A Forensic Phonetic Study of Indian English: Phonetic Features as an Indexical Marker

Sinjini Chakraborty, Dr. Grace Suneetha Didla

Department of Phonetics, The English and Foreign Languages University, Hyderabad, India

Abstract—Several researchers have documented the marked similarities and differences in both segmental and suprasegmental features across varieties of Indian English (IE). There also exists a body of literature on how these differences can be justified by the fact that the speech of a non-native English speaker carries the burden of one’s mother tongue, commonly known as mother-tongue influence, or MTI. This paper aims to establish that such segmental properties can be employed to arrive at indexical information about the speakers, like their geographical and linguistic background. This has relevance to forensic speaker identification as it assists in speaker profiling. The phonetic analysis included in this study observes the similarities and differences in segmental features of varieties of IE spoken across 6 different zones of India (East, North-East, West, Central, North and South). The study also includes a perceptual test with 10 naive listeners to identify the most distinguishable zones on the basis of their spoken IE. Finally, a comparison between the phonetic analysis and the results from the perceptual test was carried out to verify if these segmental differences resonated with their perceptual outcome. Results indicate that segmental properties can be employed as carriers of indexical information like one’s linguistic and geographical background, to a moderate precision.

Keywords—Indian English, Forensic Speaker Identification, indexical marker, segmental features, speaker profiling.

I. INTRODUCTION

Forensic linguistics is a discipline amalgamating the study of language and the law. One of the branches of forensic linguistics is forensic phonetics, which entails the identification of incriminating speech samples based on phonetic features. Since its early emergence as a distinct study in the 1990s, the police and legal bodies have always depended upon the intricacies of the voice and speech as primary evidence. With the ever-growing technological development, the number of judicial cases dealing with recorded speech, whether over the phone or the internet, has been on the rise. In view of this, the application of phonetics appears to be used as viable evidence, pertinent enough to be recognized as valid by law. Forensic phonetics focuses on the analysis of spoken communication, which includes forensic speaker identification (FSI), enhancing and decoding spoken messages, analysis of emotions in voice, authentication of recordings and related. FSI involves a process of analysing an unknown voice and categorizing the voice as belonging to a specific race, age, gender, linguistic background, geographical background, and so on [1]. Information of this kind, commonly known as indexical information, is indicated by indexical markers, which are of great importance in FSI as they are indices to an individual’s identity. Indexical markers of a speaker refer to the (i) individual identifying markers like his age and gender, and (ii) indications of his membership to specific linguistic and social groups, that is, information about his linguistic, regional, social, and cultural background. Since the objective of FSI is to identify an unknown speaker, extraction of indexical information assists in the construction of the speaker’s possible identity. This process of arriving at a blueprint of an unknown voice is called speaker profiling, and it is of paramount importance in FSI. The current study has relevance to speaker profiling based on the sub-variety of English spoken by the Indians.

In the case of bilinguals, speaker profiling is particularly of value because in-depth phonetic analysis of the unknown speech sample in a non-native language (like
English) can lead to identifying the speaker’s native language. It is known that all languages differ in their phonemic structures [2], and it is precisely for this reason that a bilingual speaker carries undertones of the native language when speaking in a non-native language. It is a language’s phonemic structure that determines to a large extent the foreign accent of a bilingual speaker when they are speaking in a non-native language [3]. It is in this capacity that mother tongue influence (MTI) becomes an important linguistic phenomenon that aids in the process of speaker profiling. MTI manifests itself to the largest extent in the segmental features of a bilingual speaker. There have been several works suggesting that variations in the consonantal, and more importantly, the vowel phonemes in different varieties of a language can be analysed to arrive at individual-identifying information about the unknown speaker [4], [5]. Furthermore, there have been several cases in FSI, when an analysis of these segmental features led to robust indexing of the speaker of that voice sample [3], [6]. Therefore it can be assumed that in the Indian-context too, an in-depth analysis of the effects of MTI on the segmental features of Indian English could lead to indexical information like one’s linguistic background, and hence, find similar uses in FSI. Additionally, it could also indicate the speaker’s possible geographical background, since the state formation in India was based on linguistic grounds [7].

II. LITERATURE REVIEW

Bailey and Robinson opine, “English is a world language; and as such it has national varieties” [8]. Truly English has been owned up by various countries and given a distinct shape to. Every variety of English spoken in the world has its own distinctive phonetic features in the form of their segmental and suprasegmental characteristics. MTI on accents, use of shibboleths and so on. India is no exception to this, where English remains the predominant language of cross-linguistic spoken communication [9]. However, only 3% of the Indian population is fluent in English [10]. As a result, English when spoken by Indians, is laden with phonetic influences from the speaker’s mother tongue. This gives rise to a broad variety called Indian English, which Verma describes as a “non-native second language variety”, and which “has a complex network of features contributed by the mother tongue of its speakers, by their cultures and also intra-language analogical processes” [11].

