An acoustic study of voiceless stops in Indian English

JAHNAVI NARKAR, *University of California–Los Angeles*
PETER STAROVEROV, *independent scholar*

Received AUGUST 2021; Revised DECEMBER 2022

ABSTRACT

This paper analyzes phonetic corpus data showing that speakers of Indian English produce word-initial and post-/s/ stops that do not differ in terms of voice onset time (VOT). They systematically produce both categories with short positive VOT which is typical of unaspirated stops. By contrast, speakers of British English do distinguish aspirated and unaspirated stops in terms of VOT in the same contexts. These results are based on VOT measurements for word-initial and post-/s/ bilabial and dorsal stops produced by 102 bilingual speakers of one of nine Indic languages and Indian English, and by 28 monolingual British English speakers. Given that Indian English is influenced by Indic languages which have phonological voiceless aspirated stops, this pattern is particularly surprising. We suggest that this cross-language category mismatch might be driven by the phonetic implementation of aspirated stops in Indic languages. More generally, if our results are representative of the stage in which Indian English and British English were in direct contact, they suggest that the existence of a richer laryngeal system in borrowing languages cannot in itself prepare the speakers to perceive and produce borrowed laryngeal contrast more faithfully. Finally, we also measured vowel onset f0 in our corpus, but found no significant differences by context – word-initial and post-/s/ stops were followed by comparable onset f0. These results suggest that intonational context may need to be further controlled for more sensitive f0 analysis.

1 Introduction

This paper establishes the acoustic correlates of voiceless stops in Indian English. It shows that English speakers of Indic (also known as Indo-Aryan) languages, which have phonological voiceless aspirated stops in their inventory, systematically produce English stops with short-lag, rather than long-lag VOT. Although this seems unexpected from a systemic phonological point of view, we argue that the phonetic implementation of aspirated stops in Indic languages likely contributes to this cross-language category mismatch (see also Sailaja (2009) on how orthography may contribute to this phenomenon).

Although our study does not directly address loanword adaptation, this pattern of adaptation in Indian English is also reported to apply when English words are borrowed into Indic languages (Paradis and LaCharité 1997). It is, therefore, a potential challenge for theories of loanword adaptation based on featural closeness (Paradis and LaCharité 1997, LaCharité and Paradis 2000, 2005), as aspirates in English and Indic are often thought to have the same features (e.g. Iverson and Salmons 1995). We speculate that the existence of a richer laryngeal system in the borrowing language does not in itself prepare the speakers to perceive and produce the English laryngeal contrast more faithfully.

Existing reports of the production of voiceless stops in Indian English are largely intuitive and for the most part, do not cite detailed acoustic data. A notable exception is Awan and Stine (2011) who compared the VOT of bilabial stops in the words *path* and *apparently* and coronal stops in the words *token* and *foretell* in Indian English and American English. They found that Indian English speakers' production was characterized by significantly shorter VOTs compared to American English speakers' production of the same stops. Awan and Stine (2011) did not include stops following /s/ in their study. Given the paucity of detailed quantitative data, reliance on intuitive reports is potentially problematic since such reports of phonological patterns do not always survive closer scrutiny (Zhang and Lai 2008, Alderete and Kochetov 2009). Moreover, cross-language adaptation patterns might not be directly accessible to intuition, and intuitions themselves can be subject to bias (Kawahara 2011). In light of these findings, our goal here is to test the existing intuitive reports

(Kelkar 1957, Nair 1996, Coelho 1997, Sailaja 2009) and document the production of English stops in Indian English with robust acoustic data.

Present-day Indian English is a conventionalized variety of English spoken by English speakers in the Indian subcontinent (Hoffmann et al. 2011). In addition to providing a phonetic description of this widely spoken yet under-studied variety of English, our study also contributes to the understanding of the processes that were involved in the formation of Indian English phonetics, when learners of British English were presumably influenced by their native languages. Thus we hope that this study will stimulate future research on synchronic cross-language adaptation patterns where English aspirates adapt to native phonetics.

Labial	Dental	Retroflex	Palatal	Velar
p b	t̪ d̪	ʈ ɖ	c ɟ	k g
p^h b^h	t̪^{h} d̪^{h}	ʈ^h ɖ^h	c^h ɟ^h	k^h g^h

Table 1. Inventory of stop consonants in relevant Indic languages.

The Indic languages that we investigate all have the stop inventory shown in Table 1 (see section 2.1 for a list of references). Dravidian languages which also have associated Indian Englishes were excluded from this study. Some Dravidian languages like Tamil do not have a four-way contrast among stops. Some others which do, like Kannada and Malayalam, do not employ it across the lexicon or registers. In Kannada, for example, voiced aspirated stops are only found in Sanskritic words, and even those are neutralized to plain voiced stops in most registers (Upadhyaya 1972). To avoid including the effect of phonetic grammars that do not encode the four-way contrast in a similar manner, we restrict our study to Indic languages only. This is not meant to serve as a claim about which languages have historically contributed to the development of present-day Indian English(es), but as a methodological simplification. Throughout this paper, we use the term "Indian Englishes" to refer to the varieties of English spoken by speakers who are bilingual in English and an Indic language.

Our study focuses on labial and dorsal stops since these two places of articulation are shared between English and Indic languages. We did not include coronal stops in our study for several reasons. First, Indic languages have two kinds of coronal stops – dental and retroflex, and neither of these places of articulation matches the English alveolar stops. Although the short-lag stop pattern likely applies to coronal stops in Indian English, documenting this phenomenon will be hindered by potential additional adjustments in place. Second, as we discuss in section 2.1, the context where word-initial /t/ appears in our dataset is likely to introduce additional variability in its production, making it even harder to assess the cues to aspiration.

	k	k^h	p	p^h
Bengali (Mikuteit and Reetz 2007)	34	73	14	-
Gujarati (Rami et al. 1999)	41	75	-	-
Hindi (Lisker and Abramson 1964)	16	84	12	63
Marathi (Lisker and Abramson 1964)	21	73	0	35
Nepali (Poon and Mateer 1985)	27	84	6	69
Pahari (Khan and Bukhari 2011)	27	77	19	66
Urdu (Hussain 2018)	35	85	18	73
Indic (average)	29	79	12	61
British English (average)	25	62	11	44
British English (Docherty 1992)	27	62	15	42
British English (Whiteside and Irving 1997)	22	62	7	45

Table 2. A comparison of the existing results on VOT (ms) in Indic languages and in British English.

Table 2 gives an overview of the VOT of voiceless aspirated and unaspirated dorsal and labial stops for Indic languages and for British English reported in the literature. Aside from British English, this table includes all languages in our dataset (see section 2.1) for which we could find data.[1]

In examining Table 2, a few important caveats should be kept in mind. First of all, the VOT values averaged across Indic languages may only be used to get a very coarse-grained view of laryngeal phonetics since each of those languages may differ from the others considerably. Second, cross-linguistic phonetic comparisons are hindered by the fact that the studies in question differ in methodology, and hence the values listed here are not obtained under comparable conditions. That said, the data in Table 2 support the idea that Indic languages have longer VOT in voiceless aspirated stops than British English, while showing similar VOT in voiceless unaspirated stops. A more direct comparison of American English with Hindi and Marathi is presented by Lisker and Abramson (1964), where the same method is used for each language. Although Lisker and Abramson have only one speaker of each Hindi and Marathi, at least for the former their results are likely representative since two other studies of Hindi yielded much higher VOT values for voiceless aspirated stops (Benguerel and Bhatia 1980, Shimizu 1989). Table 3 lists the VOT values pertaining to utterance-initial voiceless aspirated stops in Lisker and Abramson (1964). Here again, aspirated stops have a longer VOT in Indic languages than in English, usually with a substantial difference.

