# Examining the relationship between vowel quality and voice quality

**Christina M. Esposito**

Macalester College
*esposito@macalester.edu*

**Morgan Sleeper**

Macalester College
*msleeper@macalester.edu*

**Kevin Schäfer**

University of California, Santa Barbara
*kschaefer@umail.ucsb.edu*

The majority of studies on phonation types have focused on low vowels due to the minimal effects of their first formant on harmonic amplitude. In studies of multiple vowel qualities, reports on the relationship between vowel and voice quality are mixed: some show similar formant frequencies across phonation types (e.g. Abramson, Nye & Luangthongkum 2007, Khan 2012), while others show different formant frequencies depending on voice quality (e.g. Ren 1992, Kuang 2011). Results differ as to whether the degree of non-modal phonation varies (Andruski & Ratliff 2000, Kuang 2011) or does not vary (Esposito 2012, Khan 2012) across different vowel qualities. The present study draws on innovations which allow for more accurate corrections for the effects of formant frequencies on spectral measures (i.e. Hanson 1995, Iseli, Shue & Alwan 2007) to examine the relationship between vowel quality and voice quality, in eight languages – !Xóõ, Burmese, Gujarati, Jalapa de Díaz Mazatec, Mon, Santa Ana del Valle Zapotec, White Hmong, and Yi. While no significant difference in the degree of non-modal phonation due to vowel quality was found, results showed a cross-linguistic pattern in the relationship between vowel quality and voice quality: vowels with higher $log$(F1) and $log$(F2) values tended to be produced with creakier phonation, while vowels with lower $log$(F1) and $log$(F2) values tended to be produced with breathier phonation, but only on the measure H1*-H2*.

## 1 Introduction

The majority of studies on linguistically-relevant phonation type differences focus solely on low vowels, such as /a/ or /ɑ/, due to the minimal effects of their first formant on harmonic amplitude.[1] A few studies explore phonation types in wider vocalic inventories, though they report mixed results regarding the relationship between vowel quality and voice quality.

---

[1] The acoustic differences between phonation types can be quantified by measuring the difference between different harmonic amplitudes, but it is important that they either be unaffected by, or corrected for, influences from formant frequencies for the measurements to be accurate. It is fairly recent that correction formulas have been used in the study of linguistic voice qualities; these studies are indicated with an asterisk * after the year of publication.

Similar first formant frequencies (F1) are reported for breathy and modal[2] vowels in Mon (Thongkum 1987a), Nyah Kur (Thongkum 1987b), Kui (Thongkum 1986), Khmu' Rawk (Abramson, Nye & Luangthongkum 2007) and Gujarati (Khan 2012*) and for modal, breathy and creaky phonation in Jalapa Mazatec[3] (Garellek & Keating 2011*). However, when compared to modally-phonated vowels, a lower first formant frequency is found for: (i) breathy vowels in Kedang (Samely 1991), Nilotic languages (Denning 1989), Southeast Asian languages (Hombert 1978, Bradley 1982) and Xhosa (Jessen & Roux 2002*); (ii) breathy and breathy-creaky vowels in Chong (Thongkum 1987b); (iii) slack voice in Shanghainese (Ren 1992); and (iv) lax phonation in Yi (Kuang 2011*). Similarly, higher first formant frequency is reported for creaky/tense phonation in Hani (Maddieson & Ladefoged 1985) and Mpi (Denning 1989), and higher first and second formant (F2) frequencies are found for the strident[4] vowels in !Xóõ (Traill 1985).

An additional issue regarding vowel quality and phonation type is the degree of non-modal phonation for a particular vowel quality. In Green Mong (Andruski & Ratliff 2000), /a̤/ is reported to be breathier than /i̤/ and /ṳ/. And, phonation type is produced differently for different vowel qualities: the difference between the amplitude of the first and second harmonics (H1-H2) is a good indicator of the phonation difference for /i/ and /u/, but not for /a/. Yi (Kuang 2011*) also displays variation in phonation type due to vowel quality; phonation differences are smaller in low/open vowels than high/close ones. However, 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.

To summarize, there are mixed results regarding the interaction between formant frequencies and phonation type, and there is little knowledge regarding the degree of non-modal phonation for different vowel types. Due to progress in correcting for the effects of formant frequency and bandwidths on spectral measures (see Hanson 1995, 1997; Hanson & Chuang 1999; Iseli, Shue & Alwan 2007), it is now possible to focus our attention on these issues. Our goal is to investigate the relationship between formant frequency and voice quality, using data from eight unrelated languages.

## 2 Background

We will begin by providing a review of the acoustic properties associated with the production of voice quality. Following that, we present a summary of the languages studied, concentrating on vowel inventory and phonation type.

### 2.1 Acoustic properties of phonation contrasts

Various spectral measures are used to quantify phonation type differences; these include, but are not limited to, the difference between the amplitude of the first harmonic and second harmonic (i.e. H1-H2), and the difference between the first harmonic (H1) and that of harmonics exciting higher formants (e.g. H1-A1, H1-A2, and H1-A3). Values for these measures are highest for breathy phonation, smallest for creaky, and moderate for modal phonation. H1-H2, a measure of spectral balance, is one of the most common ways to measure phonation types. H1-H2 is an accurate measure of phonemic phonation types in !Xóõ (Bickley

---

[2] Modal phonation is referred to as 'clear' in Mon and other Mon-Khmer languages such as Nyah Kur, Kui, and Khmu' Rawk.

[3] Kirk et al. (1993) did find that the frequency of F1 was higher for laryngealized vowels than for their non-laryngealized counterparts, but for a much smaller set of data than examined in Garellek & Keating (2011).

[4] Strident voice quality is produced by pharyngeal constriction and vibration of the aryepiglottic folds, which affect the vibration of the true vocal folds (Traill 1985).

1982), Jalapa de Díaz Mazatec (Blankenship 2002, Garellek & Keating 2011*, Keating et al. 2010*), Chanthaburi Khmer (Wayland & Jongman 2002), Green Mong (Andruski & Ratliff 2000), Mon (Thongkum 1987a), Takhian Thong Chong (DiCanio 2009), Santa Ana del Valle Zapotec (Esposito 2010b), White Hmong (Keating et al. 2010*, Esposito 2012), Gujarati (Keating et al. 2010*, Khan 2012*), Southern Yi (Keating et al. 2010*, Kuang 2011*), and for a small set of data from Chong, Fuzhou, San Lucas Quiaviní Zapotec, Tlacolula Zapotec, and Tamang (Esposito 2006). With the exception of male speakers of Santa Ana del Valle Zapotec (Esposito 2010b), the authors are not aware of any language where H1-H2 is not a successful measure of voice quality. Fewer studies use spectral tilt measures such as H1-A1, H1-A2, and H1-A3 to successfully quantify the differences between phonemic phonation types. Example studies include: Jalapa de Díaz Mazatec (Blankenship 2002); for male speakers of Santa Ana del Valle Zapotec (Esposito 2010b), Gujarati (Khan 2012*) and Yi (Kuang 2011*).

## 2.2 About the languages

### 2.2.1 !Xóõ
!Xóõ is a Khoisan language spoken by around 4,000 people in southwestern Botswana, and by several hundred in Namibia (Lewis, Simons & Fennig 2016). It has five vowel qualities [i e ɑ u o], four lexical tones (high, mid, low, and mid-falling) and five phonation types: modal, breathy, creaky, strident, and pharyngealized. Any of these phonation types can occur with any tone. Phonation types in !Xóõ are distinguished by increased spectral noise, H1-H2, and/or other measurements of spectral tilt, depending on the speaker (Ladefoged & Antoñanzas-Barroso 1985, Gordon & Ladefoged 2001).

### 2.2.2 Burmese
Burmese is a Sino-Tibetan language of the Tibeto-Burman branch spoken by over 30 million people in Myanmar (Lewis et al. 2016). It contrasts four tones: low, high, creaky, and 'killed' (Watkins 2001). The Burmese vowel inventory is analyzed in terms of two partially-overlapping sets of vowel qualities which interact with tone: Set 1 vowels [i e ɛ a u o ɔ ə] occur in orthographically open syllables (low, high, or creaky tones), and Set 2 vowels [ɪ eɪ ɛ aɪ a ʊ oʊ aʊ] in orthographically closed syllables (killed tone, or low, high, or creaky tones with nasalization) (Watkins 2000, 2001).

The low tone is characterized by low intensity and moderate length and the high tone by high intensity, increased duration, and, for some speakers, breathiness. The creaky tone is a falling tone characterized by very high intensity, shorter duration, and higher and more fronted vowel quality. The killed tone is a falling tone with a syllable-final glottal stop, marked by high intensity and the shortest length (Bradley 1982). In terms of their phonetic properties, creaky and killed tones are very similar, though they can be distinguished by the gemination of consonants following killed tones (where the final glottal stop assimilates in place and manner to the initial consonant of the next syllable) and by the distribution of vowel qualities over the two tones – only /a/ and /ɛ/ occur in both creaky and killed tones (see Watkins 2001). Creakiness in Burmese is distinguished by higher jitter values – a measure of aperiodicity in the glottal source – compared to modal phonation (Javkin & Maddieson 1985).

### 2.2.3 Gujarati
Gujarati is an Indic language spoken by 45 million people in India, primarily in the state of Gujarat but also in Maharashtra, Rajasthan, Karnataka, Madhya Pradesh, as well as by sizable communities in the United States, Canada, and the United Kingdom (Lewis et al.

2016). Gujarati vowels [i e ɛ a ɔ o u ə] may be produced with either modal or breathy phonation. Breathy vowels are characterized by greater airflow and a higher H1-H2/H1*-H2* value when compared to modal vowels (Fischer-Jørgensen 1967, Esposito 2006, Khan 2012*), though other acoustic measures (i.e. H2*-H4*, H1*-A1*, H1*-A2*, H1*-A3*, cepstral peak prominence, and harmonics-to-noise ratio), are also reliable indicators of phonation type differences (Khan 2012*).

### 2.2.4   Jalapa de Díaz Mazatec

Jalapa de Díaz Mazatec (hereafter Jalapa Mazatec) is an Otomanguean language spoken by approximately 17,000 people in Northern Oaxaca and Veracruz, Mexico (Lewis et al. 2016). It contrasts five vowels [i æ u o a], each of which can be lexically oral or nasal, three tones (low, mid, and high), and modal, breathy, and creaky phonation (Kirk, Ladefoged & Ladefoged 1993, Silverman et al. 1995, Garellek & Keating 2011*). Vowel qualities, tones, and voice qualities are fully cross-classified (Silverman et al. 1995, Garellek & Keating 2011*). Kirk et al. (1993) show that creaky phonation is associated with higher F1 values than either modal or breathy phonation. H1-H2/H1*-H2* is a successful measure of phonation in Jalapa Mazatec (Blankenship 2002, Keating et al. 2010*, Garellek & Keating 2011*).

