JASA
EXPRESS LETTERS

https://doi.org/10.1121/10.0001585

Effects of consonantal constrictions on voice quality

Adam J. Chong,[1,a)] Megan Risdal,[2] Ann Aly,[3] Jesse Zymet,[4] and Patricia Keating[5]
[1]*Department of Linguistics, School of Languages, Linguistics and Film, Queen Mary University of London, Mile End Road, E1 4NS, London, United Kingdom*
[2]*Google, 1600 Amphitheatre Parkway, Mountain View, California 94043, USA*
[3]*Agile Six Applications, Inc., 501 West Broadway, Suite 800, San Diego, California 92101, USA*
[4]*Department of Linguistics, University of California, Berkeley, California 94720, USA*
[5]*Department of Linguistics, University of California, Los Angeles, California 90095, USA*
a.chong@qmul.ac.uk, mrisdal@gmail.com, annmalyy@gmail.com, jzymet@gmail.com, keating@humnet.ucla.edu

Abstract: A speech production experiment with electroglottography investigated how voicing is affected by consonants of differing degrees of constriction. Measures of glottal contact [closed quotient (CQ)] and strength of voicing [strength of excitation (SoE)] were used in conditional inference tree analyses. Broadly, the results show that as the degree of constriction increases, both CQ and SoE values decrease, indicating breathier and weaker voicing. Similar changes in voicing quality are observed throughout the course of the production of a given segment. Implications of these results for a greater understanding of source-tract interactions and for the phonological notion of sonority are discussed.

[Editor: Douglas D. O'Shaughnessy] Pages: EL65–EL71

Received: 13 February 2020 **Accepted:** 27 June 2020 **Published Online:** 20 July 2020

1. Introduction

It is well-known that the ease of initiating and sustaining voicing is affected by the size of the supraglottal constriction. This dependency between filter and source is due to the fact that in order to initiate and maintain voicing there must be a decrease in pressure across the larynx (e.g., van den Berg, 1958; Stevens, 1998). Voicing during stops and fricatives is notably challenging to maintain due to these aerodynamic requirements [e.g., Keating (1984), Solé (2010), (2018), and Stevens (1971), (1977)]. Thus, in the extreme, when there is full closure, as in a stop, voicing will eventually cease as oral pressure equalizes to subglottal pressure (Rothenberg, 1968; Westbury and Keating, 1986). Previous studies have thus sought to examine the ways in which speakers overcome these constraints, and have reported articulatory mechanisms (e.g., active enlargement of the oral cavity or nasal venting) that aim to reduce or slow down the build-up of oral pressure and therefore facilitate phonation [e.g., Lisker (1977), Westbury (1983), and Solé (2018)].

The difference in aerodynamic conditions due to different supraglottal constrictions has also been hypothesized to affect the way in which the vocal folds vibrate, with Halle and Stevens (1967) leaving open the possibility that these laryngeal adjustments are under a speaker's active control. Previous work has made use of physical modelling of the vocal tract to examine the rate and volume of glottal flow as a function of supraglottal resistance [e.g., Bickley and Stevens (1987)]. Fant (1997), for example, found, for a set of Swedish sounds, that voicing in voiced consonants was breathier and quieter than the voicing in vowels. Amongst voiced consonants, voiced fricatives have been argued to require spreading of the vocal folds (e.g., higher open quotient) in order to maintain the necessary airflow requirement of turbulent noise generation [e.g., Stevens (1971), Pirello *et al.* (1997), and Solé (2010)], suggesting breathier voicing. Trills have been shown to involve similar aerodynamic requirements as fricatives (Solé, 2002), with some work in the singing and clinical literature showing that trills involve a lower mean vocal fold contact quotient (CQ), suggestive of breathier voicing [e.g., Andrade *et al.* (2014) and Hamdan *et al.* (2012)], as well as a larger CQ range, suggestive of CQ oscillations during the trill.

