THE INFLUENCE OF GUJARATI ON THE VOT OF ENGLISH STOPS

Safiyyah Fadia

Thesis submitted to the University of Huddersfield for the degree of Master of Arts by Research

October 2017

ABSTRACT

This thesis is an acoustic study of three generations of women from the Gujarati-English community in Batley, West Yorkshire (United Kingdom). The literature pertaining to Gujarati is very limited, hence this thesis aims to solve some unanswered queries regarding our understanding of the influence of Gujarati on English stop productions. Research was gathered from three generations of women to judge whether there is an influence of Gujarati on English speech. The first generation was born in India and migrated to the UK during their childhood, whereas the second and third generation were born in the UK.

In order to assess this influence, the acoustic attribute of voice onset time (VOT) was studied for all stops /p b t d k g/ at word-initial level. Participants were recorded reading out pre-arranged flashcards, generating eight hundred and forty-four tokens of data, all of which were later analysed in Praat. A further objective of this research was to examine whether there are variances of VOT at different positions of an utterance.

Previous studies in the acoustic literature have studied VOT interference in bilinguals of various languages, which have yielded mixed results, including L1 influence, L2 influence, as well as no interference whatsoever. The primary results demonstrated that there is an influence of Gujarati (the L1) on English speech (the L2) in the first generation, which reduces for the second and third generation. Furthermore, a second significant finding was that there appears to be a high usage of Gujarati-like negative VOT at utterance-initial position for all three generations.

Several explanations are proposed to explain the two aforementioned findings. The main argument put forth is that the participants' VOT values mirrored the values of their dominant language. Secondly, there is perhaps a sociolinguistic change occurring in the Batley community in which negative VOT has ethnic connotations and is being used by the British-born generations as a marker of identity.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

As the fourth generation – the first great-granddaughter – of a Gujarati immigrant, I have heard countless stories of my great-grandfather's journey to the UK, to a place that he could not call his first home, where there laid traditions he did not understand, and where he then learned a language he could not easily grasp. This thesis, therefore, is inspired by my great-grandfather, and everyone else who came before me, who paved the way for me to be where I am now.

The completion of this thesis would not have been possible without the assistance and encouragement of various people. I would first like to thank my MA supervisor Dr Erica Gold of the Linguistics and Modern Languages Department at The University of Huddersfield. Her support from whilst I was an undergraduate student up until the final days of my completing this thesis has always been continuous and plentiful. Over the duration of this year, Dr Gold steered my work in the appropriate direction whilst still giving me the freedom to work independently. She has contributed towards a fulfilling postgraduate experience and for that I am truly grateful.

Additionally, I would like to thank all the participants who took part in this thesis. Though their input by means of allowing me to record them may seem insignificant to them, it was a real privilege for me to gather data from a Gujarati community of which I am a member.

Finally, thank you to my parents, the rest of my family, and my friends who have supported me through the course of this year. This Master's has been a great struggle to complete and, had it not been for my loved ones who saw my every high and (very) ugly low, I doubt I would have reached the submission deadline with a completed thesis in hand. In particular, I would like to thank my friend Maariyah. Not only have you been a friend to me this year, but you have become my sister. Thank you for being my support, for constantly comparing your MA word count to mine in the hope that it would force me to work harder, but most importantly thank you for showing me how and reminding me to switch off from my work and live a little.

Thank you.

1. INTRODUCTION

Gujarati is one of the several languages that is stitched into India's as well as England's multilingual tapestry. Nevertheless, from the various languages within the Indian subcontinent, Gujarati has rarely been at the forefront of linguistic studies. Gujarati is often omitted from research in favour of other more socially important and widely-used languages from the Indo-Aryan family, such as Urdu and Punjabi. Although the linguistic structures of Gujarati have been described (see: Pandit, 1961; Cardona, 1965) and some acoustic aspects have been discussed (see: Esposito and Khan, 2012; Khan, 2012; Kreiman et al. 2010), there is still an extensive proportion of the language as well as many dialects of Gujarati that are yet to be studied, not least in relation to other languages, such as English.

To widen our understanding of Gujarati, this thesis will investigate Gujarati speakers' voice onset time (VOT) in English speech. Voice onset time (VOT) is one of the acoustic cues, along with voicing and aspiration, that allows us to distinguish between voiced and voiceless stops. Despite VOT being characterised in the first half of the 20th century, it was only given its name in Lisker and Abramson's famous study (1964) in which they defined it as "the time interval between the burst that marks the release of the stop closure and the onset of the quasi-periodicity that reflects laryngeal vibration" (1964, p. 422). However, Lisker and Abramson in a later paper (1967) stated that, despite the importance of their earlier study, there is still a "need for a closer look at the individual languages than (their) cross-language survey afforded" (1967, p. 4).

Consequently, research concerning the interaction between languages and its influence on VOT has been heavily investigated on numerous languages, including Spanish, Polish, German and, of course, English. Whilst some studies have focused on VOT influence at a pre-pubescent age (see: Kehoe et al., 2004; Harada, 2007), others have observed participants in their adulthood (see: Newlin-Åukowicz, 2014; Lein at al. 2016). Information pertaining to the acoustic cue of VOT in Gujarati is extremely scarce, and is non-existent in terms of a comparison across different ages. This thesis, therefore, will consider Gujarati-English speakers aged from 18 to nearly 80 years old, divided into three generations.

