Received: 27 June 2019 | Revised: 6 May 2020 | Accepted: 18 May 2020

DOI: 10.1111/lnc3.12392

ARTICLE

WILEY

The cross-linguistic patterns of phonation types

Christina M. Esposito[1] | **Sameer ud Dowla Khan**[2]

[1]Department of Linguistics, Macalester College, St Paul, Minnesota, USA

[2]Department of Linguistics, Reed College, Portland, Oregon, USA

Correspondence
Christina M. Esposito, Macalester College, St Paul, MN.
Email: esposito@macalester.edu

Abstract

We provide an update on the state of research on phonation (the production of sound by the vocal folds) since Gordon and Ladefoged's, Journal of Phonetics, 2001 29, 383–406 overview, focusing on the acoustics of breathiness, creak, and other linguistic voice qualities. We highlight cross-linguistic variation, introduce measuring techniques, and discuss the relationship between phonation and other phonological dimensions (e.g., tone, vowel quality). We also review perceptual literature, an area of phonation research that has greatly expanded recently. Taken together, the studies reviewed demonstrate that phonation types indeed lie not just on a single continuum, but in a multi-dimensional space.

KEYWORDS

breathy, creaky, electroglottography, perception, phonation, register, voice quality

1 | INTRODUCTION

Phonation, or voice quality, is the production of sound by the vocal folds. A simple but useful visualization, proposed in Ladefoged (1971), is reproduced in Figure 1.

Ladefoged's (1971) unidimensional model assumes that different phonation types are produced by apertural variance between the arytenoid cartilages (roughly, varying degrees of glottal width). On each extreme are the voiceless sounds, and thus involve no voice quality per se: on one end, we see sounds that are voiceless by virtue of having maximum aperture, making

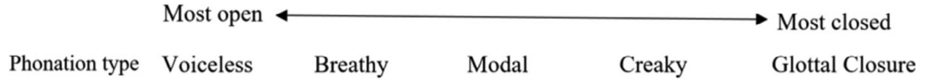

FIGURE 1 Phonation types on a continuum of glottal width, from Ladefoged (1971)

vocal fold vibration impossible, and on the other end we have complete glottal closure ([ʔ]). The three intermediate points—breathy, modal, and creaky—are all produced with vibrating vocal folds, and are arranged in order from largest to smallest vocal fold aperture.

The majority of the world's languages (see Table 3, cell 1) distinguish between one (or both) of the extremes on this continuum (i.e., voiceless sounds) and one center point (i.e., voiced sounds). For such languages, there is only modal voicing, cross-linguistically the most common voiced phonation type. In addition to modal phonation, countless languages also make distinctions within the voiced range of this continuum (e.g., breathy, creaky).

As just one example of how these phonation types might be produced, Figure 2 shows spectrograms of breathy, modal, and creaky voicing in White Hmong. In the spectrogram of the breathy example, *pog* [pɔ̤˦˨] "paternal grandmother," we see noise distributed throughout the vowel (visualized as a kind of fuzziness), particularly in the mid and higher frequencies. In the spectrogram of the creaky example *pom* [pɔ̰˨˩] "to see," we see wider, more irregular intervals between vocal pulses (visualized as vertical striations), particularly at the end of the vowel. The spectrogram of the modal example *poj* [pɔ˥˨] "female" has neither of these characteristics.

Phonation types have been studied by a wide range of disciplines over a long history, leading to the use of competing terminology for very similar phenomena. Some terms are simply conventions within particular languages or regions. For example, early studies on breathy voice used the term "murmur" (e.g., Fischer-Jørgensen, 1967; Ladefoged, 1971), especially for South Asian languages. Similarly, "clear" is used in opposition to "breathy" for Austroasiatic languages (e.g., Watkins, 2002), referring roughly to modal voice or something slightly further towards the closed end of the continuum. "Muddy" is traditionally used in the context of Sinitic languages for breathier phonations in historical analyses of Middle Chinese as well as synchronic analysis of modern Wu varieties (Gao & Hallé, 2017). Across studies, one finds "lax," "slack," "muddy," "murmured," and "whispery" on the breathier (more open) half of Ladefoged's continuum, and "tense," "stiff," "laryngealized," and "glottalized" on the creakier (more closed) half. These are shown in Figure 3, Ladefoged's (1971) continuum updated with additional labels to give a rough estimate of where they could be arranged. As we show in subsequent sections, however, these labels are simply conventional, and do not reflect discrete phonetic groupings with defined boundaries.

Still, even the more elaborate continuum in Figure 3 cannot capture languages that have a richer phonation inventory than accounted for by a unidimensional model. For example, !Xóõ utilizes supraglottalic mechanisms in addition to glottalic ones, producing not just modal, breathy, and creaky voice, but also pharyngealized voice (Garellek, 2019b; Ladefoged, 1983; Traill, 1985, 1986), along with dynamic combinations of non-modal phonations (e.g., breathy-creaky). Falsetto in Pakphanang Thai (Rose, 1997), harsh voice in Bai and Dinka (Edmondson & Esling, 2006), and faucalized or hollow voice in Dinka (Denning, 1989; Edmondson & Esling, 2006) are also difficult to visualize on a single continuum. Decades after Ladefoged's (1971) proposal, Edmondson and Esling (2006) offered a model that could account for phonation types produced both glottalically and supraglottically based on the manipulation of six "valves": (a) vocal fold adduction and abduction (i.e., the dimension on which Ladefoged's model was based), (b) ventricular incursion, (c) sphincteric compression, (d) epiglotto-pharyngeal constriction, (e) laryngeal raising, and (f) pharyngeal narrowing. Manipulation of one or more of these valves produces differences in phonation; for example, creaky voice is described as having vocal folds that vibrate slowly, with sphincteric compression, and little or no ventricular incursion, while breathy voice is produced with vocal folds that oscillate anteriorly along the ligamental glottis, while the posterior portion of the glottis remains open (see

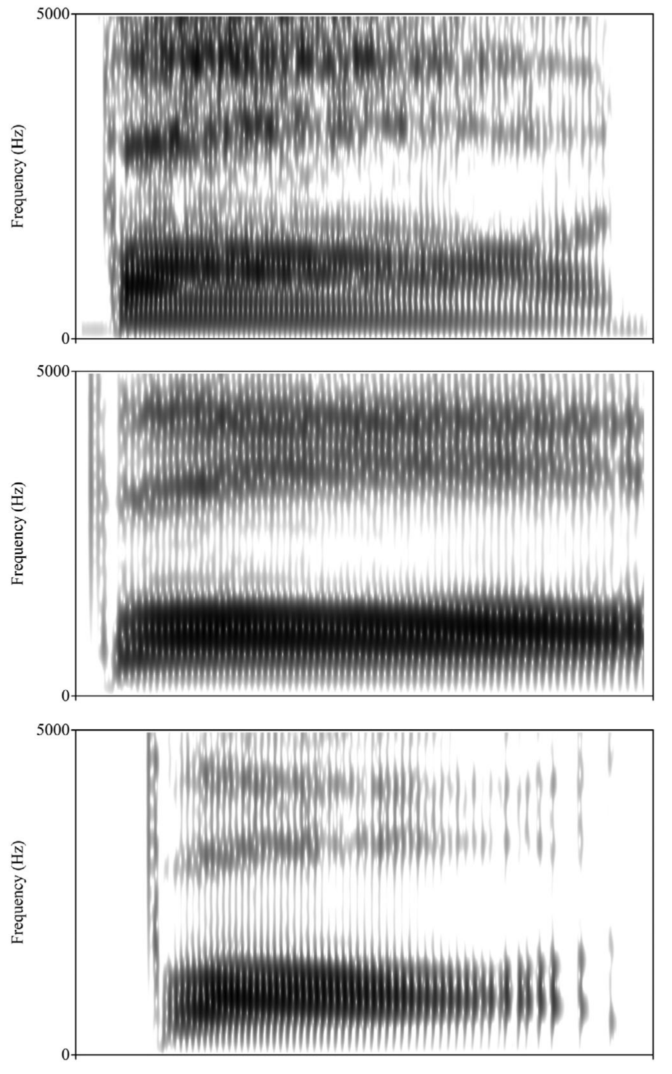

FIGURE 2 Spectrograms of the White Hmong words *pog* [pɔ̤˦˨] "paternal grandmother" (breathy, top), *poj* [pɔ˥˨] "female" (modal, middle), and *pom* [pɔ̰˨˩] "to see" (creaky, bottom)

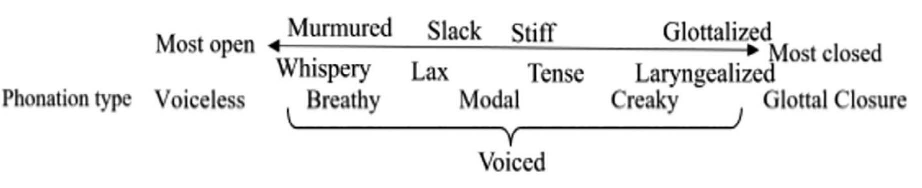

FIGURE 3 Phonation types on a continuum of glottal width from Ladefoged (1971), updated to include a selection of additional labels reflecting conventional usage. Note that their placement is somewhat arbitrary due to vague and overlapping definitions

Esling & Harris, 2005; Edmondson & Esling, 2006; Moisik, 2013 for greater details on valvular models of phonation). Nevertheless, while phonation almost certainly involves many intersecting dimensions in both articulation and acoustics (see Section 2.3 for a multidimensional approach to acoustic cues of phonation), most authors (ourselves included) approach phonation using Ladefoged's unidimensional model as a convenient starting point, especially when focusing on the cross-linguistically most common contrastive voice qualities: breathy, modal, and creaky.

1.1 | Goals of paper

The goals of the current paper are (a) to provide an update on the state of research since Gordon and Ladefoged's (2001) overview, maintaining their original focus on the acoustics of phonation as it is used for contrastive and allophonic purposes, and (b) to also expand further into studies of the perception of phonation, reflecting the considerable recent growth in that field. In the interest of space and in keeping with Gordon and Ladefoged's (2001) approach, we do not attempt to provide an extensive overview of the articulatory mechanisms involved in phonation (see Garellek, 2019a, 2019b; Kreiman & Sidtis, 2011; Laver, 1981), nor do we cover the many ways in which phonation is used as a marker of prosodic structure (Epstein, 2002; Redi & Shattuck-Hufnagel, 2001), as an index of social factors (see Kreiman & Sidtis, 2011) such as gender (Davidson, 2019a, 2019b; Zimman, 2018; Becker, Khan, & Zimman, 2017; Podesva, 2013; Yuasa, 2010), sexual orientation (Podesva, 2007), or ethnicity (Szakay, 2012), or as a result of voice pathology (see Kreiman & Gerratt, 1996, 2005; Sapienza, Hicks, & Ruddy, 2011).

