

## Acoustic Correlates of Breathy Sonorants in Marathi

Kelly Harper Berkson<sup>a\*</sup>

\*Corresponding author: [kberkson@indiana.edu](mailto:kberkson@indiana.edu), Tel: 603-568-7273

<sup>a</sup>Department of Linguistics, Indiana University Bloomington

**Abstract:** Breathy voiced sonorant consonants are typologically rare, more so than other non-modal sonorants (e.g. voiceless sonorants, which are widely attested in language families like Tibeto-Burman and Otomanguean). Similarly, they are more vulnerable to diachronic loss than voiceless sonorants (e.g. in Tibeto-Burman). The acoustic correlates of breathiness in sonorants have not been thoroughly investigated, and a question arises as to whether there is a tie between their acoustics and their typology: does the acoustic encoding of breathiness in sonorants contribute to their typological scarcity? The current study probes this question via instrumental acoustic analysis of breathy and modal obstruents and sonorants in Marathi, an Indic language. Measures standardly used to assess voice quality (F0, H1\*-H2\*, H1\*-A1\*, H1\*-A2\*, H1\*-A3\*, and Cepstral Peak Prominence) are reported. As expected, breathy voiced obstruents are associated with increased values in the H1-based measures and decreased values in Cepstral Peak Prominence (CPP), a measure which reflects the presence of noise in the signal. Sonorants show the same general trend, with higher H1-based measures and lower CPP associated with breathy than with plain sonorants, but the differences between plain and breathy consonants is greater in obstruents than in sonorants. Specifically, the acoustic correlates of breathy voice are diminished in post-sonorant contexts. It is proposed that phonemic breathy voice is not acoustically encoded as robustly in sonorant consonants as in obstruents, and that this helps explain the typology.

### Keywords

Phonation; breathy voice; obstruents; sonorants; typology

## Acoustic Correlates of Breathy Sonorants in Marathi

### 1 Introduction

While some phonation contrasts (e.g. voicing in obstruents) are typologically common, others such as the contrast between modal (“plain”) and breathy voicing are more rare. Phonemic breathiness is under-represented both typologically and in empirical research on the acoustics of phonation type distinctions. The latter is changing, however, and recent work on phonemically breathy voiced obstruents (Alam *et al.* 2008, Davis 1994, Dutta 2007, Mikuteit & Retz 2007) and vowels (Blankenship 2002; Esposito 2010, 2012; Kelterer 2017; Khan 2012; Wayland & Jongman 2003) has revealed several facts: first, breathy voice is associated with a suite of acoustic correlates, including but not limited to temporal and spectral measures. Second, the acoustic characteristics of breathy phonation—in particular, increased spectral tilt measures such as H1-H2, which relates the amplitude of the first harmonic (H1) to the amplitude of the second harmonic (H2), and increased noise as reflected by Harmonics to Noise Ratio (HNR) measures—remain consistent across segment types and languages. Very little is known, however, about the acoustic properties of phonation-type contrasts in sonorant consonants (henceforth referred to simply as sonorants).<sup>1</sup>

What phonetic studies do exist suggest that the spectral characteristics of breathy sonorants pattern with those of other breathy sounds, but these works are subject to limitations. Harris (2009) reports on three sounds (breathy bilabial and alveolar nasals and a breathy lateral) in Sumi, a Tibeto-Burman language, but data are from one female speaker.<sup>2</sup> Gautam (2012) reports on Balami, another Tibeto-Burman language, and while five speakers were recorded the results reported for individual measures are often for only one speaker. Moreover, methodological descriptions are sparse or nonexistent, rendering it impossible to determine where or how measurements were taken. Breathily nasals have been noted in some Bantu languages (Ladefoged 1971, Maddieson 2003), but there is little extant acoustic work on this topic. Demolin and Delvaux’s (2003) work on what they call “whispery nasals” in Kinyarwanda, for instance, includes aerodynamic and duration data but no spectral measures. Traill and Jackson (1988) focus on Tsonga and includes a number of acoustic measures, meanwhile, but while data are from 15 participants the focus is on the plain and breathy bilabial and alveolar nasals in a single vowel context (before the low vowel /a/). The voiceless sonorants found in a number of Tibeto-Burman languages have also been the focus of investigation: in addition to descriptions of their diachronic origin and phonological patterning (Chirkova & Handel 2015, Dantsuji 1984, Hill 2014, K. Van Bik 2009), acoustic and perception studies have also been conducted (Bhaskararao & Ladefoged 1991; Blankenship 1994; Chirkova, Bassett & Amelot 2018; Dantsuji 1984,

---

<sup>1</sup> The terms obstruent and sonorant are used here in the phonetic sense: articulatorily, obstruents are produced with a more significant obstruction in the vocal tract and sonorants with a looser constriction and relatively free airflow. Acoustically, sonorants are associated with greater acoustic energy.

<sup>2</sup> Voiceless variants are perhaps the most discussed non-modal sonorants in the Tibeto-Burman literature, but some Tibeto-Burman languages do contain breathy voiced sonorants. Sumi is one such language. The non-modal sonorants in Sumi feature voicing throughout and a breathy release. Their description as “voiced aspirates” by Sreedhar (1976) is confirmed in the acoustic study conducted by Harris (2009). Breathiness is also associated with tonal registers in a number of Tibeto-Burman languages. See, for example, descriptions of this phenomenon in Bodish languages spoken in Nepal such as Magar (Grunow-Hårsta 2008), Nar-Phu (Noonan & Hildebrandt 2016), Siklis Gurung (Ronkos 2015), and Tamang (Mazaudon 1973, Michaud & Mazaudon 2006).

1986; Gogoi & Wayland 2018; Hoffmann 2018). These sounds differ from breathy sonorants articulatorily, acoustically, and typologically, however.

The works noted above offer valuable insights on a phenomenon for which little empirical data has been reported, but the limitations of a single speaker (Harris 2009), unclear methodology and limited data (Gautam 2012), and a single manner of articulation (nasal) in a single vowel context (Traill & Jackson 1988) raise the question of how general the patterns are. To the best of my knowledge, no in-depth acoustic analysis of multiple breathy sonorants of several manners of articulation produced by many speakers in multiple vowel contexts and word positions yet exists. As such, the current study investigates the acoustic correlates of phonation type contrasts in Marathi sonorants. Two research questions are posed:

- (1) Are the acoustic correlates of breathy voice in sonorants *qualitatively* different than they are in obstruents and vowels?
- (2) Are the acoustic correlates of breathy voice in sonorants *quantitatively* different than they are in obstruents and vowels?

A qualitative difference would mean that the acoustic correlates of breathiness in sonorants are distinct from those in obstruents and vowels. A quantitative difference would mean that differences between plain and breathy obstruents exceed those between plain and breathy sonorants. Using the voiced labials as representative symbols, this can be represented as  $\{[b^h-b] > [m^h-m]\}$ . These questions are particularly interesting given the typological and diachronic patterning of breathy sonorants, which are both rarer and less diachronically stable than breathy obstruents.

Typologically, several online databases can provide insight into the patterning of phonemic breathy voice. The UCLA Phonological Segment Inventory Database (UPSID) indexes 451 languages, 13 of which (2.9% of the total) have breathy consonants. Most of these are breathy obstruents; just five of the 13 (1.1%) contain breathy voiced sonorants.<sup>3</sup> Neither the Indic nor the Tibeto-Burman language families are widely represented in UPSID, however, and breathy voiced (or voiced aspirated) and voiceless sonorants (particularly nasals) are regularly found in Indic and Tibeto-Burman languages respectively. It is therefore useful to consider PHOIBLE (Moran, McCloy, & Wright 2014), which indexes the phonemic inventories of 1672 languages and includes inventories for many Indic and Tibeto-Burman languages. The general trend in PHOIBLE aligns with UPSID: using labial consonants as an example, 13 of the 1672 languages indexed in PHOIBLE are reported to contain /b<sup>h</sup>/ and 6 are reported to contain /m<sup>h</sup>/. Tibeto-Burman voiceless nasals occur most regularly at the labial and coronal places of articulation; voiceless velar nasals, palatal nasals, laterals, and rhotics are less frequent but nevertheless well represented (DeLancey 2009, Matisoff 2003, K. Van Bik 2009). An overview of how often these sonorants are attested in PHOIBLE appears in Table 1, where counts are included for modal, voiceless, and breathy sonorants.<sup>4</sup>

---

<sup>3</sup> Another five languages contain phonemically breathy vowels.

<sup>4</sup> Some counts exceed 1672, the number of languages indexed in PHOIBLE, because some languages have multiple entries. More than one inventory is indexed for Burmese, for example, one from UPSID and one from the Stanford Phonology Archive (<http://phoible.org/languages/mya>).

**Table 1: Counts of select modal and nonmodal sonorants in PHOIBLE**

SEGMENT	PLAIN	VOICELESS	BREATHY
m	2053	45	6
n	1742	35	6
ŋ	1217	27	1
ɲ	1064	25	2
l	1421	26	2
r	815	12	1

With regards to the diachronic behavior of breathy and voiceless sonorants, and as discussed more thoroughly in Section 2.4, voiceless sonorants are not reconstructed in Proto-Tibeto-Burman (DeLancey 2009, Matisoff 2003) but are reconstructed in some of its daughter nodes, e.g. Proto-Kuki-Chin (K. Van Bik 2009). Breathily voiced stops, meanwhile—often referred to as voiced aspirates in the historical and Indic traditions—are reconstructed in Proto-Indo-European, while breathy sonorants are not (Baldi 2009, Haider 1985, Kent 1932). Historical work on the development of breathy sonorants in Indic is profoundly sparse, but Hock (2010) reconstructs them in western dialects of Middle Indo-Aryan. If this reconstruction is correct then they are retained in Marathi and Konkani but have been lost from most of the daughter languages. Taken together, these facts suggest that voiceless sonorants have been stably present in the Tibeto-Burman language family for a very long time while breathy sonorants have largely disappeared in Western Indic—in sharp contrast with the long-term stable presence of breathy obstruents in Indic languages.

In laying out the Evolutionary Phonology framework, Blevins (2004, 2006a, 2006b) notes a contrast between "natural" patterns, which have functional, phonetic roots, and "unnatural" patterns, which emerge out of factors such as diachronic change and language contact. If the way in which breathy sonorants pattern typological and diachronically is influenced by functional, phonetic facts, then they would support the following prediction about the acoustic findings: breathy sonorants are expected to be less breathy in some way than their obstruent counterparts. Put differently, using the bilabials as representative symbols, the acoustic differences between [b<sup>h</sup>] and [b] are predicted to be greater than those between [m<sup>h</sup>] and [m].

To address the research questions posed above, acoustic analysis of data collected from ten native speakers of Marathi is presented. Marathi is an understudied Indic language that contains the 4-way obstruent contrast for which Indic languages are famous (e.g. /t, t<sup>h</sup>, d, d<sup>h</sup>/: plain, aspirated, voiced, and breathily voiced) as well as a two-way contrast among numerous sonorant pairs (e.g. /n, n<sup>h</sup>/) (Dhongde & Wali 2009). Measures reported herein include fundamental frequency (F0), corrected H1-H2 (henceforth indicated with an asterisk, H1\*-H2\*), H1\*-A1\*, H1\*-A2\*, and H1\*-A3\*. In the H-based measures, spectral tilt is assessed by comparing the amplitude of the first harmonic (H1) to that of the second harmonic (H2) or elements higher in the spectrum, e.g. in the range of the first formant (A1), second formant (A2), or third formant (A3). The presence of noise in the signal is quantified via Cepstral Peak Prominence (CPP). These measures are addressed more thoroughly in Section 2.2. What has been observed in previous studies is that breathiness has relatively predictable acoustic consequences, such that breathy sounds are expected to be associated with increased spectral tilt measures and increased noise. This is predicted to be true for sonorants as well. If there is a phonetic basis for the typological pattern, however, à la Blevins (2004, 2006a, 2006b), an interaction between Phonation Type and

Obstruency is expected such that the effect of phonation type is smaller in sonorant than in obstruent contexts. That is, for a given measure, the values observed for plain and breathy obstruents will diverge more than the values observed for plain and breathy sonorants.

## 2 Background

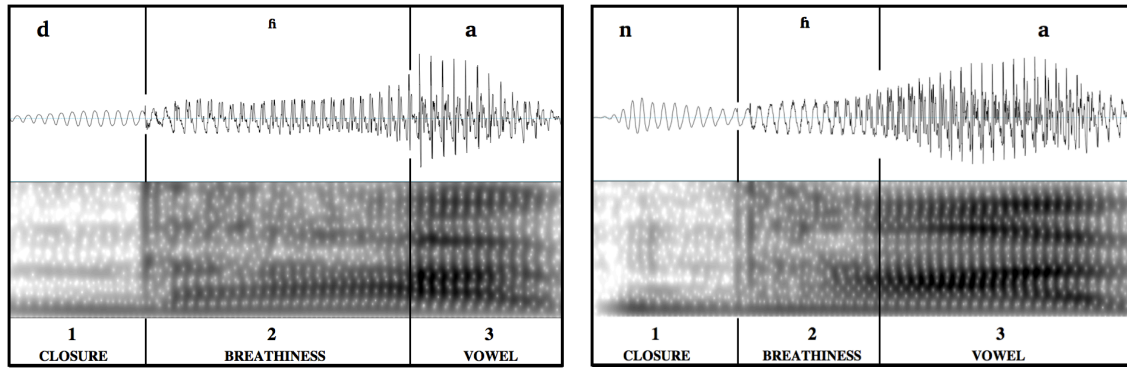
### 2.1 Production of Breathily Phonation

Some general observations about the articulatory settings for breathy phonation can be made, and for more thorough treatment see (among others) Laver (1980), Ní Chasaide and Gobl (1995), and Marchal (2009). Breathily voice tends to be produced with less tension in the vocal folds (VF) than modal voicing, which can lead to slower VF vibration and thus a lowered F<sub>0</sub>. There is also increased airflow, and the folds may never close completely. Catford's (1977: 99) description of the VF during breathily phonation as merely 'flapping in the breeze' conveys a sense of both the increased airflow and the loose vibratory mode characteristic of this phonation type (Catford 1977, Laver 1980, Marchal 2009).

As noted, phonemic breathiness is rare cross-linguistically but common in Indic languages, which are well-known for their breathily obstruents. The four-way oral stop contrast found in Marathi is typical of Indic phonetic inventories (Masica 1991). The breathily voiced sounds—also sometimes referred to as voiced aspirated or murmured (Schiefer 1987, Ladefoged & Maddieson 1996)—are characterized by vocal fold vibration during and after stop closure, and a breathily interval after the release. This post-release interval, referred to by Lisker and Abramson in their landmark 1964 article as “buzz mixed with noise” (p. 403), is illustrated in Fig. 1 with examples from the data recorded in the present study.

In Fig. 1a, the waveform and spectrogram of the initial syllable of the Marathi word [dʱaga], 'thread', has been divided into three intervals of interest. Interval 1 represents the stop closure; interval 2, labeled here as “breathily”, begins at closure release and represents the “buzz mixed with noise” referred to by Lisker and Abramson. Interval 3 represents the remainder of the vowel after the breathily begins to cease. Note that unlike the case of voiceless aspiration, where there is a lag between consonant constriction release and the onset of vowel periodicity, no such lag occurs in breathily voiced sounds because there is voicing throughout. Thus breathily and vowel periodicity overlap.

Fig. 1 also illustrates the marked similarity of breathily sonorants and breathily obstruents in Marathi. Fig. 1b shows the initial syllable from the Marathi word [nʱavi], 'barber', again taken from the data recorded for this study and broken into the same three intervals represented in Fig. 1a.: the oral constriction associated with the [n]; the breathily release, which persists for some time; and the remainder of the vowel after the breathily diminishes.



a. [d<sup>h</sup>a], excised from [d<sup>h</sup>aga], 'thread'

b. [n<sup>h</sup>a], excised from [n<sup>h</sup>avi], 'barber'

**Figure 1: Paired waveform and spectrogram of breathy voiced obstruent (left, 1a.) and sonorant (right, 1b.).** Three intervals of interest are illustrated here. Interval 1 represents the consonantal constriction; interval 2, labeled here as “breathiness”, begins at closure release and represents the “buzz mixed with noise” referred to by Lisker and Abramson. Interval 3 represents the remainder of the vowel after the breathiness begins to cease.