Nevertheless, there has been much debate on whether there can be an absolute variant called Indian English at all. Bansal had carried out seminal work in 1967 on the intelligibility of English in India [12]. However, with the ever-changing nature of language interactions, it becomes imperative to study the emerging phonetic features of the different varieties of IE. As a result, there have been several works demonstrating how each spoken variety of Indian English is shaped by influences from the speaker’s mother tongue [13], [14], [15]. Fuchs documents several distinct phonetic features of Indian English, both segmental and suprasegmental features, and states that much of the research on Indian English phonetics and phonology relies on data from speakers with Dravidian and Indo-Aryan linguistic background [16]. The only exception being Wiltshire, who presented acoustic evidence of considerable differences in segmental features of the variety of Indian English spoken by the native speakers of Tibeto-Burman languages in India [17].

Most research on segmental features of Indian English has been implicitly or explicitly based on vowels. It has been demonstrated that most vowels differ in their realisation between Indian English (IE) and British English (BE), thereby accounting for the segmental differences between these two variants of spoken English [18], [19], [20]. The most conspicuous phonetic difference between the vowels of these two variants was the merging of certain diphthongs of BE into monophthongs in IE. When it comes to consonants, however, only one potential merger contrasting IE with BE was reported; the labio-dental fricative /v/ in BE is often realised as a labio-dental approximant [ʋ] in IE, and also merged with /w/ often[21], [22]. Furthermore, it was reported that the alveolar plosives in BE /t/, /d/ were often realised as their retroflex counterparts in IE; likewise, the dental fricatives /θ/, /ð/ were replaced by dental plosives /t̪/, /d̪/, of which the voiceless phoneme is often aspirated [t̪ʰ] due to the influence of spelling in IE [19], [22]. It was also documented that the dental plosive to retroflex conversion varied both within and across speakers of IE [23].

There has been substantial research on the L1-dependent phonetic features across different varieties of IE. Most of the work documents L1-dependent similarities across the varieties of IE, thereby ear-marking some pan-Indian phonetic features [24], [25]. Additionally, it has been established that speakers of IE also display L1-dependent differences in the segmental and suprasegmental features, which indicates that IE cannot always be understood as a single, cohesive dialect of English. There exist several varieties within this broad category that can be identified reliably by their distinct phonetic features; Gujarati and Tamil speakers of IE differed in their back vowel system, rhoticity and retroflexion [26].

Even though a large body of research has documented several such robust markers of segmental
differences within and across varieties of IE, there hasn’t been enough work to establish that such distinguishable features on the phonetic level can be representative of indexical information, like one’s linguistic and geographical background. This study draws in from the body of work in varieties of IE, and verifies their application in a forensic-phonetic context. A geographical zone-wise categorisation of typical phonetic features of the variant of IE spoken can be of immense help when applied in speaker profiling, as FSI prefers analysis of speech parameters that carry individual-identifying potential. This paper attempts to indicate that an analysis of distinguishable segmental features could prove to be one such parameter in the identification of an unknown voice’s possible linguistic, and/or geographical background.

III. OBJECTIVES OF THE STUDY

This study attempts to:

- offer a comparison of segmental features across the varieties of IE,
- offer a comparison of segmental features within each variety of IE,
- offer an account of the pan-IE segmental features,
- explore whether the geographical and/or linguistic background of an IE speaker can be perceptually identified,
- offer a correlation of the phonetic analysis with their perceptual judgement.

IV. SCOPE OF THE STUDY

The objective of the research is to establish similarities and differences in solely the segmental/phonetic features of the varieties of Indian English. Since the foundation of state formation in India has been in linguistic commonality, it could be assumed that (a) languages/dialects spoken within a state will share certain common features, and likewise, (b) languages from neighbouring states will also share linguistic commonalities. Keeping this in mind, the study includes 6 zones that are each representative of (a) their geographical contiguity and (b) linguistic belonging, namely,

- North,
- North East,
- East,
- West,
- Central, and
- South.

Due to limited availability of data, this study takes into account only 12 speakers - 2 speakers per zone, but each from a different state that falls under these zones. Though this is a small sample, it serves to represent regional varieties in Indian English.

In describing the segmental features of each zone, RP has been taken as a standard for purposes of comparison as British English is not only widely used, accepted and is intelligible all over the world, but also has been the standard in English Language Teaching in India since 1835 [27].

V. METHODOLOGY

5.1 Speakers

To ensure that the subjects represented a fairly wide range of linguistic groups, a total of 12 speakers were chosen from 12 different states of India, covering the 6 aforementioned zones. All the speakers learnt English as their second language during school education, and were proficient users of English. Data were collected from students and faculty at The English and Foreign Languages University, Hyderabad, India. None of the speakers had any prior training in phonetics. The speakers included both men and women, their age ranging from 20 to 35 years. None of them reported any speech disability.

