1 Introduction

India may have the second largest population of English speakers in the world, with over 122 million citizens of India describing themselves as speakers of English in the 2011 Census. The codification of the sound system of "Indian English" as a distinct variety began approximately fifty years ago (CIEFL 1972, Kachru 1983, Bansal & Harrison 2013 [1972]), and continues to this day (Sailaja 2009, Pandey 2015). However, descriptions of the phonetics and phonology of Indian English (IndE) reveal a tension between perhaps aspirational portrayals of IndE as a unified accent with a single set of norms vs. detailed descriptions of varieties of English in India based primarily on the first languages (L1s) of the speakers (e.g., Nagarajan 1985, Jose 1992, Wiltshire & Harnsberger 2006). There are also occasional claims of regional variations (Gargesh 2004) and commonalities (Wiltshire 2005, Wiltshire 2015).

Studies on variation within IndE have documented differences in terms of consonant and vowel inventories, allophonics, phonotactics, and suprasegmentals like stress, intonation, and rhythm. On the other hand, recent research suggests that IndE, especially that of educated urban speakers, has been converging toward a more homogeneous standard. Maxwell and Fletcher (2009: 66) claim that "that there are a range of shared vowel categories across speakers of IndE of different L1 backgrounds," based on comparing results from Hindi and Punjabi L1 speakers with those from speakers of other L1 backgrounds, such as Tamil and Gujarati (Wiltshire & Harnsberger 2006). Sirsa and Redford (2013) found both segmental and prosodic similarities among speakers of two different L1s, Hindi and Telugu, summarizing their results as largely consistent with their hypothesis that IndE has phonological targets distinct from those of Indian languages.

This Element explores questions of what unites IndE accents across the nation, what distinguishes subvarieties, and, to the extent possible, what are the sources of these accent features. As English in India, like English everywhere, has developed through contact and over time, theories of language development (e.g., Mufwene 2001, Schneider 2003, 2007, Trudgill 2004) suggest sources of both uniformity and variability may be found through an examination of founder varieties, substrates, linguistic markedness, and processes such as dialect leveling, koineization, and focusing. I begin with an overall description of the linguistic situation in India and English's place in it, along with discussions of the object of study ("IndE accent") and potential factors involved in its development (Section 2). I then combine the findings of acoustic studies, including my own, on IndE sounds, from consonants and

vowels through suprasegmentals (Sections 3–5), evaluating possible unifying and distinguishing characteristics and their sources. In Section 6, I report on the ability of IndE speakers and computational analyses to perceive variation within IndE, and speakers' attitudes toward variation and identity, before discussing the overall findings and areas for future research in Section 7.

2 Linguistic Situation in India

I first provide brief descriptions of the range of languages used in India (Section 2.1), the current status of English (Section 2.2), the meanings of the term "Indian English" and its use in this Element (Section 2.3), and potential sources of uniformity and variability in the development of English in India (Section 2.4).

2.1 Languages in India

India is home to a large number of languages (Figure 1); the 2011 Census listed 121 languages with over 10,000 native speakers each (Government of India 2011), while Ethnologue[1] lists a total of 447 living languages.

The vast majority of languages fall into four distinct language families: Indo-Aryan (originally from the West), Dravidian (always in India), Tibeto-Burman, and Austronesian (both from the East). Most of India's population speaks an Indo-Aryan or Dravidian language as their L1 (see Table 1 for examples). Although historically unrelated to each other, languages from the Indo-Aryan and Dravidian families have been in contact for millennia; many have come to share linguistic features, leading to descriptions of India as a "Linguistic Area" (Emeneau 1956, Masica 1976). However, the Tibeto-Burman languages, spoken in the northeast of the country and relatively isolated from the rest of India, are phonologically and phonetically quite distinct. For example, most Tibeto-Burman languages lack retroflex consonants but have phonemic tone, while Indo-Aryan and Dravidian languages generally have retroflex consonants and lack tone. Assamese, genetically an Indo-Aryan language, is in close proximity with Tibeto-Burman languages and shares some characteristics with its neighbors, including a lack of retroflexion.

Table 1 provides background about the distribution of eighteen languages in India, with the number of speakers reporting each as a "mother tongue" (L1) from the 2011 census (Government of India 2011); these specific languages are chosen because they are referred to in the research below. This census also asked speakers to report any second or third languages, and for the L1s here I have included the number who reported English as an L2/L3.