For the purposes of the present paper, the way in which the breathy interval relates to the articulatory settings described above can be noted: lower tension in the vocal folds and a greater degree of airflow during the production of breathy voice persists into at least some portion of the subsequent vowel. This is visually apparent in Fig. 1a. and 1b.: amplitude is lower during the portion of the vowel that is heavily intermixed with breath, and the spectrogram shows darkened formants only as the breathiness weakens. Acoustic data confirms the impressionistic observation: vowels after breathy obstruents differ acoustically from those following plain obstruents, showing increased spectral tilt and noise (Dutta 2007, Khan 2012).

## 2.2 Acoustic correlates of breathy phonation

As noted, little extant acoustic research on breathy sonorants exists beyond the Sumi (Harris 2009), Balami (Gautam 2012), and Tsonga (Traill & Jackson 1988) data mentioned above. Breathiness in vowels (Andruski & Ratliff 2000; Blankenship 1997; DiCano 2012; Esposito 2010, 2012; Garellek 2012; Kelterer 2017; Khan 2012; Wayland & Jongman 2003) and Indic obstruents (Davis 1994, Dutta 2007, Mikuteit & Reetz 2007, Ohala & Ohala 1972) has been the subject of ongoing phonetic investigation, however.<sup>5</sup> Given that this body of work indicates consistent acoustic consequences of breathiness in consonant and vowel contexts (namely, increased noise in the signal and increases in spectral tilt measures like H1-H2), it is relevant to review findings related to obstruents and vowels in the following sections.

### 2.2.1 Fundamental Frequency

Across many languages, F0 in vowels is lower after voiced than after voiceless stops (Caisse 1981, Mohr 1971, Ohde 1984). This effect is phonologized in some tone languages wherein voiced segments, sometimes called “depressor consonants” (Bradshaw 2000, Ohala 1973, Tang 2008), are positionally restricted and cannot co-occur with high tones (Pearce 2009, Perkins 2013). F0 is even lower after breathy voiced than after plain

<sup>5</sup> Herein, pathological breathy voice is considered distinct from phonemic breathy voice, though the pathological literature is referred to occasionally (particularly with regards to specific acoustic measures such as CPP).

voiced stops in Hindi (Dutta 2007, Ohala 1979, Pandit 1957, Schiefer 1986). This lowering may arise from the laryngeal configuration used in production of breathy voice, less tension in the vocal folds yielding a slower basic rate of vibration (Laver 1980, Ní Chaisade & Gobl 1995, Blankenship 1997, Dutta 2007). Gautam (2012) reports similar findings for plain and breathy sonorants in Balami. While it is not clear whether the reported F0 values are sonorant-internal or post-release, breathy sounds are consistently associated with F0 values ~20Hz lower than those reported for modal sounds. The trend may not be universal, however (Pennington 2005): breathy phonation is associated with lowered F0 in vowels in Itunyoso Trique (DiCanio 2012) but not in Jalapa Mazatec (Garellek & Keating 2011).

### 2.2.2 H1-based spectral measures

Breathy phonation tends to be characterized by a dominant first harmonic and a relatively rapid drop-off in energy moving higher in the spectrum (i.e., a sharp downward spectral tilt) (Pennington 2005). This is reflected in spectral measures like H1-H2, which relate the amplitude of the first harmonic (H1) to the amplitude of elements higher in the spectrum (such as H2, the second harmonic, or A1, the harmonic that is dominant in the range of the first formant). H1-based measures are used to study many aspects of voice quality/phonation, for example the interaction of tone and phonation types (Esposito 2012, Garellek & Keating 2012, Pittayaporn & Kirby 2017), register (DiCanio 2009), age- and sex-based differences in voice quality (Iseli, Shue & Alwan 2007; Lee et al. 2016), glottalization (Garellek 2014, supplemental files; Pan 2017), ‘sweet voice’ in Japanese (Starr 2015), and voice qualities resulting from vocal pathology (Kumar, Bhat & Mukhi 2011; Narasimhan & Vishal 2017).

Breathy phonation, whether phonemic or non-phonemic, is generally associated with greater values in H1-based measures than modal phonation in languages like Balami (Gautam 2012), Chichimeco (Kelterer 2017), English (Garellek 2012), Gujarati (Khan 2012), Hindi (Dutta 2007), Hmong (Esposito 2012, Garellek 2012, Huffman 1987), Jalapa Mazatec (Garellek & Keating 2011), Khmer (Wayland & Jongman 2003), Korean (Jun & Cha 2015, Winter & Grawunder 2012), Marathi (Berkson 2013), and Shanghai Chinese (Cao & Maddieson 1992, Gao et al. 2011, Gao 2015, Tian & Kuang 2016.).

H1-H2 values are greater for breathy than for modal phonation by about 7 dB in Hmong (Huffman 1987), and successfully distinguish breathy from modal vowels in Green Mong (Andruski & Ratliff 2000) and in White Hmong (Esposito 2012, Garellek 2012) as well as in Gujarati (Esposito 2006, Khan 2012), Jalapa Mazatec (Garellek & Keating 2011), and Khmer (Wayland & Jongman 2003). Esposito (2010) presents acoustic measurements of modal and breathy vowels produced by male talkers in ten languages (Chong, Fuzhou, Green Mong, White Hmong, Mon, Santa Ana del Valle Zapotec, San Lucas Quiavini Zapotec, Tlacolula Zapotec, Tamang, and !Xóõ) and reports that H1-H2 distinguishes breathy from modal vowels in eight of the ten languages. Consonant phonation type contrasts also trigger reliable differences in H1-H2 values in subsequent vowels in Hindi (Dutta 2007). Sonorant-internally, values in Balami appear to follow the expected trend of increased H1-H2 in breathy sonorants as compared with plain sonorants (Gautam 2012). Wu dialects of Chinese contrast plain and aspirated voiceless stops with voiced stops (e.g. /p p<sup>h</sup> b/). The voiced stops, referred to as “slack-voiced” by Gao et al. (2011), are not reliably produced with voicing but are associated with a breathy release (Cao & Maddieson 1992, Gao et al. 2011, Gao 2015) that has been described as a weaker version of Hindi’s breathy voiced stops (Karlgren 1915-1926 as cited in Cao & Maddieson 1992). In the Wu dialect spoken in Shanghai, H1-H2 values are greater after the slack-voiced stops than elsewhere (Gao et al. 2011). There are age and gender effects, however:

breathiness is strongest for older male speakers and weakest for younger speakers, suggesting that breathy phonation as a correlate of the slack-voiced stops may be weakening or disappearing (Gao 2015).

Breathy phonation in obstruents and vowels is associated with greater H1-A1 values than modal phonation in Gujarati (Khan 2012). H1-A1 also distinguishes modal from breathy vowels in Chichimeco (Kelterer 2017), White Hmong (Garellek 2012), Jalapa Mazatec (Garellek & Keating 2011), Khmer (Wayland & Jongman 2003), and in three of the ten languages in Esposito's (2010) cross-language study. Breathily voiced stops in Hindi are associated with increased H1-A1 values in subsequent vowels (Dutta 2007), as are slack-voiced stops in Shanghai Chinese (Gao 2015).

Increased H1-A2 values associated with breathy phonation are also expected, and are reported for Mazatec (Blankenship 2002, Garellek & Keating 2011), Gujarati (Khan 2012), and five of the ten languages assessed in Esposito (2010). As compared with plain voiced stops, breathily voiced stops trigger increased H1-A2 values in subsequent vowels in Hindi (Dutta 2007) as do slack voiced stops in Shanghai Chinese (Gao 2015).

H1-A3 distinguishes breathy from modal vowels in Gujarati (Khan 2012) and in seven of the ten languages in Esposito (2010). While breathiness is expected to induce steeper spectral tilt, however—and therefore increased values for H1-based measures—results for H1-A3 vary: in work which preceded Esposito (2010) and excluded one of the ten languages analyzed therein (San Lucas Quiavini Zapotec), Esposito (2006) reports that H1-A3 values are greater for modal sounds in five languages but greater for breathy sounds in the remaining four. Kelterer (2017) reports similar within-language variability in Chichimeco, which contains independent tonal and phonation-type contrasts (Herrera Zendejas 2014) akin to those found in Jalapa Mazatec (Garellek & Keating 2011). In a study of four speakers, two male and two female, breathy phonation is reliably associated with increased H1-A3 values for the female speakers but not for the male speakers (Kelterer 2017). In Hindi, meanwhile, although mean H1-A3 values in subsequent vowels differ based on consonant phonation type for between 50 and 70% of the subsequent vowel, Dutta (2007) reports the presence of both inter- and intra-speaker variability.

What of the Sumi, Balami, and Tsonga data? Harris (2009) does not directly compare spectral measures for vowels following plain and breathy sonorants in Sumi; rather, contrastive sequences like [la] and [l<sup>h</sup>a] are subjected to several comparisons. Spectral slices from the midpoint of the [l] in [la] and the [l] component of [l<sup>h</sup>a] are generated and compared, with results indicating a high degree of similarity between the two in all measures reported. The procedure is repeated to compare the [l] (or consonantal) and the [h]-like (or breathy) portions of [l<sup>h</sup>a], but vowels following modal and breathy consonants are not compared. Observations can nevertheless be made about the way in which the breathy [h]-like portion differs from what precedes it. This is not a perfect comparison for our purposes, but given the paucity of existing research on the topic it is better than discarding the data altogether. Of note is that the H1-H2 values appear to be greater in the breathy than in the consonantal portion of the breathy sequences for the alveolar nasal and for the lateral, but not for the bilabial nasal. Harris's general, overall observation is that the breathy portion of the sequence does not robustly show the expected spectral correlates of breathiness (2009: 77). This contrasts with Gautam's (2012) findings for Balami, which aligned with the obstruent and vowel-related findings of increased spectral tilt associated with breathy sonorants. Again, however, the Balami data should be considered with caution. Compounding the problem of sketchy methodological

description, observations made about spectral tilt are largely descriptive and are for only one speaker.

In Traill and Jackson's 1988 investigation of Tsonga, measures of spectral tilt were calculated by hand based on spectra generated for the nasal consonants and the first 40 ms of the subsequent vowel. The measures taken were H1 minus the amplitude of the most intense harmonic around 1400 Hz and H1 minus the amplitude of the most intense harmonic above 2000 Hz. While there was some inter-speaker variation, the measure of H1 minus the amplitude of the harmonic nearest 1400 Hz consistently distinguished plain and breathy nasals for males and females. The result obtained only nasal-internally, however; the first 40 ms of the subsequent vowel did not show the same reliable differentiation.

Given these results, it is perhaps unclear what to expect with regards to Marathi sonorants. The Hindi data indicate that obstruent phonation type triggers differences in subsequent vowels (Dutta 2007). In Tsonga vowels after sonorants do not show the same differences (Traill & Jackson 1988), but this may reflect the differences between Bantu and Indic breathiness more than anything else. The Tibeto-Burman data, meanwhile, does not necessarily help clarify. The Balami and Sumi studies—both of which are limited in ways that leave comparisons difficult to make—sometimes offer conflicting observations. What is clear is that where consonant phonation type triggers differences in spectral values, breathiness should be associated with larger positive values than modal phonation. These differences may not appear as robustly or as reliably as they do in post-obstruent and vowel contexts, however.

### 2.2.3 *Measures of noise*

Increased noise associated with breathiness has been found using various harmonics-to-noise ratio (HNR) measures. Values for these measures are expected to be lower for breathy than for modal sounds due to increased airflow. This is true in vowels in Javanese (Wayland, Gargash, & Jongman 1994) and in Jul'hoansi (Miller 2007). The trend was not found for Chantaburi Khmer (Wayland & Jongman 2003), where values were generally higher for breathy than for plain values but differences attained statistical significance for only one out of five speakers. Meanwhile, differences in HNR between plain and breathy vowels in Gujarati were significant only at vowel midpoint (Khan 2012).

Cepstral Peak Prominence (CPP) is a normalized measure of harmonic amplitude that takes overall amplitude into account. During the production of breathy phonation, the vocal folds may never close completely, resulting in at least some amount of continuous airflow and increased aspiration noise in the signal as compared with modal voicing (Balasubramaniam *et al.* 2011, Catford 1977, Marchal 2009). This increased noise is visible in the cepstrum: strong peaks in the cepstrum are an indicator of less noise in the signal, a clear harmonic structure, and a very periodic sound (Keating & Esposito 2007), while the noise present in breathy phonation results in weaker cepstral peaks. Breathily sounds are thus expected to have lower CPP values than sounds produced with modal voice, a finding which has been reported for non-phonemic breathiness (Hillenbrand *et al.* 1994) as well as breathiness that results from vocal pathology (Heman-Ackah *et al.* 2003, Watts & Awan, 2011). Use of CPP to measure phonemic breathiness is becoming more common, and lowered CPP values associated with breathy phonation have been reported for Mazatec (Blankenship 2002) and Chichimeco (Kelterer 2017). CPP is also the only measure which distinguishes plain from breathy vowels in all ten of the languages assessed in Esposito (2010). In Gujarati, however, CPP values are lower in breathy than in modal

vowels, as expected, but differences are marginal overall and are significant only in the middle portion of the vowel (Khan 2012).

### 2.3 *Timecourse of breathiness*

The temporal characteristic noted above is worth highlighting. In contrasting plain and breathy vowels in Gujarati, acoustic differences are maximal in the middle of the vowel rather than at either edge (Khan 2012). The situation in Gujarati is complex: it contains breathy consonants and breathy vowels, meaning that there is a three-way contrast between modal consonant-modal vowel, breathy consonant-modal vowel, and modal consonant-breathy vowel sequences (e.g. [ba], [b<sup>h</sup>a], [b̤a]). This sort of three-way contrast is very rare, but also occurs in White Hmong, and Esposito and Khan (2012) investigated the timecourse and acoustic correlates of breathiness in Gujarati and White Hmong. The authors hypothesized that breathiness associated with consonants would be maximal at constriction release/vowel onset, because it is associated phonologically with the consonant, while breathiness associated with vowels would be more pervasive and would be localized later in the timecourse of the vowel. Both predictions were borne out.

As might be expected given the literature reviewed above, breathiness in both Gujarati and White Hmong—whether belonging to the consonant or to the vowel—is associated with increased spectral tilt measures and increased noise. The timecourse of breathiness differs, however: in both languages, breathy consonants are realized with a short period of intense breathiness in the earlier part of the vowel which decreases over the remainder of the vowel. This was heavily localized to the first 20% of the vowel in White Hmong, but was more generally localized to the first half of the vowel in Gujarati. Breathless vowels in both languages, meanwhile, showed a lower degree of breathiness than vowels after breathy consonants, and exhibited either stable breathiness throughout the first half of the vowel (White Hmong) or increasing breathiness after vowel midpoint (Gujarati) (Esposito & Khan 2012).

To summarize, breathy obstruents in Hindi trigger acoustic differences in subsequent vowels which persist through much or all of the timecourse of the vowel (Dutta 2007). Post-release differences associated with phonation contrasts in Tsonga nasals are not found (Traill & Jackson 1988). Breathiness in Gujarati—a close relative of Marathi—is complex, because the language contains breathy consonants and vowels and the two differ in terms of both degree and timecourse of breathiness. Vowels after breathy consonants have intense breathiness which peaks at the 40% mark (for H1\*-H2\* and H1\*-A3\*) and in the 40-60% interval (for CPP) (Esposito & Khan 2012: 13).

Given this ensemble of facts it is unclear what to expect for Marathi, and so no formal predictions with regards to timing of breathiness are made in Section 3.3 (Predictions) below. If Marathi patterns with its close relatives Hindi (Dutta 2007) and Gujarati (Esposito & Khan 2012), vowels after breathy consonants are expected to show the spectral and noise-related characteristics of breathiness and the effect is expected to persist across much or all of the vowel. Because Marathi does not have breathy vowels, however, there is no need to differentiate breathy vowels from vowels following breathy consonants, so although the breathiness in Marathi is phonologically associated with the consonant there is no language-internal motivation to restrict maximal breathiness to the early portion of the vowel. In fact, if the Tsonga data are an indicator that sonorants do not trigger post-release correlates of breathiness in the same way that obstruents do (Traill & Jackson 1988), then acoustic cues to breathiness after sonorants may be diminished, absent, or non-maximal in the early portion of the vowel.

