2021

The Effect of Bilingual Proficiency in Indian English on Bilabial Plosive

Taniya Chawla

West Virginia University, tc0058@mix.wvu.edu

The Effect of Bilingual Proficiency in Indian English on Bilabial Plosive

Taniya Chawla

Dissertation submitted to the College of Education and Human Services at West Virginia University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in

Communication Sciences and Disorders

Committee:

Kimberly M. Meigh, Ph.D., CCC-SLP, Chair

Melissa Luna, Ph.D., Co-Chair

Michelle W. Moore, Ph.D., CCC-SLP

Jonah Katz, Ph.D.

Department of Communication Sciences and Disorders

Morgantown, West Virginia

2021

Abstract

The Effect of Bilingual Proficiency in Indian English on Bilabial Plosive

Taniya Chawla, MASLP

Background: Bilingual speech production studies have highlighted that level of proficiency influences the acoustic-phonetic representation of phonemes in both languages (MacKay, Flege, Piske, & Schirru 2001; Zárate-Sández, 2015). The results for bilingual speech production reveals that proficint/early bilinguals produce distinct acoustic properties for the same phoneme in each language, whereas less proficient/late bilinguals produce acoustic properties for a phoneme that is closer to the native language (Flege et al., 2003; Fowler et al., 2008). Acoustic-phonetic studies for Hindi (L1) and Indian English (L2) for bilingual speakers have been understudied, and the level of proficiency has not been considered in Hindi and Indian English bilingual speakers. The present study aimed to measure the acoustic differences produced by bilingual speakers of varying proficiencies for Indian English on bilabial plosive and determine how the bilabial plosives are different from American English bilabial plosives.

Methods: The sample size for this study was twenty-four. However, only twenty participants (eleven females) between the ages of eighteen and fifty, with normal speech and hearing, were recruited. The lack of recruitment of four more participant was due to the inability to find bilingual speakers who spoke Hindi as their first language and Indian English as their second language and COVID-19 restrictions imposed on recruitment (n=4). The participants were divided into three groups based on language and proficiency: a monolingual American English group, a proficient bilingual Hindi-Indian English group, and a less-proficient bilingual Hindi-Indian English group. The bilinguals were divided into a proficient and less proficient group based on the Language Experience and Proficiency Questionnaire (Marian, Blumenfeld, & Kaushanskaya, 2007). Following the screening, participants took part in a Nonword Repetition Task. Data were analyzed using Praat and Voice Sauce software. A linear mixed-effects model using R statistics was used for the statistical analysis.

Results: Data from 20 participants (seven proficient bilingual speakers, five less-proficient bilingual speakers, and eight monolingual speakers) were included in the data analysis. Approximately four thousand repetitions were evaluated across the remaining participants. There were no significant main effects across the four dependent variables, but there was an interaction effect between group and phoneme on two dependent variables. The closure duration for proficient bilingual speakers compared to less-proficient bilingual speakers were significantly different between the voiceless unaspirated bilabial plosive (VLE) and voiceless aspirated bilabial plosive (VLH), as well as voiced unaspirated bilabial plosive (VE) and voiced aspirated bilabial plosive

(VH). For spectral tilt, there was a significant difference between the VLE and VLH for proficient bilingual speakers compared to less proficient bilingual speakers.

Discussion: The results of this study suggest that proficient bilingual speakers have faster rate of speech in both their first and second language. Therefore, it is difficult to provide information on whether this group has separate acoustic-phonetic characteristics for each phoneme for each language. In contrast, the less-proficient bilingual speakers seem to have a unidirectional relationship (i.e., first language influences the second language). Furthermore, the results of the acoustic characteristics for the control group i.e., monolingual American English speakers suggest that they may have acoustic-phonetic characteristics that represent a single acoustic-phonetic representation of bilabial plosive with their voicing contrast.

Acknowledgment

The task of completing this dissertation would not have been possible without the support and motivation of my mentor. I am indebted and grateful to you, Dr. Kimberly Meigh, for taking a chance on me and helping me navigate these four years in the Department and USA. You have molded me into a better researcher and a teacher-investigator. Thank you for having my back and being a friend in a new country.

Dr. Melissa Luna, Dr. Michelle Moore, and Dr. Jonah Katz thank you for being a part of my dissertation committee! I want to thank you all for challenging me and guiding me as I worked on my dissertation. Thank you for all your insights and feedback throughout the last year. I am grateful to all of you who put in your time to help me in this process.

Dr. Nair and Dr. Donai, thank you for your guidance and suggestions during my comps and prospectus! CSD faculty and staff members, thank you all for your support and encouragement during the past four years!

Anuradha, thank you for allowing me to recruit and collect data at the ISKON temple at Moundsville, WV.

Thank you to all my research participants! Recruiting participants is difficult during any time of the year, but COVID-19 made it much more difficult. But all my participants were so kind and understanding that collecting data was fun even with all the COVID-19 protocol.

I must thank all the lab assistants from the speech motor control lab. You guys have helped me reach the finish line. I especially have to thank Brennan for lending your voice to be part of the stimulus, Lacey, Darian, and Olivia for your hustle and support in completing all the text grid files.

Mom and Tarini -You both have always made time for Facetime or WhatsApp calls, even with the time difference. This endeavor would not have been possible without your unconditional love and support throughout the years. You are indeed my pillars of strength. Thank you for letting me take this leap in my professional life. I am forever indebted to you both!

To my family in Morgantown, Jimmy, thank you for being my biggest cheerleader and my husband! Being an international student comes with some challenges along the way, but you have helped me and supported me in the most crucial period of my education. Thank you for your constant encouragement! JoDonna and Debbie, I could not have asked for more supporting in-laws. You guys are my home away from home. Thank you for all your love and support!

Loveleen, Hansneet and Allegra, thank you for taking out time and hanging out with me! I will cherish these moments forever.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

India is known for its linguistic diversity, and bilingualism is commonplace. Hindi, with English, serves as a connecting language across the multilingual country and the world. Indian English (IE) is one of the most widely used postcolonial varieties of world English (Wells, 1982). Because of the increasing rate of emigration from the subcontinent, IE is spoken by the Indian diaspora in different parts of the world, such as Canada, Australia, and the United States (Maxwell, & Fletcher, 2009).

Proficiency and age of the second language (L2) acquisition modulate the organization of the two languages in the brain (Grosjean, 2008). For example, while proficient early bilinguals tend to produce the same phoneme with a distinct acoustic realization for both languages, this is not the case for less-proficient bilingual speakers; they may show acoustic realization for a phoneme that is closer to the native language (Flege et al., 2003; Fowler et al., 2008). Evaluating the acoustic differences in Hindi/Indian English (IE) speakers who vary in their proficiency may provide insight into models of bilingualism. Although Indian English has been an interest of study, limited research on the acoustic realization of Indian English speech sounds exists.

Studies have investigated the acoustic-phonetic characteristics of IE vowels in speakers whose first language (L1) was Punjabi and Hindi (Maxwell, & Fletcher, 2009). Moreover, several studies have directly compared the effects of different native languages on specific phonological characteristics of IE. A common theme emphasizes and supports the notion that speakers from different language backgrounds produce a similar variety of Indian English (Maxwell & Fletcher, 2009; Pickering & Wiltshire, 2000; Wiltshire & Moon, 2003; Wiltshire & Harnsberger, 2006). However, many of these studies have evaluated vowels, therefore are limited studies observing consonant context. Hence, there is a need to understand the differences and similarities in the acoustic-phonetic characteristics of Hindi and Indian English (IE) phonemes, especially consonants.

For the present study, the author aimed to measure the acoustic characteristics of Indian English bilabial plosives and how they are different from Hindi and American English for varying levels of proficiency in Indian English for bilingual speakers. Previous studies on Hindi plosive speech sounds have found a four-way contrast of plosives. This four-way contrast is a phonemic feature of many languages of the Indo-Aryan group, including Hindi (Bhaskararao, 2011).

Purpose of this study

Therefore, this study aimed to investigate how varying proficiency levels in Indian English affect different acoustic variables while producing different contrasts of bilabial plosives. In order to do so, the present study examined both Hindi and English bilabial speech sounds in a nonword repetition task. The following sections provide (a) a comprehensive review of pertinent literature

in this area and related fields and (b) a detailed description of the methodology employed in this study, (c) results, and (d) discussion.

Chapter 2

Review of the literature

Learning a second language is generally viewed in relation to linguistic processing. This processing involves acquiring a new syntax and lexicon and cognitive control systems involved in translation and switching between languages (Costa & Santesteban, 2004; Crinion et al., 2006; Kovelman et al., 2008; Kovacs & Mehler, 2009; Rodriguez-Fornells et al., 2002). Most models on bilingualism are well-defined with the cognitive planning stages and executive functions, but remain limited in their description of speech, e.g., articulation is described as a single motor output (Simmonds, Wise & Leech, 2011). Alternatively, the majority of speech motor control models are based on monolingual speakers where only the phonetic encoding and the phonetic planning stages are explained (Guenther & Hickok, 2016; Van der Merwe, 1997).

Therefore, there is a disconnect between the language and speech models, especially for bilingual speakers. Thus, it is crucial to understand how bilingualism influences speech motor control. A motor-sensory perspective on bilingualism may be beneficial when considering the influence of speech production on both languages. This includes speech acquisition, accent modification, and individuals with motor speech disorders such as dysarthria (Simmonds, Wise & Leech, 2011).

Increasing our understanding of bilingual speech production would allow us to understand how bilingual speakers have different articulation competencies in two languages, as well as how this competency differs from that of monolingual speakers (Watson, 1990). Additionally, different aspects of bilingual speech production still need to be studied (e.g., proficiency, language input). The current chapter focuses on the bilingual speech model and details how proficiency influences speech production. This model will be used to evaluate Hindi-Indian English bilingual speakers' acoustic characteristics.

Bilingual model for Speech Production and Perception
One of the most frequently cited bilingual perception-production models is the Speech Learning Model (SLM; Flege, 1995). The model is based on various premises, but for the present study, the following is highlighted. The SLM postulates that for bilingual speech production, the phonetic elements that make up the first language (L1) and second language (L2) phonetic-acoustic

characteristics exist in a "common space" and mutually influence one another (Flege, 1995). Further, the assumption in the SLM is that the learning mechanisms used to establish phonetic categories remain intact across the life span and are accessible for both L1 and, later in life, L2 acquisition. This pattern has been observed in bilingual studies examining voice onset time (VOT). VOT is an acoustic characteristic and can be described as a short voicing lag associated with voiced sounds, while a long voicing lag depicts voiceless sounds (Stevens & Klatt, 1972; Whalen et al., 1993).

Flege (1987) studied the voice onset time (VOT) for stop productions in French and English speakers. The words were spoken by native French speakers who were highly experienced in English, and there were three groups of native English subjects differing according to French-language experience. Results showed that the native French subjects who spoke English as L2 produced French /t/ with longer VOT values than French monolinguals. The most proficient speakers of French (L2) produced English /t/ with shorter VOT values than English monolinguals. Thus, the author suggests that the phonetic space of adults is restructured during L2 learning and highlight that the experienced L2 learners can produce similar L2 phones but not create new L2 phones authentically.

Similarly, MacKay, Flege, Piske, & Schirru (2001) also examined VOT in native English, Italian, and bilingual Italian-English speakers. They reported that during the stop /b/, native English speakers produced three different VOTs: with full pre-voicing that continued until stop release, with partial pre-voicing that ceased before release, or as short-lag stops. This result is in contrast to native Italian productions of /b/ produced with full pre-voicing. The Native Italian speakers could not form a new "short-lag" phonetic category for English /b/ because Italian /p/ is realized with short-lag VOT values. MacKay et al. (2001) further reported that early-arriving bilingual participant groups produced English /b/ in an English way, i.e., as short-lag stops, more often than members of the late-arriving bilingual group.

Proficiency is a variable that influences L1-L2 phonetic categories and their interaction in bilingual speakers (Piske, 2001). Flege, Schirru, and MacKay (2003) provided further evidence of L1-L2 interaction for bilingual native Italian (NI) immigrants who had lived in Canada for decades. The authors examined the production of English /eɪ / and Italian /e/. The study revealed late-immigrated bilingual speakers (Italian-English) L2 vowel formants were influenced by L1. However, earlier immigrated bilingual speakers (Italian-English), likely to have received the most English native-speaker input, produced English /eɪ / with significantly more formant movement than Native English speakers. Similarly, Stefanich & Cabrelli (2020) studied early Spanish-English bilingual speakers to measure their acoustic-phonetic representation for both languages. They studied Spanish palatal nasal /ñ/ that is separate from the similar yet acoustically distinct English /n+j/ sequence. The results at the group level revealed that the early bilingual Spanish-English group maintained separate representations for /ñ/ from /n+j/ based on measures of duration and formant trajectories of the following vocalic portion taken from /ñ/ data in Spanish mode and /n+j/ in English mode.

Recent empirical studies on the phonetic interaction in bilingual speakers have reported that traditional Galician-Spanish bilingual speakers (early bilinguals) tended to produce robust mid-vowel contrasts for both languages (Amengual & Chamorro, 2015; López-Bueno, 2017; Tomé-Lourido, 2018; Lourido & Evans, 2015). Meanwhile, late bilingual speakers (who became Galician-dominant as adults) presented open-mid vowels that overlapped in their phonetic spaces with the corresponding closed mid vowels in Spanish, highlighting that the first language influences the second for late bilingual speakers. Similarly, a study by de la Fuente Iglesias & Perez Castillejo (2020) investigated the linguistic influence on the phonetic interactions in the production of closed mid vowels of Galician-Spanish bilinguals. The study revealed that early bilinguals produced distinct phonetic categories for their four Galician mid vowels and additionally produced phonetic categories similar to their L2 (Spanish) vowels.

In addition, Zárate-Sández, G. (2015) studied the perception and production of intonation on native English speakers who learned Spanish. The speakers were divided into three groups-high proficiency, mid proficiency, and low proficiency. These three groups were compared to early bilingual speakers. This study revealed that early bilingual speakers tended to produce values in the middle range between Spanish and English. Performance of low-proficiency speakers generally approximated English monolingual speakers, while L2 speakers of very high proficiency produced values at the same level as native speakers.

The literature on foreign accents also supports the notion that there is an interaction between the first and second language acoustic-phonetic characteristics (Kartushina & Martin, 2019). Accents in speech production happen when an L2 sound assimilates to a phonetically similar L1 speech sound, such that the latter is used to produce both the L1 and the similar L2 sound (e.g., the Japanese /r/ is used to produce the English /ɹ/ and /l/ sounds; Aoyama, Flege, Guion, Akahane-Yamada, &Yamada, 2004). Chen & Mok (2019) studied the acoustic and articulatory features of English and Mandarin /ɹ/ that were learned and produced by highly proficient Mandarin-English bilinguals. Results showed that highly proficient Mandarin-English bilinguals produced English /ɹ/ without L1 influences. They used language-specific phonetic details for English and Mandarin /ɹ/. The authors suggest that language-specific phonetic realizations can be learned by highly proficient bilingual speakers (Chen, & Mok, 2019). For sequential bilingual, it is well established that for L2 speech production and pronunciation, earlier is better; simultaneous (i.e., speakers who acquired their L1 and L2 at the same time) or very early bilinguals are perceived as not having an accent in their L2 by native speakers (Flege, 1999).

The literature above is mainly associated with sequential bilingual speakers. Simultaneous bilingual speakers who have learned two languages at an early age have reported similar findings for acoustic-phonetic representation and the influence of the two languages on each other. Sundara, Polka & Baum (2006) studied the acoustic phonetics of coronal stop produced by adult simultaneous bilingual and monolingual speakers of Canadian English (CE) and Canadian French (CF). Results indicate that simultaneous bilingual, as well as monolingual adults, produce language-specific differences. Also, the study revealed that the simultaneous French-English

bilinguals produce the French /t/ and English /th/ tokens with voice onset time (VOT) like those of monolingual speakers of these languages.