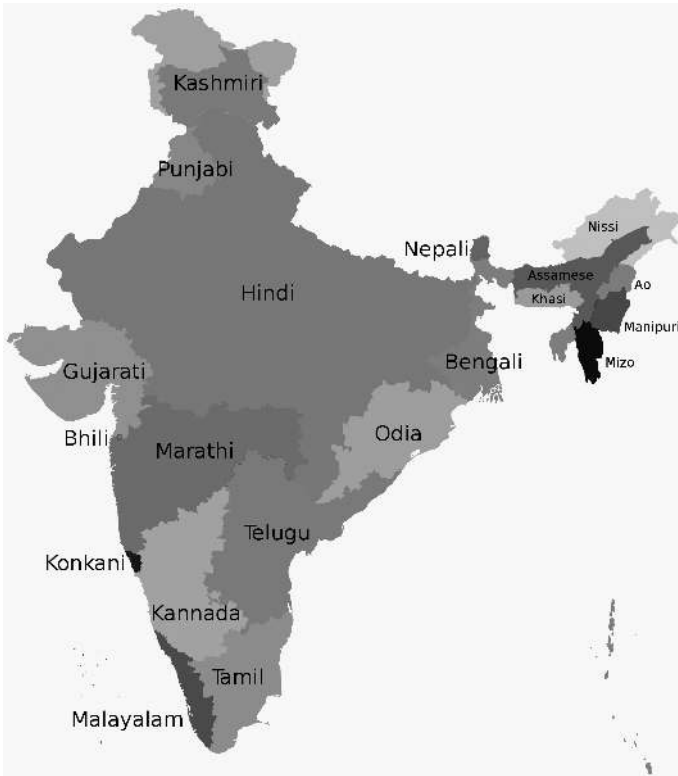[1] www.ethonologue.com/country/IN. Accessed 2020/1/16.

Figure 1 States and union territories of India by the most commonly spoken L1[2]

All numbers are self-reports and may be problematic in a variety of ways (Kohli 2017 provides a good critique). For example, it is unclear what speakers mean when describing a language as their "mother tongue"; it may be the language of their mother, whether or not the respondent speaks it best or first. There is no provision for indicating level of competence or amount of use for any of the languages listed. Those listing English as L2/L3 may range widely in both: at one extreme, people who attended English-medium schools from pre-kindergarten and currently use English daily with friends and work-colleagues, and at the other, people who grew up with a different language and first encountered English as a school subject at age twelve or later, with little use for it after leaving school.

The census calculated overall rates of bilingualism (26 per cent) and trilingualism (7.1 per cent) of the population (Government of India 2011), both of which are surely too low. Sridhar (1989) suggests reasons for the

---

[2] https://commons.wikimedia.org/w/index.php?curid=51479325

Table 1 Number of speakers of L1s and those knowing English (Government of India 2011).

Family	Language	Primarily spoken in	Number of L1 speakers	Number of English L2	Number of English L3
Indo-European	English	Throughout	259,578	76,578,017	45,554,093
Dravidian	Kannada	Karnataka	43,706,512	3,450,332	962,147
	Malayalam	Kerala	34,838,819	6,728,063	890,036
	Tamil	Tamil Nadu	69,026,881	12,325,941	783,805
	Telugu	Andhra Pradesh	81,127,740	8,074,805	2,900,566
Indo-Aryan	Assamese	Assam	15,311,351	984,983	748,554
	Bangla	West Bengal	97,237,669	4,710,845	1,812,485
	Gujarati	Gujarat	55,492,554	937,868	6,752,677
	Hindi	UP, MP	528,347,193	32,018,890	3,238,254
	Marathi	Maharashtra	83,026,680	1,395,659	10,220,047
	Oriya/Odia	Odisha	37,521,324	4,879,878	2,045,591
	Punjabi	Punjab	33,124,726	2,036,498	7,828,657
Tibeto-Burman	Angami/Tenyidie	Nagaland	152,796	62,322	11,833
	Ao	Nagaland	260,008	71,657	21,604
	Meitei/Manipuri	Manipur	1,761,079	480,817	108,473
	Mizo/Lushai	Mizoram	830,846	149,076	7,247
	Bodo (Boro)	Assam	1,482,929	31,991	38,584
Austronesian	Khasi	Meghalaya	1,431,344	237,173	16,944

underreporting: "speakers are reluctant to claim competence in a language unless they can read and write it; many languages are traditionally regarded simply as dialects of one of the major languages of the region; competence in a non-prestigeful language is not considered worth mentioning" (1989: 2). The final point raises the possibility that speakers will claim English as an L2/L3 regardless of their level of competence or use, as it is generally considered prestigious. Nonetheless, Table 1 illustrates that percentages of speakers reporting English as an L2/L3 varies widely based on L1 and region. Although the overall average is about 10 per cent, L1 Malayalam, Angami, Ao, and Manipuri speakers average over 20 per cent, while Bengali, Hindi, and Bodo speakers average under 7 per cent.