## 2.4 Marathi

Marathi is well suited for investigation of breathy voice due to its rich phonetic inventory. Spoken by more than 80 million people, primarily in the state of Maharashtra (2011 India Census), Marathi—like its Indic relatives—features a four-way obstruent contrast consisting of plain voiceless, aspirated voiceless, plain voiced, and breathy voiced categories (i.e. /t, t<sup>h</sup>, d, d<sup>h</sup>/). The breathy voiced obstruents, traditionally referred to as voiced aspirates, are generally reconstructed back to Proto-Indo-European but are retained most reliably in the Indic languages (Baldi 2009, Haider 1985, Hock 2010, Kent 1932, Kobayashi 2004). Marathi diverges from many of its Indic counterparts by extending phonemic breathy voice into the sonorants, though Konkani (spoken directly to the south of Marathi) is a notable exception as it also retains breathy sonorants (Almeida 1989, Miranda 2007).

A full treatment of the diachronic origin of these breathy sonorants does not exist, to the best of my knowledge, but they may have arisen via a series of developments in the shift from Old Indo-Aryan to Middle Indo-Aryan (IA). The relevant facts are as follows. A regular /h/-nasal metathesis process in Middle IA that turned Sanskrit *cihna-* ‘mark’ into Pāli *cinha*, for example (Cardona & Luraghi 2009: 374), resulted in sonorant-/h/ sequences. Metathesis also operated over sequences where an /h/ was derived from lenition of \*s, such that Sanskrit \**snā-* ‘bathe’ becomes *nha-* in Middle IA Prakrit (Hock 2010: 92). Hock (2010), with reference to Turner (1962-1966) and von Hinüber (1986), argues that the resulting sonorant-/h/ sequences were treated differently in different branches of Middle IA, with Western dialects of Middle IA (from which Marathi descends) treating them as aspirated sonorants. This claim is strongly supported by the fact that they behave as a single segment (e.g. as [n<sup>h</sup>]) in early Pāli poetry. Masica (1991: 170), in describing innovations from Old IA to Middle IA, also notes that “the new aspirated nasals mh, nh/ṅh [might be] unit phonemes (rather than clusters), since they occur initially (at least in several forms of Prakrit, barely in Pali) where otherwise only single consonants are permitted.” Masica is cautious, however, emphasizing that such sequences occur more commonly in medial than in initial position in Middle IA, where they may be better analyzed as sequences. He concludes that the “evidence is conflicting” (p. 174) with regards to the phonemic status of breathy sonorants in Middle IA.

Setting questions about their exact origin aside, Masica (1991) lists the breathy sonorants as phonemes in the inventories of several New Indo-Aryan languages, including Marathi. In line with Masica (1991) and Dhongde and Wali (2009), and contra Pandharipande (1997), breathy sonorants are included in the Marathi consonant inventory which appears in Table 2. Bloch (1970), Kavadi and Southworth (1965), and Southworth (2000) were also consulted in constructing this inventory, additional commentary about which is as follows. The voiceless aspirated labial stop, /p<sup>h</sup>/, is undergoing a fairly widespread change to /f/ in Marathi and in other Indic languages.<sup>6</sup> The retroflex sibilant which appears in parentheses in Table 2 is what Dhongde and Wali call a “marginal phoneme”—it appears only in Sanskrit words and is most often produced as a postalveolar/palatal [ʃ]. The retroflex consonants differ from the other consonants in at least two ways: (i) there is no breathy voiced retroflex lateral consonant in Marathi; and (ii), with the exception of the velar nasal, which can only occur in coda position, all consonants can serve as onsets except for the retroflex nasals /ŋ, ŋ<sup>h</sup>/ and lateral /l/.

---

<sup>6</sup> In the data recorded for this study, the male speakers produced [f] almost exclusively. The female speakers produced [p<sup>h</sup>] occasionally, but [f] dominated.

**Table 2: Marathi consonant chart**

		Place of Articulation													
		Labial		Dental		Alveolar		Retroflex		Post-alveolar/Palatal		Velar		Glottal	
Manner of Articulation	Stop	p p <sup>h</sup>	b b <sup>h</sup>	t t <sup>h</sup>	d d <sup>h</sup>			t̪ t̪ <sup>h</sup>	ɖ ɖ <sup>h</sup>			k k <sup>h</sup>	g g <sup>h</sup>		
	Nasal		m m <sup>h</sup>		n n <sup>h</sup>				ɳ ɳ <sup>h</sup>				ŋ		
	Affricate					ts	dz dz <sup>h</sup>			tʃ tʃ <sup>h</sup>	dʒ dʒ <sup>h</sup>				
	Fricative					s		(ʂ)		ʃ					h
	Approximants		ʋ ʋ <sup>h</sup>		l l <sup>h</sup>				ɭ		j				
	Rhotic										r r <sup>h</sup>				

With the facts outlined above in mind, and because the current study focuses on the acoustic correlates of phonation type of obstruents and sonorants in syllable onset position, the labial, dental, and post-alveolar/palatal voiced sounds serve as the target consonants herein. They are highlighted with boxes in Table 2. The palatal approximant /j/ does not have a breathy counterpart, and so it was excluded.

Marathi contains six major monophthongs, as presented in Table 3; three additional monophthongs (/æ ɜ ɔ/) which occur only in borrowed words are omitted here (Dhongde and Wali 2009:9, Ghatage 2013, Yardi 1998).

**Table 3: Marathi Monophthongs**

	Front	Central	Back
High	i		u
Mid	e		o
Low-mid		ə	
Low		a	

Dhongde and Wali (2009:10) note that vowel nasalization is allophonic, not phonemic, and that—contra Pandharipande (1997)— phonemic length distinctions are not retained. Long versions of /i/ and /u/ remain as positional variants in final syllables, and Marathi orthography still represents length distinctions, but these distinctions are neutralized in spoken language.

### 3 Methods

#### 3.1 Materials

To investigate the acoustic correlates of breathy phonation in obstruents and sonorants, this study focused on the voiced labials, dentals, and postalveolar/palatals in Marathi

(highlighted with boxes in Table 2). This subset of consonants was selected in order to allow for the comparisons that are central to the present work: namely, modal vs. breathy phonation and obstruents vs. sonorants. Selection was also influenced by the phonotactics of Marathi—retroflexes were excluded because there is no breathy voiced retroflex lateral and the retroflex sonorants are disallowed in onset position.

In order to enhance generalizability, target sounds included sonorants from multiple places and manners of articulation occurring in multiple word-positions (initial and medial). Multiple vowel contexts were used to account for the fact that harmonic amplitudes may be influenced by formant frequencies and bandwidths. Accordingly, the stimuli list developed for the present study includes target consonants before the low vowel [a] and before the higher vowel [e]. In the [a] context, F1 is greatest and the acoustic correlates of phonation type are expected to be strongest. Language-internal morphological facts guided the selection of the higher vowel context, as several regular morphological processes involve an [a]~[e] alternation. For example, Marathi nouns inflect for three grammatical genders (masculine, feminine, neuter), and singular masculine forms often end in /a/ while neuter forms end in /e/. Masculine nouns ending in /a/ are pluralized by changing the /a/ to an /e/ [e.g. [dʌba] *box* vs. [dʌbe] *boxes*]. The full stimuli list is in Appendix A.

### 3.2 *Speakers*

Participants included ten native speakers of Marathi, five female (mean age 37 yrs) and five male (mean age 45 yrs). All were multilingual, speaking English and Hindi or other Indic languages in addition to Marathi, and resided in Lawrence, Kansas or in the Kansas City area at the time of recording. Demographic background data is presented in Appendix B.

### 3.3 *Corpus*

Participants produced target words embedded in a carrier sentence. The stimuli (see Appendix A) included a total of 64 words: 16 consonants x 2 word positions (initial and medial) x 2 vowel contexts (before [a] and before [e]). Five of the targets were nonce words created to fill accidental gaps in the lexicon. The voiceless obstruents were also recorded but are excluded here, in order to focus on sonority effects. Thus, the total number of tokens analyzed is 1920 (10 speakers X 16 consonants X 2 word positions X 2 vowel contexts X 3 repetitions).

### 3.4 *Procedure*

Speakers were recorded onto a flash disc using an Electro-Voice N/D 767 cardioid microphone and a Marantz Portable Solid State Recorder (PMD 671) in the anechoic chamber at the University of Kansas, Lawrence. Half of the participants were recorded at a sampling rate of 22.05 kHz, and the other half at 44.1 kHz, both with 16-bit resolution.

### 3.5 *Annotation and Measures*

Sound files were annotated in PRAAT (Boersma & Weenink 2010), where textgrids were used to demarcate the vowel in each syllable of interest.<sup>7</sup> The breathy

---

<sup>7</sup> A reviewer correctly noted that a comparison of segment-internal measures in plain and breathy sonorants would be valuable, given that little work on this topic exists. I strongly concur. However, as the question investigated here centers on the interaction between phonation type contrasts and obstruency, the focus is on segment-external measures (that is, measures taken in the following vowel) because the same measures can be taken after sonorants and after obstruents.

interval in breathy sounds was included as part of the vowel, meaning that the intervals marked as 2 and 3 in Fig. 1a and 1b were analyzed as a single vocalic interval. VoiceSauce (Shue *et al.* 2011), a free application implemented via Matlab® (The Mathworks, 2011), was used to measure F0, CPP, H1\*-H2\*, H1\*-A1\*, H1\*-A2\*, and H1\*-A3\* in the vowel after each consonant of interest. F0 was measured using the Straight algorithm developed by Kawahara *et al.* (1998). Harmonic amplitudes—H1, H2, A1, A2, and A3—are computed in VoiceSauce across a four-pitch-period window. A pitch-synchronous window is used to avoid the variability that may result from a fixed window. These amplitude measures may be affected by vowel formants and thereby rendered inaccurate, and so the algorithm developed by Iseli & Alwan (2004) is used to correct for the effect of formants on spectral amplitude measures. The resulting corrected measures are those denoted by an asterisk (H1\*-H2\*). CPP is calculated using the algorithm from Hillenbrand *et al.* (1994). Again, a variable window is used, this one equal to five pitch periods.<sup>8</sup>

The output generated by VoiceSauce included measurements calculated at every millisecond of the demarcated intervals, and these were averaged into five equal portions to allow for presentation and analysis of time-normalized data. These averaged windows represent 0-20% of the segment (Window 1), 21-40% of the segment (Window 2), 41-60% of the segment (Window 3), 61-80% of the segment (Window 4), and 81-100% of the segment (Window 5). 0% marks consonantal constriction release and 100% marks the end of the vowel. The first and last 12 ms of each vowel were excluded (as they were calculated across a four-pitch-period window, they included some information from the preceding or following segment).

Data were analyzed using R (R Core Team, 2012). The code used for statistical analysis and can be found at <https://osf.io/u2ghp/>.

### 3.6 Research Questions and Predictions

Several predictions can be made about the research questions addressed in this work, which are repeated here for convenience.

- (1) Are the acoustic correlates of breathy voice in sonorants *qualitatively* different than they are in obstruents and vowels?
- (2) Are the acoustic correlates of breathy voice in sonorants *quantitatively* different than they are in obstruents and vowels?

Regarding (1), the acoustic correlates of breathiness in sonorants are predicted to align with those found in obstruents and vowels: F0 and CPP values are expected to be lower, and the H-based spectral tilt measures (H1\*-H2\*, H1\*-A1\*, H1\*-A2\*, H1\*-A3\*) greater, for breathy than for plain sounds. This prediction is based on the holistic observation drawn from previous studies of phonation contrasts in sonorants that differences, when found, generally align with the results found for obstruents and vowels.

The predictions for question (2) are made with both the previous literature and the typological and diachronic facts laid out in the introduction in mind. Breathily obstruents and sonorants in Marathi are realized with the prototypical Indic interval of intense post-release breathiness, and as such they are expected to pattern like Hindi (Dutta 2007) and Gujarati (Esposito and Khan 2012): contra Tsonga, wherein sonorant phonation type did

---

<sup>8</sup> Data were spot-checked for accuracy. While formant measures are not reported here, taking formant measures of breathy sounds in Marathi proves challenging for VoiceSauce (and for Praat) because the breathiness renders formant structure difficult for the tracking algorithms to identify. No similar difficulty arose with the measures reported herein.

not trigger differences in subsequent vowels (Traill & Jackson 1988), consonant phonation type in Marathi is expected to trigger acoustic differences that persist across much or all of the subsequent vowel. Traill and Jackson's findings do raise questions about how robustly the correlates of breathiness are realized in post-sonorant contexts, however, and the Sumi (Harris 2009) and Balami (Gautam 2012) data raise questions about how reliably sonorant phonation contrasts trigger the expected acoustic consequences of breathiness. The typological scarcity of breathy sonorants and their susceptibility to diachronic loss further motivate a prediction that the acoustic correlates of phonation type are more robustly realized in obstruent contexts than in sonorant contexts (that is, [b<sup>h</sup>-b] is predicted to exceed [m<sup>h</sup>-m]).

### 3.7 Analysis

The analysis was designed to probe the interaction of key interest here, which is between the variables PhonationType (breathy vs. plain) and Obstruency (obstruent vs. sonorant). Furthermore, because of the dynamic nature of breathiness (Khan 2012) and questions about how differences arising from consonant phonation type are realized in subsequent vowels, the three-way interaction between PhonationType, Obstruency, and Window (which out of the five portions of the vowel the measurement was averaged over) was also considered.

These effects were estimated using linear mixed effect models (Baayen, Davidson, & Bates, 2008).<sup>9</sup> Each measure was first sphered to have a mean of 0 and a standard deviation of 1; this was done to help the models converge more easily. All predictors were centered (PhonationType was coded with -0.5 for plain and 0.5 for breathy; Obstruency was coded with -0.5 for sonorant and 0.5 for obstruent)<sup>10</sup> so that both main effects and interactions could be observed in the model coefficients. After this, models were fit regressing each measure on PhonationType, Obstruency, Window, and all interactions between these variables. Several additional covariates (speaker sex [male or female], vowel context [e or a], position of the consonant within the word [initial or medial], place of articulation of the consonant [labial, alveolar, or postalveolar/palatal], and repetition number [1, 2, or 3]) were included but were not permitted to interact with the variables of interest, in order to determine whether observed effects are robust enough to emerge across diverse contexts (e.g. places of articulation, vowel contexts). By-subject random effects for the intercept and the variables of interest (PhonationType, Obstruency, Window, and their interactions) and random correlations were also fitted; a maximal random effects structure (Barr et al., 2013) was not used because this would require fitting a very large number of parameters.

Six measures were tested, and each had two parameters that could be taken as support for the predictions (the three-way interaction between PhonationType, Obstruency, and Window, and the two-way interaction not including Window). Accordingly, the alpha level was set to .05/12. One-tailed tests were used for the two-way interactions. PhonationType was coded with Plain as the baseline level, so positive or negative coefficients could be predicted based on the known patterns for breathiness in each measure. Obstruency was coded with Sonorant as the baseline level. Expected behavior, and expected polarity of

---

<sup>9</sup> Code used for statistical analysis is available at <https://osf.io/u2ghp/>.

<sup>10</sup> Therefore, the coefficient for PhonationType would represent how much higher plain is than breathy, in the given measure; a positive coefficient would indicate plain>breathy and a negative coefficient would indicate plain<breathy. Likewise, the coefficient for Obstruency would indicate how much higher obstruent is than sonorant, with a positive coefficient indicating obstruent>sonorant and a negative coefficient indicating obstruent<sonorant. The interpretation of the interaction coefficients is explained below.

interaction terms, is outlined here and recapitulated in Table 4. With regards to PhonationType, breathy consonants are expected to be associated with lower F0 and CPP values (and therefore negative coefficients are predicted) and with greater values than plain consonants for all of the H1-based measures (and therefore positive coefficients are predicted).

With regards to Obstruency, differences between breathy and plain obstruents are expected to exceed differences between breathy and plain sonorants. Thus, for measures where a negative coefficient is predicted for PhonationType (e.g., we expect F0 to be lower after breathy consonants than after plain consonants), a negative interaction term would indicate that the negative-going effect is even more negative-going in obstruent contexts than in sonorant contexts. Table 4 presents an overview of expected behavior and interaction term polarity. Voiced labial stops and nasals are used as representative symbols.

