority hierarchy. pages 131–145, Chicago: Chicago Linguistic Society.

Heffner, R.-M. S. (1950). *General Phonetics*. The University of Wisconsin Press, Madison.

Hothorn, T., Hornik, K., and Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics*, (3):651–674.

Jany, C., Gordon, M., Nash, C. M., and Takara, N. (2007). How universal is the sonority hierarchy?: A cross-linguistic study. Paper presented at ICPhS XVI.

Miller, B. (2012). Sonority and the larynx. In Parker, S., editor, *The Sonority Controversy*, pages 257–288. Walter de Gruyter.

Mittal, V. K., Yegnanarayana, B., and Bhaskararao, P. (2014). Study of the effects of vocal tract constriction on glottal vibration. *Journal of the Acoustical Society of America*, 136(4):1932–1941.

Murty, K. S. R. and Yegnanarayana, B. (2008). Epoch extraction from speech signals. *IEEE Transactions on Audio, Speech, and Language Processing*, 16(8):1602–1613.

Nathan, G. S. (1989). Preliminaries to a theory of phonological substance: The substance of sonority. In Corrigan, R., Noonan, M., and Eckman, F., editors, *Linguistics categorization*, volume 61 of *Current Issues in Linguistic Theory*, pages 55–68. John Benjamins, Amsterdam.

Ohala, J. J. (1983). *The Origin of Sound Patterns in Vocal Tract Constraints*.

Parker, S. (2002). *Quantifying the Sonority Hierarchy*. PhD thesis, University of Massachusetts Amherst.

Shue, Y. L., Keating, P. A., Vicenik, C., and Yu, K. (2011). Voicesauce: A program for voice analysis. In *Proceedings ICPhS 17*, pages 1846–1849.

Solé, M.-J. (2015). Acoustic evidence of articulatory adjustments to sustain voicing during voiced stops. In *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK: the University of Glasgow.

Stevens, K. (2000). *Acoustic Phonetics*. Current Studies in Linguistics Series. MIT Press.

Tehrani, H. (2012). Eggworks. `http://www.linguistics.ucla.edu/faciliti/sales/software.htm`.

Westbury, J. and Keating, P. (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics*, 22:145–166.