Many of the aforementioned scholars (e.g. Kehoe at al., 2004, Newlin-Åukowicz, 2014), as well as others (e.g. Sancier and Fowler, 1997), have observed language interference or transfer in which speakers reinterpret features of one language and implement them in another language that they speak. Although this can apply to many linguistic aspects, with regards to VOT scholars have found transfer from an L1 to an L2, (see: Caramazza et al., 1973; Flege & Efting 1987a), bi-directional L1-L2 interference in which both languages influence one another (see: Flege & Efting 1987b, Fowler et al. 2008), and have also found the third possibility of an L2 suppressing L1 interference completely (see: Flege et al., 2002). This thesis will solely focus on the VOT influence Gujarati (the L1) has on English (the L2) at word-initial level across each generation and in different positions of an utterance. This thesis will attempt to answer the following three research questions:

1. What are the word-initial VOT values for /p b t d k g/ of the data across the three generations, and do such values correspond more with English or with Gujarati VOTs?
2. Are there any variances in word-initial VOT at different places in an utterance, and is there any significance of these variances?
3. Are there any sociolinguistic stimuli/influences that may affect the transfer of VOT for any of the participants?

In order to answer these research questions, fieldwork was carried out on British-Gujarati speakers residing in the West Yorkshire town of Batley. All participants involved in this thesis are female bilinguals and were all asked to articulate the same target words in utterances that placed the word as utterance-initial, utterance-medial and utterance-final.

All of the preliminary information provided in this introduction will be expanded in the following chapters. Chapter 2 gives the background information needed to contextualise this thesis. This will include a brief history of Gujarati, migration from Gujarat to Yorkshire as well as statistical reports of Gujarati in the UK. Next, phonetic information will be provided regarding Gujarati and English, followed by scholarly literature of VOT and language contact.

Subsequently, Chapter 3 describes the methods that were carried out in order to complete and analyse the fieldwork. This includes the grounds on which participants were recruited, details of the activity that they were asked to complete, as well as the approach taken to acoustically measure the data.

Chapter 4 displays all of the findings that arose from the fieldwork and presents these results in numerous graphs to best illustrate the patterns between each generation and each position of the target word. Continuing from this, Chapter 5 examines the main patterns that were presented in Chapter 4. It provides potential justifications for the data and why certain patterns may have emerged. This chapter will refer to literature from Chapter 2 as well as other relevant studies to aid in the understanding of the results and to assess how Gujarati-English VOTs fit in amongst previous research.

Lastly, Chapter 6 concludes the thesis and reviews the main findings of this work. It will explicitly answer each of the three research questions that were outlined earlier in this chapter. It will then consider the limitations of the study and will finally propose how this thesis as well as VOT in Gujarati speakers of English more generally can be expanded for future research.

Overall, this thesis aims to investigate whether VOT transfer exists across three different generations of British-Gujarati speakers. Its main motivation is to add to the growing literature in the fields of phonetics and sociolinguistics.

2. LITERATURE REVIEW

2.1. Introduction

This chapter of the dissertation will present the literature that pertains to the various topics of this thesis.

2.2. Gujarati

Gujarati *(also spelt as: Gujrathi, Gujerathi and Gujerati)*[1] is an Indo-Aryan language from the Indo-European family that originates from the state of Gujarat in India. Located in the westernmost part of the Indian peninsula, Gujarat is also known as the Jewel of Western India due to its vibrant and flamboyant culture (Discoveredindia.com, 2017). According to The Eighth Schedule of the Indian Constitution (2003), Gujarati is one of the twenty-two *scheduled languages* of India. Such languages within this list are entitled to receive extensive development and growth under the obligation of the Government of India[2]. The repute of Gujarati is increased by the fact that it holds official status in Gujarat (Ministry of Minority Affairs, 2014). Over 99% of the language's speakers reside in various cities within Gujarat as well as other bordering Indian states, such as Maharashtra and Rajasthan (Lewis, 2009). Smaller communities can be found in neighbouring countries within the Indian subcontinent (e.g. Pakistan and Bangladesh) and South and East Africa. However, the most sizeable and long-established communities from outside of India are notably located in the UK and USA. In total, Gujarati is spoken by 46.6 million people worldwide (Lewis, 2009), classifying it as the 23rd most commonly spoken language (Grimes, 1996). Figure 1 below illustrates the location of Gujarat within India as well as the official language of each geographical state. However, Gujarati, as well as all the other official languages, are not confined to solely one state and are typically spoken in other places as well (Lewis, 2009).

It is also worth noting that the statistics gathered regarding the number of Gujarati speakers in India only refer to native speakers. There may well be large numbers of second-language speakers of Gujarati that have been unaccounted for.

[1] For the purposes of clarity, only the spelling *Gujarati* will be used for the remainder of this thesis.
[2] The Government of India is required to take measures to ensure that scheduled languages 'grow rapidly in richness and become effective means of communicating modern knowledge' (Official Languages Resolution, 1968).

Fig. 1: A map of the languages that are spoken in India. The Gujarat state, located in Western India, has been circled. (Sarver, 2014).

As indicated above, Gujarati is located on the western coast of India. As it is the sixth largest state in the country, there are undoubtedly a number of dialects spoken within this area. However, assessing the number of dialects that make up Gujarati is rather unclear. A common phrase often spoken by the older generations is “the dialect changes after every twelve villages” (Fatimah Rawat, p.c.). In other words, the dialects of Gujarati slowly yet gradually change from village to village. This creates a dialect chain in which linguistic differences between dialects can only be recognised when comparing those that are separated by a wide distance. Despite this, Gujarat can still be divided into separate districts, as decided by the Government of Gujarat. Such decisions were originally based on wider geographic regions and their respective climates and cultures. These regions are: Kutch (dry and desert-like climate), Saurashtra (dry but soil with high moisture), North Gujarat (dry air and less moisture in soil), Central Gujarat (similar to North Gujarat but much less farming) and South Gujarat (humid air, albeit low) (discoveredindia.com, 2017). However, more recently, the formation of districts has been based upon political and administrative gains. For example, in 2013, the Indian government established 7 new districts in order to set up more administrative units and “decentralise the administration and make it more efficient and speedy” (Ghosh, 2013). In total, there are 33 districts that make up Gujarat. Each district is divided into sub-districts called Talukas of which there are 249 in total. Figure 2 below is a map of the 33 districts within Gujarat. The district Surat which is the focus for this thesis, has been circled.