**Table 4: Predicted Findings and Interaction Term Polarity**

MEASURE	EXPECTED BEHAVIOR BY PHONATIONTYPE	EXPECTED BEHAVIOR BY OBSTRUENCY	INTERACTION TERM SHOULD BE:
F0 AND CPP	breathy < plain	$[b^h-b] > [m^h-m]$	negative ( $< 0$ )
H1*-H2*, H1*-A1*, H1*-A2*, H1*-A3*	breathy > plain	$[b^h-b] > [m^h-m]$	positive ( $> 0$ )

For measures where a negative coefficient is predicted for PhonationType (meaning that breathy sounds are associated with lower values) we predict a negative interaction term. For measures where a positive coefficient is predicted for PhonationType, we predict a positive interaction term.

Two-tailed tests were used for the three-way interaction since specific predictions for how the effect might change over the course of the syllable were not made.

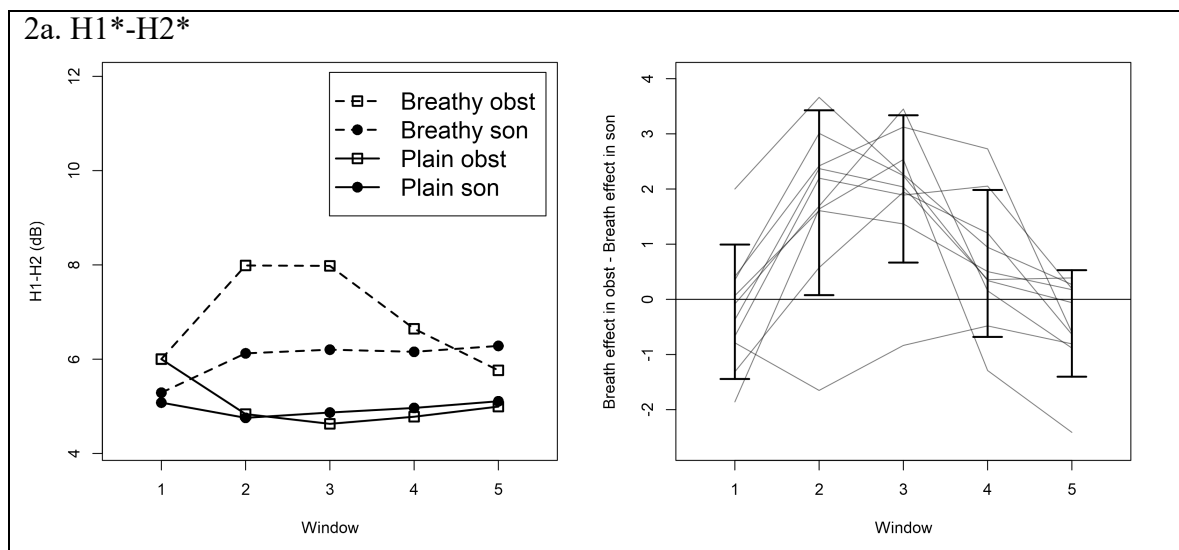
## 4 Results

Figure 2 shows the pattern of differences, and the timecourse of this pattern, for each measure. Mean values for each acoustic measure are shown for the timecourse of vowels (five windows) following each of the four consonant types under investigation (plain voiced obstruents, breathy voiced obstruents, plain sonorants, and breathy sonorants). The overall pattern is illustrated in the left panel by plotting mean values in vowels after the four consonant types; lower values are expected after breathy than after plain sounds for CPP and F0, and greater values are expected after breathy sounds for the remaining spectral measures. For all measures, differences between plain and breathy obstruents are expected to exceed differences between plain and breathy sonorants (that is, as delineated in Table 3, the value resulting from  $[b^h - b]$  is expected to exceed the value resulting from  $[m^h - m]$ ). Overall, the predicted patterns seem to emerge in all measures except F0, particularly around the vowel midpoints. The right panel of each plot illustrates how robust the pattern is or is not, across subjects.

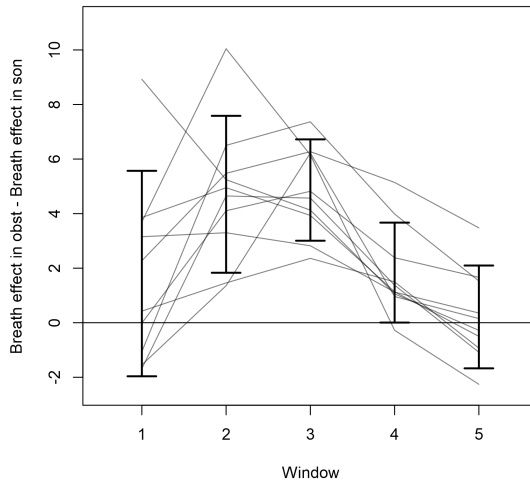
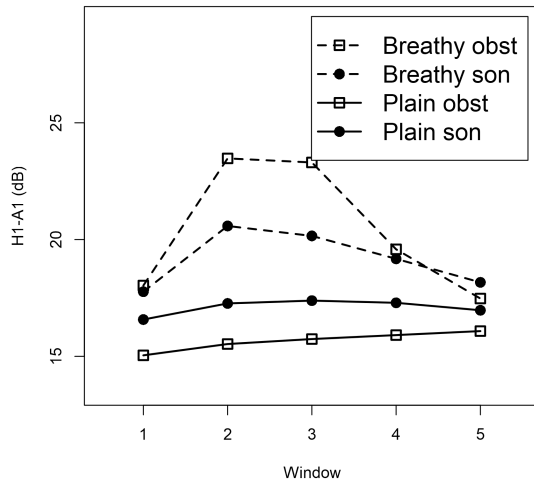
These plots are described more thoroughly with reference to Figure 2a, which shows results for H1\*-H2\*. The left panel of 2a. shows that H1\*-H2\* values in vowels after breathy consonants (dashed lines) are greater than those in vowels after plain consonants (solid lines), as expected. More importantly, this breathy vs. plain difference (dashed vs. solid lines) is bigger after obstruents (open squares) than it is after sonorants (filled circles). There is a *difference between the differences*, in other words: while breathy consonants are associated with greater H1\*-H2\* values than plain consonants overall, the difference

between plain and breathy obstruents exceeds the difference between plain and breathy sonorants. It is this *difference-between-differences* that is plotted for individual speakers' data in the right panel of 2a. Each line in 2a. represents one speaker and shows the difference-between-differences value at each window, calculated by subtracting the difference resulting from {mean H1\*-H2\* after breathy sonorants minus mean H1\*-H2\* after plain sonorants} from the difference resulting from {mean H1\*-H2\* after breathy obstruents minus mean H1\*-H2\* after plain obstruents}. The formula used to calculate the *difference-between-differences* can be illustrated more simply as follows, again utilizing the voiced labial symbols as representative: {[b<sup>h</sup>-b] - [m<sup>h</sup>-m]}. For measures where post-breathy vowels are expected to have lower values than post-plain vowels (e.g., CPP and F0), this difference-between-differences is expected to be negative if the pattern is stronger in obstruents than sonorants: [b<sup>h</sup>-b] is predicted to yield a large negative number for such measures, [m<sup>h</sup>-m] is predicted to yield a small negative number, and a large negative number minus a small negative number yields a negative number. For measures where post-breathy vowels are expected to have higher values than post-plain vowels (e.g., the H1-based measures), this difference-between-differences is expected to be positive: [b<sup>h</sup>-b] is predicted to yield a large positive number, [m<sup>h</sup>-m] is predicted to yield a small positive number, and a large positive number minus a small positive number yields a positive number. Therefore, a positive difference-between-differences (for spectral tilt measures) or a negative difference-between differences (for CPP or F0) indicates that obstruents displayed greater acoustic differentiation than sonorants, at least numerically. A value near zero indicates that obstruents and sonorants patterned similarly.

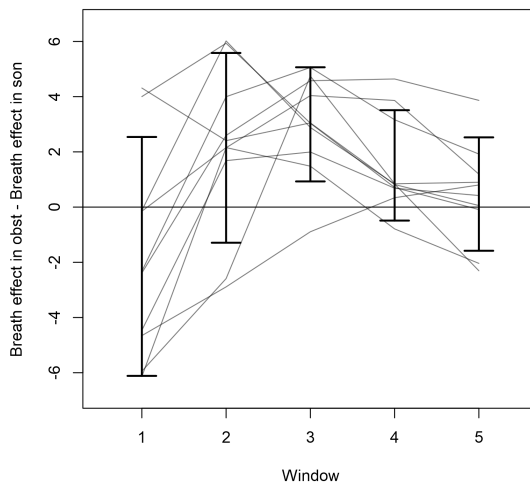
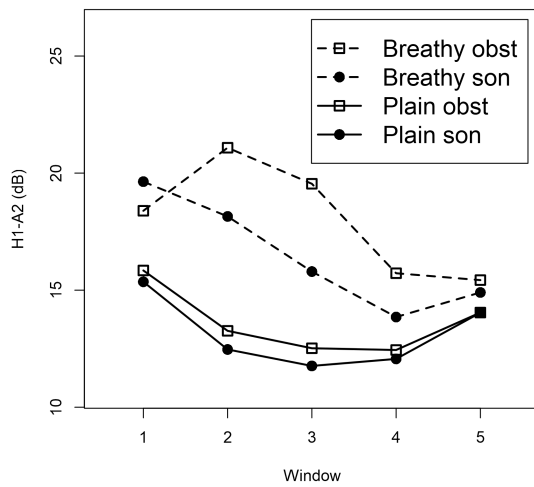
Again consider H1\*-H2\* as an example. H1\*-H2\* was predicted to be higher for breathy than for plain sounds, so the expectation is for each of the differences (i.e., breathy mean minus plain mean in obstruents, and breathy mean minus plain mean in sonorants) to be positive numbers. Since the contrast is also expected to be more robust (i.e., the difference is expected to be greater) in obstruents than in sonorants, then the difference between breathy obstruents and plain obstruents should be larger (more positive) than the difference between breathy sonorants and plain sonorants. The difference-between-differences is therefore also expected to be positive. The right panel of Fig. 2a. shows that this is the overall pattern observed. It also makes clear that the effect is mainly driven by the middle portion of the vowel, in line with Khan's (2012) findings for Gujarati vowels.



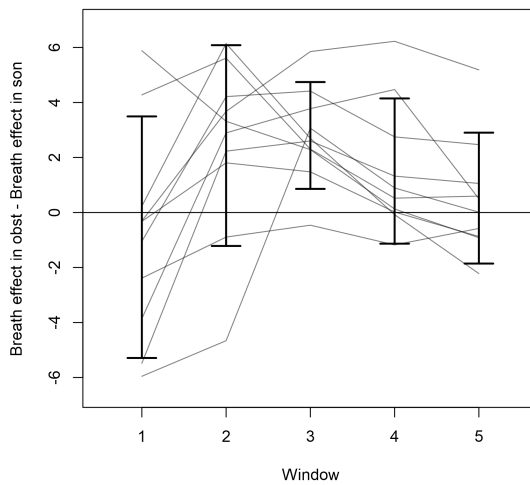
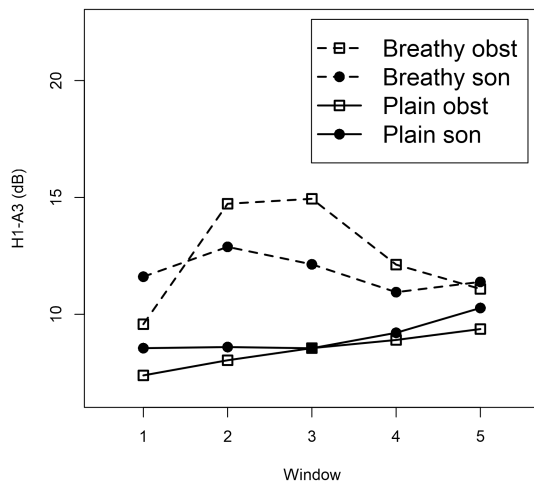
2b. H1\*-A1\*

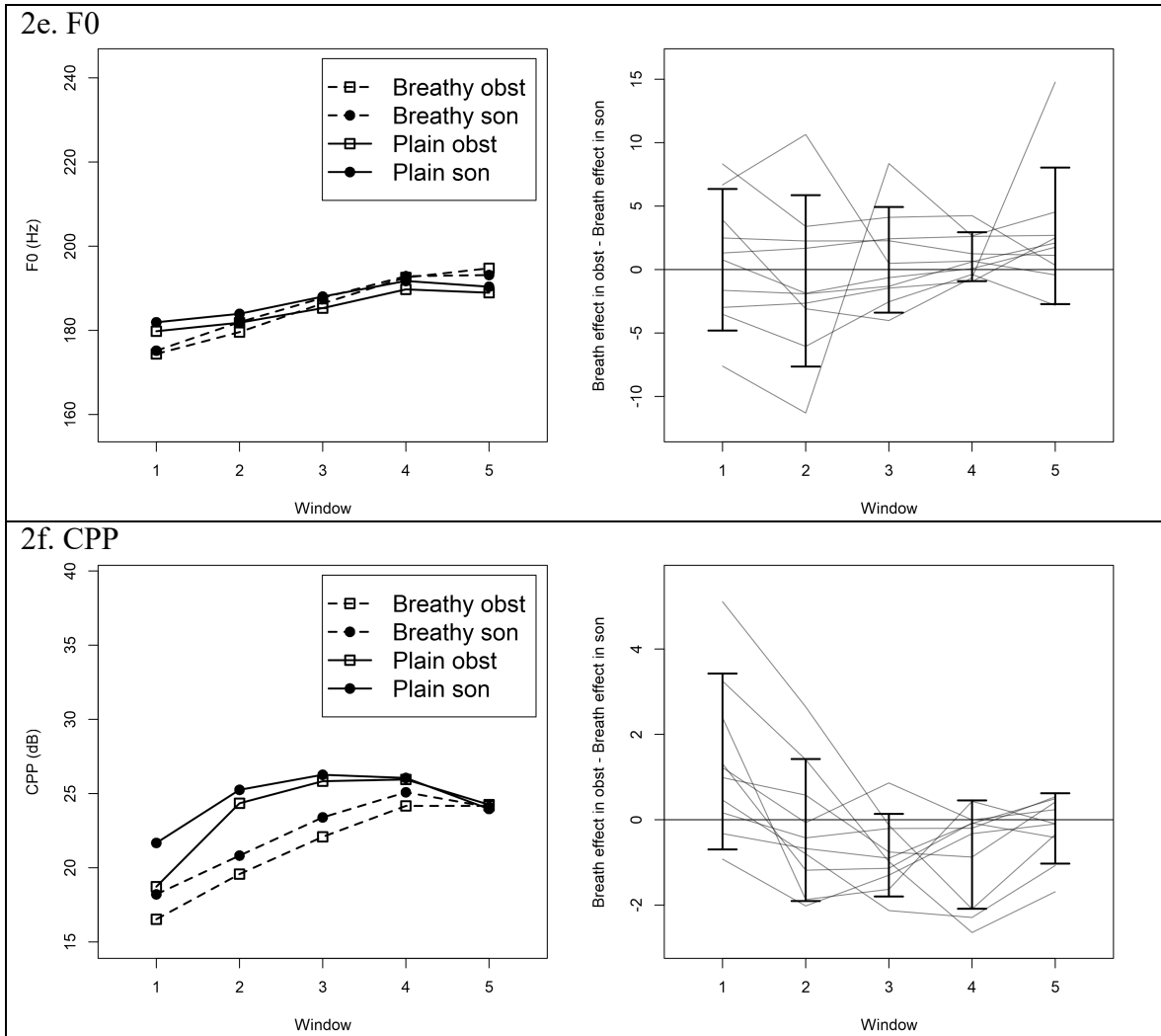


2c. H1\*-A2\*



2d. H1\*-A3\*





**Figure 2: Timecourse of Acoustic Measures in Vowels by Obstruency and PhonationType (left panel) and Robustness of Pattern across Subjects (right panel)**

Left panels present mean values plotted over the timecourse of the vowel following each of the four consonant types: plain voiced obstruents (solid line, open squares), breathy voiced obstruents (dashed line, open squares), plain sonorants (solid line, filled circles), and breathy sonorants (dashed line, filled circles).

