Identifying acoustic cues for dialect profiling: Policing in multilingual communities of India

Ravina Toppo[*] and Sweta Sinha

Department of Humanities and Social Sciences, Indian Institute of Technology Patna, India

ABSTRACT

A multilingual country such as India with numerous languages and dialects provides fertile grounds for evasive language crimes. From threat letters to ransom demands, the scope of crime is huge. The cases of illegal immigrants have only added to the fragility of international boundaries especially, during political upheavals. This leads to further vulnerability of society and also creates challenges for the police and law enforcement agencies towards timely intervention. The purpose of the study is to exhibit dialectal variation in Indian English by comparing two varieties. The current paper is based on the acoustic analysis of Indian English spoken by two distinct groups with different mother tongues. Ten native speakers of Hindi and Bangla were recorded in an anechoic chamber. A phonetically balanced passage was selected to be read. The analysis is based on Native Language Influence Detection (Perkins & Grant, 2018) to derive acoustic phonetic correlates that can be used as significant identifying markers to distinguish Indian English speakers of Bangla and Hindi speech communities. The paper highlights that dialect profiling in the Indian context can be efficiently correlated with formant frequencies and Voice Onset Time for speech data. Acoustic analysis was done on PRAAT. PRAAT was used in this study because it has often been used by other similar studies to measure desired acoustic parameters simultaneously. Formant frequencies were measured at the midpoint of the vowels in the PRAAT using the LPC formant measurement algorithm. The normalization procedure was applied to the measured formant frequencies of vowels. The research affirms that acoustic analysis can provide verifiable cues for NLID. The framework can be used in the detection of native language influence in speech-centric criminal cases. The acoustic analysis shows that Indian English has subvarieties that could help in dialect profiling. The variation in Indian English vowel patterns could be due to the influence of the native language of the speakers.

Keywords: Bark difference method; dialect profiling; formant frequency; Indian English; VOT

INTRODUCTION

For the past couple of decades, the police and other law enforcement agencies seem to have relied upon criminal profiling to narrow down the number of suspects. It is believed that the idea of such profiling originated in the Behavioral Science labs of the FBI, where psychologists and criminologists worked together to create a framework to point out certain types of perpetrators (Cooley, 2012; Winerman, 2004). However, until recently, such profiling has not been derived from sociolinguistics, which provides cues about age, gender, geographical origin, and socio-economic status, to name a few.

Profiling in a forensic context has been labeled and explained differently by different researchers. Geberth (2015) perceives a criminal personality profile as "an educated attempt to provide investigative agencies with specific information as

* Corresponding Author
  Email: ravina.toppo@gmail.com

to the type of individual who may have committed a certain crime" (p. 965). Ainsworth (2001) described offender profiling as "the process of using all the available information about a crime, a crime scene, and a victim to compose a profile of the (as yet) unknown perpetrator" (p. 7). The identification of a particular offender is not indicated in profiling. It is established on the likelihood that somebody with a specific set of traits is likely to have committed a specific kind of offense. Profiling in this case operates on two premises: modus operandi (methods of operation) and behavior. The idea behind the premises is that the offender leaves behind detectable clues based on how they operate and behave. The premises when collected can narrow down the list of suspects based on significant observations. This is derived from the rationale that personality is reflected in behavior (verbal/non-verbal), and by observing behavior, the investigator can determine the kind of person responsible for the crime (Chifflet, 2015).

Basing the research on the behavioral aspect, this paper attempts to identify verifiable acoustic correlates that can be used to decipher the geographical origin of a suspect/witness. This is alternatively also referred to as dialect profiling. It involves the task of finding acoustic evidence of the region in which a particular speaker spent their life before adulthood. It can help in identifying the provenance and background of a speaker by their mother tongue. Speaker profiling requires analysis of a recorded speech of a speaker to extract as much information as possible about them (Foulkes & French 2001; Foulkes et al., 2019; Jessen, 2007). French and Harrison (2006) and French and Stevens (2013) listed ten indexical dimensions that can be gleaned from the speech: age, gender, social and educational background, regional background, ethnic group characteristics, influence of the first language, spoken or read sample, presence of disguise speech/language pathology, and intoxication. Speaker's non-native language is usually influenced by the native language of the speaker. In dialect profiling, linguists aim to determine the information about a speaker based on their acoustic features or accents of voice. The issues arising out of illegal immigration in India have been studied primarily from the vantage point of social sciences, as recent as that of the Rohingyas fleeing Myanmar and entering India (Chaudhury & Samaddar, 2018; Shamshad, 2017). Forensic research for the detection of native language influence should have generated an adequate framework to bolster policing in India, but the lack of research in this direction, especially in the Indian context, is starkly evident by the lack of relevant literature.

Native language influence detection (NLID) attempts to indicate the speaker's native language (Perkins & Grant, 2018). In the area of dialect profiling, NLID holds much potential to influence law enforcement agencies in investigating language-related crimes (Grant, 2008; Malmasi & Dras, 2017). NLID holds much potential as multilingualism is becoming more widespread and multilinguals are dominating the number of monolinguals all over the world (Thomason, 2001). NLID can be done by investigating the acoustic features of speech sounds and analyzing the physical properties of sounds produced by speakers (Schuller et al., 2016; Zampieri et al., 2017). A majority of speakers who speak a second language (L2) uphold some mark of native language (L1) in their speech (Oyama, 1976). The arena of NLID is determined by the theory of interlanguage and cross-linguistic effect, established from second language acquisition theories from pedagogical perspectives. This paper will try to present findings based on NLID exhibiting that NLID can possibly demonstrate a sociolinguistic-based method to indicate the possible features for dialect profiling. Moreover, this study also is an attempt to show how linguists can help on developing practical applications that can be used in criminal investigations.

In the present study, NLID was used to examine the acoustic properties of the target language. Acoustic phonetics studies the acoustic characteristics of speech, analysis, and description of physical properties of speech such as formant frequency, intensity, and duration (Kent & Read, 2002). Acoustic phonetic properties which show dialectal differences are formant frequencies, duration, intensity, pitch, and Voice Onset Time (Das & Hansen, 2004; Jessen, 2008; Rose, 2002). The primary approach used to investigate the acoustic properties of the vowel is formant frequency analysis. The first two formants display vowel distributions in vowel space, and vowel distribution depends on dialects. Previous research has suggested that formant analysis of vowels shows significant variations in accents of Indian English due to native language influences. This study is an attempt to exhibit dialectal variation of Indian English through acoustic analysis.

