
eScholarship
University of California

UCLA Working Papers in Phonetics, No. 104, 82-106

Breathy Nasals and /Nh/ Clusters in Bengali, Hindi, and Marathi

Christina M. Esposito, esposito@humnet.ucla.edu
Sameer ud Dowla Khan, sameer@humnet.ucla.edu
Alex Hurst ahurst@ucla.edu

Abstract

Previous work on breathiness in Indic languages has focused primarily on the acoustic properties of breathy (also known as aspirated) oral stops in languages like Hindi ([baːl] 'hair' vs. [bʱaːl] 'forehead') or Bengali ([bati] 'bowl' vs. [bʱ ːl[‘borth vst þ

In a few languages of South Asia, such as Marathi, Konkani, and Newari, the

2 Language Background

2.1 Bengali

Bengali (Bangla) is an Indo-European language spoken by approximately 171 million people in the People's Republic of Bangladesh and in the eastern Indian states of West Bengal, Assam, and Tripura. Bengali is also used as a second language by over 40 million Bangladeshis and Indians (Gordon 2005). Speakers of Bengali are typically fluent in more than one dialect; most Bengalis speak both the regional variety alongside the Standard Kolkata dialect (often called Choltibhasha), which is used in all media and formal education. All data in this study is pronounced in Choltibhasha by speakers born in the Dhaka Division of Bangladesh. Bengali, along with many other eastern Indic languages, is written with the Bengali script, a syllabic alphabet derived from the ancient Brahmi script.

2.2 Hindi

Hindi is an Indo-European language spoken primarily in a large belt across north-central India, including the states of Uttar Pradesh, Bihar, Madhya Pradesh, Rajasthan, Haryana, Himachal Pradesh, and the Union Territory of Delhi. Roughly 180 million people use it as a first language, with 120 million second-language speakers (Gordon 2005). While a number of different regional varieties of Hindi exist (such as Braj and Bhojpuri), the form examined here is Modern Standard Hindi, the language of literature, the press, and standard education. While this form of Hindi originally existed mainly as a lingua franca for speakers of the regional varieties, it is now spoken natively by many people, especially those from urban areas (Masica 1991). Hindi is written with the Devanagari script.

2.3 Marathi

Marathi is an Indo-European language spoken in the state of Maharashtra, located in the northwestern part of the Deccan plateau in south-central India. Major cities found in this region include Pune (Poona), Mumbai (Bombay), and Nagpur. Marathi has roughly 68 million native speakers, with 300,000 second-language speakers (Gordon

1984; Hinskens and van de Weijer 2003). No previous research has been conducted on Bengali /Nh/.

4 Current Study

The goal of the current study is to determine if /Nh/ sequences in Bengali and Hindi are breathy nasals, like Marathi [N̤], or if they are [Nh] clusters instead. To answer this question, simultaneous audio, aerodynamic, and electroglottographic (EGG) recordings were made for Bengali, Hindi, and (to provide a point for comparison) Marathi. Both within- and cross-language comparisons were made. In addition, we examined some phonological evidence to see if /Nh/ in Bengali and Hindi behave like clusters or single segments (expanded on in Part 7). We will begin by discussing the within-language comparisons.

5 Within-language comparisons

For the within-language comparisons, voiced aspirated oral stops ([D^h]) were compared to potentially breathy nasal ones (/Nh/) in Bengali and Hindi. If breathy nasals do exist in Bengali and Hindi, they might share some features with voiced aspirated oral stops (except, of course, for nasality).

5.1 Methods

5.1.1 Speakers

For Bengali, three adult male and two adult female speakers were recorded for this study. For Hindi, one adult female speaker was recorded. All speakers spoke English in addition to the language in question.

5.1.2 Speech Materials

The Bengali and Hindi speakers produced words that included the sounds given in Table 3:

1. Nasals	[N]
2. Potentially Breathy Nasals (nasals followed by [h])	/Nh/
3. Voiced Unaspirated Stops	[D]
4. Voiced Stops Followed by [h][4]	[Dh]
5. Voiced Aspirated Stops	[D^{h}]
6. Glottal Fricative[5]	[h]

Table 3: A list of the consonants (and clusters) recorded for both Bengali and Hindi.