The right panel in each plot presents the difference-between-differences for individual talkers. Each line represents one subject and shows the sum that results from subtracting the {mean after breathy sonorant minus mean after plain sonorant} from the {mean after breathy obstruent minus mean after plain obstruent}. Using the labials as representative symbols, this difference-between-differences can be conceptualized as  $\{[b^h-b] - [m^h-m]\}$ .

For the HI-based measures (2a-d), each of these differences (i.e., mean after breathy obstruent minus mean after plain obstruent, and mean after breathy sonorant minus mean after plain sonorant) are expected to be positive numbers, because breathy consonants are expected to be associated with greater values than plain consonants. The contrast is expected to be more robust (i.e., breathy and plain are expected to diverge more) in obstruent than in sonorant contexts, so the difference-between-differences (or  $\{[b^h-b] - [m^h-m]\}$ ) is expected to be positive and to fall above the zero-line.

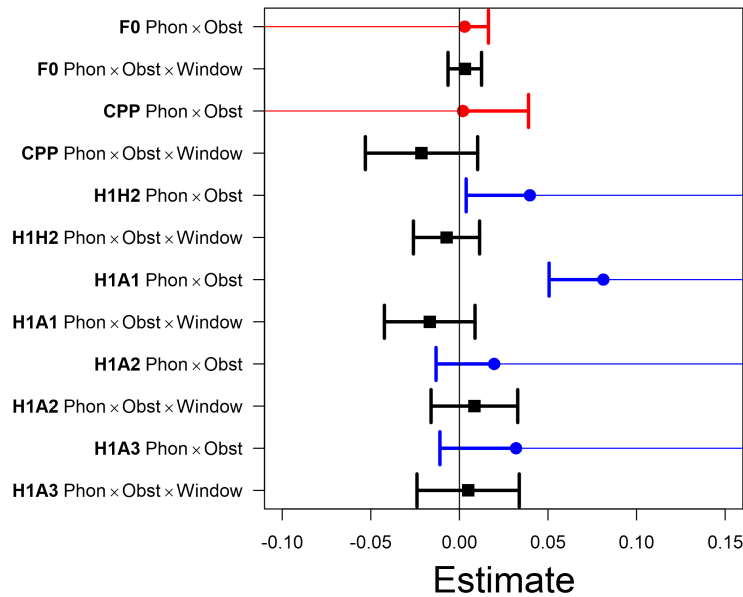
For F0 and CPP (2e-f), differences are expected to be negative because breathy consonants are expected to be associated with lower values than plain consonants. The contrast is expected to be more robust (i.e., with greater acoustic differentiation) in obstruent contexts than in sonorant contexts, the difference-between-differences is expected to be negative and to fall below the zero-line.

Overall, the expected patterns were supported in the planned statistical analysis. First, as a simple manipulation check, I examined whether the effects of phonation were in the direction predicted for each measure. For F0 and CPP (data in 2e. and 2f), vowels following breathy segments were expected to have lower values overall and therefore have negative coefficients (see section 3.7 and footnote 10 for an explanation of the coefficient coding in the regression model); this was the case for cepstral peak prominence ( $\beta=-0.34$ ), but only negligibly so for F0 ( $\beta=-0.01$ ). For H1-based measures (data in 2a-d), vowels after breathy segments were expected to have higher values (and positive coefficients), and this pattern held (H1\*-H2\*:  $\beta=0.13$ ; H1\*-A1\*:  $\beta=0.20$ ; H1\*-A2\*:  $\beta=0.20$ ; H1\*-A3\*:  $\beta=0.20$ ).

Of greater interest in the statistical analysis were interactions between phonation type and obstruent/sonorant status, as reflected by the difference-between-differences which were shown in the right panels of Figure 2. In Figure 3, point estimates and confidence intervals for each interaction of interest are shown. The point estimate is the model's estimate of the difference-between-differences for that measure (described above), which is expected to be negative for F0 and CPP and positive for the other spectral measures. The confidence intervals are best interpreted as the range of values from which the point estimate does not differ significantly; e.g., a confidence interval from 0.5 to 1.0 would indicate that the estimate is significantly bigger than 0.5 and significantly smaller than 1.0, but not significantly different from any value inside this range. Since the question of theoretical interest is whether the predicted effect (that the breathy-plain difference is bigger in obstruents than sonorants), I focus on measures where the estimate is significantly different from *zero*, i.e., where the confidence interval does not include zero. Furthermore, in cases where there was a directional prediction (described above), I use one-tailed confidence intervals, which correspond to one-tailed statistical tests. These confidence intervals extend to infinity in the other direction, since a one-tailed test cannot make claims about statistical significance in the unpredicted direction (e.g., if one predicts a difference to be above zero [positive] and performs a one-tailed test, one can only prove that the test result is significantly bigger than zero, not significantly below zero). For F0 and CPP I predicted a negative difference-between-differences, and thus use one-tailed confidence intervals that show what the estimate is significantly lower than but not what the estimate is significantly higher than. For spectral tilt measures I predicted a positive difference-between-differences, and thus use one-tailed differences that show the opposite (what the estimate is significantly higher than, and not what the estimate is significantly lower than).

First consider the results without taking Window into account (i.e., averaged across the vowel duration). Since I had directional predictions for each of these differences-between-differences (described above), each of these is shown with a one-tailed confidence interval. Spectral tilt measures seem to be the most consistent with the prediction that breathy-plain differences are greater in obstruents than sonorants. There are significant two-way PhonationType\*Obstruency interactions (confidence intervals not including zero) in H1\*-H2\* and H1\*-A1\*, while in H1\*-A2\* and H1\*-A3\* there are near-significant trends in the expected direction.

The basic prediction is borne out, then: breathy (as compared with plain) segments in Marathi are associated with decreased CPP values and with increased values in H1-based measures, and PhonationType and Obstruency interact in the spectral tilt measures.



**Figure 3: Point Estimates and Confidence Intervals for Interactions of Interest** *Point estimates and 99.6% (1-.05/12) confidence intervals for the two-way (PhonationType x Obstruency) and three-way (PhonationType x Obstruency x Window) interactions of interest. Intervals are red for coefficients where a negative sign was predicted, blue where a positive sign was predicted, and black where there was no directional prediction.*

Now consider the interactions with Window. The interactions discussed above represent differences-between-differences (i.e., the breathy-plain difference in obstruents is bigger than the breathy-plain difference in sonorants). The interactions that include Window represent a difference between those differences-between-differences, meaning that the difference described above is bigger at the end of the vowel than at the beginning of the vowel, bigger in the middle of the vowel than at the beginning or end, and so forth. The difference is not static, in other words, holding steady across the entirety of the vowel, but changes over time in some way. Since I did not have a directional prediction about these three-way interactions (i.e., no *a priori* expectation that the predicted difference-between-differences would be largest at any particular part of the vowel), I tested these with two-tailed confidence intervals, as shown in Figure 3. None of these crosses zero, indicating that none of the three-way interactions with Window were significant in the planned analysis—i.e., for H1\*-H2\* and H1\*-A1\*, the breathy-plain difference is significantly bigger in obstruents than sonorants, but this pattern does not significantly change over the course of the vowel.

The failure of interactions with Window to reach significance may have been because of the non-monotonic nature of this effect (meaning that the nature of the effect is not in one consistent direction, e.g. neither steadily increasing nor steadily decreasing but rather increasing and then decreasing). Consider again the results for H1\*-H2\*. As shown in Figure 3 there was a non-significant negative-going three-way interaction (PhonationType\*Obstruency\*Window) for this measure. As explained above, the PhonationType\*Obstruency interaction for this measure is expected to be positive—i.e., with greater values associated with vowels following breathy consonants than with vowels following plain consonants. This difference is also predicted to be greater in obstruent than in sonorant contexts. If the PhonationType\*Obstruency interaction's interaction with

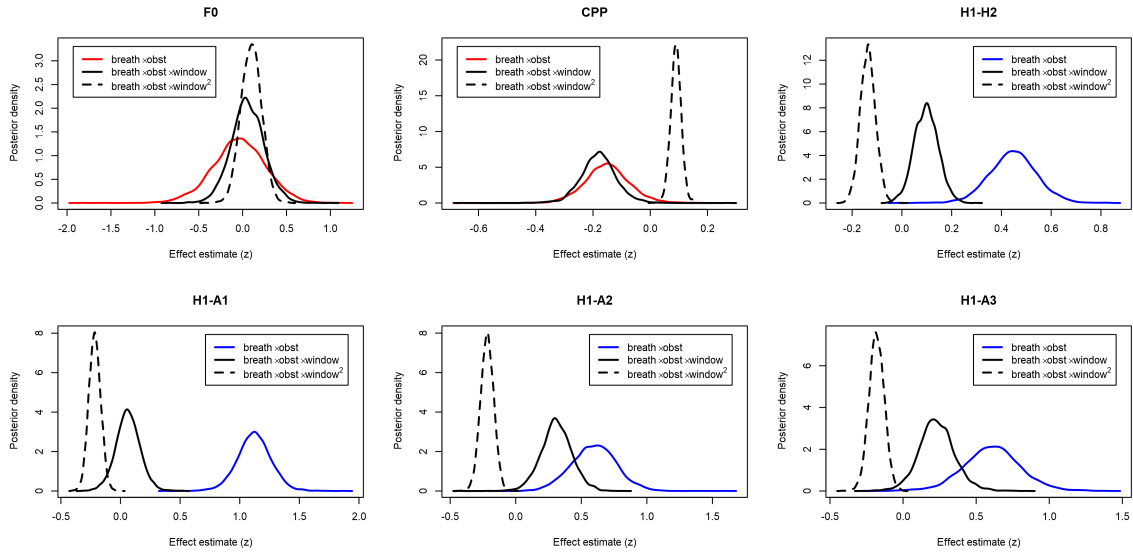
Window is numerically negative, that means that the positive PhonationType\*Obstruency interaction gets smaller as the vowel progresses. This is visible in Fig. 2b: while Window 1 breaks this trend, the PhonationType\*Obstruency interaction is indeed biggest at Windows 2 and 3 and then decreases going through Windows 4 and 5. It is an oversimplification of the pattern visible in Figure 2 to state that the effect decreases across the duration of the vowel, however. In fact, across all the spectral measures, post-obstruent differences appear to be small at vowel onset (Window 1) but are maximal and exceed post-sonorant differences at vowel mid-point (Window 3). This effect is non-monotonic (meaning that it increases and then decreases again), and thus additional modeling with polynomial terms (allowing the model to fit a curved line to the data) would be necessary to statistically evaluate it. The original statistical model did not attempt to fit the polynomial terms that would be necessary in order to capture the non-monotonic interactions with Window, because this pattern was not predicted. That original model, though, may not be capturing the data well, because it is essentially attempting to fit a straight line through this inverted U-shaped effect. Therefore, I present the results of an exploratory analysis below. The exploratory analysis uses a new model which includes a quadratic term for Window, which may better capture the way that the interaction effect initially increases and then decreases.

#### 4.1 Exploratory Analysis

As *p*-values and confidence intervals are not strictly meaningful for an exploratory (as opposed to confirmatory) analysis, instead of showing confidence intervals like in Figure 3 I instead summarize the model using Bayesian posterior densities, shown in Figure 4.<sup>11</sup> The model specification for the results shown here is available at <https://osf.io/u2ghp/>. The primary consequence of adding the quadratic term into the model is that the two-way PhonationType\*Obstruency interactions look likely to be significant (i.e., most of their posterior densities, representing where the true effect probably lies, are very far away from zero) for all spectral measures. For example, for H1\*-H2\*, almost all of the posterior density is above zero, meaning that there is almost no chance (based on the data, the prior, and the modeling assumptions) that the true effect size for the interaction in the population is zero. Likewise, the three-way interactions with quadratic terms (those shown with dotted lines, PhonationType\*Obstruency\*Window<sup>2</sup>) look likely to be significant for all measures except F0. The two-way interaction for CPP also looks like it would come close to being significant. These results support the trend observed in Figs. 2 and 3: phonation-type contrasts appear to be more robustly cued in obstruent than in sonorant contexts, and, based on the significant interactions with the window terms, this pattern is stronger in some parts of the vowel than others (particularly, as we see from Figure 2, it is strongest near the center of the vowel).

---

<sup>11</sup> For ease of comparison with Fig. 3, a supplementary figure with confidence intervals for this exploratory model is also shown at <https://osf.io/u2ghp/>. Conclusions about statistical significance should not be made based on such intervals; these are merely presented for ease of exposition. This supplementary figure is consistent with Figure 4 shown here: the confidence intervals suggest that the PhonationType\*Obstruency two-way interactions and PhonationType\*Obstruency\*Window<sup>2</sup> three-way interactions, for all measures except F0, would be significant, or close to significant, if this had been a planned confirmatory analysis.



**Figure 4: Bayesian Posterior Densities for Interactions of Interest in the Exploratory Analysis** Each curve shows the posterior density of the effect, i.e., where the effect probably lies in the population; the higher the curve, the higher probability that the true effect lies at that point. These roughly correspond to confidence intervals: a significant effect will have a peak far away from zero, and at zero its posterior density will be very small. As in Fig. 2, effects where a negative coefficient was predicted are shown with red curves, and those where a positive coefficient was predicted are shown with blue curves; effects without a directional prediction are shown with black curves. The three-way interactions without a quadratic term are shown with a solid black line. The three-way interactions with a quadratic term, meaning those that allow for a non-monotonic effect, are shown with a dashed black line.

To summarize, the results for the H-based measures confirm that consonant phonation type in Marathi obstruents and sonorants indeed affects subsequent vowels, as Dutta (2007) found for Hindi. More importantly, the effect of phonation type is larger after obstruents than it is after sonorants. The effect of breathiness, as well as the difference-between-differences, is generally greatest at vowel midpoint, and is significant or near-significant, as shown in Figure 3 and also supported by the exploratory analysis shown in Figure 4.

The simplest prediction for CPP was that values after breathy sounds would be lower than after plain sounds, and this is confirmed. The PhonationType\*Obstruency interaction for CPP is both non-significant and numerically in the unexpected direction, however. This is shown in Figure 3, although in the exploratory analysis shown in Figure 4 this interaction is in the expected direction and is probably close to significant. In fact, in the exploratory analysis most of the three-way interactions that included a quadratic term were likely significant. Predictions about how the effect of breathiness would change over time were not made, but these data suggest that the effect of breathiness in Marathi is non-monotonic, as was also visible in the plots in Figure 2. The acoustic consequences of breathy consonants in Marathi tend to grow stronger towards the midpoint of subsequent vowels before decreasing again. The timecourse of breathiness in Marathi is therefore similar to that reported by Khan (2012) for Gujarati, albeit with a caveat. Recall that in Gujarati breathiness is contrastive in both vowels and voiced obstruents, and that breathiness associated with consonants is maximal earlier in the vowel, near the consonant release, while breathiness associated with vowels is pervasive throughout the vowel and

increases over time (Esposito & Khan 2012). Marathi consonants pattern like Gujarati vowels, then, in that the effect of consonant breathiness not maximal at vowel onset, but instead increases and is maximal at vowel midpoint. This may suggest that because Marathi does not have contrastive breathiness in vowels it is not crucial for breathiness to be temporally adjacent to the consonants with which it is phonologically associated. As for F0, this measure did not appear to vary based on breathiness in any reliable way. Differences in F0 values in vowels following breathy vs. plain sounds were negligible, and no meaningful interactions emerged.

## 5 Discussion

The results presented in Section 4 reveal that phonation type contrasts in Marathi consonants trigger acoustic differences in subsequent vowels, that the differences persist across much or all of the vowel, and that differences related to consonant phonation type are often maximal at vowel midpoint. This was most notable for the spectral tilt measures. The first research question addressed herein asked whether the acoustic correlates of breathy voice in sonorants are *qualitatively* different than they are in obstruents and vowels. We can now state that the answer is no: breathiness in sonorants, like in obstruents and vowels, is associated with increases in H1-based spectral measures and in noise (as reflected by lower CPP values). F0 values are numerically lower, which is the expected direction, but these differences do not attain statistical significance.

The second research question asked whether the acoustic correlates of breathy voice in sonorants are *quantitatively* different than they are in obstruents and vowels, and again the answer to this question is yes. Breathily sonorants trigger the expected acoustic patterns, but the differences are reduced (or less robust) in sonorant contexts as compared with obstruent contexts. The extent to which they are less robust—e.g., whether this attenuation generalizes across vowel contexts, word positions, and speaker sex, or whether the difference is driven by certain contexts—is a matter for future investigation.