	/p^h/	/k^h/
Hindi	63	84
Marathi	35	73
English	28	43

Table 3. Summary of the average VOT (ms) for Hindi, Marathi, and English reported by Lisker and Abramson (1964).

The fact that Indic languages have a larger stop inventory may also suggest that aspirated stops will exhibit less variability in VOT than their English counterparts. However a recent detailed study of Hindi by Hauser (2019) did not find a distinction in VOT variability between these two languages.

Since Indic languages have a richer set of laryngeal contrasts than English, perceptually these contrasts may be supported by multiple kinds of cues. First, there is evidence that all four categories of Indic stops cannot be distinguished based on VOT alone, and the literature proposes a number of other durational cues to laryngeal contrasts (Mikuteit and Reetz 2007, Rami et al. 1999, Davis 1994, Dutta 2007, Berkson 2012, Dmitrieva and Dutta 2020, Clements and Khatiwada 2007, Schwarz et al. 2019). In this article, one of the cues we investigate is VOT. While this cue is the best studied cue to voicing cross-linguistically, it should be kept in mind that the actual durational cues employed by Indic speakers may be more complex.

The second cue we investigate is onset f0 following the stops. f0, which is a secondary cue to voicing and/or aspiration is some languages, may be pointing in different directions in Indic languages versus in British English. Tonal effects of aspiration are less studied than VOT, so we can only draw some preliminary conclusions from the literature here. Compared to voiced unaspirated stops in English, voiceless aspirates are reported to be associated with rising f0 in the following vowel (House and Fairbanks 1953, Lehiste and Peterson 1961, Hombert 1976, Ohde 1984, Kingston and Diehl 1994). However, in these studies, aspiration co-varies with voicing, so the conclusions we can draw are only preliminary. Studies of Indic languages comparing effects on f0 of aspirated versus unaspirated voiceless stops report a relatively small local f0 *lowering* with aspiration. This is found for f0 of vowels following /k/ versus /k^h/ in Marathi (Dmitrieva and Dutta 2020), Hindi/Urdu (Dutta 2007, Schertz and Khan 2020), and Nepali (Clements and Khatiwada 2007). Thus, although tonal effects of aspiration are understudied, the existing literature suggests that aspiration in English voiceless stops may tend to raise f0 of the following vowel, while in Indic languages it may have a lowering effect.

To summarize, existing results on the acoustics of Indic voiceless stops suggest two important differences compared to English. In terms of VOT, Indic aspirated stops are more aspirated than English aspirated stops, while Indic unaspirated stops, like their English counterparts, are short-lag

[1] Note that Mikuteit and Reetz (2007) did not measure the VOT of [p^h] since their speakers consistently produced these as [f].

and have comparable VOTs. English aspirated stops are also reported to correspond to higher f0 values in the following vowel, while Indic aspirated stops correspond to lower f0 values. Assuming this current situation is reflective of the state of British English and Indic languages during the first substantial stages of language contact, we may hypothesize that when British English aspirated stops were perceived by native Indic speakers, they may not have been 'aspirated enough' in terms of VOT, and may have had f0 effects inconsistent with native effects in Indic languages. For these reasons, English aspirated stops may have been categorized with native unaspirated stops, as models of non-native speech perception would predict (Best 1995, Best and Tyler 2007, Flege 1995). If this is correct, we expect that word-initial voiceless stops in Indian English have shorter VOT values compared to other varieties of English, and they may lack lowering f0 effects on the following vowel. The aim of our study is to test if these acoustic predictions hold of actual Indian English productions, and hence to document the acoustics of voiceless stops in a corpus of Indian English. The paper is organized as follows. Section 2 presents our methods, followed by results (section 3), discussion (section 4), and concluding remarks (section 5).

2 Method

2.1 Materials and participants

In this study, we analyzed audio recordings of English speech from the Speech Language Archive (Weinberger 2014) – a web-based corpus containing recordings of the same English passage read by multiple speakers of different English varieties, both native and non-native. From the archive (accessed on October 10, 2019), we selected all recorded speech samples from adult speakers of Indian English who were also native speakers of one of nine Indic languages listed in Table 4.[2] These languages were selected because they all have the same inventory of stop consonants, as shown in Table 1.

Twenty-eight adult monolingual speakers of British English (BE), sixteen males and twelve females, functioned as the comparison group. These speakers were all speakers of the same variety, southern British English. The speakers selected were English speakers on the archive who were born and resided in England. Sex, present residence, and birthplace were self-reported by the speakers.

The archive uses self-reported native languages, and hence distinguishes Hindi and Urdu. Following this, we treat Hindi and Urdu as separate languages in our analysis, while not necessarily supporting the claim of these being two different languages. Although multilingualism is common in India, we unfortunately do not have enough information on other languages spoken by the subjects. However, speech samples produced by speakers who spent their childhood and/or adolescence (anytime under the age of eighteen years) in a primarily English-speaking country other than India were excluded.

Language	Number of speakers	Grammatical description
Bengali	14	Khan (2010)
Gujarati	11	Mistry (1997)
Hindi	28	Ohala and Ohala (1972)
Konkani	3	Almeida (2003)
Marathi	8	Dhongde and Wali (2009)
Nepali	14	Acharya (1991)
Oriya	2	Majumadāra (1970)
Pahari	2	Jouanne (2014)
Urdu	13	Ohala and Ohala (1972)
Total	102	

Table 4. Number of speakers and the referenced grammatical description for each language analyzed.

[2] Other major Indic languages like Kashmiri, Punjabi and Sinhala were not included since these languages do not have the four-way laryngeal contrast.

Each participant read the same passage, and the recordings are therefore comparable to each other, although speech rate and recording conditions are not controlled for. The passage read by each participant is given in (1). The Speech Accent Archive also provides transcriptions for some of the recordings, but our study did not rely on these transcriptions.

(1) Passage read by participants on the Speech Accent Archive
"Please call Stella. Ask her to bring these things with her from the store: Six **spoons** of fresh snow **peas**, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the **kids**. She can **scoop** these things into three red bags, and we will go meet her Wednesday at the train station."

The words analyzed in this study include *kids* and *peas* (word-initial stops), and *spoons* and *scoop* (post-/s/ stops). These words are highlighted in bold in (1). These words were chosen since they are all content words occurring in prominent contexts and were least likely to be reduced. All the target words are nouns, except for the verb *scoop*, which is the only instance of unaspirated [k] before a stressed vowel.

We decided not to include the word-initial /t/ of the word *toy* (comparing it, for example, to /t/ in *store*) for two reasons. First, as discussed in section 1, Indic languages do not have alveolar stops, and hence Indian English productions of /t/ might differ not only in aspiration, but also of place, creating a confound. Second, a preliminary examination of a subset of the recordings revealed that the word *toy* occurred in a garden path context and elicited unnatural productions from a number of speakers. In other words, a number of speakers appeared to first parse *toy* as head of its NP, then noticed that in fact *frog* was the head, and *toy* the specifier. This often resulted in elongated and hesitant pronunciations of *toy* and a pause between *toy* and *frog*. These mispronunciations would create high variability in our phonetic data, and make the results for coronals less comparable to other places of articulation. For these reasons, coronal stops were excluded from consideration.[3]

2.2 Annotations and measurements

The words selected for this study allowed us to compare stops that are aspirated in British English (*peas*, *kids*) to unaspirated stops (*spoons*, *scoop*). Each speaker contributed one token of the words in each category in our study. In six cases, a speaker did not produce a given word or misspoke (for example, produced *swoop* instead of *scoop*). Such tokens were excluded.

