Wubuy coronal stop perception by speakers of three dialects of Bangla

Rikke Louise Bundgaard-Nielsen¹,², Brett Baker³, Olga Maxwell³, Janet Fletcher³

¹ MARCS Institute/University of Western Sydney
² La Trobe University
³ University of Melbourne

rikkelou@gmail.com, bjbaker@unimelb.edu.au, omaxwell@unimelb.edu.au, j.fletcher@unimelb.edu.au

Abstract

We tested native speakers from three major dialect groups of Bangla, on their discrimination of a four-way coronal stop contrast from the Australian Indigenous language Wubuy. Bangla is generally assumed to have a two-way contrast in coronal stops, with an additional place distinction in affricates. The results show that Bangla speakers are able to discriminate the Wubuy contrasts, but also that certain contrasts are more difficult to discriminate than others. We discuss these results with respect to the Bangla coronal inventory, and importantly, with respect to the variation in the phonetic realisation of coronals between the dialects of Bangla. We argue that the phonetic realisation of what is regarded to be the 'same' phonemic inventory can have implications for the perceptual behaviour of speakers.

Index Terms: perception, coronal place contrasts, phonology, variation

1. Introduction

Research has shown that multiple coronal stop contrasts are difficult to discriminate, even for native speakers [1], [2]. However, little research has investigated whether having such experience necessarily allows listeners to do better than those without an elaborated series of coronal place contrasts. In other words, does an experience of multiple contrasts within the coronal set of gestures by itself lead to better discrimination of coronal contrasts in other languages, or is particular experience with specific contrasts necessary? Moreover, is it the case that the phonological behaviour and the phonetic correlates of the phonemic inventory of coronals play a role in the behaviour of listeners?

We tested these questions with a two-part discrimination study with speakers of Bangla, an Indo-Aryan language spoken by approximately 180 million people primarily in West Bengal (India) and Bangladesh (discussed in 1.1 below). Bangla has a consonant inventory contrasting multiple coronals: a ‘dental’ and ‘retroflex’ series of stops, and an alveopalatal stop or affricated stop series. Our stimulus materials are taken from a language with an even more complex coronal contrast set, the Australian Indigenous language Wubuy [3], discussed in 1.2, below.

We expected the Bangla listeners to be able to discriminate coronal contrast pairs easily, to the extent that they matched their native (L1) phonemic categories, and to the extent that there is overlap in the phonetics of the realisation in each language. We also expected that they would find non-native contrasts without native language correspondences more difficult, just as most theories of cross- and second language (L2) speech perception predict (PAM/PAM-L2; [4],[5],[6]). However, there is disagreement in the literature about what, exactly, are the articulatory and acoustic correlates of the Bangla coronal segments. There are also reported to be differences between the various dialects of Bangla in the implementation of the contrasts, and their phonetic correlates (discussed in 1.1), such that speakers in the north of Bangladesh are said to have largely 'alveolar' (i.e. non-retracted, perhaps) realisations of what are described as ‘retroflex’ segments in the south of the country. We therefore also expected to find some differences between groups in the extent to which they could discriminate coronal contrast pairs depending in addition on the phonetics of the contrast in their particular dialect.

We first present a discussion of Bangla coronal phonology and phonetics (1.1), followed by a discussion of Wubuy (1.2). In 1.3, we present predictions for the discrimination of Wubuy coronal stops by Bangla listeners, before we present our method, results and discussion (in 2., 3., and 4., respectively).

1.1. Bangla coronal phonology and phonetics

Existing studies of Bangla, [7], [8], [9] argue for a two-way phonemic contrast in coronal stops, in terms of place of articulation. These stops are typically described as ‘dental’ and ‘retroflex’, but it is unclear from the available descriptions what are the articulatory and acoustic correlates of these coronals. [10] labels the three series ‘lamino-dental’, ‘apico-postalveolar’ and ‘lamino-postalveolar’, and also notes that the retroflex is rare, word-initially. [11] argues that the lamino-postalveolar affricate in Indo-Aryan languages is simply a third coronal position of lamino-postalveolar/alveopalatal, as in Wubuy, with the affricated release being the inevitable consequence of the nature of the constriction.