1.2 | Association of phonation

Phonation types can be phonologically associated to vowels and/or consonants, though association to the former has received the majority of researchers' attention. Tables 1 and 2 provide example words illustrating how contrastive breathy and creaky voice quality (respectively) can be associated to consonants, vowels, or both within a given language. Table 3 provides a list of

TABLE 1 Example words from languages (columns) illustrating the cross-linguistic variation in the association (rows) of contrastive breathy phonation

Association of breathy phonation	White Hmong	Santa Ana del Valle Zapotec	Gujarati	Marathi
None (i.e., modal)	da˥˨ “yellow”	bǎd “duck”	baɾ “twelve”	maɾ “beat”
Consonantal	dʱa˧ “separate”		bʱaɾ “burden”	m̤aɾ “(a caste)”
Vocalic	da̤˥˨ “lie, fool”	bâ̤d “Tlacolula”	ba̤ɾ “outside”	

TABLE 2 Example words from languages (columns) illustrating the cross-linguistic variation in the association (rows) of contrastive creaky phonation

Association of creaky phonation	White Hmong	Santa Ana del Valle Zapotec	K'iche'
None (i.e., modal)	pɔ˥˨ “female”	bǎd “duck”	bas(o) “glass”
Consonantal			b̰aʔ “then”
Vocalic	pɔ̰˨˩ “to see”	bâ̰d “scabies”	

TABLE 3 Languages arranged in terms of whether voice quality contrasts are associated to consonants (rows) and/or to vowels (columns)

	No voice quality contrast on vowels	**Voice quality contrast on vowels**
No voice quality contrast on consonants	(1) Arabic Cantonese Yue English Japanese Manange Mandarin Navajo Punjabi Spanish Standard Khmer Swedish Thai Yueyang Xiang	(2) Bai Burmese Chanthaburi Khmer Chong Dinka Green Mong Kedang Mazatec lgs. Mon Mpi Suai Tamang Vietnamese Wu Yi lgs., for example, Bo Zapotec lgs.
Voice quality contrast on consonants	(3) Bura Hausa Javanese Margi Most Indo-Aryan lgs., for example, Bengali Newar Tsonga	(4) Gujarati Ju|'hoansi Wa White Hmong !Xóõ

languages that represent the four possibilities of what types of segments (i.e., consonants vs. vowels) can host a voice quality contrast.

Many examples of contrastive vocalic phonations (Table 3, cell 2) come from Austroasiatic (or Mon-Khmer) languages of Southeast Asia, for example, Mon (clear vs. breathy: Abramson, Tiede, & (Luang)-Thongkum, 2015) and Suai (clear vs. breathy: Abramson, (Luang)-Thongkum, & Nye, 2004); neighboring languages of the Sino-Tibetan family, for example, Sgaw Karen (modal vs. breathy vs. creaky: Brunelle & Finkeldey, 2011), Bai (Edmondson & Esling, 2006), and Yi (Loloish) languages such as Bo, Southern Yi (Kuang & Cui, 2018; Kuang & Keating, 2014), and Hani (Maddieson & Ladefoged, 1985); Nilotic languages of East Africa, for example, Dinka (Denning, 1989; Edmondson & Esling, 2006; Remijsen & Manyang, 2009); and Otomanguean languages of southern Mexico, for example, Jalapa Mazatec (modal vs. breathy vs. creaky/laryngealized: Garellek & Keating, 2011; Blankenship, 2002; Silverman, Blankenship, Kirk, & Ladefoged, 1995; Kirk, Ladefoged, & Ladefoged, 1993) and Santa Ana del Valle (SAV) Zapotec (modal vs. breathy vs. creaky: Esposito, 2010b).

As for languages with contrastive phonation on consonants (cell 3), breathy obstruents (e.g., [bʱ]) are documented in almost all Indo-Aryan languages, for example, Hindi (Dixit, 1989) and Bengali (Khan, 2010; Mikuteit & Reetz, 2007), and to a lesser extent in other South Asian languages, for example, Malayalam (Namboodiripad & Garellek, 2016). Breathy obstruents are also found in the Owerri variety of Igbo (Ladefoged, 1964). A contrast between lax/slack and tense/stiff stops is described in Javanese (Brunelle, 2010; Thurgood, 2004). Notably, because obstruents allow for a negligible degree of airflow during the consonant articulation itself, any associated phonation on the breathier half of the continuum will be primarily realized on the adjacent vowels. This means that the phonation will be understood as phonetically vocalic but phonologically consonantal; see Section 2.4 for further discussion of this issue.

Non-modal sonorants include breathy nasals (e.g., [m̤] or [mʱ]) in Marathi (Berkson, 2019) and Tsonga (Traill & Jackson, 1988) and creaky approximants [j̰] in Chadic languages (Ladefoged, 1964) including Hausa (Lindau, 1984; Lindsey, Hayward, & Haruna, 1992), Bura, and Margi. Note that while the articulatory complications of producing breathier phonations during an obstruent are not present in sonorants (due to their continuous airflow through either the oral or nasal passages), non-modal sonorants appear to be less common than their obstruent counterparts (Berkson, 2019), though this may be due to simply being underreported in many phonetic and phonological studies.

Of note, very few languages (cell 4) contrast phonation types on both consonants and vowels; in fact, this seems to be limited to five languages !Xóõ (Traill, 1985), Ju|'hoansi (Miller, 2007), Wa (Watkins, 2002), White Hmong (Esposito & Khan, 2012), and Gujarati (Esposito & Khan, 2012)—all of which contrast breathy-voiced aspirated obstruents in addition to breathy vowels (often alongside other non-modal phonations).

Some languages are ambiguous or at least complex in the association of non-modal phonation. Tense and lax phonations in Yi languages, for example, are measurable on both sonorant onsets and vowel nuclei: one interpretation is that the phonation is associated to the vowel but allophonically spreads to the consonant, while another interpretation views phonation as associated to the entire syllable (Garellek, Ritchart, & Kuang, 2016). In Wu Chinese varieties, the presence of non-modal phonation on vowels (and occasionally on onset consonants) is generally considered a property of the tone class: higher-register *yīn* tones are associated with generally higher pitch across the rime, modal voicing across the vowel, and voicelessness in obstruent onsets, while the lower-register *yáng* tones are associated with generally lower pitch across the rime, slack voicing in the vowel, and weak voicing in obstruent onsets (e.g., Zee & Xu, 2017; Zhang & Yan, 2018). Lastly, laryngealization (*stød*) in Danish can appear on consonants or vowels, but its distribution is restricted to the end of a long vowel or on a sonorant consonant preceded by short vowel; Basbøll (2005) and Grønnum (1998, 2014) thus analyze *stød* as a feature of the end of a bimoraic syllable, rather than a feature of either the vowel or consonant. Crucially, unlike languages in cell 4, Wu, Yi, and Danish do not have a *contrast* between a syllable with non-modal phonation associated to the vowel versus an otherwise identical syllable with non-modal phonation associated to the consonant.

2 | PRODUCTION OF PHONATION

To best describe the physical properties of phonation, we begin by surveying approaches for measuring phonation type differences, both acoustically from an audio signal (Section 2.1) and using articulatory instruments, in particular an electroglottograph (Section 2.2), highlighting developments that have taken place since Gordon and Ladefoged (2001). We then overview variation in how phonation is produced (Section 2.3) and where exactly it is localized with respect to its associated segment (Section 2.4), followed by a focus on how phonation can be expressed partially through duration (Section 2.5).

2.1 | Acoustic measurements

Researchers have relied mainly on spectral balance and spectral tilt measures to quantify the acoustic signal. The most common spectral balance measure—the difference between the amplitude of the first and second harmonics (H1–H2)—reflects the open quotient, that is, the proportion of the glottal cycle during which the glottis is open (Holmberg, Hillman, Perkell, Guiod, & Goldman, 1995). H1–H2 has been used to successfully measure phonation types in a wide variety of languages such as !Xóõ (Bickley, 1982; Garellek, 2019a, 2019b), Coatzospan Mixtec (Gerfen & Baker, 2005), Jalapa Mazatec (Blankenship, 2002; Garellek & Keating, 2011; Kirk et al., 1993), Chanthaburi Khmer (Wayland & Jongman, 2003), Phnom Penh Khmer (Kirby, 2014), Green Mong (Andruski & Ratliff, 2000), White Hmong (Esposito, 2012), Marathi (Berkson, 2019), Gujarati (Khan, 2012), Mon (Abramson et al., 2015), Takhian Thong Chong (DiCanio, 2009), SAV Zapotec (Esposito, 2010b), Sgaw Karen (Brunelle & Finkeldey, 2011), Yi (Kuang, 2011), Trique (DiCanio, 2012, 2014), and so forth. Other studies have relied on spectral tilt measures, quantifying the amplitude between the first harmonic (H1) and the harmonics exciting higher formants (e.g., H1–A1, H1–A2, H1–A3); these are reported to correlate with the abruptness of vocal fold closure (Stevens, 1977). Spectral tilt measures have been used successfully in: Krathing Chong (Blankenship, 2002), Takhian Thong Chong (DiCanio, 2009), Gujarati (Khan, 2012), Marathi (Berkson, 2019), Mon (Abramson et al., 2015), SAV Zapotec (Esposito, 2010b), !Xóõ (Ladefoged, 1983), Chanthaburi Khmer (Wayland & Jongman, 2003), and Phnom Penh Khmer (Kirby, 2014). A newer low frequency measure, H2–H4, was proposed by Kreiman, Gerratt, and Antoñanzas-Barroso (2007) and has been used to measure phonation in American English (Garellek & Seyfarth, 2016); Chong, Fuzhou, Mon, and San Lucas Quiaviní Zapotec (Esposito, 2006); English and Mandarin (Bishop & Keating, 2012); Southern Yi (Kuang, 2011); and Gujarati (Khan, 2012). For all spectral balance and tilt measures, a higher value indicates a breathier phonation, while a lower value indicates a creakier one. Pennington (2005) provides a thorough discussion on the benefits and drawbacks of narrow-band spectral measures such as H1–H2, mid-band spectral measures such as H1–A3, and wide-band spectral measures (which are very rarely used in the literature on linguistic voice quality).

Other studies have used measures of periodicity (roughly, modal-ness), such as cepstral peak prominence (CPP) (e.g., Krathing Chong [Blankenship, 2002], Gujarati [Khan, 2012], Mazatec [Blankenship, 2002; Garellek & Keating, 2011], Southern Yi [Kuang, 2011], Marathi [Berkson, 2019], and Mon [Abramson et al., 2015]) and harmonics-to-noise (HNR) ratio (e.g., Ju|'hoansi [Miller, 2007], Gujarati [Khan, 2012], and Javanese [Wayland, Gargash, & Jongman, 1994]). CPP and HNR measures are all higher for modal voice, and lower in both breathier and creakier phonations due to weaker harmonics and stronger noise in the signal. For more information on acoustic measures of phonation as they relate to both production and perception of voice, see Kreiman, Gerratt, Garellek, Samlan, and Zhang (2014).