Our statistical analyses involved calculating the VOT and f0 difference between the initial stops and stops after /s/. For those calculations, we had to also exclude all tokens that didn't have pairs. For example, if a speaker produced *swoop* instead of *scoop*, we excluded the token *swoop*, and also did not analyze the word *kids* from that speaker since there was no #s_ token to compare it to. We followed a similar procedure for words that had to be excluded from f0 measurements because of creaky voice.

	BE	IE-Ben	IE-Guj	IE-Hin	IE-Kon	IE-Mar	IE-Nep	IE-Ori	IE-Pah	IE-Urd	Total
Dor	56	26/24	22	52/50	5	16/14	22	4	4	26/24	176
Lab	56	28	22	52	6	18/16	24	4	4	26	184

Table 5. Number of tokens analyzed for VOT (first number) and f0 (second number, if different), excluding speaker errors and instances of creaky voice.

Table 5 shows the total number of tokens analyzed for each context (recall that each speaker produced just two tokens for each place of articulation). Here and in what follows, we abbreviate the language names: BE – British English; Ben – Bengali; Guj – Gujarati; Hin – Hindi; Kon – Konkani; Mar – Marathi; Nep – Nepali; Ori – Oriya; Pah – Pahari; Urd – Urdu.[4] Note that for the

[3] Our preliminary acoustic measurements based on a subset of Indian English data, nonetheless, confirm that the VOT of /t/ in *toy* is, like /p/ and /k/, consistently short-lag.

[4] Recall that first language is reported by the speakers themselves and hence Hindi and Urdu are considered different languages.

Indic languages, the language being analyzed is the variety of English associated with a particular Indic language. Whenever the token counts for VOT and f0 are different, the count for f0 is given as a second number in Table 5.

Acoustic measurements were done in Praat (Boersma and Weenink 2019). VOT was measured manually, as the interval between the beginning of the release burst and the onset of quasi-periodicity.

Since f0 is one of the cues to laryngeal distinctions, fundamental frequency was measured for the vowel immediately following the target stops. f0 measurements were taken using a modified version of a Praat script by Katherine Crosswhite[5] starting at the first point immediately at the onset of voicing at which periodicity could be detected, and then at four additional equally spaced points within the first half of vowel duration. This procedure was adopted for two reasons. First, f0 depression has been found to be local; that is f0 following voiced and aspirated stops is lowered immediately following the stop with lowering effects quickly tapering off beyond this point. Second, measuring f0 at equidistant points in the vowel meant differences in the vowel lengths produced by different speakers in different contexts did not affect the f0 measurement.

Self-reported speaker sex from the Speech Accent Archive was used to set the measurement parameters in the script. In a few cases, f0 could not be measured at a given point because of creaky voice – these cases were inspected manually to confirm the instance of aperiodicity before excluding these individual measurements. Automatic f0 measurements were corrected manually to fix errors which typically result from pitch-halving and pitch-doubling errors of Praat's autocorrelation algorithm (Boersma 1993). f0 outliers were identified by visually inspecting the f0 distribution at each measurement point. For the outlier tokens, pitch was measured manually by taking the duration of a single glottal pulse as the duration of one cycle of the periodic waveform and taking its inverse (see also Dmitrieva and Dutta 2020).

2.3 Summary

To summarize, this study investigated the VOT and f0 of voiceless stops in a corpus of Indian English, that is, the English of speakers whose native languages are Indic. The measurements for Indian English speakers were compared to those from speakers of British English. We expected that British English speakers would produce robustly different VOTs in voiceless stops following /s/ vs. word-initially. On the other hand, we did not expect a robust difference in VOT for Indian English speakers. For VOT, the expectation thus involved the distinction between initial voiceless stops vs. stops after initial /s/. To address these distinctions, we subtracted the VOT of stops after /s/ (in spoons, scoop) from the VOT of initial stops (in peas, kids) for each place of articulation and for each speaker. The resulting VOT difference was our dependent variable in what follows.

Our expectations regarding f0 were complicated by the fact that fundamental frequency may vary with intonational context as well as with laryngeal setting of the preceding consonant. Intonational context of each word is not fully controlled for in our corpus. The target words with initial stops are phrase-final (*peas*, *kids*), but the others (*spoons*, *scoop*) occur in phrase-medial position. Comparing the words with a word-initial stop to words with a stop after /s/, we therefore expected to see both intonational differences and, potentially, differences based on aspiration. The latter effect may be highly local and small in size (cf. Dmitrieva and Dutta 2020, for Marathi) or masked entirely by the intonational effects.

From the point of view of intonation, we may expect phrase-final words to have lower pitch than phrase-medial ones across the board. As discussed in section 1, the effects of aspiration on f0 of the following vowel are often quite subtle but based on the studies of British English we may preliminarily expect that aspirates will have a local raising effect on f0 of the following vowel in this language. This expectation is only tentative since most studies to date compare English voiced unaspirated stops to voiceless aspirated ones, and hence the raising effect on f0 may follow from voicing, not just aspiration (House and Fairbanks 1953, Lehiste and Peterson 1961, Hombert 1976, Ohde 1984, Kingston and Diehl 1994). On the other hand, for Indian English our expectation was that the stops would not differ from each other in aspiration, and therefore the two categories of stops would not show a difference in effects on the f0 of the following vowel.

---

[5] The script is available at: http://phonetics.linguistics.ucla.edu/facilities/acoustic/formant_logging.txt

In sum, we expected to find intonational differences based on context, and these differences were expected to be similar across languages. By contrast, due to effects of aspiration, f0 in the initial portion of the vowel was expected to be raised after aspirates in British English but not in Indian English. Like VOT, this prediction is best addressed by calculating f0 difference between vowels after initial stops vs. stops after /s/. This f0 difference was our dependent variable for statistical modeling.

3 Results

3.1 VOT

Table 6 shows the mean VOT we obtained for British English compared to findings from two previous studies on British English. Given that these other studies included fewer subjects and employed different methodologies, the mean VOT values we obtained for British English are a decent match with them, serving as a sanity check for our measurements and methodology.

Phone	Present study (n=28)	Docherty (1992) (n=5)	Whiteside and Irving (1997) (n=10)
k	25	27	22
k^h	74	62	62
p	14	15	7
p^h	58	42	45

Table 6. Comparison of mean British English VOT (in ms) with previous studies.

Table 7 shows the mean VOT we obtained for Indian English, and compares these values to previous results on VOT in Indic languages. The only comparable study of Indian English itself (Awan and Stine 2011) reports a mean VOT of 33ms for [p^h] – a result that matches our data.[6]

	k	k^h	p	p^h
Bengali (Mikuteit and Reetz 2007)	34	73	14	-
Gujarati (Rami et al. 1999)	41	75	-	-
Hindi (Lisker and Abramson 1964)	16	84	12	63
Marathi (Lisker and Abramson 1964)	21	73	0	35
Nepali (Poon and Mateer 1985)	27	84	6	69
Pahari (Khan and Bukhari 2011)	27	77	19	66
Urdu (Hussain 2018)	35	85	18	73
Indic (average)	29	79	12	61
	post-/s/ k	word-initial k	post-/s/ p	word-initial p
Indian English (present study)	47	47	18	28
British English (present study)	25	74	14	58

Table 7. A comparison of the existing results on VOT (ms) in Indic languages and in British English.