The prediction that was made based on the typological and diachronic patterning of breathy sonorants was that the differences between plain and breathy obstruents would exceed the differences between plain and breathy sonorants (e.g. {[b<sup>h</sup>-b] > [m<sup>h</sup>-m]}), and this too was borne out. Typologically, it is of course the case that grammatical systems are influenced by many factors. Non-optimal contrasts exist; synchronic inventories are influenced by, for example, production and perception limitations, cognitive and historical factors, and language contact (de Boer 2000). Nevertheless, functional explanations can account for at least some of the attested patterns. Blevins (2004, 2006a, 2006b), in laying out the framework of Evolutionary Phonology, notes a distinction between "natural" patterns arising from functional roots and "unnatural" patterns arising from diachronic change, language contact, and so forth. The current finding that phonation type contrasts are more robustly encoded by obstruents than by sonorants suggests that the typological profile of breathy sonorants falls into the former category, arising naturally from functional phonetic roots.<sup>12</sup> These same functional, phonetic roots presumably also contribute to the diachronic pattern. As noted by Wright (2004):

---

<sup>12</sup> I do not claim that functional forces are the only ones relevant. Genetic relationships are also involved—breathily obstruents (but not sonorants) were passed from Proto-Indo-European on to many daughter languages, while breathily sonorants did not develop until much later. Discussion of this and other factors contributing to the typological pattern are outside the scope of this paper.

In an ideal setting, there is no background noise or distractions, and the listener is so riveted by what the talker is saying that he/she gives the signal undivided attention. Under normal conditions it is rare for speech to occur in the absence of at least some form of environmental masking. What this means for speech is that a robustly encoded phonological contrast is more likely to survive signal degradation or interference in reception.... (Wright 2004:42)

Of the contrasts investigated herein, the obstruents (e.g. /b/ vs. /b<sup>h</sup>/) are the more robustly encoded contrasts. As such they are more resistant to the cumulative effects of signal degradation, and therefore less vulnerable to diachronic loss.

As has been pointed out, voiceless sonorants do not pattern like breathy sonorants. Rather, they have been stably present in the Tibeto-Burman language family for a very long time and appear not to be subject to the same diachronic loss that affects breathy sonorants. Blevins (2018) raises an intriguing point by suggesting that voiceless sonorants may be “contextually salient” due to the very paucity of their segment-internal acoustic information: to wit, their salience may hinge on “contrasting with surrounding voiced sounds...giving an overall whispered effect” (p. 32). Perception experiments will be needed to probe this proposal, but if it is indeed the case that voicelessness in sonorants can be contextually salient then breathy sonorants may be at a disadvantage precisely because they do not feature an extreme change in voicing. This suggestion is unintuitive, because breathy sonorants contain segment-internal acoustic information in addition to post-release, segment-external information. The data presented here, however, suggest that the post-release information does not render the distinction between plain and breathy sounds as acoustically robust in sonorants as in obstruents.

A question may be raised as to how direct the link is between the acoustic patterns reported here and the perceptual consequences thereof. This question remains unanswered at present, although pilot data reported in Berkson (2016) are suggestive. In a simple identification task focused on the dental consonants of Marathi ([d] vs. [d<sup>h</sup>], [n] vs. [n<sup>h</sup>], and [l] vs. [l<sup>h</sup>]), seven native Marathi listeners heard CV syllables (e.g. [da] vs. [d<sup>h</sup>a]) and were asked to identify the consonant they heard. Stimuli were presented in randomized blocks: one block contained only the [d] and [d<sup>h</sup>] stimuli, another only the [n]/[n<sup>h</sup>] stimuli, and a third only the [l]/[l<sup>h</sup>] stimuli. Participants correctly identified all plain consonants with similar accuracy rates (~90%). The breathy obstruent [d<sup>h</sup>] was identified correctly 93% of the time, but accuracy was lower for the breathy sonorants: 78% for [n<sup>h</sup>] and 74% for [l<sup>h</sup>] (Berkson 2016). If this pattern holds in a larger study, then the indication is clear: phonation contrasts in sonorants are subject to more perception errors than similar contrasts in obstruents.

Moreover, with regards to the nasals in particular, previous work has established that nasal consonants may be inherently more breathy than other types of consonants. Garellek et al. (2016) investigate the interaction of breathiness and nasality, noting a diachronic connection between breathy and nasal sounds in some languages which previous researchers have proposed is due to the perceptual similarity of breathiness and nasality. Garellek and colleagues analyze data from three languages in the Yi branch of Lolo-Burmese (Bo, Luchun Hani, and Southern Yi). These languages have a tense/lax contrast, wherein vowels produced with lax phonation are breathier than vowels produced with tense phonation. Acoustic and electroglottographic investigation of nasal consonants, lateral consonants, and oral and nasal vowels reveals that breathier phonation is often found in nasal consonants (as compared with lateral consonants). Moreover, vowels after nasals

are often breathier than vowels after laterals. In Luchun Hani, lax vowels (which involve phonologically specified breathy phonation) are breathier than tense vowels overall. There is a split within lax vowels, too, however: lax vowels are breathier after nasal consonants than after lateral consonants. These findings raise questions about the breathy nasals in Marathi. If the nasal consonants of Marathi pattern like those of the Yi languages analyzed by Garellek et al. (2016), vowels after plain nasals may be slightly breathy to begin with. If so, the diminished difference between vowels after plain and breathy nasals may result not only from weaker breathiness in post-sonorant contexts (as compared with post-obstruent contexts) but also from a higher baseline level of breathiness. The diminished acoustic differentiation in post-sonorant contexts may result from two factors, in other words: vowels after [m] may be *more* breathy than vowels after [b], and vowels after [m<sup>h</sup>] may be *less* breathy than vowels after [b<sup>h</sup>]. In the left panel of Figure 2b, which shows mean H1\*-A1\* values, plain sonorants are slightly higher (i.e., slightly breathier) than plain obstruents across the entire duration of the vowel. The same observation can be made about the beginning of the vowel in Fig. 2d.: with regards to mean H1\*-A3\*, vowels after plain sonorants are slightly breathier at vowel onset than vowels after plain obstruents. Future work analyzing the nasals separately from the approximants will help prove (or disprove) whether the idea outlined above has merit.

Another question that can be raised has to do with why voiceless sonorants in Tibeto-Burman languages pattern quite differently than breathy sonorants. The historical paths leading to the formation of phonemically voiceless sonorants—and presumably breathy sonorants as well—are profoundly limited (Blevins 2018). Representing segments with a glottal spread gesture with *H* and sonorants with *R*, Blevins notes that voiceless sonorants arise mainly in two ways: from coarticulation between *HR* and *RH* sequences, or due to final devoicing (see Blevins 2018 p. 39-46). As described in Section 2.4, Hock (2010) proposes that breathy sonorants arose in Western Middle Indic via coarticulation: metathesis yielded *RH* sequences, and the coarticulated sequence was then reinterpreted as a single segment. Reanalysis of coarticulated sonorants as voiceless segments was also one source of voiceless sonorants in TB, for example after prefixation of historical affixes like \**s*- and \**ʔ*- to sonorant-initial roots. In terms of the development of non-modal sonorants in Western Middle Indic and Tibeto-Burman, then, the segmental processes involved were similar. Why, then, do voiceless sonorants remain robustly attested in TB while breathy sonorants have all but disappeared from most western Indic languages, being retained primarily in Marathi and Konkani?

What may be relevant here is that while the segmental processes contributing to the presence of non-modal sonorants in the two language families may be similar, there is a grammatical component in Tibeto-Burman that is absent in Indic. Specifically, some TB voiceless sonorants derive from historical prefixes with grammatical content. One relatively straightforward example is the \**s*- causative prefix from Proto-Tibeto-Burman.<sup>13</sup> This is illustrated in Table 5 with data drawn from three Tibetan languages: Themchen, a conservative dialect of Amdo Tibetan (AT) (Haller 2004); Lhasa, an innovative dialect of Central Tibetan (CT) (Chang 1971); and Derge, a dialect of Kham Tibetan (KT) which is

---

<sup>13</sup> As noted by a reviewer, nonmodal sonorants do not have identical sources in all TB languages and are realized differently across TB languages: they are voiceless in some languages, creaky in others, and breathy in still others. Furthermore, the processes that yielded nonmodal sonorants in some languages have very different reflexes in other languages (e.g. tonal contrasts, breathiness or murmur associated with the syllable, and so forth). Variation across TB languages, which is intensely interesting and remains understudied, lies beyond the scope of the present paper.

intermediary between the innovative and conservative dialects (Bielmeier et al. 2008).<sup>14</sup> The historical form is preserved orthographically in Written (or Classical) Tibetan. In conservative Themchen, the causative is variably produced either with an *s*-nasal cluster or with a non-modal nasal. Derge and Lhasa are representative of a pattern found in a number of tonal dialects of Tibetan, where high tone is associated with historically voiceless singleton consonants (both aspirated and unaspirated) and low tone with historically voiced singleton consonants (Sun 1997; Tournadre & Dorje 2003). In intermediary Derge, the historical *s*- is lost and the causative is produced with both a nonmodal nasal and high tone. (Forms are written here as they appeared in the source material; the superscript [h] is used here to indicate voicelessness, following Bielmeier *et al.*) In Lhasa, which is representative of the most innovative dialects of Tibetan, the voiceless sonorants are not retained but the tonal reflex of the \**s*- (high tone associated with voiceless consonants) is present.

**Table 5: Tibetan \**s*- causative correspondences**

LANGUAGE	NON-CAUSATIVE	CAUSATIVE	SOURCE
Written Tibetan	<i>nyal</i> ‘lie down’	<i>s-nyol</i> ‘cause to lie down’	
AT-Themchen <i>conservative</i>	ɲa	ɣ̥ɲu ~ ɲu	(Haller 2004)
KT-Derge <i>intermediary</i>	ɲɛ: (L)	ɲ <sup>h</sup> ɛ: (H)	(Bielmeier et al. 2008)
CT-Lhasa <i>innovative</i>	ɲɛ: (L)	ɲɛ: (H)	(Chang 1971)

The pattern illustrated in Table 5 is not restricted to Tibetan. Voiceless sonorants, for example, are widely attested in TB, appearing in languages from the Himalayish, Qiangic, Lolo-Burmese, Nungish, and Kamarupan supgroups (Matisoff 2003: 37). In Hakha Lai, for example, we find pairs like [mit] ‘go out (*light*)’ ~ [m̥it] ‘extinguish (*light*)’ and [rɪl] ‘roll (*intransitive*)’ ~ [r̥ɪl] ‘roll (*transitive*)’ (data from K. Van Bik 2009). Hakha Lai is a Central Chin language from the Kuki-Chin subgroup, quite distant from the Tibetan languages, but as noted by Mortenson (2013: 281) these pairs “appear to reflect stems with and without the historical \**s*- causative prefix, respectively”.

The crucial point is that at least some of the voiceless sonorants in TB languages retain grammatical saliency as residue of the prefixes from which they derive (K. Van Bik, p.c.). If the developmental path proposed by Hock (2010) is correct, the same cannot be said for the breathy sonorants in Marathi. The diachronic source for these sounds was a phonological process of metathesis. A grammatical element akin to that of the \**s*-causative was absent, in other words. Thus there may be a grammatical component to the stability of voiceless nasals in Tibeto-Burman languages.

The grammatical origin of some of the Tibeto-Burman voiceless sonorants also speaks to the general matter of opportunity: a similar phonological environment could yield nonmodal sonorants in both TB and in Indic, but this environment could be created by a regular grammatical process only in Proto-Tibeto-Burman (PTB). Any sonorant-initial

<sup>14</sup> Volume 2 of the invaluable *Comparative Dictionary of Tibetan Dialects* (CDTD) by Bielmeier et al., which focuses on verbs, was published very recently and I do not yet have access to it. Volume 1 is not yet available. I have worked with (and cited) the preprint, which has been available since 2008.

stem to which an appropriate prefix could attach yielded such an environment in PTB, while the environments where metathesis could yield breathy sonorants in Middle Indic were presumably more limited. Is there a way to probe this point more rigorously, however? Phonotactic frequencies in modern languages may provide one source of information—as pointed out by a reviewer, if the environments where breathy sounds could develop were limited then such sounds are not expected to be frequent. An analysis of phonotactic frequencies in the Marathi portion of the EMILLE/CIIL (Enabling Minority Language Engineering/Central Institute of Indian Languages<sup>15</sup>) corpus suggests that this is true: breathy sonorants represent fewer than 1% of the consonants in the 2.2-million word corpus, whether calculated by type (0.3%) or by token (0.9%) frequency (Berkson & Nelson 2017). Plain sonorants, meanwhile, represent 46% of the consonants in the corpus by type and 45% by token frequency. The corpus is drawn from various written sources, and while analysis of a spoken corpus would unquestionably be more informative the data from EMILLE are valuable because no comparable spoken corpus yet exists and because the Devanagari writing system used for Marathi is relatively phonetically transparent.

To the best of my knowledge no published works on the phonotactic frequency of voiceless sonorants in a modern TB language exist. Some clues about language-internal representation of voiceless sonorants can be gathered, however. These data are valuable because I am unaware of any existing alternative, but they should be interpreted with caution because the data sources differ dramatically from the corpus-based frequencies reported for Marathi. Consider Burmese, which contains voiceless counterparts of the nasal and lateral phones (Watkins 2001). Analysis of the first 10,000 lines of the Burmese version of Wikipedia reveals that plain nasals and laterals account for approximately 16% of the onset consonants therein while voiceless nasals and laterals account for approximately 4% (F. M. Tyers, p.c.). This distribution is closely echoed in data I culled from the searchable online database affiliated with STEDT, the *Sino-Tibetan Etymological Dictionary and Thesaurus* (Matisoff 2016). Of the 10,000 written Burmese entries that appear in a basic search of STEDT, 17% begin with a plain nasal or lateral and 7% with a voiceless nasal or lateral.<sup>16</sup> Of the 1310 Hakha Lai lexical entries in STEDT, 17% start with a plain sonorant and 8.7% with a voiceless sonorant. Again, the imbalance in both amount and type of data contributing to the Marathi observations and the Burmese/Hakha Lai observations mean that they must be interpreted carefully. For example, the numbers for Marathi are drawn from all the consonants in the corpus, while the Burmese and Hakha Lai data drawn from STEDT are for absolute initial onsets only. The data are suggestive, however, that non-modal sonorants may well be better represented in Tibeto-Burman languages than breathy sonorants are in Marathi.

It is important to note that the acoustic facts, as illuminated by the current study, may also influence the phonotactic frequency data from EMILLE/CIIL: it may be true that there were fewer opportunities for breathy sonorants to develop in the first place, but they may also have been more vulnerable to the cumulative effects of signal degradation, and therefore to diachronic loss, à la Wright (2004). Evidence gleaned from comparing Marathi

---

<sup>15</sup> EMILLE/CIIL is a suite of corpora which are free for academic and research use. In addition to Marathi, it also contains materials for Assamese, Bengali, Gujarati, Hindi, Kannada, Kashmiri, Malayalam, Oriya, Punjabi, Sinhala, Tamil, Telegu, and Urdu. Additional details are available at [http://catalog.elra.info/product\\_info.php?products\\_id=696](http://catalog.elra.info/product_info.php?products_id=696).

<sup>16</sup> Other varieties of Burmese are included in STEDT, but Written Burmese was selected for analysis because it had many times more lexical items listed than any of the other varieties of Burmese indexed therein.

and Konkani suggests that this may be plausible. As has been noted, Konkani is one of the few other daughter languages of Western Middle Indic that retains contrastive breathy sonorants, though Almeida (1985) notes that “aspiration in Konkani is light compared to that found in many Indo-Aryan languages” (p. 22). In Marathi breathy sonorants occur in some very high-frequency function words, specifically [am<sup>h</sup>i] ‘we’ and [tum<sup>h</sup>i] ‘y’all’. The cognates in Konkani do not contain breathy sonorants: [ami] ‘we’ and [tumi] ‘y’all’ (data from Almeida 1989, Miranda 2007).