Watson, Trouvain, & Barry (2007) applied the finding from the Speech Learning Model to simultaneous bilingual speakers. The study measured the voicing contrast of French and English for bilingual speakers and compared it to French and English monolingual speakers, respectively. Results suggest simultaneous bilingual speakers produced distinct VOT values in both languages; however, their production differed from the monolingual group. Similarly, Strandberg, Gooskens, & Schüppert (2021) investigated the acoustic-phonetic transfer in the production of long mid front vowels [øː] and [œː] for simultaneous bilingual Finnish and Finland-Swedish speakers in Finland. The authors compared the height and fronting of [øː] and [œː] produced by bilingual and monolingual Finland-Swedish speakers, as the phoneme /ø/ can be realized as the allophones [ø] and [œ] in Swedish, while in Finnish, only [ø] is produced. The results suggest an effect of Finnish (L1) influence on the distinction of the phonetic variants while producing the Swedish phoneme for simultaneous Finland-Swedish bilinguals. Further, the results indicate an increased overlap of [øː] and [œː] in the vowel spaces of bilingual speakers, particularly in informal speech.

Cummings & Montrul (2020) assessed the rhotic production by bilingual Spanish-English speakers. The groups observed were sequential bilingual speakers and simultaneous bilingual speakers. The results of this study suggested that simultaneous bilingual speakers maintained two separate phonetic categories for the alveolar tap and trill, suggesting that they had been exposed to enough Spanish and English input in childhood to distinguish the phonetic differences between these two sounds. In contrast, the sequential bilinguals demonstrated a non-target-like production of the English alveolar approximant, possibly due to influence from the Spanish rhotic system.

Simultaneous bilinguals who have continued to systematically use both the languages, produce both L1 and L2 speech in a manner that does not differ phonetically from monolingual speakers of their respective languages (MacLeod et al., 2009; Guion, 2003). Importantly, simultaneous bilinguals produce similar cross-language sounds distinctly (Fowler et al., 2008; Guion, 2003). Similar results have also been reported in very early bilinguals who acquired their L2 before the age of 3 (Barlow, Branson, & Nip, 2013) and grew up in bilingual communities (e.g., French-English communities of Canada, see MacLeod et al., 2009). Consistent with the literature on sequential bilingual speakers, studies that measured speech production in simultaneous and early bilingual speakers also support the idea that simultaneous and very early bilinguals have separate representations for the L1 and L2.

Both sequential and simultaneous bilingual speech production studies suggest that very early experience with two languages allows bilinguals to partition their acoustic-phonetic space to be able to (1) accommodate the phonetic categories of both their languages, (2) keep these categories separate, and (3) allow interaction between the two languages. Bilingual speech production studies have compared early and late bilingual speakers, and the distinction is based on the age of arrival in a second language (L2) country (predominantly English-speaking countries).

The second language proficiency level is generally understood as a relative and continuous phenomenon, varying from one individual to the next (Birdsong, 2006). Thus, reliable methods to assess bilinguals' speech performance and precise methods to define linguistic proficiency are needed.

Defining Proficiency

Language proficiency is one of the core measures of bilingualism. Studies observing neuroimaging and bilingualism have suggested that language networks are affected by proficiency level. Abutalebi &Green (2007) demonstrated that activity in prefrontal regions reduces as the level of proficiency increases. This suggests that the processes required to produce language become more spontaneous, requiring less domain-general executive control as the language becomes more familiar (Simmonds, Wise & Leech, 2011). Several studies have shown greater activation for lower proficiency (less proficient bilinguals than more proficient bilinguals; Chee et al., 2001; Stein et al., 2009).

Additionally, the results of the second language (L2) foreign accent studies support the view that the earlier in life one learns a second language, the better L2 speech will be pronounced (Flege, 1988; Patkowski, 1990; Thompson, 1991; Flege et al., 1995; Moyer, 1999). Also, studies have suggested that age-related changes in the degree of foreign accent in L2 might be from the nature and the extent of the interaction between a bilingual's first language (L1) and second language (L2) systems (Oyama, 1979; Flege, 1995; Bialystok, 1997). According to this approach, age is an index of the state of development of the L1 system. The more developed the L1 system is when L2 learning commences, the greater the influence L1 has over the L2 production (Piske et al., 2001).

Chakraborty et al. (2008) observed kinematic characteristics for Bengali and English speech sounds for Bengali Indian English bilingual speakers. The results revealed that adult bilingual speakers produced equally consistent speech movement patterns in their production of L1 and L2. Highly proficient speakers were marginally more variable in their L1, and movement durations were longer for less proficient speakers in L1 and L2. Also, the results highlighted that proficiency in L2 provides bilingual speakers with greater flexibility in speech motor planning. (Chakraborty et al., 2008). Nip & Blumenfeld (2015) studied 29 native English speakers with low or high proficiency in Spanish. This study revealed less proficient bilingual speakers differed in their speech motor control and performance from proficient bilingual speakers. Speakers with proficient bilingual speakers had less speech movement variability, shorter durations, greater maximum speeds, and greater ranges of movement. Speakers with less L2 proficiency may have entrenched L1 articulatory movement patterns that interfere with L2 production (Scovel, 1981).

Despite all the literature, defining proficiency has been complex. Different measures have been used for defining proficiency including studies that evaluated age of arrival and length of residence report changes in L1 and L2 phonetic variables (Flege, 1995; Oyama, 1976). Age of arrival has been used as a predictor for proficiency level. Subjects' age of first exposure to the L2, or what is referred to as the age of learning, has been indexed as the age at which the subjects first

arrived as immigrants in a predominantly L2-speaking country (i.e., Canada or the United States of America; Flege, Schirru, and MacKay, 2003; Piske, 2001; Yeni-Komshian et al. 2000).

Piske, Flege, MacKay & Meador (2002) used age of arrival in the L2 country as one of three indirect proficiency measures for a group of seventy-two (72) Italian-English bilinguals residing in Ottawa, Canada. Other variables included the first exposure to the L2, length of residence in an L2-speaking country, and frequency of L2 use. Results highlighted that the bilinguals who were most likely to self-rate their abilities to be higher in Italian (L1) than English (L2) arrived in Canada as young adults and used Italian often. The relationship between age of arrival and degree of the second language (L2) foreign accent is linear until adolescence (Flege et al., 1999; Oyama, 1976).

Mack (1989) studied English French bilinguals and adult English monolingual speakers. The results revealed that the early bilingual speakers' production of /d/ and /t/ and/I/ did not differ significantly from monolingual speakers. The author suggests that early bilingualism can yield monolingual performance in at least one of the bilingual speaker's languages. Another study has also shown that very early bilinguals with the mean age of arrival of 2.4 years do not show altered L1 production (Barlow, 2014). Only later bilinguals with the mean age of arrival of 8.3 years apply the acoustic-phonetic characteristics of the L1 when speaking their L2 (Barlow, 2014).

Another variable used to differentiate proficiency while studying speech production for bilingual speakers was the length of residence (LOR) in an L2 country. Riney & Flege (1998) found that non-native speakers in their senior year of college in the USA were rated higher for accurately producing English speech sounds than their freshman year. There is a significant correlation between length of residence (LOR) in the United States of America and degree of accent; adult learners with LOR's of less than one year will have a stronger accent as compared to a group with LOR's of several years (Flege & Fletcher, 1992). Flege & Liu (2001) assessed the Chinese participants' identification of word-final English consonants and their scores on grammatical judgment test. The participants were assigned to one of the four groups based on LOR in the United States and their primary occupation (students vs. non-students). Participants with longer LOR had significantly higher scores than those with relatively short LORs in all three experiments.

Proficiency in a language depends on many factors (Piske et al., 2001) and affects production (Flege et al., 1995; Fowler et al., 2008; Oyama, 1976; Piske et al., 2002). There is a need to understand the relationship between proficiency in a second language and speech production. The need to understand the relationship between proficiency and speech production for bilingual speakers is two folds:

First, the speech learning model and many bilingual speech production studies are based on sequential bilingual speakers. However, a substantial bilingual population learns English as a second language (sequentially and simultaneously) without arrival in an L2-speaking country. For example, English is a second language in India, and the status of English in India is not viewed as a foreign language compared to other countries (Buschfeld et al., 2014; Melchers et al., 2019).

Indian English is now a fundamental part of the education system in India. It is increasingly becoming the medium of teaching in schools in India starting as early as elementary schools (Piller & Skillings, 2005), although more so in urban areas, private schools, and certain regions of India (Fuchs, 2016). Second, the current bilingual literature has cited Italian, French, Korean, and Spanish, to name a few languages (Piske, Flege, MacKay & Meador, 2002; Mack, 1989; Yeni-Komshian et al., 2000; Riney & Flege, 1998). However, there is a paucity of bilingual studies measuring the acoustic-phonetic characteristic for Hindi and Indian English. The proficiency levels have not been considered in Hindi and Indian English speakers.

Hindi

Hindi is one of the Indo-Aryan languages and is the focus of the present study. This language has a ten-vowel system and consists of three short vowels and seven long vowels (Ohala, 1999). There are twenty-seven consonants in Hindi. These consonants are categorized according to the place and manner of articulation, aspiration, voicing, and nasalization (Manjula & Sharma, 2014). Hindi has aspiration as a phonemic feature, including eight aspirated plosives and two aspirated fricatives (Samudravijaya, Rao & Agrawal, 2000). Aspiration and voicing are independent of each other in Hindi (Shapiro, 1989). In Hindi, bilabial plosives are produced in four ways: voiced plosives, voiceless plosives, voiced aspirated plosives, and voiceless aspirated plosives.

According to Dixit (1987), the essential factor in producing aspiration is the higher-than-normal rate of airflow from the glottis and an unobstructed supra-glottal vocal tract. Further, voicing and aspiration are laryngeally controlled. For voiced aspirated, there is an approximation of the glottis for a part of the closure interval, and the glottis is open for the part of the closure interval and for the entire noise interval. The "voiced aspirated" plosive in Hindi is both voiced and aspirated. The voiceless aspirated plosives are produced with glottal noise following the closure release. Voiced aspiration, in contrast, requires a continuous vibration during the aspiration phase. The spectrogram reveals that the acoustic noise for both voiceless aspirated and voiced aspirated is found in the same frequency region as the upper resonance of the following vowel, revealing that the noise (aspiration) source is the glottis (Dixit, 1987).

Ohala & Ohala (1992) measured the aspiration for Hindi stops. Aspiration was evaluated based on vowel duration and voice onset time. The results highlight that vowel duration following the aspirated stops was longer for voiced aspirated stops than voiceless aspirated stops. Voice onset time (VOT) for the aspirated voiced stop was unmeasurable. However, the overall mean VOT for the voiceless unaspirated stop was 19.4 msec and 84.8 msec for the voiceless aspirated stop. Although the voice onset time is an apt measure for evaluating the voicing differences for various stops, it might not be the most feasible for evaluating the four contrast of stops present in Hindi as there is an overlap of the VOT for unaspirated and aspirated stops (Shimizu, 1989). Due to the four-way stop contrasts in Hindi, acoustic characteristics other than VOT are measured to

provide more detailed information for this contrast. The acoustic characteristics used to describe this four-way contrast are closure duration, fundamental frequency, and spectral tilt.

Dutta (2007) studied the four-way contrast for bilabial plosives. The study highlights that closure duration for aspirated stops is shorter than unaspirated stops regardless of voicing. However, voiced aspirated stops have several other defining features that distinguish them from other Hindi stops. The mean fundamental frequency is lowered for voiced aspirated stops. The voiced aspirated stops are produced with a greater difference in the amplitude of the first and second harmonics. The overall difference between the amplitude of the first harmonic and peak amplitude of the first formant for voiced aspirated stops is higher. The vowels following a voiced aspirated stop have a lower mean F0 than the vowels following a voiceless aspirated stop. The extent of lowering of F0 for vowels following voiced stops is such that the lowering can be seen till 30% of the vowel (Dutta, 2007).

Benguerel & Bhatia (1980) measured closure duration for Hindi stops and differentiated the aspirated versus unaspirated stops using acoustics. The closure duration (CD) for voiceless unaspirated (140 msec) at the intervocalic position is longer than voiced aspirated (112 msec). CD ranged from voiceless unaspirated (154 msec)> voiceless aspirated (140 msec)>voiced unaspirated (130 msec)>voiced aspirated (112 msec). Evaluating the four-way contrast for stops in Hindi reveals that closure durations are longer for unaspirated consonants than aspirated consonants (Durvasula & Luo, 2012).

Acoustic-phonetic characteristics of aspiration were measured for Tamil (L1) and Hindi (L2) for bilingual speakers. They evaluated the acoustic characteristics of both voiceless and voiced aspirate speech sounds in Hindi. The study revealed that there is lag-VOT, which serves as a contrast for the voiceless aspirated stops. In the spectral tilt, the H1-H2 and SNR measures are most effective for the aspirated voiced stops, and there is a greater difference in the spectral tilt for voiced-aspirated stops (Patil & Rao, 2016).

In a cross-language study, the results give an insight into the differences in the production of a stop. This study suggests that there is a lowering of F0 for aspirated stops and are more distinctive and pronounced as compared to the voiced stops for Hindi (Shimizu, 1989). Four-way contrast for stops is present in another Indo-Aryan language (e.g., Nepali). The author acoustically evaluated the four contrasts in Nepali. The study highlights that there is a lowering of F0 for voiced aspirated consonants. Moreover, the voiced and voiceless aspirated sounds are produced with greater glottal opening than unaspirated sounds (Khatiwada, 2009).

Indian English

Indian English (IE) is a variety of English that was developed on the Indian subcontinent (Sirsa, & Redford, 2013). Majority of these IE speakers are native speakers of a regional Indian language such as Hindi or Punjabi. These IE bilingual speakers are first exposed to the language in English medium schools and are educated in Indian English from primary school onwards (age 6), secondary school, or even higher secondary school (Sirsa, & Redford, 2013). The present study

investigates the influence of the native language (Hindi) on Indian English as spoken by Indians who are educated in English medium schools. The acoustic characteristics of Hindi plosives described above for aspiration (e.g., fundamental frequency) should influence L2 speakers.

Sirsa & Redford (2013) studied the native language influences on Indian English spoken by Hindi and Telugu speakers as L2. They measured acoustic similarities and differences between Indian English produced by Hindi (L1) speakers and Telugu (L1) speakers. Some segmental and suprasegmental characteristics were chosen for analysis, including vowel quality (F1-F2), degree of retroflexion for /ʈ/ and /ɖ/, the extent of aspiration for voiceless stops. The results revealed an influence of L1 on Indian English. The results revealed that there were effects of L1 on Indian English in vowel production, articulation of /s/, and on final lengthening. Some language influences like stop aspiration are pronounced for both the native languages Hindi and Telugu, and it was significant for Indian English as well. However, there was less aspiration for Hindi than Telugu. The study also highlighted that Hindi speakers produced speech sounds with a lower average frequency in Hindi-English than Telugu speakers. Thus, the literature for Indian English supports the idea that speakers from different native language backgrounds in India produce a similar IE variety, though some L1-dependent differences are also documented.

Limited studies have measured the acoustics of Indian English consonants, including the influence of aspiration from the first language on Indian English (L2). Malhotra and Vogelaar (2004) have suggested that aspirated Hindi plosives are more highly aspirated than their English counterparts. Therefore, the individual usually chooses to substitute the non-aspirated form of the phoneme instead for Indian English.