### 2.2.5   Mon

Mon is a Mon-Khmer language spoken by 800,000 people, primarily in southern Myanmar but also in Thailand (Lewis et al. 2016). It contrasts nine vowels [i e ɛ ə a ɑ u o ɔ] along with two registers: a high-level tone with modal phonation, and a low-level tone with breathy phonation. All vowel qualities can occur with both registers, with the exception of [ɔ ɑ] and the diphthongs [ao ɑe], which only occur in the first register (Huffman 1990). There is some dispute over the distinctive phonetic features of these registers, however. Lee (1983) finds the difference to be one of duration and pitch, not phonation. Thongkum (1987b, 1990), however, concludes that phonation type does play a part in the Mon register system; while most speakers use modal phonation for the first register and breathy phonation for the second, some speakers contrast degrees of breathiness (quantified by H1-H2 differences), with the second register more breathy than the first, while others use creaky phonation for the first register and breathy for the second. A more recent study, Abramson, Tiede & Luangthongkum (2015*), finds that modal and breathy vowels are best distinguished by H1*-H2*, H1*-A2*, H1*-A3* and closed quotient differences (as measured by an electroglottograph).

### 2.2.6   Santa Ana del Valle Zapotec

Santa Ana del Valle Zapotec is an Otomanguean language spoken in Santa Ana del Valle, Oaxaca, Mexico, classified by The Ethnologue (Lewis et al. 2016) as part of the San Juan Guelavía Zapotec subgroup, which has approximately 28,000 speakers; the exact number of Santa Ana del Valle Zapotec speakers is unknown. Santa Ana del Valle Zapotec has six vowels [i e ɨ a u o], each of which can have one of three contrastive phonation types: modal, breathy, or creaky. Tone is contrastive on modal vowels, which may have a high or a rising tone. Breathy and creaky vowels are both produced with a falling tone (Esposito 2010b). Breathy and creaky vowels are best distinguished by H1-A3 for males and H1-H2 for females (Esposito 2010b).

### 2.2.7   White Hmong

White Hmong is a Hmongic language spoken by approximately two million people in Laos, Thailand, Vietnam and southwestern China, and by communities in the United States

(Lewis et al. 2016). It contrasts [i e a ɔ u ɨ ẽ ã ɔ̃] and has seven lexical tones. The contrast between modal, breathy, and creaky phonation plays a part in the tonal system: the rising (45), mid (33), low (22), mid-rising (24), and high-falling (52) tones are modal, while the low-falling (21) tone is creaky and the mid-to-high falling (42) tone is breathy (Smalley 1976, Ratliff 1992, Esposito 2012*).

Phonation type in White Hmong is distinguished by spectral measures (specifically, the amplitudes of the first harmonic (H1*) and H1*-H2*), measurements of periodicity (cepstral peak prominence), and electroglottographic measures (closed quotient and Derivative-EGG Closure Peak Amplitude) (Esposito 2012*).

### 2.2.8 Yi

Yi, a Tibeto-Burman language, is spoken in the Yunnan, Sichuan and Guizhou provinces of China by approximately two million speakers (Lewis et al. 2016). Yi contains seven vowels [i ɛ a z ə o u], where [z] represents a vowel which has assimilated to the preceding consonant. (While not part of the International Phonetic Alphabet, the symbol has been widely used by Chinese phoneticians – Jianjing Kuang pc, 14 November 2016). In addition, Yi contrasts three tones (high, mid, and low); the mid and low tones are produced with either a lax or tense voice quality. Phonation type contrasts are best distinguished by contact quotient (as measured by an electroglottograph), H1*-H2* and H1*-A1* (Kuang 2011*).

## 3 Method

### 3.1 Speakers and speech materials

All data used in this study came from one of three sources: the UCLA Phonetics Lab Archive, an online collection of fieldwork recordings of over 200 languages (http://archive.phonetics.ucla.edu/), the Production and Perception of Linguistic Voice Quality project (http://www.phonetics.ucla.edu/voiceproject/voice.html), or, in the case of Santa Ana del Valle Zapotec, recorded by one of the authors as part of a previous study on phonation type. From the two electronic databases, languages with non-modal phonation types on vowels were selected if (1) data from more than one native speaker were accessible, and (2) if those data were transcribed to indicate voice quality. (This last criterion excluded some languages for which data was available but transcription did not indicate phonation type.) Individual tokens were selected for analysis as part of minimal or near-minimal sets contrasting phonation types. All tokens that fit this criterion were examined. The number of tokens varied across languages; wordlists are available in the appendix. Transcriptions were taken from the source material, unless otherwise noted. Tone and coda consonant were not controlled, but tokens with aspirated onset consonants were excluded because of their tendency to alter voice quality of the following vowel (Garellek & Keating 2011*, Esposito & Khan 2012*). Data were either digitally recorded or digitized as part of the archival process (with a 44.1 kHz sampling rate). Tokens were extracted from longer recordings in Audacity (Audacity Team 2014).

### 3.2 Summary of language data

Table 1 summarizes the relevant phonetic information (i.e. the acoustic measures of voice quality) presented in Section 2.2 and provides information on the recordings used for each

**Table 1** Table of relevant language data, including both phonetic and recording information.

| Language | Recorded by | Initial recording medium | Location and date of recording | Number and gender of speakers | Vowels (examined in current study) | Voice qualities (examined in current study) | Acoustic measures of phonation (reported in previous research) |
|---|---|---|---|---|---|---|---|
| !Xóõ | P. Ladefoged & T. Traill | Reel-to-reel tape | Botswana 1979 | 42 speakers, all male | [ɑ] | Modal, breathy, creaky | Spectral noise; H1–H2 |
| Burmese | Two recordings: 1. P. Ladefoged; 2. Unknown | 1. Reel-to-reel tape; 2. Cassette tape | 1. Pune, India 1992 2. Soundproof booth at the UCLA Phonetics Laboratory 1983 | One male, one female | [i e ɛ a u o ɔ ə][ɪ eɪ ɛ aɪ a ʊ oʊ] | Modal, creaky | Jitter |
| Jalapa Mazatec | P. Ladefoged & P. Kirk | Reel-to-reel tape | Jalapa de Díaz, Oaxaca, Mexico 1980s | Two male, two female | [i æ u o a] | Modal, breathy, creaky | H1–H2/H1*–H2* |
| Gujarati | S. Khan | Digital | Soundproof booth at the UCLA Phonetics Lab 2008-2009 | Three male, seven female | [i e ɛ a ɔ o u ə] | Modal, breathy | H1*–H2*, H2*–H4*, H1*–A1*, H1*–A2*, H1*–A3*, cepstral peak prominence, harmonics-to-noise ratio |
| Mon | P. Ladefoged | Reel-to-reel tape | Multiple recordings locations: unknown; Thailand 1981–1986 | Five male speakers | [i e ɛ ə a u o ie eə ai oi oa ua] | Modal, breathy | H*1–H2* H1*–A3* H1*–A2* |
| Santa Ana del Valle Zapotec | C. Esposito | Digital | Soundproof booth at the UCLA Phonetics Lab; other locations around Los Angeles, California 2003-2005 | Two female speakers | [i e a uo] | Modal, breathy, creaky | H1–H2 and H1–A3 |
| White Hmong | C.Esposito, J. Ptacek & S. Yang | Digital | Hmong American Partnership (St. Paul, Minnesota) 2008 | 11 female, 21 male | [i a ɔ u] | Modal, breathy, creaky | H1* and H1*–H2* |
| Yi | J. Kuang | Digital | Xinping and Jiangcheng, China 2009 | Three female, three male | [i ɛ a z ə o u] | Modal, lax, tense | H1*–H2* and H1*–A1* |

language, including number of speakers and vowels available[5] for study from the audio sources. We do not have access to specific details about the nature of the recordings, such as microphone type, amplifiers and/or filters applied.

### 3.3 Procedure

For each word, the vowel was labeled for vowel quality and phonation type in Praat text grids (Boersma & Weenink 2008) according to the transcriptions in the original data, unless otherwise noted in the appendix. The frequency of the first (F1) and second (F2) formants, a measure of spectral balance, H1*-H2*, and a measure of spectral tilt, H1*-A3*, were measured automatically using VoiceSauce (Iseli et al. 2007, Shue et al. 2011). (More details on how measurements are calculated in VoiceSauce can be found at http://www.phonetics.ucla.edu/voiceproject/Publications/Shue-etal_2009_ASA-poster.pdf; Shue, Keating & Vicenik 2009). Spectral measures were corrected for surrounding formant frequencies and bandwidths (Hanson 1995, Iseli et al. 2007); corrected harmonic amplitudes are indicated by an asterisk (e.g. H1*-H2*). (See Iseli et al. 2007 for the spectral magnitude correction formula.) Formant frequencies and their amplitudes were calculated using the Snack Sound Toolkit (Sjölander 2004). Fundamental frequency (f0) was calculated by the STRAIGHT algorithm (Kawahara, de Cheveigné & Patterson 1998).

Each acoustic measure was calculated at every millisecond of the target vowel, and then averaged across three timepoints. Analysis was conducted on the timepoints where phonation contrasts were localized in a particular language based on previous research and/or examination of the spectrogram. Care was taken to avoid consonant transitions at the beginning and end of the vowel. For Burmese and Mon, a few diphthongs were included in the analysis because they were available in the recordings. For these diphthongs, only the onglide vowel, which is where the phonation was localized, was measured.

# 4 Results

### 4.1 Cross-language results

Two separate linear mixed-effects models were fit to the data using R package *lme4* (Bates et al. 2015) with F1, F2, and phonemic phonation type (modal, breathy, or creaky) as fixed effects and either H1*-H2* or H1*-A3* as the dependent variable. Speaker and language were nested random effects. The *log* function was used to rescale F1 and F2 measures; hereafter, referred to as *log*(F1) and *log*(F2). Results for each measure (H1*-H2* and H1*-A3*) will be discussed separately. (Sex of speaker was not included as a fixed effect. For White Hmong, Yi, Jalapa Mazatec, and Gujarati, previous research (Keating et al. 2010*) shows that there are no significant interactions of phonation with gender for these languages. The Burmese data included only one male and one female speaker, making it difficult to draw a generalization regarding gender differences. The remaining languages in our study – !Xóõ, Mon, and Santa Ana del Valle Zapotec – included data from only one gender.)

---

[5] There were gaps in the recordings, such that not all vowels for a particular language were available for analysis. In Table 1, we list the vowels that we analyzed in the current study. Diphthongs are included when present in the data. The recordings for !Xóõ contained mostly [α], with very few examples of phonation distinctions on other vowels. For this reason, and because we tried to use minimal sets whenever possible, the data here is limited to the vowel [ɑ]. The [ɑɑ] in the transcribed wordlist indicates a long vowel.