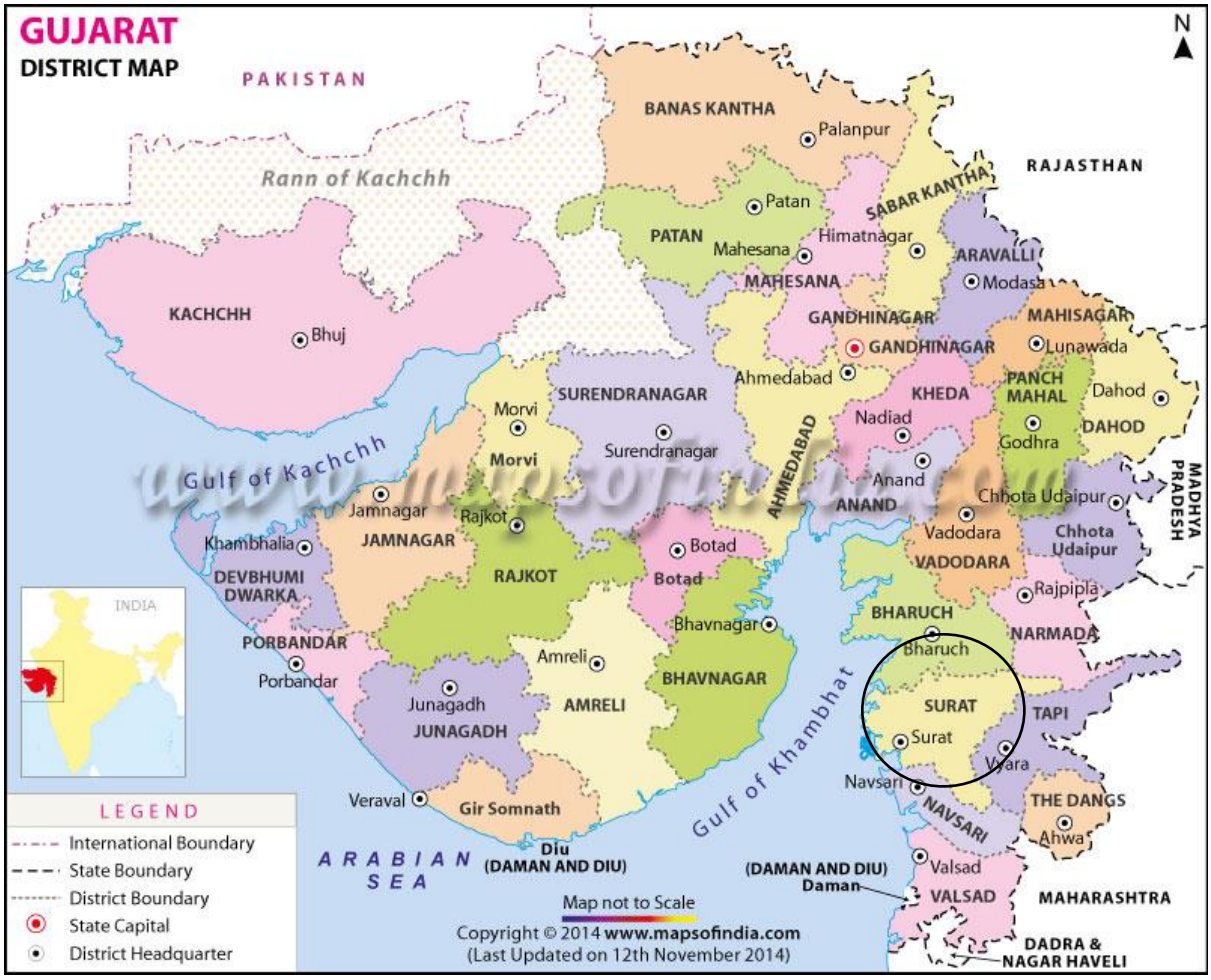

Fig. 2: A map of all the districts in the Gujarat state. The district of Surat has been circled (mapsofindia, 2014).

In Figure 2 above, Standard Gujarati has been the most discussed dialect by linguists. Though Standard Gujarati is not exclusive to any certain state, this dialect is heavily used in Ahmedabad and Ghandinagar[3] and therefore is widely recognised and utilised within media-related, governmental as well as educational domains. Following a thorough review of the linguistic literature, no prior research has examined Surati Gujarati – the dialect of Surat.

2.3. Gujarati in the UK

As recorded by the 2011 UK Census, Gujarati is the sixth most spoken language in England and Wales, as 213,000 people (0.4% of the population) reported it as their main language (ONS, 2013a). This is the fourth most popularly-spoken Indic language following Punjabi (c. 273,000; 0.5%), Urdu (c. 269,000; 0.5%) and Bengali (c. 221,000; 0.4%) (ONS, 2013a). However, according to the National Congress of Gujarati Organisations, the number of

[3] Ahmedabad is the former capital district of Gujarat whereas Ghandinagar is now the current capital. The standard form of Gujarati is used within these districts due to their administrative, educational and political positions.