These previous studies have primarily examined specific segmental classes. An exception is Mittal *et al.* (2014), who examined the effect of different degrees of oral constrictions on glottal vibration. They compared strength of excitation (SoE) of six different consonants spanning five degrees of constriction [z, ɣ, r, l, n, ŋ] (here [r] is a trill) relative to an [a] vowel. SoE is a measure of the relative amplitude of the impulse-like excitation at the instant of significant excitation

[a)]Author to whom correspondence should be addressed.

during voicing and thus of the relative amplitude of voicing, independent of noise in the signal and largely unaffected by differences in the absorption of energy by the vocal tract itself across time (Murty and Yegnanarayana, 2008; Murty *et al.*, 2009). Mittal *et al.* (2014) found that compared to a vowel, [r] and [z] resulted in a decrease in SoE (i.e., weaker voicing). They found smaller differences among the other consonants. With [r], specifically, they also found oscillations in SoE values patterning with the open and close phases of the trill. Their study, however, is limited in that it examined only two speakers and a limited range of segment classes, such that a statistical analysis was not possible. Therefore, it is unclear which differences in SoE are statistically robust. In this study, we extend the study of Mittal *et al.* by examining not only SoE, but also using electroglottography (EGG) to examine CQ, an articulatory measure. We address the following research questions: (1) do the strength and quality of voicing differ in consonants with different oral constrictions, and if so, (2) does voicing change during a segmental constriction?

Finally, we also consider what implications source-filter interactions might have for the phonological notion of sonority [see Parker (2017) for review] which has been argued to play an explanatory role in a variety of phonological patterns. A traditional sonority scale with the inclusion of "flaps" (= taps) and trills (Parker, 2002) is as follows: vowels > glides > liquids > flaps > trills > nasals > obstruents. The phonetic correlates of sonority, however, are still not settled. Most commonly, sonority is equated with audibility or loudness [e.g., Fletcher (1972)] or acoustic intensity (Parker, 2002). These parameters depend on the vocal tract more than on the glottal source [e.g., high vowels have lower intensity than low vowels because the vocal tract shape dampens the signal: Lehiste and Petersen (1959)]. Others have emphasized the importance of the oral constriction aperture size [e.g., Clements (2009)], while conceding the potential influence the source, i.e., voicing, can have in enhancing resonance by providing, for example, "a strong and efficient excitation source" [Clements (2009), p. 167]. They typically, however, make no explicit reference to inherent source-filter interactions. Even when effects of the source are considered, these are often divorced from the effects of aperture size (Miller, 2012). Thus, our study has the potential to shed light on how source-filter interactions might relate to sonority.

2. Methods

2.1 Materials, participants, and procedure

To extend the investigation of Mittal *et al.* (2014), we examined the production of 14 voiced consonants with different degrees of constriction from a traditional phonological sonority scale: (1) glides ([j, w], (2) liquids [l, ɹ]), (3) trill and tap ([r, ɾ]), (4) nasal ([n]), (5) fricatives ([ð, ɤ, ʁ, z]), and (6) affricates and stop ([d͡ʒ, g͡ɤ, d]). We also included seven vowels ([i, y, e, ø, a, o, u], but for present purposes they have been pooled together for analysis. The consonants in groups (1–4) are sonorant consonants; vowels are also sonorant sounds. In contrast, the consonants in groups (5 and 6) are obstruents. Consonants were placed in a [aˈCa] context, following Mittal *et al.* (2014), whereas vowels were placed in a [ˈwV] context. Five out of the 21 total segments ([z, ɤ, n, r, l]) were examined by Mittal *et al.* (2014); no stops or affricates were examined in their study.