These consonants and consonant clusters were produced intervocalically. Six words per consonant/cluster were produced. Each word contained one and only one of the consonants/clusters in question. A word never contained more than one type of consonant/cluster being studied.

5.2 Procedure

Each word was repeated three times in the carrier sentence [ʃe ____ bollo] "S/he said ____" for Bengali, and [ab ____ kahie] "Please say ____ now" for Hindi. Simultaneous electroglottographic and audio recordings were made for each speaker. Tokens were digitized and analyzed at a sampling rate of 22 kHz using AcQuirer software (Scicon RD). Figure 1 shows the audio, oral flow, nasal flow, and EGG signals, respectively, for Hindi [paka:na:] "to cook (transitive)" as displayed in AcQuirer (Scicon RD). (The oral and nasal flow data are relevant to

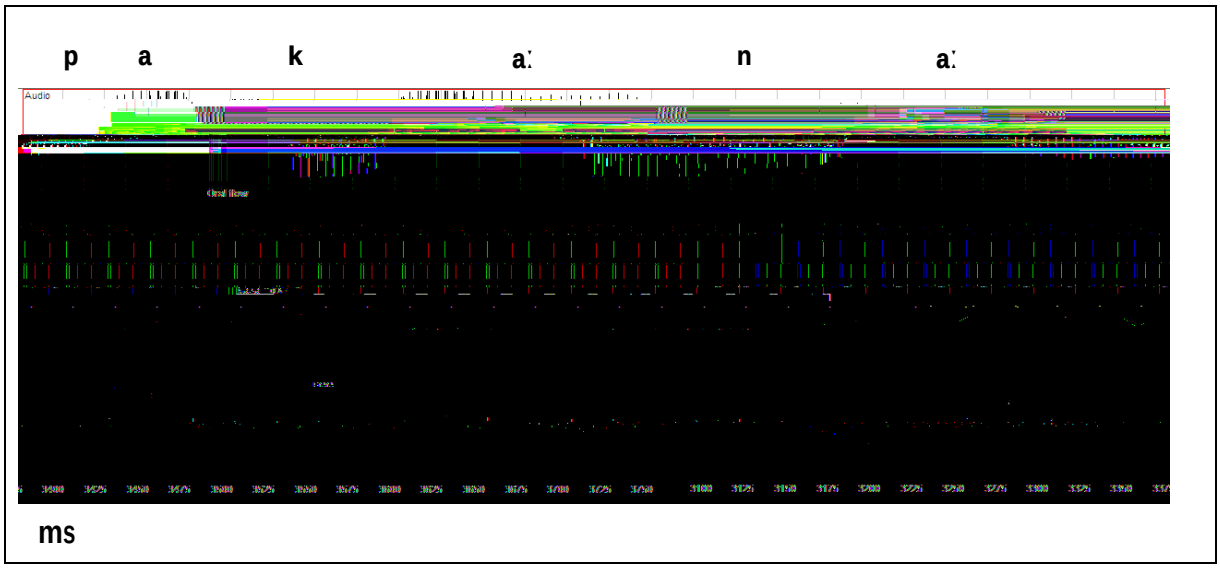


Figure 1: Audio, oral flow, nasal flow, and EGG of Hindi [paka:na:] "to cook (transitive)"

Duration and closed quotient were measured for each token. In the next section, we detail how and where each measurement was taken.

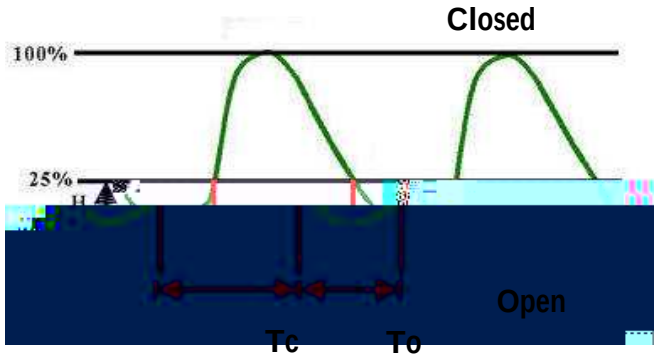
5.2.1 Description of measures

5.2.1.1 Duration

For each segment, the duration of the modal/unaspirated portion and the duration of the breathy/aspirated portion were measured in milliseconds (ms) using spectrograms created from the audio signal. For the unaspirated voiced stops ([D]) and the modal nasals ([N]), the duration of the entire segment was measured.