Gujarati speakers in the UK is estimated to be 600,000 (NCGO, 2012). It is uncertain which of these statistics is the closest approximation of the actual number. However, it is probable that the higher figure of the two is more precise, due to the fact that the census only allowed for one main language. This consequently discounted those people who can speak and use Gujarati but do not perceive it to be their main language.

As this thesis will consider age as a variable, it is also worth looking at the statistics regarding the relationship between minority languages and age of such languages' speakers. Figure 3 below shows the percentage of all UK residents who did not consider English to be their main language.

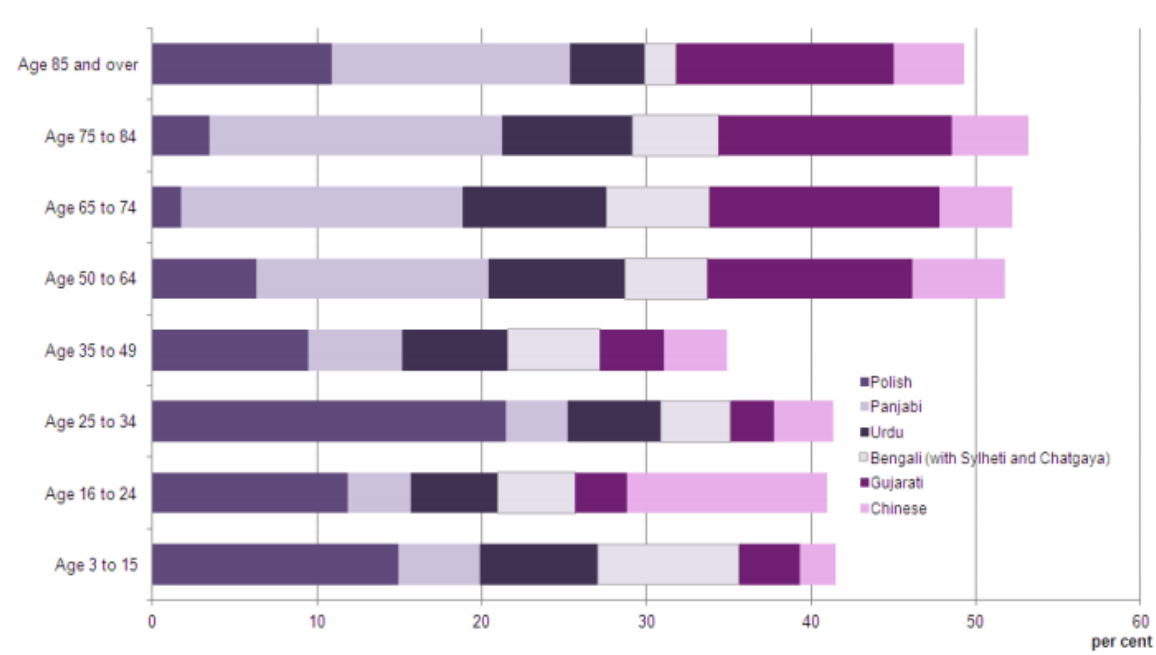

Fig. 3: A graph illustrating the percentage of residents whose main language is not English (ONS, 2013b)

As illustrated in Figure 3, the percentage of residents that consider Gujarati to be their main language is significantly higher within the 50 to 85 age brackets than it is for the residents aged 3 to 49. It is most likely that the majority of these speakers from the older generations are either immigrants who travelled from Gujarat to the UK or they are the first UK-born Gujarati residents. Gujarati was almost certainly the only language used within the household for these residents. Consequently, the older generations are more likely to consider Gujarati as their main language than those residents who have been raised bilingually. The pattern that is displayed regarding Gujarati bears the closest resemblance to Punjabi as this too illustrates the same result.

2.4. An overview of migration from Gujarat to the UK

For centuries, Gujarat has been renowned for its pivotal role within the merchandising domain. The natural harbours in the state, particularly in the Gulf of Khambatt, as well as the rivers that course down to the Arabian Sea have meant that Gujarat has notably

sustained trade production (Mattausch, 2001). Historical data explains commercial ties between Gujarat and several countries, such as Egypt, Bahrain and the Persian Gulf dating as far back as 1000 to 750 BC (Rao, 1985; Mapsofindia.com, 2017). The trading port in Surat, the second biggest city in Gujarat, had the first exporting link with Britain via the East India Company in the 17th century (Chaudhuri, 1999). A combination of these merchandising connections, the labour shortages after World War II in the UK (Messina, 2001), the release of The Indian Independence Act 1947[4] as well as the creation of the British Nationality Act 1948[5] all prompted significant migration from India to the UK (Communities and Local Government, 2009). The earliest migrants from Gujarat in the 1950s were males who sought employment to become financially stable enough to later bring over their families (Communities and Local Government, 2009). Such migrants favoured employment within the textile and mill industry, and thus settled in towns within Leicestershire, Lancashire and Yorkshire, such as Leicester, Preston, Bradford and Batley (see: NCGO, n.d.; Ember et al., 2005; The Yorkshire Post, 2016). Once the demand for textile industries started to weaken in the 1970s, a few Gujaratis dispersed and resettled in other parts of the country whilst many others chose to remain in the original town to which they migrated (Communities and Local Government, 2009). It is for this reason why Leicester has the highest proportion of Gujarati speakers in the UK (ONS, 2013c).