Reports also suggest that there are systematic differences in the realisation of these stops between the regions of Bangladesh (e.g., [8], [9], [12]). While this regional variation has yet to be systematically examined (acoustically, articulatorily or perceptually), the variety spoken in the capital Dhaka and in the southern regions is generally considered similar to the variety spoken in West Bengal, India, in and around Kolkata [8], [9]. The participants tested in this study also belong to two further dialect groups: Rangpur and Nalitabar in the north and north-east, and Sylhet, in the far east of Bangladesh.
1.2. Wubuy coronal phonology and phonetics

Wubuy is an Australian Indigenous language, spoken in the community of Numbulwar in the Northern Territory of Australia. Like many other Australian languages, it has a four-way coronal contrast /t ʈ t̪ ȶ/ in both stops and nasals, with two laminals (lami-no-dental and lami-no-alveopalatal) and two apicals (apico-alveolar and apico-retroflex) [3], [13], [14]. Unlike many other languages with an apical contrast, however, Wubuy contrasts these segments word-initially, allowing us to test discrimination in both the word-medial and word-initial environment, using stimuli derived from real words.

Wubuy coronal stops are systematically differentiated, to varying extents, by acoustic information intrinsic and extrinsic to the consonants themselves. In terms of intrinsic differentiation, data collected from a single speaker [13] suggest that the four coronal stops of Wubuy differ in terms of the spectral characteristics of the stops, at least impressionistically. According to [13], the stops are set apart by very subtle differences in the burst noise and in formant transitions into and out of the consonants. Dental /t/ reportedly has a slightly higher burst frequency than alveolar /t̪/, while the difference in burst frequency between /ʈ/ and the retroflex /ȶ/ is reportedly almost non-existent, with the majority of place-information carried on the preceding vowel (if present). The lami-no-palatal /h/ is characterised by a lower burst frequency, and a burst that remains relatively high in intensity for a longer period, than the other three Wubuy coronal stops.

Other data from a recent study of the temporal characteristics of Wubuy coronal stop closures using 3D electromagnetic articulometry (EMA) further indicate that the stops differ in terms of the duration of the relatively stable ‘hold’ phase (consonant ‘nucleus’) of the stop relative to the duration of the entire consonant gesture [15]. According to this study, the retroflex stop exhibits a significantly shorter normalised closure period relative to the other coronals, and the two apical stops (alveolar and retroflex) have shorter relative consonant nucleus durations than the two laminal stops (dental and palatal). The EMA data [16], [17] also suggest dynamic differences involved in the articulation of the coronals. As suggested by [16] as well as other authors [18], the primary difference between the apicals is not constriction location, but gestural dynamics. While the apico-alveolar involves a simple tongue tip raising gesture, the apico-retroflex instead involves a curling back of the tip followed by a back-to-front ‘flapping out’ gesture wherein the tip makes contact with the alveolar ridge. The extent of variability in the realisation of apicals in Wubuy has not been examined. But we know from other Australian languages [19], as well as the reports of fieldworkers, that apico-retroflexes in particular appear to have highly variable realisations. In particular, at least some of their realisations appear to be very similar to the canonical apico-alveolar.

In terms of acoustic phonetic cues, the stops are differentiated by vowel transitions into and out of the stops: Wubuy coronal stops are differentiated by F1 and F3 in an /Ca/ context in the preceding vowel [14], in a manner consistent with previous accounts of other Indigenous Australian languages, which have indicated that these three coronal stops are not reliably differentiated in the F2 of the preceding vowel transitions [20]. The stops are also differentiated by F2 and/or to various degrees by F1 and F3 in the following vowel in the /Ca/ and /Ca/ contexts, in a manner generally consistent with previous findings that place of articulation information is typically carried in the F2 transition in the following vowel in a /CV/ syllable (e.g., [21], [22], [23]. For a thorough review, see [24]).