Perhaps the greatest recent accomplishment in measuring phonation has been the advancement of algorithms to correct for formant frequency and bandwidths (Hanson, 1995; Iseli, Shue, & Alwan, 2007). Prior to these corrections, the majority of studies on phonation were restricted to low vowels, to minimize the effects of the first formant on harmonic amplitude. The use of correction algorithms in research on phonation types did not become popular until

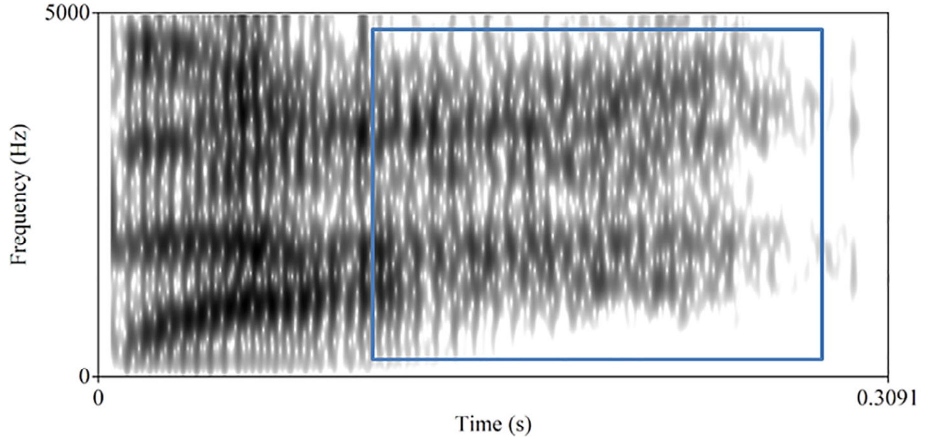

FIGURE 4 Spectrogram of the SAV Zapotec word [dậ] "powder" with a box illustrating the concentration of breathy phonation at the end of the vowel

the advent of VoiceSauce (Shue, Keating, Vicenik, & Yu, 2011), a free program that computes a variety of acoustic measurements including f0, formant frequencies, and harmonic amplitudes with corrections for surrounding formant frequencies and bandwidths.

There is evidence to suggest that H1–H2 may be a (near-)universal acoustic measure of phonation: in a study on phonation in Gujarati, White Hmong, Jalapa Mazatec, and Southern Yi, of all the acoustic measures tested, only H1–H2 was successful in all four languages (Keating, Esposito, Garellek, Khan, & Kuang, 2010). In fact, the authors are not aware of any language where H1–H2 is not a successful measure of voice quality in at least a subset of the language; languages for which H1–H2 is not *always* the best measure include (a) SAV Zapotec (Esposito, 2010b), where only female speakers' productions show a robust distinction in H1–H2; (b) Burmese (Gruber, 2011), where H1–H2 is only successful in the expected direction in phrase-final position following a high-toned word; (c) Takhian Thong Chong (DiCanio, 2009), where H1–A3 and H1–A2 distinguish the breathy registers from the non-breathy ones, while H1–H2 only distinguishes tense registers from the non-tense ones; (d) Shanghainese Wu, where H1–H2 was a successful measure of breathiness in Cao and Maddieson (1992), Ren (1992), and Gao and Hallé (2017), but not in Tian and Kuang (2019); and (e) Marathi, where H1–H2 is indeed the main acoustic cue for breathiness for male speakers but only marginally so for female speakers—this is attributed to the fact that a high H1 is confusable with nasalization for female speakers, making it less of an ideal cue for phonation (Berkson, 2013).

2.2 | Electroglottography

Electroglottography is a non-invasive technique for indirect assessment of vocal fold activity. The device, an electroglottograph (EGG), transmits an electrical impulse between two electrodes placed on either side of a participant's larynx. The device works by quantifying signal conduction: higher conduction occurs when there is greater vocal fold contact, so as the vocal folds move apart, the signal is impeded by the gap between the folds. EGG analysis has been used to measure phonation types in Maa (Guion, Post, & Payne, 2004), Vietnamese (Michaud, 2004), SAV Zapotec (Esposito, 2005), Tamang (Mazaudon & Michaud, 2006, 2008), Takhian Thong Chong

(DiCanio, 2009), Gujarati (Esposito & Khan, 2012; Khan, 2012), White Hmong (Esposito, 2012; Esposito & Khan, 2012), Yi (Kuang & Cui, 2018; Kuang & Keating, 2014), and Bo, Gujarati, Luchan Hani, White Hmong, Mandarin, Black Miao, Southern Yi, Santiago Matatlán Zapotec and San Juan Guelavía Zapotec (Keating, Kuang, Esposito, Garellek, & Khan, 2011, 2012).

The most common measure derived from an EGG signal is one of vocal fold contact during a vibratory cycle, variably referred to as contact quotient, closed quotient, or closing quotient, but generally abbreviated CQ (Baken & Orlikoff, 2000; Rothenberg & Mahshie, 1988). CQ is a ratio of the duration of the vocal fold contact phase to the total duration for a complete vibratory cycle. Phonation types produced with wider vocal fold aperture (e.g., breathy, lax) have lower CQ values compared to phonations types produced with greater vocal fold contact (e.g., creaky, tense), with modal voicing in between. Studies comparing acoustic and EGG measures (DiCanio, 2009; Esposito, 2012) found that CQ was inversely correlated with H1−H2, indicating that both reflect glottal aperture.

Other EGG measures assess the speed of vocal fold activity. One common measure, Derivative-EGG Closure Peak Amplitude (DECPA), corresponds to the amplitude of the positive peak in the derivative of the EGG signal, which is the highest rate of increase of vocal fold contact. Phonations produced with faster glottal closure have greater DECPA values than phonations produced with slower glottal closure. And, while breathy phonation is produced with vocal folds that have less abrupt closure compared to other phonation types (Childers & Lee, 1991; Klatt & Klatt, 1990), one unusual finding is that DECPA values are higher for breathy phonation than for creaky and/or modal phonation in White Hmong (Esposito, 2012) and Yi (Kuang & Keating, 2014), as well as for breathy-aspirated stops in Gujarati (Esposito & Khan, 2012). Furthermore, DECPA is only weakly correlated with both spectral slope and balance measures (Esposito, 2012), or not correlated at all (Keating et al., 2010). One suggestion for the unexpected DECPA values is that the vocal folds are further apart in breathy phonation than in other phonation types, and therefore must move more quickly to return to their initial position (Esposito, 2012; Keating et al., 2010).

2.3 | Variation in production

As with many phonological features, phonation types are variable, both within and across languages. We see evidence for variation as a function of (a) phrasal position, as in SAV Zapotec (Esposito, 2010a, 2010b) and Burmese (Gruber, 2011), (b) gender, as women are breathier than men in both Jalapa Mazatec (Blankenship, 2002) and Chanthaburi Khmer (Wayland & Jongman, 2003), or (c) speaker: in Coatzospan Mixtec, creaky/laryngealized vowels manifest as audible creak, subtle laryngealization, or without any audible creakiness (Gerfen & Baker, 2005).

Furthermore, across languages, there is variation in what is labeled "breathy" or "creaky" voice. Keating, Garellek, and Kreiman (2015) investigated the wide range of phonation types that fall under the category of "creak," such as vocal fry, multiply pulsed voice, aperiodic voice, non-constricted creak, and tense/pressed voice. These differed from prototypical creak along the following parameters: low f0, irregular f0, glottal constriction, damped pulses, and/or multiple subharmonics.

Similarly, in their studies on breathiness, Tian, Zhou, and Kuang (2019) and Tian and Kuang (2019) find support for three types of "breathier" voice qualities: slack/lax voice, dominated by changes in spectral cues (as in Southern Yi), whispery voice, produced with noise as the dominant acoustic feature (as in Shanghainese Wu), and [true] breathy voice, produced with both (as in Gujarati and White Hmong); as the authors note, these should be interpreted as overlapping regions in a continuous phonetic space and not distinct divisions.

A cross-linguistic acoustic study (Keating et al., 2010) of multiple phonation types in four languages—Gujarati, Jalapa Mazatec, White Hmong, and Yi—found that phonation types mapped out in a multidimensional space did not cluster by category (e.g., breathy vowels in White Hmong were distinct from breathy vowels in Gujarati). This may be the result of language, speaker, and/or recording differences in acoustic data which were greater than phonation category differences. However, some patterns did emerge along the dimension representing H1–H2: the creaky/tense phonations all had negative values, the modal phonations values were between −1 and 0, and the breathy/lax phonations values were all greater than 0, providing support for Ladefoged's continuum of glottal width serving as at least one dimension of the cross-linguistic variation in voice quality.

Taken together, these studies suggest that phonation contrasts in a language are typically cued in multiple dimensions, an issue we return to in Section 4 with respect to how listeners weigh multiple cues in perception.

2.4 | Localization of non-modal phonation

In some languages, such as Yi (Keating et al., 2010), non-modal phonation associated to a vowel persists for the entirety of the segment. More typically, however, non-modal phonation is localized to a portion of a vowel: in Jalapa Mazatec (Silverman, 1997; Blankenship, 2002; Keating et al., 2010; Garellek & Keating, 2011) and Chanthaburi Khmer (Wayland & Jongman, 2003), non-modal phonation is at the beginning of the vowel, in Gujarati (Keating et al., 2010; Khan, 2012) it is most robust in the middle, while in SAV Zapotec (Esposito, 2005) non-modal phonation is restricted to the end of the vowel (as seen in Figure 4). Sometimes, the localization of non-modal phonation varies within a language: in White Hmong, breathiness is localized to beginning of the vowel, while creakiness is strongest at the end (Keating et al., 2010), and in !Xóõ, breathiness persists across the vowel, while creakiness and glottalization are restricted to the end, and harsh/strident and pharyngealized qualities to the middle (Garellek, 2019a, 2019b).

Localization of non-modal phonation is also crucial in languages with complex "clusters" (i.e., phonetic sequences) of phonation type, which involve at least one non-modal phonation localized to a portion of the vowel. These include: !Xóõ, which distinguishes clusters of breathy-creaky, pharyngealized-creaky, and breathy-pharyngealized phonation in addition to modal, breathy, creaky, and pharyngealized voice (Ladefoged, 1983; Traill, 1985, 1986) and Takhian Thong Chong (DiCanio, 2009), which distinguishes modal phonation from clusters of modal-tense, breathy-tense, and breathy-modal.

Silverman (1997) proposed that localization of non-modal phonation increases perceptual salience, especially for tone, as f0 is most recoverable during the modally-phonated portion of a non-modal vowel.