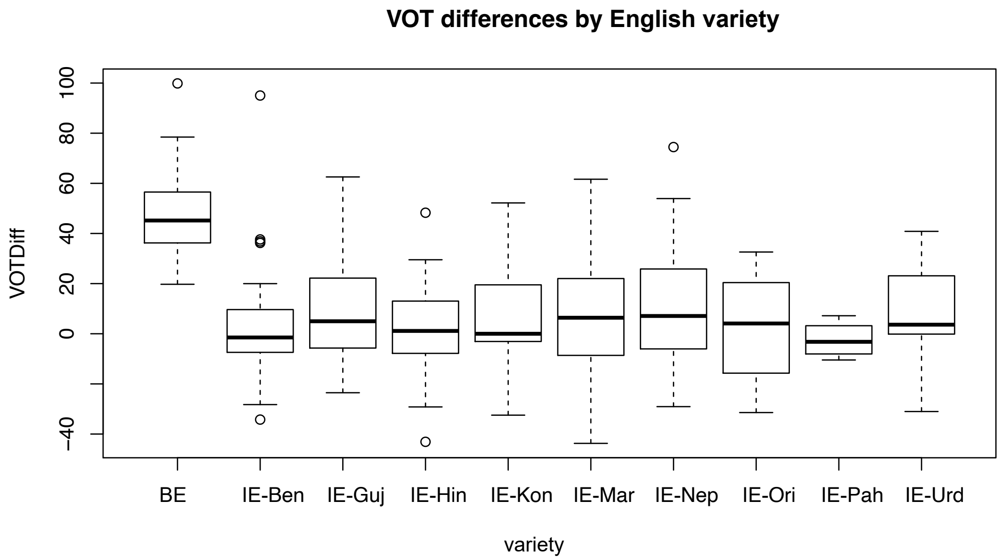
The VOT values for Indian English word-initial stops were close to the VOT of unaspirated stops for a given place of articulation in the Indic languages. Thus, Table 7 shows that when speaking Indian English, Indic speakers produce stop VOTs that more closely approximate Indic VOTs of the unaspirated category.

Furthermore, the differences between initial and post-/s/ stops in Indian English were very small. Figure 1 plots these differences in VOT for each variety of Indian English associated with the corresponding Indic language. It can be observed that all Indian Englishes have a difference between

[6] Note that Awan and Stine's (2011) study included speakers of a slightly different set of languages – their set included Punjabi, Kannada, and Telugu but did not include Oriya and Pahari.

initial and post-/s/ stops that is close to zero, while the difference in British English is quite substantial (mean = 46.7ms.; s.d. = 15.3ms).

Figure 1. VOT differences between initial stops and stops after /s/ by English variety. See section 2.2 for IE variety abbreviations.



We analyzed the VOT differences with a linear mixed effects regression model using consonant place (dorsal, labial), and speaker language (as self-reported in our dataset) as fixed effects and the speaker as a random effect. Recall that each speaker produced only one token of each word, so we could not separate the effect of speakers from the effect of individual tokens by adding item as a random effect. Our first model, Model 1, summarized in Table 8, included all the Indian Englishes as well as British English. Taking British English VOT differences as baseline, Model 1 found a significant effect for each Indian English variety.

	β	SE	df	t-value	*p*(>|*t*|)
LanguageBen	–43.3	6.9	203.7	–6.2	<0.001
LanguageGuj	–43.9	7.4	202.5	–6	<0.001
LanguageHin	–51.4	6	205.6	–8.5	<0.001
LanguageKon	–69.1	15.1	208.1	–4.6	<0.001
LanguageMar	–46.7	8.3	204.6	–5.6	<0.001
LanguageNep	–40.4	7.4	204	–5.5	<0.001
LanguageOri	–60.8	15.2	202.5	–4	<0.001
LanguagePah	–52.4	15.2	202.5	–3.5	<0.001
LanguageUrd	–40.9	7.1	203.8	–5.7	<0.001
placeLab	–4.997	5	99.5	–1	n.s.
LanguageBen:placeLab	0.3	8.8	102	0.03	n.s.
LanguageGuj:placeLab	13.3	9.4	99.5	1.4	n.s.
LanguageHin:placeLab	13.3	7.5	106.9	1.8	<0.1
LanguageKon:placeLab	48.8	18.1	119	2.7	<0.01
LanguageMar:placeLab	13	10.4	104.1	1.2	n.s.
LanguageNep:placeLab	13.4	9.3	102.7	1.4	n.s.
LanguageOri:placeLab	32.9	19.4	99.5	1.7	<0.1
LanguagePah:placeLab	6.6	19.4	99.5	0.3	n.s.
LanguageUrd:placeLab	4.4	9	102.3	0.5	n.s.

Table 8. Model 1 results for VOT differences among British English and Indic languages.

In addition, taking dorsal place as baseline, Model 1 found a significant interaction between labial place and the language being Konkani. These interactions were marginal for Hindi and Oriya. We only have two speakers for Oriya and three for Konkani, so the effects for these languages still need to be tested with a larger dataset. A marginal interaction between language being Hindi and labial place was somewhat unexpected, but note that VOT differences in Hindi were still very small (mean –3.4ms for dorsals; 6.1ms for labials).

Although Model 1 supported the difference between British and Indian English as two groups, the potential differences between individual Indian Englishes were harder to assess with this model. Therefore, we ran a second model which included the same factors but addressed only data from Indian English speakers, excluding British English speakers. Model 2, in which Bengali was the baseline, is summarized in Table 9. Unlike Model 1, this model found no significant main effect of language. In fact, in Model 2 only the interaction of language being Konkani with labial place turned out to be significant. Thus, we did not find robust differences in VOT between the different Indian Englishes based on speaker language or consonant place. The significant interaction for Konkani should be interpreted with caution since we only have three speakers for this language.

	β	SE	df	t-value	$p(>\|t\|)$
LanguageGuj	–0.53	9.1	150.3	–0.1	n.s.
LanguageHin	–8	7.9	152.2	–1	n.s.
LanguageKon	–26	16.8	154.2	–1.5	n.s.
LanguageMar	–3.3	10	151.5	–0.3	n.s.
LanguageNep	2.9	9.1	151.2	0.3	n.s.
LanguageOri	–17.5	16.9	149.8	–1	n.s.
LanguagePah	–9	16.9	149.8	–0.5	n.s.
LanguageUrd	2.4	8.9	151.1	0.3	n.s.
placeLab	–4.7	7.7	76.1	–0.6	n.s.
LanguageGuj:placeLab	13	11.5	74.6	1.1	n.s.
LanguageHin:placeLab	13	9.8	78.8	1.3	n.s.
LanguageKon:placeLab	48.7	20.1	86.9	2.4	<0.05
LanguageMar:placeLab	12.7	12.5	77.1	1	n.s.
LanguageNep:placeLab	13.1	11.4	76.4	1.1	n.s.
LanguageOri:placeLab	32.6	21.4	73.7	1.5	n.s.
LanguagePah:placeLab	6.2	21.4	73.7	0.3	n.s.
LanguageUrd:placeLab	4.1	11.2	76.2	0.3	n.s.

Table 9. Model 2 results for VOT differences among Indic languages.