McCullough (2013) investigated acoustic characteristics like voice onset time to differentiate the production and perception of non-native English speakers. The results revealed a shorter voice onset time for Indian English bilingual talkers than other Mandarin-English bilingual talkers and monolingual American English talkers for the production task. The author noted shorter VOT was for both voiceless and voiced stops for Indian English speakers. There was also a significant difference in the vowel production for Indian English talkers compared to monolingual American English talkers. Similarly, Davis and Beckman (1983) reported that Indian English speakers produced English voiceless stop targets with short lag VOT and additionally found that most English voiced stop targets were produced with lead voicing. Awan and Stine (2011) found that speakers of Indian English from a variety of L1 backgrounds, including Hindi, had VOT values of 33ms for /p/ and 40ms for /t/, as in word-initial position opposed to 69ms and 77ms, respectively, for American English speakers. Word-medially, Indian English speakers had VOT values of 34ms for /p/ and 39ms for /t/, compared to 67ms and 87ms for American English speakers.

Indian English can be differentiated from American English and British English based on closure duration. Closure duration for /p/ in Indian English is shorter than American English /p/ and British English /p/ except when the syllable is stressed. In the instances of stress, the pattern

is reversed (Bush, 1967). Das and Hansen (2004) state that Indo-Aryan and Dravidian languages have more than six variations of the plosives versus only two in American English (AE), leading to confusion for those Indian English (IE) speakers who have had limited exposure to native American English speech.

Aspiration does not work similarly for Indian English (IE) as it does in Received Pronunciation (RP) or American English (Sailaja, 2009). Aspiration in IE is non-contrastive as it is in RP or American English (AE). Aspiration in Indian English may be spelling dependent (Sailaja, 2009). Moreover, the scarcity of objective data and limited acoustic variables being measured makes it difficult to conclude how aspiration is characterized acoustically for Indian English bilabial plosives.

American English

Most bilingual studies (Yeni-Komshian, Flege, & Liu, 2000) have compared the acoustic differences of bilinguals with a monolingual group. The monolingual group is usually American English speakers. American English is generally described as having a voiced versus voiceless stop contrast at each of three places of articulation: bilabial, alveolar, and velar (Ladefoged, 1999). The most common measurement used to describe the stop contrast is the voice onset time.

Zue (1976) revealed that the mean VOT of voiced stops was found to be 20.6 msec. Specifically, /b/ had a mean VOT of 13 msec. The mean VOT for voiceless stops was found to be 67.5 msec. On average, the burst duration for a voiceless aspirated stop is 15msec longer than a voiced stop (Zue, 1976). Similarly, another study highlights that voiced plosives with prestress have short VOT values (20-30 msec) and a significant formant transition following voice onset, whereas voiceless plosives with a prestress have longer VOT values (greater than 50 msec; Stevens & Klatt, 1974). However, the researchers report that VOT is not the only cue useful for differentiation of the plosives. Formant transitions and other acoustic cues are essential for comparing and discriminating different plosive sounds in English (Stevens & Klatt, 1974). There is significant variability in the VOT values observed for utterance-initial /b,d,g/ when these stops are prevoiced; glottal pulsing is not observed to occur in the interval preceding stop release (Lisker & Abramson, 1964; Zlatin, 1974).

Thus, various studies measuring the American English bilabial plosives have used acoustic characteristics like closure duration and mean fundamental frequency. The closure duration for stop /p/ on average had a greater closure duration than /t/ and /k/ (Yao, 2007; Zue, 1976). The closure duration for alveolar stops is slightly but consistently shorter than bilabials and velars, while bilabials and velars have very similar closure durations (Crystal and House, 1988).

Lisker (1957) studied closures duration for English /p/ and /b/ in the intervocalic position for isolated words spoken at moderate conversational speed. The results revealed that the closure duration is greater for /p/ than for/b/. Moreover, closure durations for /p/ fall in the 90-140 msec range, with an average value of about 120 msec, while values for /b/ vary from 65 to 90 msec, with an average of 75 msec. Closure duration is a major cue to the voiced-voiceless distinction in

intervocalic stops (Lisker, 1957). Stathopoulos & Weismer (1983) revealed that voiceless stops are generally produced with longer closure durations and are unaffected by place-of-articulation. For the medial position, the data shows that the magnitude of the difference for closure duration between the two speech sounds is 13ms.

Ohde (1984) measured fundamental frequency (F0) as an acoustic correlate for voicing in stop sounds in English. The results show that F0 contours are nearly identical for voiceless unaspirated and voiceless aspirated stops. Also, voiceless stops had a significantly higher F0 value than voiced stops. Edwards (1981) also highlighted that F0 of a vowel following the voiceless stops /p, t, k/is typically higher than the average F0 of the same vowel succeeding the voiced stops /b, d, g/. Thus, both the studies evaluating American English stops have revealed that voiceless stops have a higher mean fundamental frequency than voiced stops.

Spectral tilt measure is generally used to investigate the effect of voicing and aspiration on neighboring vowels (Garellek, 2019). A comparison for the spectral amplitude for plosives and fricatives for American English found that spectral tilt was higher for alveolar than labial plosives by about 4 dB for voiced plosives and 14 dB for voiceless plosives, but approximately the same for fricatives (Alwan, Jiang & Chen, 2011).

There is conflicting literature that American English is an aspirating language with an absent feature as the phonological feature of contrast (Honeybone, 2005). Hunnicutt & Morris (2016) found consistent phonetic voicing differences during closure in utterance-medial contexts for stops in English as spoken in Southern America. American English bilabial plosives in word-medial, non-pre-stress position can be differentiated in the voicing contrast, mainly by measuring acoustic characteristics like VOT, fundamental frequency, and closure duration (Lisker, 1980).

Need, Aim and Hypotheses:
The Speech Learning Model (Flege, 1995) and the literature provided above for the acoustic-phonetic representation for bilingual speakers summarizes that *less proficient* bilingual speakers will have acoustic-phonetic characteristics similar to their mother tongue when producing a phoneme in the second language and *proficient* bilingual speakers have acoustic-phonetic representation of phonemes in both languages, (López-Bueno, 2017; Tomé-Lourido, 2018; Sundara, Polka & Baum, 2006). Proficiency has been frequently measured in two ways, age of arrival and the length of residence in L2 country (predominantly English-speaking countries; Flege & Eefting, 1988; Flege 1995). There are bilingual speakers who learn a second language simultaneously or at an early age without entering an L2 speaking country, for example English is spoken and taught as a second language in various previously British colonized countries (e.g., India). Thus, there is a need to study the acoustic-phonetic representation in early bilingual speakers like Hindi-Indian English bilingual speakers who learn the second language without

arriving in the L2 speaking country. Further, the literature on bilingual speech production for Hindi and Indian English is understudied.

There is acoustic and linguistic data available for Hindi and Indian English (Ohala and Ohala, 1992; Dutta, 2007; McCullough, 2013). Hindi has a four-way contrast for bilabial plosives as compared to English which has two bilabial plosives (Dixit, 1987). The acoustic literature on the four-way contrast in Hindi has revealed longer voice onset time, lower mean fundamental frequency, and shorter closure duration for aspirated bilabial plosives for Hindi (Ohala and Ohala, 1992; Benguerel & Bhatia, 1980; Dutta 2007). In Indian English, the plosives have been acoustically characterized as having shorter voice onset time and shorter closure duration (McCullough, 2013; Bush 1967).

A unique phonetic marker, such as aspiration may provide insight into how second-language proficiency may influence the acoustic properties of language-specific sounds. The present study aims to measure the acoustic differences produced by proficient and less proficient bilingual speakers for Indian English and Hindi bilabial plosives and determine how they differ from American English bilabial plosives. The following are the hypotheses:

1. The proficient bilingual speakers will produce both **aspirated** and **unaspirated** stops with acoustic-phonetic characteristics (VOT, Closure Duration, F0, and Spectral Tilts) similar for both Hindi and Indian English bilabial plosives, respectively.
2. Less-proficient speakers will produce acoustic-phonetic characteristics (VOT, Closure Duration, F0, and Spectral Tilts) similar to Hindi **aspirated** bilabial plosives for Indian English **unaspirated** bilabial plosives.
3. Monolingual American English speakers will produce acoustic-phonetic characteristics (VOT, Closure Duration, F0, and Spectral Tilts) similar to English **unaspirated** bilabial plosives for Hindi **aspirated** bilabial plosives.

Chapter 3

METHODS

Participants

The sample size for this study was twenty-four (effect size=0.3, power=0.8). However, only twenty participants (eleven females) between the ages of eighteen and fifty, with normal speech and hearing, were recruited. The lack of recruitment of four more participant was due to the inability to find bilingual speakers who spoke Hindi as their first language and Indian English as their second language (n=4). The participants were divided into three groups. The first group consisted of monolingual English speakers. Monolingual speakers were American English speakers who had English as their native language and went through the English school system

in West Virginia. The participants predominantly used English daily (85-100%)[1], including speaking to friends and family. The average age of the eight-monolingual speakers was 25 years (range 19 to 37 years; SD= 6.63). Initially, the bilingual group included pure bilingual speakers who spoke Hindi as their first language (L1) and Indian English as their second language (L2). However, it was getting difficult to recruit pure bilingual speakers. Thus, the inclusion criteria were expanded to include multilingual speakers from India, whose first language was Hindi, and the second language was Indian English. This population consisted of Indian international students, employees, and immigrants in and around Morgantown, West Virginia, and Pennsylvania. The average age of the twelve bilingual speakers was 33.33 years (range = 24 to 49 years; SD=8.42).

Bilingual language history and proficiency were measured using the Language Experience and Proficiency Questionnaire- Qualtrics Survey (LEAP-Q; Marian, Blumenfeld, & Kaushanskaya, 2007[2], Appendix A). Bilingual speakers were categorized into a proficient and less proficient group according to the aggregate of their responses to the questions about proficiency in speaking, understanding spoken language, and reading for English (L2). The questionnaire evaluated specific measures associated with each language: ages of acquisition; ages of attained fluency; estimates of proficiency in speaking, reading, understanding; the extent of exposure to the language in different contexts; and degree of accent (Kaushanskaya, Blumenfeld & Marian, 2019). The above questionnaire was chosen for the present study as it can be administered across all age groups (14 to 80 years of age) and across all levels of proficiency (Blumenfeld, Schroeder, Bobb, Freeman & Marian, 2016; Krizman, Skoe, Marian & Kraus, 2014). The rating scales make the LEAP-Q sensitive to fluctuations in bilingual experience, and as a result, it can be used by highly proficient bilingual speakers as well as by very inexperienced L2 learners (Conrad, Recio & Jacobs, 2011; Mercier, Pivneva & Titone, 2014; Nip & Blumenfeld, 2015; Pelham & Abrams, 2014).

Table 1. Description of the speaker group with mean of the aggregate response in proficiency for L1 and L2, and approximate age of acquisition (AOA) of first and second language.

Speaker Group	Average Age with SD	L1	L2	Proficiency-L1	Proficiency-L2	AOA-L1	AOA L1
Monolingual	25;6.63	AE		9.625		3	
Bilingual	33;8.42	Hindi	IE	8.52	7.55	5	5-8[3]

---

[1] This range accounts for foreign language training.
[2] https://bilingualism.northwestern.edu/leapq/
[3] This is the data from the participants that completed the questionnaire

Screening

Participants reported no history of speech-language or hearing disorders. All speech and language screening procedures were conducted remotely via zoom with the participant in a controlled and sanitized laboratory. The hearing screening took place in person in the laboratory using a GSI-18 portable audiometer. All the participants passed hearing screening with thresholds below 30dB hearing level at 250Hz, 500Hz, 1000Hz, 2000Hz, and 4000Hz for both ears (American Speech-Language-Hearing Association, 1990). Oral motor examination was not administered to comply with COVID-19 guidelines, which would not allow for proximity to the participant's face. All participants completed the Language Experience and Proficiency Questionnaire (LEAP-Q) through a Qualtrics survey before the appointment.

Monolingual participants were screened on Test of Minimal Articulation Sentence and Reading Screening subtests (Secord, 1981) to rule out speech disorders. Bilingual participants passed a brief interview with basic questions to assess their conversational skills in Hindi. Questions such as "Where do you work?" and "Tell me about your favorite Hindi movies" were asked. In addition, participants were asked to name as many animals in English and fruits in Hindi as possible in one minute (timed by the researcher). Name generating tasks included different categories for each language to control for the priming effect. The average number of names generated by bilinguals for the English category was 17.16, with a standard deviation (SD) of 5.27. The average number of names generated by bilinguals for the Hindi category was 11.16, with a standard deviation (SD) of 2.72. This task has been used in other studies to evaluate fluency in Hindi and fluency in English (Grover, 2016; Hurks et al., 2006).

All participants passed the Photo Articulation Test-Hindi (AYJNISHD-Mumbai, India) with 100% on the percentage of correct consonant (PCC) score. The Test of Minimal Articulation Sentence and Reading Screening subtest (Secord, 1981) also were conducted to rule out speech disorders. The following dialectal variations were kept in mind during the screening. First, Indian English (IE) has several notable phonological features, for example reduced vowel inventory compared to the Received Pronunciation (RP) of British English, second, the substitution of retroflex/dental stops ($/ʈ/$ or $/t̪ʰ/$ and $/ɖ/$) for British English alveolar stops and the omission of some fricative sounds /θ/and /ð/, and a labio-dental approximant $/ʋ/$ is used by some speakers in place of both $/v/$ and$/w/$(Sirsa, & Redford, 2013; Sailaja, 2012; Bansal, 1976; Wells, 1982). Additionally, although suprasegmental features are not standardized, IE's rhythms are notably different from those of most other English dialects.

These participants were recruited, and data was collected from both the onsite (in and around WVU campus) and offsite (ISKON Temple-WV) location. These locations for recruitment and data collection were approved by the West Virginia University Institutional Review Board. Participants provided written consent in compliance with the West Virginia University Institutional Review Board. Participants who passed the screening and completed the task were thanked for their time and were given a gift certificate.

Stimuli

<u>Recording:</u>

All stimuli were recorded in a sound booth by a native Hindi and American English speaker knowledgeable in phonetic transcription and trained to maintain a slow natural speech rate. English nonwords were recorded with the targeted speech sounds with English unaspirated bilabial plosives (voiceless /p/ and voiced /b/). Hindi nonwords were recorded with aspirated bilabial plosives (voiceless-aspirated /pʰ/ and voiced-aspirated /bʰ/). The target bilabial plosive was in the medial position in a Consonant (C) Vowel (V) Consonant (C) Vowel (V) combination (CV**C**VC). The above nonword structure was chosen as the target phonemes for both languages can occur in the medial position in a word, and to control the aspiration in English [pʰ], which occurs in the initial position (Lisker, 1980; Dixit 1987; Ohala, 1983; Hewlett & Beck, 2013; Baken & Orlikoff, 2000; Honeybone, 2005).

The medial position has been chosen for the target sound as Hindi, and English contrasts occur in the medial place for both languages, except /bʰ/ as this speech sound is absent in the English phonetic repertoire. All nonwords were constructed with the front close vowel /i/. This vowel was chosen as it occurs in both the languages (English and Hindi) and controls for geminate (i.e., lengthening of the consonant) that can occur with /pʰ/ in the medial position (Ohala, 1994). All English nonwords composed of the target phoneme followed English phonotactic rules had low bi-phone probability and positional probability to control for linguistic frequency effects. Hindi nonwords were assessed by three Native Hindi speakers for phonotactic constraints and rated for perception and word likeliness. There was a consensus that these two sounds were easy to produce and mostly did not sound like a Hindi word.

Table 2. Nonwords for Hindi and American English

Hindi Nonword	American English Nonword
/mipʰim/	/mipim/
/mibʰim/	/mibim/

The practice set of stimuli were also created and included five English nonwords and five Hindi nonwords. These practice nonwords were different from the actual stimulus but had the target speech sounds.