The national language of India is Hindi, though the constitution designates English as a co-official language. States have generally been designed along linguistic lines, and may choose their own languages for official status along with the national languages Hindi and English. Thus Gujarati is the official language of Gujarat, Telugu of Andhra Pradesh, and so on. However, some states, especially those with a great deal of internal linguistic diversity such as Nagaland, chose (Indian) English as an official state language and regional lingua franca. Other states list English as one of their official languages, including Meghalaya, Goa, and Tripura, and many list it as an "additional official language" (Haryana, Karnataka, Mizoram, etc.). Although English plays a role in government, businesses, and schools across India, the rise in English has not led to widespread language shift (Sahgal 1991: 300); instead English provides an additional resource in the multilingual repertoire of its users. Over 99 per cent of Indians list an Indian language as their L1, and few, approximately 0.02 per cent, list English. Nonetheless, English has been described as "a major player in the language ecology of contemporary India" (Sridhar 1989: xiv), so I turn now to its place within India.

2.2 English in India

English was introduced to India by the British, beginning in the 1600s; for more detail on its past, see the overviews provided in Schneider (2007), Mukherjee (2007), and Sharma (2017). The term 'British' encompasses a wide range of varieties, including not only the most prestigious standard forms in England, but also regional and social dialects, along with other British varieties such as Scottish, Welsh, and Irish. Bernaisch and Koch (2016) point out that even Americans had a small presence in the early days of colonization. The potential effects of these varieties on the development of IndE are discussed further in Section 2.4.

Currently English performs a range of public functions as the language of the national legislature, the legal system, and some state governments, as well as being widely used in business and higher education, especially in the sciences, medicine, and technology (Sailaja 2009). English is primarily used with other Indian speakers of English, rather than outsiders, and Kachru (1976) observes that in India "the English language is used to 'integrate' culturally and linguistically pluralistic societies. 'Integration' with the British or American culture is not the primary aim" (1976: 229). More recently, Mallikarjun (2020: 166) reaffirms that the primary motivation for learning English continues to be for use as a lingua franca inside India, not with external speakers. English use in India is largely an urban rather than rural phenomenon, and the concentration of the population into urban centers has been increasing over the past six decades (Kohli 2017: 31). Kohli further notes that the urban vs. rural division correlates with privilege, including opportunities to learn and use English, and Agnihotri and Khanna (1997) commented that "those who form their impressions based on their experiences in Bombay and Delhi are likely to have a misleading picture of the use of English in India" (1997: 70). Furthermore, use of English is more common in the public domain than the private. Agnihotri and Khanna's survey of 1,128 urban users of English reported home use as only 36 per cent on average, although it also showed that "in urban metropolitan India, English is making serious claims as a language of peer group communication" (1997: 67).

While English is not widely replacive of Indian languages, nor widely used as a home language, two journalists over the past decade have described a relatively new phenomenon in which English is both. In mixed marriages in urban settings, English has become the home language and the L1 for children in those homes, according to Rai (2012) and Pai (2018). Rai describes communities in Bangalore, where she writes that a "generation of urban children is growing up largely monolingual – speaking, thinking and dreaming only in English." Pai (2018) similarly reports on a new "caste" based primarily on advanced English skills, a group she describes as "affluent, urban, highly-educated, usually in intercaste or inter-religious unions." The size of this emerging group is quite small; even Pai estimates that only the top 1 per cent of people who use English in India fall into this group, which may be the same small subset of the urban population Krishnaswamy and Burde (1998: 127) described as having lost their mother tongue and cultural roots in favor of English. However, even the top 1 per cent of English users in India means over a million people, and as both urban populations and mixed marriages rise, this is a phenomenon to monitor.

As for popular culture, English has made only small inroads into movies, music, and the performing arts (Sailaja 2009), but has successfully increased its

presence in both publishing and TV since the opening of the economy in the early 1990s (Sailaja 2009, Chand 2010, Kohli 2017). Kohli (2017: 34–35) documents the presence of English in published media, with a thriving English newspaper industry and a third-place ranking, after only the USA and the UK, in English book publishing. Furthermore, Kohli (2017: 35) observes that "English also has a powerful and growing presence in Indian television and cable news channels some of which provide round-the-clock programs," while Chand (2010) writes that "Cable TV channels based in India, e.g. NDTV, have been influential in de-stigmatizing various non-RP Indian accents through talk shows and other programs in IndE" (2010: 25).