To check if the model with only Indian Englishes would have found a significant difference with a different baseline language, we ran a model comparison of Model 2 with one that omitted that effect of language. This model comparison is shown in Table 10. This comparison shows that including the effect of language does not improve the model fit. That is, the model with language as a main effect is not significantly different from one in which the effect of language is not included. This indicates that VOT difference is not affected by the linguistic background of the Indian English speakers.

df2	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
model2	4	1564.75	1577.37	–778.38	1556.75	*NA*	*NA*	*NA*
baseline	20	1582.93	1645.99	–771.56	1542.93	13.83	16	0.61

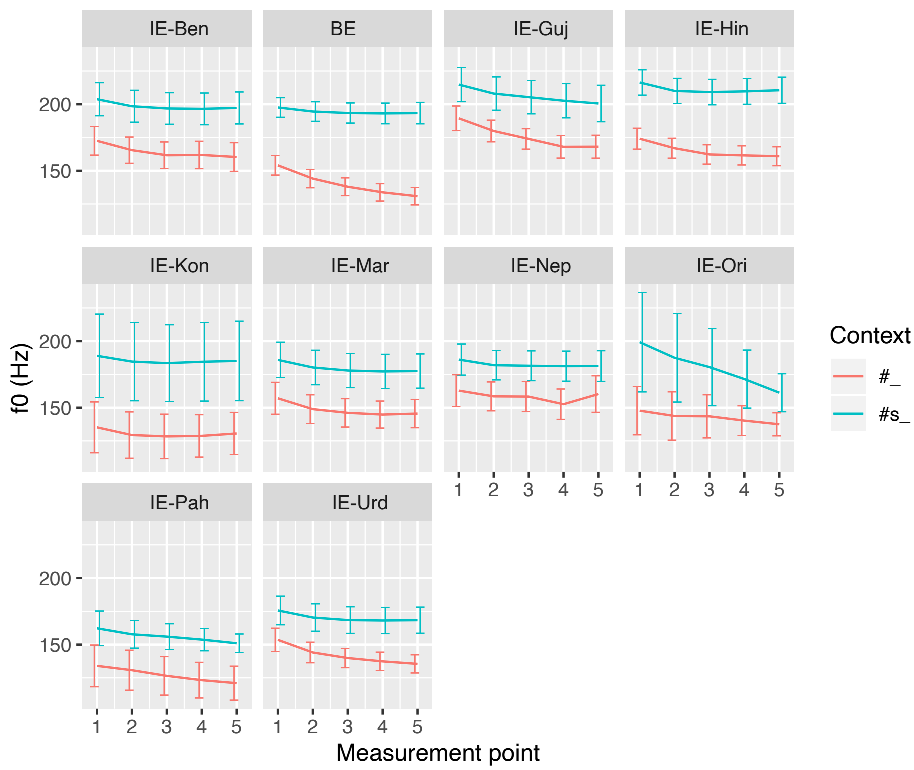
Table 10. Model comparison results for VOT differences among Indic languages.

To summarize, the VOT distinction between aspirated and plain stops was sizeable in our British English data but close to zero in Indian English data (see Figure 1). British English was also significantly different from Indian English as a whole, but the Indian Englishes did not differ from each other in terms of stop VOT distinctions. Our complete dataset also exhibited an interaction of place with the language being Konkani.

3.2 Fundamental frequency

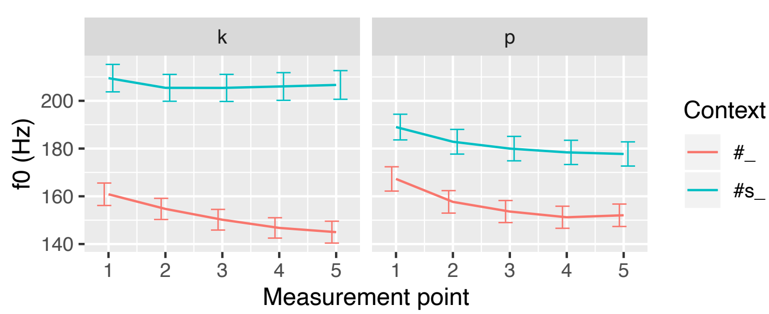
Figure 2 summarizes mean f0 within the first half of vowel duration by language. Here British English did not appear to differ from Indian English in how vowel f0 patterns with aspiration. For all languages, the overall mean f0 after word-initial stops (aspirated in British English) was lower than that after post-/s/ stops (unaspirated in British English), and the size of the difference was comparable across languages. Relatively large standard error (seen in error bars) for Oriya, Konkani, and Pahari follows from a small size of our corpus for these languages.

Figure 2. Mean f0 values for each measurement point by language. Error bars show standard error. See section 2.2 for language name abbreviations.



The variation in vowel f0 by place of the preceding consonant is examined in Figure 3. Here, the difference between f0 after the two kinds of stops seems larger after /k/ than after /p/. f0 in the words scoop and spoons is also higher than in the words kids and peas respectively. It is likely that this is due to the effect of prosodic context. The words kids and peas are phrase-final in the passage we analyzed, and therefore it is expected that these words may have a lower pitch than the phrase-medial scoop and spoons (Pierrehumbert 1980, Beckman and Pierrehumbert 1986).

Figure 3. Mean f0 values for each measurement point. Error bars show standard error.



The greater f0 difference between vowels following /k/ on the one hand and /p/ on the other, however, was surprising. It is possible that small differences in the prosodic properties of the contexts in which the words occurred led to this difference. The word *kids* was almost always produced at the end of an intonational phrase (IP) while *peas* was variably produced at the end of an intonational or intermediate phrase (ip). There were many instances where *peas* was produced with a rising ip boundary tone typically associated with listed items. Moreover, the word *scoop* was produced with a pitch accent in most cases, but the word *spoons* did not always receive a pitch accent. These are post-hoc speculations and need to be confirmed with quantitative analyses of our data. The conclusion we want to draw at this point is that, predictably, intonational effects had a notable effect on f0 differences.

Additionally, as expected, intonational effects had an influence on the entire first half of the vowel, rather than a localized effect at the onset of the vowel. Thus, an examination of the data suggests that the intonational context of our items, perhaps unsurprisingly, affected vowel f0. On the other hand, there was no clear indication of an effect of context on f0 – the patterning of f0 by context did not seem to differ substantially by language, although for some varieties of Indian English the difference between vowel f0 after initial and post-/s/ stops appeared to be smaller than in English. It may also be the case that local f0 effects based on context were masked by larger-scale intonational effects. However, note that since we did not find any differences in aspiration by context (word-initial versus post-/s/), comparable f0 following these stops was expected in the case of Indian English.

Recall that we expected a potential local raising of f0 with aspiration in British English. This, however, was not borne out. Figure 2 shows that vowel f0 after initial stops was consistently lower than vowel f0 after post-/s/ stops. We also predicted a smaller difference in f0 by aspiration in Indian English than in British. In order to assess this latter prediction, we ran a linear mixed effects regression model (Model 3) analyzing the f0 differences by aspiration. The dependent variable for this model is the difference in f0 between vowels after initial stops versus vowels after post-/s/ stops. This difference is negative for all languages analyzed, as seen in Figure 2. Our model targets f0 differences at the very first measurement point since previous research shows that aspiration effects on f0 are local, appearing to be the largest early on and quickly fading away (Dmitrieva and Dutta 2020). The model included speaker native language (with British English as baseline) and place (with /k/ as baseline) as main effects and individual speakers as a random effect. Model 3 results are presented in Table 11.

	β	SE	df	t-value	$p(>\|t\|)$
LanguageBen	23.5	12.8	204.3	1.8	<0.1
LanguageGuj	30.8	13.2	200.7	2.3	<0.05
LanguageHin	–1.8	10.2	202.2	–0.2	n.s.
LanguageKon	16.4	27	209.5	0.6	n.s.
LanguageMar	15.3	15.7	204.3	1	n.s.
LanguageNep	24.9	13.2	203	1.9	<0.1
LanguageOri	–10.4	27.2	200.7	–0.4	n.s.
LanguagePah	27.1	27.2	200.7	1	n.s.
LanguageUrd	24.6	12.8	202.7	1.9	<0.1
placeLab	32.1	8.7	103.8	3.7	<0.001
LanguageBen:placeLab	–18.1	15.6	109	–1.2	n.s.
LanguageGuj:placeLab	–25.4	16.4	103.8	–1.6	n.s.
LanguageHin:placeLab	7.9	12.7	106.9	0.6	n.s.
LanguageKon:placeLab	–16.3	31.6	122.7	–0.5	n.s.
LanguageMar:placeLab	–9.8	19.1	109	–0.5	n.s.
LanguageNep:placeLab	–7.4	16.2	106.8	–0.5	n.s.
LanguageOri:placeLab	4.9	33.7	103.8	0.1	n.s.
LanguagePah:placeLab	–23.6	33.7	103.8	–0.7	n.s.
LanguageUrd:placeLab	–7.8	15.7	106.5	–0.5	n.s.

