Breathy Nasals in Indic Languages

Christina M. Esposito (1), Sameer Khan (2), Alex Hurst (3)

UCLA Dept. of Linguistics

1. esposito@humnet.ucla.edu  2. sameer@humnet.ucla.edu  3. ahurst@ucla.edu
### Background: Breathiness in Indic Languages

- In addition to having a four-way oral stop contrast, breathiness in some Indic languages (like Marathi) extends to nasal stops.

<table>
<thead>
<tr>
<th></th>
<th>Bengali</th>
<th>Hindi</th>
<th>Marathi</th>
</tr>
</thead>
<tbody>
<tr>
<td>vd asp.</td>
<td>बैठि ‘kiln’</td>
<td>बाल ‘forehead’</td>
<td>ओरधा ‘half’</td>
</tr>
<tr>
<td>vd unasp.</td>
<td>बाति ‘bowl’</td>
<td>बाल ‘hair’</td>
<td>कांडा ‘onion’</td>
</tr>
<tr>
<td>vl asp.</td>
<td>प्हाति ‘I burst’</td>
<td>प्लाल ‘knife blade’</td>
<td>त्साॅव्था ‘fourth’</td>
</tr>
<tr>
<td>vl unasp.</td>
<td>पाटि ‘mat’</td>
<td>पाल ‘take care of’</td>
<td>राठा ‘road’</td>
</tr>
</tbody>
</table>

- Bengali and Hindi do have sequences of phonemic [N]+[h]. However, it is not clear if these are single breathy nasals or [Nh] clusters.
Note on terminology

- Traditionally, segments like $[\text{b}^\text{h}]$ are referred to as “voiced aspirated”.
- We will use the term “breathy” for the nasal counterparts.

<table>
<thead>
<tr>
<th>Terminology</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Nh}$</td>
<td>a nasal stop followed by an $[\text{h}]$</td>
</tr>
<tr>
<td>$\text{Dh}$</td>
<td>a voiced oral stop followed by an $[\text{h}]$</td>
</tr>
<tr>
<td>$\text{D}^\text{h}$</td>
<td>voiced aspirated</td>
</tr>
<tr>
<td>$\text{Th}$</td>
<td>voiceless aspirated</td>
</tr>
</tbody>
</table>

Terminology
Previous work on breathiness in Indic languages has focused on oral stops (Ohala 1983, Dixit 1987, to name a few.)

Breathy nasals in a closely related language:
  – Urdu (Aziz 2002)

  • Speakers were asked to produce word-initial, word-medial, and word-final [N]+[h] sequences.
    – Word-Initial – Spectrographic evidence revealed that there was no breathiness on the nasal, and that speakers consistently inserted a schwa between the nasal and the [h].
    – Word-Medial – Speakers were inconsistent in their pronunciation of [h].
    – Word-Final – Speakers nasalized the preceding vowels creating a [\tilde{v}(h)] sequence. (The nasal was always deleted.)
Current Study

- There is disagreement in the literature as to whether the Hindi [Nh] is single breathy nasals or clusters.
  - clusters (Ohala 1983; Botma 2004)
  - single breathy nasals (Maddieson 1984; Hinskens and van de Weijer 2003)

- No previous research has been done on Bengali [Nh].

Do [N]+[h] behave like single breathy nasals or [Nh] clusters in Bengali/Hindi?

- To answer this question, simultaneous audio, aerodynamic, and electroglottographic (EGG) recordings were made of Bengali, Hindi, and Marathi speakers.
Specifically, the following comparisons were made:

1) **Within-language comparison** - Compare the duration, airflow, and EGG data of Bengali and Hindi voiced aspirated consonants to potentially breathy nasal ones.

2) **Cross-language comparison** - Compare the duration, airflow, and EGG data of Marathi phonemically breathy nasals to Bengali/Hindi.