Even if we chose not to consider the Konkani forms, however, their occurrence in high frequency words should not be considered insurance against breathy sonorants disappearing from Marathi. Consider the interdental fricatives /θ ð/, for instance: though typologically uncommon, they are phonemic in English and are contained in some high-frequency function words yet have been lost from many varieties of Modern English (generally due to merging with either /f v/ or /t d/, but see Blevins 2006b for an overview of this phenomenon). As with any high-frequency items, there are questions waiting to be investigated with regards to the acoustic characteristics of these function words in running speech. Is the contrast between plain and breathy sonorants maintained in running speech in Marathi—especially in those high frequency function words? This question cannot be addressed at present, but analysis of more naturalistic running speech (as compared with the methodology employed in the current study, wherein target words were presented in a carrier sentence) will prove informative.

In addition, recall that in this paper a number of factors (speaker sex [male or female], vowel context [e or a], position of the consonant within the word [initial or medial], place of articulation of the consonant [labial, alveolar, or postalveolar/palatal], and repetition number [1, 2, or 3]) were included in the models but were not permitted to interact with the variables of PhonationType, Obstruency, or Window. This methodological choice was made in order to determine whether observed effects are robust enough to emerge across diverse contexts (e.g. places of articulation, vowel contexts). Each of these factors are worthy of future investigation, however.

Moreover, the larger question of generalizability remains open: work with other languages will help to determine whether the pattern reported here is language-specific or representative. Tibeto-Burman languages which indeed contain breathy voiced (as compared with voiceless) sonorants, like Sumi or Balami, are options, as is Konkani. One benefit inherent in focusing on Konkani in future work is that similar lexical items could be compared across Konkani and Marathi. The voiceless nasals in languages from the Tibeto-Burman or Otomanguean families, or in Icelandic, may also provide fruitful territory for study. Though the voiced/voiceless contrast differs from the one investigated herein, nonmodal sonorants in general are of interest because the larger question here has to do with the acoustic properties of typologically uncommon sounds. Voiceless sonorants may also prove fruitful in future work because they pattern so differently than breathy sonorants. While breathy sonorants are retained in very few western Indic languages, voiceless sonorants seem to resist diachronic loss.

## 6 Conclusion

In this paper I presented an acoustic analysis of plain and breathy voiced obstruents and sonorants in Marathi in order to test whether breathy phonation in sonorants is associated with the acoustic correlates of breathiness found for vowels and obstruents. The results indicate that the answer is yes. Breathly sonorants are associated with increased

spectral tilt as reflected by several H-based measures and with increased noise. However, the acoustic differences between plain and breathy consonants are diminished in sonorant contexts: [b] differs from [b<sup>h</sup>] more than [m] differs from [m<sup>h</sup>]. I posited that the acoustic facts can help explain the typological scarcity of breathy sonorants, and posed questions about why voiceless sonorants (e.g. in Tibeto-Burman languages) appear to resist diachronic loss while breathy sonorants appear to be vulnerable to it. Several possible explanations (the grammatical origin of the voiceless nasals in TB, perceptual salience) were entertained. Future work probing these factors will help determine whether they are valid, and analyses of breathy sonorants in other languages will help clarify the generalizability of the Marathi data.

## References

- Alam, F., S. M. Habib, & M. Khan. (2008). Acoustic analysis of Bangla consonants. Avail. at <http://dSPACE.bracu.ac.bd/xmlui/bitstream/handle/10361/669/ACOUSTIC%20ANALYSIS%20OF%20BANGLA%20CONSONANTS.pdf?sequence=1>
- Almeida, M. (1985). *A description of Konkani*. PhD dissertation. Georgetown University.
- Almeida, M. (1989). *A description of Konkani*. Panaji, Goa: Thomas Stephens Konkni Kendr.
- Andruski, J. & M. Ratliff. (2000). Phonation types in production of phonological tone: The case of Green Mong. *Journal of the International Phonetic Association*, 30: 37-61.
- Balasubramaniam, R.K, J. S. Bhat, S. Fahim III, & R. Raju III. (2011). Cepstral analysis of voice in unilateral adductor vocal fold palsy. *Journal of Voice*, 25.3: 326-329.
- Baayen, R. H., D. J. Davidson, & D. M. Bates. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59.4: 390-412.
- Baldi, P. (2009). Indo-European languages. in B. Comrie (ed.), *The World's Major Languages, Second Edition*. Routledge: New York, NY, pp. 23-50.
- Barr, D. J., R. Levy, C. Scheepers, & H. J. Tily. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68.3: 255-278.
- Bengueral, A. & T. Bhatia. (1980). Hindi stop consonants: An acoustic and fibroscopic study. *Phonetica*, 37:134-148.
- Berkson, K. (2012). Capturing a four-way distinction: Durational measures of oral stops in Marathi. *Kansas Working Papers in Linguistics*, 33: 27-46.
- Berkson, K. (2013). *Phonation Types in Marathi: An Acoustic Investigation*. PhD dissertation. University of Kansas, Lawrence, KS. Available at <https://kuscholarworks.ku.edu/handle/1808/12339>
- Berkson, K. (2016). Production, perception, and distribution of breathy sonorants in Marathi. *Formal Approaches to South Asian Languages Vol. 2*, M. Menon and S. Syed (eds). p. 4-14.
- Berkson, K. and M. Nelson. (2017). Phonotactic frequencies in Marathi. *IULC Working Papers*, 17.1: 1-25.
- Bhaskararao, P., & Ladefoged, P. (1991). Two types of voiceless nasals. *Journal of the International Phonetic Association*, 21.2: 80-88.
- Bielmeier, R., F. Haller, K. Häsler, B. Huber & M. Volkart. (2008). *The Comparative Dictionary of Tibetan Dialects (Preprint)*. Vol. 2: Verbs. Bern.

- Blankenship, Barbara. 1994. Phonetic structures of an endangered language: Khonoma Angami. *The Journal of the Acoustical Society of America*, 95.5: 2875–2876.
- Blankenship, B. (1997). The time course of breathiness and laryngealization in vowels. PhD dissertation, UCLA.
- Blankenship, B. (2002). The timing of nonmodal phonation in vowels. *Journal of Phonetics*, 30: 163-191.
- Blevins, J. (2004). *Evolutionary phonology: The emergence of sound patterns*. Cambridge, UK: Cambridge University Press.
- Blevins, J. (2006a). A theoretical synopsis of Evolutionary Phonology. *Theoretical linguistics*, 32.2: 117-166.
- Blevins, J. (2006b). New Perspectives on English Sound Patterns: “Natural” and “Unnatural” in Evolutionary Phonology. *Journal of English Linguistics*, 34.1: 6-25.
- Blevins, J. (2018). Evolutionary Phonology and the life cycle of voiceless sonorants. In S. Cristofaro & F. Zúñiga (eds.) *Typological hierarchies in synchrony and diachrony*. Amsterdam: John Benjamins Publishing Company, pp. 29-58.
- Bloch, J. (1970). *The formation of the Marathi language*, originally printed (1920). (D. R. Chanana, Trans.). Delhi: Motilal Banarsidass Press.
- Boersma, P. & D. Weenink. (2010). Praat: doing phonetics by computer [Computer program]. Version 5.1.32. Available at [www.praat.org](http://www.praat.org).
- Bradshaw, M. M. (2000). *A crosslinguistic study of consonant-tone interaction*. PhD dissertation. The Ohio State University, Columbus, Ohio.
- Caisse, M. (1981). Cross-linguistic differences in fundamental frequency perturbation induced by voiceless unaspirated stops. *The Journal of the Acoustical Society of America*, 70.S1: S76-S77.
- Cao, J. & Maddieson, I. (1992). An exploration of phonation types in Wu dialects of Chinese. *Journal of Phonetics*, 20.1: 77-92.
- Cardona, G. & S. Luraghi. (2009). Indo-Aryan languages. in B. Comrie (ed.), *The World's Major Languages, Second Edition*. Routledge: New York, NY, pp. 373-379.
- Catford, J. (1977). *Fundamental problems in phonetics*. Edinburgh: Edinburgh University Press.
- Census of India. (2011). *Statement 1: Abstract of speakers' strength of languages and mother tongues*. Census of India Website: Office of the Registrar General & Census Commissioner, India. Online resource available at <<http://www.censusindia.gov.in/2011Census/Language-2011/Statement-1.pdf>>. Accessed June 10, 2018.
- Chang, B. S. (1971). The Tibetan causative: phonology. *Bulletin of the Institute of History and Philology* 42, pp. 623-765.
- Chirkova, K & Z. Handel. (2015). Diachronic developments of voiceless nasals: the case of Ersu, Lizu, and related languages. Originally presented at 46th International Conference on Sino-Tibetan Languages and Linguistics Dartmouth, August 8-10, 2013. Available at <<https://hal.archives-ouvertes.fr/hal-01180886/document>>.
- Chirkova, K., P. Basset & A. Amelot. (2018). Voiceless nasal sounds in three Tibeto-Burman languages. *Journal of the International Phonetic Association*, 1-32. doi:10.1017/S0025100317000615
- Dantsuji, M. (1984). A study on voiceless nasals in Burmese. *Studia Phonologica* 18: 1-14.

- Dantsuji, M. (1986). Some acoustic observations on the distinction of place of articulation for voiceless nasals in Burmese. *Studia Phonologica* 20: 1-11.
- Davis, K. (1994). Stop voicing in Hindi. *Journal of Phonetics*, 22.2: 177-193.
- De Boer, B. (2000). Self-organization in vowel systems. *Journal of Phonetics*, 28.4: 441-465.
- DeLancey, S. (2009). Sino-Tibetan languages. in B. Comrie (ed.), *The World's Major Languages, Second Edition*. Routledge: New York, NY, pp. 693-702.
- Demolin, D., & V. Delvaux. (2001). Whispy voiced nasal stops in Rwanda. In *Seventh European Conference on Speech Communication and Technology – Proceedings of Eurospeech 2001*. Accessed online at <https://pdfs.semanticscholar.org/f0e3/44722be6de02621af03e3b6f389a654e65f0.pdf>.
- Dhongde, R.V. & K. Wali. (2009). *Marathi*. John Benjamins Publishing Co.: Amsterdam.
- DiCanio, C. T. (2009). The phonetics of register in Takhian Thong Chong. *Journal of the International Phonetic Association*, 39.2: 162-188.
- DiCanio, C. T. (2012). Coarticulation between tone and glottal consonants in Itunyoso Trique. *Journal of Phonetics*, 40.1: 162-176.
- Dmitrieva, O., & Dutta, I. (2018). Acoustic correlates of the four-way laryngeal contrast in Marathi stops. *The Journal of the Acoustical Society of America*, 143.3: 1756-1756.
- Dutta, I. (2007). Four-way stop contrasts in Hindi: An acoustic study of voicing, fundamental frequency and spectral tilt. Ph.D. dissertation, University of Illinois at Urbana-Champaign.
- Esposito, C.M. (2006). *The Effects of Linguistic Experience on the Perception of Phonation*. Ph.D. dissertation, University of California at Los Angeles.
- Esposito, C.M. (2010). The effects of linguistic experience on the perception of phonation. *Journal of Phonetics*, 38.2: 306-316.
- Esposito, C.M. (2012). An acoustic and electroglottographic study of White Hmong tone and phonation. *Journal of Phonetics*, 40: 466-476.
- Esposito, C. M., & S. 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.
- Gao, J. (2015). *Interdependence between Tones, Segments and Phonation types in Shanghai Chinese: acoustics, articulation, perception and evolution*. Ph.D dissertation, Université Sorbonne Nouvelle - Paris 3.
- Gao, J., P. Hallé, K. Honda, S. Maeda, & M. Toda. (2011). Shanghai slack voice: acoustic and ePGG data. *Proceedings of the 17th International Congress of Phonetic Science (ICPhS)*, 719–722. Oral presentation. August 2011, Hong Kong, China.
- Garellek, M. (2012). The timing and sequencing of coarticulated non-modal phonation in English and White Hmong. *Journal of Phonetics*, 40.1: 152-161.
- Garellek, M., & Keating, P. (2011). The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *Journal of the International Phonetic Association*, 41.2: 185-205.
- Garellek, M., Ritchart, A., & Kuang, J. (2016). Breathy voice during nasality: A cross-linguistic study. *Journal of Phonetics*, 59: 110-121.
- Gautam, B. R. Phonation types of Balami consonants: An acoustic analysis. *Nepalese Linguistics*, 27: 42-53.

- Ghatage, M. M. (2013). Pronunciation problems of the Marathi speakers. *Language in India*, 13(4), 107-115.
- Gogoi, P., & Wayland, R. (2018). Perception of voiceless nasals in Mizo and Angami. *The Journal of the Acoustical Society of America*, 144.3, 1720-1720.
- Grunow-Hårsta, K. (2008). *A descriptive grammar of two Magar dialects of Nepal: Tanahu and Syangja Magar*. Ph.D. dissertation. University of Wisconsin-Milwaukee. Available from ProQuest Dissertations & Theses Global. (304451693).
- Haider, H. (1985). The fallacy of typology: remarks on the PIE stop-system. *Lingua*, 65.1-2: 1-27.
- Haller, F. (2004). *Dialekt und Erzählungen von Themchen: sprachwissenschaftliche Beschreibung eines Nomadendialektes aus Nord-Amdo*. Bonn: VGH Wissenschaftsverlag.
- Harris, T. (2009). *An acoustic and articulatory study of sonorant phonation in Sumi*. Unpublished honours thesis; The University of Melbourne School of Languages and Linguistics. Melbourne, Australia.
- Heman-Ackah, Y., R. Heuer, D. Michael, R. Otrowski, M. Horman, M. Baroody, J. Hillenbrand, & R. Sataloff. (2003). Cepstral peak prominence: a more reliable measure. *The Annals of Otology, Rhinology & Laryngology*, 112.4: 324-344.
- Herrera Zendejas, E. (2014). *Mapa fónico de las lenguas mexicanas: Formas sonoras 1 y 2*. El Colegio de Mexico AC.
- Hildebrandt, K. A. (2007). Tone in Bodish languages: Typological and sociolinguistic contributions. *Trends in Linguistics Studies and Monographs*, 189, 77.
- Hill, N.W. (2014). Proto-Kuki-Chin initials according to Toru Ohno and Kenneth Van Bik. *Journal of the Southeast Asian Linguistics Society (JSEALS)*, 7: 11-30
- Hillenbrand, J., R. Cleveland, and R. Erickson. (1994). Acoustic correlates of breathy vocal quality. *Journal of Speech and Hearing Research*, 37: 769-778.
- Hock, H. H. (2010). Middle Indo-Aryan "Aspirate" Clusters Revisited. *Studia Orientalia Electronica*, 108: 87-102.
- Hoffmann, M. (2018). Inter-and intra-speaker variability in the production of voiceless nasal consonants. *The Journal of the Acoustical Society of America*, 143.3, 1968-1968.
- Huffman, M. K. (1987). Measures of phonation in Hmong. *The Journal of the Acoustical Society of America*, 81: 495-504.
- Hinüber, O. (1986). *Das ältere Mittelindisch im Überblick*. Wien: Verlag der Österreichischen Akademie der Wissenschaften. (2<sup>nd</sup> edition 2001).
- Iseli, M., & Alwan, A. (2004, May). An improved correction formula for the estimation of harmonic magnitudes and its application to open quotient estimation. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04)*.
- Iseli, M., Shue, Y., & Alwan, A. (2007). Age, sex, and vowel dependencies of acoustic measures related to the voice source. *The Journal of the Acoustical Society of America*, 121.4: 2283-2295.
- Jun, S. A., & Cha, J. (2015). High-toned [il] in Korean: Phonetics, intonational phonology, and sound change. *Journal of Phonetics*, 51: 93-108.
- Karlgren, B. (1915-1926). Études sur la phonologie Chinoise. *Archives d'études Orientales*, Vol. 15 (in 4 parts). Leiden: E. J. Brill; Uppsala: K.W. Appelberg.