<u>Perceptual and Spectrographic Judgment:</u>

The author perceptually evaluated all recorded nonwords for syllable stress and pronunciation accuracy. Two raters (undergraduate students whose major is speech-language

pathology and who have passed speech science courses) rated the spectrograms to identify the bilabial plosives for Hindi and English nonwords. The raters labeled voice onset time for all contrasts except /bʰ/. The annotators labeled voiceless plosive stop-gap (period of silence; voiceless), voice bar for voiced plosive speech sound, and the burst for aspiration. Moreover, they identified the fundament frequency, closure duration, and VOT.
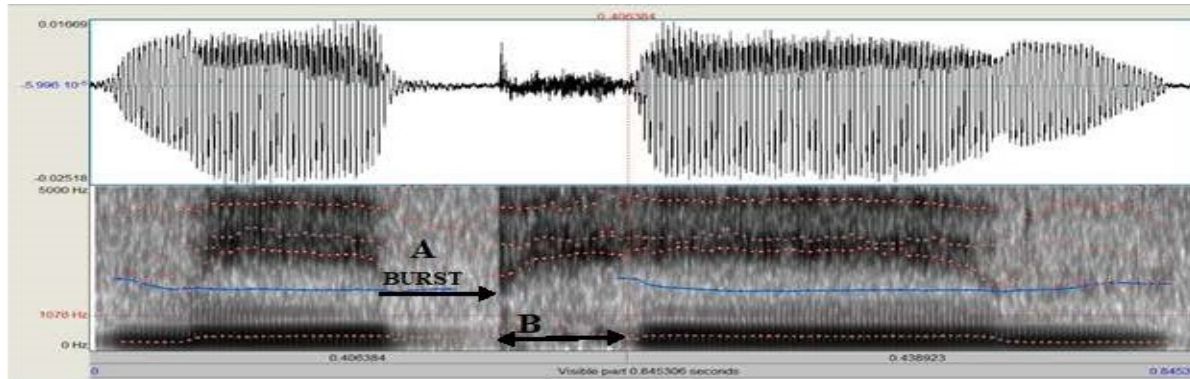
Figure 1. Spectrogram of the American English nonword with voiceless bilabial plosive (American English) shows the Stop Gap (A), Release Burst, and Voice Onset Time (VOT, B).

Stimuli were rated and controlled for loudness and duration (M = 907.517ms, SD = 41.81ms). The loudness was controlled with RMS equalization by using MATLAB (2010). Stimuli were chosen based on the Intra Class Correlation reliability. Excellent reliability was found between the two raters for the fundamental frequency and duration values across all phases of the spectrogram for all the four target speech sounds. In the case of the fundamental frequency, the average measure for a two-way mixed ICC with the absolute agreement was 1.00 (p<.000, 95% CI .999 and 1.000). Additionally, excellent reliability was found between the two raters for all the duration values across all phases of the spectrogram for all four target speech sounds. The average measure for two-way mixed ICC with an absolute agreement for the duration was .997 (p<.000, 95% CI .993 and .999).

The experiment took place amid COVID-19, due to which extra precaution was taken during data collection. The WVU-IRB approved COVID-19 procedure included the following steps below for both onsite and offsite locations for data collection. All the procedures were the same for both the locations.

Lab Daily Activities

All lab personnel maintained a log of when they were present in the lab and testing areas. Masks were worn by all lab personnel in all lab spaces. Only one person was allowed in 807-G Allen Hall (Speech Motor Control Lab) at one time. Other spaces were used for lab meetings or work as needed and will accommodate at least 200 sq ft per lab person as required by the university. All lab materials were cleaned using a disinfectant (Oxivir TB) before leaving the lab,

including pens, laptops, microphones, keyboards, and any other items touched during the time in the lab.

Data Collection Activities

Before the participant's arrival, all surfaces were cleaned using disinfectant (Oxivir TB) (e.g., mic's, keyboards, desks, pens, and any other items touched during the time in the lab). The participant was met outside the building. The participant was asked questions regarding COVID symptoms and was screened for temperature (Appendix B). If the participant passed this screening procedure, they were allowed into the building to complete the experiment.

Procedure
During the experiment:

The researcher wore a mask, a face shield, and gloves throughout the experiment. The table and chairs set up for the experiment maintained a 6-feet distance between the participant and the researcher. The participants were asked to use hand sanitizer before handling any equipment. A plexiglass glass screen was placed between the participant and researcher. Only minimal equipment was placed in front of the participant (e.g., headphones for hearing screening, microphone, and laptop for experimental procedures). The participant kept their mask on from behind the plexiglass screen during the consent process, the hearing screening, and when leaving the testing area. They were asked to remove their mask to repeat nonwords into a microphone during the experimental procedure. During this time, the researcher was not in the room. Following data collection, all the surfaces were cleaned using disinfectant (Oxivir TB).

Nonword repetition task:

Nonword repetition tasks have been used to assess speech production and language skills in bilingual populations (Duncan & Paradis, 2016; Edwards et al., 2004; Munson et al., 2005; Windsor et al., 2010). In various previous studies, nonword repetition has been used to measure lexical access, speech production, motor planning, phonological processing, and phonological memory (Coady & Evans, 2008). Further, research has shown that this task is primarily free of cultural bias, unlike other tasks that are more dependent on language experience (Oetting & Cleveland, 2006; Rodekohr & Haynes, 2001).

The experiment was broken down into two nonword repetition tasks: the practice set and the main task. All stimuli were presented through E-Prime (Schneider, Eschman, & Zuccolotto, 2002). During the nonword repetition task, each participant received visual instruction about the nonword repetition task through E-Prime. Each trial started with a text warning to "listen," which lasted 1000 msec. Immediately after that, the auditory presentation of a single nonword occurred. After the nonword presentation, a text to "repeat" appeared on the laptop screen for 5000ms. During this time, the participant's response was captured by the Shure microphone (SM81-LC Cardioid Condenser microphone). This procedure occurred twenty times within an experimental block for ten blocks. A total of two hundred trials were recorded. Each target

phoneme had fifty trials for each subject. The nonwords were randomized for each block, and each block was pseudorandomized to avoid order artifacts. The nonwords were randomized to control phonetic convergence[4] (Clopper & Dossey, 2020; Nielsen, 2011). No feedback was provided on performance or accuracy. After each block, there was a rest period for the subject.

Data Analysis

 Nonwords were separated, trimmed, labeled, and saved using Audacity (v. 2.3.2) for each block and each participant. Once each sound was trimmed and saved, a text grid file was made using Praat software v.5.3 (Boersma, Weenink, 2013). The text grid file was created to set up the Onset and Offset boundaries for the Consonant Vowel (CV) transition, closure phase, voice onset time for unaspirated bilabial plosives/release burst for aspirated bilabial plosives, and Vowel Consonant (VC) transition. The second formant energy and the vowel duration were used to calculate the boundaries (Cho & Ladefoged, 2002; Patil & Rao, 2013). The starting point for the onset boundary of VC transition was the mid-point of the second formant frequency for the first vowel, and the offset boundary ended at the end of the second formant for the first vowel. The closure phase started from the offset of the VC transition until the beginning of the Release Burst. For CV transition, the onset and offset boundary was selected at the offset of the VOT, and the offset for CV transition ended at the midpoint of the second formant frequency for the second vowel.
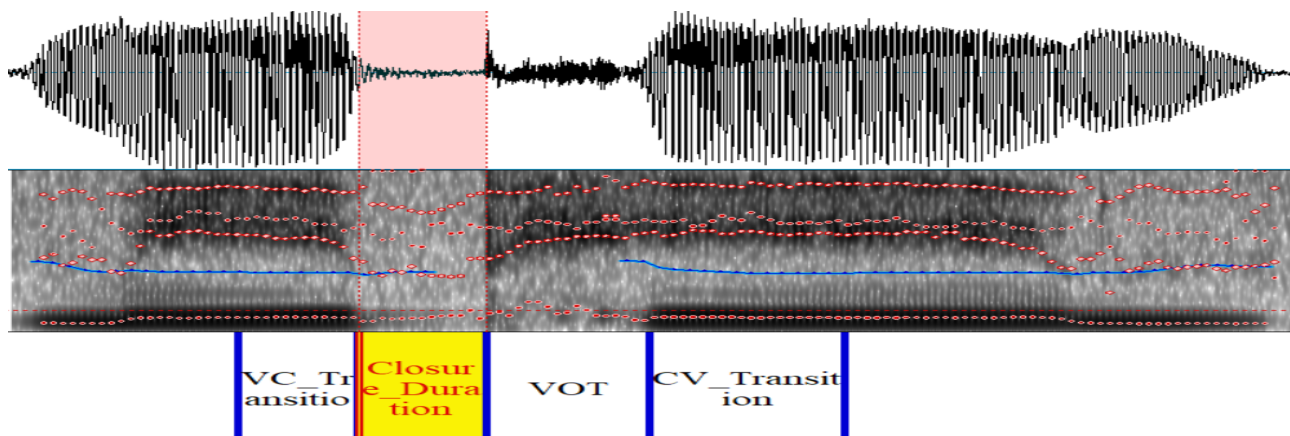
Figure 2: Example of a text-grid file with labeled boundaries VC_Transition, Closure_Duration, VOT and CV_Transition in Praat

 The process of marking the boundaries was first demonstrated to the annotators, who later helped form the text grid files. These annotators were seniors in the undergraduate degree of communication sciences and disorders who took speech science and phonetics courses. After all boundaries were marked, text-grids were formed for each nonword, for all participants, and a Praat script was run on software v.5.3. (Boersma, Weenink, 2013) for calculating the closure duration

---

[4] Non-native speakers imitate the model talker on a shadowing task

and voice onset time. All the durations were logged transformed. The mean fundamental frequency was calculated by using a script[5]. Spectral tilt was calculated using VoiceSauce (Shue, Keating, Vicenik & Yu, 2011). The average was taken of all the spectral tilt values for CV transition for each nonword. The H1-A1 values were used for statistical analysis. H1A1, refers to an estimation of the first formant bandwidth, in the case of aspiration. This spectral tilt measure is expected to be higher than unaspirated stop at the first formant bandwidth measures based on the spectral intensity measure (Dutta 2007; Bali 1999).

Statistical Analysis

Interrater reliability analyses were conducted for fundamental frequency and durations using intraclass correlation coefficient (ICC) estimates. Data were checked for reliability for both monolingual and bilingual groups. Twenty percent of the data from each group were assessed for reliability. The raters randomly selected the participant. Each rater divided the total blocks in half and assessed five blocks per participant. Further, the third trial of each nonword from each block was picked to assess fundamental frequency and duration. Each analysis was based on an absolute-agreement, 2-way mixed-effects model using averaged measures. There was good reliability between the two raters for the fundamental frequency and duration values across all phases of the text-grid files for all four nonwords. For fundamental frequency, the average measure was .897 (p<.000, 95% CI .871 and .917). For duration the average measure was .797 (p< .000, 95% CI .746 and .837).

The ICC for the bilingual speaker group, the fundamental frequency, and duration for all the phases were based on an absolute-agreement, 2-way mixed-effects model using averaged measures. The results show good reliability for both the mean fundamental frequency and duration between the two raters across all phases for all four nonwords. The average measure was .815 (p<.000, 95% CI .778 and .846) for the mean fundamental frequency. The average measure was .831 (p<.000, 95% CI .798 and .859) for the duration.

The linear mixed-effects regression analysis was calculated using the version (version 1.4.1717) of the lme4 package in R (Bates et al., 2014). The four dependent acoustic variables were the voice onset time for the voiceless plosives contrast, the mean fundamental frequency for CV transition, closure duration, and spectral tilt for CV transition. This model estimates the main effect of type of speaker group, type of phoneme, and the interaction between type of speaker group*type of phoneme. The phonemes measured were voiceless bilabial plosive-American English (VLE), voiced bilabial plosive-American English (VE), voiceless aspirated bilabial plosive-Hindi (VLH), and voiced aspirated bilabial plosive-Hindi (VH). The participants were included as a random variable. The type of speaker and type of phoneme were dummy coded. Dummy coding is a way of comparing two or more categorical predictor variables. The monolingual speaker group with a target phoneme was set as a baseline and compared with the

---

[5] https://www.fon.hum.uva.nl/praat/manual/Scripting/fundamental

other two speaker groups and the other three phonemes. The monolingual group was chosen as the baseline as they were the control group in this experiment.

Dependent variable = $\beta 0$ (Type of Speaker) + $\beta 1$ (Type of phoneme) + $\beta 3$ (Type of speaker* Type of phoneme) + error.

Chapter 4

Results

Data from 20 participants (seven proficient bilingual speakers, five less-proficient bilingual speakers, and eight monolingual speakers) were included in the data analysis. The reduced number of participant was due to the inability to find bilingual speakers who spoke Hindi as their first language and Indian English as their second language, and COVID-19 restrictions imposed on recruitment (n=4). Approximately four thousand repetitions were evaluated across the remaining participants. The overall results of this study revealed no significant main effects for the bilingual group across the four dependent variables (refer to the tables for each dependent variable). There was a statistically significant interaction effect between proficiency speaker and phoneme type, which will be described in detail below in reference to each dependent variable.

1. Durational Difference

Two acoustic characteristics were assessed for durational differences for unaspirated and aspirated bilabial plosives. The first is the Voice Onset Time (VOT), and the second is the closure duration. For VOT, the Voiceless bilabial plosive-American English (VLE) was compared to Voiceless Aspirated bilabial plosive-Hindi (VLH). The second acoustic characteristic-closure duration was compared across voiceless unaspirated (VLE) and voiceless aspirated (VLH). The second comparison was voiced unaspirated (VE) and voiced aspirated (VH). Figure 3 highlights the closure duration differences between the speaker group and phoneme type. There were no significant main effects for bilingual speakers, but there was a significant difference in the interaction between the bilingual speaker*phoneme type.
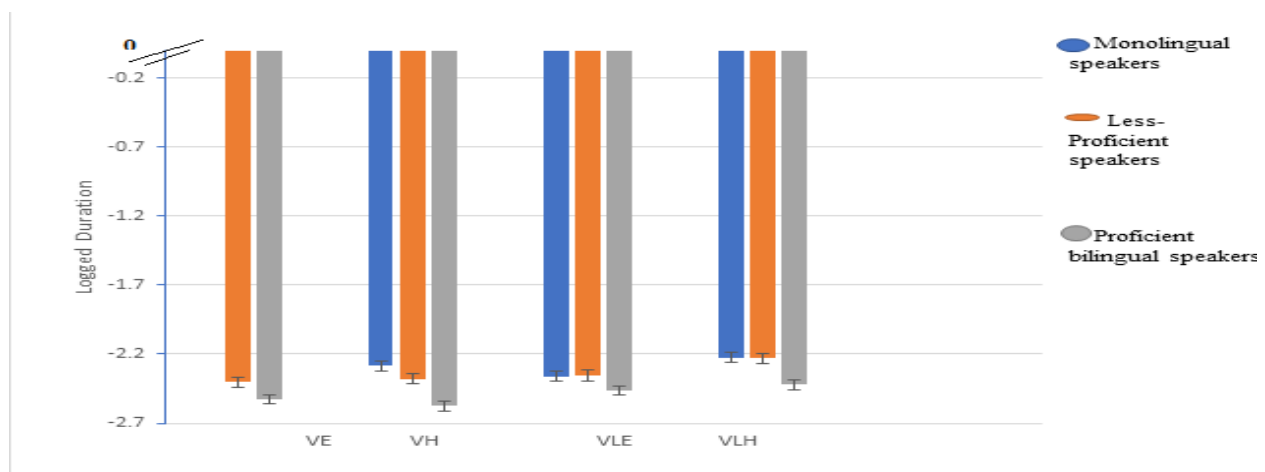

Figure 3: Logged closure duration compared across speaker group per phoneme

1a. Monolingual speakers

There was a statistical difference for VOT between VLE and VLH for monolingual speakers (0.36968, p=.0001). The main effects and interaction effects for the VLE-VLH difference, where VLE is the baseline, are provided in Table 3. The closure duration results highlighted a statistical difference between VLE and VLH for monolingual speakers (0.07998, p=0.001). The main effects and interaction effects for closure duration for baseline monolingual VLE-VLH comparison and between the speaker group are reported in Table 4. In addition, there was a statistical difference for closure duration between voiced bilabial plosive VE and voiced aspirated bilabial plosives VH for monolingual speakers (0.03708, p=0.01). See Table 5 for the main effects and interaction effects for baseline monolingual speaker VE-VH comparison and between the speaker group.