The findings also clearly suggest that the different stop contrasts differ in the degree to which the contrast is differentiated more by preceding or following vowel transitions, consistent with other reports on place of articulation information in vowel transitions into and out of consonants. Moreover, differences in the position of a stop in the word may systematically affect the way in which place of articulation differences can be measured [14].

1.3. Discrimination predictions

On the basis of the above characterisations of Wubuy and the three Bangla dialects (Dhaka; Eastern; and Northern) it is possible to make a number of predictions for the discrimination of Wubuy coronal stops by Bangla listeners. Firstly, we predict that all participants will find discrimination of any contrast involving the lamino-alveopalatal easy, as all three dialects maintain a robust distinction between a stop (or affricated stop) of this kind and at least one other coronal place of articulation.

With respect to the remaining contrasts, we predict that all speakers will discriminate /t ʈ t̪ ȶ/ more accurately than the other two contrasts, due to greater acoustic differentiation [14]. In the case of the listeners from Dhaka, this prediction is also commensurate with the claim that this is a native phonological contrast produced with canonical Bangla phonetic realisations. However, as the ‘retroflex’ category is often realised with an alveolar place of constriction [25], we also suggest that speakers may treat the apical contrast /t ʈ/ as allophonic (and thus find it difficult to discriminate). With respect to the two other dialect groups, we have too little information to enable firm predictions. However, in sections 3 and 4 we offer some tentative proposals linking the discrimination accuracy of these groups to potential differences in their coronal inventories and to the phonetic realisations of coronal stops across these dialects.

2. Method

2.1. Stimuli

The stimuli for the experiment were based on the audio recordings of three female native speakers of Wubuy (aged 5-61 years), born and raised in the Numbulwar area by native speakers of Wubuy. The participants were also able to understand and speak the community language Roper Kriol [27] to some extent; and two participants spoke the neighbouring language Enindhilyakwa [28] with relatives. All three speakers had acquired English as a second language in a classroom setting and had some basic linguistic training.

Each participant produced the three target consonants /t ʈ t̪/ in two prosodic contexts: word-medially in an /Ca/ context, and word-initially in the context /Ca/ (i.e. following pause). In addition, they produced the lamino-alveopalatal /h/ in the intervocalic /Ca/ context. The target words were all real Wubuy words, selected to provide a symmetrical vowel context on either side of the consonants in the /Ca/ context and imbedded in carrier phrases (see [14] for details). Five correct utterances were recorded for each target, resulting in a total of 45 correct utterances (5 tokens per 4 targets in the /Ca/ context, and 5 tokens per 3 targets in the /Ca/ context). Recordings of targets containing any dysfluencies or reading errors were discarded.
The recordings were made in a sound-attenuated professional recording studio at MARCS Institute in Sydney with a Shure SM10A headset cardiod microphone, an EDIROL UA-25 USB audio interface, and a laptop computer with Cool Edit 2000. All recordings had a 16-bit sampling depth with a sampling rate of 44.1 KHz. Subsequently, the target /aCa/ and /##Ca/ sequences were excised using a Praat script and checked by the first and second author. Each excised token was given a 20 ms ramp-in and a 10 ms ramp-out using Praat.

2.2. Stimulus presentations

The Bangla listeners were presented with two separate randomised cross-speaker categorical XAB discrimination tasks (Study 1: the /aCa/ context, Study 2: the /##Ca/ context) programmed in Pyscope X, with the stimuli presented over headphones from a MacBookPro laptop computer. The XAB discrimination task was explained to the participants as a task in which a ‘teacher’ (the first voice heard) was being imitated by a ‘good student’ and a ‘bad student’ (voices 2 and 3), and it was the task of the participant to indicate (with a key press on the keyboard) which of the two students was the ‘good student’ who copied correctly what the teacher had said. We provided this explanation to make the experimental paradigm meaningful to our participants and increase the likelihood that they would understand the task at hand and be able to complete the experiments.