2.5. From Gujarat to Kirklees

Within a decade of The Indian Independence Act 1947, Gujarati males immigrated to West Yorkshire, particularly within the metropolitan area of Kirklees, which includes, but is not confined to, Batley, Dewsbury and Huddersfield (IMWS, 2007). The earliest statistics regarding Batley reports that in 1957 there were as few as 40 Gujarati males residing in the town (Communities and Local Government, 2009). The reason why Batley, an extremely small town that is often omitted from most maps, became the core for various expatriate communities to settle in was entirely on the premise of its employment opportunities. The town was at the forefront of textile manufacturing. Batley was deemed to be pioneers of the 'shoddy trade'[6] (Jubb, 1860, p. 3), both on a national and global level, which made the prospects of gaining work greater than other areas. Starting with 40 males in 1957, the Gujarati community has rapidly increased, as has other ethnic populations. Though there are no exact numbers released regarding Batley alone, we can determine the number of British Gujaratis from data regarding Kirklees as a whole. According to the 2011 UK Census, 4.9% of the population in Kirklees (20,797 people) are, or at least consider themselves to be, 'Asian British: Indian' (ONS, 2013d). This is 2.4% higher than the national average in England

[4] The Indian Independence Act (1947) triggered several issues, which include civil unrest, economic hardships and religious (Hindu and Muslim) discord. This increased migration to the UK in order to escape from such hardships (see Bates, 2003).
[5] The British Nationality Act (1948) gave all citizens of any British colony the right to work and reside in the UK.
[6] The shoddy trade is the name given to the industry in which old woollen rags were recycled to make a new material called *shoddy.*

and Wales. With regards to the geographical area in which these Gujarati speakers reside, Figure 4 below illustrates the distribution of this ethnic group.

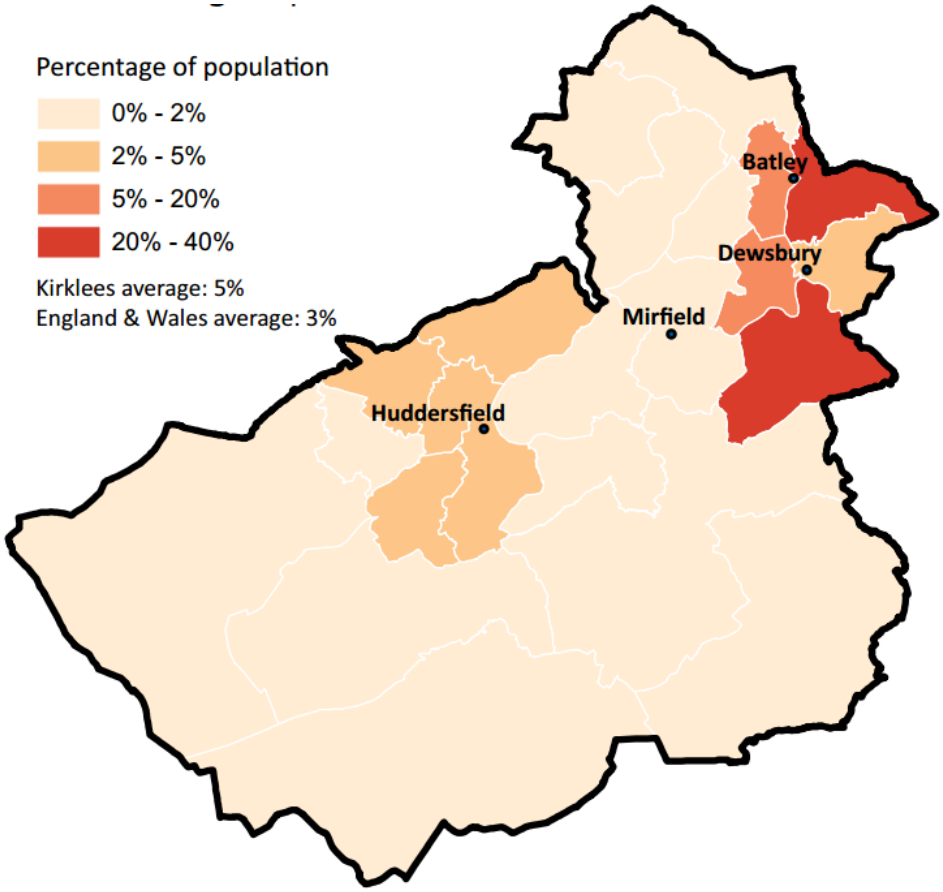

Fig. 4: A map illustrating the geographical distribution of Indians within the metropolitan borough of Kirklees (Jivraj and Finney, 2013)

This map of Kirklees in Figure 4 clearly indicates that the towns and surrounding areas of Batley and Dewsbury are home to the highest proportion of Indians within Kirklees. The Indian ethnic group make up the largest proportion of the population in Batley East (37%) and Dewsbury South (23%) (Jivraj and Finney, 2013), indicating that the majority of the 20,797 people who identified themselves as British Indian reside within Batley or Dewsbury. Nearly all of the Gujarati people residing in these towns are either from the Gujarati districts of Surat or Baruch (Communities and Local Government, 2009), and speak Surati or Baruchi Gujarati respectively. The lack of geographical scattering of Indian people across Kirklees, and instead the high-density clusters of such people in Batley and Dewsbury, has resulted in the establishment of an extremely strong community in which the residents have established their own businesses, mosques, shops and schools. These clusters of South Asians have meant that Batley and Dewsbury have the greatest proportion of people in

Kirklees who cannot speak English well (Jivraj and Finney, 2013). This potentially indicates that the strong community bonds allow people to get by without ever acquiring English.

2.6. Linguistic sketch of Gujarati

In order to understand the general overview of Gujarati, a brief linguistic sketch of the language needs to be presented in addition to the aforementioned demographic qualities.

2.6.1. Gujarati sound system

The phoneme inventory of Gujarati is comprised of eight vowels[7] and approximately thirty-seven consonants[8], the latter of which is displayed below in Table 1.