Table 11. Model 3 results for differences in vowel f0 at the first measuring point.

Model 3 found a significant effect of labial place, as expected based on Figure 3 since the f0 difference for /p/ was smaller than for /k/. A significant main effect of language was found for Gujarati; Bengali, Nepali, and Urdu are marginal. These effects seem to go in the predicted direction since the f0 difference was smaller (closer to zero) for these languages than for British English. However, it seems premature to conclude that a smaller f0 difference in some Indian Englishes is due to the implementation of aspiration because f0 differences are affected by a number of additional factors in our dataset, and because there is no consistent effect across the Indian Englishes.

4 Discussion

We compared the VOT of word-initial and post-/s/ stops in Indian English, as well as the influence of these stops on f0 of the following vowel. Both of these acoustic variables may be affected by aspiration, and in addition f0 is affected by a number of contextual factors. VOT is a primary cue to aspiration, and it was found that British English exhibited a substantial difference in VOT between aspirated (word-initial) and unaspirated (post-/s/) stops while all Indian Englishes seem to have a very small VOT difference (Figure 1). In line with this, Model 1 (Table 8) that included British English as baseline and compared it to Indian Englishes found a significant effect of language while Model 2 (Table 9) that included only Indian varieties of English found no main effect of language. These findings are in line with our initial expectation that word-initial stops are unaspirated in Indian English. In this way, our study serves to document the implementation of Indian English stops with corpus data.

In our dataset, the size of aspiration difference by place also seems to vary, and in particular Model 1 (Table 8) finds a significant place x language interaction for Konkani and a marginal one for Hindi and Oriya. We believe that more data would be needed to properly interpret these results, since especially for Konkani and Oriya our corpus has limited data.

f0 of the vowel following stops was found to be lower for word-initial stops than for post-/s/ stops for all varieties of English in our study (see Figure 2). This was particularly unexpected for British English where previous studies lead one to expect a potential raising effect of aspiration on f0. We believe our findings are likely explained by contextual effects, and this is indirectly confirmed by the fact that f0 in the words *kids* and *peas* is lower than for *scoop* and *spoons* at all measurement points, not just at the onset of the vowel. The words *kids* and *peas* are phrase-final in the passage we analyzed, and they may have a lower pitch than the phrase-medial *scoop* and *spoons* (Pierrehumbert 1980, Beckman and Pierrehumbert 1986).

In sum, the direction of f0 difference between vowels after aspirates and vowels after unaspirated stops in British English was not consistent with the expected aspiration effect. We suspect that the contextual influences described above may be masking the more subtle effect of aspiration. More generally, the f0 difference between initial stops and post-/s/ stops varies with place of articulation of the stop and, to some extent, with the Indic language in our dataset. Figure 3 suggests a larger difference for vowels after /k/ than for vowels after /p/, and this difference corresponds to a significant effect of place on the f0 differences at the first measurement point in Model 3 (Table 11). We do not have an immediate explanation for this place effect, and leave this as a topic of investigation for future research.

As for the other language spoken by the participants, several Indian Englishes show a smaller f0 difference than British English. Model 3 (Table 11) found a significant effect of language for Gujarati, and a marginal one for Bengali, Nepali, and Urdu. A smaller difference in vowel f0 would be expected based on the fact that Indian English stops are unaspirated, but given the overall patterning of our f0 data we do not think these effects support this conclusion.

To summarize, our analysis of vowel f0 did not find support for the expected differences. We take this null result to suggest that aspiration effects on f0 are very subtle, and hence likely undiscoverable in a corpus study with lots of additional variables like ours. The effect of aspiration on f0 needs to be reevaluated in a study that controls for intonational context and vowel quality.

In terms of VOT, present-day Indian English exhibits an extremely interesting pattern where both categories of English stops were historically adapted to the native unaspirated category, despite the fact that Indic languages also have aspirates. Existing impressionistic reports also suggest that this pattern extends to loan adaptation (Paradis and LaCharité 1997). In parallel to this pattern, the native VOT of stops in Indic languages differs substantially from that of most varieties of English:

Indic aspirated stops have a much longer VOT than English aspirated stops, while Indic unaspirated stops have VOT values comparable to those in English (see section 2.3). While the VOT of aspirated English stops does not perfectly match those of unaspirated Indic stops, it seems plausible that English aspirated stops could have historically been perceived as unaspirated by a native Indic listener. That is, such a listener might have found English stops to not be ‘aspirated enough’ by Indic acoustic benchmarks. In other words, it is possible that the category boundary for aspirated versus unaspirated stops in Indic was, at least historically, higher than that in English.

Sailaja (2009) suggests that the lack of aspiration in Indian English voiceless stops may also be conditioned in part by the lack of aspiration marking in English orthography. While this explanation is plausible, it is unlikely to be the only factor involved. There is evidence suggesting that perceptual assimilation plays a role in the early stages of second language acquisition (Best 1995, Flege 1995, Best and Tyler 2007) and loanword adaptation (Silverman 1992, Peperkamp 2005, Peperkamp et al. 2008, Boersma et al. 2009). While orthography has also been shown to be a contributing factor (Vendelin and Peperkamp 2006), it has been shown to be a factor when perception is indeterminate (Daland et al. 2015). Thus, it is possible that both our account and Sailaja’s are correct. The acoustic mismatch we observed in our study likely contributed to a perceptual mismatch which led to Indian English speakers referencing the orthography to resolve this resulting ambiguity.

It might be useful to consider the adaptation of English dental fricatives in Indian English to evaluate the effect orthography has on adaptation patterns. The English interdental fricatives /θ/ and /ð/ are adapted as dental stops /t̪/ or /t̪ʰ/ and /d̪/ respectively (Kelkar 1957, Sailaja 2009). Notice that both these phonemes are represented in the orthography with “th” (e.g., as in *think* and *this* respectively). While the adaptation of /θɪŋk/ as [t̪ʰɪŋk] can be explained by the orthography, the explanation for /ðɪs/ being adapted as [d̪ɪs] is potentially problematic for this account. An acoustic account, however, can explain both these adaptation patterns. In the adaptation of /θ/ as [t̪ʰ], aspiration was employed by Indic speakers to preserve the acoustic friction present in the English fricative. This is also observed in the adaptation of /f/ as [pʰ] in some Indian Englishes (Kelkar 1957). Additionally, the reason that /ð/ isn’t adapted as [d̪ʰ] likely has to do with conflicting f0 cues. Lowering of f0 is especially prominent following breathy stops like [d̪ʰ] (Dmitrieva and Dutta 2020, Schertz and Khan 2020), and could be a major contributing factor to the perceptual identification of breathy stops (Dmitrieva and Dutta 2020, Schertz and Khan 2020). This acoustic difference between Indic plain voiced and breathy voiced stops leads to the adaptation of /ðɪs/ as [d̪ɪs]. Thus, in this case, due the greater acoustic mismatch, there was no resulting perceptual ambiguity. Listeners, therefore, did not consult the orthography leading to the lack of aspiration in such words that are spelled with <h>.

This account, which assumes assimilation originating in acoustic similarity, can also explain why Indian English speakers who are native speakers of languages that do not have contrastive aspirates such as Ao, Angami, Mizo, and Tamil do produce aspiration in the appropriate contexts where it appears in “native” varieties of English (Wiltshire 2005, Wiltshire and Harnsberger 2006). The unpredictable adaptation of /θ/ as either aspirated or unaspirated is reflective of the competing acoustic parameters – f0, whose lowering is less pronounced after voiceless, aspirated stops, and the attempt to preserve friction. In summary, while we do not completely rule out the influence of orthography on Indian English, we do point out that specific phonetic differences could have contributed to the adaptation pattern seen in Indian English.