Previous studies have used formant measurements to assess vowel quality in several extensive studies on dialects and sociolects. To study the regional varieties of North American English, Labov et al. (2005) analyzed the first two formants of the vowels. Clopper and Paolillo (2006) investigated American English vowels from six dialect regions by analyzing formant frequencies and duration. Various studies have been conducted on the variations in accent and the effect of native language influence on Indian English. Wiltshire and Harnsberger (2006) investigated the phonetic and phonological influences of L1, Tamil, and Gujarati on Indian English. Their study reveals that the Indian English accent exhibits both phonetic and phonological influences of their native language.

Kalashnik and Fletcher (2007) studied vowel analysis for North Indian English, which shows that there are many different varieties of English spoken in India and the Indian English vowel system is highly influenced by L1. Maxwell and Fletcher (2009) analyzed acoustic and durational characteristics of Indian English to provide a disparity for native speakers of Hindi and Punjabi. Kulshreshtha et al. (2012) studied acoustic characteristics of dialects of Hindi for forensic speaker identification, which illustrates that a speaker's vowel quality, quantity, tone, and intonation could be potential characteristics for dialect identification. Phull and Kumar (2016) categorized Indian English into four different accents exhibiting prominent differences in the accents by formant analysis.

Previous research studies have made an effort to explain the perceptual and acoustic features of accents in speakers and Voice Onset Time (VOT) has been found to be the foremost characteristic that helps in determining whether a speech is accented or not (Das & Hansen, 2004; Yavas, 2002). This study aims to identify potential variations in VOT between speakers of Indian English produced by native Bangla and Hindi speakers. Borden et al. (2007) found that English voiced /b/, /d/, and /g/ phonation starts very shortly after the stop consonant release, while for the production of voiceless stops /p/, /t/, and /k/, there is a delay of 50ms or more after the release.

Previous research has inevitably suggested that Indian English has absorbed many features from the L1s of India, but these studies have limited their scope (Kalashnik & Fletcher, 2007; Kulshreshtha et al., 2012; Maxwell & Fletcher, 2009). In this research study, an attempt has been made to suggest that dialect profiling can be done using acoustic analysis and can be used in criminal profiling. These studies have compared various L1s and their effects on Indian English. However, a comparative study of Indian English produced by native speakers of Hindi and Bangla has not been conducted.

The Present Study
Indian English is an established variety of English spoken by millions of speakers in India. The English language came to India with the British people in the seventeenth century (Schneider, 2007; Sharma, 2017). Indian English is taught as a second language and used as the medium of instruction in the education system. It is used in day-to-day life with other speakers who are non-native speakers of English. It is also called a transplanted variety. Thus, Indian English has its own varieties which are the outcome of complex contact situations. Therefore, a comparative study has been made between the varieties of L2 English and L1 English, to distinguish the transfer of influence. Indian English produced by native speakers of Hindi and

Bangla was selected for the present study because both languages belong to the Indo-Aryan language family. These languages share geographical boundaries in India. Both Hindi and Bangla are syllable-timed languages (Mostafa, 2010).

A majority of people in India speaks Hindi. There are ten vowel sounds in Hindi. They are /ə/, /a/, /i/, /i:/, /u/, /u:/, /e/, /æ/, / o/, / ɔ/ (Tiwari, 1966). The vowels /ə/, /i/, /u/, are always short vowels while the vowels /a/, /i:/, /u:/, /e/, /æ/, / o/, / ɔ/ are always long vowels in length.

Over 175 million people in eastern India speak Bangla (Dasgupta, 2003). There are seven vowel inventories in Bangla with non-contrastive lengths supplemented by quality differences in the short and long high vowels. They are /i/, /e/, /æ/, /a/, /ɔ/, /o/, /u/ (Chatterji, 2002). In Bangla vowels, there is no distinction between short and long vowel sounds. Thus, a change in vowel length leads to no effect on the meaning of the words in which they occur (Rahman, 2008). Acoustic studies have suggested that Bangla vowels are centralized with respect to the corresponding cardinal vowels (Datta, 2018).

Linguists have also done a comparative study of acoustic measurements of Received Pronunciation (RP) vowels to provide empirical evidence of sociophonetic inferences. RP is used by a very small percentage of the British people. It is described as an accent of England, allied with educated members of society and considered to be the standard accent (Wells, 1982). This study attempts to make a comparative study of British English RP with Indian English. Present-day data (Indian English) was assembled to make a comparison of formant frequencies with that of Wells (RP) (1962). Wells (1962) collected acoustic measurements of RP speakers. Data were derived from speakers, who were of university age.

For dialect profiling, an acoustic analysis of a speech sample is made to show the acoustic differences between languages or language varieties. Though previous literature suggests that Indian English is highly influenced by the native language spoken in India, a comparative study exhibiting dialectal variation between Hindi and Bangla has not been conducted. This study is an attempt to show dialect variations of Indian English spoken by Bangla and Hindi speakers. This study is an effort to prove the assertion that dialectal variation can be seen through formant means analysis, formant space analysis, vowel space area, and VOT calculation.

The National Crime Records Bureau (2021) has revealed data showing that crimes have increased in places where Hindi and Bangla are spoken. An enormous rise in crime is also evident with the rapid advancement of technology. We must look for clues in Indian English, a language that is widely spoken throughout India in this era of evolving language-related crimes. Therefore, to solve language-related crimes, it is important to

investigate the native language influence of Hindi and Bangla in the acoustic patterns of vowels in Indian English.

METHOD
Speakers
Ten native speakers, both male and female, of Hindi and Bangla, in the closed age group of 25-35 years who were uniformly exposed to English as their second language as the medium of instruction for school and college education, took part in the data collection for this research study. Ten speakers (five Hindi and five Bangla) belonging to two different linguistic regions were selected for data collection. Each speaker signed a consent form indicating their willingness to participate in the research.

Speech Materials and Recording Procedures
Speech recordings were made in an anechoic chamber. A phonetically balanced passage was selected to be read. The English-language passage was presented to them on a sheet of paper. Each participant was given the reading in advance for familiarity. This was done to ensure that each participant was relaxed and comfortable. Gaps were allowed between sentences and they were allowed to repeat any words where a mistake was made. The continuous speech recording method was used as it might produce a more natural production than having the speakers repeat alone words (Sharf & Masur, 2002). For acoustic analysis of sounds, PRAAT (Boersma & Weenink, 2008) was used. The normalization technique was used for the measurements of the first three formants for a set of vowel tokens.

Acoustic Analysis
The waveform editor PRAAT, used for acoustic analysis, has speech synthesis and analysis functionality. It was used to see the waveforms and spectrograms, as well as to listen to the sound files.