The localization of non-modal phonation may play a role in distinguishing between similar, contrastive phonation types within a language. Both Gujarati and White Hmong possess breathy-aspirated obstruents and breathy vowels, which are potentially homophonous because the breathiness associated to obstruents is realized on the following vowel (Esposito & Khan, 2012). However, production evidence suggests that timing of non-modal phonation may play a role in distinguishing between these two types of breathiness: breathy consonants are characterized by an early period of breathiness at beginning of the vowel followed by decreased breathiness, while breathy vowels showed stable (Gujarati) or increasing (White Hmong) breathiness throughout the vowel (Esposito & Khan, 2012). In the spectrograms of Gujarati shown in Figure 5, the breathy voice "fuzziness" obscuring the otherwise-clear striations of modal voicing is visible across the entire breathy vowel in the word [ba̤ɾ] "outside" (middle panel), while it is restricted to the onset of the modal vowel following the breathy consonant in the word [bʱaɾ] "burden" (bottom panel).

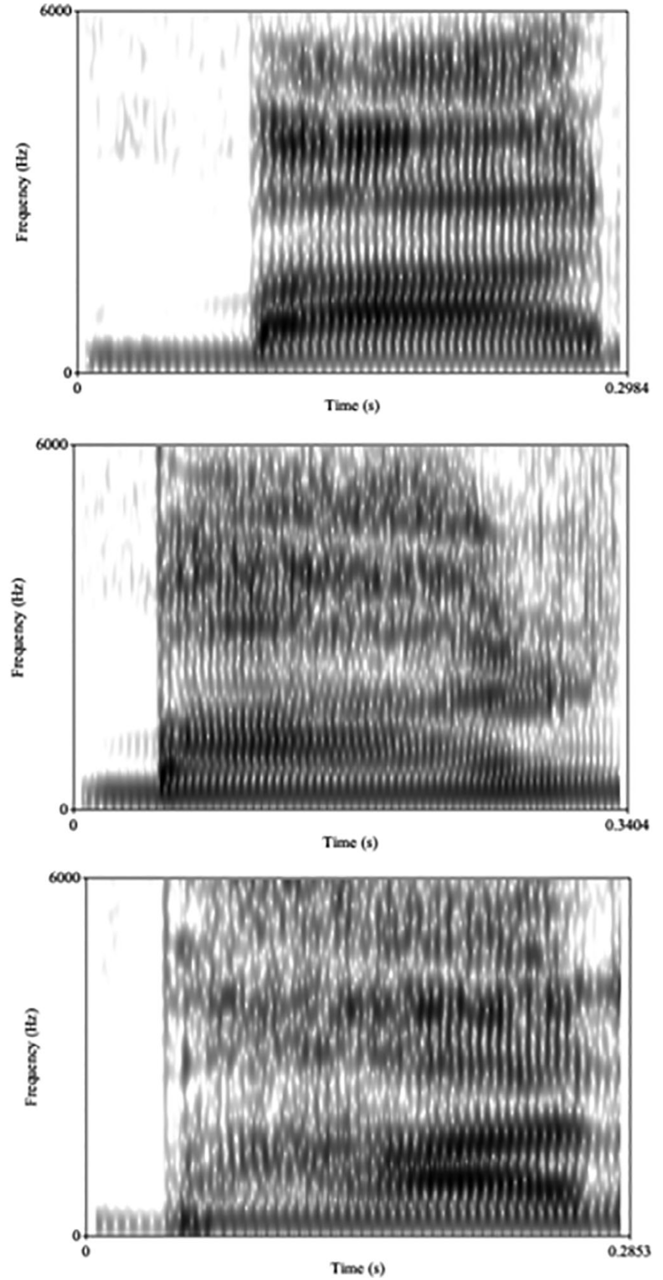

FIGURE 5 Spectrograms of Gujarati words બાર [baɾ] "twelve" (top panel, modal voicing), બહાર [ba̤ɾ] "outside" (middle panel, breathy vowel), and ભાર [bʱaɾ] "burden" (bottom panel, breathy-aspirated consonant)

2.5 | Duration of non-modal phonation

Duration of non-modal phonation varies with language, with almost every possible pattern being attested. Non-modal vowels are longer than their modal counterparts in Kedang (Samely, 1991), Jalapa Mazatec (Kirk et al., 1993), Chanthaburi Khmer (Wayland & Jongman, 2003), Khmu' Rawk (Abramson, Nye, & (Luang-)Thongkum, 2007), and Gujarati (Fischer-Jørgensen, 1967), while in other languages, non-modal vowels are shorter than their modal counterparts, for example, the creaky vowels in Hmong (Andruski & Ratliff, 2000; Esposito, 2012) and Coatzospan Mixtec (Gerfen & Baker, 2005) and the two phonation clusters involving tense voice in Chong (DiCanio, 2009). And in languages like Yalálag Zapotec (Avelino, 2010) and Suai (Abramson et al., 2004), there is no difference in duration between non-modal and modal vowels.

3 | RELATIONSHIP TO OTHER PHONOLOGICAL CATEGORIES

In many languages, phonation contrasts interact with contrasts in other phonological dimensions, such as airstream mechanism (Section 3.1), tone (Section 3.2), and vowel quality (Section 3.3). Non-modal phonation can also arise as a result of coarticulation (Section 3.4). As we explain below, when phonation categories are hard to separate from tone and/or vowel quality categories, the term "register" (Henderson, 1952) is often used to label this multidimensional feature.

3.1 | Relationship to airstream mechanism

Two classes of non-pulmonic consonants involve a narrowing and downward or upward movement of the glottis: glottalic ingressive (i.e., implosive) and glottalic egressive (i.e., ejective) consonants, respectively. Due to the glottal constriction required for these non-pulmonic sounds, they are often associated with creakier phonations in many languages, and by extension, pulmonic sounds can be associated with breathier phonations to help enhance their contrast with glottalic sounds.

As an example of the former case, creaky obstruents (e.g., [b̰]) are documented as an optional realization of implosives (e.g., [ɓ]) in West African languages such as Bura, Hausa, Margi, Kalabari, and Igbo (Ladefoged, 1964) and Mayan languages such as K'iche', Kaqchikel, Q'eqchi', Tz'utujil, and Poqomchi' (Pinkerton, 1986). And as an example of the latter case, slack voice is described in the Nguni languages of southern Africa (Rycroft, 1980), for example, Xhosa (Jessen & Roux, 2002) and Zulu (Traill, Khumalo, & Fridjon, 1987), as a frequent (but optional and inconsistent) realization on vowels following voiced pulmonic consonants, to better distinguish them from implosives.

3.2 | Relationship to tone

Cross-linguistically, languages vary with respect to the contrastive role played by f0 (e.g., tone or pitch accent contrast) and by non-modal phonation on vowels. Selected examples are displayed in Table 4, with the contrastive role of f0 shown in rows and the contrastive role of vowel phonation shown in columns.

Ignoring cell 1, for which there is neither a tone/pitch accent nor phonation contrast on vowels, we can focus on languages that have at least one of these kinds of lexical contrasts. In cell 4a, both tone and vowel phonation are independently contrastive, and thus different combinations of tone and vowel phonation are attested; for example, in Jalapa Mazatec, the three tones (low, mid, high) cross with the three phonations (modal, breathy, creaky/laryngealized) to give nine attested combinations of these two suprasegmental dimensions (Garellek & Keating, 2011; Kirk et al., 1993; Silverman et al., 1995). Dinka (Andersen, 1993; Denning, 1989; Edmondson & Esling, 2006; Remijsen & Manyang, 2009), Mpi (Blankenship, 2002), Yalálag Zapotec (Lancia, Avelino, & Voigt, 2013), Yi languages (Kuang & Keating, 2014), and !Xóõ (Garellek, 2019a, 2019b; Traill, 1985) also allow for tone and phonation to be (nearly) fully cross-classified, demonstrating the orthogonality of the two dimensions in these languages.

Arguably more interesting, however, are the two ways in which the line between the tone and phonation can be blurred. Languages in cells 2a and 2b are described as having contrastive phonation on vowels, but no contrast in tone, even though the contrastive phonations in 2b include f0 specifications as part of their realization. Languages in cells 3a and 3b are described as having contrastive tone, but no contrast in phonation, and yet the contrastive tones in 3b are described as having voice qualities allophonically associated with them. In cells 4a and 4b, tone and phonation interact in all words, but while languages in cell 4a suggest two orthogonal

TABLE 4 Languages arranged by the lexically contrastive nature of f0/tone/pitch accent (rows) and voice quality on vowels (columns). Further distinctions are made where tone and voice quality interact (2b, 3b, 4b)

	No voice quality contrast on vowels		**Voice quality contrast on vowels**	
No tone/pitch accent contrast	(1) *neither f0 nor phonation contrast* Australian Aboriginal lgs. Most Austronesian lgs. Most Indo-European lgs. Berber, Cushitic, and Semitic lgs. Standard Khmer Turkic lgs.		(2a) *f0 is not a cue for phonation* Danish Gujarati (2b) *f0 is a cue for phonation* Chanthaburi Khmer Chong Javanese Kedang Mon Suai Wa	
Tone/pitch accent contrast	(3a) *phonation is not a cue for tone* Japanese Navajo Punjabi Manange Most W. Afr. Lgs. Swedish Central Thai	(3b) *phonation is a cue for tone* Cantonese Yue Khmu' Rawk Mandarin Pakphanang Thai Ph.Penh Khmer Yueyang Xiang	(4a) *phonation & tone are orthogonal* Dinka Mazatec lgs. Mpi Yalálag Zapotec Yi lgs., for example, Bo	(4b) *phonation & tone are fused* Black Miao Burmese Green Mong Most Zapotec lgs. Tamang Vietnamese White Hmong Wu

dimensions of tone versus phonation, languages in cell 4b suggest that these are two aspects of one dimension. Thus, as we explain more fully below, three of these six cells (2b, 3b, and 4b) involve languages in which tone and phonation are hard if not impossible to separate.

Languages in cell 3b have lexical tones and no lexical vowel phonation, but a subset of the tone categories have been reported to be associated with optional non-modal phonation; for example, the low(-dipping) tone 3 of Mandarin Chinese is usually (but not always) accompanied with creak (Kuang, 2017), and the extra low(-falling) tone 4 of the Cantonese variety of Yue Chinese is often accompanied by creaky voice (Yu & Lam, 2014). On the other extreme of the pitch range, the Yueyang variety of Xiang Chinese (Peng & Zhu, 2010) and the Pakphanang variety of Thai (Rose, 1997) use falsetto for the highest tones.