5.2 Listeners

A separate set of 10 listeners were chosen for the experiment; each belonging to a different state of India, ensuring linguistic heterogeneity. The listeners were, at the time, all resident students at The English and Foreign Languages University, Hyderabad for at least 2 years. This was to ensure that all the listeners had a fair exposure to different accents of Indian English, as the location of the study was a multicultural, metropolitan space. The group comprised both men and women, their age ranging from 20 to 27 years. None of the listeners was diagnosed with any hearing disability, nor did they have any prior training in phonetics.

5.3 Text

The text chosen for this study was the Rainbow passage, which is a phonetically-balanced text, where the ratios of the various phonemes reflect the ratios of those phonemes in normal unscripted speech [28]. It contains all the 44 sounds of English in either word-initial, -medial or -final positions, and is used by phoneticians across the globe for accent checking purposes.
5.4 Speaking Task

When collecting data, the speakers were given a few minutes to familiarise themselves with the passage, and then asked to read it aloud, like they were reading it out to someone. Data was collected separately for each speaker, ensuring no external linguistic bias.

5.5 Perceptual Task

The listeners were briefed on the purpose of the study, explained the process of the experiment and introduced to the questionnaire (see Fig. 1). All doubts were satisfactorily clarified before the task began. The listeners were played a total of 12 samples; each sample was a 45-seconds long excerpt from the middle of each speaker sample. It was ensured that the listeners were not familiar with the speakers.

The samples were played in an arbitrary order to each listener separately. Each sample was played for a maximum of three times, if requested by the listener. After each sample was played, the listeners were asked to note their responses for two questions. The listeners were asked to note their responses to the questions (3 questions per speech sample). Each session lasted for 15-20 minutes, ensuring no listener fatigue.

![Fig. 1: Sample of the Questionnaire](https://dx.doi.org/10.22161/ijels.53.2)

5.6 Analysis

5.6.1 Phonetic Analysis: A detailed phonetic analysis of the recorded speech samples was carried out. It involved transcription of the 44 target phonemes as uttered by each of the speakers. These were listed against their RP realisations for comparison.

These transcriptions were then analysed for similarities and differences in phonetic features, both within and across the varieties of Indian English, as represented by their corresponding zones. The data were also analysed to arrive at a set of pan-IE segmental features.

5.6.2 Perceptual Analysis: Listeners’ responses from the questionnaires were tabulated separately for each zone; noting the number of correct and incorrect positive identifications for each zone. The correct positive identifications were then gathered to arrive at the zones that were most and least identifiable by naive perceptual judgement.

5.6.3 Correlation of Perceptual Judgement with Phonetic Analysis: The results from the perceptual task were correlated with the findings of the phonetic analysis to offer an account of the phonetic similarities and differences across the varieties of Indian English included in this study.

Additionally, an indication to distinguishable phonetic features within a zone, and across zones (that is, pan-Indian phonetic features) have been provided.

VI. OBSERVATIONS

6.1 Phonetic Analysis of Segmental Features

As can be seen in Fig. 2, the vowels and diphthongs of IE seem to exhibit more differences with BE than the consonants in Fig. 3. Since these figures display only the set of phonemes that have at least one instance of difference between their IE and BE realisations, it can be understood that the monophthongs and diphthongs differ more than the consonants; 17 out of 20 vowels listed in the figures, as opposed to only 16 out of 24 consonants. In both the figures, the instances where IE realisations are the same as BE realisations have been indicated by a tick mark (√). Additionally, instances where IE realisations of phonemes are unique to a particular zone have been indicated with a red block in the figures (2 and 3).

Fig. 2 shows that the vowels /ɒ/ and /ɔ:/ differed greatly for all zones from their BE realisations; but IE realisations were significantly consistent across the zones. Likewise, the central vowel /ʌ/ was consistently replaced by /ɑ:/ and the long vowel /u:/ was replaced by /ʊ/ across several zones, the exception being in North East which conformed to the BE realisations of both these vowels. It must be noted that the monophthongs /a:, ʊ/ and the diphthongs /aɪ, aʊ/ were realised as in BE across all the zones in this study.

However, diphthongs proved to be most distinguishable for IE speakers as compared to BE, both in this study and others before. Most of the diphthongs merged into a monophthong: the diphthongs /eɪ, eə/ merged into /e/ across all zones; the BE diphthong /au/ merged into either /o/ or /ɒ/. The diphthong /au/ in IE was mostly consistent with its BE realisation, excepting some
zones, wherein it was realised as /ɑɪ/. It can be noted that except for North East, the second element /ə/ in the diphthongs /ɪə/ and /ʊə/ were consistently lengthened to the back open vowel /ɑ/.