It was used to make an orthographic transcription to measure vowel formants and VOT. After performing an auditory analysis of the target tokens and visual analysis of the formant tracks, vowels were plotted automatically on the spectrogram by PRAAT. A Linear Predictive Coding (LPC) formant tracking algorithm was carried out to measure the formants at the midpoint of the vowels for formant mean analysis, vowel space analysis, vowel space area measurement, and vowel normalization.

For the measurement of VOT, boundaries were manually marked and labeled at the initial burst of stop consonant and the onset of the periodicity of the vowel, and the duration was noted.

## FINDINGS AND DISCUSSION
Formant Mean Analysis
To examine the distinctive features of vowel trajectories, formant frequencies F1, F2, and F3 were taken. The first two formants are the most distinguishing acoustic parameters that can identify vowel quality (Ladefoged & Johnson, 2011). Formant frequencies easily correspond to vowels.

Major articulatory dimensions used in classifying vowels are the frontness and backness of the tongue, the height of the tongue, and lip rounding. The frontness and backness of vowels are related to F2 and F1 co-relates to vowel height. The difference between rounded and unrounded vowels is demonstrated by F3. F3 is used less in vowel categorization and its importance appears to differ across languages. While for some languages the first two formants are adequate for listeners to identify vowel sounds (Adank, 2003).

The average of F1 of Hindi Indian English (HIE) and Bangla Indian English (BIE) has been shown in Figure 1. Figure 1 shows that the vowel /a/ of BIE has the highest F1 and the vowel /iː/ has the lowest F1 value. Vowels /i/, /iː/, and /u/ have almost same F1 values for Hindi speakers. The difference between the formant values of F1 for vowels /i/, /u/, /uː/, /ɔ/, and /æ/ of HIE and BIE is less.

Figure 1.
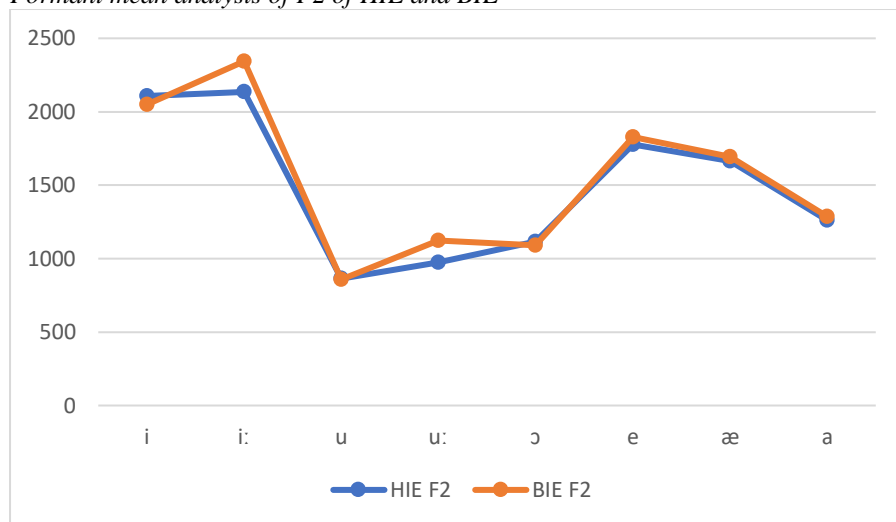*Formant mean analysis of F1 of HIE and BIE*
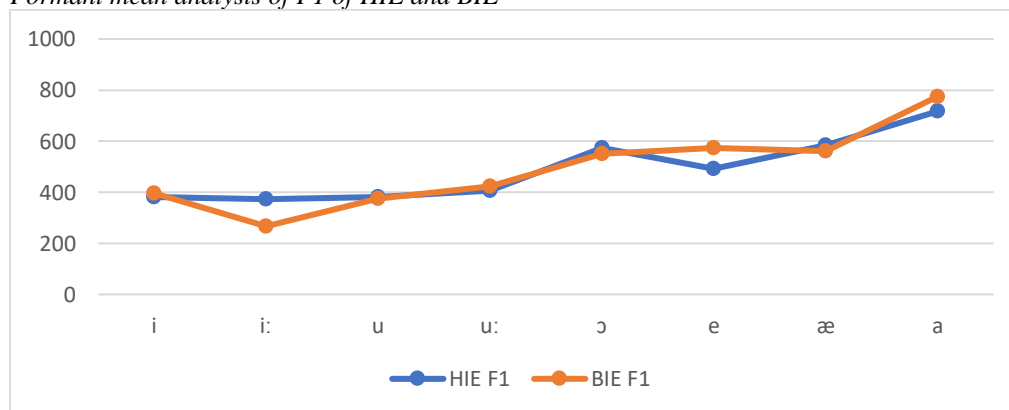
The average of F2 of HIE and BIE has been shown in Figure 2. In Figure 2, the vowel /i:/ of BIE has the highest F2, and the vowel /u/ has the lowest F2. Again, F2 of vowels like /i/, /u/, /ɔ/, /æ/, and /a/ of both varieties have a minute difference. However, BIE has the highest F2.

Figure 2.
*Formant mean analysis of F2 of HIE and BIE*

The average of F3 formants of HIE and BIE has been shown in Figure 3. Figure 3 shows, that BIE has the highest F3 values for /i:/ while HIE has the lowest F3 for vowel /a/. The vowel /u/ of HIE is more rounded than other rounded vowels like /u:/ and /ɔ/. HIE rounded vowels are more rounded than BIE rounded vowels. Although, the formant value F3 of Hindi and Bangla speakers of /ɔ/ is very close to each other, which implies that /ɔ/ is equally rounded for both dialects.

It can be deduced from Figure. 1, 2, and 3 that Bangla speakers have the highest F1, F2 and F3 values. BIE speakers have the highest and also lowest F1 and F2 values. Again, BIE has the lowest F1 and F2 values, whereas HIE has the lowest F3 values.