Inversely, languages in cell 2b are generally considered to have contrastive vowel phonation but not lexical tone, and yet f0 is a robust acoustic property distinguishing the phonation categories in these languages; for example, breathy vowels in Wa (Watkins, 2002) and Kedang (Samely, 1991) consistently begin with a lower f0 than their clear/modal counterparts, and breathy vowels in Chanthaburi Khmer (Wayland & Jongman, 2003) have a slightly higher f0 than their clear counterparts. This systematic role of f0 is why members of cell 2b are sometimes called "register languages."

Languages in cell 4b, most of which are concentrated in Mainland Southeast Asia, are also often called "register languages," as it is impossible to state unequivocally they have lexical "tones" as opposed to lexical "phonations," or vice versa. In such languages, suprasegmental categories are consistently produced with both a particular f0 specification as well as a voice quality specification. For example, mid-rising *ngã* tone of Vietnamese generally involves glottalization vowel-medially (Brunelle, Nguyễn, & Nguyễn, 2010; Nguyễn & Edmondson, 1998; Phạm, 2001) and the low-falling *huyền* tone is described as having a lax or breathy quality (Phạm, 2001, 2003), at least in isolation (see Brunelle et al., 2010 for continuous speech), though the low-falling pitch is also been reported to accompanied by creak (Michaud, 2004; Nguyễn & Edmondson, 1998); the *m* tone of White Hmong is produced with creak and falling pitch, and the *g* tone is produced with breathiness and a high-falling pitch, while the other five tones are all produced with modal voice but different pitch contours (Esposito, 2012); each of the four tonal categories in Burmese (high, low, creaky, killed) has specifications of pitch, voice quality, vowel quality, and duration that vary by phrasal position (Gruber, 2011); the low tone 4 and low-rising tone 3 of Tamang are breathy (Mazaudon & Michaud, 2008), while the cognate tones in closely-related Manange are consistently modal (Hildebrandt, 2005); and in SAV Zapotec, syllables fall into one of four categories: breathy phonation with falling tone, creaky phonation with low falling tone, modal phonation with a high tone, and modal phonation with rising tone (Esposito, 2010b). Black Miao uses tense voice in the extra-high tone (55) and vocal fry in the extra-low tone (11), and uses breathy voice to distinguish the mid tone (33) from the nearby low (22) and high (44) tones, which are modal (Kuang, 2013). Kuang (2013, 2017) distinguishes these uses of voice quality on lexical tone based on their intrinsic association with f0: while the use of tense voice in the extra-high tone and vocal fry in the extra-low tone in Black Miao can be considered "pitch-driven non-modal phonation," in that the non-modal phonation is an epiphenomenon of producing tones at the two extremes of the f0 range, the use of breathy voice in the mid tone cannot be explained as such; instead, Kuang calls the latter a "pitch-independent non-modal phonation," part of the phonological specification of the tone.

It should thus become clear that for cells 2b, 3b, and 4b, it can be seen as arbitrary to classify a language as having a "tone contrast" versus a "phonation contrast," or something that involves both equally; we return to this issue in Section 5 on perception. Curiously, for those languages in which vowel phonation plays a systematic role in lexical contrasts (cells 2a, 2b, 3b, 4a, 4b), we note that f0 also plays a major role in virtually every example; the only exceptions we are aware of (cell 2a) are Danish, for which Grønnum (2014) demonstrates that f0 perturbations are not a reliable cue for laryngealization (*stød*), and Gujarati, for which Khan (2012) argues f0 plays no systematic role in enhancing breathy phonation. It is noteworthy though, that even for this latter language, older studies in the form of impressionistic description (Pandit, 1957) and instrumental work (Dave, 1967; Fischer-Jørgensen, 1967) report a small and statistically inconsistent relation between f0 and breathy phonation, suggesting at least a subtle and complex interplay between the two. We return to this curiosity in Section 4.1.

3.3 | Relationship to vowel quality

Formant frequency may differ based on phonation type, but this is language-specific. Compared to their modal counterparts, a lower first formant frequency (F1) is found for: (a) breathy vowels in Kedang (Samely, 1991), Nilotic languages (Denning, 1989), various Mainland Southeast Asian languages (Bradley, 1982; Hombert, 1978), and Xhosa (Jessen & Roux, 2002), (b) breathy and breathy-creaky vowels in Krathing Chong ((Luang)-Thongkum, 1987), and (c) lax phonation in Yi (Kuang, 2011), meaning that these breathier phonations may be associated to higher vowel quality. Similarly, higher F1 (i.e., lower vowel quality) is reported for creaky/tense phonation in Hani (Maddieson & Ladefoged, 1985) and Mpi (Blankenship, 2002). However, in languages such as Mon ((Luang)-Thongkum, 1987), Nyah Kur ((Luang)-Thongkum, 1986), Suai ((Luang)-Thongkum, 1986), Khmuʹ Rawk (Abramson et al., 2007), Gujarati (Khan, 2012), Jalapa Mazatec (Garellek & Keating, 2011), F1 frequencies are largely unaffected by phonation type. While the results are mixed when we compare across languages, a study on the relationship between voice quality and vowel quality in eight languages—!Xóõ, Burmese, Gujarati, Jalapa Mazatec, Mon, White Hmong, Yi, and Zapotec—did show a cross-linguistic pattern in the relationship between vowel quality and voice quality; when using H1–H2 as the measure of voice quality, vowels with higher F1 and F2 (i.e., lower, fronter vowels) tend to be produced with creakier phonation, with voice quality contrasts more robustly distinguished, while vowels with lower F1 and F2 (i.e., higher, backer vowels) tend to be breathier, with voice quality contrasts less dispersed (Esposito, Sleeper, & Schafer, 2019 to appear). At the greatest extreme, Luanyjang Dinka has a breathy versus modal ("brassy") distinction across six vowel qualities, except that for the highest, backest vowel, there is only a breathy /ṳ/ and no modal counterpart (Remijsen & Manyang, 2009). It remains for future research whether this connection is the direct result of an intrinsic connection between voice quality and vowel quality, or due to speakers' efforts to enhance the voice quality contrasts in vowel qualities that otherwise lend themselves to more ambiguity in phonation (cf. vowel harmony as enhancement of weak roundedness in lower/fronter vowels: Kaun, 2004). Some support for the former theory comes from additional studies that suggest that there could be a perceptual benefit to the covariation of phonation type and vowel quality: listeners consistently perceive breathier productions as higher vowel quality (Lotto, Holt, & Kluender, 1997) and creakier productions as lower (Brunner & Żygis, 2011).

In a related sense, degree of phonation may vary by vowel quality. In Green Mong (Andruski & Ratliff, 2000), the low vowel /a̤/ is reported to be breathier than /i̤/ and /ṳ/. The opposite is true of Yi (Kuang, 2011), however, where phonation differences are less apparent in open vowels than close ones. Moreover, in other languages, such as White Hmong

(Esposito, 2012) and Gujarati (Khan, 2012), there is no significant difference between phonation types across different vowel qualities.

3.4 | Coarticulation

Non-modal phonation can arise due to the effect of glottal consonants adjacent to vowels, for example, creaky vowels [V̰] resulting from the coalescence of /VʔV/ in K'iche' (Baird, 2011) or as a realization of glottal constriction as English (Dilley, Shattuck-Hufnagel, & Ostendorf, 1996; Garellek & Seyfarth, 2016); and breathy vowels [V̤] as a realization of /h/-adjacent vowels in English (Epstein, 1999; Garellek, 2012; Ladefoged, 1983). While the bulk of existing research on non-modal phonation has focused on its contrastive uses, future work may reveal that more languages vary phonation allophonically.

4 | PERCEPTION OF PHONATION

The question of how listeners perceive phonation differences has received much attention in the linguistic community in recent years, in particular due to technological developments facilitating (re)synthesis of acoustic features. Research in this vein underscores the observation that phonation is multidimensional not just in its production, but also in its perception; listeners can draw from a range of cues to a particular phonological contrast, from f0 to duration to the various acoustic measures of phonation discussed in Section 2.1 above.

4.1 | Perceiving non-modal phonation

Many studies on the perception of phonation focus on determining which acoustic cues are most salient to speakers. The earliest studies utilized speakers of Gujarati, a language well-suited for this task because it contrasts phonation, but not tone; results showed that phonation types where the spectrum was dominated by the first harmonic (H1) were consistently judged to be breathy, regardless of whether the stimulus was naturally produced (Fischer-Jørgensen, 1967) or synthesized (Bickley, 1982). More recently, the question of the perceptual salience of acoustic cues has resurfaced, finding that H1–H2 (or a correlated EGG measure, such as CQ) cued phonation type differences in Green Mong (Andruski, 2006), Yi (Kuang, 2011) and Itunyoso Trique (DiCanio, 2014), but, in this last case, only when durational cues were ambiguous. In White Hmong (Garellek, Keating, Esposito, & Kreiman, 2013) listeners relied on both H2–H4 and H1–H2, to distinguish breathy from modal phonation; it is not surprising that listeners relied on multiple acoustic cues, as phonation employs a multi-dimensional acoustic space.

Other studies have looked at the role of native language background (i.e., speaking a language with (a) only modal phonation, (b) contrastive non-modal phonation, or (c) allophonic non-modal phonation) in sensitivity to the perception of non-modal phonation. Results showed that different listener populations all rely on H1–H2, but to differing degrees. For example, Gujarati listeners were aided by their native phonemic contrast in phonation in that they did better at distinguishing breathy from modal vowels, compared to English and Spanish speakers, though all three populations relied on H1–H2 (Esposito, 2010a). Kreiman, Gerratt, and

Khan (2010) showed that Gujarati listeners were more sensitive to H1–H2 differences, even small ones, than listeners of English or Thai. Likewise, Mandarin listeners were more sensitive to changes in H1–H2 than English listeners (Kreiman & Gerratt, 2010). Taken together, these studies show that listeners are sensitive to H1–H2 regardless of language background, but that the degree of sensitivity is language-dependent.

Another main area of focus is the role of phonation in the perception of contrastive tone. In Cantonese Yue (Yu & Lam, 2014), Mandarin (Yang, 2011), and Green Mong (Andruski, 2006), non-modal phonation associated with certain tones increases the accuracy of tone identification; in the case of Mandarin it also increases the speed of identification (Belotel-Grenié & Grenié, 1994, 2004). And, in Coatzospan Mixtec, listeners are able to utilize small pitch and amplitude perturbations timed in the middle of the vowel (in the absence of spectral cues) to determine the difference between creaky/laryngealized versus modal vowels (Gerfen & Baker, 2005).