Twelve participants (6 M, 6 F) were recorded producing three repetitions of each consonant and vowel. Since our segment set goes beyond the inventory of any one language, and we additionally wanted voicing to be maintained through the consonantal gesture, the participants were all trained phoneticians, all of whom were proficient in English (7 native American English speakers; 2 native Singapore English speakers; 1 each of Japanese, Mandarin Chinese, and Russian). Audio signal recordings were made using a high-quality B & K microphone, with simultaneous EGG signal recordings using a Glottal Enterprises EG2-PCX electroglottograph. Both signals were obtained at a sample frequency of 22 kHz using PC-QUIRER (Tehrani, 2015b) in a sound-attenuated recording booth. Two other participants were also recorded, but their data were excluded from the analysis due to weak EGG signals, and/or lack of voicing in stops.

2.2 Data analysis

The audio recordings were segmented manually in PRAAT (Boersma and Weenink, 2015), by identifying target consonant intervals where voicing was maintained through the constriction for at least three glottal pulses. Tokens without at least three glottal pulses ($n = 112$ out of 897) were excluded, leaving 785 tokens in the analysis. Affricates were segmented as stop closure ("cl") and fricative release ("rel") separately. These were included with the stops and fricatives, respectively, in the analysis below. In the data below, the closures of stop [d] and affricate [d͡ʒ] are both coded as d-closures. For vowels, a sustained portion around the midpoint was identified which excluded transitions from the preceding glide [w]. For the trill [r], the second full closure was chosen. For taps, the entire contact interval was used.

EGG and acoustic measurements were extracted automatically from the EGG and audio signals using EGGWORKS (Tehrani, 2015a) and VOICESAUCE (Shue *et al.*, 2011). Means were taken over the entire segmented interval, and measures were scaled and centered by speaker using the *scale* function in R (R Core Team, 2015). Below, we report on two measures to examine the strength and quality of voicing. Contact quotient (CQ) is a measure, derived from the EGG signal, of the proportion of the glottal vibratory cycle where the vocal fold contact is greater than a specified threshold. Here, we use the hybrid method (Howard *et al.*, 1990): the contacting phase begins at the positive peak in the dEGG signal, and the decontacting phase ends when the EGG signal crosses the 25% threshold (Orlikoff, 1991). Herbst (2004) showed that this version of CQ performed as well as, or better than, other methods, and it has since been shown to best reflect differences in phonation in the modal-to-breathy range (Kuang, 2011). Additionally, we report on the strength of excitation (SoE) measure developed by Murty and Yengnanarayana (2008) and Murty *et al.* (2009). SoE is related to RMS energy but does not reflect energy absorption by the vocal tract, or energy contributed by noise. It is also related to the closing peak in dEGG, but according to Mittal *et al.* (2014) (p. 1935), it "may reflect changes in both the source and vocal tract system characteristics." SoE is thus a measure of the strength of voicing. There is no equivalent EGG measure.[1]

We use conditional inference trees (CITs) (Hothorn *et al.*, 2006) to examine whether segment classes of differing constriction degrees show differences in the quality and strength of voicing. CITs use an unsupervised algorithm that recursively partitions the observations into different subsets on the basis of significant differences on predictor variables. This approach does not require any *a priori* description of the number of groupings to be found. We submitted both SoE and CQ to CIT analyses, with manner class as the predictor, using the *ctree*() function from the *party* package in R (Hothorn *et al.*, 2019). Duration, which could be a factor in determining voice quality, was also included as a predictor in initial analyses. This, however, was not significant for the major manner classes of interest, therefore all analyses below have duration omitted.

3. Results

3.1 Global segmental distinctions

CITs for SoE and CQ are shown in Fig. 1 (see supplementary materials online for results including other measures H1-H2 and energy).[2] SoE divides the segments into six groups: vowels, nasal, liquids, glides/trill, fricatives/tap, stops. In general, sonorants have higher values than obstruents. In more detail, SoE tracks vocal tract constriction to some extent, with vowels having the highest values, and voiced stop closures the lowest. SoE makes distinctions among four groups of sonorants. In contrast, CQ distinguishes only four groups: vowels, glides/liquids/nasal, stops/tap, and fricative/trill. CQ is not highly related to vocal tract constriction degree, since the trill and voiced fricatives have the lowest values, indicating less vocal fold contact. This breathier voicing accords with previous work [e.g., Keyser and Stevens (2006) and Stevens (1971)] that suggests the vocal folds need to be somewhat spread for voiced fricatives, and that trills are aerodynamically like voiced fricatives (Solé, 2002).