Figure 3.
Formant Mean Analysis of F3 of HIE and BIE
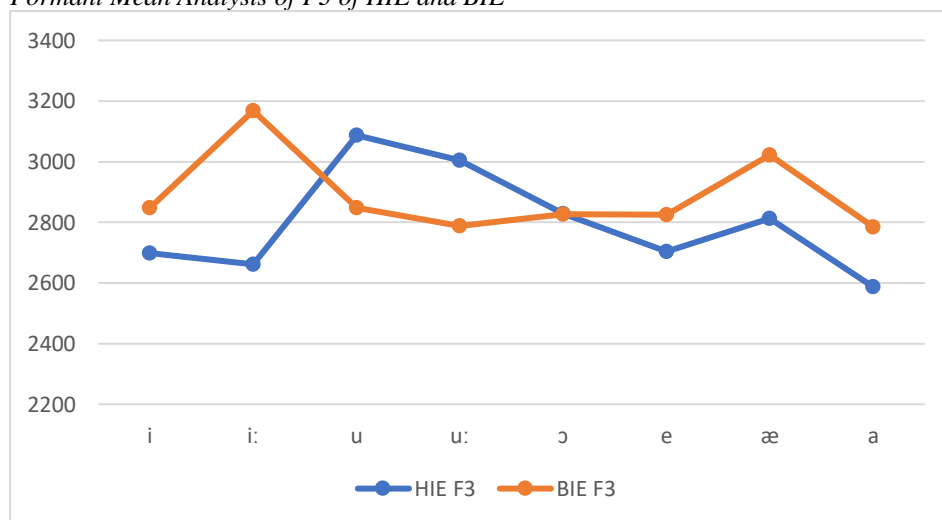
Vowel Space Analysis
A two-dimensional graph is used to display the main parameters of vowel production, where the two dimensions denote tongue advancement and height. F1 vs. F2 representation can be efficiently used to analyze variations in age, gender, and accent variability. To determine the vowel space, the measured values of F1 and F2 are plotted in a vowel quadrilateral with F1 plotted on the x-axis and F2 plotted on the y-axis.

Figure 4.
F1 vs F2 Formant space analysis for vowels of HIE and BIE
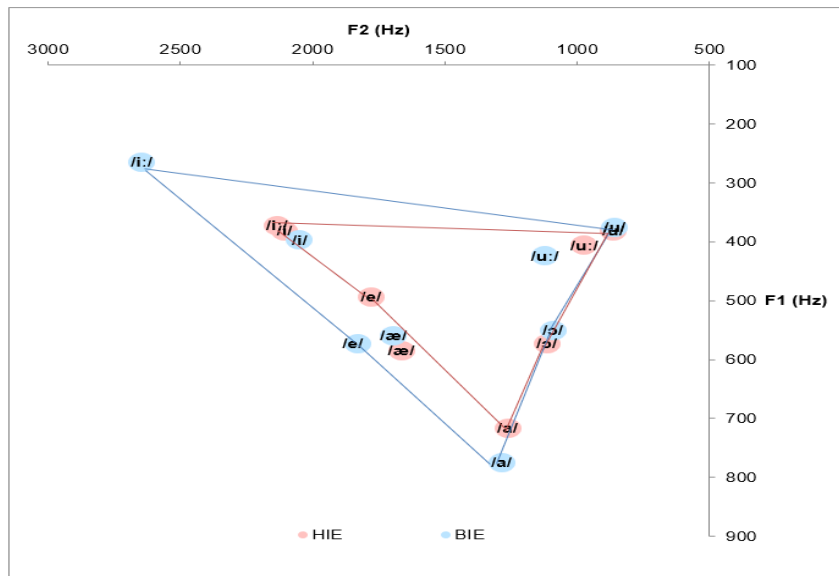
F1 vs. F2 formants space of English spoken by Hindi and Bangla L1 speakers is shown in Figure 4. When compared with one another, both Indian English accents exhibit variation. The vowel /iː/ of Indian English spoken by Bangla speakers is more frontal and raising as compared to that of Indian English spoken by Hindi speakers. Again, in both varieties of Indian English, the vowel /u/ is raising and more towards the back. Vowel /e/ of HIE is higher and more towards the center while vowel /e/ of BIE is more towards the front. Vowels /iː/ and /u/ of BIE are more closed as compared to HIE. The vowel /a/ of BIE is more open than that of HIE. Vowels like /ɔ/, /æ/, and /i/ of both varieties of Indian English share almost the same vowel space, i.e., their formant values are close to each other. Vowels /i/ and /iː/ of HIE have similar formant space. It can be noticed that both varieties of Indian English have different vowel patterns. It is evident that, from the analysis of formant frequencies, a comprehensible variation in vowel space can be portrayed.

Although Bangla vowels are not marked for length, BIE speakers tend to produce both long and short high vowels. This very phenomenon might be possible because Bangla orthography has the provision of short and long vowel symbols (Barman, 2009). Bangla speakers can distinguish long and short vowels in orthography, thus, they tend to produce both long and short high vowels. Therefore, it can be said that though Bangla lacks length-contrastive vowel sounds, it does have length-contrastive vowel symbols. On the other hand, Hindi is marked for both length and quality. Even then, data demonstrate that HIE speakers tend to produce short vowels. This might be possible because Hindi has allophonic variations.

The above analysis of formant frequencies shows significant variations in the varieties of Indian English, which is due to the native language influence. Several researchers have employed formants for the identification of dialects (Adank et al., 2007; Yan & Vaseghi, 2003). Olagbaju et al., (2010) in their study calculated the mean and standard deviation of the first three formants of vowels produced by Mandarin and Hindi speakers. Their findings highlighted that analysis of formant frequencies can contribute to a richer understanding of accent variation. By analyzing acoustic characteristics like formant frequencies and duration in regional dialects of Hindi, Sinha et al. (2019) investigated the influence of dialectal variation. Their findings also suggest that with the help of acoustic analysis, it can be postulated as the substance for differentiating and identifying speakers from different dialects.

Vowel Space Area

The Vowel Space Area (VSA) is generally determined by calculating the Euclidean distances (EDs) between the first and second formant coordinates of the corner vowels /i/, /u/, and /a/ in triangular VSA on the F1– F2 plane (e.g., Blomgren et al., 1998; Liu et al., 2005).

The vowels chosen for the calculation of extended areas of vowel space for HIE and BIE were /iː, a, and u/. The majority of the vowel space used by the speakers across both dialects was encompassed by these vowels. The vowel areas of the /iː, a, u/ triangle were calculated using the F1 and F2 values of the vowels /iː, a, u/ of each speaker.

VSA is mathematically stated as follows (adopted from Blomgren et al., 1998):

$$VSA = \sqrt{(S(S - EDiu)(S - EDia)(S - EDau))}$$

Where, $S = \frac{(EDiu + EDia + EDau)}{2}$

Where, $EDiu = \sqrt{((F1i - F1u)^2 + (F2i - F2u)^2)}$

$EDia = \sqrt{((F1i - F1a)^2 + (F2i - F2a)^2)}$

$EDau = \sqrt{((F1a - F1u)^2 + (F2a - F2u)^2)}$

Again, F1i is the first formant of the vowel /i:/, F2i is the second formant of the vowel /i:/, and so forth. EDiu is the Euclidean distance between the vowels /i:/ and /u/, EDia is the Euclidean distance between the vowels /i:/ and /a/, and EDau is the Euclidean distance between the vowels /a/ and /u/. Thus, the EDs: EDiu, EDia, and EDau define the triangular VSA with the vowels /i:/, /u/, and /a/.