5.2.1.2 Closed Quotient

Closed quotient (CQ) was measured automatically with AcQuirer by dividing the amount of vocal fold closure (Tc) by the sum of the amount of vocal fold closure plus the amount of vocal fold opening (To) for each glottal pulse. Figure 2 shows this equation and how it was derived from a glottal pulse from the EGG signal.



Closed quotient = Total closure / (Total closure + Total open)

Figure 2: Closed quotient and how it was calculated from the EGG signal. H represents a default 25% threshold. Tc = total closure. To = total open

CQ was measured at two points: 1) the middle of the modal/unaspirated portion 2) the middle of the breathy/aspirated portion. Again, spectrograms were used to determine the modal/unaspirated and breathy/aspirated portions of the segment.

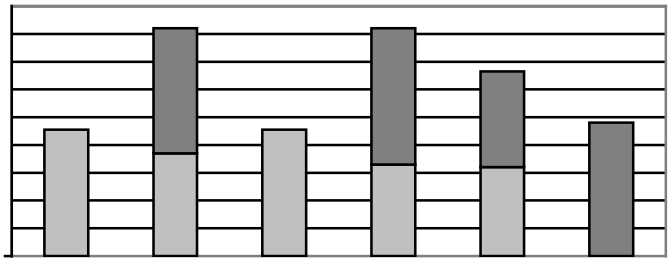
5.3 Results and Discussion

We begin by presenting the results and discussion of the within-language comparison for Bengali, and then move on to Hindi.

5.3.1 Bengali

5.3.1.1 Duration

Figure 3 is a graph of the average duration of the breathy/aspirated and modal/unaspirated portion of the segments [N], /Nh/, [D], [Dh], [D^{h}], and [h]. Duration is graphed on the y-axis in ms. Due to typographical limitations, breathiness/aspiration is represented by an apostrophe (') in all graphs in this paper.



of breathy nasals in particular, the CQ values of /Nh/ pattern with both [D^{ɦ}] and [Dh], making the data inconclusive for Bengali.

5.3.1.3 Summary of results for Bengali

The within-language comparison of the Bengali data, in general, is inconclusive. While duration indicates that /Nh/ is more like the cluster [Dh] than like [D^{ɦ}], other evidence, such as CQ, is ambiguous, at best. The CQ data indicates that /Nh/ behaves like both the cluster [Dh], and the single segment [D^{ɦ}]. Thus, it is not clear from the data whether /Nh/ is a cluster or a single segment.

We will now move on the results and discussion of data for Hindi.

5.3.2 Hindi

5.3.2.1 Duration

In Hindi, underlying sequences of /D/+/h/ were consistently split up by schwa epenthesis, yielding [Dəh]. A few of the /N/+/h/ tokens were also split up in this same way, yielding some productions of [Nəh]. All tokens with schwa epenthesis were excluded from the analysis, including all the graphs below. Figure 5 is a graph of the average duration of the breathy/aspirated and modal/unaspirated portion of the segments [N], /Nh/, [D], [D^{h}], and [h]. Duration is graphed on the y-axis in ms.