Although media exemplifies a range of accents to Indian users of English, it is the choice of pedagogical model that has garnished the most attention from academics. As English came to India with the British, the presumed model had been some form of British English (BrE) for generations, and many teachers were foreigners. Since the late 1960s, however, there has been a movement toward using a more local English model (CIEFL 1972, Bansal & Harrison 2013[1972], Nihalani et al. 1979). Nihalani, Tongue, and Hosali claim "this view that the only suitable model for Indian learners is British Received Pronunciation is not shared by the majority of the people in the country, not even by many distinguished teachers of English" (1979: 204). Newer models were prescribed, based on descriptions of English as spoken by proficient speakers from around India, with the explicit goal of devising a form of English that serves as "a socially acceptable pronunciation devoid of regional peculiarities" (Pandey 1981: 11). The principle behind such models of English have been accepted to the extent that currently, English teachers in India are primarily speakers of a local English.

The development of a local model belongs in Stage 4 of the Schneider (2003, 2007) model of Postcolonial English development. This stage, "endonormative stabilization," accompanies political and psychological independence, where-upon a local variety of English begins to be seen as an expression of a new and independent identity. These developments should foster an increased accept-ance of the local English, as the "existence of a new language form is recognized and this form has lost its former stigma and is positively evaluated" (Schneider 2007: 50). An increasingly positive attitude toward IndE within India has been documented over the last four decades, from Kachru (1976) to Bernaisch and Koch (2016).

Kachru (1976) reports on a large-scale survey of 700 students in BA/MA programs, 196 English college/university teachers, and 29 heads of English Departments. These respondents ranked BrE as their preferred teaching model (averaging 66.7 per cent for faculty, 67.6 per cent for students), well

over IndE (26.7 per cent, 22.7 per cent), although more reported that they themselves speak IndE (55.6 per cent overall). Approximately fifteen years later, Sahgal (1991) surveyed 45 speakers in elite areas of Delhi, and their choice of model showed a large swing to a preference for "ordinary Indian English" (47 per cent) over others, including the English of AIR/TV (All-India Radio/TV) announcers (27 per cent) as "it was felt by some of my informants that AIR/TV announcers imitated the BBC pronunciation and had not evolved their own identity" (Sahgal 1991: 304). Sahgal summarized the results as showing both increased awareness of IndE and more acceptance of its distinct norms. Shortly thereafter (1993–4), Agnihotri and Khanna con-ducted their large (n = 1128) urban survey, finding English-medium education more widespread in the younger generation, which claims higher levels of proficiency and more positive attitudes toward both English and English-speaking Indians (Agnihotri & Khanna 1997: 98). Furthermore, "English is perceived as one of the Indian languages by nearly 75 per cent of the informants in this study" (Agnihotri & Khanna 1997: 90).

Later surveys continue to elicit answers that some form of BrE is the best teaching model (Hohenthal 2003, Padwick 2010), yet show positive attitudes toward IndE, especially for use within India. Hohenthall's thirty participants agreed overall with statements like "I like speaking English" (80 per cent yes) and "English is important to India as a whole" (90 per cent yes), while 55 per cent of Padwick's fifty participants chose IndE, when asked which variety should be spoken in India, over second place "don't know/mind" (21 per cent). As pointed out in Bernaisch and Koch (2016), even the choice of a different model for the classroom does not mean that the participants lack a positive attitude toward their own English. Bernaisch and Koch (2016) found positive attitudes toward IndE in their survey that asked participants to rate how well thirteen words pairs, related to competence, power, solidarity, and status, characterized IndE, BrE, AmE, and Sri Lankan English. Based on ninety-four responses from highly-educated urban participants, the overall results show positive attitudes toward all varieties. Though BrE is rated higher on most categories, IndE rates higher on solidarity attributes 'friendly' and 'humble', suggesting covert prestige. Bernaisch and Koch also found that young women both use IndE structures and have the most positive attitudes toward IndE, and, as women often lead linguistic change, they suggest that positive attitudes would continue to grow among IndE speakers.

These studies indicate a growing awareness and acceptance of local IndE as playing a role in India, with at least covert prestige but continued mixed feelings about the choice of a model. Issues of model choice may also relate to how English in India is perceived: as a unified goal or as a set of acceptable varieties, an issue to which I now turn.