- Kavadi, N. B., & F. C. Southworth. (1965). *Spoken Marathi. Book I*. University of Pennsylvania Press: Philadelphia, PA.
- Kawahara, H., A. D. Cheveigne, & R. D. Patterson. (1998). An instantaneous-frequency-based pitch extraction method for high-quality speech transformation: revised TEMPO in the STRAIGHT-suite. In *Proceedings of the Fifth International Conference on Spoken Language Processing (ICSLP '98)*.
- Keating, P. & C. Esposito. (2007). Linguistic voice quality. *UCLA Working Papers in Phonetics*, 105: 85-91.
- Kelterer, A. (2017). Non-modal voice quality in Chichimeco--" Hablamos más con la garganta". Unpublished MA thesis; Lund University Centre for Language and Literature. Lund, Sweden.
- Kent, R. G. (1932). The sounds of Latin. A Descriptive and Historical phonology. *Language*, 8.3: 11-216.
- Khan, S. D. (2012). The phonetics of contrastive phonation in Gujarati. *Journal of Phonetics*, 40: 780-795.
- Kobayashi, M. (2004). *Historical phonology of old Indo-Aryan consonants* (No. 42). Tokyo: Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Kumar, B. R., Bhat, J. S., & Mukhi, P. (2011). Vowel harmonic amplitude differences in persons with vocal nodules. *Journal of Voice*, 25.5: 559-561.
- Ladefoged, P. (1971). *Preliminaries to Linguistic Phonetics*. University of Chicago Press, Chicago.
- Ladefoged, P. & I. Maddieson. (1996). *Sounds of the World's Languages*. Blackwell Publishers, Inc.: Cambridge, MA.
- Laver, J. (1980). *The phonetic description of voice quality (Cambridge Studies in Linguistics 31)*. Cambridge University Press: London, UK.
- Lee, S. J., Cho, Y., Song, J. Y., Lee, D., Kim, Y., & Kim, H. (2016). Aging effect on Korean female voice: Acoustic and perceptual examinations of breathiness. *Folia Phoniatria Et Logopaedica*, 67.6: 300-307.
- Lisker, L. & A. Abramson. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20: 384-422.
- Maddieson, I. (2003). The sounds of the Bantu languages. in D. Nurse and G. Philippson (eds.), *The Bantu Languages*. Routledge: New York, NY, pp. 15-41.
- Marchal, A. (2009). *From Speech Physiology to Linguistic Phonetics*. John Wiley & Sons, Inc.: Hoboken, NJ.
- Masica, C. (1991). *The Indo-Aryan languages*. Cambridge University Press: Cambridge, UK.
- Mazaudon, M. (1973). *Phonologie Tamang: étude phonologique du dialecte tamang de Risianku, langue tibéto-birmane du Népal* (Vol. 4). Société d'études linguistiques et anthropologiques de France.
- MathWorks, Inc. (2011). MATLAB version 7.12.0.635. Natick, MA.
- Matisoff, J. A. (2003). *Handbook of Proto-Tibeto-Burman: system and philosophy of Sino-Tibetan reconstruction*. Berkeley and Los Angeles, CA: University of California Press.
- Matisoff, J. A. (2016). *The Sino-Tibetan etymological dictionary and thesaurus*. Berkeley: The Regents of the University of California. Online searchable database available at <http://stedt.berkeley.edu/~stedt/cgi/rootcanal.pl>.

- Michaud, A., & M. Mazaudon, M. (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). In *Proceedings of Speech Prosody 2006, Dresden* (pp. pp-819). Technische Universität Darmstadt Press.
- Mikuteit, S. & H. Reetz. (2007). Caught in the ACT: The timing of aspiration and voicing in Bengali. *Language and Speech*, 50.2: 247-277.
- Miller, A. L. (2007). Guttural vowels and guttural co-articulation in Ju|'hoansi. *Journal of Phonetics*, 35.1: 56-84.
- Miranda, R. (2007). Konkani. in D. Jain & G. Cardona (eds.), *The Indo-Aryan Languages*. Routledge: ebook published 2007. Retrieved from <http://ebookcentral.proquest.com>. pps. 729-765 in the 2003 print edition.
- Mohr, B. (1971.) Intrinsic variations in the speech signal. *Phonetica* 23: 65-93.
- Moran, S., McCloy, D., and Wright, R. (eds.). (2014). PHOIBLE Online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (<http://phoible.org>)
- Mortensen, D. R. (2011). Lexical prefixes and Tibeto-Burman laryngeal contrasts. In *Annual Meeting of the Berkeley Linguistics Society*, 37.1: 272-286.
- Narasimhan, S. V., & Vishal, K. (2017). Spectral measures of hoarseness in persons with hyperfunctional voice disorder. *Journal of Voice*, 31.1: 57-61.
- Ní Chasaide, A. & C. Gobl. (1995). Voice source variation. in W. Hardcastle & J. Laver (eds.), *The Handbook of Phonetic Sciences*. Blackwell: Cambridge, MA, pp. 427-461.
- Noonan, M. & K. Hildebrandt. (2016). Nar-Phu. in R. J. LaPolla & G. Thurgood (eds.), *The Sino-Tibetan Languages 2<sup>nd</sup> Edition*. Routledge: New York, NY, pp. 534-556.
- Ohala, J. J. (1973). The physiology of tone. *Consonant Types and Tone* (No. 1), 1-14.
- Ohala, M. & J. Ohala. (1972). The problem of aspiration in Hindi phonetics. *Annual Bulletin of the Research Institute of Logopedics and Phoniatics, Faculty of Medicine, University of Tokyo, Japan*. No. 6: 39-46.
- Ohala, M. (1979). Phonological features of Hindi stops. *Proceedings of the South Asian Languages Analysis Roundtable*, 1: 79-88.
- Ohde, R. N. (1984). Fundamental frequency as an acoustic correlate of stop consonant voicing. *The Journal of the Acoustical Society of America*, 75.1: 224-230.
- Pan, H. (2017). Glottalization of Taiwan Min checked tones. *Journal of the International Phonetic Association*, 47.1: 37-63.
- Pandharipande, R. V. (1997). *Marathi*. Routledge: London, UK.
- Pandit, P . B. (1957). Nasalization, aspiration and murmur in Gujarati. *Indian Linguistics*, 17: 165-172.
- Pearce, M. (2009). Kera tone and voicing interaction. *Lingua*, 119(6), 846-864.
- Pennington, M. (2005). *The phonetics and phonology of glottal manner features*. PhD dissertation, University of Indiana.
- Perkins, J. M. (2013). *Consonant-tone interaction in Thai*. Ph.D. dissertation. Rutgers University: New Brunswick, NJ.
- Pittayaporn, P., & Kirby, J. (2017). Laryngeal contrasts in the tai dialect of Cao Bang. *Journal of the International Phonetic Association*, 47.1: 65-85.
- R Core Team. (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. available at <http://www.R-project.org>.

- Ronkos, D. A. (2015). *Obstruent voicing and tone in Siklis Gurung*. MA thesis. The City University of New York: New York, NY. Available from ProQuest Dissertations & Theses Global. (1689405761).
- Schiefer, L. (1986). F0 in the production and perception of breathy stops: Evidence from Hindi. *Phonetica*, 43: 43-69.
- Schiefer, L. (1987). 'Voiced aspirated' or 'breathy voiced' and the case for articulatory phonology. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München*, 27: 257-278.
- Shue, Y., P. Keating, C. Vicenik, & K. Yu. (2011). VoiceSauce: A Program for Voice Analysis. available at <http://www.ee.ucla.edu/~spapl/voicesauce/>.
- Southworth, F. (2000). Review of *Marathi* by R. Pandharipande. *Lingua*, 110: 215-224.
- Sreedhar, M.V. (1976). *Sema Phonetic Reader*. Mysore: Central Institute of Indian Languages.
- Starr, R. L. (2015). Sweet voice: The role of voice quality in a Japanese feminine style. *Language in Society*, 44(1), 1-34.
- Sun, J. T. S. (1997). The typology of tone in Tibetan. *Chinese Languages and Linguistics IV: Typological studies of languages in China*. pp. 485-521.
- Tang, K. E. (2008). *The phonology and phonetics of consonant-tone interaction*. PhD dissertation. University of California-Los Angeles.
- Tian, J., & Kuang, J. (2016). Revisiting the Register Contrast in Shanghai Chinese. *Tonal Aspects of Languages 2016*, 147-151.
- Tournadre, N. & S. Dorje. (2003). *Manual of standard Tibetan: language and civilization*. Ithaca, NY: Snow Lion Publications.
- Traill, A. & M. Jackson. (1988). Speaker variation and phonation type in Tsonga nasals. *Journal of Phonetics*, 16.4: 385-400.
- Turner, R. L. (1962-1966). *A comparative dictionary of Indo-Aryan languages*. London: Oxford University Press.
- UCLA Phonological Segment Inventory Database. [http://web.phonetik.uni-frankfurt.de/cgi-bin/upsid\\_sounds.cgi](http://web.phonetik.uni-frankfurt.de/cgi-bin/upsid_sounds.cgi).
- Van Bik, K. (2009). *Proto-Kuki-Chin: A Reconstructed Ancestor of the Kuki Chin Languages*, STEDT Monograph Series, volume 8. Berkeley: STEDT.
- Watkins, J. W. (2001). Burmese. *Journal of the International Phonetic Association*, 31.2: 291-295.
- Watts, C. & S. Awan. (2011). Use of spectral/cepstral analyses for differentiating normal from hypofunctional voices in sustained vowel and continuous speech contexts. *Journal of Speech, Language and Hearing Research*, 54: 1525-1537.
- Wayland, R., S. Gargash, & A. Jongman. (1995). Acoustic and perceptual investigation of breathy voice. *The Journal of the Acoustical Society of America*, 97.5: 3364.
- Wayland, R. & A. Jongman. (2003). Acoustic correlates of breathy and clear vowels: The case of Khmer. *Journal of Phonetics*, 31: 181-201.
- Winter, B., & Grawunder, S. (2012). The phonetic profile of Korean formal and informal speech registers. *Journal of Phonetics*, 40(6), 808-815.
- Wright, Richard. (2004). A review of perceptual cues and cue robustness. In Hayes B., R. Kirchner, & D. Steriade eds., *Phonetically Based Phonology*, 34-57. Cambridge University Press, New York.
- Yardi, V. V. (1998). English Pronunciation for Marathi Speakers. *Saket publications, Aurangabad*.

## **Acknowledgments**

In addition to three anonymous reviewers, I am indebted to the following people (in alphabetical order) for their valuable input: Namrata Barve, Nancy Caplow, Kenneth de Jong, Stefon Flego, Allard Jongman, Goun Lee, Hyunjung Lee, Vitor Leongue, Stephen Politzer-Ahles, Joan Sereno, Kenneth Van Bik, and Jie Zhang. I am also indebted to the ten people who participated in the production study, for their time and patience.

## Appendix A – Stimuli List

word initial							
<i>before [a]</i>					<i>before [e]</i>		
1	b	[badzu]	बाजू	side	[be[e]	बेटे	islands
2	b <sup>h</sup>	[b <sup>h</sup> atsa]	भाचा	nephew	[b <sup>h</sup> egʌ]	भेग	crack (crevice)
3	m	[masa]	मासा	fish	[meʂ]	मेष	(a zodiac sign - Aries)
4	m <sup>h</sup>	[m <sup>h</sup> atre]	म्हात्रे	surname	[m <sup>h</sup> etʌr]	म्हेतर	a sweeper
5	v	[vara]	वारा	wind	[vedʌ]	वेड	passion, obsession
6	v <sup>h</sup>	[v <sup>h</sup> ave]	व्हावे	let it happen	[v <sup>h</sup> egʌ]	व्हेग	an ancient tree*
7	d	[dada]	दादा	elder brother	[dehʌ]	देह	body
8	d <sup>h</sup>	[d <sup>h</sup> aga]	धागा	thread	[d <sup>h</sup> enu]	धेनु	cow
9	n	[natu]	नातू	grandson	[nete]	नेते	leaders
10	n <sup>h</sup>	[n <sup>h</sup> avi]	न्हावी	barber	[n <sup>h</sup> e[i]	न्हेटी	farming implement*
11	l	[lagu]	लागू	applicable (also, a last name)	[leki]	लेकी	daughters
12	l <sup>h</sup>	[l <sup>h</sup> asa]	ल्हासा	capital of Tibet	[l <sup>h</sup> ete]	ल्हेते	wood-carving tool*
13	dʒ	[dʒagruk]	जागस्कू	alert	[dʒet <sup>h</sup> e]	जेथे	where (locative)
14	dʒ <sup>h</sup>	[dʒ <sup>h</sup> apʌd]	झापड	slap	[dʒ <sup>h</sup> epʌ]	झेप	leap
15	r	[raja]	राजा	king	[reti]	रेती	sand
16	r <sup>h</sup>	[r <sup>h</sup> as]	ऱ्हास	deterioration	[r <sup>h</sup> ekʌ]	ऱ्हेक	ancient style of planting*
word medial							
1	b	[dʌba]	डबा	box	[dʌbe]	डबे	boxes
2	b <sup>h</sup>	[ub <sup>h</sup> a]	उभा	standing (masc)	[ub <sup>h</sup> e]	उभे	standing (neut)
3	m	[ɾʌma]	रमा	(proper name)	[vime]	विमे	insurance policies
4	m <sup>h</sup>	[tum <sup>h</sup> a]	तुम्हा	to you	[tum <sup>h</sup> e]	तुम्हे	to you
5	v	[hʌva]	हवा	air	[hʌve]	हवे	wanted
6	v <sup>h</sup>	[giv <sup>h</sup> a]	गीव्हा	tongue	[nʌv <sup>h</sup> e]	नव्हे	is not
7	d	[gʌda]	गदा	mace (weapon)	[pʌde]	पदे	couplets
8	d <sup>h</sup>	[gʌd <sup>h</sup> a]	गधा	donkey	[gʌd <sup>h</sup> e]	गधे	donkeys
9	n	[mʌna]	मना	mind	[sone]	सोने	gold
10	n <sup>h</sup>	[pun <sup>h</sup> a]	पुन्हा	again	[pʌn <sup>h</sup> e]	पन्हे	mango drink
11	l	[pela]	पेला	glass/tumbler	[pele]	पेले	glasses/tumblers
12	l <sup>h</sup>	[dʒil <sup>h</sup> a]	जिल्हा	district	[dʒil <sup>h</sup> e]	जिल्हे	districts
13	dʒ	[mʌdʒa]	मजा	fun	[mʌdʒet]	मजेत	in fun
14	dʒ <sup>h</sup>	[kidʒ <sup>h</sup> a]	किझा	Malaysian fruit*	[kidʒ <sup>h</sup> e]	किझे	Malaysian fruits*
15	r	[kʌra]	करा	do	[tare]	तारे	stars (pl.)
16	r <sup>h</sup>	[tʌr <sup>h</sup> a]	तऱ्हा	the way (of doing something)	[tʌr <sup>h</sup> e]	तऱ्हे	ways (of doing something)

\*Indicates a nonce word

## Appendix B – Demographic Data

<b>Female Participants</b>					
<b>Subject ID</b>	<b>W1</b>	<b>W3</b>	<b>W4</b>	<b>W5</b>	<b>W6</b>
<b>Age</b>	29	40	25	56	36
<b>Language spoken at home</b>	Marathi	Marathi	Marathi, Konkani	Marathi	Marathi
<b>Other languages spoken</b>	Hindi, English	Hindi, English, Kannada	Hindi, English, Konkani	Hindi, English	Hindi, English
<b>Hometown</b>	Pune	Wordha	Mumbai, Pune	Jalgaon, Pune	Pune
<b>Time in the US</b>	3 yrs	3 yrs	1 yrs	5 yrs	>5 yrs
<b>Amount of Marathi spoken per day</b>	a lot, every day	30% daily language use	50% daily language use	50% daily language use	30% daily language use
<b>Parents' languages</b>	Marathi	Marathi	Konkani, Marathi	Marathi	Marathi
<b>Male Participants</b>					
<b>Subject ID</b>	<b>M1</b>	<b>M4</b>	<b>M5</b>	<b>M6</b>	<b>M7</b>
<b>Age</b>	30	34	43	64	56
<b>Language spoken at home</b>	Marathi, Hindi, English	Marathi	Marathi	Marathi	Marathi
<b>Other languages spoken</b>	Hindi, English, Bengali, French	Hindi, English	Hindi, English	Hindi, English	Hindi, English
<b>Hometown</b>	Mumbai	Pune	Pune	Asoda/Pune	Mumbai
<b>Time in the US</b>	2 years 8 months	2.5 yrs	3.5 years	5 years	>27 yrs
<b>Amount of Marathi spoken per day</b>	15% of daily language use	10 min/week	30% daily language use	25% daily language use	2 - 4 hrs/day
<b>Parents' languages</b>	Marathi	Marathi	Marathi	Marathi	Marathi