Table 1
The Edi:u, Edi:a, Edau, and VSA values for HIE and BIE

		mean	Standard deviation	df	t	p
Edi:u	BIE	1792.400	412.329			
				8	2.573	0.033
	HIE	1270.600	188.805			
Edi:a	BIE	1456.800	352.294			
				8	2.735	0.026
	HIE	946.000	224.391			
Edau	BIE	588.200	120.290			
				8	0.400	0.699
	HIE	541.000	234.483			
VSA	BIE	389448.600	168943.685			
				8	2.002	0.080
	HIE	216319.200	94152.283			

Table 1 shows the mean values, standard deviation, df (degree of freedom), t value, and p-value of Edi:u, Edi:a, Edau, and VSA among BIE and HIE speakers. We compared the VSA size, Edi:u, Edi:a, and EDau of the speakers of BIE and HIE. Table 1 shows that the mean values of VSA, Edi:u, Edi:a, and EDau for the BIE speakers were slightly higher than that for HIE speakers. To test if the differences in mean values of VSA, Edi:u, Edi:a, and EDau among BIE and HIE speakers were significant, a t-test was used. The t-value was found to be significant for Edi:u (t value=2.573, p<0.033) and Edi:a at 0.05 level (t value=2.735, p<0.026), which confirms that Edi:u and Edi:a differ significantly among BIE and HIE speakers. The calculated mean of the vowel space area shows that the shape of the vowel spaces varies across dialects. Also, the mean values demonstrate that the vowels produced by BIE speakers occupy a greater vowel space area as compared to the HIE vowel space area, though there was no significant difference between them. The mean difference in vowel space area is likely due to the far front position of the /i:/ vowel of BIE. The positional difference across dialects can also be found for the vowel /a/, suggesting regional variation in the vowel.

Vowel Normalization

Vowel normalization is a technique used to eliminate variation due to physiological differences among speakers. No two speakers have the same dimensions of vowel tracts; thus, the same vowel produced by different speakers shows different formant frequencies. Vowel normalization preserves dialectal variation in vowel quality and phonological differences between vowels. It is used to model the cognitive process of listeners to normalize vowels produced by various speakers (Thomas, 2002).

The Bark Difference method was selected for the normalization procedure. It is amongst the most viable approaches for investigation in dialect identification. The Bark difference method normalizes the values of F1 and F2 through a series of mathematical conversions using the F1, F2, and F3 values (Thomas, 2011). Using the formula, NORM transforms the formant values to Bark.

$$Z_i = 26.81 \times \frac{Fi}{1960 + Fi} - 0.53$$

Where F_i is the value for particular formant *i*, it then calculates the difference between Z3-Z1, Z3-Z2, and Z2-Z1. Z3-Z2 is used to plot the normalized front-back dimension, and Z3-Z1 is used to plot the normalized height dimension.

The formant values were converted to z-scores using normalization procedures to investigate whether the significance of gender-related variances resulting from anatomical and/or physiological factors can be eliminated. Normalization measures were executed using the NORM suite.

Table 2.
Mean Vowel Formant Values Bark Difference Normalized

Speaker	Vowel	Z3-Z1	Z3-Z2	Z2-Z1
1hie	/i/	11.166	1.643	9.523
2hie	/i:/	11.153	1.461	9.692
3hie	/u/	12.025	8.196	3.829
4hie	/u:/	11.625	7.332	4.293
5hie	/e/	10.153	2.785	7.368
6hie	/æ/	9.638	3.487	6.151
7hie	/ ɔ/	9.770	6.114	3.656
8hie	/a/	8.072	4.747	3.325
9bie	/i/	11.365	2.178	9.187
10bie	/i:/	13.381	1.162	12.219
11bie	/u/	11.566	7.718	3.848
12bie	/u:/	10.975	5.983	4.992
13bie	/e/	9.763	2.886	6.877
14bie	/æ/	10.305	3.834	6.471
15bie	/ ɔ/	9.958	6.240	3.718
16bie	/a/	7.536	4.511	3.025

According to earlier explanations, the high and low quality of vowels are inversely correlated with F1 or Z3-Z1. Similarly, the front and back qualities of vowels are directly correlated with F2 or Z3-Z2. From Table 2, it can be deduced that the average Z3-Z1 values for HIE vowels, /i/ and /i:/, appear to be almost equal. The average Z3-Z2 values seem to be very close to each another, suggesting that the speaker's front part of the tongue is raised to nearly the same height while producing these vowels. The HIE vowels /i/ and /i:/ has similar vowel height and frontness. It can be said that HIE speakers produce these vowels from such a position that in English inventory it could be considered a halfway position between /i/ and /i:/. The above Z3-Z1 and Z3-Z2 values reveal that vowels /i/ and /i:/ are non-contrastive in terms of vowel height and frontness.

The average Z3-Z1 values for HIE vowels /u/ and /u:/ have a difference between them. The Z3-Z1 values show that the vowel /u/ is relatively higher than the vowel /u:/ in vowel space. Their mean values of Z3-Z2 also seem to be different. This suggests that the vowel /u:/ is more towards the front as compared to the vowel /u/, while in the English vowel inventory, the vowel /u/ is more central and fronted as compared to /u:/. The findings of Deterding (2003) suggest that the vowel /u:/ is placed further back in Singapore English. Cruttenden (2001) also showed that British English accepts the fronted variety of vowel /u:/ and even Bradford English has fronted /u:/ (Watt & Tillotson,

2001). The above Z3-Z1 and Z3-Z2 values show that vowels /u/ and /u:/ are contrastive in terms of vowel height and frontness.

While in BIE the average value of Z3-Z1 and Z3-Z2 for vowels /i/ and /i:/ and /u/ and /u:/ have a major difference between them. The Z3-Z1 values of vowels /i:/ and /u/are comparatively higher than vowels /i/ and /u:/. The value of Z3-Z1 of vowel /i:/ is greater than the value of vowel /i/. This suggests that the vowel /i:/ is higher than the vowel /i/. Again, the value of Z3-Z2 of vowel /i/ is greater than the value of vowel /i:/. This proposes that the vowel /i:/ is more towards the front as compared to the vowel /i/.

Similarly, the average value of Z3-Z1 of the BIE vowel /u/ is higher than the value of /u:/. This shows that the vowel /u/ is comparatively higher than the vowel /u:/. Again, the Z3-Z2 value of the vowel /u/ is greater as compared to the value of the vowel /u:/. Further, suggesting that the vowel /u:/ is more fronted and central as compared to the vowel /u/.

The Indian English vowel produced by native speakers of Hindi, Bangla, and RP British English appears in Figure 5. Each vowel denotes the mean F1 and F2 of HIE and BIE taken at the temporal midpoint. The formant values of Received Pronunciation British English (RP English) were taken from the study of Wells (1962).