IPA Consonants				Place of Articulation															
				Bilabial		Labiodental		Alveolar		Post alveolar		Retroflex		Palatal		Velar		Glottal	
Gujarati - Marathi				VL	V	VL	V	VL	V	VL	V	VL	V	VL	V	VL	V	VL	V
Manner of Articulation	Plosive	NA	IPA	p	b			t	d			ʈ	ɖ			k	g		
			G - M	પ - प	બ - ब			ત - त	દ - द			ટ - ट	ડ - ड			ક - क	ગ - ग		
		A	IPA	pʰ	bʱ			tʰ	dʱ			ʈʰ	ɖʱ			kʰ	gʱ		
			G - M	ફ - फ	ભ - भ			થ - थ	ધ - ध			ઠ - ठ	ઢ - ढ			ખ - ख	ઘ - घ		
	Nasal	NA	IPA		m				n				ɳ		ɲ		ŋ		
			G - M		મ - म				ન - न				ણ - ण		ઞ - ञ		ङ - ङ		
	Trill	NA	IPA						r										
			G - M						ર - र										
	Tap / Flap	NA	IPA						ɾ										
			G - M						ર - र										
	Fricative	NA	IPA			f		s		ʃ								h	ɦ
			G - M			ફ - फ		સ - स		શ - श								હ - ह	હ - ह
			G - M							ષ - ष									
	Approximant	NA	IPA				ʋ								j				
			G - M				વ - व								ય - य				
	Lateral Approximant	NA	IPA						l				ɭ						
			G - M						લ - ल				ળ - ळ						
Affricates		NA	IPA	ts/tʃ	dz	dz	kʃ		A	IPA	tsʰ								
			G - M	ચ - च	જ - ज	ઝ - झ	ક્ષ - क्ष			G - M	છ - छ								

VL-Voiceless	NA-Non Aspirated	G-Gujarati
V-Voiced	A-Aspirated	M-Marathi

Table 1: An IPA description of all the consonants in Gujarati and Marathi[9] (Patil et al. 2012, p. 180)

Table 1 indicates that Gujarati has an extensive phoneme inventory (Patil et al., 2012, p. 180). This includes sixteen stops, five nasals, five fricatives and five affricates. As this thesis will concentrate exclusively on stop sounds, it is important to note that 43% of all the sounds within the language are plosives. This exceptionally high percentage makes clear that Gujarati includes many plosives. Like many other Indic languages, such as Urdu, Hindi and Marathi, Gujarati makes use of the retroflex manner of articulation, which further increases the depth of the plosive, phonemic inventory.

Furthermore, the Gujarati language continues to mirror the majority of the other Indo-Aryan among Indo-European languages in terms of its four-way contrast in voicing and aspiration for stops and affricates. As Table 1 illustrates, Gujarati distinguishes stop and

[7] Linguists report that the conservative dialects have eight vowels, whereas other dialects, such as Saurashtra, have six vowels (Firth, 1967, p. 231-232).

[8] There are disputes between linguists regarding the exact number of consonants, particularly with regards to the glottal fricative /h/, the palatal nasal /ɲ/ and the velar nasal /ŋ/.

[9] The consonants for both Gujarati and Marathi are exactly the same. The only difference displayed in Table 1 is the phonetic transcription of each sound.

affricate sounds as voiced aspirated, voiced unaspirated, voiceless aspirated and voiceless unaspirated, all of which are significant in altering meaning within speech. Though a minimal quadruplet does not exist which presents all four contrasts, Gujarati does have numerous sets of minimal pairs that clearly display the voicing and aspiration distinction. For example, with regards to the voiced unaspirated, Gujarati speakers would be heard using:

Gujarati - [બાર]
Transliteration - /bar/
Grammar - num.sg.
Translation - 'twelve'

Such speakers would also be heard articulating the first phoneme with aspiration in order to change its meaning, as shown below:

Gujarati - [ભાર]
Transliteration -/b^{h}ar/
Grammar - n.sg.
Translation - 'burden'

Likewise, in relation to the voiceless phonemes, the following minimal aspiration-contrast pair can be found to be used in the language.

For voiceless unaspirated:

Gujarati - [કાન]
Transliteration - /kan/
Grammar - n.sg.masc.
Translation - 'ear'

And for voiceless aspirated:

Gujarati - [ખાન]
Transliteration - /k^{h}an/
Grammar - n.sg.masc.
Translation - 'Khan[10]'

[10] A popular surname used widely in the Indian sub-continent.

2.7. Linguistic sketch of English

This section examines the sound system of the second language concerning this thesis: English.

2.7.1. English sound system

English is made up of twenty-four consonants and twenty vowels, the former of which is demonstrated below in Table 2.

	Bilabial	Labiodental	Dental	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Plosive	p b			t d			k g	
Affricate					tʃ dʒ			
Nasal	m			n			ŋ	
Fricative		f v	θ ð	s z	ʃ ʒ			h
Approximant	(w)				r	j	w	
Lateral approximant				l				

Table 2: An IPA Chart of all the consonants in British English (Roach, 2004).

Unlike Gujarati which has sixteen stops, English has only 6 (nearly three times less than Gujarati). Although it is only the retroflex position that is not present in the English language, the reason for such a contrast between the number of stop sounds in English from Gujarati is due to the fact that the former does not mark a contrast between aspiration in order to change meaning. Thus, each stop sound, like most sounds from other manners of articulation, are divided into voiceless and voiced. Furthermore, English does not form a contrast in terms of tone or phonation. It is “characterised only by consonantally conditioned, prosodic and sociolinguistic variations in voice quality” (Krieman et al.) This means that English only differs its voice quality on a supra-lexical level. Thus, English does not use voice quality to differ otherwise identical lexical items, but rather speakers of English use voice quality as intonational features, which vary depending on the sociolinguistic context.