The inter-stimulus interval (ISI) between stimuli within each XAB trial was 500 ms. The response window was presented for 2 seconds following the playing of the third target, with the trial-interval of 1 second. If the participant did not respond within the 2-second window, the trial was replayed later.

In both studies, the participants were presented with 6 unique triads per contrast type, with a total of 12 contrasts (all combinations of the four consonants /t ʈ t̪ ȶ/ as there might be differences in discriminability depending on the order of presentation) in the /aCa/ context (72 trials total per listener) and 6 contrasts (all combinations of the three consonants /t ʈ t̪/, again allowing for differences in discriminability due to the order of presentation) in the /##Ca/ context (36 trials total per listener). The participants first completed the discrimination of the target consonants in the phrase-medial context (Study 1).

2.3. Participants

10 female and 11 male native speakers (age range 19-38 years, except for one male speaker in his early seventies) of Bangladeshi Bangla participated in the study. All speakers reported having normal hearing. 14 of the participants were from Dhaka, nearby Narayanganj and Mathkola, Sathkira near the border with India, or Chittagong. Another three participants were from the north-eastern region of Bangladesh near the border with the Indian state of Meghalaya (Sylhet), and two participants were from the northern and north-western parts of the country (Nalitabari, Rangpur).

All participants spoke English, though their competence varied from intermediate to advanced. 15 also spoke Hindi; 6 could read and understand Arabic; and one spoke French and Spanish. All went to Bangla-medium schools and had acquired English as a second language from between the ages of 4 and 12 years. All were living in Melbourne at the time of the experiment and reported using Bangla with family and in most social contexts. The testing took place in a professional recording studio at the University of Melbourne. All procedures were explained to the participants in English by the third author. Two female speakers were excluded from the analyses below as they failed to reach our discrimination criterion of an individual average discrimination accuracy (over the nine contrasts tested in Study 1 and Study 2 combined) above chance (50% correct discrimination).

3. Results

Due to the suspected differences in the phonetic realisations of their coronal stops in different varieties of Bangla, we analysed the discrimination results from the three distinct listener groups separately. The results from the Dhaka (and nearby Western dialect) listeners are presented in 3.1. The preliminary results from the other participants are presented in 3.2, in appropriately cautious fashion, given the very low numbers of participants from each of these groups.

3.1. Dhaka listeners

The discrimination results from Study 1 and 2 for the 14 Bangla participants from Dhaka and nearby regions of Bangladesh are presented in Figure 1.

To assess whether the Bangla listeners were able to discriminate the Wubuy coronal stop contrasts in Studies 1 (/aCa/) and 2 (/##Ca/), we first conducted a series of one-sample t-tests against chance performance. The results (see Table 1, below) indicate that the discrimination accuracy of the Bangla listeners was significantly above chance performance (50% correct discrimination).

<table>
<thead>
<tr>
<th>Context</th>
<th>Contrast</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>/aCa/</td>
<td>/t t̪/</td>
<td>5.486</td>
<td>27</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>/aCa/</td>
<td>/t t̪/</td>
<td>17.062</td>
<td>27</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>/aCa/</td>
<td>/t ʈ/</td>
<td>6.809</td>
<td>27</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>/aCa/</td>
<td>/t ʈ/</td>
<td>14.932</td>
<td>27</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>/aCa/</td>
<td>/t̪ ʈ/</td>
<td>12.312</td>
<td>27</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>/aCa/</td>
<td>/t̪ ʈ/</td>
<td>15.377</td>
<td>27</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>/##Ca/</td>
<td>/t t̪/</td>
<td>5.037</td>
<td>27</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>/##Ca/</td>
<td>/t t̪/</td>
<td>5.229</td>
<td>27</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>/##Ca/</td>
<td>/t t̪/</td>
<td>4.46</td>
<td>27</td>
<td>&lt; .001</td>
</tr>
</tbody>
</table>

We then conducted two separate One-Way ANOVAs to assess if there were any differences in discriminability of the
Wubuy contrasts in the two studies. The results show that there was a main effect of contrast in Study 1 (the /aCa/ context); F(5,162) = 7.866, p < .001. Bonferroni-corrected post hoc comparisons revealed that the main effect of ‘contrast’ in Study 1 was due primarily to /t ʈ/ being discriminated less accurately than most other contrasts (see Table 2, below), along with the apical contrast /t̪ ʈ/. There was no main effect of contrast in Study 2, indicating that all three contrasts were equally well discriminated by the listeners, though the discrimination accuracy was only ~70%.