The front vowel /i:/ of BIE is more frontal than HIE and proves to be more closure to RP English. Whereas, the vowel /i/ of both varieties have modest differences from RP English. The BIE speakers have acquired vowel /i/ and /i:/ distinctions, while the front vowel of HIE has similarities. Vowels /e/ and /æ/ of both varieties are more toward the center as they differ substantially from RP English, where/æ/ is a considerably lower vowel. Vowel /a/ of HIE has a modest difference from RP English while in BIE it appears to be more open and lower. While vowel /u:/ of BIE is more similar to RP English and in HIE, it is more towards the back. The back vowel /u/ of both varieties differ markedly from RP English in being further back. In HIE and BIE, the vowel /u/ is backed, while RP English fronts this vowel. Vowel /ɔ/ in both varieties appears to be more towards the center as compared to RP English demonstrating significant differences, and suggesting a lack of open mid-back vowels in both varieties. HIE and BIE vowel /ɔ/ have frontness, while in RP English, this vowel backs further.

The first three formants produced by different speakers remove physiological differences. Overall, the outcomes of mean normalized vowel space exhibited substantial dialectal variation in the vowel space area. These variations were existing even after normalizing the speaker's gender effects. Although, dialects may exhibit some inter-speaker variability.

Figure 5

The Vowel Spaces of HIE (1hie in red color), BIE (2bie in red color), and RP English (3rp vowel in green color) are plotted in Barks. (note: i = /i/, I = /i:/, e = /e/, A = /æ/, a = /a/, o = /ɔ/, u = /u/, and U = /u:/).
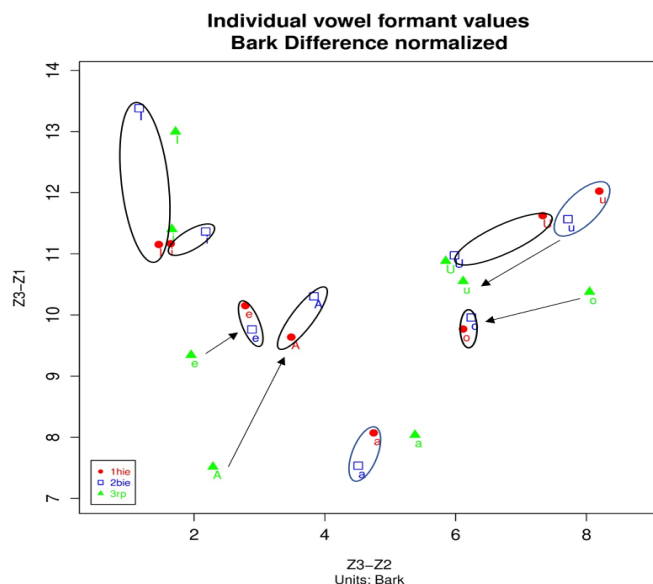
Voice Onset Time

Using PRAAT, VOT was assessed. A sudden amplitude difference on the wave marked the point of the stop consonant release. The beginning of voicing was marked by the beginning of periodicity. VOT was measured as the elapsed time between the boundaries located at the places of stop consonant release and voice onset was displayed.

Figure 6 shows the VOT of /p/, /t/, and /k/ among HIE and BIE speakers. The finding of this study indicates that /p/ is aspirated in BIE, while in HIE /k/ is more aspirated. Stop consonant /t/ has a lower mean VOT in the speech pattern of both HIE and BIE speakers. Stop consonants with shorter VOTs tend to have a weaker release of airflow because voicing starts so quickly after the stop release. This concludes that there is no or less aspiration or reduction of aspiration. Therefore, we can say that this study indicates that /t/ is unaspirated or less aspirated in the Indian English produced by Bangla and Hindi speakers.

Figure 6.

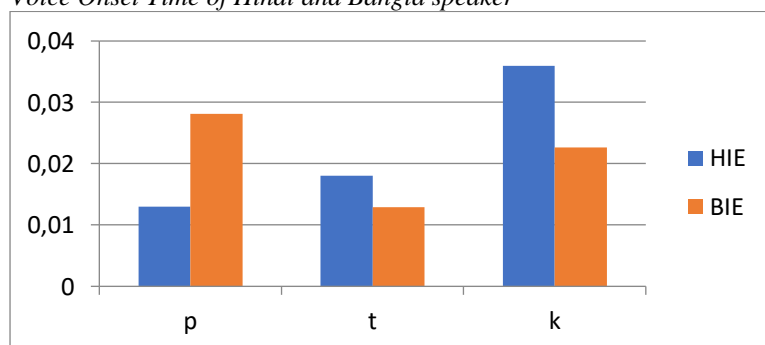*Voice Onset Time of Hindi and Bangla speaker*

Table 3 shows the mean VOT of /p/, /t/, and /k/ among HIE and BIE speakers. It can be deduced from Table 3 that voiced stops /p/, /t/, and /k/ are produced with a short lag in HIE and BIE. HIE speakers produce /t/ and /k/ with average VOT values of 18ms for /t/ and 35ms for /k/, which is substantially different from BIE speakers. BIE speakers, by contrast, produce VOTs of /t/ and /k/ in the short lag range, with values even shorter than those of HIE speakers. However, /p/ produced by BIE speakers has a higher VOT as compared to HIE speakers. Awan and Stine (2011) in their study have found that Indian speakers tend to produce lower mean VOTs. Therefore, short VOTs produced by HIE and BIE speakers support the previous studies. According to Malhotra and Vogelaar (2004), aspirated Hindi stop consonants are more highly aspirated than their English equivalents, so speakers usually tend to use the non-aspirated form of the phoneme instead. Das and Hansen (2004) suggested that Indian English speakers have little exposure to other varieties of English, which might lead to confusion as Indo-Aryan languages have six variations of the stop consonants.

Table 3.
Mean VOT of HIE and BIE

	/p/	/t/	/k/
HIE	12ms	18ms	35ms
BIE	28ms	12ms	22ms

CONCLUSION

Indian English produced by Bangla and Hindi native speakers appears to have vowel patterns that suggest they could be considered a separate sub-variety of Indian English. The preliminary acoustic analysis shown here suggests that dialect profiling can be done using formant analysis and by calculating VOT. This study confirms that specific features of the vowels of Indian English may be distinctive for speakers with a variety of native language backgrounds in India. The present study indicates that the difference in the mean value of vowel space area due to the variation in the formant frequency of vowels is reflected in the dialect variation. This study illustrates that native language influence detection can be understood by acoustic analysis. A detailed analysis of the acoustic features of BIE and HIE varieties displays marked differences that can be used to derive the speaker's geographical regions. This study has many implications in the field of forensic phonetics and language analysis for the determination of origin. However, a larger data set would be more helpful to establish more robust acoustic features to distinguish these two varieties.