1b. Proficiency v/s bilingual speakers

The closure duration results highlighted a statistical difference for closure duration between VLE and VLH for proficient compared to the less-proficient bilingual group (-0.08209, p<0.01). There was also a statically significant difference in closure duration for the VE-VH comparison VE and VH comparisons for proficient bilingual speakers compared to less-proficient bilingual speakers (-0.07340, p<0.01). See Table 4 (VLE-VLH) and 5 (VE-VH) for the interaction effects for closure duration for both proficient and less proficient groups. Table 6 (VLE-VLH) and 7 (VE-VH) provide correlation for fixed effects for both proficient and less proficient speakers and the compared to the stimulus and monolingual speakers. There does not seem to be a correlation between closure duration difference for both VLE-VLH and VE-VH comparison for proficient bilingual speakers compared to less-proficient bilingual speakers.

Table 3. The Main Effects and Interaction Effects for VOT (log-transformed) of VLE compared to VLH for monolingual and bilingual speakers.

	Estimate	Standard Error	Df	t value
Intercept	-2.85043	0.06924	17.85928	-41.170*
Less Proficient	-0.14364	0.11131	17.65196	-1.290
Proficient	-0.05462	0.11415	17.54666	-0.478
VLH	0.36968	0.03969	3787.61072	9.315*
Less Prof*VLH	0.02648	0.06353	3786.01717	0.417
Proficient*VLH	0.07833	0.06497	3785.02993	1.206

*p =0.000

Table 4. Main Effects and interaction for closure duration (log-transformed) of VLE compared to VLH for monolingual and bilingual speakers.

	Estimate	Standard Error	df	t value
Intercept	-2.33091	0.10309	17.59027	-22.611*
Less Proficient	-0.05974	0.15590	21.53863	-0.383
Proficient	-0.04388	0.13888	41.86135	-0.316
VLH	0.07998	0.01280	3961.79037	6.247*
Less Prof*VLH	0.04245	0.02547	3961.72165	1.728**
Proficient*VLH	-0.08209	0.02740	3962.74220	-2.996***

*p=0.000, **p=0.0, ***p=0.01

Table 5. Main Effects and interaction for closure duration (log-transformed) of VE compared to VH for monolingual and bilingual speakers.

	Estimate	Standard Error	df	t value
Intercept	-2.32973	0.10310	17.59574	-22.598*
Less Proficient	-0.11049	0.15589	21.53338	-0.709
Proficient	-0.06083	0.13887	41.83595	-0.438
VH	0.03708	0.01358	3960.78464	2.730**
Less Prof*VH	-0.01140	0.02488	3960.72116	-0.458
Proficient*VH	-0.07340	0.02731	3960.69467	-2.688**

*p=0.000, **p=0.1, ***p=0.01

Table 6. The correlation of fixed effects and interaction for closure duration (log-transformed) of VLE compared to VLH for monolingual and bilingual speakers.

	Intercept
Less Proficient	-0.661
Proficient	0.000
VLH	-0.057
Less Prof*VLH	-0.030
Proficient*VLH	0.000

Table 7. The correlation of fixed effects and interaction effects for closure duration (log-transformed) of VE compared to VH for monolingual and bilingual speakers.

	Intercept
Less Proficient	-0.661
Proficient	0.000
VH	-0.058
Less Prof*VH	0.031
Proficient*VH	0.000

2. Fundamental Frequency

2a. Monolingual speakers

The difference in the mean fundamental frequency between VLE and VLH was statistically significant for monolingual speakers (-9.176, p=0.001). The main effects and interaction effects for the VLE-VLH difference, where VLE is the baseline, are provided in Table 8. There was no statistical difference in the mean fundamental frequency between VE and VH for monolingual speakers. The main effects and interaction effects for the VE-VH difference, where VE is the baseline, are provided in Table 9.

2b. Bilingual speakers

There were no significant main effects and interaction effects found for proficient and less-proficient bilingual speakers.

Table 8. Main Effects and interaction for Mean Fundamental Frequency (Hz) of VLE compared to VLH for monolingual and bilingual speakers.

	Estimate	Standard Error	df	t value
Intercept	199.411	17.350	17.043	11.493*
Less Proficient	-2.343	27.976	17.043	-0.084
Proficient	-11.299	28.735	17.042	-0.393
VLH	-9.176	1.002	3962.00	-9.158*
Less Prof*VLH	2.668	1.616	3962.00	1651
Proficient*VLH	2.331	1.658	3962.00	1.406

*p=0.000

Table 9. Main Effects and interaction for Mean Fundamental Frequency (Hz) of VE compared to VH for monolingual and bilingual speakers.

	Estimate	Standard Error	df	t value
Intercept	192.9613	17.3502	17.0426	11.122*
Less Proficient	-3.5813	27.97663	17.0425	-0.128
Proficient	-6.5629	28.7346	17.0423	-0.228
VH	-0.6429	1.0019	3962.001	-0.642
Less Prof*VH	-0.5185	1.6147	3962.000	-0.321
Proficient*VH	0.1099	1.6572	3962.000	0.066

*p=0.000

3. Spectral Tilt

3a. Monolingual speakers

The spectral tilt (H1A1) results for VE compared to VH for monolingual speakers highlighted a statistical difference (-4.33975, p=.001). The main effects and interaction effects for the VLE-VLH difference, where VLE is the baseline, are provided in Table 10.

3b. Proficiency v/s bilingual speakers

A statistical difference was observed for the spectral tilt between VLE and VLH for proficient bilingual speakers compared to less-proficient bilingual speakers (-1.6410, p=0.05). Additionally, a significant difference was observed between VLE and VLH for less-proficient speakers than monolingual speakers (2.2398, p=0.001). The main effects and interaction effects for the VLE-VLH difference, where VLE is the baseline, are provided in Table 10. There was a statistical difference between the less proficient bilingual and monolingual groups in the VE and VH comparison (1.57920, p=0.05). The main effects and interaction effects for the VE-VH difference, where VE is the baseline, are provided in Table 11. Figure 4 highlights the spectral tilt differences between the speaker group and phoneme type. There were no significant main effects for bilingual speakers, but there was a significant difference in the interaction between the bilingual speaker*phoneme type. Table 12 (VLE-VLH) and 13 (VE-VH) provide correlation for fixed effects and interaction for both proficient and less proficient compared to the stimulus and monolingual speakers. There does not seem to be a correlation between spectral tilt difference for both VLE-VLH and VE-VH comparison for proficient bilingual speakers compared to less-proficient bilingual speakers.

Table 10. Main Effects and interaction for spectral tilt (H1A1) difference of VLE compared to VLH for monolingual and bilingual speakers.

	Estimate	Standard Error	df	t value
Intercept	14.253	2.6265	17.2807	5.427*
Less Proficient	-1.6758	4.2351	17.2801	-0.396
Proficient	-2.3435	4.3499	17.2797	-0.539
VLH	-0.6331	0.3877	3966.0008	-1.633
Less Prof*VLH	2.2398	0.6245	3966.0005	3.587*
Proficient*VLH	-1.6410	0.6408	3966.0002	-2.561**

*p=0.001, **p=0.05

Table 11. Main Effects and interaction for spectral tilt (H1A1) when VE compared to VH for monolingual and bilingual speakers.

	Estimate	Standard Error	df	t value
Intercept	17.27117	2.62650	17.27972	6.576*
Less Proficient	-0.67553	4.23510	17.27972	-0.160
Proficient	-4.81597	4.34990	17.27972	-1.107
VH	-4.33975	0.38669	3966.00007	-11.223*
Less Prof*VH	1.57920	0.62352	3966.00007	2.533**
Proficient*VH	0.05947	0.64042	3966.00007	0.093

*p=0.001, **p=0.05

Table 12. The correlation of fixed effects and interaction effects for spectral tilt of VLE compared to VLH for monolingual and bilingual speakers.

	Intercept
Less Proficient	-0.620
Proficient	0.000
VLH	-0.074
Less Prof*VLH	0.046
Proficient*VLH	0.000

Table 13. The correlation of fixed effects and interaction effects for spectral tilt of VE compared to VH for monolingual and bilingual speakers.

	Intercept
Less Proficient	-0.620
Proficient	0.000
VH	-0.074
Less Prof*VH	0.046
Proficient*VH	0.000
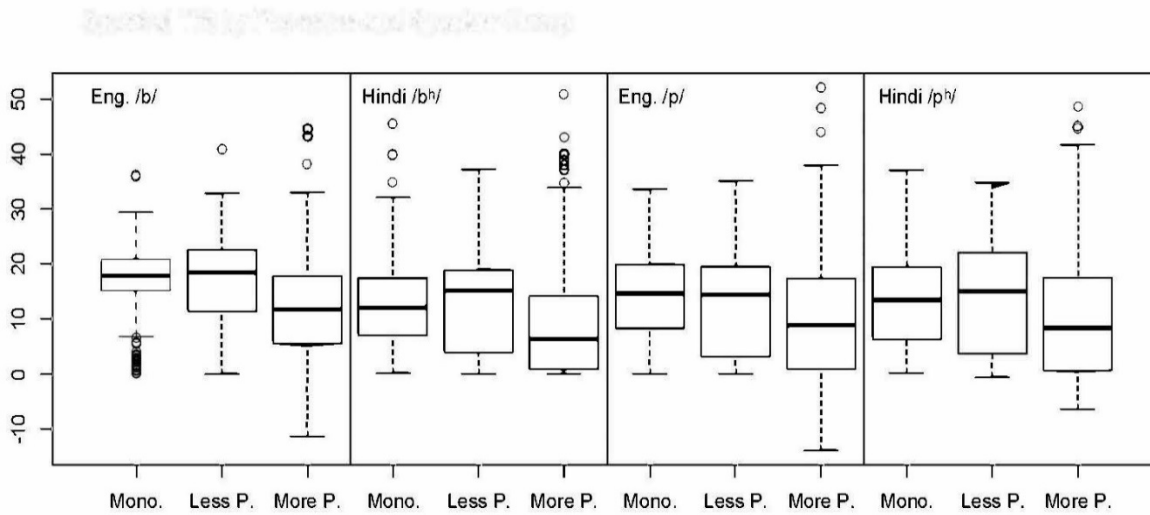
Figure 4: Spectral tilt compared across speaker group per phoneme.

Chapter 5

Discussion

This study aimed to measure the acoustic-phonetic characteristics of bilabial plosives of bilingual Hindi speakers that vary in their Indian English proficiency and monolingual American English speakers. The literature on acoustic-phonetic differences in bilingual speakers has highlighted that a proficient bilingual speaker can produce speech sounds for both languages while maintaining the acoustic-phonetic characteristic for that specific speech sounds in that language (Flege, 1995). In addition, for less proficient bilingual speakers, L2 speech sounds are influenced by L1(Flege, 1995). To measure the above differences for bilingual speakers, three hypotheses were postulated.

Proficient bilingual speakers

The first hypothesis was that proficient bilingual speakers would produce both aspirated and unaspirated bilabial plosives with acoustic-phonetic characteristics similar to Hindi and Indian English bilabial plosives. This hypothesis was rejected as only two of the four acoustic-phonetic characteristics had significant differences in the interaction effect for bilabial plosives similar to Hindi and Indian English bilabial plosives. The results revealed that proficient bilingual speakers have a shorter closure duration for aspirated and unaspirated bilabial plosives. This was an unexpected finding as we anticipated that proficient bilingual speakers would have a shorter closure duration only for aspirated bilabial plosive secondary to typical Hindi speakers producing bilabial plosives with shorter closure duration (Dutta, 2007).

One explanation might be the rate of speech for proficient bilingual speakers. The durational results from the present study indicate that proficient bilingual speakers had a faster rate of speech than less-proficient bilingual speakers. Similar results have been reported by Chakraborty et al. (2008), where the author examined the speech rate as a function of L2 proficiency. The findings revealed that the high proficient speakers produced a faster rate of speech than low-proficient speakers. However, the author reported that the highly proficient group produced sentences in Bengali and English at an equivalent rate of speech compared to sentences produced in Bengali and English by the low proficient group. This finding suggests that the rate of speech and proficiency could be dissociated. Recent evidence highlights that bilingual speaker with increased proficiency may have more flexibility to switch between L1 and L2 articulatory patterns (Nip & Blumenfeld, 2015). Hence, although the proficient speakers may have a faster rate of speech (as suggested by closure duration), this may be due to their proficiency level and flexibility in speaking L1 and L2 and not just in the rate of speech itself.

The second dependent variable (spectral tilt) results indicate that proficient bilingual speakers have a lowered spectral tilt for aspirated and unaspirated bilabial plosives. Again, this was an unexpected finding as we anticipated that proficient bilingual speakers would have lowered spectral tilt-only for unaspirated bilabial plosive as typical Hindi speakers produce aspirated bilabial plosives with higher spectral tilt (Dutta, 2007). One of the reasons for this finding could be that the speech rate might be influencing this acoustic characteristic, i.e., a faster rate of speech causing shorter durational values and more spectral energy in lower frequencies for aspirated and unaspirated plosives. There is evidence that demonstrates that changes in rate of speech can cause changes in segmental duration, articulatory displacement, and velocity and intrasyllabic coarticulation (Gay, 1981; Tasko, & McClean, 2004).

Bilingual speech production studies have documented that speech rate can influence acoustic characteristics like voice onset time and closure duration (Stölten, Abrahamsson & Hyltenstam, 2015; Schmidt & Flege, 1996). Schmidt and Flege (1996) analyzed VOT in word-initial stops produced at three different speaking rates by early and late Spanish-English bilinguals and Spanish and English monolinguals. The finding was that the effects of speaking rate on VOT in the early bilinguals were similar to those found in American English monolinguals; VOT increased

significantly as the speaking rate decreased. The late bilinguals, in contrast, showed less effect of speaking rate on VOT than did the English monolinguals. Guion et al. (2000); Baese-Berk & Morrill (2015) have revealed that the proficiency and experience in a second language influence the rate of speech; the more proficient and experienced the bilingual speakers are in their second language (L2), they have a faster rate of speech in their L2.

Thus, the significant interaction effects for closure duration and spectral tilt for both unaspirated and aspirated plosives as well as correlation results reveal that rate of speech plays a significant part in proficient bilingual speaker's speech production.

Less Proficient bilingual speakers

The second hypothesis was that less-proficient speakers would produce acoustic-phonetic characteristics like Hindi **aspirated** bilabial plosives for American English **unaspirated** bilabial plosives. This hypothesis was partially accepted as two out of four acoustic-phonetic characteristics met the criteria, and there was a uni-directional relationship (L1 influencing L2). The closure duration results highlight that less-proficient bilingual speakers had a shorter closure duration for unaspirated and aspirated bilabial plosives. Studies have highlighted that the Hindi speakers produce aspirated plosives with a shorter closure duration than unaspirated plosive (Durvasula & Luo, 2012; Dutta, 2007). Thus, a shorter closure duration for unaspirated bilabial plosive for Indian English may suggest that less-proficient Indian English speakers were influenced by their L1 (Hindi).