3) **Vowel lengthening** -
   1) **Bengali**: Vowels are longer in open syllables than in closed syllables. Are vowels preceding a [Dh] or [Nh] more like vowels in an open or closed syllable?
   2) **Hindi**: Vowels are longer before breathy consonants than before modal ones (Maddieson & Gandour 1975). Are vowels longer before [Nh] than before [N]?
About the Languages

- **Bengali**
  - spoken by approximately 171 million speakers in Bangladesh and India

- **Hindi**
  - spoken by approximately 180 million speakers in northern India

- **Marathi**
  - spoken by approximately 68 million speakers in central India
Within- and Cross-Language Comparison Methods

- **Bengali**: 1 adult male and 1 adult female speaker
- **Hindi**: 1 adult female speaker
- **Marathi**: 1 adult female speaker

**Speech Materials**
- **Marathi**: The speaker produced 6 words per type, each repeated 3 times.

<table>
<thead>
<tr>
<th>Descriptions</th>
<th>Symbol Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Nasals</td>
<td>N</td>
</tr>
<tr>
<td>2. Breathy nasals</td>
<td>Ñ</td>
</tr>
<tr>
<td>3. Voiced Aspirated Stops</td>
<td>Dʱ</td>
</tr>
<tr>
<td>4. Glottal Fricative</td>
<td>h</td>
</tr>
</tbody>
</table>

• Uttered in the carrier sentence: [to _____ bda wats] “He said the word ___”

Each sound was either word-initial or intervocalic
- **Bengali and Hindi**: The speakers produced 6 words per type, each repeated 3 times.

<table>
<thead>
<tr>
<th>Descriptions</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Nasals</td>
<td>N</td>
</tr>
<tr>
<td>2. Potentially Breathy Nasals (nasals followed by [h])</td>
<td>Nh</td>
</tr>
<tr>
<td>3. Voiced Unaspirated Stops</td>
<td>D</td>
</tr>
<tr>
<td>4. Voiced Stops Followed by [h]</td>
<td>Dh</td>
</tr>
<tr>
<td>5. Voiced Aspirated Stops</td>
<td>Ḍ</td>
</tr>
<tr>
<td>6. Glottal Fricative</td>
<td>h</td>
</tr>
</tbody>
</table>

- **Carrier sentences**: **Bengali**: [ʃe ____ bollo] “S/he said ____” ; **Hindi**: [əb ____ kəhie] “Please say ____ now”

- About schwa insertion: speakers were not explicitly told to avoid inserting schwas between an [N] or [D] and the following [h].
Procedure

- Simultaneous aerodynamic, 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).

Audio, oral flow, nasal flow, and EGG of Marathi [puṇa] “again”
Procedure

The following measurements were taken for each language:

- **Duration**
  - duration of the modal/unaspirated portion (for [D] and [N], this would be the duration of the entire segment)
  - duration of the breathy/aspirated portion (for [h], this would be the duration of the entire segment)

<table>
<thead>
<tr>
<th>N</th>
<th>Nh or Ñ</th>
<th>D</th>
<th>Dh</th>
<th>Ñn</th>
</tr>
</thead>
</table>

- duration of peak oral and nasal airflow (as shown in the figure below)
- **Airflow**
  
  The peak **airflow** (oral or nasal, depending on the consonant being measured) was taken at two points:

  1) modal/unaspirated  2) breathy/aspirated.

  (for [D] and [N] there would only be one measurement, at the non-aspirated portion, and for the [h] only the aspirated portion)

- **Closed Quotient (CQ)**
  
  was measured automatically using AcQuirer at two points:

  1) modal/unaspirated  2) breathy/aspirated

Closed quotient = \( \frac{T_c}{T_c + T_o} \)

(From http://www.lpl.univ-aix.fr/~ghio/pedago-EggUK.htm)
Results – **Within-language comparison**

**Duration**

**Bengali**

- The duration of [Nh] is more like that of [Dh] than that of [Dʰ].
The modal half of [Nh] is just as modal as [N], or the first half of [D̄] or [Dh]

The breathy half of [Nh] is just as breathy as [h], or the second half of [D̄] or [Dh]
The airflow in [Nh] is mostly nasal, but does have a small amount of oral flow.
Results – Within-language comparison

Duration

Hindi

- In Hindi, schwas were occasionally inserted between 1) N and h and always between 2) D and h. Tokens with schwa were excluded from the analysis.