### 4.1.1 H1*-H2*

Using the *MuMIn* package (Bartoń 2015), the model showed an adjusted $R^2$ value of .4275 (adjusted $R^2 = .1313$ for fixed effects alone). Random effects (language and speaker) thus predictably accounted for most, but by no means all of the variance (language: std. dev. = 2.132, speaker: std. dev. = 2.800). There was a significant interaction of all fixed effects ($p < .05$, df = 4, $\chi^2 = 11.93$). The model also contained significant main effects of $log$(F1) ($p < .001$, df = 1, $\chi^2 = 737.09$), $log$(F2) ($p = .025$, df = 1, $\chi^2 = 5.00$), and phonation type ($p < .001$, df = 5, $\chi^2 = 1872.61$) on H1*-H2*.

### 4.1.2 H1*-H2*: *log*(F1)

There was a significant main effect of $log$(F1) on H1*-H2* ($p < .001$), and a significant interaction with phonation type ($p < .001$). The effect of $log$(F1) on H1*-H2* is graphed in Figure 1. (For all figures, the labels 'breathy', 'modal' and 'creaky' represent phonemic phonation type, while H1*-H2* values represent phonetic phonation. Thus, what is labeled 'breathy' in Figure 1 refers to the phonemic phonation category, not the H1*-H2* value.) For all three phonation types, vowels with higher $log$(F1) frequencies (i.e. low/open vowels) had lower H1*-H2* values, suggesting a creakier phonation than vowels with lower $log$(F1) values (i.e. high/closed vowels).
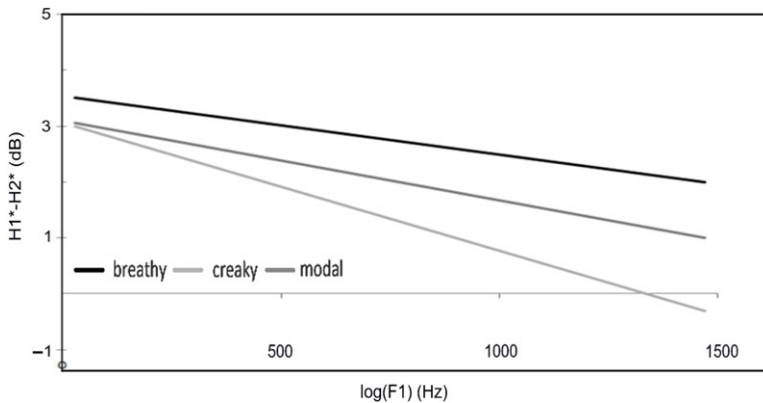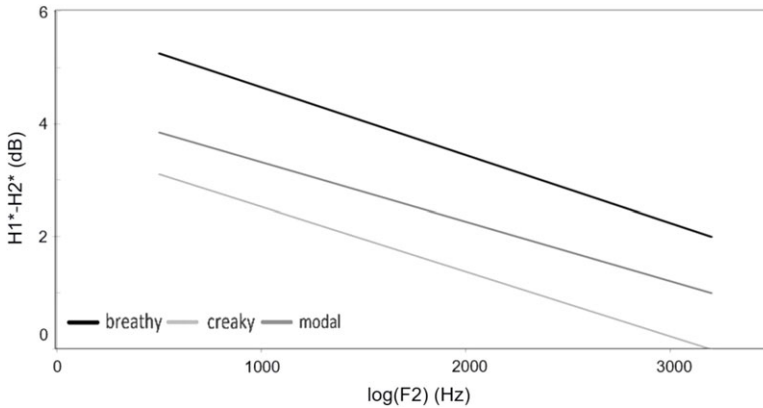


**Figure 1** Graph of the effect of the first formant (*log* (F1), Hz) on H1*-H2* (dB) for three phonation types (breathy (solid black line), creaky (light gray line) and modal (dark gray line)).

### 4.1.3 H1*-H2*: *log*(F2)

There was a significant main effect of $log$(F2) on H1*-H2* ($p = .02$)[6] and a significant interaction with phonation type ($p < .001$). The effect of $log$(F2) on H1*-H2* is graphed in Figure 2. For all three phonation types, there was a positive relationship between $log$(F2) and spectral tilt: vowels with greater $log$(F2) frequencies (i.e. front vowels) had lower spectral tilt values, suggesting a creakier phonation type for front vowels than their non-front counterparts.

---

[6] A significant interaction of both $log$(F1)1* and $log$(F2)2* on H1*-H2* ($p < .001$) suggests that the correlations of $log$(F1) and $log$(F2) with H1*-H2* are not additive.

**Figure 2** Graph of the effect of the second formant (*log*(F2), Hz) on H1*-H2* (dB) for three phonation types (breathy (solid black), creaky (light gray) and modal (dark gray)). Note: *log*(F2) frequency increases from left to right.

### 4.1.4  H1*-A3*

For H1*-A3* the model showed an adjusted $R^2$ value of .6042 (adjusted $R^2$ = .0761 for fixed effects alone). Random effects (speaker and language) thus predictably accounted for most, but by no means all of the variance (speaker: std. dev. = 2.133). There were highly significant interactions of F1:F2 ($p$ < .001, df = 1, $\chi^2$ = 97.75) and F2:Phonation ($p$ < .001, df = 2, $\chi^2$ = 14.39), and a marginally significant interaction of F1:Phonation ($p$ < .05, df = 2, $\chi^2$ = 9.17). The model also contained significant main effects of H1*-A3* for formant values, with only a marginally significant main effect of phonemic phonation type: (F1) ($p$ < .001, df = 1, $\chi^2$ = 106.19), *log*(F2) ($p$ = .001, df = 1, $\chi^2$ = 90.30), and marginally significant phonation type ($p$ < .05, df = 2, $\chi^2$ = 8.50).

### 4.1.5  H1*-A3*: *log*(F1)

There was a significant main effect of *log*(F1) on H1*-A3* ($p$ < .001). *log*(F1) was positively correlated with H1*-A3* ($\beta$ = 133.26), such that low *log*(F1) values (i.e. high/closed vowels) tended to have lower spectral tilt values. However, given significant interactions of *log*(F1) with *log*(F2), the effect size for *log*(F1) was effectively only sufficiently large when *log*(F2) values were at the low end of the range ($\beta$ < 23.79, $\beta$ < 27.29, and $\beta$ < 17.24 for breathy, creaky, and modal phonation, respectively[7]). When *log*(F2) was highest, there was essentially no effect of *log*(F1) ($\beta$ < −2.85, $\beta$ < −4.14, and $\beta$ < −2.73 for breathy, creaky, and modal phonation, respectively[8]). The effect of *log*(F1) on H1*-A3* when *log*(F2) was low and when *log*(F2) was high are graphed in Figures 3 and 4, respectively. For all three phonation types, vowels with lower *log*(F1) frequencies (i.e. high/close vowels) had lower H1*-A3* values, suggesting a creakier phonation than vowels with higher *log*(F1) values (i.e. low/open vowels), though only when *log*(F2) values were low.

### 4.1.6  H1*-A3*: *log*(F2)

The significant ($p$ < .001) main effect of *log*(F2) showed a positive relationship with H1*-A3* ($\beta$ = 103.66); however, considering the range of the data and the interactions in

---

[7] Values when *log*(F2) = 770.6 Hz.
[8] Values when *log*(F2) = 3886.0 Hz.
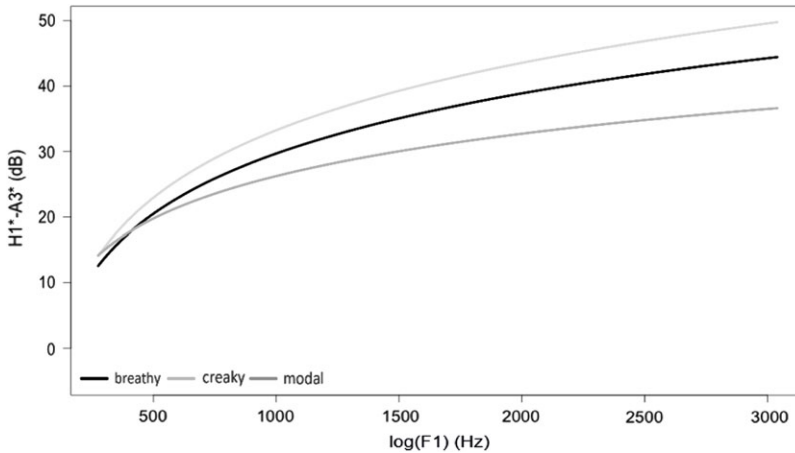
**Figure 3** Graph of the effect of the first formant (*log*(F1), Hz) on H1*-A3* (dB) for three phonation types (breathy (solid black line), creaky (light gray line) and modal (dark gray line)) when *log*(F2) values are low.
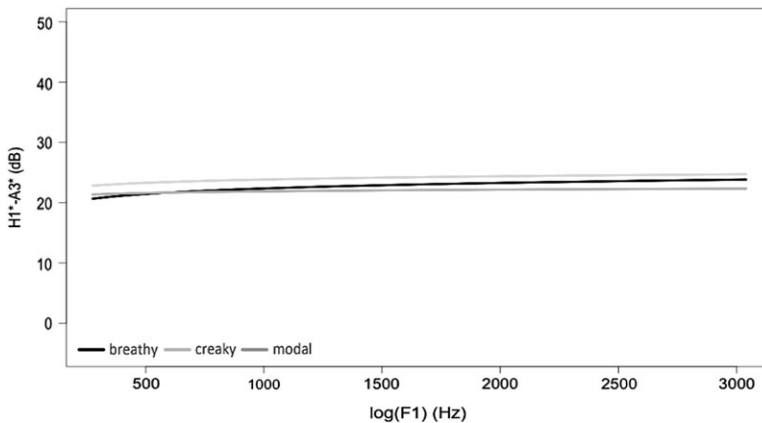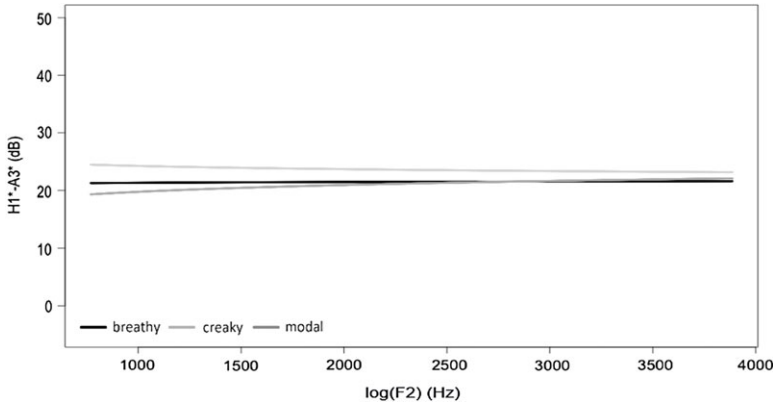


**Figure 4** Graph of the effect of the first formant (*log*(F1), Hz) on H1*-A3* (dB) for three phonation types (breathy (solid black line), creaky (light gray line) and modal (dark gray line)) when *log*(F2) values are high.

the model, the actual effect was more complex. As with $log$(F1) above, significant interactions of $log$(F1) and $log$(F2) meant that the effect size for $log$(F2) was much smaller when $log$(F1) values were low. When $log$(F1) was at the bottom of the range, $log$(F2) was positively correlated with H1*-A3* ($\beta < 11.15$, $\beta < 12.10$, and $\beta < 9.91$ for breathy, creaky, and modal phonation, respectively[9]). The effect of $log$(F2) on H1*-A3* when $log$(F1) is low is graphed in Figure 5. For all three phonation types, vowels with lower $log$(F2) frequencies (i.e. back vowels) had lower H1*-A3* values, suggesting a creakier phonation than vowels with higher $log$(F2) values (i.e. front vowels), though this effect was very small, and only applies when $log$(F1) values were low.