However, not all tonal contrasts are cued by phonation differences. Garellek et al. (2013) showed that phonation cues in White Hmong are important for identifying the high-falling breathy tone (*g* tone), but not the falling creaky tone (*m* tone), suggesting that breathiness is key to distinguishing between similar tones, while creakiness is secondary to pitch. Brunelle and Finkeldey (2011) find a similar split in Sgaw Karen: while voice quality differences are important in distinguishing tone, their contrastive role is restricted to distinguishing modal/breathy phonation from creaky/glottalized. And in Black Miao (Kuang, 2013), which contrasts five level tones (11, 22, 33, 44, 55), we see a similar pattern in which only a subset of the tones are distinguished by phonation: while tones 55 (tense) and 11 (vocal fry) are distinguished by both pitch and phonation, 33 (breathy) is distinguished from 22 and 44 (both modal) by phonation type differences, while 22 and 44 are both modal and thus distinguished only by pitch difference, making them the most confusable.

Not surprisingly, phonation contrasts are cued in multiple dimensions and the relative importance of these cues are language-specific. This is also true of other phonological contrasts, perhaps most famously stop consonant voicing, which even in a single language like English is cued by multiple dimensions including f0, F1 frequency at onset, burst intensity, closure duration, and preceding vowel duration, in addition to voice onset time (Kingston & Diehl, 1994), but these cues and their relative importance is language-specific (Kingston & Diehl, 1994) and lie on a continuum (e.g., Keating, Mikoś, & Ganong, 1981). Voice quality contrasts can as such be seen as doubly "multidimensional": in one sense, modulations of vocal fold vibration are acoustically and articulatory multidimensional in nature, as they are produced and perceived through a range of cues that can be weighted differently, across languages (e.g., Esposito, 2010a, 2010b) and even within a single language (e.g., DiCanio, 2014; Zhang & Yan, 2018). In another sense, the phonological categories associated with these voice quality contrasts are also multidimensional, in that they often incorporate tone, vowel quality, duration, and other features to create highly complex contrasts as described in detail in Section 3.

4.2 | Perception and linguistic change

As can be expected given this doubly multidimensional nature of voice quality contrasts, perception studies have revealed that some languages, such as Southern Yi (Kuang & Cui, 2018) and the Kuai dialect of Suai (Abramson et al., 2004) are in a state of transition. Studies on these languages suggest that their register distinction is being converted from one which is distinguished primarily by phonation to one in which pitch is dominant, as in Suai (Abramson et al., 2004), Kammu (Svantesson & House, 2006), or Vietnamese (Thurgood, 2002), or to one in which vowel quality is dominant, as in Southern Yi (Kuang & Cui, 2018) and Standard Khmer (Wayland & Jongman, 2002). In a slightly different situation, Phnom Penh Khmer has lost consonant+rhotic clusters but developed a tone distinction in which breathiness plays a perceptual role (Kirby, 2014). Indeed, many Austroasiatic languages in particular are described as having undergone or as currently undergoing some form of evolution to generate a phonation contrast, which in many cases has further morphed into a tonal contrast (Sidwell & Rau, 2015). The reverse, in which a phonation contrast arises out of what was historically a tonal contrast, is also argued for in Quiaviní Zapotec (Uchihara, 2016).

The results of perception studies on phonation have offered insight into the distribution of non-modal phonation. Breathy sonorants, in particular breathy nasals, are rare, both within and across languages when compared to their obstruent counterparts (Berkson, 2019). There may be a perceptual explanation for this: a study on Marathi revealed that listeners identified breathiness in obstruents with significantly greater accuracy than in sonorants (Berkson, 2016). And, another study on the typologically rare contrast—breathy-voiced consonants versus breathy vowels—found that acoustic differences between these two types of segments were not robust enough to be perceptually salient for listeners of Gujarati (Esposito, Khan, Berkson, & Nelson, 2019). While it had been hypothesized that differences in timing and magnitude of phonation could play a role in distinguishing these phonation types (Esposito & Khan, 2012), but Esposito, Khan, et al. (2019) demonstrate that this may not be the case. Listeners' inability to reliably distinguish between these types of consonants might explain why these contrasts are so rare.

5 | CONCLUSION

In the nearly two decades since Gordon and Ladefoged's (2001) review of cross-linguistic phonation types, there has been a surge of high-quality research on the topic. Acoustic, articulatory, and perceptual evidence has emerged to support the idea that while phonation contrasts do exist in a continuous space (as laid out by Ladefoged 30 years prior), that space is more accurately conceived of as multidimensional rather than as a single continuum of glottal width. This complex of phonetic realizations of voice quality shows wide variation within and across languages, and can be measured using a constellation of acoustic and EGG parameters, and yet we see evidence for the near-universality of one acoustic measure (H1–H2) and its corresponding EGG measure (CQ) across languages, alongside CPP and other highly useful measures of periodicity. On top of the acoustic and articulatory multidimensionality of voice quality narrowly defined, the phonological categories associated with phonation contrasts are also often expressed through yet another kind of multidimensionality, in which these contrasts incorporate aspects of vowel quality, duration, and other complex phonetic features along with contrasts produced through laryngeal modulation itself.

The increased interest in voice quality research has also further emphasized its intricate interactions with phonological concepts like tone and vowel quality, often treated as multiple phonetic aspects of a single register feature. In fact, the relation between voice quality and tone is so pervasive, that (aside from some languages in which phonation and tone are fully orthogonal) practically every language that uses voice quality phonemically also uses variation in f0 to mark voice quality contrasts. And conversely, listeners generally expect voice quality to be a marker of lexical tone, as researchers find faster, more accurate lexical identification when non-modal voice qualities are predictably realized with tone.

Lastly, perceptual experiments have also shed light on the looming question of why it is so rare to find contrasts in the association of non-modal phonation, that is, a contrast between non-modal vowels versus non-modal consonants, within the same language. Only five languages (to our knowledge) are documented to have this contrast in association. Results from perception work suggest that, while non-modal phonation is fairly common across many different languages and language families, distinguishing non-modal phonations associated to different segments is highly error-ful and subject to reduction to a two-way modal versus non-modal distinction rather than a three-way distinction between modal, non-modal consonant, and non-modal vowel distinction.

Even still, more work remains to be done to expand our understanding of non-modal phonation. While the current review seeks to connect the acoustics and perception of phonemic and allophonic uses of non-modal voice quality, more connections to studies of prosodic, sociolinguistic, and pathological aspects of non-modal voice quality in a diverse range of language varieties are necessary to truly understand this multidimensional, dynamic feature of the voice.

ORCID

Christina M. Esposito https://orcid.org/0000-0002-9357-1560
Sameer ud Dowla Khan https://orcid.org/0000-0001-8678-2872

REFERENCES

Abramson, A. S., (Luang)-Thongkum, T., & Nye, P. W. (2004). Voice register in Suai (Kuai): An analysis of perceptual and acoustic data. *Phonetica*, *61*, 147–171.

Abramson, A. S., Nye, P. W., & (Luang)-Thongkum, T. (2007). Voice register in Khmu': Experiments in production and perception. *Phonetica*, *64*, 80–104.

Abramson, A. S., Tiede, M. K., & (Luang)-Thongkum, T. (2015). Voice register in Mon: Acoustics and electroglottography. *Phonetica*, *72*, 237–256.

Andersen, T. (1993). Vowel quality alternation in Dinka verb inflection. *Phonology*, *10*, 1–42.

Andruski, J. (2006). Tone clarity in mixed pitch/phonation-type tones. *Journal of Phonetics*, *34*, 388–404.

Andruski, J., & Ratliff, M. (2000). Phonation types in production of phonological tone: The case of green Mong. *Journal of the International Phonetic Association*, *30*, 37–61.

Avelino, H. (2010). Acoustic and electroglottographic analyses of nonpathological, nonmodal phonation. *Journal of Voice*, *24*, 270–280.

Baird, B. O. (2011). Phonetic and phonological realizations of 'broken glottal' vowels in K'ichee. In *Proceedings of formal approaches to Mayan linguistics*. Cambridge, MA: Massachusetts Institute of Technology Press.

Baken, R. J., & Orlikoff, R. F. (2000). *Clinical measurement of speech and voice*. San Diego, CA: Singular Publishing Group.

Basbøll, H. (2005). *The phonology of Danish*, Oxford: Oxford University Press.

Becker, K., Khan, S. D., & Zimman, L. (2017). Creaky voice beyond binary gender. Paper presented at: The Annual Meeting of the American Dialect Society, Austin.

Belotel-Grenié, A., & Grenié, M. (1994). Phonation types analysis in Standard Chinese. Paper presented at: Proceedings of the 3rd International Conference on Spoken Language Processing, Yokohama.

Belotel-Grenié, A., & Grenié, M. (2004). The creaky voice phonation and the organisation of Chinese discourse. Paper presented at: International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages. Beijing.

Berkson, K. (2013). Phonation Types in Marathi: An Acoustic Investigation. (PhD dissertation). The University of Kansas, Lawrence.

Berkson, K. H. (2016). Production, perception, and distribution of breathy sonorants in Marathi. *Proceedings of Formal Approaches to South Asian Languages*, *6*, 4–14.

