Representing a four-way contrast: Nepali, voiced aspirates, and laryngeal realism
Martha Schwarz, Morgan Sonderegger, Heather Goad (McGill University)

Theories of laryngeal realism (Honeybone 2005, Iverson & Salmons 1995, Beckman et al. 2011) argue for a tight correspondence between a segment’s phonetic cues and the (laryngeal) phonological features that represent it. Consequently, the ‘p’/’b’ contrast in French, expressed phonetically by vocal fold vibration during the stop closure is represented by a [voice] feature whereas the ‘p’/’b’ contrast in English, expressed phonetically by long and short lag VOT, is represented by a [spread glottis] feature. However, this theory predicts potential conflicts for segments specified for multiple laryngeal cues, i.e. voicing and aspiration. We therefore analyze Nepali’s four-way contrast between voiceless, voiced, voiceless aspirated, and voiced aspirated stops, the type of system Iverson and Salmons (1995) propose is represented by the features in (1) (assumes privative features).

(1) voiceless (T): voiceless aspirated (Th): voiced (D): voiced aspirated (Dh):
[ ] [spread] [voice] [spread, voice]

By examining a) presence of acoustic cues, b) diagnostics for feature specification by Beckman et al. (2011, 2013) of speech rate and intervocalic passive voicing, and c) final neutralization patterns, we find support for the representation in (1), with the caveat that the voicing contrast is “stronger” than the aspiration contrast. We propose that the strength of [voice] over [spread] is not represented through the architecture of a feature geometry but through temporally ordering the [voice] and [spread] features within a segment, and propose a Q theory representation (Inkelas & Shih, to appear) to do so.

Methods: Data was collected from 17 native Nepali speakers in Sikkim, India and consists of 20 target words with the relevant stops in initial, medial, and final position produced in a carrier phrase (n = 1490). Acoustic annotation/analysis was conducted in Praat (Boersma & Weenik 2016) for several phonetic cues argued to differentiate stops in Indo-Aryan languages, beyond VOT: prevoicing duration, ACT (‘After Closure Time’: period between release burst and onset of voicing of following vowel) and SA (‘Superimposed Aspiration’: period following release burst characterized by glottalic pulsing and friction noise).

Results: We first examine presence of cues and their capacity to distinguish one class from another word-initially. We find that prevoicing duration significantly distinguishes the voiced from the voiceless classes, supporting the representation of D and Dh with a [voice] feature as in (1). Post-release duration proved to be significantly longer on the aspirated classes than the unaspirated classes, seemingly supporting the specification of [spread] on Th and Dh. However, the primary cue is ACT for Th and SA for Dh. Although this is problematic for a strict reading of laryngeal realism since the [spread] feature corresponds to different cues on different classes, this can be potentially reconciled with Ridouane et al.’s (2011) expanded definition of [spread] that combines sounds like Th and Dh due to their acoustic, if not articulatory, similarity.

We next apply diagnostics for feature specification from Beckman et al.’s (2011, 2013) proposal that cues corresponding to specified features (i.e. aspiration on Th and Dh) are controlled by speakers, but that cues corresponding to unspecified features (i.e. aspiration on T and D) are automatic. They provide two contexts in which to diagnose this: the effect of speech rate on VOT durations word-initially, and passive voicing during stop closure intervocically. For speech rate, Beckman et al. (2011) finds that cue durations corresponding to specified features increase as speech rate slows but that cues corresponding to unspecified features do not. For passive voicing, Beckman et al. (2013) finds that stops specified for [voice] in Russian (a voicing language) are voiced throughout the closure in intervocalic position, but stops specified for [spread] in German (aspirating language) block passive voicing from the surrounding vowels.

Initial position - speech rate: Using mixed-effect models that evaluate the effects of speech rate, laryngeal class, and their interaction on duration of prevoicing, ACT, and SA, we find mixed results for this diagnostic in Nepali. Speech rate effects are as predicted by Beckman et al. for prevoicing
duration on both D and Dh classes ($\beta=-.025$, $p<.0001$). For ACT and SA duration, the interaction between class and speech rate goes in the expected direction ($\beta = -.016$ and -.191 respectively), but the effect is not significant ($p = .64$ and .18). Medial position - passive voicing: As seen in (1), Nepali has stops doubly-specified for both [spread] and [voice]. Thus, in contrast to the languages examined by Beckman et al. Nepali’s stops pose conflicting predictions. We find that voiced aspirated segments behave like a stop specified for [voice], not [spread], suggesting that the [voice] feature may be stronger than the [spread] feature.

Finally, we examine patterns of neutralization in final position. Previous descriptions report that Dh variably neutralizes to D, while the other classes remain distinct (Khatiwada 2009). We too find that voicing contrast is maintained but that the aspiration contrast is weaker in word final position, particularly within the voiced class, again suggesting that [voice] is stronger than [spread].

**Representation:** In many two- and three-way contrast languages laryngeal neutralization involves loss of all laryngeal contrast (Iverson & Salmons 2011), but Nepali exhibits loss of aspiration contrast independent of voicing. We account for this with a Q Theory representation. Q Theory (Inkelas & Shih, to appear) divides every segment into three subsegment feature matrices that correspond to the temporally ordered internal phases of segments. Building on Q Theory’s representation of [k] as (k k k) and [kʰ] as (k k h) as well as the option for a segment like [ts] with two phases to be realized as either (t s s) or (t t s) depending on the phonetic and phonological properties of the language, we revise Iverson and Salmon’s representation of Nepali’s stop classes in (1) by proposing the following (strikethrough shows effect of word-final feature loss):

<table>
<thead>
<tr>
<th>T</th>
<th>D</th>
<th>Th</th>
<th>Dh</th>
</tr>
</thead>
<tbody>
<tr>
<td>(t t t)</td>
<td>[ ] [ ] [ ]</td>
<td>(d d d)</td>
<td>(t h h)</td>
</tr>
<tr>
<td>[voice]</td>
<td>[voice]</td>
<td>[voice]</td>
<td>[voice]</td>
</tr>
<tr>
<td>[spread]</td>
<td>[spread]</td>
<td>[spread]</td>
<td></td>
</tr>
</tbody>
</table>

This representation allows for loss of one feature independent of the other on Dh because the features are associated with separate subsegments. If neutralization targets the features at the rightmost edge due to impoverished cues in this position (Steriade 1997) it predicts correctly that Dh loses [spread] rather than [voice] since [spread] aligns farther to the right – reflecting that phonetically prevoicing precedes aspiration. To account for the loss of [spread] on Dh but not Th, Dh is represented with only one [spread] segment while Th has two, making it robust to right-edge feature loss. This is phonetically motivated by the shorter burst duration of Dh compared to Th word-initially. That two thirds of Dh is specified for [voice] further explains the presence of passive voicing through intervocalic Dh which is expected of a [voice] (sub)segment but not [spread].

This paper thus identifies potential challenges Nepali poses for laryngeal realism and finds that some aspects of the theory transfer to doubly-specified segments without issue, some raise questions for further study (whether speech rate effects hold up for all specified features), and proposes that others (passive voicing, final neutralization) may be solved with a representation in which the features [voice] and [spread] are temporally ordered rather than simultaneously specified. The proposal makes several predictions in the spirit of laryngeal realism and cue to feature correspondence that should be tested in other languages with doubly-specified segments. For example, in a language with phonological (e.g. final neutralization) reasons to specify Dh as (d h h) the aspiration on Dh will not be shorter than on the language’s Th (t h h) stop, and that in doubly-specified segments it should always be the feature corresponding to the rightmost phonetic cues that neutralizes independent of the other.