This paper has attempted to investigate a vital area in forensic linguistics using NLID and PRAAT in an endeavor to equip law enforcement agencies for better and more efficient crime control. The cases that involve linguistic evidence reflect the socio-political climate of the state to a great extent. Majorities of language crimes are either through computer-mediated communication systems or a clear manifestation of the complexities of multilingualism. Both these factors make the perpetrators obscure. Therefore, forensic linguistic examinations and research provide a huge scope and relevance in the current age of evolving language-related crimes. However, it could be erroneous to assume that the study presented here provides an acoustic-phonetic cue exhaustively. It is definitely a building block that can lead to more exhaustive research that can provide a more robust framework with more verifiable acoustic cues.

REFERENCES

Adank, P. (2003). *Vowel normalization: a perceptual-acoustic study of Dutch vowels* [Doctoral thesis, Catholic University of Nijmegen). Radbound Repository. https://hdl.handle.net/2066/19286

Adank, P., Van Hout, R., & Velde, H. V. D. (2007). An acoustic description of the vowels of northern and southern standard Dutch II: Regional varieties. *The Journal of the Acoustical Society of America*, *121*(2), 1130-1141. https://doi.org/10.1121/1.2409492

Ainsworth, P. (2001). *Offender profiling and crime analysis* (1st ed.). Willan. https://doi.org/10.4324/9781843924630

Awan, S. N., & Stine, C. L. (2011). Voice onset time in Indian English-accented speech. *Clinical Linguistics & Phonetics*, *25*(11-12), 998-1003. https://doi.org/10.3109/02699206.2011.619296

Barman, B. (2009). A contrastive analysis of English and Bangla phonemics. *Dhaka University Journal of Linguistics*, *2*(4), 19-42. https://doi.org/10.3329/dujl.v2i4.6898

Blomgren, M., Robb, M., & Chen, Y. (1998). A note on vowel centralization in stuttering and nonstuttering individuals. *Journal of Speech, Language, and Hearing Research*, *41*(5), 1042-1051. https://doi.org/10.1044/jslhr.4105.1042

Boersma, P., & Weenink, D. (2008). *Praat: Doing phonetics by computer* (Version 5.0.25). Praat. http://www.praat.org/

Borden, G., Harris, K., & Raphael, L. (2007). *Speech science primer: Physiology, acoustics, and perception of speech* (5th ed.). Lippincott Williams & Wilkins.

Chatterji, S. K. (2002). *The origin and development of the Bengali language*. Rupa.

Chaudhury, S.B.R., & Samaddar, R. (Eds.). (2018). *The Rohingya in South Asia: People without a state*. Routledge. https://doi.org/10.4324/9780429467677

Chifflet, P. (2015). Questioning the validity of criminal profiling: An evidence-based approach. *Australian & New Zealand Journal of Criminology*, *48*(2), 238-255. https://doi.org/10.1177/0004865814530732

Clopper, C. G., & Paolillo, J. C. (2006). North American English vowels: A factor-analytic perspective. *Literary and linguistic computing*, *21*(4), 445-462. https://doi.org/10.1093/llc/fql039

Cooley, C. M. (2012). Criminal profiling on trial: The admissibility of criminal profiling evidence. In B. E. Turvey (Eds.), *Criminal profiling: An introduction to behavioral evidence analysis* (pp. 628-652). Elsevier.

Cruttenden, A. (2001). *Gimson's pronunciation of English* (6th ed.). Oxford University Press.

Das, S., & Hansen, J. H. (2004). Detection of voice onset time (VOT) for unvoiced stops (/p/, /t/, /k/) using Teager Energy Operator (TEO) for automatic detection of accented English. In J. Tanskanen (Eds.), *Proceedings of the 6th Nordic Signal Processing Symposium, 2004.* (pp. 344–347). IEEE.

Dasgupta, P. (2003). Bangla. In G. Cardona & D. Jains (Eds.), *The Indo-Aryan languages*, (pp. 351-390). Routledge.

Datta, A. K. (2018). *Acoustics of Bangla speech sounds*. Springer. https://doi.org/10.1007/978-981-10-4262-1

Deterding, D. (2003). An instrumental study of the monophthong vowels of Singapore English. *English World-Wide*, *24*(1), 1-16. https://doi.org/10.1075/eww.24.1.02det

Foulkes, P., & French, P. (2001). Forensic phonetics and sociolinguistics. In R. Mesthrie (Ed.), *The concise encyclopedia of sociolinguistics.* Elsevier.

Foulkes, P., French, P., & Wilson, K. (2019). LADO as forensic speaker profiling. In P. Patrick, M. Schmid, & K. Wilson (Eds.), *Language analysis for the determination of origin*, (pp. 91-116). Springer.

French, P., & Harrison, P. (2006). Investigative and evidential applications of forensic speech science. In A. Heaton- Armstrong, E. Shepherd, G. Gudjonsson & D. Wolchover (Eds.), *Witness testimony: Psychological, investigative and evidential perspectives*, (pp. 247-262). Oxford University Press.

French, P., & Stevens, L. (2013). Forensic speech science. In M. Jones & R. Knight (Eds.), *Bloomsbury companion to phonetics*, (pp. 183-197). Bloomsbury Academic. http://doi.org/10.5040/9781472541895.ch-012

Geberth, V. J. (2015). *Practical homicide investigation: Tactics, procedures, and forensic techniques*. CRC Press. https://doi.org/10.4324/9781003095835

Grant, T. (2008). Approaching questions in forensic authorship analysis. In J. Gibbons & M. Turell (Eds.), *Dimensions of forensic linguistics* (pp. 215-229). John Benjamins. https://doi.org/10.1075/aals.5.15gra

Jessen, M. (2007). Speaker classification in forensic phonetics and acoustics. In C. Muller (Ed.), *Speaker classification I* (pp. 180-204). Springer. https://doi.org/10.1007/978-3-540-74200-5_10

Jessen, M. (2008). Forensic phonetics. *Language and linguistics compass*, *2*(4), 671-711. https://doi.org/10.1111/j.1749-818X.2008.00066.x

Kalashnik, O. & Fletcher, J. (2007). An acoustic study of vowel contrasts in North Indian English. In J. Trouvain & W.J. Barry (Eds.), *Proceedings of the 16th international congress of phonetic sciences* (pp. 953–956).

Kent, R. D., & Read, C. (2002). *The acoustic analysis of speech* (2nd ed.). Singular.