Berkson, K. H. (2019). Acoustic correlates of breathy sonorants in Marathi. *Journal of Phonetics*, *73*, 70–90.
Bickley, C. (1982). Acoustic analysis and perception of breathy vowels. (Speech Communication Group Working Papers), Research Lab of Electronics, MIT, 73–93.
Bishop, J., & Keating, P. (2012). Perception of pitch location within a speaker's range: Fundamental frequency, voice quality and speaker sex. *Journal of the Acoustical Society of America*, *132*, 1100–1112.
Blankenship, B. (2002). The timing of non-modal phonation in vowels. *Journal of Phonetics*, *30*, 163–191.
Bradley, D. (1982). Register in Burmese. In D. Bradley (Ed.), *Tonation* (Pacific Linguistics Series A-62) (pp. 117–132). Canberra, Australia: Australian National University.
Brunelle, M. (2010). The role of larynx height in the Javanese tense–lax stop contrast. In R. Mercado, E. Potsdam, & L. D. Travis (Eds.), *Austronesian and theoretical linguistics* (pp. 7–24). Amsterdam, the Netherlands: John Benjamins.
Brunelle, M., & Finkeldey, J. (2011). Tone perception in Sgaw Karen. Paper presented at: Proceedings of the 17th International Congress of Phonetic Sciences, Hong Kong, 372–5.
Brunelle, M., Nguyễn, D. D., & Nguyễn, K. H. (2010). A laryngographic and laryngoscopic study of northern Vietnamese tones. *Phonetica*, *67*, 147–169.
Brunner, J. & Żygis, M. (2011). Why do glottal stops and low vowels like each other? Paper presented at: Proceedings of the International Congress of Phonetic Sciences, Hong Kong, 376–9.
Cao, J., & Maddieson, I. (1992). An exploration of phonation types in Wu dialects of Chinese. *Journal of Phonetics*, *20*, 77–92.
Childers, D. G., & Lee, C. K. (1991). Vocal quality factors: Analysis, synthesis, and perception. *Journal of the Acoustical Society of America*, *90*, 2394–2410.
Dave, R. V. (1967). A formant analysis of the clear, nasalized, and murmured vowels in Gujarati. *Indian Linguistics*, *28*, 1–30.
Davidson, L. (2019a). Perceptual coherence of creaky voice qualities. Paper presented at: Proceedings of the International Congress of Phonetic Sciences, Melbourne.
Davidson, L. (2019b). The effects of pitch, gender, and prosodic context on the identification of creaky voice. *Phonetica*, *76*, 235–262.
Denning, K. (1989). The diachronic development of phonological voice quality, with special reference to Dinka and the other Nilotic languages. (PhD dissertation). Stanford University.
DiCanio, C. T. (2009). The phonetics of register in Takhian thong Chong. *Journal of the International Phonetic Association*, *39*, 162–188.
DiCanio, C. T. (2012). Coarticulation between tone and glottal consonants in Itunyoso Trique. *Journal of Phonetics*, *40*, 162–176.
DiCanio, C. T. (2014). Cue weight in the perception of Trique glottal consonants. *Journal of the Acoustical Society of America*, *119*, 3059–3071.
Dilley, L., Shattuck-Hufnagel, S., & Ostendorf, M. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics*, *24*, 423–444.
Dixit, R. P. (1989). Glottal gestures in Hindi plosives. *Journal of Phonetics*, *17*, 213–237.
Edmondson, J. A., & Esling, J. H. (2006). The valves of the throat and their functioning in tone, vocal register and stress: Laryngoscopic case studies. *Phonology*, *23*, 157–191.
Epstein, M. (1999). A Comparison of Linguistic and Pathological Breathiness Using the LF Model. (MA thesis). UCLA.
Epstein, M. (2002). Voice Quality and Prosody in English. (PhD dissertation). UCLA.
Esling, J. H., & Harris, J. (2005). States of the glottis: An articulatory phonetic model based on laryngoscopic observations. In W. J. Hardcastle & J. Beck (Eds.), *A figure of speech: A festschrift for John Laver* (pp. 347–383). Mahwah, NJ: Lawrence Erlbaum.
Esposito, C. M. (2005). An acoustic and electroglottographic study of phonation in Santa Ana del Valle Zapotec. Paper presented at: The 79th Meeting of the Linguistic Society of America, San Francisco, California.
Esposito, C. M. (2006). The Effects of Linguistic Experience on the Perception of Phonation. (PhD dissertation). UCLA.
Esposito, C. M. (2010a). The effects of linguistic experience on the perception of phonation. *Journal of Phonetics*, *38*, 306–316.

Esposito, C. M. (2010b). Variation in contrastive phonation in Santa Ana del Valle Zapotec. *Journal of the International Phonetic Association*, *40*, 181–198.
Esposito, C. M. (2012). An acoustic and electroglottographic study of white Hmong tone and phonation. *Journal of Phonetics*, *40*, 466–476.
Esposito, C. M., & Khan, S. D. (2012). Contrastive breathiness across consonants and vowels: A comparative study of Gujarati and white Hmong. *Journal of the International Phonetic Association*, *42*, 123–143.
Esposito, C. M., Khan, S. D., Berkson, K. H., & Nelson, M. (2019). Distinguishing breathy consonants and vowels in Gujarati. *Journal of South Asian Languages and Linguistics*, *6*, 214–243.
Esposito, C. M., Sleeper, M., & Schafer, K. (2019). Examining the relationship between vowel quality and voice quality. *Journal of the International Phonetic Association*, 1–32. DOI: https://doi.org/10.1017/S0025100319000094.
Fischer-Jørgensen, E. (1967). Phonetic analysis of breathy (murmured) vowels in Gujarati. *Indian Linguistics*, *28*, 71–139.
Gao, J., & Hallé, P. (2017). Phonetic and phonological properties of tones in Shanghai Chinese. *Cahiers de Linguistique Asie Orientale*, *46*, 1–31.
Garellek, M. (2012). The timing and sequencing of coarticulated non-modal phonation in English and white Hmong. *Journal of Phonetics*, *40*, 152–161.
Garellek, M. (2019a). The phonetics of voice. In W. F. Katz & P. F. Assmann (Eds.), *The Routledge handbook of phonetics* (pp. 75–106). London: Routledge.
Garellek, M. (2019b). Acoustic discriminability of the complex phonation system in !Xóõ. *Phonetica*, *77*, 131–160. https://doi.org/10.1159/000494301
Garellek, M., & Keating, P. (2011). The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *Journal of the International Phonetic Association*, *41*, 185–205.
Garellek, M., Keating, P., Esposito, C. M., & Kreiman, J. (2013). Voice quality and tone identification in white Hmong. *Journal of the Acoustical Society of America*, *133*, 1078–1089.
Garellek, M., Ritchart, A., & Kuang, J. (2016). Breathy voice during nasality: A cross-linguistic study. *Journal of Phonetics*, *59*, 110–121.
Garellek, M., & Seyfarth, S. (2016). Acoustic differences between English /t/ glottalization and phrasal creak. Paper presented at: Proceedings of Interspeech 2016, 1054–8. San Francisco.
Gerfen, C., & Baker, K. (2005). The production and perception of laryngealized vowels in Coatzospan Mixtec. *Journal of Phonetics*, *33*, 311–334.
Gordon, M., & Ladefoged, P. (2001). Phonation types: A cross-linguistic overview. *Journal of Phonetics*, *29*, 383–406.
Grønnum, N. (1998). Danish. *Journal of the International Phonetic Association*, *28*, 99–105.
Grønnum, N. (2014). Laryngealization or pitch accent: The case of Danish stød. In Campbell, N., Gibbon, D., & Hirst, D. (eds.). Speech Prosody 7. Paper presented at: Proceedings of the 7th international conference on Speech Prosody: Social and Linguistic Speech Prosody, 804–8. Dublin.
Gruber, J. F. (2011). An articulatory, acoustic, and auditory study of Burmese tone. (PhD dissertation). Georgetown University.
Guion, S., Post, M., & Payne, D. (2004). Phonetic correlates of tongue root vowel contrasts in Maa. *Journal of Phonetics*, *32*, 517–542.
Hanson, H. M. (1995). Glottal characteristics of female speakers. (PhD dissertation). Harvard University.
Henderson, E. J. A. (1952). The main features of Cambodian pronunciation. *Bulletin of the School of Oriental and African Studies, University of London*, *14*, 149–174.
Hildebrandt, K. (2005). A phonetic analysis of Manange segmental and suprasegmental properties. *Linguistics of the Tibeto-Burman Area*, *28*, 1–36.
Holmberg, E. B., Hillman, R. E., Perkell, J. S., Guiod, P., & Goldman, S. L. (1995). Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech, Language, and Hearing Research*, *38*, 1212–1223.
Hombert, J. (1978). Consonant types, vowel quality and tone. In V. Fromkin (Ed.), *Tone: A linguistic survey* (pp. 77–112). New York, NY: Academic Press.
Iseli, M., Shue, Y.-L., & Alwan, A. (2007). Age, sex, and vowel dependencies of acoustical measures related to the voice source. *The Journal of the Acoustical Society of America*, *121*, 2283–2295.

Jessen, M., & Roux, J. (2002). Voice quality differences associated with stops and clicks in Xhosa. *Journal of Phonetics*, *30*, 1–52.

Kaun, A. R. (2004). The typology of rounding harmony. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically based phonology*, Cambridge: Cambridge University Press.

Keating, P., Esposito, C. M., Garellek, M., Khan, S. D., & Kuang, J. (2010). Phonation contrasts across languages. *UCLA Working Papers in Phonetics*, *108*, 188–202.

Keating, P., Esposito, C. M., Garellek, M., Khan, S. D., & Kuang, J. (2011). Phonation contrasts across languages. Paper presented at: Proceedings of the International Congress of Phonetic Sciences, Hong Kong. 1046–9.

Keating, P., Garellek, M., & Kreiman, J. (2015). Acoustic properties of different kinds of creaky voice. Paper presented at: Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow. 0821.1–0821.5.

Keating, P., Kuang, J., Esposito, C. M., Garellek, M. & Khan, S. D. (2012). Multi-dimensional phonetic space for phonation contrasts. Poster presented at: LabPhon13 in Stuttgart, Germany.

Keating, P., Mikoś, M. J., & Ganong, W. F. (1981). A cross-language study of range of voice onset time in the perception of initial stop voicing. *Journal of the Acoustical Society of America*, *70*, 1251–1271.

Khan, S. D. (2010). Bengali (Bangladeshi standard). *Journal of the International Phonetic Association*, *40*, 221–225.

Khan, S. D. (2012). The phonetics of contrastive phonation in Gujarati. *Journal of Phonetics*, *40*, 780–795.

Kingston, J., & Diehl, R. (1994). Phonetic knowledge. *Language*, *70*, 419–454.

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

Kirk, P. L., Ladefoged, J., & Ladefoged, P. (1993). Quantifying acoustic properties of modal, breathy and creaky vowels in Jalapa Mazatec. In A. Mattina & T. Montler (Eds.), *American Indian linguistics and ethnography in honor of Laurence C. Thompson*. Missoula, MT: University of Montana Press.

Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America*, *87*, 820–857.

Kreiman, J., & Gerratt, B. R. (1996). The perceptual structure of pathologic voice quality. *Journal of the Acoustical Society of America*, *100*(3), 1787–1795.

Kreiman, J., & Gerratt, B. R. (2005). Perception of aperiodicity in pathological voice. *Journal of the Acoustical Society of America*, *117*, 2201–2211.

Kreiman, J., & Gerratt, B. R. (2010). Perceptual sensitivity to first harmonic amplitude in the voice source. *Journal of the Acoustical Society of America*, *128*, 2085–2089.

Kreiman, J., Gerratt, B. R., & Antoñanzas-Barroso, N. (2007). Measures of the glottal source spectrum. *Journal of Speech, Language, and Hearing Research*, *50*, 595–610.

Kreiman, J., Gerratt, B. R., Garellek, M., Samlan, R., & Zhang, Z. (2014). Toward a unified theory of voice production and perception. *Loquens*, *1*, e009.

Kreiman, J., Gerratt, B. R., & Khan, S. D. (2010). Effects of native language on perception of voice quality. *Journal of Phonetics*, *38*, 588–593.

Kreiman, J., & Sidtis, D. (2011). *Foundations of voice studies*. Oxford, England: Wiley-Blackwell.

Kuang, J. (2011). Production and perception of the phonation contrast in Yi. (MA thesis). UCLA.

Kuang, J. (2013). The tonal space of contrastive five level tones. *Phonetica*, *70*, 1–23.

Kuang, J. (2017). Covariation between voice quality and pitch: Revisiting the case of mandarin creaky voice. *Journal of the Acoustical Society of America*, *142*, 1693–1706.

Kuang, J., & Cui, A. (2018). Relative cue weighting in production and perception of an ongoing sound change in southern Yi. *Journal of Phonetics*, *71*, 194–214.