As can be seen in Fig. 3, consonants seem to display several variations both within and across zones. Unlike the data on vowels and diphthongs, the data on consonants did not suggest an overt pattern in segmental differences between zones. However, the segmental features of consonants in IE show several differences from that of BE; mostly in fricatives. The dental fricatives /θ/ and /ð/ are consistently replaced with their respective dental plosive counterparts, mostly along with aspiration. Additionally, even though most voiceless fricatives of BE were realised as is in IE across zones, their voiced counterparts consistently show several allophonic realisations in IE; but their reliability across or within zones could not be established from such a small set. Most plosives in IE across zones were consistent with those in BE, excepting the case of North East, wherein the alveolar plosives were replaced with the alveolar tap. It must be noted that speakers from the South zone consistently replaced voiced fricatives and affricates with their voiceless counterparts, and /r/ with the retroflex flap. Though the speakers had no problem with articulating the glide /j/ /w/ was always realised as the approximant /ʋ/ in IE across the zones, excepting North East. However, all the speakers across zones pronounced most plosives and nasals like in BE.

Combining the data in Figs. 2 and 3, it can be suggested that the segmental features of IE in the South zone display maximum instances of radical inter-zone differences (marked in red), followed by the North East zone, when compared to the rest of the zones; East IE and West IE show the least. On the other hand, IE in North East seems to bear another form of segmental difference than the rest of the zones: data from North East IE displays maximum instances of BE realisations for both vowels and consonants (marked with ✔) as compared to the variants of IE in other zones. On the contrary, the Central IE displays the least similarity with BE segmental features.
It must be noted that all the listeners actively responded to the zone-identification question (Q1) on the questionnaire (see Fig.1), but were hesitant to respond to the state-identification question (Q2), on the grounds of not being certain. Also, when the listeners were asked to identify the zones, they were more often than not confused between the North and the Central zones. Speakers from the East were sometimes misidentified for West or North; and West for South or Central. Speakers from North East were mostly identified correctly, but the speaker from Assam (belonging to East) was often misidentified as North East. As for the reason for their identification (Q3), the listener responses were quite varied. Most listeners responded with subjective observations like “sounded similar to a native speaker of Hindi”, “sounds Bihari”, “sounds anglicised”, “the ‘r’ and ‘n’ sound South-Indian”, etc. However, some participants were quite specific in their reasons for identification, somewhat technical in fact, like “the ‘r’ was more rolled” and “the ‘z’ was pronounced as ‘j’”.

VII. DISCUSSION

Drawing in from the Figs. 2 and 3, it can be suggested that vowels and diphthongs seem to be more reliable in establishing differences across the zones included in this study. This is consistent with prior observation that vowel properties tend to exhibit inter-speaker differences more than consonants [2]. Similarly, in keeping with prior research, the data shows that the back vowel system is quite different for IE across all zones, as compared to BE; but not so much the front vowels. For both vowels and consonants, the speakers of both West IE and East IE displayed the lowest number of distinctive phonetic features, and correspondingly, these two zones had the least percentage of positive identifications in the perceptual test. Speech samples from the South and North East, on the other hand, ensured the highest percentages of positive identification by naive listeners. This can be explained by the fact that the segmental features of IE in both these zones exhibited very distinct phonetic features, like consistent use of retroflexes and flaps, respectively, that are otherwise not observed in the IE of other zones.

Furthermore, quite a few pan-IE segmental features were noted in the data, that consistently differed from BE. Figs. 2 and 3 show that the diphthongs /əe/ and /ea/ were always pronounced as /e/ across the zones. Similarly, the BE dental fricatives /θ/ and /ð/ were consistently realised as dental plosives and the semi-vowel /w/ as the approximant /v/ in IE. Most BE long vowels were often consistently shortened (/u:/ to /ʊ/), or fronted (/ɔ:/ to /a:/) across the zones. Another notable trend across the IE zones was the consistent replacement of the /a/ component to the back open vowel /ɑ/ in all the BE diphthongs BE that consist of a /ə/, suggesting that IE speakers show minimal instances of the vowel /ə/ in their speech.

VIII. CONCLUSION

To conclude, it can be said that the geographical background of an Indian can indeed be identified by a phonetic analysis of their speech sample, to a moderate precision. Though this is a small representative data, it serves to indicate that segmental differences can be phonetically analysed to arrive at dependable indexical information about an unknown speaker. However, a more thorough phonetic, perceptual and acoustic analysis is required on a larger population to verify the dependability of segmental characteristics of IE speakers as a parameter in forensic identification or speaker profiling. Even though India is a country of multiple languages, a verified set of pan-Indian features could aid in distinguishing Indians from non-Indians in a forensic scenario. Similarly, a description of distinct phonetic features of IE spoken in each zone in India could contribute to forensic litigation, in case it involves identification of one IE speaker from hundreds of others.

REFERENCES