The differences among sounds on the two measures can be seen in Fig. 2(L), which plots SoE by CQ by segment type. In general, collapsing over all tokens [Fig. 2(L), inset], SoE and CQ are moderately positively correlated ($r(738) = 0.41$, $p < 0.001$): more vocal fold contact results in stronger excitation. Since more vocal fold contact generally means more harmonic energy, this is expected. However, plotting by individual segments shows that the relation between the two measures is more nuanced. Within the obstruents, these two measures are negatively correlated: voiced fricatives and trill have low CQ but medium SoE, while voiced stop closures have low SoE but medium CQ. This is presumably because, as noted above, voiced fricatives and trill show the greatest glottal adjustment, allowing both voicing and sufficient airflow for generating frication or trilling the tongue tip. In fact, along the CQ dimension (x axis) the sonorants (consonants and vowels) are all very similar, at the far right of the plot, while the obstruents occupy most of the dimension. That is, CQ makes distinctions among the obstruents more than among the sonorants. Conversely, along the SoE dimension (y axis), we can see that the sonorants are more spread out than the obstruents. That is, SoE makes distinctions among the sonorants more than among the obstruents.

3.2 Timecourse of voicing measures: SoE and CQ

We next turn to the timecourse of both voicing measures to examine how the quality of voicing changes during a consonantal constriction, focusing on the quality of voicing in stops and trills, as compared to nasals (as a representative sonorant). We focus on a qualitative discussion of the general patterns observed in the changes in the strength of voicing as indexed by SoE [following Mittal *et al.* (2014)], and the amount of vocal fold contact as indexed by CQ, as seen in Fig. 2(R) with three representative speakers. Note that the timecourse of these voicing measures does not show individual voicing pulses. Our speakers show consistently stable (and strong) voicing