Table 2. Bonferroni post hoc comparisons of the discriminability of the Wubuy contrasts, α = .003 to correct for multiple comparisons.

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Mean diff.</th>
<th>S.E.</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>/t̪ t/ /t ʈ/</td>
<td>-1.18452</td>
<td>0.04293</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>/ʈ t̪/ /ʈ ʈ/</td>
<td>-1.1762</td>
<td>0.04293</td>
<td>0.001</td>
</tr>
<tr>
<td>/t ʈ/ /t̪ ʈ/</td>
<td>-1.1667</td>
<td>0.04293</td>
<td>0.002</td>
</tr>
<tr>
<td>/t̪ t/ /t ʈ/</td>
<td>-1.19643</td>
<td>0.04293</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>/t̪ ʈ/ /t ʈ/</td>
<td>-1.16071</td>
<td>0.04293</td>
<td>0.004</td>
</tr>
</tbody>
</table>

3.2. Preliminary results from North-eastern and North-western Bangla listeners

The preliminary discrimination results from the three Bangla participants from the north-eastern (Sylhet) region of Bangladesh are presented in Figure 2, while the preliminary results from the two participants from the north-western region are presented in Figure 3. Due to the low Ns in both these groups, we did not conduct any statistical testing of the average discrimination results.

Although we have limited information on the phonetic realisation of the stop contrast in North-western dialects of Bangla, the behaviour of these two speakers suggests that their ‘retroflex’ category might be realised with canonical alveolar articulation, most of the time. Their discrimination accuracy on the /ʈ t̪/ contrast was the strongest of the three groups, and would be commensurate with a phonological system where the phonetic realisation was close to that of the Wubuy pair. Like the other two groups, they also found the /t ʈ/ contrast relatively easy, although the /t̪ ʈ/ was difficult for them, suggesting that these sounds are allophones in this dialect.

4. Discussion

The present studies show that speakers with multiple coronal places of articulation can recruit their native language experience to discriminate phones from other languages with elaborated coronal inventories. The results also suggest that successful discrimination does not depend solely on shared phonology but also on the specific phonetic characteristics of the languages/dialects in question, as phonetic overlap—or lack thereof—directly and systematically influences coronal stop discrimination. For the Dhaka listeners, the relatively poor discrimination of /ʈ t̪/ suggests that listeners’ behaviour is shaped by a phonetic overlap between /t ʈ/ as /ʈ/ can have alveolar-like realisations. There also appears to be a difference in the discrimination across contexts, utterance-initial /t ʈ/ is discriminated poorly than utterance-medial /t ʈ/. This may be related to differences in frequency in the input, as in the case of Wubuy listeners [2]. Although our information on the phonetics of the coronals in other Bangla dialects is sketchy, the Northern dialect is said to be ‘part of the continuum from Bengali to Assamese’ [8] where (at least according to some descriptions, e.g. [26]), the coronal contrast is /ʈ t̪/ rather than /t ʈ/ as in other Indo-Aryan varieties. This would fit with the discrimination behaviour of these speakers: where /t ʈ/ are discriminated much better than the other groups. Listeners from the Eastern dialect Sylhet may be under the influence of neighbouring Austroasiatic and Tibeto-Burman languages which have a single coronal stop category, and may thus find all coronal contrasts difficult.

5. Acknowledgements

We would like to thank our Wubuy speakers and Bangla participants and the University of Melbourne for a Faculty Research Grant. We would also gratefully acknowledge our collaborators Cathi Best, Mark Harvey and Christian Kroos for their gracious permission to use the Wubuy material funded by grants to Best, Baker and Harvey.
6. References