Kuang, J., & Keating, P. (2014). Glottal articulations in tense vs. lax phonation contrasts. *Journal of the Acoustical Society of America*, *136*, 2784–2797.

Ladefoged, P. (1964). *A phonetic study of west African languages: An auditory-instrumental study*, Cambridge: Cambridge University Press.

Ladefoged, P. (1971). *Preliminaries to linguistic phonetics*. Chicago, IL: University of Chicago.

Ladefoged, P. (1983). The linguistic use of different phonation types. In D. Bless & J. Abbs (Eds.), *Vocal fold physiology: Contemporary research and clinical issues*. San Diego, CA: College-Hill Press.

Lancia, L., Avelino, H., & Voigt, D. (2013). Measuring laryngealization in running speech: Interaction with contrastive tones in Yalálag Zapotec. Paper presented at: Proceedings of Interspeech, Lyon 602–606.

Laver, J. (1981). *Phonetic description of voice quality*. Cambridge: Cambridge University Press.

Lindau, M. (1984). Phonetic differences in glottalic consonants. *Journal of Phonetics*, *12*, 147–155.

Lindsey, G., Hayward, K., & Haruna, A. (1992). Hausa glottalic consonants: A laryngographic study source. *Bulletin of the School of Oriental and African Studies*, *55*, 511–527.

Lotto, A., Holt, L., & Kluender, K. (1997). Effect of voice quality on perceived height of English vowels. *Phonetica*, *54*, 76–93.

(Luang)-Thongkum, T. (1986). An acoustic study of the register complex in Kui (Suai). *The Mon-Khmer Studies Journal*, *15*, 1–20.

(Luang)-Thongkum, T. (1987). Another look at the register distinction in Mon. *UCLA Working Papers in Phonetics*, *67*, 132–165.

Maddieson, I., & Ladefoged, P. (1985). Tense and lax in four minority languages of China. *Journal of Phonetics*, *13*, 433–454.

Mazaudon, M, & Michaud, A. (2006) Pitch and voice quality characteristics of the lexical word-tones of Tamang, as compared with level tones (Naxi data) and pitch-plus-voice-quality tones (Vietnamese data). Paper presented at: Proceedings of Speech Prosody, Dresden 823–6.

Mazaudon, M., & Michaud, A. (2008). Tonal contrasts and initial consonants: A case study of Tamang, a 'missing link' in tonogenesis. *Phonetica*, *65*, 231–256.

Michaud, A. (2004). Final consonants and glottalization: New perspectives from Hanoi Vietnamese. *Phonetica*, *61*, 119–146.

Mikuteit, S., & Reetz, H. (2007). Caught in the ACT: The timing of aspiration and voicing in east Bengali. *Language and Speech*, *50*, 247–277.

Miller, A. L. (2007). Guttural vowels and guttural co-articulation in Ju|'hoansi. *Journal of Phonetics*, *35*, 56–84.

Moisik, S. R. (2013). The Epilarynx in speech. (PhD thesis). University of Victoria.

Namboodiripad, S., & Garellek, M. (2016). Malayalam (Namboodiri dialect). *Journal of the International Phonetic Association*, *47*, 109–118.

Nguyễn, V. L., & Edmondson, J. (1998). Tones and voice quality in modern northern Vietnamese: Instrumental case studies. *Mon-Khmer Studies*, *28*, 1–18.

Pandit, P. B. (1957). Nasalization, aspiration and murmur in Gujarati. *Indian Linguistics*, *17*, 165–172.

Peng, J. G., & Zhu, X. N. (2010). Falsetto in Yueyang dialect (Chinese). *Journal of Contemporary Linguistics*, *1*, 24–32.

Pennington, M. (2005). The phonetics and phonology of glottal manner features. (Ph.D. dissertation). Indiana University.

Phạm, A. H. (2001). Vietnamese tone: Tone is not pitch. (PhD dissertation). University of Toronto.

Phạm, A. H. (2003). *Vietnamese tone: A new analysis*. New York, NY: Routledge.

Pinkerton, S. (1986). Quichean (Mayan) glottalized and nonglottalized stops: A phonetic study with implications for phonological universals. In J. J. Ohala & J. J. Jaeger (Eds.), *Experimental phonology* (pp. 125–139). Orlando, FL: Academic Press.

Podesva, R. J. (2007). Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of SocioLinguistics*, *11*, 478–504.

Podesva, R. J. (2013). Gender and the social meaning of non-modal phonation types. *Annual Meeting of the Berkeley Linguistics Society*, *37*, 427–448.

Redi, L., & Shattuck-Hufnagel, S. (2001). Variation in the realization of glottalization in normal speakers. *Journal of Phonetics*, *29*, 407–429.

Remijsen, B., & Manyang, C. A. (2009). Luanyjang Dinka. *Journal of the International Phonetic Association*, *39*, 113–124.

Ren, N. (1992). Phonation types and stop consonant distinctions: Shanghai Chinese. (PhD dissertation). University of Connecticut.

Rose, P. (1997). A seven-tone dialect in southern Thai with super-high: Pakphanang tonal acoustics and physiological inferences. In A. S. Abramson (Ed.), *Southeast Asian linguistic studies in honour of Vichin Panupong* (pp. 191–208). Bangkok, Thailand: Chulalongkorn University Press.

Rothenberg, M., & Mahshie, J. J. (1988). Monitoring vocal fold abduction through vocal fold contact area. *Journal of Speech and Hearing Research*, *31*, 338–351.

Rycroft, D. K. (1980). The depression feature in Nguni languages and its interaction with tone. Communication No. 8, Department of African Languages, Rhodes University, Grahamstown.

Samely, U. (1991). Kedang (Eastern Indonesia): Some Aspects of its Grammar. Forum phoneticum.

Sapienza, C., Hicks, D. M., & Ruddy, B. H. (2011). Voice disorders. In N. B. Anderson & G. H. Shames (Eds.), *Human communication disorders: An introduction* (8th ed., pp. 202–237). Boston, MA: Pearson.

Shue, Y.-L., Keating, P., Vicenik, C., & Yu, K. M. (2011). VoiceSauce: A program for voice analysis. Paper presented at: Proceedings of the 17th International Congress of Phonetic Sciences, Hong Kong. 1846–9.

Sidwell, P., & Rau, F. (2015). Austroasiatic comparative-historical reconstruction: An overview. In M. Jenny & P. Sidwell (Eds.), *The handbook of Austroasiatic languages*. Boston, MA: Brill.

Silverman, D. (1997). Laryngeal complexity in Otomanguean vowels. *Phonology*, *14*, 235–262.

Silverman, D., Blankenship, B., Kirk, P., & Ladefoged, P. (1995). Phonetic structures of Jalapa Mazatec. *Anthropological Linguistics*, *37*, 70–88.

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

Svantesson, J.-O., & House, D. (2006). Tone production, tone perception, and Kammu tonogenesis. *Phonology*, *23*, 309–333.

Szakay, A. (2012). Voice quality as a marker of ethnicity in New Zealand: From acoustics to perception. *Journal of SocioLinguistics*, *16*, 382–397.

Thurgood, E. (2004). Phonation types in Javanese. *Oceanic Linguistics*, *43*, 277–295.

Thurgood, G. (2002). Vietnamese and tonogenesis. *Diachronica*, *19*, 333–363.

Tian, J., & Kuang, J. (2019). The phonetic properties of the non-modal phonation in Shanghainese. *Journal of the International Phonetic Association*, 1–27. https://doi.org/10.1017/S0025100319000148

Tian, J., Zhou, Y., & Kuang, J. (2019). Cross-linguistic variation in the phonetic realization of breathy voice. Paper presented at: Proceedings of 19th International Congress of Phonetic Sciences, Melbourne.

Traill, A. (1985). *Phonetic and phonological studies of !Xóõ bushman*. Hamburg, Germany: Buske.

Traill, A. (1986). The laryngeal sphincter as a phonatory mechanism in !Xóõ bushman. In R. Singer & J. K. Lundy (Eds.), *Variation, culture and evolution in African populations: Papers in honor of Dr. Hertha de Villiers* (pp. 123–131). Johannesburg, South Africa: Witwatersrand University Press.

Traill, A., & Jackson, M. (1988). Speaker variation and phonation type in Tsonga. *Journal of Phonetics*, *16*, 385–400.

Traill, A., Khumalo, J. S. M., & Fridjon, P. (1987). Depressing facts about Zulu. *African Studies*, *46*, 255–274.

Uchihara, H. (2016). Tone and registrogenesis in Quiaviní Zapotec. *Diachronica*, *33*, 220–254.

Watkins, J. (2002). *The phonetics of Wa: Experimental phonetics, phonology, orthography and sociolinguistics*. Canberra, Australia: Pacific Linguistics.

Wayland, R., Gargash, S., & Jongman, A. (1994). Acoustic and perceptual investigation of breathy voice. *Journal of the Acoustical Society of America*, *97*, 3364.

Wayland, R., & Jongman, A. (2002). Registrogenesis in Khmer. *Mon-Khmer Studies*, *32*, 101–114.

Wayland, R., & Jongman, A. (2003). Acoustic correlates of breathy and clear vowels: The case of Khmer. *Journal of Phonetics*, *31*, 181–201.

Yang, R.-X. (2011). The phonation factor in the categorical perception of mandarin tones. Paper presented at: Proceedings of the International Congress of Phonetic Sciences, Hong Kong. 2204–7.

Yu, K. M., & Lam, H. W. (2014). The role of creaky voice in Cantonese tonal perception. *Journal of the Acoustical Society of America*, *136*, 1320–1333.

Yuasa, I. P. (2010). Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech*, *85*, 315–337.

Zee, E., & Xu, L. (2017). Shanghainese. In G. Thurgood & R. J. LaPolla (Eds.), *The Sino-Tibetan languages* (2nd ed.,), London: Routledge.

Zhang, J., & Yan, H. (2018). Contextually dependent cue realization and cue weighting for a laryngeal contrast in Shanghai Wu. *Journal of the Acoustical Society of America*, *144*, 1293–1308.

Zimman, L. (2018). Transgender voices: Insights on identity, embodiment, and the gender of the voice. *Language and Linguistics Compass*, *12*, e12284.

AUTHOR BIOGRAPHIES

Christina M. Esposito is an associate professor of Linguistics at Macalester College. She researches the production and perception of phonation in various languages, such as White Hmong, Zapotec, and Gujarati.

Sameer ud Dowla Khan is an associate professor of Linguistics at Reed College. His research interests include intonation, phonation, and reduplication across various languages, with a focus on South Asia.

How to cite this article: Esposito CM, Khan SD. The cross-linguistic patterns of phonation types. *Lang Linguist Compass*. 2020;e12392. https://doi.org/10.1111/lnc3.12392