Kulshreshtha, M., Singh, C. P., & Sharma R. M. (2012). Speaker profiling: The study of acoustic characteristics based on phonetic features of Hindi dialects for forensic speaker identification. In H. Patil, & A. Neustein (Eds.), *Forensic speaker recognition* (pp. 71-100). Springer. https://doi.org/10.1007/978-1-4614-0263-3_4

Labov, W., Ash, S., & Boberg, C. (2005). *Atlas of North American English: Phonetics, phonology and sound change*. De Gruyter Mouton. https://doi.org/10.1515/9783110167467

Ladefoged, P., & Johnson, K. (2011). *A course in phonetics* (6th ed.). Cengage Learning.

Liu, H. M., Tsao, F. M., & Kuhl, P. K. (2005). The effect of reduced vowel working space on speech intelligibility in Mandarin-speaking young adults with cerebral palsy. *The Journal of the Acoustical Society of America*, *117*(6), 3879-3889. https://doi.org/10.1121/1.1898623

Malhotra, H. K., & Vogelaar, R. (2004). Accent reduction for Asian Indians. *Advance for Speech-Language Pathologists and Audiologists*, *14*(49), 10-23.

Malmasi, S., & Dras, M. (2017). Multilingual native language identification. *Natural Language Engineering*, *23*(2), 163-215. https://doi.org/10.1017/S1351324915000406

Maxwell, O. & Fletcher, J. (2009) Acoustic and durational properties of Indian English vowels. *World Englishes, 28*(1), 52–69. https://doi.org/10.1111/j.1467-971X.2008.01569.x

Mostafa, T. (2010). Problems bangladeshi learners face in pronouncing certain English phonemes. *BRAC University Journal,* Special Issue 01, 130–138. http://hdl.handle.net/10361/5165

National Crime Records Bureau. (2021). Crimes in India. https://ncrb.gov.in/sites/default/files/CII-2021/CII_2021Volume%201.pdf

Olagbaju, Y., Barkana, B. D., & Gupta, N. (2010, April). English vowel production by native Mandarin and Hindi speakers. In S. Latifi (Ed.), *Proceedings of the Seventh International Conference on Information Technology: New Generations* (pp. 343-347). Conference Publishing Service.

Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research*, *5*(3), 261-283. https://doi.org/10.1007/BF01067377

Perkins, R., & Grant, T. (2018). Native language influence detection for forensic authorship analysis: Identifying L1 Persian

bloggers. *International Journal of Speech Language and the Law*, *25*(1), 1-20. https://doi.org/10.1558/ijsll.30844

Phull, D. K. & Kumar, G. B. (2016). Vowel analysis for Indian English. *Procedia Computer Science 93,* 533-538. https://doi.org/10.1016/j.procs.2016.07.264

Rahman, A. M. (2008). A comparison between English and Bangla vowel systems. *Khulna University Studies*, *9*(1), 9-16. https://doi.org/10.53808/KUS.2008.9.1.0829-A

Rose, P. (2002). *Forensic speaker identification*. Taylor & Francis. https://doi.org/10.1201/9780203166369

Schneider, E. W. (2007). *Postcolonial English: Varieties around the world*. Cambridge University Press.

Schuller, B., Steidl, S., Batliner, A., Hirschberg, J., Burgoon, J.K., Baird, A., Elkins, A.C., Zhang, Y., Coutinho, E., & Evanini, K. (2016). The INTERSPEECH 2016 computational paralinguistics challenge: Deception, sincerity & native language. *Proceedings of the Annual Conference of the International Speech Communication Association* (pp. 2001-2005). https://doi.org/10.21437/Interspeech.2016-129

Shamshad, R. (2017). *Bangladeshi migrants in India: Foreigners, refugees, or infiltrators?* Oxford University Press.

Sharf, G., & Masur, H. (2002). Voice onset time in normal speakers of a German dialect: Effects of age, gender, and verbal material. In F. Windsor, M. L. Kelly, & N. Hewlett (Eds.), *Investigations in clinical phonetics and linguistics.* Psychology Press.

Sharma, D. (2017). English in India. In A. Bergs & L. Brinton (Eds.), *Volume 5 Varieties of English* (pp. 311-329). De Gruyter Mouton. https://doi.org/10.1515/9783110525045-016

Sinha, S., Jain, A., & Agrawal, S. S. (2019). Empirical analysis of linguistic and paralinguistic information for automatic dialect classification. *Artificial Intelligence Review*, *51*(4), 647-672. https://doi.org/10.1007/s10462-017-9573-3

Tiwari, B. (1966). *Hindi bhasha*. Kitab Mahal.

Thomas, E. R. (2002). Instrumental phonetics. In J. K. Chambers, Peter Trudgill, and Natalie Schilling-Estes (Eds.), *The handbook of language variation and change* (pp. 168-200). Oxford.

Thomas, E. R. (2011). *Sociophonetics: An introduction.* Palgrave Macmillan.

Thomason, S. G. (2001). *Language contact*. Edinburgh University Press.

Watt, D., & Tillotson, J. (2001). A spectrographic analysis of vowel fronting in Bradford English. *English World-Wide, 22*(2), 269–303. https://doi.org/10.1075/eww.22.2.05wat

Wells, J. C. (1962). *A study of the formants of the pure vowels of British English* [Unpublished Doctoral dissertation]. University of London.

Wells, J. C. (1982). *Accents of English: Volume 1*. Cambridge University Press.

Wiltshire, C. R., & Harnsberger, J. D. (2006). The influence of Gujarati and Tamil L1s on Indian English: A preliminary study. *World Englishes*, *25*(1), 91-104. http://doi.org/10.1111/j.0083-2919.2006.00448.x

Winerman, L. (2004, July). Psychological sleuths--criminal profiling: The reality behind the myth. *Monitor on Psychology*, *35*(7). https://www.apa.org/monitor/julaug04/criminal

Yan, Q., & Vaseghi, S. (2003). Analysis, modeling, and synthesis of formants of British, American, and Australian accents. *Proceedings of the IEEE international conference on acoustics, speech, and signal processing* (pp. 712-715).

Yavas, M. (2002). Voice onset time patterns in bilingual phonological development. In F. Windsor, M. L. Kelly, & N. Hewlett (Eds.), *Investigations in clinical phonetics and linguistics* (pp. 327–339). Mahwah, Erlbaum. https://doi.org/10.4324/9781410613158

Zampieri, M., Malmasi, S., Ljubešić, N., Nakov, P., Ali, A., Tiedemann, J., Scherrer, Y., & Aepli, N. (2017). Findings of the VarDial evaluation campaign 2017. *Proceedings of the Fourth Workshop on NLP for Similar Languages, Varieties, and Dialects*, (pp. 1-15). Association for Computational Linguistics.