Thus, native Indic phonetics seems to have affected Indian English in a predictable way. As noted previously, this pattern also reportedly extends to loanword adaptation, and here the observed correlation between native phonetics and borrowed adaptations is particularly surprising from the point of view of the feature-based theories of loanword adaptation (Paradis and LaCharité 1997, LaCharité and Paradis 2005). If cross-language adaptation is based on features and contrast, we might expect that the native Indic contrasts would be mapped onto the English contrast system in loans, and in Indian English. Thus, a case where this apparently does not happen is a potential challenge to the feature-based theories. On the other hand, the perception-based approaches to loanwords and non-native phonetics (Silverman 1992, Yip 1993, Peperkamp et al. 2008) make predictions about the perceptual processing of English aspirates in Indic speakers, and these predictions can be tested more directly in future perceptual experiments.

Given our results, which are undeniably only half the puzzle, we propose that in the early stages of Indian English, listeners encountered an acoustic mismatch between their native Indic categories and those in the source language (some variety of British English).[7] Perceptual ambiguity resulted from the mismatch in VOT of word-initial stops in the source language and the VOT of aspirated stops in the listeners' native Indic languages. Since phonological parsing by means of perceptual factors alone was indeterminate, English orthography might have influenced the adaptation of voiceless stops in Indian English as proposed by Daland et al. (2015). Needless to say, this proposed perceptual adaptation pattern needs to be confirmed by a perceptual experiment. It must also be noted that since this pattern presumably occurred in the early stages of Indian English, it is possible that the perceptual category boundaries of the present-day Indian English speaker have shifted, and are, therefore, not a reliable metric to assess this proposed adaptation mechanism.

5 Conclusion

This study documents the acoustics of voiceless stops in Indian English in comparison to British English. Our main finding, in line with previous reports and findings, is that speakers of Indian English show a very small difference in VOT between initial stops and stops after /s/ while British speakers show quite a substantial difference for these stops in the same context. The absolute VOT values of Indian English stops approximate the values for unaspirated stops, and thus our corpus shows evidence of Indian English voiceless stops being always produced without aspiration as suggested by some impressionistic and other acoustic studies. On the other hand, we did not find any evidence to suggest that f0 effects of aspiration may be different in British vs. Indian English.

Our study also shows how naturalistic corpus data can be used to document subtle phonetic phenomena in cross-language adaptation and phonetics. Of course, the design of our study and our reliance on corpus data comes with certain limitations. We hope that this research will stimulate further studies addressing these issues. First, we acknowledge that we do not know if the adaptation pattern described here happens as part of active synchronic grammar of Indian English or exists as a historical pattern established and conventionalized for the existing words but not necessarily implemented grammatically. (Flege and Eefting 1987 make a similar point for Puerto Rican Spanish.) Indian English is a conventionalized variety, and many present speakers acquire it without being exposed to British or American English, and thus it may be that the adaptation of aspirated stops as unaspirated was present as a grammatical pressure only when Indian English was established, and is now carried over as a property of this variety. Future studies with non-words may clarify this.

Second, the acoustic data may not be sufficient to establish the correct featural representation of voiceless stops in Indian English. It could be, for example, that Indian English distinguishes two categories of stops based on the feature [voice] rather than on aspiration features, but in order to confirm this hypothesis we may need additional data on the acoustics of Indian English voiced stops, and potentially on the perception of this contrast.

Finally, although our data suggest that Indic laryngeal phonetics affects the implementation of VOT in Indian English, our study does not allow us to exclude other potential factors such as orthography. This may need to be investigated further, as it is a potential contributing factor along with native phonetics. Our study also highlights the complex interactions between linguistic and extra-linguistic factors in loanword adaptation and non-native perception.

Acknowledgements

We would like to thank two anonymous reviewers and the editor for very helpful comments on this paper. Many thanks to audiences at the 24th Annual Midphon Conference, the 37th West Coast Conference on Formal Linguistics, and (Formal) Approaches to South Asian Languages 10 where work related to this paper was presented. Thanks also to Ljiljana Progovac, Natalia Rakhlin, and Megha Sundara for helpful comments and suggestions. We also wish to thank the Linguistics Program at Wayne State University where most of this work was done.

[7] It might be appropriate to view this early stage through the predictions made by models of non-native speech perception, such as the Perceptual Assimilation Model (or PAM) (Best 1995), and L2 acquisition, such as the Speech Learning Model (Flege 1995) and PAM-L2 (Best and Tyler 2007). See Narkar (2021) for a discussion of the Indian English adaptation pattern in terms of PAM.

References

Acharya, Jayaraj. 1991. *A descriptive grammar of Nepali and an analyzed corpus.* Georgetown University Press.

Alderete, John and Alexei Kochetov. 2009. Japanese mimetic palatalization revisited: Implications for conflicting directionality. *Phonology* 26(3):369–388.

Almeida, Matthew. 2003. *A description of Konkani*. Ph.D. thesis, UMI Ann Arbor.

Awan, Shaheen N. and Carolyn L, Stine. 2011. Voice onset time in Indian English-accented speech. *Clinical linguistics & phonetics* 25(11-12):998–1003.

Beckman, Mary E. and Janet B. Pierrehumbert. 1986. Intonational structure in Japanese and English. Ph*onology* 3:255–309.

Benders, Titia, Sujal Pokharel, and Katherine Demuth. 2019. Hypo-articulation of the four-way voicing contrast in Nepali infant-directed speech. *Language Learning and Development* 15(3):232–254.

Benguerel, A.-P. and T. K. Bhatia. 1980. Hindi stop consonants: an acoustic and fiberscopic study. *Phonetica* 37:134–148.

Berkson, Kelly. 2012. Capturing breathy voice: Durational measures of oral stops in Marathi. *Kansas Working Papers in Linguistics* 33.

Best, Catherine T. 1995. A direct realist view of cross-language speech perception. *Speech perception and linguistic experience*:171–206.

Best, Catherine T. and Michael D. Tyler. 2007. Nonnative and second-language speech. *Language experience in second language speech learning: In honor of James Emil Flege* 17:13.

Boersma, Paul. 1993. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *Proceedings of the Institute of Phonetic Sciences* 17:97–110.

Boersma, Paul and Silke Hamann. 2009. Loanword adaptation as first-language phonological perception. *Loanword Phonology*: 11–58.

Boersma, Paul and David Weenink. 2019. Praat: Doing phonetics by computer. Software.

Clements, George N. and Rajesh Khatiwada. 2007. Phonetic realization of contrastively aspirated affricates in Nepali. In *Proceedings of the 16th International Congress of Phonetic Sciences*: 629–632.

Coelho, Gail M. 1997. Anglo-Indian English: A nativized variety of Indian English. *Language in Society* 26(4):561–589.

Daland, Robert, Mira Oh, and Syejeong Kim. 2015. When in doubt, read the instructions: Orthographic effects in loanword adaptation. *Lingua* 159:70–92.

Davis, Katharine. 1994. Stop voicing in Hindi. *Journal of Phonetics* 22(2):177–193.

Dhongde, Ramesh Vaman and Kashi Wali. 2009. *Marathi*, vol. 13. John Benjamins Publishing.

Dmitrieva, Olga and Indranil Dutta. 2020. Acoustic correlates of the four-way laryngeal contrast in Marathi. *Phonetica* 77:209–237.

Docherty, Gerard J. 1992. *The timing of voicing in British English obstruents*. New York: Foris.