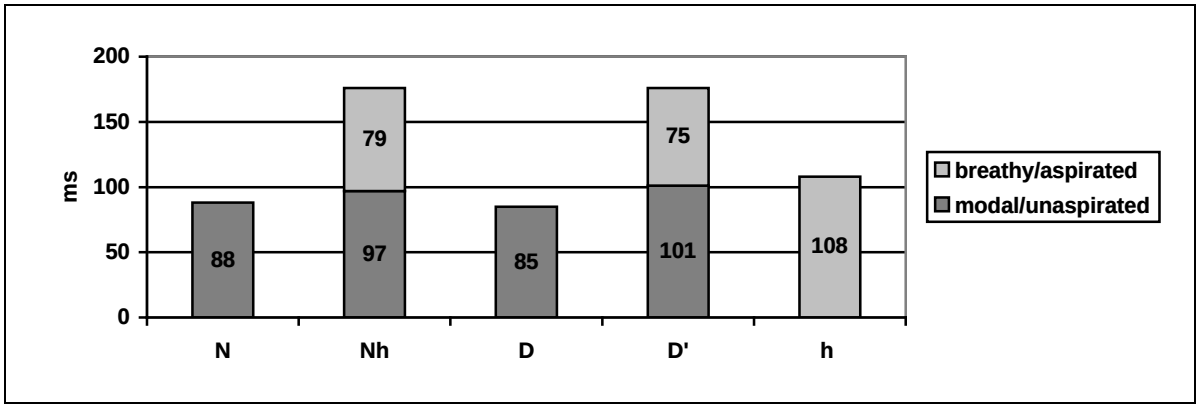


Figure 5: Graph of the average duration of the breathy/aspirated and modal/unaspirated portion of [N], /Nh/, [D], [D^{h}], and [h] for Hindi.

While the duration of /Nh/ is similar to [D^{ɦ}], suggesting the former might be a single consonant, the fact that schwas occasionally break up /Nh/ make it look more like a cluster (a single segment like [D^{ɦ}] cannot be split up in a similar manner). It is hard to draw a firm conclusion from the duration data, however, as true [Dh] clusters were not produced without schwa epenthesis.

5.3.2.2 Closed quotient

Figure 6 is a graph of the average CQ value for Hindi [N], /h/, [D], [D

5.3.2.3 Summary of results for Hindi

To summarize, the within-language comparison of the Hindi data is inconclusive. While duration indicates that Hindi /Nh/ is more like the single segment [D^{h}], the schwa epenthesis suggests it is more like a cluster. However, the CQ data is ambiguous, indicating that /Nh/ behaves like both a cluster and a single segment.

We will now present the cross-language comparison.

6 Cross-language comparison

For the cross-language comparison, the potentially breathy nasals in Bengali and Hindi were compared to the phonemically breathy nasals in Marathi.

6.1 Methods 8["G"(resfsW()]5F[:W[GRW)G:8["G"(resfs([(“]F“:W[GRW)G:8["G"(resfs:S:

6.1.2.2 Bengali and Hindi

The [N] and /Nh/ words from part 5 of this study were used again in this part of the study.

6.2 Procedure

For Marathi, each word was repeated three times in the carrier sentence [to ____ ʃəbda wats] "He said the word _____". For Bengali and Hindi, the same carrier sentences that were used in Part 5 of the study were used here. Simultaneous aerodynamic, EGG, and audio recordings were made for each speaker. (For Bengali, aerodynamic data could only be collected for one male and one female speaker.) Tokens were digitized and analyzed at a sampling rate of 22 kHz using AcQuirer software (Scicon RD). CQ was taken following the same procedure used in Part 5. In additional, peak airflow and

6.2.2 Duration of peak airflow

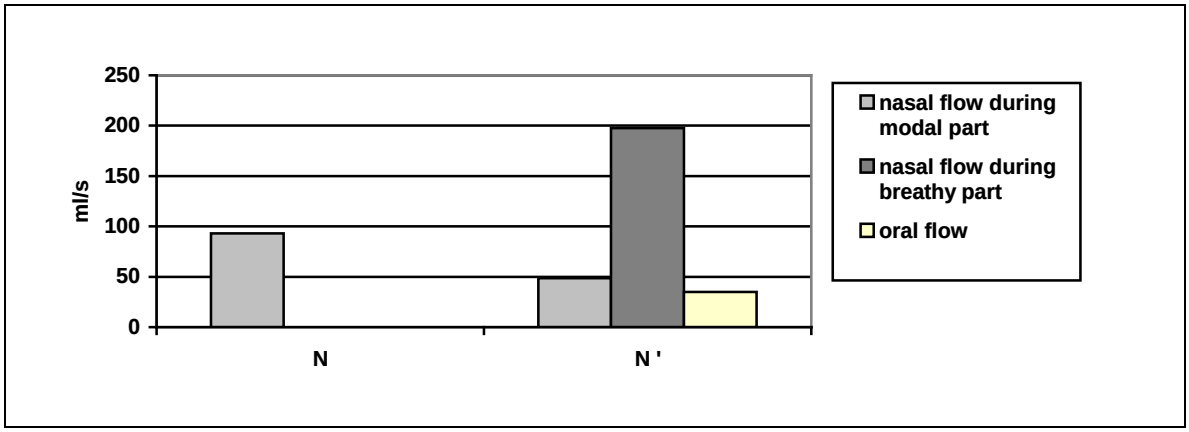


Figure 8: Graph of the average peak oral and nasal flow of [N] and [N̩] for Marathi.