When $log$(F1) was highest, the effect was reversed, such that $log$(F2) had a negative effect ($\beta < -28.38$, $\beta < -34.54$, and $\beta < -19.70$ for breathy, creaky, and modal phonation, respectively[10]). Vowels with higher $log$(F2) frequencies (i.e. front vowels) had lower

---

[9] Values when $log$(F1) = 275.3 Hz.
[10] Values when $log$(F1) = 3035 Hz.

**Figure 5** Graph of the effect of the second formant (*log*(F2), Hz) on H1*-A3* (dB) for three phonation types (breathy (solid black line), creaky (light gray line) and modal (dark gray line)) when *log*(F1) values are low.
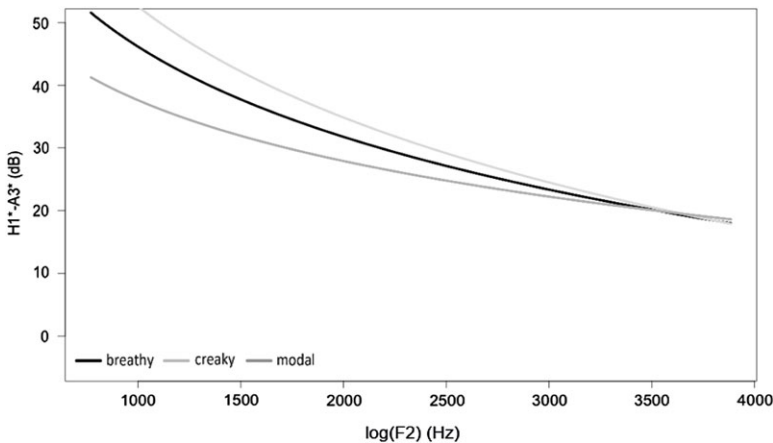


**Figure 6** Graph of the effect of the second formant (*log*(F2), Hz) on H1*-A3* (dB) for three phonation types (breathy (solid black line), creaky (light gray line) and modal (dark gray line)) when *log*(F1) values are high.

H1*-A3* values, suggesting a creakier phonation than vowels with lower *log*(F2) values (i.e. back vowels), though only when *log*(F1) values were high. This interaction is graphed in Figure 6.

## 4.2 Within-language results

When looking at the relationship between spectral measures and formants (regardless of vowel quality) within a language, there was a negative relationship between *log*(F1) and H1*-H2* for all languages[11] tested (see Figure 7, graph on the left). Vowels with greater *log*(F1) frequencies had lower H1*-H2* values, more typical of creaky phonation, suggesting that lower/more open vowels within each language were more creaky and less breathy. For the second formant (see Figure 7, graph on the right), all languages except Jalapa Mazatec, for which there was no effect of *log*(F2) on H1*-H2*, had a negative effect of *log*(F2) on

---

[11] The !Xóõ data were excluded from this analysis because they consisted only of tokens of [ɑ].
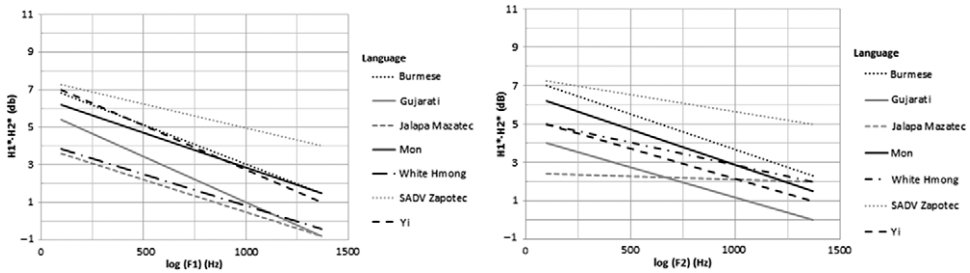
**Figure 7** Graphs of the effect of *log*(F1),Hz) on H1*-H2* (dB) (left), and *log*(F2),Hz) on H1*-H2* (dB) (right) for Burmese, Gujarati, Jalapa Mazatec, Mon, White Hmong, Santa Ana del Valle Zapotec and Yi.

H1*-H2*, suggesting that front vowels within each of these languages were more creaky and less breathy.

Moving onto the second acoustic measure, there was a positive relationship between *log*(F1) and *log*(F2) and H1*-A3* for all the languages tested except Jalapa Mazatec (see Figure 8). For six of the seven languages tested, vowels with higher *log*(F1) and higher *log*(F2) values had H1*-A3* values associated with breathiness. For Jalapa Mazatec, there was a positive relationship between *log*(F1) and H1*-A3*; however, there was no interaction between *log*(F2) and H1*-A3*.
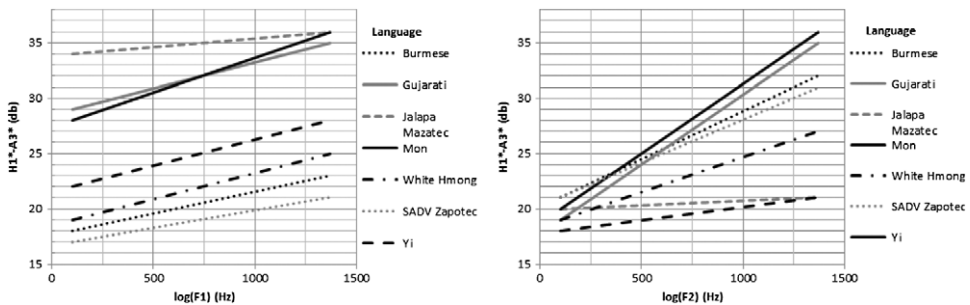


**Figure 8** Graphs of the effect of *log*(F1),Hz) on H1*-A3* (dB) (left), and *log*(F2),Hz) on H1*-A3* (dB) (right) for Burmese, Gujarati, Jalapa Mazatec, Mon, White Hmong, Santa Ana del Valle Zapotec and Yi.

### 4.3 Language-specific differences in voice quality as a function of vowel quality

In this section, we take a closer look at the results within a subset of the languages studied (Burmese, Jalapa Mazatec, and Santa Ana del Valle Zapotec) to see if vowel quality affects the degree of non-modal phonation. Gujarati, Mon, White Hmong, and Yi were excluded from this part of the study because the relationship between vowel quality and voice quality has already been documented in Khan (2012*), Abramson et al. (2015*), Esposito (2012*) and Kuang (2011*), respectively. !Xóõ was excluded because the data set contains only [ɑ]. We choose to include Jalapa Mazatec in this section because, while it has been the focus of two previous studies on voice quality (Kirk et al. 1993, Garellek & Keating 2011*), neither provided data on phonation type for individual vowel qualities.

Separate two-way repeated measures ANOVAs (one for each language) were performed, with vowel quality and voice quality as independent factors. Results showed that for both H1*-H2* and H1*-A3* there was not a significant difference in phonation due to vowel quality. The individual results for each vowel quality tested in Burmese, Jalapa Mazatec, and Santa Ana del Valle Zapotec are graphed in Figure 9 (H1*-H2*) and Figure 10 (H1*-A3*).
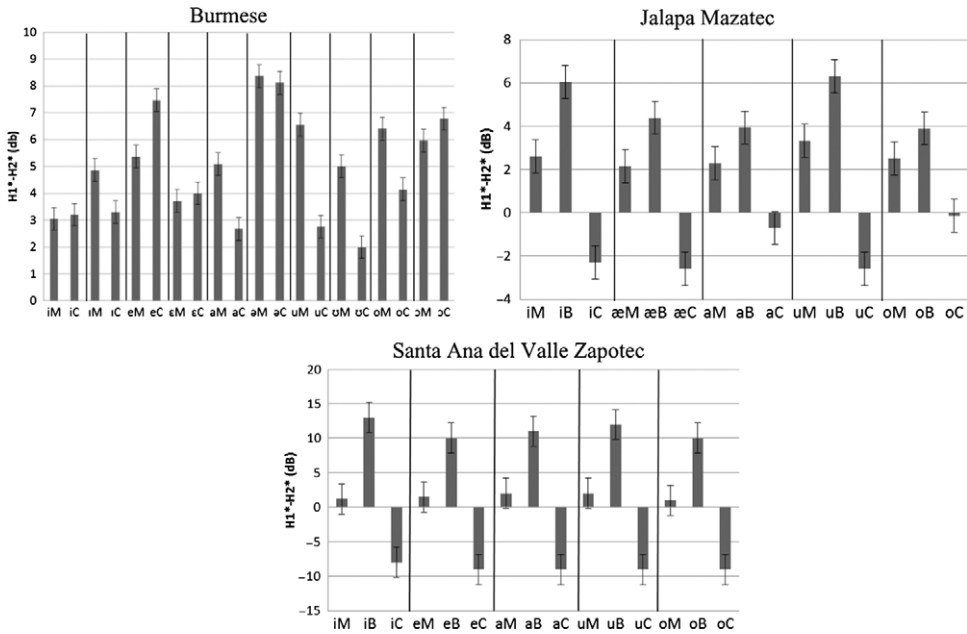
**Figure 9** Graphs of the average H1*-H2* (in dB) value for individual vowel qualities in, clockwise from top row: Burmese, Jalapa Mazatec; bottom row: Santa Ana del Valle Zapotec. Vowel qualities are transcribed using the International Phonetic Alphabet along the x-axis. Phonation type is indicated with an 'M' for modal, 'B' for breathy and 'C' for creaky. Error bars are standard error. There was no statistically significant difference in the phonation type between different vowel qualities.
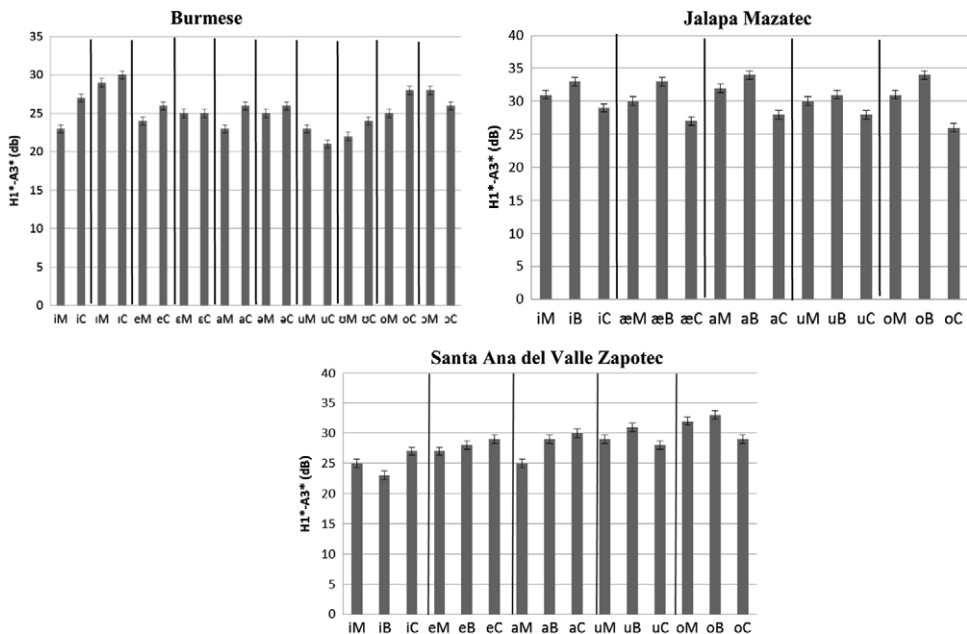