https://doi.org/10.1121/10.0001585

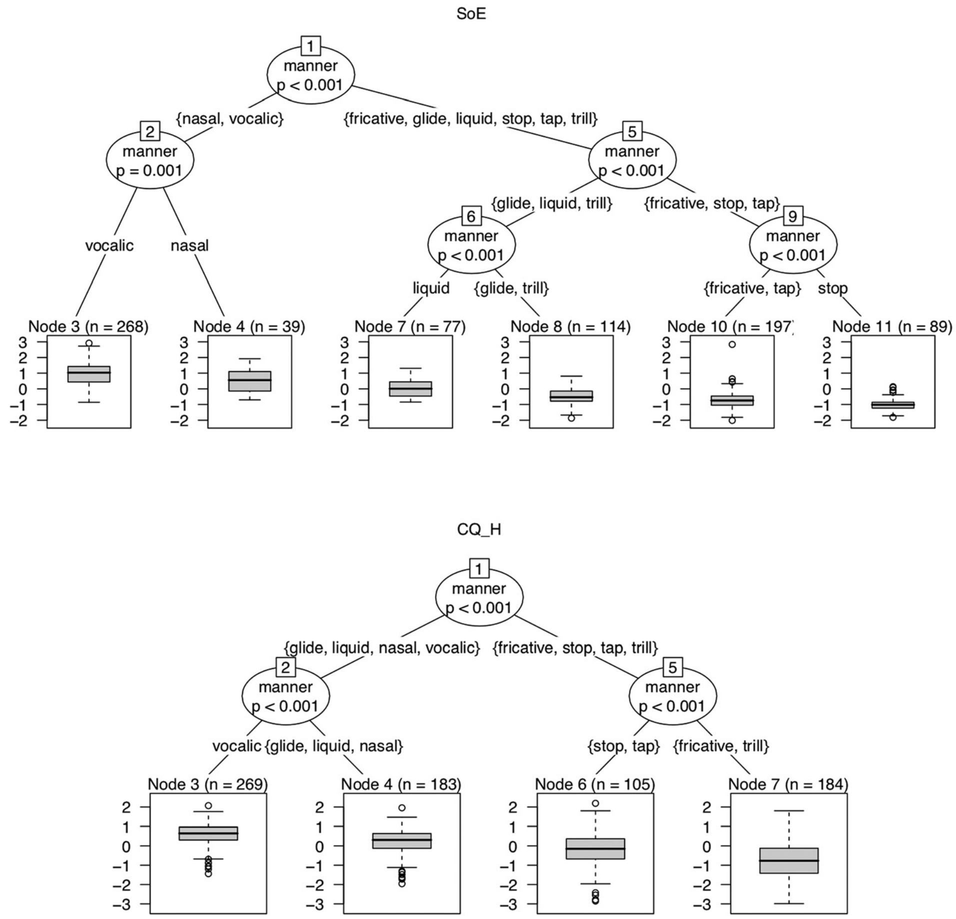

Fig. 1. Conditional inference trees for SoE (upper) and CQ_H (lower) by manner.

throughout the nasal articulation. With stops, however, speakers show two types of patterns. Some speakers show stable values of the two measures throughout the closure. But many show a drop in both measures throughout the duration of voicing as voicing becomes more difficult and weaker, sometimes dying out completely. This is in line with previous findings and is presumably due to increase in supraglottal pressure [see Solé (2018), and other references above]. Most interestingly, trills, which involve both open and closed oral articulations, show different degrees of voicing strength during each phase. For most speakers, both SoE and CQ oscillate during the trills, with open phases showing stronger voicing than closed phases. For SoE, speakers were uniform in this behavior, showing only variability in the amplitude of each oscillation. Thus,

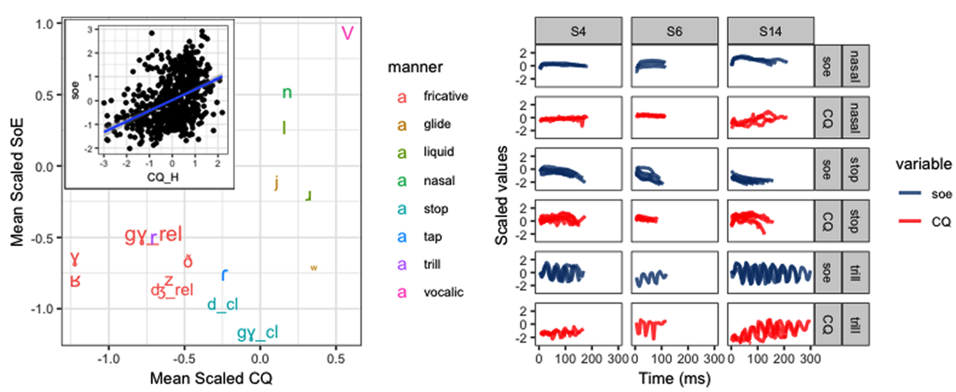

Fig. 2. (Color online) (L) Two-dimensional space of scaled SoE by scaled CQ_H by segment (inset: collapsed across segments). Size of segment label indicates standard deviations. (R) Timecourse of scaled CQ_H and SoE for nasal, stops, and trill (three representative speakers).

https://doi.org/10.1121/10.0001585

differences in voicing measures observed across segmental categories, especially in SoE, are also seen during the articulation of a single segment, as conditions for voicing change.