2.8. Voice Onset Time

As this thesis will specifically consider the influence of Gujarati on English plosives, more details need to be discussed regarding voice onset times of the plosives in each language. Firstly, however, a broad overview of voice onset time is provided below.

During the second half of the twentieth century, researchers began to heavily explore the voicing contrast in stops (see: Fischer-Jorgensen, 1954; Lisker and Abramson, 1964; Docherty, 1992; Bohn and Flege, 1993). Lisker and Abramson's (1964) famous cross-language study of word-initial stops across eleven languages provided us with one of the earliest explanations of voice onset time. They defined it as "the time interval between the burst that marks release of the stop closure and the onset of quasi-periodicity that reflects laryngeal vibration" (Lisker and Abramson, 1964, 422). Following on from their cross-language report, several other scholars have continued to investigate the topic further, including Keating et al. (1983) who explored voice onset time in fifty-one languages, and Cho and Ladefoged (1999) who studied eighteen languages.

Lisker and Abramson (1964) proposed that voice onset time, based on the eleven languages in their study, can be divided into three groups: -125ms to -75ms, 0ms to +25ms, and +60ms to +100ms (1964, p. 403), which can also be labelled as prevoicing/voicing lead, short lag and long lag respectively[11]. However, as this arrangement clearly shows, there are gaps between the three ranges, and thus not all languages can be incorporated into one of the three classifications. Consequently, Cho and Ladefoged (1999) considered the groupings of VOT in further detail, and proposed that this acoustic cue is more accurately organised within four categories: unaspirated (mean VOT=30ms), slightly aspirated (mean VOT=50ms), aspirated (mean VOT=90ms), and highly aspirated (mean VOT=over 90ms) (Cho and Ladefoged, 1999, p. 223).

Researchers have used these aforementioned studies as the premise and foundation for their own reports. English, due to its undeniably widespread presence on a global level, has been the most investigated language regarding VOT. However, scholars have also examined other languages, some of which include French (Caramazza et al. 1973), Spanish (Flege and Hammond, 1982; Fellbaum, 1996) and Arabic (Khattab, 2000; Flege and Port, 1981). In spite of VOT being a prevalent acoustic parameter to research, there is a lack of VOT literature for Gujarati. From the limited Gujarati data that is accessible as well as from the substantial information regarding English, details of the categories into which the two languages fit will be discussed in Sections 2.7.2. and 2.8.2. respectively.

In spite of voice onset time being a prevalent acoustic measurement used to calculate the timing of voicing in stops, there are several scholars (see Caramazza et al., 1973; Bohn and Flege, 1993) who have proposed that VOT may not be an essential acoustic key to the perception of voicing in stops as it was originally deemed to be. Docherty (1992) claims that VOT is an excessively limited acoustic cue as it merely concentrates on voicing in

[11] Lisker and Abramson (1964) defined the terms voicing lead and voicing lag. The former entails VOT measurements that occur before the release burst and thus have negative values, whereas the latter entails VOT measurements that occur after the release burst and thus have positive values.

word-initial stops, and thus disregards those stops that are in word-medial and/or word-final positions. Nonetheless, even with the studies that suggest the inadequacies of voice onset time, the wider literature still deem it to be a chief acoustic cue that phonologically distinguishes between voiced and voiceless stops. As this thesis will exclusively consider stops in word-initial position, Docherty's (1992) criticism will not apply nor affect this thesis.

Using the above understanding of voice onset time, the following two sections (2.9 and 2.10) will outline the acoustic qualities of VOT in Gujarati and English.

2.9. The acoustics of VOT in Gujarati

Although the phonemic characterisation of Gujarati exists (see: Nair, 1979; Pandit, 1961; Turner, 1921), scholars have only begun to study Gujarati on a phonetic level within the last few decades. This is partially due to Gujarati having fewer speakers than its Indian sub-continental counterparts, as well as the fact that most speakers of the language are concentrated within the state of Gujarat as opposed to being geographically dispersed, making it difficult to study the language within bilingual contexts. The principal focus of the studies that have been undertaken relate to phonation, due to the fact that Gujarati is one of only two languages that marks a contrast between modal and breathy phonation across vowels and consonants (see: Esposito & Khan, 2012; Khan, 2012; Vupalla and Bhaskararoa, 2014; Fischer-Jørgensen, 1967). Examples of minimal pairs that show phonation contrast in vowels are displayed below in Table 2.

Modal voice		**Breathy Voice**	
IPA	**Gloss**	**IPA**[12]	**Gloss**
/pɔr/	Last year	/pɔ̤r/	Dawn
/mɛl/	Dirt	/mɛ̤l/	Palace
/ar/	Obstruction	/a̤r/	Bones

Table 3: Minimal pairs of modal and breathy voice in Gujarati.

As Table 2 illustrates, Gujarati possesses two sets of vowels. Both sets are entirely the same with regards to the position from which the vowels are articulated. However, one set is for modal phonation whilst the other is for breathy phonation. With regard to VOT, there is a possibility that the vowel phonations may affect consonants, therefore this has to be born in mind through the course of this study.

The majority, albeit not all, of the literature that concentrates on phonation in Gujarati is rather recent. As a result of such a distinctive characteristic being present in the language, scholars have paid thorough attention on this topic. However, other acoustic

[12] The IPA diaeresis of two dots under the vowel represents the breathy nature of the phoneme. Note that this diacritic is only for breathy vowels. The diacritic for breathy consonants is X^h.

attributes such as fundamental frequency (f0), intensity and segmental duration, although they may be common across all languages, are yet to be explored within Gujarati.