**Figure 10** Graphs of the average H1*-A3* (in dB) value for individual vowel qualities in, top row: Burmese, Jalapa Mazatec; bottom row: Santa Ana del Valle Zapotec. Vowel qualities are transcribed using the International Phonetic Alphabet along the x-axis. Phonation type is indicated with an 'M' for modal, 'B' for breathy and 'C' for creaky. Error bars are standard error. There was no statistically significant difference in the phonation type between different vowel qualities.

## 4.4 Summary

The current study examined phonation types in a wide range of vowels across eight unrelated languages. For the acoustic measure H1*-H2*, results showed a tendency for vowels with higher $log(F1)$ and higher $log(F2)$ values to be creakier, while vowels with lower $log(F1)$ and lower $log(F2)$ values were breathier. These results are represented in Figure 11. This trend holds true even when we collapsed phonation type differences and looked at the relationship between just vowel quality and H1*-H2*. For all languages examined, except Jalapa Mazatec, vowels, regardless of phonemic phonation type, with higher formant frequencies were creakier, while vowels with lower formant frequencies were breathier. For Jalapa Mazatec, there was no interaction between the second formant and the H1*-H2*.
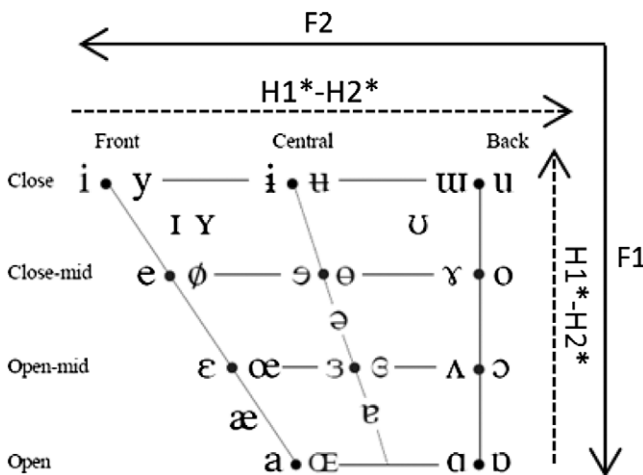


**Figure 11** The relationship between vowel quality and voice quality as measured by H1*-H2*. F1 and F2 are indicated with solid-line arrows. H1*-H2* values are indicated with dashed arrows. Vowels with higher formants values have a creakier H1*-H2* value. Vowels with lower formant values display a breathier H1*-H2* value. The Vowel Chart is from the IPA Chart (http://www.internationalphoneticassociation.org/content/ipa-chart, available under a Creative Commons Attribution-Sharealike 3.0 Unported License. Copyright ©2015 International Phonetic Association.)

Results for H1*-A3* were more complicated. For H1*-A3*, results showed a tendency for vowels with higher $log(F1)$ and higher $log(F2)$ values to be breathier, while vowels with lower $log(F1)$ and lower $log(F2)$ values were creakier. However, the relationship between phonemic phonation type and spectral tilt was not highly significant without interactional effects from $log(F1)$ and $log(F2)$. The positive relationship between H1*-A3* and formant frequency held true even when we collapsed phonation type differences and looked at the relationship between just vowel quality and H1*-A3*. Once again, results were different for Jalapa Mazatec, where there was no interaction between the second formant and H1*-A3*.

A question arises as to why Jalapa Mazatec patterned differently from the other languages tested? One reason might be because, in Jalapa Mazatec, voice quality and tone are fully cross-classified, unlike any of the other languages tested.[12] Perhaps this difference played a role in the results.

In regards to degree of non-modal phonation for a particular vowel quality, we found no significant difference in phonation type due to vowel quality in any of the languages tested

---

[12] This is also true of !Xóõ, but it was excluded from the part of the study looking at language-specific differences in voice quality as a function of vowel quality.

for either H1*-H2* or H1*-A3*. Other studies, such as Esposito (2012*) on White Hmong and Khan (2012*) on Gujarati, show the same pattern: phonation type does not vary as a function of vowel quality. However, this result contrasts with the reported results for Green Mong (Andruski & Ratliff 2000) and Yi (Kuang 2011*), making these the only two languages reported to have a phonation type difference due to vowel quality. This is particularly interesting given that White Hmong and Green Mong are two mutually intelligible varieties of Hmong, with similar phonation and vowel inventories; thus one might have expected them to pattern the same way.

## 5 Discussion and conclusion

Why does a relationship between formants and voice quality exist? One reason may be vocal tract length, which effectively alters formant frequency. During breathy phonation, the larynx may be lowered,[13] lengthening the vocal tract and, thus, lowering formant frequencies. For creaky phonation, on the other hand, the vocal tract is shortened due to a raised larynx (Laver 1980) and/or the engagement of the aryepiglottic sphincter (Esling 2005). The shorter vocal tract length in turn raises formant frequencies. This explanation fits with the results of this study along the H1*-H2* measure, where vowels with higher formant frequencies are creakier, and vowels with lower formant frequencies are breathier. However, these results do not hold true for H1*-A3*, where results remain more complicated, and in some cases (i.e. in terms of $log$(F1) when $log$(F2) is low, and of $log$(F2) when $log$(F1) is low) contradict the observed results from H1*-H2*. One reason for this discrepancy could be due to issues with using H1*-A3* as a measure of phonation in these eight languages. While H1*-H2* (or H1-H2) has been shown to be a successful measure of phonation type for all of these languages in previous literature (apart from Burmese, for which no studies we are aware of examine H1-H2), H1*-A3* (or H1-A3) has been previously shown to quantify phonation in only three of the languages, Gujarati, Mon, and Santa Ana del Valle Zapotec – and for the latter, only male speakers (Esposito 2010b). In fact, in looking at Figures 3–5, we see that along the H1*-A3* measure, the three phonation categories do not vary in the expected direction: the breathy phonation category has values that fall in between creaky and modal. However, for H1*-H2* (Figures 1 and 2), the values vary as expected for the three voice qualities – highest for breathy, intermediate for modal, and lowest for creaky. Bearing this in mind, along with the reduced explanatory power of the H1*-A3* model relative to the H1*-H2* model, and the fact that the effects of $log$(F1) and F2 on H1*-A3* are localized to specific range of values, we take the H1*-H2* measure to be most useful in determining the relationship between vowel quality and voice quality in these languages.

Furthermore, there could be a perceptual benefit to the covariation of phonation type and vowel quality. Evidence supporting this idea comes from two studies examining the effects of phonation type on the perception of vowel quality. In both studies, listeners' judgements were influenced by phonation type: listeners consistently identified breathier productions of vowels as having higher vowel quality (Lotto, Holt & Kluender 1997) and creakier productions as lower (Brunner & Żygis 2011). These results are particularly interesting because neither listening group (speakers of English and German, respectively) spoke a language with contrastive phonation, and yet both were still perceptually attuned to the relationship between vowel quality and voice quality seen along the H1*-H2* measure in the present study.

Thus far, the explanations we have presented for the relationship between vowel quality and voice quality assume that changes in phonation type induce changes in vowel quality.

---

[13] Though in a study observing the vocal tract during different phonation types, breathy phonation is observed with either a lowered or neutral laryngeal position (Edmondson & Esling 2006).

However, from the current research, it is unclear which of the two variables[14] is the causative one. There is some evidence from historical reconstructions of the register distinction in Mon-Khmer languages that suggests that laryngeal-height specifications that accompany non-modal phonation are the causative variable. In a diachronic account of the origin of register, Huffman (1976) proposes a stage where the breathy- and modally-phonated voiced onset consonants are reanalyzed as two separate registers (i.e. breathy and non-breathy). The phonation type of the onset consonant conditions a change in the vowel quality; when the voice quality differences between the registers were lost, the vowel quality distinctions remained.

Regardless of the cause and/or conditioning factor, any account of the relationship between vowel quality and voice quality must address the fact that similar first formant frequencies are reported for breathy and modal vowels in Mon[15] (Thongkum 1987a), Nyah Kur (Thongkum 1987b), Kui (Thongkum 1986), Khmu' Rawk (Abramson et al. 2007) and Gujarati (Khan 2012*). While the first four of these languages are all Mon-Khmer, the presence of Gujarati – an Indo-European language – on this list indicates that this phenomenon is not simply an isolated areal feature of this language family. One theory that could account for the similar first formant frequencies reported among breathy and modal vowels is that non-modal phonation types may be produced without changing the size of the vocal tract (e.g. with a neutral laryngeal position); this may explain why we do not see formant frequencies varying as a function of phonation in these five languages. There are many ways to produce non-modal phonation besides varying laryngeal height (including varying pharyngeal width (Traill 1985, Samely 1991) and/or tongue advancement (see Denning 1989 for an overview)), and future research could use articulatory tools to illuminate whether this is the case in these languages.