2.3 "Indian English" as a Variety/Set of Varieties

In Schneider's (2003, 2007) dynamic model, English in India has likely passed stage 3, "nativisation," in which "the shape of English is a strongly localized one, a characteristic which is due to some extent to the fact that learners have approximated not inaccessible external models but rather local ones" (2007: 167), and entered stage 4 "endonormative stabilization," in which the localized norms provide the target for acquisition. Stabilization is normally followed by differentiation among varieties, whether regional or social in stage 5, although Schneider (2003) warns that the ideal of homogeneity at stage 4 is usually somewhat mythical. Mukherjee (2007) argues that "the situation in which Indian English finds itself today could be seen as a stable, productive steady state in the evolutionary process in which there is an equilibrium between conflicting forces of progression and conservativism" (2007: 157), mixing some stage 3 traits (complaint tradition) with some of stage 5 (heterogeneity developed and developing).

An "Indian English accent" is easily recognized, both by foreigners and IndE speakers (Bush 1967, Bansal 1976, Chand 2009, McCullough 2013, Fuchs 2015), suggesting that there are characteristics that distinguish it as a whole from other varieties of English. Bush (1967) found that her twelve AmE listeners correctly identify IndE words and sentences when presented with clear speech samples from four speakers each of AmE (Midwest), BrE (RP), and IndE (L1 Hindi from U.P.). These same listeners were not as accurate distinguishing between AmE and BrE samples. McCullough (2013) showed that even very short samples, a consonant-vowel sequence, were enough for her twenty-eight AmE listeners to distinguish IndE from the other English accents presented (AmE, Korean, Mandarin). Fuchs (2015) created versions of a three-sentence recording from BrE and IndE speakers, manipulated to contain only certain cues (segmental, pitch, rhythm), and asked listeners (17IndE, 17BrE) to judge the speech as British/somewhat British/somewhat Indian/Indian. Both types of listeners clearly distinguished the accents, and all types of cues influenced their decisions, with segmental cues being the most important. These studies all suggest that there are qualities that contribute to the common perception of an accent as IndE. Additionally, while users of IndE have long reported that they can distinguish varieties within IndE and even guess where the speaker is from, there is now some evidence that speakers of IndE can distinguish at least among Northern, Southern, and Northeastern varieties of IndE (Sirsa & Redford 2013, Sitaram et al. 2018, Chakraborty & Didla 2020; see Section 6.1 for further discussion).

Given the potential uniformity and variability, opinions range widely about the proper use of the term "Indian English" and its relationship to variation.

Some use "Indian English" as the name for a particular kind of English from which other varieties deviate, others use it as an umbrella term for the whole set of varieties spoken in India, and still others argue that there is no use for the term because there is only "English in India," not "Indian English." After reviewing some of these positions, I also describe which Indian English this Element examines for potentially relevant characteristics.

Early writings are oriented toward promulgating a standard, with variations of English spoken in India treated as substandard in some way, often as learners' errors or fossilizations falling short of the "target." Such works attempt to identify an "educated Indian English" to treat as a model, sometimes justified as the most likely to be intelligible (e.g., Bansal 1976, Bansal & Harrison 2013 [1972]). It is not always clear whether this "educated" version actually exists; for example, CIEFL (1972: 15–16) provides a long prescriptive list of potential difficulties based on sixteen L1 backgrounds and marks speakers of all backgrounds as having problems with some of their recommended standards. Bansal and Harrison (2013 [1972]: 4), on the other hand, claim that "in every region there are people who have shaken off the gross features of regional accent and speak a more 'neutral' form of Indian English." Pandey (2015) seems to take this viewpoint as well when he writes, "English-medium education as well as higher education has helped reduce the variation to the extent that a more general variety has emerged as an acceptable standard across the subcontinent" (2015: 301). Thus, from this perspective, there is a model IndE accent, and only a lack of proper education or successful learning contributes to variation. Domange (2015) objects that seeing variation only as the result of incomplete learning ignores the range of factors involved in normal varietal development: "IE varieties are not conceptualised as emerging from interaction between the speakers, let alone as developing from one generation to the next" (2015: 535).

The term "Indian English" has also widely been seen as a term covering a whole range of varieties. For example, Chand (2009: 307) writes that "Indian English is an umbrella term for multiple English varieties spoken in India by speakers of varying fluency, nativity, ethnic, regional, and linguistic backgrounds." Sharma (2017) also notes that it is useful as an "umbrella term" for related varieties, and claims that although the term Indian English is generally used to refer to an acrolectal style, "certainly no single variety or standard is shared across the North or the South, much less across the entire country" (2017: 326), raising the possibility that there is more than one acrolectal variety, based on geographical or social factors. As a version of the umbrella viewpoint, "Indian English" may be a cover term for specific characteristics that are shared across varieties, which seems to be the position of Nihalani et al. (1979) when they suggest that "Indian English, in spite of all the variety that one notices from