- The duration of [Nh] is similar to [\(D^h\)].
Results – Within-language comparison
Closed Quotient (CQ)

The CQ of the [Nh] is similar to the individual segments [N]+[h].

The [h] of [Nh] is less breathy than the aspirated part of [D^6].
Results – Within-language comparison

Airflow

Hindi

- [Nh] has both oral and nasal flow during its breathy ([h]) portion. It is likely that the nasal portion of the [h] is simply due to co-articulation.
Within-Language Comparison

Bengali

- **Duration:** \([Nh]\) is more like the **cluster** \([Dh]\) than like \([D^h]\) and other single segments.
- **CQ:** \([Nh]\) behaves like both the **cluster** \([Dh]\), and the **single** \([D^h]\).
- **Airflow:** The airflow in \([Nh]\) is mostly nasal suggesting **singleton**, but does have a small amount of oral flow.

Hindi

- **Duration:** \([Nh]\) is similar to the **singleton** \([D^h]\).
- **CQ:** \([Nh]\) has similar CQ values to the individual segments \([N]\)+\([h]\), suggesting a **cluster**.
- **Airflow:** \([Nh]\) has both oral and nasal flow during its breathy (\([h]\)) portion suggesting it is more like a **cluster**.
Results – Cross-language comparison

Closed Quotient (CQ)

- In Marathi, the [N] is overall breathier (i.e. has a lower CQ) than the [N].

- This is not true of Bengali or Hindi, where the [Nh] has a lower CQ only during the [h] portion.
Results – *Cross-language comparison*  
*Duration of Peak Airflow of [N]/[Nh]*

- **Marathi** – The [N] is composed mostly of nasal flow, with a small amount of oral flow near the end of the segment.

- **Hindi and Bengali** – The [Nh] is composed of similar parts nasal and oral flow, both taking place near the end of the segment.
Results – Cross-language comparison

Airflow

In **Marathi** and **Bengali** the [N]/[Nh] has mostly nasal flow. For Bengali, since the [h] is mostly nasal, it suggests that the [Nh] is realized as [N].

However, in the **Hindi** [Nh] there is considerable oral flow (in additional to nasal), suggesting co-articulation.
Summary

Cross-Language Comparison

Bengali

- **Duration of peak airflow:** The \([\text{Nh}]\) is composed of similar parts nasal and oral flow at the end of the segment, as opposed to Marathi, where the \([\text{N}]\) is has mostly of nasal flow. This suggests the Bengali \([\text{Nh}]\) is more like a **cluster**.

- **CQ:** The \([\text{Nh}]\) is breathy only during the \([h]\) portion, unlike Marathi, where the \([\text{N}]\) is breathier for most of the segment. This suggests the Bengali \([\text{Nh}]\) is more like a **cluster**.

- **Airflow:** In Marathi and Bengali the \([\text{N}]\)/\([\text{Nh}]\) has mostly nasal flow. This suggests that the Bengali \([\text{Nh}]\) is realized as a **singleton** \([\text{N}]\).

Hindi

- **Duration of peak airflow:** The \([\text{Nh}]\) is composed of similar parts nasal and oral flow, as opposed to Marathi, where the \([\text{N}]\) is has mostly of nasal flow. This suggests the Hindi \([\text{Nh}]\) is more like a **cluster**.

- **CQ:** The \([\text{Nh}]\) is breathy only during the \([h]\) portion, unlike Marathi, where the \([\text{N}]\) is breathier for most of the segment. This suggests the Hindi \([\text{Nh}]\) is more like a **cluster**.