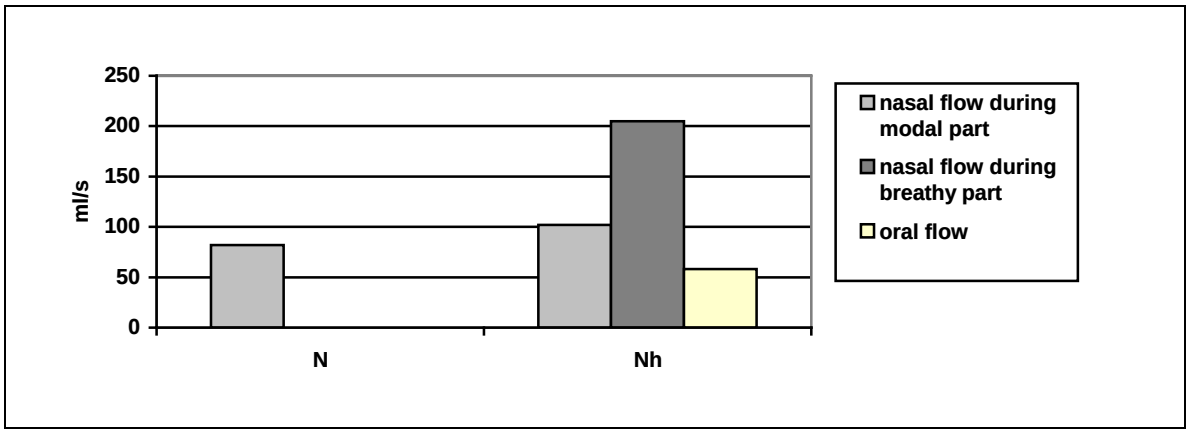
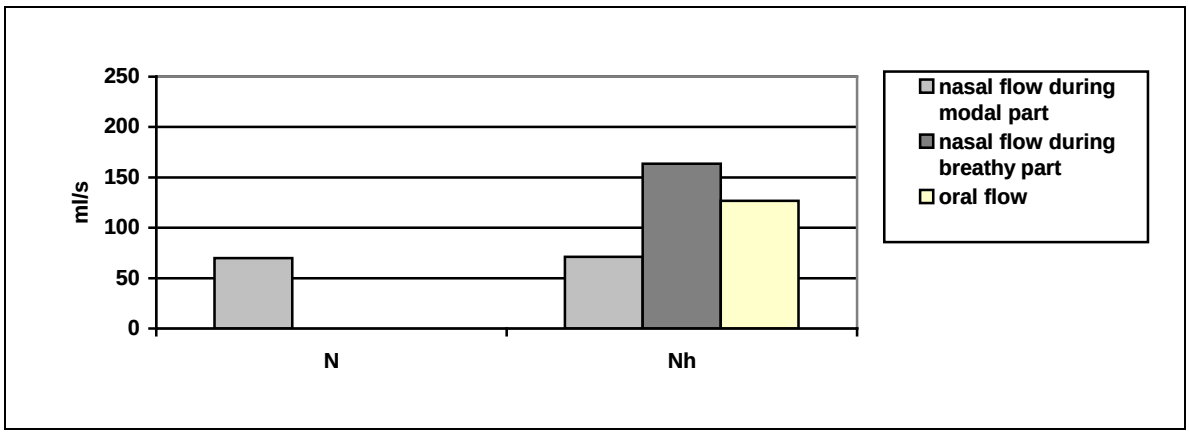