4. Discussion and conclusion

In this study we examined voicing in consonants with different degrees of oral constriction, extending a previous study by Mittal *et al.* (2014) by examining a wider range of consonants. The CIT analyses show that major classes of segments differ significantly in the strength and quality of voicing. Voiced obstruents show the weakest and breathiest voicing, whereas vowels show the strongest and least breathy voicing. Thus, when it is harder to sustain voicing, we observe lower SoE (less strength in voicing) and lower CQ (breathier voicing) across broad manner classes of segments. One notable result is that fricatives and trills show the breathiest voicing (lowest CQ value), while showing differences in SoE. This has implications for segmental typology and sound change, most notably, providing further evidence of the link between trills and breathy voicing. For example, Kirby (2014) showed that in some languages, such as Khmer, trills have developed diachronically into breathy voicing. Furthermore, it helps explain why breathy-modal contrasts are extremely rare in fricatives[3] and trills.[4]

Our results provide some support for the idea that the effect of oral constrictions on glottal configurations is passive and not speaker-controlled. While Mittal *et al.* (2014) assume that these are involuntary, Halle and Stevens (1967) suggest that vocal fold positioning and vibratory patterns are parameters that a speaker may adjust overtly to maintain voicing with supraglottal constriction. In this connection, Dhananjaya *et al.* (2012) and Mittal *et al.* (2014) call attention to SoE oscillations during trills along with openings and closings of the oral constriction, such that voicing is weaker during the closure phases. Our SoE data are in line with this, but alone do not differentiate if voicing changes are active or passive; conceivably the variation could reflect changes in the supraglottal contribution to the SoE measure. In contrast, CQ reflects only the glottal state, and in our data oscillates in trills similarly to SoE (though less clearly). Thus, the glottal state varies with changes in oral constriction during trills. We take the fast, cyclic oscillations (20–30 Hz) in CQ to be suggestive of passive (vs active) responses to rapid changes in the oral cavity, as assumed by Mittal *et al.* (2014). Halle and Stevens (1967) did not consider such evidence from trills. Future work would further examine simultaneously the formation of the oral constrictions along with glottal state.

Finally, our examination of source-filter interactions has possible implications for phonological sonority. Unlike previous work, our study examines measures (CQ and SoE) that are more focused on the glottal source than on the vocal tract. CQ does not distinguish enough segment classes to account for the degrees of sonority. SoE distinguishes more classes, but does not reproduce the ranking of the sonority hierarchy, in part because our glides were less vowel-like than expected, and in part because our trills have more energy than our taps. Nonetheless, the new measures examined here might help explain some of the sonority reversals observed cross-linguistically. For example, sonority reversals often involve obstruents [stops > fricatives, Jany *et al.* (2007)], which accords with the low CQ of fricatives.

In sum, our current study has examined the extent to which the strength and quality of voicing is affected by consonantal constrictions of different degrees. We have shown that, broadly speaking, voicing becomes weaker and breathier as the degree of consonantal constriction increases. We have also shown that the strength and quality of voicing changes over the course of a consonantal articulation, presumably due to changes in aerodynamic factors. Future work would seek to examine a wider range of speakers from different language backgrounds as well as a fuller set of segmental contrasts.

Acknowledgments

We would like to thank members of the 2015 Fall Speech Production course at UCLA who contributed the data to this study. This study has benefited from comments and feedback from three anonymous reviewers, an Associate Editor, Marc Garellek, and audiences at the UCLA Phonetics Lab, CUNY Phonology Forum 2016, and LabPhon 2016. We would also like to thank Soo Jin Park for coding the 2008 SoE measure of Murty and Yagnanarayana, and Yen-Liang Shue for adding it into VOICESAUCE.

References and links

[1]In our data, SoE is not strongly related to our EGG measure peak increase in contact, which is the amplitude of the closing peak in the dEGG signal, thereby giving the moment when SoE is measured.

[2]See supplementary material at https://doi.org/10.1121/10.0001585 for full set of analyses including H1-H2 and energy. We also provide the results of an analysis using only the native English speakers (n = 9) in our corpus, and on English-only coronal segments to control for segments not in English and for any possible place of articulation effects. The results are qualitatively similar as what we have presented here with all our speakers and segments.

https://doi.org/10.1121/10.0001585

[3]UPSID-451 (Maddieson and Precoda, 1989; Reetz, 1999) contains only two languages with such contrasts.