Dutta, Indranil. 2007. *Four-way stop contrasts in Hindi: An acoustic study of voicing, fundamental frequency and spectral tilt*. PhD dissertation, University of Illinois at Urbana-Champaign Urbana, IL.

Flege, James E. 1995. Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research* 92:233–277.

Flege, James E. and Wieke Eefting. 1987. Production and perception of English stops by native Spanish speakers. Journal of Phonetics 15(1):67–83. Hauser, Ivy. 2019. Effects of phonological contrast on within-category phonetic variation. Doctoral Disser- tations. 1673.

Hoffmann, Sebastian, Marianne Hundt, and Joybrato Mukherjee. 2011. Indian English an emerging epicentre? A pilot study on light verbs in web-derived corpora of South Asian Englishes. *Anglia-Zeitschrift für englische Philologie* 129(3-4):258–280.

Hombert, Jean-Marie. 1976. The effect of aspiration on the fundamental frequency of the following vowel. In *Annual Meeting of the Berkeley Linguistics Society*, vol. 2: 212–219.

House, Arthur S. and Grant Fairbanks. 1953. The influence of consonant environment upon the secondary acoustical characteristics of vowels. *The Journal of the Acoustical Society of America* 25(1):105–113.

Hussain, Qandeel. 2018. A typological study of voice onset time (VOT) in Indo-Iranian languages. *Journal of Phonetics* 71:284–305.

Iverson, Gregory and Joseph Salmons. 1995. Aspiration and laryngeal representation in Germanic. *Phonology* 12(3):369–396.

Jouanne, Thomas Christian Nicolas Marie. 2014. *A preliminary analysis of the phonological system of the western Pahari language of Kvar.* Master's thesis, University of Oslo.

Kawahara, Shigeto. 2011. Japanese loanword devoicing revisited: A rating study. *Natural Language and Linguistic Theory* 29(3):705–723.

Kelkar, Ashok R. 1957. "Marathi English": a study in foreign accent. *Word* 13(2):268–282.

Khan, Abdul Qadir and Nadeem Haider Bukhari. 2011. An acoustic study of VOT in Pahari stops. *Kashmir Journal of Language Research* 14(1).

Khan, Sameer ud Dowla. 2010. Bengali (Bangladeshi Standard). *Journal of the International Phonetic Association* 40(2):221–225.
Kingston, John and Randy L. Diehl. 1994. Phonetic knowledge. *Language* 70(3):419–454.
LaCharité, Darlene and Carole Paradis. 2000. Phonological evidence for the bilingualism of borrowers. In *Proceedings of the 2000 Annual Conference of the Canadian Linguistics Association*: 221–232. Cahiers linguistiques d'Ottawa Ottawa, Ontario.
LaCharité, Darlene and Carole Paradis. 2005. Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic Inquiry* 36(2):223–258.
Lehiste, Ilse and Gordon E. Peterson. 1961. Some basic considerations in the analysis of intonation. *The Journal of the Acoustical Society of America* 33(4):419–425.
Lisker, Leigh and Arthur Abramson. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20:384–422.
Majumadāra, Pareśacandra. 1970. *A Historical Phonology of Oṛiyā*, vol. 66. Calcutta: Sanskrit College.
Mikuteit, Simone and Henning Reetz. 2007. Caught in the ACT: The timing of aspiration and voicing in East Bengali. *Language and Speech* 50(2):247–277.
Mistry, Purushottam J. 1997. Gujarati phonology. *Phonologies of Asia and Africa* 2:653–673.
Nair, N. Gopalakrishnan. 1996. *Indian English phonology: A case study of Malayalee English*. Prestige Books.
Narkar, Jahnavi. 2021. The curious absence of aspiration in Indian English: The role of phonetics in adaptation. In *37th West Coast Conference on Formal Linguistics*: 189–196. Cascadilla Proceedings Project.
Ohala, Manjari and John Ohala. 1972. The problem of aspiration in Hindi phonetics. *Project on Linguistic Analysis Reports*: 63–70.
Ohde, Ralph N. 1984. Fundamental frequency as an acoustic correlate of stop consonant voicing. *The Journal of the Acoustical Society of America* 75(1):224–230.
Paradis, Carole and Darlene LaCharité. 1997. Preservation and minimality in loanword adaptation. *Journal of Linguistics* 33(2):379–430.
Paradis, Carole and Antoine Tremblay. 2009. Nondistinctive features in loanword adaptation. In C. A. Wetzels (Ed.), *Loan Phonology*. *Current Issues in linguistics Theory* 307:211–224.
Peperkamp, Sharon. 2005. A psycholinguistic theory of loanword adaptations. In *Annual Meeting of the Berkeley Linguistics Society*, vol. 30: 341–352.
Peperkamp, Sharon, Inga Vendelin, and Kimihiro Nakamura. 2008. On the perceptual origin of loanword adaptations: experimental evidence from Japanese. *Phonology* 25(1):129–164.
Pierrehumbert, Janet Breckenridge. 1980. *The phonology and phonetics of English intonation*. Ph.D. thesis, Massachusetts Institute of Technology.
Poon, Pamela G. and Catherine A. Mateer. 1985. A study of VOT in Nepali stop consonants. *Phonetica* 42:39–47.
Rami, Manish K., Joseph Kalinowski, Andrew Stuart, and Michael P. Rastatter. 1999. Voice onset times and burst frequencies of four velar stop consonants in Gujarati. *The Journal of the Acoustical Society of America* 106(6):3736–3738.
Sailaja, Pinjali. 2009. *Indian English*. Edinburgh: Edinburgh University Press.
Schertz, Jessamyn and Sarah Khan. 2020. Acoustic cues in production and perception of the four-way stop laryngeal contrast in Hindi and Urdu. *Journal of Phonetics* 81:100979.
Schwarz, Martha, Morgan Sonderegger, and Heather Goad. 2019. Realization and representation of Nepali laryngeal contrasts: Voiced aspirates and laryngeal realism. *Journal of Phonetics* 73:113–127.
Shimizu, Katsumasa. 1989. A cross-language study of voicing contrasts of stops. *Studia phonologica* 23:1–12.
Silverman, Daniel. 1992. Multiple scansions in loanword phonology: evidence from Cantonese. *Phonology* 9(2):289–328.
Upadhyaya, U. Padmanabha. 1972. *Kannada Phonetic Reader*. Mysore, Central Institute of Indian Languages.
Vendelin, Inga and Sharon Peperkamp. 2006. The influence of orthography on loanword adaptations. *Lingua* 116(7):996–1007.
Weinberger, Steven. 2014. The speech accent archive. Data corpus. George Mason University.
Whiteside, Sandra P. and Caroline J. Irving. 1997. Speakers' sex differences in voice onset time: Some preliminary findings. *Perceptual and Motor Skills* 85(2):459–463.
Wiltshire, Caroline R. 2005. The "Indian English" of Tibeto-Burman language speakers. *English World-Wide* 26(3):275–300.
Wiltshire, Caroline R. and James D. Harnsberger. 2006. The influence of Gujarati and Tamil L1s on Indian English: A preliminary study. *World Englishes* 25(1):91–104.
Yip, Moira. 1993. Cantonese loanword phonology and optimality theory. *Journal of East Asian Linguistics* 2(3):261–291.
Zhang, Jie and Yuwen Lai. 2008. Phonological knowledge beyond the lexicon in Taiwanese double reduplication. In Y.-E. Hsiao, H.-C. Hsu, L. Wee, and D.-A. Ho, (eds.), *Interfaces in Chinese Phonology: Festschrift in Honor of Matthew Y. Chen on His 70th Birthday*: 183–222. Taiwan: Academia Sinica.