The results of this study shed light on how proficiency plays a role in speech production for bilingual speakers. In the case of the less proficient bilingual speakers, the acoustic characteristics are suggesting that L1 (Hindi) influences L2 (Indian English). These results align with prevalent data on bilingual speakers, which have shown that the first language influences the second language for less proficient bilingual speakers (Flege,1995; Flege, MacKay & Meador (1999); Piske, Mackay & Flege, 2001; Amengual & Chamorro, 2015; López-Bueno, 2017).

This influence of L1 might be due to the acoustic similarity of bilabial plosives in L1 and L2. Thus, in the present study, having the bilabial contrast might have confused the less-proficient bilingual speakers, and therefore, they produced aspirated bilabial plosives, which are present in their L1 repertoire. Similar results have been found by (Kartushina & Frauenfelder, 2014). In this study, the author measured the compactness of specific vowel productions by quantifying the spatial distribution, that is, the spread of produced vowels in the F1-F2 vowel space. The author studied Spanish French bilingual speakers. The results revealed that Spanish speakers whose L1 productions were less variable produced French (L2) sounds were more precise than those whose productions were more variable. The author suggests that the speakers with more variable L1 speech productions have little acoustic space available for the formation of new sounds in their L2, therefore they are more likely to occur within the space of existing L1 sounds and to be confused with them (Kartushina & Frauenfelder, 2014). Thus, it is possible that the compactness

of L1 sound categories might influence the perception and production of the L2 speech sounds (Kartushina & Frauenfelder, 2014).

Second, these durational values were longer for less-proficient bilingual speakers than proficient bilingual speakers. This may be because less-proficient bilingual speakers had a slower speech rate as they anticipated the change between two languages, thus increasing the duration. Similar results have been suggested by Deuchar et al. (2014), where they analyzed the spontaneous bilingual codeswitching in the Bangor Miami Corpus. The results suggest that in anticipation of switching languages, Spanish-English bilinguals produced slowed speech rate and cross-language phonological influence on the voice onset time of plosives. As noted above, the bilingual speakers with higher proficiency in L2 have more flexibility in switching between L1 and L2 and a faster rate of speech in both L1 and L2 (Nip & Blumenfeld, 2015; Chakraborty et al., 2008). Thus, bilingual speakers with less proficiency in L2 may be less flexible and have a slower rate of speech.

Further increased processing demand is associated with L2 speech production (Nip & Blumenfeld, 2015; Green, 1998). In the case of the less-proficient speakers, the more dominant L1(Hindi), which may be more automated, makes it difficult for the less-proficient bilingual speaker to suppress their L1 when producing phonemes in their L2 (Indian English). This would also require more cognitive control to suppress L1 while producing L2 nonwords, possibly requiring a longer time to produce the target nonword/phoneme, resulting in longer durational values in less proficient bilingual speakers than proficient bilingual speakers. The acoustic difference suggests that less-proficient bilingual speakers may be aspirating for Indian English plosives. These results align with the spectral tilt characteristics of the aspirated plosives produced by typical Hindi speakers, i.e., aspirated plosive has a higher spectral tilt than unaspirated plosive (Durvasula & Luo, 2012; Dutta, 2007). Both the acoustic characteristics demonstrate that less proficient bilingual speakers might be influenced by their L1 (Hindi) when producing the Indian English bilabial plosive. Thus, results from both the dependent variable (closure duration and spectral tilt) suggest that Hindi (L1) might be influencing speech production in Indian English (L2) for the less-proficient bilingual speakers.

Monolingual American English speakers

The third hypothesis was that the monolingual American English speakers would produce acoustic-phonetic characteristics similar to **unaspirated** bilabial plosives for **aspirated** bilabial plosives. This hypothesis was accepted as all of the four acoustic-phonetic characteristics met the criteria. The results for the third hypothesis highlight longer closure duration and longer VOT for voiceless aspirated bilabial plosive and a higher mean fundamental frequency for both VLE and VLH. Acoustic studies have highlighted that VOT is longer for voiceless plosive and has a higher mean fundamental frequency (Stevens & Klatt, 1974; Yao, 2007). For the VLE and VLH contrast, there was a longer closure duration for VLE (Stathopoulos & Weismer, 1983). The closure duration was shorter for VE as compared to VH. Thus, revealing that monolingual speakers were producing American English- unaspirated bilabial plosives. The acoustic characteristics results of the monolingual American English speakers suggest that they may have acoustic-phonetic

characteristics that represent a single phonetic representation of bilabial plosive with their voicing contrast.

Monolingual speakers were included because the author assumed that the proficient bilingual speakers would produce unaspirated bilabial plosives similar to the monolingual American English group as both the varieties of English have bilabial contrast and differ in voicing. Yet, the bilingual speakers produced unaspirated bilabial plosives for American English nonword. The results of the acoustic characteristic suggest that they were not similar to the production of monolingual American English speakers. This could be because bilingual speakers may not perceive and produce the bilabial contrast in the same way as monolingual American English speakers do.

Summary and Conclusion

The results of this study were unexpected for the bilingual speaker group, especially the proficient bilingual speaker group. The results were not as clear-cut. Thus, it is difficult to suggest that proficient bilingual speakers have distinct acoustic-phonetic characteristics for each phoneme in each language. Four acoustic variables were measured, i.e., voice onset time, closure duration, mean fundamental frequency, and spectral tilt. There were no significant main effects or interaction effects for both proficient and less proficient bilingual speakers in the case of voice onset time and mean fundamental frequency, and this may be due to decreased sample size in the present study. However, both closure duration and spectral tilt results prove that the rate of speech plays an essential role in bilingual speech production. In addition, the two dependent variables indicate that language proficiency influences bilingual speech production in unique and specific ways.

The results suggest there is an interlanguage influence for less-proficient bilingual speakers, the first language influences the second language. Literature on bilingual speech production has provided empirical data to support the cross-linguistic phonetic transfer observed in this study (Flege et al., 1999, Baker and Trofimovich, 2005; Strandberg, Gooskens & Schüppert, 2021). The SLM postulates that L2 learners may differ from monolingual speakers due to interactions between L1 and L2 subsystems, which occur due to the L1 and L2 sounds existing in a common phonetic space (Flege, 1995). This interaction is described as bidirectional, as L1 and L2 sound linked to one another may resemble each other in production (Flege, 1995).

Although Flege's model provides a theoretical framework for sequential bilingual speakers to understand *how* a bilingual speaker's L1 and L2 acoustic representations may coexist and interact, it does not explain the acoustic-phonetic representation and their interaction in the case of early or simultaneous bilingual speakers. In the present study, the interlanguage influence (i.e., uni-directional) for early bilingual speakers differing in their L2 proficiency may be due to the following reasons. The first might be the age of acquisition of the speaker since all the bilingual speakers learned Hindi early (approximately between 3-5 years). It is common for bilinguals to speak their second language with a detectable foreign accent, that is, produce speech that is detectably different from that of native speakers of the language. Thus, this unidirectional

influence may be due to the age of acquisition and the input provided while learning a second language. Substantial evidence suggests that the extent of the differences between bilingual production of L1 and L2 phones from those of monolinguals depends on the age at which each language is acquired (Flege, 1987; Aoyama et al., 2004; MacKay, Flege, Piske, & Schirru 2001; Zárate-Sández, 2015).

To conclude, the benefits from this present study are multifold. First, this study supports and fills in the gaps in the literature for bilingual speech production. Second, the results of this study help understand how proficiency in a second language plays a role in speech production for early bilingual speakers. Third, the results of this study provide acoustic data for Hindi and Indian English bilingual speakers. Furthermore, the acoustic data provided by this study might be helpful in clinical assessments for bilingual speakers, especially in speech sound disorders and speech acquisition in bilingual speakers. This acoustic information on Hindi-Indian English bilingual speakers is helpful as there is a growing population of Hindi English bilingual speakers (i.e., an approximate forty-two percent increase in bilingual Hindi and English speakers in the US) in a growing bilingual population in the US (Zeigler & Camarota, 2019).

Limitation

This study had the following limitations that are outlined below.

Input: This present study has highlighted differences between aspirated and unaspirated bilabial plosives for each language category when produced by the two groups of bilingual speakers. However, the comparison between the American English monolingual speakers is not valuable. This is due to how bilingual speakers perceive and produce speech, which differs from monolingual American English speakers. The difference may be due to the language input, as it varies significantly from monolingual American English speakers. In the case of Indian English, it is taught by non-native English speakers in India. The second language immersion for Indian English bilingual speakers also differed from the American English monolingual speakers. Flege & Bohn (2021) have recognized that it's impossible to equate speech production for a second language learner to a monolingual speaker of the L2 as various factors play a role in learning a second language. This may include the type of input received, duration of the input received, and frequency with which the input is received. In addition, future studies need to include these different types of input and note whether native or non-native speakers of L2 gave this input in L2.

Proficiency:

1. Speakers: most bilingual speakers in the present study are early bilingual speakers who learned Indian English at a young age. These speakers were students in higher education and employees working in the USA, requiring a certain level of proficiency in L2. Thus, recruiting bilingual speakers based in India will help get a diverse pool of proficiencies in L2 for future studies. Also, a control for a third language would be helpful as a majority of the Hind-Indian English participants in this study were multilingual.

32

2. Measure: In the present study, LEAP-Q was used for determining proficiency. This questionnaire is based on self-perception. Thus, participants' motivation and self-esteem can play a role in determining their proficiency. Proficiency plays a vital role in speech production. Thus, it needs to be free from bias that includes (participants' and investigators' bias). The bias can be controlled by adding another assessment or tool that the researcher can administer to calculate proficiency and the proficiency questionnaire completed by the participants to determine proficiency.

Stimuli: In the present study, four nonwords were asked to be repeated. These nonwords were randomized; however, there was a possibility of phonetic convergence due to limited nonwords. Hence, to avoid phonetic convergence in future studies, the stimuli need to be pseudo-randomized, or filler words could be added. Nasal phoneme was used as a control for first and final position in the nonword as they are present in both American English and Hindi. However, nasal phoneme can affect acoustic characteristics like closure duration for aspirated plosives in specific languages (Downing & Hamann, 2021). Thus, it might be helpful in future studies to either have a different phoneme as a control for both languages or measure the effect of nasality on acoustic characteristics for aspiration.

Fundamental Frequency: In the present study, the fundamental frequency was one of the dependent variables. There is enough physiological and acoustic evidence to highlight that fundamental frequency is dependent on gender. However, in this study, gender differences were not measured. Thus, it is crucial to measure the gender differences for future studies.

Future Direction

There are several areas of further research that could be examined based on this research study. First, the present study results revealed that the acoustic characteristics for unaspirated and aspirated bilabial plosives for bilingual speakers, especially the proficient group, are masked by the rate of speech. Thus, for future studies, it is critical to control for the rate of speech, evaluate the acoustic characteristics (example: VOT, closure duration, and spectral tilt) and kinematic characteristics (example: range, velocity, and spatio-temporal index) for bilingual speakers, and observe whether proficient bilingual speakers have separate acoustic-phonetic/articulatory-phonetic characteristics for each phoneme in each language.

Second, this study highlights that speech rate plays a crucial role in bilingual speech production. Thus, for future studies, the rate of speech needs to be controlled for all the production studies for bilingual speakers. Especially if we want to observe the interlanguage influence (uni-directional and bi-directional) on both L1 and L2 in bilingual speech production. In the present study, we observe that the first language influences speech production in L2. However, it is not apparent from the results if there is an interaction between the languages for proficient bilingual speakers. Kartushina, Hervais-Adelman, Frauenfelder, & Golestani (2016) studied the phonetic drift of L1 towards L2 for native and non-native vowel production during a short period. These participants were monolingual French speakers who learned Danish vowels. The results revealed that the production of native vowels was affected by the L2, with a drift towards non-native vowels. The Speech Learning Model postulates that "L1 categories are not fossilized but continue to

develop over the lifespan "to reflect the properties of all L1 or L2 phones identified as a realization of each category" (Flege, 1995, p 239). Furthermore, studies have highlighted that L1 and L2 coexist and constantly interact for bilingual speakers (Grosjean, 1989; Piske et al., 2001; Sancier & Fowler, 1997). Thus, to have a better understanding of the interlanguage influence, it would be interesting to measure both acoustic characteristics (example: VOT, closure duration, formant frequencies) and kinematic variables (example: range, velocity, and spatio-temporal index) for phonemes from both L1 and L2 to observe the bi-directional interaction or phonetic drift of languages with varying levels of proficiency for Hindi-Indian English early bilingual speakers after controlling for the rate of speech.

Third, examining the relationship between the perception and production of aspirated and unaspirated bilabial plosives for Hindi-Indian English bilingual speakers. A perceptual assimilation study with a goodness of fit task within the framework of Perceptual models like PALM-L2 (Best &Tyler, 2007) and the revised Speech Learning Model (SLM-r; Flege & Bohn, 2021) will provide a better understanding of the relationship between the two processes. The results of this study highlight that for less-proficient bilingual speakers, the first language influences speech production in their second language. Thus, it is essential to measure how perception plays a role in the interlanguage influence in early or simultaneous bilingual speakers with varying proficiency levels. Especially, Hindi has specific phonemes which are absent in the English phonetic repertoire and vis-versa. For example, a perception-based study observed the identification and discrimination accuracy of the /v/ and /w/ comparison, based on the length of residence of Hindi listeners in the USA. Results revealed that bilingual Hindi-Indian English speakers performed significantly less accurately than American English (AE) controls. The results highlight that American English (AE) exposure in the US did not improve perceptual performance (Grover, Shafer, Campanelli, Whalen, & Levy, 2021). Thus, it would be interesting to measure how aspiration affects the identification and discrimination of bilabial plosives in Hindi and Indian English.

Fourth, the present study indicates that proficiency influences the acoustic-phonetic representation, and there is an interlanguage influence for both the languages in bilingual speakers. Future studies could vary the speech rate in a nonword repetition task and use kinematic variables to measure if proficiency still influences phonetic representation at an articulatory-phonetic level for bilingual speakers. Also, along with proficiency, cognitive and linguistic complexities play a role in the speech motor control for bilingual speakers (Nip & Blumenfeld, 2015). Therefore, in the future, researchers can measure the effects of linguistic complexities by varying the length of nonword and having more than one task (nonword and sentences) can help understand the relationship between linguistic complexities and speech motor control in bilingual speakers. Kinematic measures of speech motor performance, such as duration, spatio-temporal index, and range of movement, may provide information that may help in bridging that relationship between how linguistic complexity affects L2 speech production, especially in the case of early bilingual speakers.

Fifth, a longitudinal study measuring the bilingual speakers who are learning English as a second language in their adulthood may provide information on the role of input in the acoustic-phonetic relationship between the first and second languages. This experiment should include sessions that control the type of input received, duration of the input received, and frequency with which the input is received. This study could measure the acoustic differences in L1 and L2 speech production before the first session and after the last sessions received to see what type of changes can be observed with respect to acoustic-phonetic characteristics in each language. Also, this type of study could provide information on how much second language input matters in second language acquisition for bilingual speakers. Literature has shown that L1 and L2 phonetic systems are dynamic and continue to change (Chang, 2019; Tobin, Nam & Fowler, 2017; Chang, 2012). Thus, investigating the phonetic drift based on the input received would provide us with the acoustic-phonetic information and changes in L1 and L2.

Lastly, replication of studies also provides insight into whether effects can be generalized to a larger population. Thus, researchers may want to investigate the current findings in the future, with an increased sample size for bilingual speakers (e.g., 20-30 participants). Also, they might want to examine the acoustic differences and similarities between the aspirated and unaspirated bilabial contrast in Hindi and Indian English, providing a detailed description of the acoustic characteristics for each language. Thus, providing more acoustic evidence for the two languages and the phonetic contrast in those two languages.