- **Airflow:** In the Hindi \([\text{Nh}]\) there is considerable oral flow (in additional to nasal), suggesting co-articulation of nasal \([\text{N}]\) and oral \([h]\). This suggests the Hindi \([\text{Nh}]\) is more like a **cluster**.
Vowel Lengthening

**Bengali**
- In Bengali, vowels are longer in open syllables than in closed syllables.
  - Are vowels preceding a [Nh] more like vowels in an open or closed syllable?

**Hindi:**
- In Hindi, vowels are longer before breathy consonants than before modal ones (Maddieson & Gandour 1977).
  - Are vowels longer before [Nh] than before [N]? (or is the vowel as it would be before other clusters)
Procedure

Methods

– **Speakers** – 2 adult Bengali speakers; 1 adult Hindi speaker
– **Speech Materials** - The speakers were asked to produce words with VC₀V sequences, where the C₀ was one of the following:

  - voiced aspirated oral stop [VDʰV]
  - voiced unaspirated oral stop [VDV]
  - potentially breathy nasal stop [VNʰV]
  - modal nasal stop [VNV]
  - a cluster involving any of the above [VD.CV], [VN.CV], [VD.hV], etc.

Procedure

– Measured the duration (ms) of vowels before each consonant (or consonant sequence).
– Controlled for vowel quality and stress
Vowel Lengthening - Bengali

- In Bengali, vowels followed by a single segment ([D], [D̂], [N]) are longer than those followed by clusters ([D.C], [D.h], and [N.C]).

- Vowels preceding [Nh] are just as long as those preceding [D], [D̂], or [N]; in this sense, [Nh] behaves more like a single segment than a cluster of [N.C], [D.C], or [D.h].
Vowel Lengthening – Hindi

- Vowels before [Nh] are actually shorter than ones before [N].
- This is the exact opposite of what is found with vowels before [Dʰ], which are longer than those before [D].
**Conclusion - Do [N]+[h] behave like single breathy nasals or [Nh] clusters in Bengali?**

<table>
<thead>
<tr>
<th></th>
<th>Within Language Comparison</th>
<th>Cross-Language Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration</td>
<td>Cluster</td>
<td>Cluster</td>
</tr>
<tr>
<td>CQ</td>
<td>Inconclusive</td>
<td>Cluster</td>
</tr>
<tr>
<td>Duration of Airflow</td>
<td>Single</td>
<td>Single</td>
</tr>
<tr>
<td>Airflow</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vowel Lengthening</td>
<td>Single</td>
<td></td>
</tr>
</tbody>
</table>

- The evidence is split.
- At this point, it is not clear if Bengali [Nh] is a single segment or a cluster.
**Conclusion - Do \([N]+[h]\) behave like single breathy nasals or \([Nh]\) clusters in Hindi?**

<table>
<thead>
<tr>
<th></th>
<th>Within Language Comparison</th>
<th>Cross-Language Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Duration</strong></td>
<td>Single</td>
<td>Cluster</td>
</tr>
<tr>
<td><strong>CQ</strong></td>
<td>Cluster</td>
<td>Cluster</td>
</tr>
<tr>
<td><strong>Duration of Peak Airflow</strong></td>
<td>Cluster</td>
<td>Cluster</td>
</tr>
<tr>
<td><strong>Airflow</strong></td>
<td>Cluster</td>
<td>Cluster</td>
</tr>
<tr>
<td><strong>Vowel Lengthening</strong></td>
<td>Cluster</td>
<td>Cluster</td>
</tr>
</tbody>
</table>

- With the exception of duration, the evidence for Hindi points toward **cluster**.
- It is debatable where duration is a good indication of cluster status.
We would like to thank

– our speakers: Anuradha Deo, Farida Amin Khan, Shafiq ud Dowla Khan, Anoop Mahajan, Gyanam Mahajan

– the UCLA Phonetics Lab for their comments and suggestions