In conclusion, the results of this study, using data from eight unrelated languages and a wider vocalic inventory than previously investigated, show the following with regards to the interaction of vowel quality and voice quality: First, using the H1*-H2* measure, there exists a clear, cross-linguistic relationship between formant frequency and voice quality, with vowels with higher $log$(F1) and $log$(F2) values being creakier, and vowels with lower $log$(F1) and $log$(F2) values being breathier. Results are more complicated (and, in some cases, contradictory) for H1*-A3*, though because of the reduced explanatory power of the model using H1*-A3* and the fact that it has only previously been shown to successfully measure phonation type in three of the eight languages examined here, we regard H1*-H2* as the more useful measure in determining the relationship between vowel quality and voice quality. Second, for both H1*-H2* and H1*-A3*, vowel quality was shown not to significantly influence the degree of non-modal phonation in these languages, aligning with previous work on White Hmong (Esposito 2012*) and Gujarati (Khan 2012*), but contrasting with reported results for Green Mong (Andruski & Ratliff 2000) and Yi (Kuang 2011*).

## Acknowledgements

---

[14] An additional possibility is that vowel quality affects the accuracy of formant measurements and, therefore of formant correction, calling into question the accuracy of acoustic measurements. This concern could be addressed with more articulatory studies on voice.

[15] Mon is perplexing because different results were found in the current study; several factors could be at play, such as dialect variation or that the current study uses corrected measuring techniques.

## Appendix: Wordlists from the languages examined in the present study

Transcriptions and translations are retained from original wordlists, unless otherwise noted. Token order preserves the order from the source/archive material from which they were taken.

*!Xóõ*

In the transcription below, [ɑː] is rendered as <ɑɑ>.

| qà̰ɑ | 'long ago' |
| ǁáɑ | 'camelthorn tree' |
| ǀa̰ˀje | 'wait for him/it' |
| ǀgà̰ˀje | 'bend' |
| ǁn̥a̰ˀɑm | 'be damp' |
| !n̥a̰ˀɑm | 'evade' |
| ˀ!nà̰le | 'lower it' |
| !a̰le | 'climb up' |
| nǂɢà̰li | 'species of tree' |
| ǂqà̰n | 'smooth' |
| ǂɑbá | 'peg' |
| !a̰lɑ | 'species of tree' |
| ǀa̰ˀɑbe | 'hide it' |
| !na̰ˀ'ɑbe | 'surround it' |

*Burmese*

| poʊ | 'be surplus' |
| poʊ̰ | 'send' |
| doʊ̰ʔ | 'stick' |
| pam̥õʊ̰ŋ | 'bread' |
| ḛɪʔ | 'a bag' |
| ḛɪ | 'to sleep' |
| lɛ̰ʔ | 'hand' |
| mən̥ɛ̰ʔ | 'morning' |
| təbɛ̰ | 'pupil' |
| lɛ | 'reserve' |
| tɪ̰ʔ | 'one' |
| wɪ̰ʔ | 'to wear' |
| sɪ̰ʔ | 'war' |
| twɪ̰ | 'to found' |
| za̰ʔjoʊ̃ŋ | 'play-shed, theatre' |
| ŋa̰ | 'five' |

| dim̥a̤ | 'just here' |
| la | 'come' |
| so̤ʔ | 'birth-story, play' |
| sa̤ɪʔjwɛ | 'paper' |
| pjɔ | 'appear, come to light' |
| mjɛ̃ʔna | 'face' |
| tṳ̃ | 'be wrinkled' |
| tʊ̃ŋ | 'to push' |
| sə̤ | 'a play' |
| təgã | 'one animal' |
| əsa | 'piece' |
| ɲɛ | 'intellect' |
| ɲ̥ɛʔ | 'squeeze' |
| jɛ̄ | 'for one' |
| jɛꞏ | 'possessive' |
| l̥ṳ | 'to be pretty' |
| gəjṳ | 'importance' |
| lu | 'man' |
| lṳꞏ | 'almost, nearly' |
| tũː | 'to push' |
| tṳ̃ | 'to crow' |
| tṳ̃ꞏ | 'to be wrinkled' |
| wa̤ | 'to be fat' |
| ja | 'thing which, place where' |
| əɲaða̤ | 'man of the upper country' |
| deða̤ | 'a place' |
| wi̤ji̤ja̤ | 'vigour' |
| tʃa | 'to be long in duration' |
| tʃa̤ꞏ | 'to fall' |
| mjĩmma̤ | 'mare' |
| tã | 'to be worth' |
| tã̤ꞏ | 'to be clean' |
| tʃī | 'to cut' |
| tʃi̤ꞏ | 'to look' |
| pjɪ(n)naji̤ | 'scholar' |
| tədi̤ | 'one bushel' |
| lē | 'wind (N)' |

| | |
|---|---|
| le̤· | 'habit' |
| pɔ̃ | 'above, on' |
| pɔ̤· | 'certainty' |
| lẽĩ | 'twist' |
| lḛ̃ĩ̤· | 'roll' |
| kõũ | 'completeness' |
| kõ̤ṳ̃· | 'pick it up' |
| ʃoʊʔ | 'to be confused' |
| tãĩɲoʊ̤ | 'cry' |

*Gujarati*

| | |
|---|---|
| kan | 'ear' |
| ka̤n | 'Krishna' |
| paɖʋũ | 'to draw, to cause to fall' |
| pa̤ɖ | 'mountain' |
| bar | 'twelve' |
| ba̤r | 'outside' |
| baɳ | 'arrow' |
| ba̤nũ | 'excuse' |
| malik | 'boss, god' |
| ma̤raɹʒ | 'priest, emperor' |
| rab | 'gruel, baby food' |
| ra̤bər | 'guide' |
| cʃalaki | 'cleverness, trick' |
| səla̤kar | 'advisor' |
| ʋan | 'complexion' |
| ʋa̤n | 'vehicle' |
| ləʋaro | 'prattle, gossip' |
| ʋjəʋa̤ru | 'feasible' |
| səməgrə | 'whole, entire' |
| sətjagrə̤ | 'insistence on truth' |
| ḏərəḏ | 'disease, pain' |
| sərə̤ḏ | 'frontier, border' |
| keʋũ | 'of what kind' |
| kɛ̤ʋũ | 'to say' |
| ḏeʃi | 'native, villager' |
| ḏɛ̤ʃə̤ṯ | 'apprehension' |

| | |
|---|---|
| pelũ | 'that' |
| pe̥lũ | 'first' |
| benəmun | 'unparalleled' |
| bɛn | 'sister' |
| beroɹʒgar | 'unemployed' |
| bɛraʃ | 'deafness' |
| mɛ̃ | 'I' |
| mɛman | 'guest' |
| mɛl | 'dirt, filth' |
| me̥l | 'palace' |
| ʋɛr | 'revenge, enmity' |
| ʋe̥r | 'sawdust' |
| ʋeʋəj | 'father-in-law of one's child' |
| ʋe̥ʋar | 'daily interaction' |
| ʃerɖi | 'sugarcane' |
| ʃɛr | 'city' |
| cʃinṯa | 'worry' |
| cʃi̥nə | 'mark, trait' |
| ʋiʋrəɳ | 'criticism, commentary' |
| ʋi̥ʋəl̪ | 'puzzled and upset' |
| kɔgl̥o | 'mouthful of liquid, gargle' |
| kɔ̥ʋũ | 'to rot' |
| ɖol̥o | 'eyeball' |
| ɖol̥ʋũ | 'to swing' |
| ɖɔl̥ũ | 'dirty, polluted' |
| ɖɔl̥ʋũ | 'to stir in' |
| ɖ̪oro | 'thread, necklace' |
| ɖ̪ɔ̥ʋũ | 'to milk' |
| mokl̥ũ | 'spacious, open' |
| mɔ̃ | 'mouth, face' |
| mɔko | 'favorable opportunity' |
| mɔ̥k | 'charming' |
| locʃən | 'eye' |
| lo̥cʃumbək | 'magnet' |
| soɖ̥o | 'trade' |

| | |
|---|---|
| sɔ | 'hundred' |
| sɔ̰dər | 'born of the same mother' |
| nəivɐd̪ | 'food offered to god' |
| nəjʋə̰t̪ | 'slightly' |
| bɔ̰w | 'many, much' |

*Jalapa Mazatec*

| | |
|---|---|
| nã²ɲã¹ʃĩ³ | 'hunting dog' |
| nã²ɲã¹ʃḭ̃³ | 'male dog' |
| ho² nũ¹ | 'two years' |
| ho² nṵ̃¹ | 'two vines' |
| ngu² ʃi² kã² | 'one is twenty' |
| ngu² ʃi² kã̰² | 'one lives alone' |
| tʃa¹ | 'old' |
| tʃa̰¹ | 'load' |
| ʔjũ¹ nda² | 'very good' |
| ʔjũ¹ nda̰³² | 'very hard' |
| ti²ɸa̰¹hĩ² | 'choose' |
| ndæ³ | 'horse' |
| ndæ̰³ | 'buttocks' |
| ja̰² tʃi³nka² | 'he wears a shirt' |
| ja̰² tʃi³nka² | 'he carries a shirt' |
| nk²ja² | 'house' |
| tʃu³jæ³ | 'turtle' |
| jæ³ | 'manure' |
| jæ³ | 'snake' |
| hæ²ngḭ² | 'went' |
| tʃḭ ²kḭ ² | 'he paints' |
| nĩ²ʔja²ndḭ³ | 'grass-roofed house' |
| tsḭ ³ | 'drunk' |
| sḭ ³ | 'you drink (imperative)' |
| tsḭ ³ | 'nausea' |
| hi¹tsi³ | 'yours' |
| ja̰¹nḭ̃² | 'carries (on his back)' |
| tsḭ̃ | 'fever' |
| nĩ² | 'red' |
| nḭ̃³ | 'kernels of corn' |

| | |
|---|---|
| ʔjũ¹ ndʒæ̰²¹ | 'thief' |
| ʔjũ¹ ntæ̰¹ | 'very tame' |
| nga² jæ²³ | 'everything' |
| ti¹mã²ndzæ̰² | 'visible' |
| hæ² | 'green ear of corn' |
| na²m̩mi² ntæ² | 'he doesn't hear' |
| tæ² | 'ten' |
| (n)tæ² nta³ | 'I hear his voice' |
| tʃæ² | 'lazy' |
| tʃu³næ̰³ | 'braids' |
| kʷi²tʃæ³ | 'his' |
| tʃæ³tʃæ³ | 'new' |
| mæ̰̃² | 'he wants' |
| kæ̰̃² | 'dead' |
| nĩ²mæ̰̃² | 'bumblebee' |
| ndʒa̰¹ ʃu¹ | 'chocolate drink' |
| ʔi³ja²ntʃa̰¹ | 'fence' |
| ti¹βa̰² | 'he hits' |
| ngʷa̰³ | 'he puts on' |
| kʷa̰³ | 'it will happen' |
| ka²ma²ɲa̰² | 'gather together' |
| tʃa̰³² | 'armadillo' |
| næ²ʃa¹ta̰²¹ | 'we (EXCL) work for' |
| ɲã̰² | 'nine' |
| nã̰¹ | 'tongue' |
| ʃã̰³ | 'poor' |
| nã³ | 'mother' |
| jo̰² | 'meat' |
| (ʔjũ¹) ndʒo̰² | '(a lot of) dust' |
| βo̰² | 'hungry' |
| ti¹ʔo² | 'grinds' |
| tʃõ̰³ | 'thunder, lightning' |
| ka³ma³ndṵ | 'it became long' |
| ka³ma³ndʒṵ | 'it became tomorrow' |
| tʃṵ¹ | 'blouse' |
| ti²jṵ² | 'it is stopped' |

| | |
|---|---|
| ti²ʔju² | 'you (PL) drink' |
| ta³βa³ʃi¹ɲṵ̃² | 'ear' |
| ɲṵ̃² | 'four' |
| nũ¹ | 'year' |
| nṵ̃¹ | 'vine' |
| ti¹βa̰¹ʔa³ | 'weave' |
| ti¹ɸʷa̰² | 'diarrhea' |
| ni²ɸʷa² | 'clear, clean' |