[4]It has been reported that the contrast between the two trills of Czech is one of voice quality [modal /r/ vs breathy /r̝/; Howson *et al.* (2014)]; but given that fricatives are inherently breathy, as we have shown here, their results are also consistent with /r̝/ being a fricative trill. In order to distinguish between these possibilities, it would be necessary to compare the trill directly to the Czech voiced fricatives.

Andrade, P. A., Wood, G., Ratcliffe, P., Epstein, R., Pijper, A., and Svec, J. G. (**2014**). "Electroglottographic study of seven semi-occluded exercises: Lax vox, straw, lip-trill, tongue-trill, humming, hand-over-mouth, and tongue-trill combined with hand-over-mouth," J. Voice **28**(5), 589–595.

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

Boersma, P., and Weenink, D. (**2015**). "Praat: Doing phonetics by computer" [computer program], version 5.3, http://www.praat.org/ (Last viewed November 8, 2015).

Clements, G. N. (**2009**). "Does sonority have a phonetic basis? Comments on the chapter by Vaux," in *Contemporary Views on Architecture and Representations in Phonological Theory*, edited by E. Raimy and C. Cairns (MIT Press, Cambridge), pp. 165–175.

Dhananjaya, N., Yegnanarayana, B., and Bhaskararao, P. (**2012**). "Acoustic analysis of trill sounds," J. Acoust. Soc. Am. **131**(4), 3141–3152.

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

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

Halle, M., and Stevens, K. (**1967**). "On the mechanism of glottal vibration for vowels and consonants," Q. Prog. Rep. Res. Lab. Electron. MIT **85**, 267–271.

Hamdan, A., Nassar, J., Al Zaghal, Z., El-Khoury, E., Bsat, M., and Tabri, D. (**2012**). "Glottal contact quotient in Mediterranean tongue trill," J. Voice **26**(5), 669e11–669e15.

Herbst, C. (**2004**). "Evaluation of various methods to calculate the EGG contact quotient," Diploma Thesis in Music Acoustics, KTH Speech, Music and Hearing, Stockholm, Sweden.

Hothorn, T., Hornik, K., Strobl, C., and Zeileis, A. (**2019**). "Party: A Laboratory for Recursive Part(y)itioning, 2006," R package version 0.9-0, http://CRAN.R-project.org/ (Last viewed 1 September 2019).

Hothorn, T., Hornik, K., and Zeileis, A. (**2006**). "Unbiased recursive partitioning: A conditional inference framework," J. Comput. Graph. Stat. **15**, 651–674.

Howard, D. M., Lindsey, G. A., and Allen, B. (**1990**). "Toward the quantification of vocal efficiency," J. Voice **4**(3), 205–212.

Howson, P., Komova, E., and Gick, B. (**2014**). "Czech trills revisited: An ultrasound and EGG study," J. Int. Phon. Assoc. **44**(2), 115–132.

Jany, C., Gordon, M., Nash, C. M., and Takara, N. (**2007**). "How universal is the sonority hierarchy?: A cross-linguistic study," in *Proceedings International Congress of Phonetic Sciences*, p. 16.

Keating, P. A. (**1984**). "Phonetic and phonological representation of stop consonant voicing," Language **60**(2), 286–319.

Keyser, S. J., and Stevens, K. (**2006**). "Enhancement and overlap in the speech chain," Language **82**(1), 33–63.

Kirby, J. P. (**2014**). "Incipient tonogenesis in Phnom Penh Khmer: Acoustic and perceptual studies," J. Phonetics **43**, 69–85.

Kuang, J. (**2011**). "Phonation in tonal contrasts," M.A. thesis, University of California, Los Angeles, CA.

Lehiste, I., and Petersen, G. E. (**1959**). "Vowel amplitude and phonemic stress in American English," J. Acoust. Soc. Am. **31**(4), 428–435.

Lisker, L. (**1977**). "Factors in the maintenance and cessation of voicing," Phonetica **34**(4), 304–306.