References

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of neurolinguistics*, *20*(3), 242-275. https://doi.org/10.1016/j.jneuroling.2006.10.003

Alwan, A., Jiang, J., & Chen, W. (2011). Perception of place of articulation for plosives and fricatives in noise. *Speech communication*, *53*(2), 195-209.

Amengual, M., & Chamorro, P. (2015). The effects of language dominance in the perception and production of the Galician mid vowel contrasts. *Phonetica*, *72*(4), 207-236. DOI: 10.1159/000439406

American Speech-Language-Hearing Association. (1990). Guidelines for screening hearing impairment and middle-ear disorders. ASHA, 32(2), 17–24.

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese/r/and English/l/and/r. *Journal of Phonetics*, *32*(2), 233-250. https://doi.org/10.1016/S0095-4470(03)00036-6

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

Baese-Berk, M. M., & Morrill, T. H. (2015). Speaking rate consistency in native and non-native speakers of English. *The Journal of the Acoustical Society of America*, *138*(3), EL223-EL228. https://doi.org/10.1121/1.4929622

Baken, R. J., & Orlikoff, R. F. (2000). *Clinical measurement of speech and voice*. Cengage Learning. 225-295

Baker, W., & Trofimovich, P. (2005). Interaction of native-and second-language vowel system (s) in early and late bilinguals. *Language and speech*, *48*(1), 1-27. https://doi.org/10.1177/00238309050480010101

Bansal, R. K. (1976). The Intelligibility of Indian English, Hyderabad: Central Institute of English and Foreign Languages. *Eric Report Ed*, *177849*.

Barlow, J. A. (2014). Age of acquisition and allophony in Spanish-English bilinguals. *Frontiers in psychology*, *5*, 288. DOI https://doi.org/10.3389/fpsyg.2014.00288

Barlow, J. A., Branson, P. E., & Nip, I. S. (2013). Phonetic equivalence in the acquisition of/l/by Spanish–English bilingual children. *Bilingualism: Language and Cognition*, *16*(1), 68-85. DOI https://doi.org/10.1017/S1366728912000235

Bates, D., Maechler, M., Bolker, B., Walker S. (2014). lme4: Linear Mixed-Effects Models Using Eigen and S4. R package version 1.1-7, URL http://CRAN.R-project.org/package= lme4

Benguerel, A. P., & Bhatia, T. K. (1980). Hindi stop consonants: an acoustic and fiberscopic study. *Phonetica*, *37*(3), 134-148.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech. *Language experience in second language speech learning: In honor of James Emil Flege*, *17*, 13.

Bhaskararao, P. (2011). Salient phonetic features of Indian languages in speech technology. *Sadhana*, *36*(5), 587-599.

Bialystok, E. (1997). The structure of age: In search of barriers to second language acquisition. *Second language research*, *13*(2), 116-137. https://doi.org/10.1191/026765897677670241

Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language learning*, *56*, 9-49. **https://doi.org/10.1111/j.1467-9922.2006.00353.x**

Blumenfeld, H. K., Schroeder, S. R., Bobb, S. C., Freeman, M. R., & Marian, V. (2016). Auditory word recognition across the lifespan: Links between linguistic and nonlinguistic inhibitory control in bilinguals and monolinguals. *Linguistic Approaches to Bilingualism*, *6*(1-2), 119-146. https://doi.org/10.1075/lab.14030.blu

Boersma, P., & Weenink, D. (2013). Praat software. *Amsterdam: University of Amsterdam*.

Buschfeld, S., Hoffmann, T., Huber, M., & Kautzsch, A. (Eds.). (2014). *The evolution of Englishes: The dynamic model and beyond*. John Benjamins Publishing Company.

Bush, C. N. (1967). Some acoustic parameters of speech and their relationships to the perception of dialect differences. *TESOL Quarterly*, *1*(3), 20-30.

Chakraborty, R., Goffman, L., & Smith, A. (2008). Physiological indices of bilingualism: Oral–motor coordination and speech rate in Bengali–English speakers. *Journal of Speech, Language, and Hearing Research*. DOI: 10.1044/1092-4388(2008/024)

Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics*, *40*(2), 249–268. https://doi.org/10.1016/j.wocn.2011.10.007

Chang, C. B. (2019). Language change and linguistic inquiry in a world of multicompetence: Sustained phonetic drift and its implications for behavioral linguistic research. *Journal of Phonetics*, *74*, 96-113. DOI https://doi.org/10.1016/j.wocn.2019.03.001

Chee, M. W., Hon, N., Lee, H. L., & Soon, C. S. (2001). Relative language proficiency modulates BOLD signal change when bilinguals perform semantic judgments. *Neuroimage*, *13*(6),1155-1163. https://doi.org/10.1006/nimg.2001.0781

Chen, S., & Mok, P. P. K. (2019). Speech production of rhotics in highly proficient bilinguals: Acoustic and articulatory measures. In *Proceedings of ICPhS*.

Cho, T., Jun, S. A., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of phonetics*, *30*(2), 193-228. DOI https://doi.org/10.1006/jpho.2001.0153

Clopper, C. G., & Dossey, E. (2020). Phonetic convergence to Southern American English: Acoustics and perception. *The Journal of the Acoustical Society of America*, *147*(1), 671-683.

Coady, J. A., & Evans, J. L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *International Journal of Language & Communication Disorders*, *43*(1), 1-40. **https://doi.org/10.1080/13682820601116485**

Conrad, M., Recio, G., & Jacobs, A. M. (2011). The time course of emotion effects in first and second language processing: a cross cultural ERP study with German–Spanish bilinguals. *Frontiers in psychology*, *2*, 351. https://doi.org/10.3389/fpsyg.2011.00351

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of memory and Language*, *50*(4), 491-511. https://doi.org/10.1016/j.jml.2004.02.002

Crinion, J., Turner, R., Grogan, A., Hanakawa, T., Noppeney, U., Devlin, J. T., ... & Usui, K. (2006). Language control in the bilingual brain. *Science*, *312*(5779), 1537-1540. **https://doi.org/10.1126/science.1127761**

Crystal, T. H., & House, A. S. (1988). The duration of American-English stop consonants: An overview. *Journal of Phonetics*, *16*(3), 285-294.

Cummings Ruiz, L. D., & Montrul, S. (2020). Assessing Rhotic Production by Bilingual Spanish Speakers. *Languages*, *5*(4), 51. **https://doi.org/10.3390/languages5040051**

Das, S., & Hansen, J. H. (2004, June). Detection of voice onset time (VOT) for unvoived stops (/p/,/t/,/k/) using the Teager energy operator (TEO) for automatic detection of accented English. In *Proceedings of the 6th Nordic Signal Processing Symposium, 2004. NORSIG 2004.* (pp. 344-347). IEEE.

Davis, K., & Beckman, M. (1983). Production and perception of the voicing contrast in Indian and American English. *Working Papers of the Cornell Phonetics Laboratory*, *1*(7790), 108.

de la Fuente Iglesias, M., & Perez Castillejo, S. (2020). Phonetic interactions in the bilingual production of Galician and Spanish/e/and/o. *International Journal of Bilingualism*, *24*(2), 305-318. https://doi.org/10.1177/1367006919826868

Deuchar, M., Davies, P., Herring, J. R., Couto, M. C. P., & Carter, D. (2014). 5. Building Bilingual Corpora. In *Advances in the Study of Bilingualism* (pp. 93-110). Multilingual Matters. https://doi.org/10.21832/9781783091713-008

Dixit, P. (1987). Mechanisms for voicing and aspiration. *UCLA Working Papers in Phonetics*, *67*, 49-102.

Dixit, R. P. (1987, June). In defense of the phonetic adequacy of the traditional term 'voiced aspirated'. In *Proceedings of the Eleventh International Congress of Phonetic Sciences* (Vol. 2, pp. 145-148).

Duncan, T. S., & Paradis, J. (2016). English language learners' nonword repetition performance: The influence of age, L2 vocabulary size, length of L2 exposure, and L1 phonology. *Journal of speech, language, and hearing research*, *59*(1), 39-48. https://doi.org/10.1044/2015_JSLHR-L-14-0020

Durvasula, K., & Luo, Q. (2012, October). Voicing, aspiration, and vowel duration in Hindi. In *Proceedings of Meetings on Acoustics 164ASA* (Vol. 18, No. 1, p. 060009). Acoustical Society of America.

Dutta, I. (2007). *Four-way stop contrasts in Hindi: An acoustic study of voicing, fundamental frequency and spectral tilt*. Urbana, IL: University of Illinois at Urbana-Champaign. http://hdl.handle.net/2142/82651

Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of speech, language, and hearing research*. https://doi.org/10.1044/1092-4388(2004/034)

Edwards, T. J. (1981). Multiple features analysis of intervocalic English plosives. *The Journal of the Acoustical Society of America*, *69*(2), 535-547.

Flege, J. E. (1987). A critical period for learning to pronounce foreign languages?. *Appliedlinguistics*, *8*(2),162-177. https://doi.org/10.1093/applin/8.2.162

Flege, J. E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *The Journal of the Acoustical Society of America*, *84*(1), 70-79. https://doi.org/10.1121/1.396876

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

Flege, J. E., & Bohn, O. S. (2021). The revised speech learning model (SLM-r). *second language speech learning: theoretical and empirical progress*, 3-83.

Flege, J. E., & Eefting, W. (1988). Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation. *The Journal of the Acoustical Society of America*, *83*(2), 729-740. https://doi.org/10.1121/1.396115

Flege, J. E., & Fletcher, K. L. (1992). Talker and listener effects on degree of perceived foreign accent. *The Journal of the Acoustical Society of America*, *91*(1), 370-389. https://doi.org/10.1121/1.402780

Flege, J. E., & Liu, S. (2001). THE EFFECT OF EXPERIENCE ON ADULTS' ACQUISITION OF A SECOND LANGUAGE. *Studies in second language acquisition*, *23*(4), 527-552. DOI 10.1017/S0272263101004041

Flege, J. E., Munro, M. J., & MacKay, I. R. (1995). Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America*, *97*(5), 3125-3134. https://doi.org/10.1121/1.413041

Flege, J. E., Schirru, C., & MacKay, I. R. (2003). Interaction between the native and second language phonetic subsystems. *Speech communication*, *40*(4), 467-491. doi:10.1016/S0167-6393(02)00128-0

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of memory and language*, *41*(1), 78-104. https://doi.org/10.1006/jmla.1999.2638

Fowler, C. A., Sramko, V., Ostry, D. J., Rowland, S. A., & Hallé, P. (2008). Cross language phonetic influences on the speech of French–English bilinguals. *Journal of Phonetics*, *36*(4), 649-663. https://doi.org/10.1016/j.wocn.2008.04.001

Fuchs, R. (2016). The historical and social context of Indian English. In *Speech Rhythm in Varieties of English* (pp. 9-34). Springer, Berlin, Heidelberg.

Garellek, M. (2019). The phonetics of voice. *The Routledge handbook of phonetics*, 75-106.

Gay, T. (1981). Mechanisms in the control of speech rate. *Phonetica*, *38*(1-3), 148-158. 10.1159/000260020

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and cognition*, *1*(2), 67-81. https://doi.org/10.1017/S1366728998000133

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and language*, *36*(1), 3-15. DOI https://doi.org/10.1016/0093-934X(89)90048-5

Grosjean, F. (2008). *Studying bilinguals*. Oxford University Press.

Grover, V. (2016). Perception and Production of/v/and/w/in Hindi speakers *(Doctoral dissertation, CUNY Academic Works.)*. https://academicworks.cuny.edu/gc_etds/1545

Grover, V., Shafer, V. L., Campanelli, L., Whalen, D. H., & Levy, E. S. (2021). Perception of American English consonants/v/and/w/by Hindi speakers of English. *Journal of Second Language Pronunciation*. DOI https://doi.org/10.1075/jslp.20033.gro

Guenther, F. H., & Hickok, G. (2016). Neural models of motor speech control. In *Neurobiology of language* (pp. 725-740). Academic Press. https://doi.org/10.1016/B978-0-12-407794-2.00058-4

Guion, S. G. (2003). The Vowel Systems of Quichua-Spanish Bilinguals. *Phonetica*, *60*(2), 98–128. https://doi.org/10.1159/000071449

Guion, S. G., Flege, J. E., Liu, S. H., & Yeni-Komshian, G. H. (2000). Age of learning effects on the duration of sentences produced in a second language. *Applied Psycholinguistics*, *21*(2), 205-228. https://doi.org/10.1017/S0142716400002034

Hewlett, N., & Beck, J. M. (2013). *An introduction to the science of phonetics*. Routledge.

Honeybone, P. (2005). Diachronic evidence in segmental phonology: the case of obstruent laryngeal specifications. *The internal organization of phonological segments*, *319*, 54.

Hunnicutt, L., & Morris, P. A. (2016). Prevoicing and aspiration in Southern American English. *University of Pennsylvania Working Papers in Linguistics*, *22*(1), 24.

Hurks, P. P. M., Vles, J. S. H., Hendriksen, J. G. M., Kalff, A. C., Feron, F. J. M., Kroes, M., ... & Jolles, J. (2006). Semantic category fluency versus initial letter fluency over 60 seconds as a measure of automatic and controlled processing in healthy school-aged children. *Journal of Clinical and Experimental Neuropsychology*, *28*(5), 684-695. https://doi.org/10.1080/13803390590954191

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. Routledge.

Kartushina, N., & Frauenfelder, U. H. (2014). On the effects of L2 perception and of individual differences in L1 production on L2 pronunciation. *Frontiers in psychology*, *5*, 1246. https://doi.org/10.3389/fpsyg.2014.01246

Kartushina, N., & Martin, C. D. (2019). Talker and acoustic variability in learning to produce nonnative sounds: evidence from articulatory training. *Language Learning*, *69*(1), 71-105. **https://doi.org/10.1111/lang.12315**

Kartushina, N., Hervais-Adelman, A., Frauenfelder, U. H., & Golestani, N. (2016). Mutual influences between native and non-native vowels in production: Evidence from short-term visual articulatory feedback training. *Journal of Phonetics*, *57*, 21-39. https://doi.org/10.1016/j.wocn.2016.05.001

Kaushanskaya, M., Blumenfeld, H. K., & Marian, V. (2019). The Language Experience and Proficiency Questionnaire (LEAP-Q): Ten years later. *Bilingualism: Language and Cognition*, 1-6. https://doi.org/10.1017/S1366728919000038

Khatiwada, R. (2009). Nepali. *Journal of the International Phonetic Association*, *39*(3), 373-380.

Klatt, D. H., & Stevens, K. N. (1972, April). Sentence recognition from visual examination of spectrograms and machine-aided lexical searching. In *Proceedings 1972 Conference on Speech Communication and Processing* (pp. 315-318).

Kovács, Á. M., & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences*, *106*(16), 6556-6560. https://doi.org/10.1073/pnas.0811323106

Kovelman, I., Baker, S. A., & Petitto, L. A. (2008). Bilingual and monolingual brains compared: a functional magnetic resonance imaging investigation of syntactic processing and a possible "neural signature" of bilingualism. *Journal of cognitive neuroscience*, *20*(1), 153-169. https://doi.org/10.1162/jocn.2008.20011

Krizman, J., Skoe, E., Marian, V., & Kraus, N. (2014). Bilingualism increases neural response consistency and attentional control: Evidence for sensory and cognitive coupling. *Brain and language*, *128*(1), 34-40. https://doi.org/10.1016/j.bandl.2013.11.006

Ladefoged, P. (1999). American English. *Handbook of the international phonetic association*, *41-44*

Lisker, L. (1957). Closure duration and the intervocalic voiced-voiceless distinction in English. *Language*, *33*(1), 42-49.