*Mon*

The original handwritten transcription uses [ɑ] in place of [a]; based on the recordings and analogous transcriptions in Huffman (1990), we have changed them to [a] here. The transcription otherwise remains unchanged from the source. In the transcription below, [j] is rendered as <y>.

| | |
|---|---|
| kloik | 'pig' |
| klo̰ik | 'waistcloth' |
| kleəŋ | 'to bind' |
| klḛəŋ | 'to carry on the shoulder' |
| kloiŋ | 'wolf' |
| klo̰iŋ | 'long' |
| kləŋ | 'to come' |
| klə̰ŋ | 'boat' |
| kleʔ | 'to do last' |
| klḛʔ | 'short' |
| kleh | 'to wear over shoulder' |
| klḛh | 'to carve meat' |
| to | 'cotton' |
| to̰ | 'center' |
| toa | 'hand' |
| to̰a | 'to sleep' |
| nṵa | 'crest of bird' |
| pua | 'religious festival' |
| pṵa | 'as much as' |
| ɓoa | 'bean' |
| ɓo̰a | 'older sister' |
| mṵa | 'one' |
| wɛʔ | 'a Wa (an ethnic group)' |

| hwę̰ʔ | 'corpse' |
|---|---|
| ŋiək | 'small parrot' |
| ŋi̤ek | 'tooth' |
| sạy | 'to separate' |
| sạy | 'skinny' |
| loik | 'moss' |
| lọik | 'book' |
| wạiŋ | 'mint' |
| hwaiŋ | 'bifurcate (road)' |
| hwạik | 'serrated' |
| pay | 'reject' |
| pạy | 'spouse of elder sister' |
| pay | 'to separate' |
| ɓạy | (term of address) |

*Santa Ana del Valle Zapotec*

In the transcription below, a falling tone on a breathy vowel is rendered with a grave accent.

| páʔp | 'potato' |
|---|---|
| kárː | 'car' |
| páɡ | 'pay' |
| fáb | 'detergent' |
| táp | 'lid' |
| lát | '(tin) can' |
| rán | 'frog' |
| bál | 'stick' |
| bá | 'cow' |
| káʔmː | 'bed' |
| nậʔtʃ | 'rough' |
| dăd | 'father' |
| băd | 'duck' |
| lăʔs | 'skinny' |
| stămp | 'stamp' |
| kwăʔtʃ | 'twin' |
| măn | 'animal' |
| dămː | 'owl' |
| dʒăn | 'saint' |

| | |
|---|---|
| fǎld | 'skirt' |
| mǎrkw | 'mark' |
| rập | 'have' |
| tập | 'four' |
| bậd | 'Tlacolula' |
| sậk | 'price' |
| bậdʒ | 'tepache' |
| dậny | 'mountain' |
| rập | 'care' |
| nːậʃ | 'chocolate' |
| lậdy | 'clothes' |
| nậl | 'thick' |
| nận | 'thick' |
| nậd | 'to stick' |
| dạʔạ | 'petate' |
| lạʔạ | 'huaje bean' |
| lậts | 'field' |
| dʒậʔp | 'girl' |
| bạʔạ | 'tomb' |
| bḙlː | 'meat' |
| bḙdz | 'rabbit' |
| bḙʔ | 'mushroom' |
| gḙl | 'dawn' |
| bḙtz | 'child' |
| líbr | 'book' |
| tíly | 'Spanish' |
| díʒ | 'language' |
| lím | 'lime' |
| ʃíl | 'wing' |
| tʃĩntʃ | 'bedbug' |
| tʃĩkl | 'gum' |
| tĩʔt | 'Tito (proper name)' |
| zǐt | 'stranger' |
| bĩʔtʃ | 'small animal' |
| zịt | 'far away' |
| dʒị̀ | 'day' |

| | |
|---|---|
| nì̱ʒ | 'delicious' |
| bâ̱d | 'scabies' |
| ʒâ̱n | 'butt' |
| bâ̱ʔt | 'skunk' |
| gâ̱ʔa̱ | 'nine' |
| dâ̱ts | 'empty' |
| gâ̱ts | 'hidden' |
| béd | 'Peter' |
| prés | 'prince' |
| tʃétʃ | 'Moses' |
| tʃéʔk | 'check' |
| méskl | 'mortar' |
| mĕs | 'table' |
| bĕrd | 'green' |
| bĕn | 'twenty' |
| mĕɾg | 'row of agave plants' |
| mĕl | 'Ishmael' |
| dè̱ts | 'back' |
| gè̱t | 'tortilla' |
| gè̱l | 'milpa' |
| bè̱l | 'fish' |
| dè̱ | 'powder' |
| ʃi̱g | 'jicara' |
| nì̱z | 'dried corn cob' |
| rî̱ʔi̱ | 'jar' |
| nî̱ʔi̱ | 'foot' |
| nî̱s | 'water' |
| ʒî̱t | 'far away' |
| gî̱ts | 'paper' |
| sóʔbr | 'envelope' |
| lór | 'parrot' |
| ór | 'hour' |
| kótʃ | 'car' |
| bób | 'drool' |
| ʃŏp | 'six' |
| sŏp | 'soap' |

| | |
|---|---|
| tŏp | 'speed bump' |
| bŏt | 'canister' |
| kŏk | 'coconut' |
| tʃo̰n: | 'three' |
| do̰b | 'maguey' |
| ko̰b | (a type of alcoholic drink) |
| lo̰ | 'face' |
| do̰p | 'short' |
| ɾo̰ʔ | 'big' |
| do̰n | 'if' |
| bo̰ʔ | 'carbon' |
| ʒo̰b | 'corn' |
| do̰p | 'flatulence' |
| lúk | 'Lucas' |
| lúdʒ | 'tongue' |
| ʃúk | 'upper arm' |
| kúrs | 'course' |
| lús | 'Luis' |
| lŭpy | 'Lupe' |
| mŭsk | 'music' |
| lŭʒ | (a type of edible rodents) |
| pŭnt | 'top' |
| grŭp | 'group' |
| ɾṳ | 'cough' |
| ʂṳ | 'earthquake' |
| sṳdy | 'a cut' |
| dṳl: | 'fault' |
| ʃṳn | 'eight' |
| ɾṳʔ | 'mouth' |
| dṳʔ | 'rope' |
| gṳn | 'bull' |
| lṳ | 'you' |
| gṳʔ | 'sweet potato' |

*White Hmong*

| | |
|---|---|
| ca[45] | 'haul or tow' |
| ka[45] | 'insect' |

| | |
|---|---|
| pɔ$^{45}$ | 'ball-like' |
| tɔ$^{45}$ | 'very deep' |
| ca$^{33}$ | 'log' |
| pɔ$^{33}$ | 'spleen' |
| tɔ$^{33}$ | 'puncture' |
| ca$^{22}$ | 'why, how' |
| ka$^{22}$ | 'maggots' |
| ki$^{22}$ | 'to infect' |
| kɔ$^{22}$ | 'stem' |
| pa$^{22}$ | 'stick' |
| pɔ$^{22}$ | 'thorn; cover' |
| pu$^{22}$ | 'escape with all or part of the trap/arrow still attached' |
| qa$^{22}$ | 'disgust, sicken' |
| ta$^{22}$ | 'finished, end, done' |
| ti$^{22}$ | 'wing' |
| tɔ$^{22}$ | 'to wait' |
| tu$^{22}$ | 'open-minded, stable, peaceful' |
| za̠$^{22}$ | 'color, paint' |
| ca$^{52}$ | 'ridge' |
| ka$^{52}$ | 'dawn, bright' |
| pa$^{52}$ | 'flower' |
| pɔ$^{52}$ | 'female' |
| tɔ$^{52}$ | 'ahead' |
| ca$^{24}$ | 'argue, disagree' |
| ka$^{24}$ | 'stem, stalk' |
| pɔ$^{24}$ | 'throw' |
| tɔ$^{24}$ | 'to mix' |
| ca̠$^{21}$ | 'to stick; to argue' |
| ci̠$^{21}$ | 'mark' |
| ka̠$^{21}$ | 'allow' |
| ki̠$^{21}$ | 'expensive' |
| ku̠$^{21}$ | 'a proper name (male)' |
| pa̠$^{21}$ | 'bridge; blanket' |
| pi̠$^{21}$ | 'vagina' |
| pɔ$^{21}$ | 'to see' |
| pu̠$^{21}$ | 'to see' |

| | |
|---|---|
| tạ²¹ | 'represent; sharpen (as in a knife)' |
| tị²¹ | 'because of it' |
| tọ²¹ | 'to bite' |
| tụ²¹ | 'to stack' |
| zạ²¹ | 'avoid' |
| cạ⁴² | 'root; origin' |
| cị⁴² | 'light up' |
| cụ⁴² | 'collect' |
| kụ⁴² | 'very poor' |
| pọ⁴² | 'paternal grandmother' |
| qạ⁴² | 'axle; swivel' |
| tạ⁴² | 'done' |
| tị⁴² | 'turn; steer' |
| tọ⁴² | 'chair' |
| tụ⁴² | 'whose; properly' |
| zạ⁴² | 'sentence' |

*Yi*

| | |
|---|---|
| bi²¹ | 'give' |
| bi̱²¹ | 'rupture' |
| sz²¹ | 'twig' |
| sẕ²¹ | 'new/thirsty' |
| sʅ¹²¹ | 'seven' |
| sʅ³³ | 'screw' |
| bu³³ | 'worm' |
| bu̱³³ | 'full' |
| dɛ³³ | 'cut' |
| dɛ̱³³ | 'plain' |
| bə²¹ | 'mountain' |
| bə̱²¹ | 'hoof' |
| bɚ²¹ | 'bottle' |
| la²¹ | 'light' |
| la̱²¹ | 'scoop up' |
| ɣo²¹ | 'sell' |
| ɣo̱²¹ | 'obtain' |
| li³³ | 'heavy' |

| | |
|---|---|
| li̠³³ | 'to market' |
| dʐ̩¹²¹ | 'wine' |
| dʐ̩̠¹²¹ | 'sew' |
| fu³³ | 'flow out' |
| fu̠³³ | 'maggot' |
| kɛ³³ | 'pick up' |
| kɛ̠³³ | 'scold' |
| də²¹ | 'sun set' |
| də̠²¹ | 'to fruit' |
| zɚ²¹ | 'pillar' |
| to³³ | 'put' |
| to̠³³ | 'chop' |

## References

Abramson, Arthur S., Patrick W. Nye & Therapan Luangthongkum. 2007. Voice register in Khmu': Experiments in production and perception. *Phonetica* 64(2–3), 80–104.