Maddieson, I., and Precoda, K. (**1989**). "Updating UPSID," J. Acoust. Soc. Am. **86**(S1), S19–S19.

Miller, B. (**2012**). "Sonority and the larynx," in *The Sonority Controversy*, edited by S. Parker (Walter de Gruyter, Berlin), pp. 257–288.

Mittal, V. K., Yegnanarayana, B., and Bhaskararao, P. (**2014**). "Study of the effects of vocal tract constriction on glottal vibration," J. Acoust. Soc. Am. **136**(4), 1932–1941.

Murty, K. S. R., and Yegnanarayana, B. (**2008**). "Epoch extraction from speech signals," IEEE Trans. Audio Speech Lang. Proc. **16**(8), 1602–1613.

Murty, K. S. R., Yegnanarayana, B., and Joseph, M. A. (**2009**). "Characterization of glottal activity from speech signals," IEEE Sign. Process. Lett. **16**(6), 469–472.

Orlikoff, R. F. (**1991**). "Assessment of the dynamics of vocal fold contact from the electroglottogram: Data from normal male subjects," J. Speech Lang. Hear. Res. **34**(5), 1066–1072.

Parker, S. (**2002**). "Quantifying the sonority hierarchy," Ph.D. thesis, UMass, Amherst, MA.

Parker, S. (**2017**). "Sounding out sonority," Lang. Ling. Compass **11**(9), e12248.

Pirello, K., Blumstein, S., and Kurowski, K. (**1997**). "The characteristics of voicing in syllable-initial fricatives in American English," J. Acoust. Soc. Am. **101**(6), 3754–3765.

R Core Team (**2015**). "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/ (Last viewed 25 October 2018).

Reetz, H. (**1999**). "Simple UPSID interface," http://web.phonetik.uni-frankfurt.de/upsid.html (Last viewed 1 November 2019).

Rothenberg, M. (**1968**). *The breath-stream dynamics of simple-released-plosive production (Bibliotheca Phonetica No. 6)* (Karger, Basel/New York).

https://doi.org/10.1121/10.0001585

Shue, Y. L., Keating, P. A., Vicenik, C., and Yu, K. (**2011**). "Voicesauce: A program for voice analysis," in *Proceedings International Congress of Phonetic Sciences 17*, pp. 1846–1849.

Solé, M.-J. (**2002**). "Aerodynamic characteristics of trills and phonological patterning," J. Phon. **30**(4), 655–658.

Solé, M.-J. (**2010**). "Effects of syllable position on sound change: An aerodynamic study of final fricative weakening," J. Phon. **38**(2), 285–305.

Solé, M.-J. (**2018**). "Articulatory adjustments in initial voiced stops in Spanish, French and English," J. Phon. **66**, 217–241.

Stevens, K. N. (**1971**). "Airflow and turbulence noise for fricative and stop consonants," J. Acoust. Soc. Am. **50**(4B), 1180–1192.

Stevens, K. N. (**1977**). "Physics of laryngeal behavior and larynx modes," Phonetica **34**(4), 264–279.

Stevens, K. N. (**1998**). *Acoustic Phonetics* (MIT Press, Cambridge).

Tehrani, H. (**2015a**). "Eggworks," http://www.linguistics.ucla.edu/faciliti/sales/software.htm (Last viewed November 1, 2015).

Tehrani, H. (**2015b**). "PCQuirerX," http://www.linguistics.ucla.edu/faciliti/sales/software.htm (Last viewed November 1, 2015).

Westbury, J. (**1983**). "Enlargement of the supraglottal cavity and its relation to stop consonant voicing," J. Acoust. Soc. Am. **73**(4), 1322–1336.

Westbury, J., and Keating, P. (**1986**). "On the naturalness of stop consonant voicing," J. Ling. **22**(1), 145–166.

van den Berg, J. (**1958**). "Myoelastic-aerodynamic theory of voice production," J. Speech Hear. Res. **1**(3), 227–244.