Lisker, L. (1978). In qualified defense of VOT. *Language and Speech*, *21*(4), 375-383.

Lisker, L. (1980). On generalizing the rabid-rapid distinction based on silent gap duration. *The Journal of the Acoustical Society of America*, *67*(S1), S52-S52.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*(3), 384-422. https://doi.org/10.1080/00437956.1964.11659830

López-Bueno, L. (2017). *Galician-Spanish mid vowel production in Southern Galicia: using language mode to explore bilingual speakers' underlying phonological categories* (Doctoral dissertation, Cardiff Metropolitan University).

Lourido, G. T., & Evans, B. G. (2015). Switching language dominance for ideological reasons: A study of Galician new speakers' speech production and perception. *In Proceedings of the 18th 702 International Congress of Phonetic Sciences, 703 ed. by The Scottish Consortium for ICPhS. Glasgow: University of Glasgow*

Mack, M. (1989). Consonant and vowel perception and production: Early English-French bilinguals and English monolinguals. *Perception & Psychophysics*, *46*(2), 187-200. DOI: 10.3758/bf03204982

MacKay, I. R., Flege, J. E., Piske, T., & Schirru, C. (2001). Category restructuring during second-language speech acquisition. *The Journal of the Acoustical Society of America*, *110*(1), 516-528. DOI: 10.1121/1.1377287

MacLeod, A. A., Stoel-Gammon, C., & Wassink, A. B. (2009). Production of high vowels in Canadian English and Canadian French: A comparison of early bilingual and monolingual speakers. *Journal of Phonetics*, *37*(4), 374-387. DOI https://doi.org/10.1016/j.wocn.2009.07.001

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

Manjula, R., & Sharma, N. (2014). 14 Motor Speech Disorders in Languages of the Indian Subcontinent: Some Perspectives from Hindi and Kannada. *Motor Speech Disorders: A Cross-Language Perspective*, *12*, 207.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*. DOI: 10.1044/1092-4388(2007/067)

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**

McCullough, E. A. (2013) Perceived foreign accent in three varieties of non-native English. OSUWPL Volume 60, 51-66.

Melchers, G., Shaw, P., & Sundkvist, P. (2019). *World Englishes*. Routledge.

Mercier, J., Pivneva, I., & Titone, D. (2014). Individual differences in inhibitory control relate to bilingual spoken word processing. *Bilingualism: Language and Cognition*, *17*(1), 89-117. https://doi.org/10.1017/S1366728913000084

Moyer, A. (1999). Ultimate attainment in L2 phonology: The critical factors of age, motivation, and instruction. *Studies in second language acquisition*, 81-108. https://doi.org/10.1017/S0272263199001035

Munson, B., Kurtz, B. A., & Windsor, J. (2005). The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/1092-4388(2005/072)

Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics*, *39*(2), 132-142.

Nip, I. S., & Blumenfeld, H. K. (2015). Proficiency and linguistic complexity influence speech motor control and performance in Spanish language learners. *Journal of Speech, Language, and Hearing Research*, *58*(3), 653-668. https://doi.org/10.1044/2015_JSLHR-S-13-0299

Oetting, J. B., & Cleveland, L. H. (2006). The clinical utility of nonword repetition for children living in the rural south of the US. *Clinical Linguistics & Phonetics*, *20*(7-8), 553-561. https://doi.org/10.1080/02699200500266455

Ohala, M. (1983). *ASPECTS OF HINDI PHONOLOGY* (Vol. 2). Motilal Banarsidass Publisher. 37-76.

Ohala, M. (1994). Hindi. *Journal of the International Phonetic Association*, *24*(1), 35-38.

Ohala, M. (1999). Hindi. Handbook of the International Phonetic Association: a guide to the use of the International Phonetic Alphabet.

Ohala, M., & Ohala, J. J. (1992). Phonetic universals and Hindi segment duration. In *Second International Conference on Spoken Language Processing*.

Ohde, R. N. (1984). Fundamental frequency as an acoustic correlate of stop consonant voicing. *The Journal of the Acoustical Society of America*, *75*(1), 224-230.

Olson, D. J. (2016). The role of code-switching and language context in bilingual phonetic transfer. *Journal of the International Phonetic Association*, *46*(3), 263-285. DOI https://doi.org/10.1017/S0025100315000468

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

Oyama, S. (1979). The concept of the sensitive period in developmental studies. *Merrill-Palmer Quarterly of Behavior and Development*, *25*(2), 83-103.

Patil, V. V., & Rao, P. (2016). Detection of phonemic aspiration for spoken Hindi pronunciation evaluation. *Journal of Phonetics*, *54*, 202-221.

Patil, V., & Rao, P. (2013). Acoustic features for detection of phonemic aspiration in voiced plosives. In *INTERSPEECH* (pp. 1761-1765).

Patkowski, M. S. (1990). Age and accent in a second language: A reply to James Emil Flege. *Applied linguistics*, *11*(1), 73-89.

Pelham, S. D., & Abrams, L. (2014). Cognitive advantages and disadvantages in early and late bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(2), 313. https://doi.org/10.1037/a0035224

Pickering, L., & Wiltshire, C. (2000). Pitch accent in Indian-English teaching discourse. *World Englishes*, *19*(2), 173-183.

Piller, B., & Skillings, M. J. (2005). English Language Teaching Strategies Used by Primary Teachers in One New Delhi, India School. *Tesl-Ej*, *9*(3), n3.

Piske, T., Flege, J. E., MacKay, I. R., & Meador, D. (2002). The production of English vowels by fluent early and late Italian-English bilinguals. *Phonetica*, *59*(1), 49-71. DOI: 10.1159/000056205

Piske, T., MacKay, I. R., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of phonetics*, *29*(2), 191-215. DOI:10.1006/jpho.2001.0134

Riney, T. J., & Flege, J. E. (1998). Changes over time in global foreign accent and liquid identifiability and accuracy. *Studies in Second Language Acquisition*, *20*(2), 213-243.

Rodekohr, R. K., & Haynes, W. O. (2001). Differentiating dialect from disorder: A comparison of two processing tasks and a standardized language test. *Journal of Communication Disorders*, *34*(3), 255-272. https://doi.org/10.1016/S0021-9924(01)00050-8

Rodriguez-Fornells, A., Rotte, M., Heinze, H. J., Nösselt, T., & Münte, T. F. (2002). Brain potential and functional MRI evidence for how to handle two languages with one brain. *Nature*, *415*(6875), 1026-1029. DOI: 10.1038/4151026a

Sailaja, P. (2009). *Indian English*. Edinburgh University Press. 17-39.

Samudravijaya, K., Rao, P. V. S., & Agrawal, S. S. (2000). Hindi speech database. In *Sixth International Conference on Spoken Language Processing*.

Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, *25*(4), 421–436. https://doi.org/10.1006/jpho.1997.0051

Schmidt, A. M., & Flege, J. E. (1996). Speaking rate effects on stops produced by Spanish and English monolinguals and Spanish/English bilinguals. *Phonetica*, *53*(3), 162-179. https://doi.org/10.1159/000262196

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-prime (version 2.0). *Computer software and manual]. Pittsburgh, PA: Psychology Software Tools Inc*.

Scovel, T. (1981). The recognition of foreign accents in English and its implications for psycholinguistic theories of language acquisition. In *Proceedings of the 5th Congress of AILA* (pp. 389-401). University of Laval Press Laval.

Secord, W. (1981). *T-MAC: Test of minimal articulation competence*. Charles E. Merrill Publishing Company.

Shapiro, M. C. (1989). *A primer of modern standard Hindi*. Motilal Banarsidass Publisher.

Shimizu, K. (1989). A cross-language study of voicing contrasts of stops. Studia Phonoligica 23.1-12

Simmonds, A. J., Wise, R. J., & Leech, R. (2011). Two tongues, one brain: imaging bilingual speech production. *Frontiers in Psychology*, *2*, 166. https://doi.org/10.3389/fpsyg.2011.00166

Sirsa, H., & Redford, M. A. (2013). The effects of native language on Indian English sounds and timing patterns. *Journal of phonetics*, *41*(6), 393-406. DOI: 10.1016/j.wocn.2013.07.004

Stathopoulos, E. T., & Weismer, G. (1983). Closure duration of stop consonants. *Journal of Phonetics*, *11*(4), 395-400.

Stefanich, S., & Cabrelli, J. (2020). Shared or Separate Representations? The Spanish Palatal Nasal in Early Spanish/English Bilinguals. *Languages*, *5*(4), 50. **https://doi.org/10.3390/languages5040050**

Stein, M., Federspiel, A., Koenig, T., Wirth, M., Lehmann, C., Wiest, R., & Dierks, T. (2009). Reduced frontal activation with increasing 2nd language proficiency. *Neuropsychologia*, *47*(13),2712-2720. DOI: 10.1016/j.neuropsychologia.2009.05.023

Stevens, K. N., & Klatt, D. H. (1974). Role of formant transitions in the voiced-voiceless distinction for stops. *The Journal of the Acoustical Society of America*, *55*(3), 653-659.

Stölten, K., Abrahamsson, N., & Hyltenstam, K. (2015). Effects of age and speaking rate on voice onset time: The production of voiceless stops by near-native L2 speakers. *Studies in Second Language Acquisition*, *37*(1), 71-100. https://doi.org/10.1017/S0272263114000151

Strandberg, J. A. E., Gooskens, C., & Schüppert, A. (2021). Simultaneous bilingualism and speech style as predictors of variation in allophone production: Evidence from Finland-Swedish. *Journal of Phonetics*, *88*, 101095.

Sundara, M., Polka, L., & Baum, S. (2006). Production of coronal stops by simultaneous bilingual adults. *Bilingualism: Language and Cognition*, *9*(1), 97–114. https://doi.org/10.1012/S1366728905002403

Tasko, S. M., & McClean, M. D. (2004). Variations in articulatory movement with changes in speech task. *Journal of Speech, Language and Hearing Research*, 47 (1), 85-100. https://doi.org/10.1044/1092-4388(2004/008)

Thompson, I. (1991). Foreign accents revisited: The English pronunciation of Russian immigrants. *Language learning*, *41*(2), 177-204.

Tobin, S. J., Nam, H., & Fowler, C. A. (2017). Phonetic drift in Spanish-English bilinguals: Experiment and a self-organizing model. *Journal of Phonetics*, *65*, 45-59. DOI https://doi.org/10.1016/j.wocn.2017.05.006

Tomé Lourido, G. (2018). *The role of social factors in bilingual speech processing: The case of Galician New Speakers* (Doctoral dissertation, UCL University College London).

Van der Merwe, A. (1997). A theoretical framework for the characterization of pathological speech sensorimotor control. *Clinical management of sensorimotor speech disorders*, *2*, 3-18.

Watson, I. (1990). Acquiring the voicing contrast in French: A comparative study of monolingual and bilingual children. In J. N. Green & W. Ayers-Bennett (eds.), Variation and change in French: Essays presented to Rebecca Posner on the occasion of her sixtieth birthday, pp. 37–60. London: Routledge

Watson, I., Trouvain, J., & Barry, W. J. (2007). Simultaneous bilinguals and Flege's speech learning model. In *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 1533-1536).

Wells, J. C. (1982). *Accents of English: Volume 3: Beyond the British Isles*. Cambridge University Press.

Whalen, D. H., Abramson, A. S., Lisker, L., & Mody, M. (1993). F 0 gives voicing information even with unambiguous voice onset times. *The Journal of the Acoustical Society of America*, *93*(4), 2152-2159. DOI : 10.1016/j.wocn.2017.05.002

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.

Wiltshire, C., & Moon, R. (2003). Phonetic stress in Indian English vs. American English. *World Englishes*, *22*(3), 291-303. **https://doi.org/10.1111/1467-971X.00297**

Yao, Y. (2007). Closure duration and VOT of word-initial voiceless plosives in English in spontaneous connected speech. *UC Berkeley PhonLab Annual Report*, *3*(3).

Yeni-Komshian, G. H., Flege, J. E., & Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean–English bilinguals. *Bilingualism: Language and cognition*, *3*(2), 131-149. https://doi.org/10.1017/S1366728900000225

Zárate-Sández, G. (2015). Perception and production of intonation among English-Spanish bilingual speakers at different proficiency levels *(Doctoral dissertation, Georgetown University)*.

Zeigler, K., & Camarota, S. (2019). 67.3 million in the United States spoke a foreign language at home in 2018. *Center for Immigration Studies*. https://cis.org/Report/673-Million-United-States-Spoke-Foreign-Language-Home-2018

Zlatin, M. A. (1974). Voicing contrast: Perceptual and productive voice onset time characteristics of adults. *The Journal of the Acoustical Society of America*, *56*(3), 981-994. https://doi.org/10.1121/1.1903359

Zue, V. W. (1976). Acoustic characteristics of stop consonants: A controlled study. *Technical Report*. Massachusetts Inst of Tech Lexington Lincoln Lab. https://apps.dtic.mil/sti/pdfs/ADA034600.pdf

Appendix A

The language experience and proficiency questionnaire (Marian, Blumenfeld, & Kaushanskaya, 2007) is a self-reported questionnaire where the participants provide answers based on the following questions below. This questionnaire can be administered to both bilingual and multilingual speakers. The questionnaire is based on the number of languages known, and each language has a set of questions that the participant needs to answer.

1. Age of learning of language: the participant reported the age of acquisition for each language, respectively, for bilingual and multilingual speakers.
2. Order of dominance in a language: the participant self-ranked the orders of the dominance of each language, respectively, for both bilingual and multilingual speakers.
3. Proficiency in a language: the participants reported proficiency in speaking, reading, and understanding spoken language. The participant had to answer the proficiency question for each language known. The self-rating proficiency scale is from 0 to 10, and the rating ranges from 0, which means no proficiency, and 10 is excellent proficiency.
4. Immersion duration: the participant reported the duration they spent in the first language and second language speaking environment (e.g., school and job).
5. Contributors to language learning/ acquisition: the participant self-rated the contributors like friends, tv shows, etc., on a scale of 1-10 that helped in language learning.
6. The extent of exposure in each language: After completing the contribution part of the questionnaire, participants answered the degree to which they are exposed to these contributors for each language.
7. Self-reported foreign accent: the participants had to self-rate themselves on their perception of an accent in each language. Also, they self-rated how much others identify them as non-native speakers in the languages they know, based on their accent.

Appendix B
WVU Speech and Hearing Clinic Screening Questions Prior to Entering Building

Adapted from the WVU Medicine - Morgantown Campus Screening Questions at Entry Points effective 6-15-20

1. In the two weeks prior to today, did you have contact with someone who was diagnosedwith COVID-19?

2. In the two weeks prior to today, did you live in or visit a place where COVID-19 is spreading (e.g., a nursing home, city, or other facility)?

If the answer to either or both questions is "Yes," the client or guardian will need to reschedulethe appointment in alignment with current guidelines (e.g., 2-weeks after the fever has discontinued).

Check for Symptoms

1. I am going to take your temperature. (Must be below 100.4 degrees)

2. Do you/Does your child have a fever or have you/he/she felt hot or feverish in the last two

   days?

3. Do you have any of the following symptoms?
   - Shortness of breath

   - Cough

   - Runny or stuffy nose

   - Muscle aches, body aches, or headaches

   - New loss of taste or smell

   - Fatigue or malaise

   - Nausea, vomiting or diarrhea

If the client has a fever (100.4 degrees or higher), answers yes to any of the questions, or is observed to exhibit any of the symptoms listed in Question 3 of Part 2, the client will be askedto reschedule in accordance with current guidelines (e.g., 2-weeks after symptom(s) have ceased).