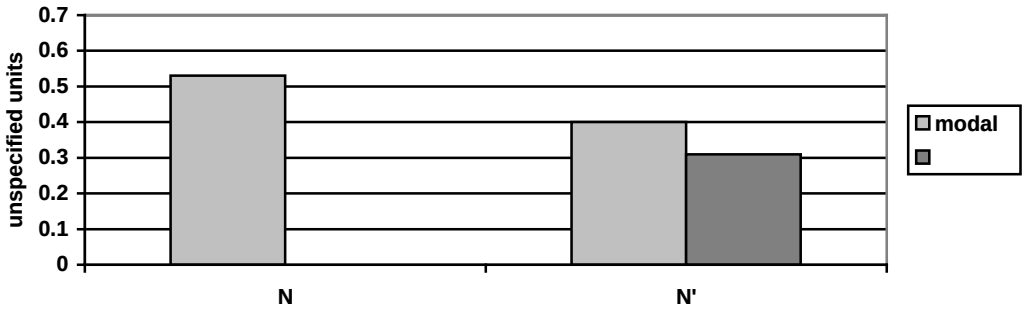
Figure 9: Graph of the average peak oral and nasal flow of [N] and /Nh/ for Bengali.



The [N̰] of Marathi and the /Nh/ of the Bengali have mostly nasal flow, with only a small amount of oral flow, suggesting that the /Nh/ of Bengali is more similar to [N̰]. On the other hand, Hindi /Nh/, in addition to the nasal flow, has a great deal of oral flow. This additional oral flow suggests the presence of two distinct segments, [N] and [h], in Hindi.

6.3.2 Closed Quotient

Figures 11, 12, and 13 are graphs of the average CQ value of the [N] and [N̰] for Marathi, and [N] and /Nh/ for Bengali and Hindi.



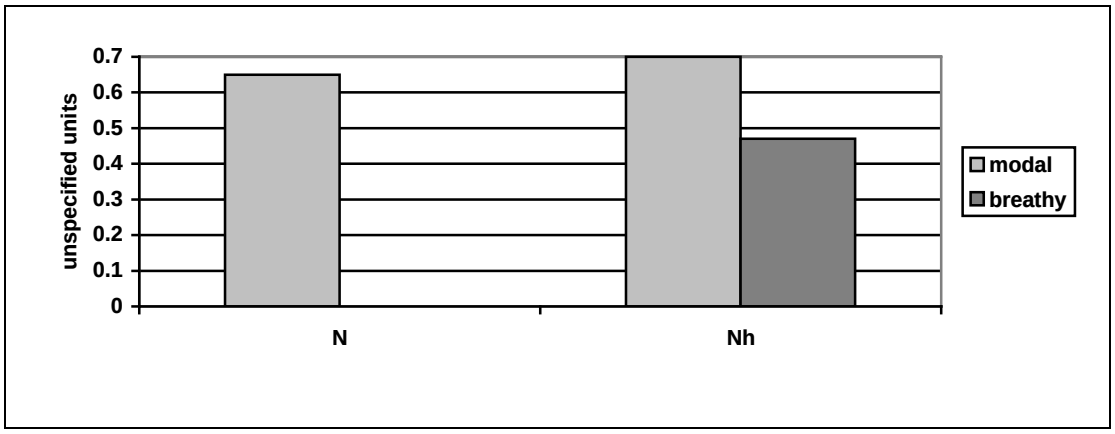


Figure 13: Graph of the average closed quotient value of [N] and /Nh/ in Hindi. The arrow is pointing in the direction of increased breathiness.

In Marathi, [N] has an overall lower CQ (i.e., is breathier) than modal [N]. This is not true of Bengali or Hindi, where the /Nh/ has a lower CQ only during the [h] portion;