Abramson, Arthur S., Mark K. Tiede & Therapan Luangthongkum. 2015. Voice register in Mon: Acoustics and electroglottography. *Phonetica* 72(2–3), 237–256.

Andruski, Jean & Martha Ratliff. 2000. Phonation types in production of phonological tone: The case of Green Mong. *Journal of the International Phonetic Association* 30(1/2), 37–61.

Audacity Team. 2014. Audacity(R): Free audio editor and recorder, Version 2.0.0 [computer program]. http://audacity.sourceforge.net/ (24 April 2014).

Bates Douglas, Martin Maechler, Benjamin Bolker & Steven Walker S. 2015. *lme4*: Linear mixed-effects models using Eigen and S4. R package version 1.1-8, http://CRAN.R-project.org/package=lme4 (12 January 2015).

Bartoń, Kamil. 2015. MuMIn: Multi-model inference. R package version 1.15.1. http://CRAN.R-project.org/package=MuMIn (12 January 2015).

Bickley, Corine. 1982. Acoustic analysis and perception of breathy vowels. S*peech Communication Group Working Papers* 1, 73–93. [Research Laboratory of Electronics, MIT]

Blankenship, Barbara. 2002. The timing of nonmodal phonation in vowels. *Journal of Phonetics* 30, 163–191.

Boersma, Paul & David Weenink. 2008. Praat: Doing phonetics by computer. http://www.praat.org/ (25 April 2014). [Computer program]

Bradley, David. 1982. Register in Burmese. In David Bradley (ed.), *Tonation* (Pacific Linguistics Series A-62), 117–132. Canberra: Australian National University.

Brunner, Jana & Marzena Żygis. 2011. Why do glottal stops and low vowels like each other? *Proceedings of the 17th International Congress of Phonetic Sciences* (ICPhS XVII), Hong Kong, 376–379.

Denning, Keith. 1989. *The diachronic development of phonological voice quality, with special reference to Dinka and the other Nilotic languages*. Ph.D. dissertation, Stanford University.

DiCanio, Christian T. 2009. The phonetics of register in Takhian Thong Chong. *Journal of the International Phonetic Association* 39(2), 162–188.

Edmondson, Jerold A. & John H. Esling. 2006. The valves of the throat and their functioning in tone, vocal register and stress: Laryngoscopic case studies. *Phonology* 23(2), 157–191.

Esling, John H. 2005. There are no back vowels: The laryngeal articulator model. *Canadian Journal of Linguistics* 50, 13–44.

Esposito, Christina M. 2006. *The effects of linguistic experience on the perception of phonation*. Ph.D. dissertation, University of California, Los Angeles.

Esposito, Christina M. 2010a. The effects of linguistic experience on the perception of phonation. *Journal of Phonetics* 38(2), 306–316.

Esposito, Christina M. 2010b. Variation in contrastive phonation in Santa Ana del Valle Zapotec. *Journal of the International Phonetic Association* 40(2), 181–198.

Esposito, Christina M. 2012. An acoustic and electroglottographic study of White Hmong tone and phonation. *Journal of Phonetics* 40, 466–476.

Esposito, Christina M. & Sameer ud Dowla Khan. 2012. Contrastive breathiness across consonants and vowels: A comparative study of Gujarati and White Hmong. *Journal of the International Phonetic Association* 42(2), 123–143.

Fischer-Jørgensen, Eli. 1967. Phonetic analysis of breathy (murmured) vowels in Gujarati. *Indian Linguistics* 28, 71–139.

Garellek, Marc & Patricia A. Keating. 2011. The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *Journal of the International Phonetic Association* 41(2), 185–205.

Gordon, Matthew & Peter Ladefoged. 2001. Phonation types: A cross-linguistic overview. *Journal of Phonetics* 29, 383–406.

Hanson, Helen M. 1995. *Glottal characteristics of female speakers*. Ph.D. thesis, Harvard University.

Hanson, Helen M. 1997. Glottal characteristics of female speakers: Acoustic correlates. *The Journal of the Acoustical Society of America* 101(1), 466–481.

Hanson, Helen M. & Erika S. Chuang 1999. Glottal characteristics of male speakers: Acoustic correlates and comparison with female data. *The Journal of the Acoustical Society of America* 106(12), 1064–1077.

Hombert, Jean-Marie. 1978. Consonant types, vowel quality and tone. In Victoria Fromkin (ed.), *Tone: A linguistic survey*, 77–112. New York: Academic Press.

Huffman, Franklin. 1976. The register problem in fifteen Mon-Khmer languages. In Philip. N. Jenner (ed.), *Austroasiatic studies 1*, 575–589. Honolulu, HI: University of Hawai'i Press.

Huffman, Franklin. 1990. Burmese Mon, Thai Mon, and Nyah Kur: A synchronic comparison. *Mon-Khmer Studies* 16–17, 31–84.

Iseli, Markus, Yen-Liang Shue & Abeer Alwan. 2007. Age, sex, and vowel dependencies of acoustic measures related to the voice sources. *The Journal of the Acoustical Society of America* 121(4), 2283–2295.

Javkin, Hector & Ian Maddieson. 1985. An inverse filtering study of creaky voice. *UCLA Working Papers in Phonetics* 57, 115–125.

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

Kawahara, Hidei, Alain de Cheveigné & Roy D. Patterson. 1998. An instantaneous-frequency-based pitch extraction method for high-quality speech transformation: Revised TEMPO in the STRAIGHT-suite. *Proceedings of the 5th International Conference on Spoken Language Processing*, Paper 0659.

Keating, Patricia A., Christina [M.] Esposito, Marc Garellek, Sameer ud Dowla Khan & Jianjiang Kuang. 2010. Phonation contrasts across languages. *UCLA Working Papers in Phonetics* 108, 188–202.

Khan, Sameer ud Dowla. 2012. The phonetics of contrastive phonation in Gujarati. *Journal of Phonetics* 40(6), 780–795.

Kirk, Paul, Jenny Ladefoged & Peter Ladefoged. 1993. Quantifying acoustic properties of modal, breathy and creaky vowels in Jalapa Mazatec. In Anthony Mattina & Timothy Montler (eds.), *American Indian linguistics and ethnography: In honor of Laurence C. Thompson*, 435–450. Missoula, MT: University of Montana Press.

Kuang, Jianjing. 2011. Production and perception of the phonation contrast in Yi. Master's thesis, University of California, Los Angeles.

Ladefoged, Peter & Norma Antoñanzas-Barroso. 1985. Computer measures of breathy voice quality. *UCLA Working Papers* 61, 79–86.

Laver, John. 1980. *Phonetic description of voice quality*. Cambridge: Cambridge University Press.

Lee, Thomas. 1983. An acoustical study of the register distinction in Mon. *UCLA Working Papers in Phonetics* 57, 79–96

Lewis, Paul M., Gary F. Simons & Charles D. Fennig (eds.). 2016. *Ethnologue: Languages of the world*, 17th edn. Dallas, TX: SIL International. http://www.ethnologue.com.

Lotto, Andrew, Lori Holt & Keith Kluender. 1997. Effect of voice quality on perceived height of English vowels. *Phonetica* 54, 76–93.

Maddieson, Ian & Peter Ladefoged. 1985. Tense and lax in four minority languages of China. *Journal of Phonetics* 13, 433–454.

Ratliff, Martha. 1992. *Meaningful tone: A study of tonal morphology in compounds, form classes, and expressive phrases in White Hmong* (Center for Southeast Asian Studies Monograph series; Northern Illinois University, Special report 27).

Ren, Nianqi. 1992. *Phonation types and stop consonant distinctions: Shanghai Chinese*. Ph.D. dissertation, University of Connecticut.

Samely, Ursula. 1991. *Kedang (Eastern Indonesia): Some aspects of its grammar*. Hamburg: Helmut Buske.

Shue, Yen-Liang, Patricia A. Keating & Chad Vicenik. 2009. VoiceSauce: A program for voice analysis. *The Journal of the Acoustical Society America* 124(4), 2221.

Shue, Yen-Liang, Patricia A. Keating, Chad Vicenik & Kristine Yu. 2011. *VoiceSauce: A program for voice analysis*. *Proceedings of the 17th International Congress of Phonetic Sciences* (ICPhS XVII), Hong Kong, 1846–1849.

Silverman, Daniel, Barbara Blankenship, Paul Kirk & Peter Ladefoged. 1995. Phonetic structures in Jalapa Mazatec. *Anthropological Linguistics* 37, 70–88.

Sjölander, Kåre. 2004. *The snack sound toolkit*. KTH Stockholm, Sweden. http://www.speech.kth.se/snack/ (1 June 2014).

Smalley, William. 1976. The problem of consonants and tone: Hmong (Meo, Miao). In William Smalley (ed.), *Phonemes and orthography: Languages planning in ten minority languages in Thailand* (Pacific Linguistics Series C no. 43). Canberra: Australian National University.

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

Thongkum, Theraphan. 1987a. Another look at the register distinction in Mon. *UCLA Working Papers in Phonetics* 67, 132–165.

Thongkum, Theraphan. 1987b. Phonation types in Mon-Khmer languages. *UCLA Working Papers in Phonetics* 67, 9–48.

Thongkum, Theraphan. 1990. The interaction between pitch and phonation type in Mon: Phonetic implications for a theory of tonogenesis. *Mon-Khmer Studies* 16–17, 11–24.

Traill, Anthony. 1985. *Phonetic and phonological studies of !Xóõ Bushman*. Hamburg: Buske.

Watkins, Justin W. 2000. Notes on creaky and killed tone in Burmese. *SOAS Working Papers in Linguistics* 10, 139–149.

Watkins, Justin W. 2001. Burmese. *Journal of the International Phonetic Association* 31(2), 291–295.

Wayland, Ratree & Allard Jongman. 2002. Registrogenesis in Khmer. *Mon-Khmer Studies* 32, 101–114.