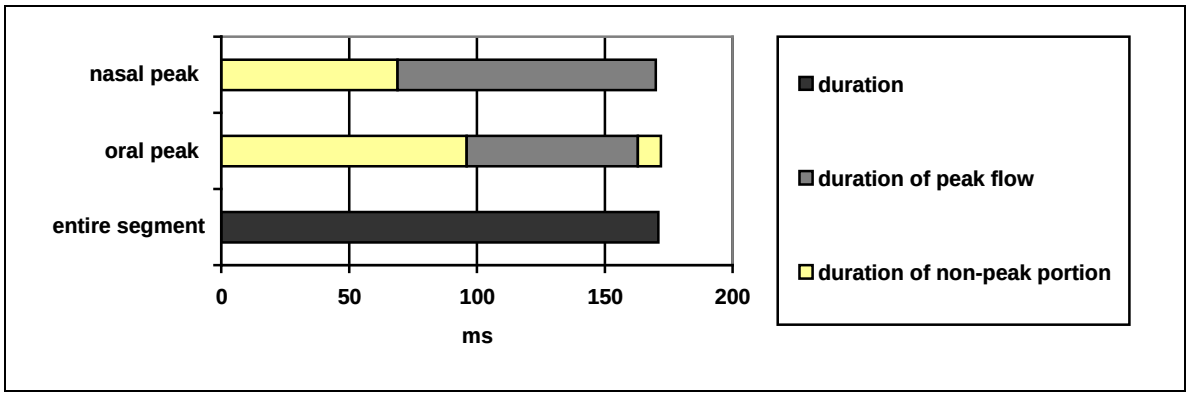


Figure 15: Graph of the average duration of peak airflow in Bengali /Nh/.

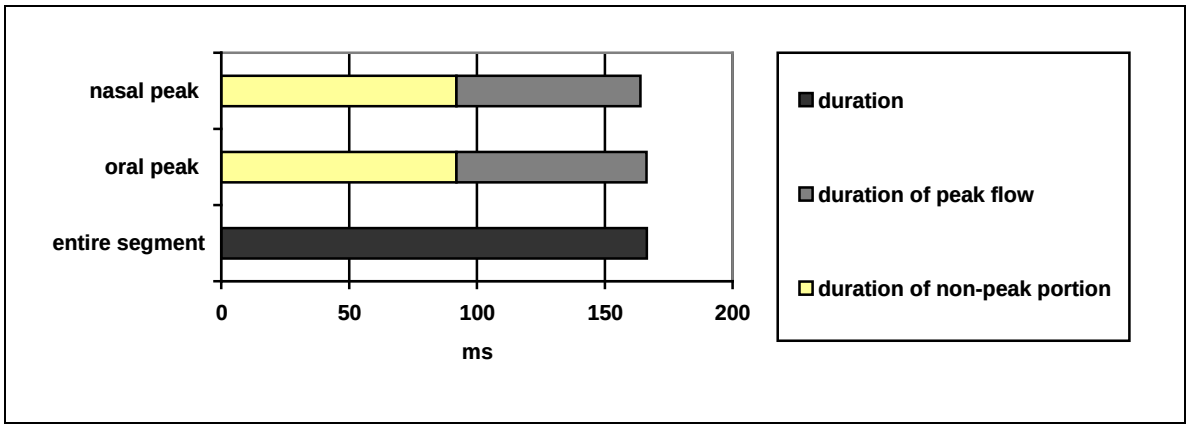


Figure 16: Graph of the average duration of peak airflow in Hindi /Nh/.

In Marathi, the [N̰] is composed mostly of nasal flow, with a small amount of oral flow near the end of the segment. This is not the case for Bengali and Hindi. In these languages, the /Nh/ is composed of similar parts nasal and oral flow, both taking place near the end of the segment, suggesting that these languages lack the type of breathy nasal found in Marathi.

6.3.4 Summary of Cross-language comparison

In sum, the Bengali and Hindi data do not pattern like that of Marathi, suggesting that the /Nh/ sequences in these languages are not breathy nasals. For example, when looking at duration of peak airflow, the Bengali and Hindi /Nh/ is composed of similar parts nasal and oral flow at the end of the segment, as opposed to Marathi, where the [N̰]

7.1 Methods

7.1.1 Speakers

The same Bengali and Hindi speakers used in the first part of the study were used in this section.

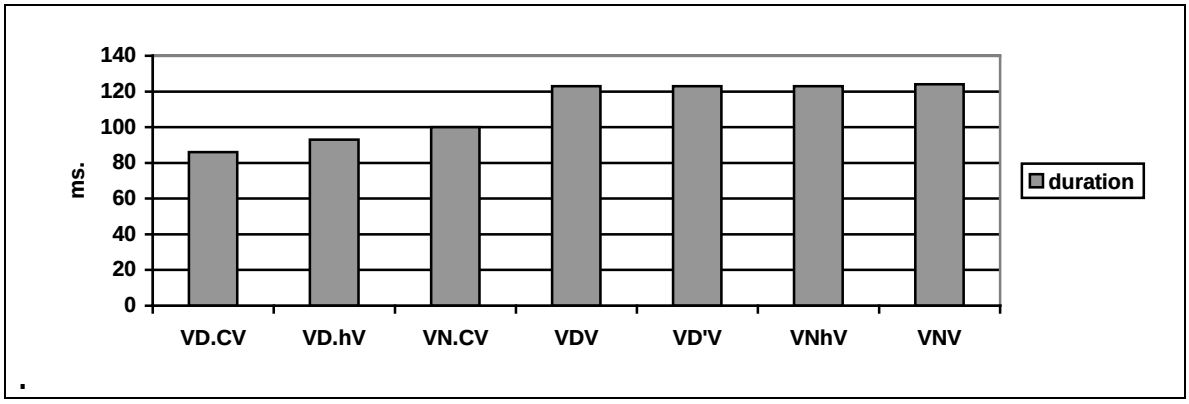
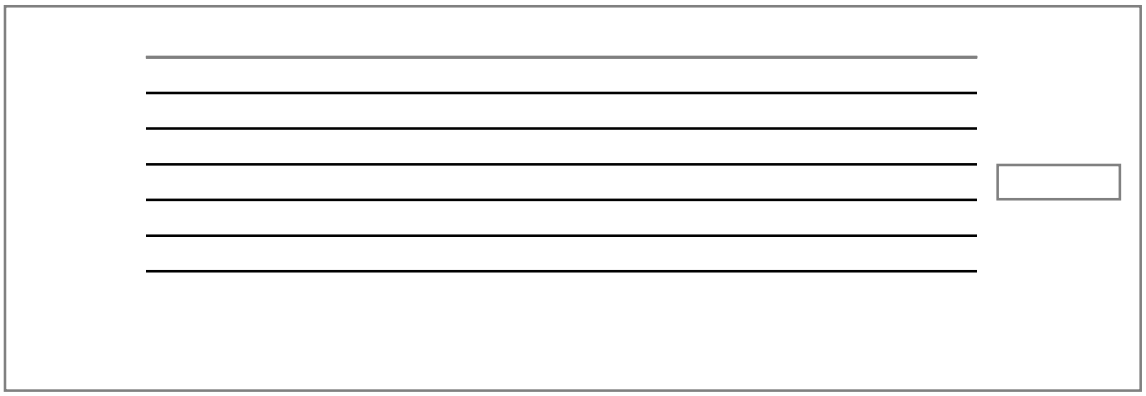


Figure 17: Graph of the vowel duration in the sequences [VD.CV], [V.DhV], [VN.CV], [VDV], [VD^{h}V], /VNhV/, and [VNV].

As stated previously, Bengali vowels followed by a single segment ([D], [D^{h}], [N]) are longer than those followed by clusters ([D.C], [D.h], and [N.C]). The duration of the vowel is longer than 120 ms in each of the environments where a single segment followed the vowel, while the duration of the vowel is consistently shorter than 100 ms in all of the closed-syllable environments (i.e., where the vowel is followed by a sequence of two consonants). It turns out that vowels followed by /Nh/ are in fact longer than 120 ms, thus suggesting that /Nh/ is behaving more like a single segment than like a cluster of [N.C], [D.C], or [D.h].

7.3.2 Hindi

Figure 18 is a graph of the average duration of [ʊ] in [VDhV], [VDCV], [VNCV], [VDV], [VNV], /VNhV/, and [VD^{h}V]. Duration is graphed on the y-axis in ms.



From the results found here, it appears that there is little reason to posit breathy nasals for standard Hindi; neither phonetic nor phonological evidence points to their existence. Bengali, on the other hand, may have a phonological breathy nasal that simply lacks some of the phonetic attributes of its Marathi counterpart.

The results obtained in this study raise an interesting issue about the typology of breathy nasals in South Asia. Based on the evidence examined here, it appears that there might be three different treatments of /Nh/'s in Indic languages: phonemic breathy nasals (as in Marathi [N̤]), [Nh] clusters (as in Hindi), and a possible transitional state between the two (as in Bengali). It is necessary to examine /Nh/ data from a variety of other Indic languages (e.g., Nepali, Gujarati, or Punjabi) to see if this predicted three-way classification holds true for the rest of the family.

References

Aziz, O. 2002. “Nasal Aspirates in Urdu”. Center fo