One of the acoustic attributes that is virtually absent from the phonetic literature of Gujarati, and yet is "the most extensively investigated and most voluminously reported" (Hewlett & Beck, 2006) attribute in other languages, is voice onset time. The only research paper that has been found within the literature is a short report by Rami et al. (1999) who investigated the burst frequency and voice onset time of four velar stops in Gujarati: the unaspirated /k/and /g/ and the aspirated /k^h/ and /g^h/. Despite the narrow focus on only one place of articulation, the results of their study can still be used as a base rate with which we can compare the findings of this thesis, or at least the findings for the same phonemes. Rami et al. (1999) discovered that the mean VOT for the aspirated voiceless plosive /k^h/ was considerably longer than the unaspirated voiceless stop /k/. However, the findings for the voiced plosives /g/ and /g^h/ were too close together to suggest any statistical difference. Table 3 below displays the mean VOT of their results.

	Velar stop consonant			
	Voiced		Voiceless	
	/g/	/g^h/	/k/	/k^h/
Voice onset time	−37.3 (9.8)	−29.2 (7.3)	40.6 (21.1)	75.0 (33.2)
Burst frequency	1555 (198)	1531 (176)	1451 (146)	1427 (140)

Table 4: Mean VOT (and standard deviations in brackets) for the four velar Gujarati stops (Rami et al., 1999, p. 3737).

As can be seen in Table 3, both the voiced unaspirated and aspirated plosives /g/ and /g^h/ respectively have negative VOT values and thus is voicing lead. Whereas, the voiceless unaspirated and aspirated /k/ and /k^h/ are situated on the positive region of the VOT continuum, and thus are voicing lag. According to Cho and Ladefoged's (1999) classification of stop sounds, the unaspirated phoneme /k/ lies almost exactly in between the 'unaspirated' (mean VOT of 30ms) and the 'slightly aspirated' (mean VOT of 50ms) categories. The aspirated /k^h/, however, is situated between the 'slightly aspirated' and 'aspirated' (mean VOT of 90ms) groups, yet there is a slight leverage towards the latter. Thus, it is evident that even Cho and Ladefoged's (1999) categorisation of VOT values are still too wide to fully understand where Gujarati stop sounds are positioned on the VOT continuum.

Furthermore, it is indeterminate whether Rami et al.'s (1999) findings will be replicated in this thesis due to the lack of control of the demographic variables of gender and age. With regards to the former, Rami et al. (1999) used six males in their investigation

as opposed to only two females; for the latter, the participants had a mean average age of 33.6 years.

Despite the acoustic literature of Gujarati being rather scarce, other genetically-related languages[13] from South Asia and their acoustic properties have been extensively investigated. Lisker and Abramson (1964), for example, in their cross-language study of voice onset time examined Hindi as well as Marathi, both of which are Indo-Aryan languages. Table 4 below shows the voice onset time measurements that were found for Hindi in their study.

			Voice Onset Time (ms)
Voiced Stop Consonants	/b/	Unaspirated	-85
	/b^h/	Aspirated	-61
	/d/	Unaspirated	-87
	/d^h/	Aspirated	-87
	/g/	Unaspirated	-63
	/g^h/	Aspirated	-75
Voiceless Stop Consonants	/p/	Unaspirated	13
	/p^h/	Aspirated	70
	/t/	Unaspirated	15
	/t^h/	Aspirated	67
	/k/	Unaspirated	18
	/k^h/	Aspirated	92

Table 5: Mean measurements of Voice Onset Time in Hindi (Lisker and Abramson, 1964)[14].

Table 4 illustrates that all voiced stops, regardless of whether they are aspirated or not, have negative voice onset time values. What is interesting in Lisker and Abramson's (1964) study is that their findings for Marathi were extremely similar to those of Hindi. Thus, there may be a possibility that, due to Hindi and Marathi – two linguistically-related languages – bearing the same results, Gujarati would too yield similar results.

Furthermore, the results from the pilot study that was undertaken at an undergraduate level is a further insight into what results may ensue from this thesis. The pilot study – which explored the production of only English alveolar stops by British-Gujarati speakers across three generations – found that, as the generations became younger, the VOT for voiceless stops became longer. Additionally, the target words articulated in isolation had a higher VOT than when articulated within a sentence. The results for the voiced stops, however, were much less clear. Nevertheless, this experimental study does suggest that there may be a negative correlation between VOT and age.

[13] This refers to the languages within the Indo-Aryan branch of Indo-European languages. The Indo-Aryan branch is the dominant language family within the Indian subcontinent.

[14] The findings for the retroflex phonemes /ʈ/, /ʈ^h/, /ɖ/ and /ɖ^h/ have been omitted from the table as they will not be analysed in this thesis.

2.10. The acoustics of VOT in English

More so than other languages, researchers have investigated English in relation to various acoustic parameters. With regards to voice onset time, Keating (1984) reports that "English divides up the VOT continuum with some lead values but mainly short lag vs. long lag" (1984, p. 43). This reinforces the findings of Lisker and Abramson (1964) who also states that English stops present values on both sides of the VOT continuum.

Table 6 below displays the proposed VOT values for English voiced and voiceless stop sounds. The measurements of both American English (Lisker and Abramson, 1964) and British English (Docherty, 1992) are shown in order to fully explain the ranges that span across the VOT continuum as well as to recognise the differences between the two dialects.

		Voice Onset Time (ms)	
		Lisker and Abrasmon (1964) [AE]	**Docherty (1992) [BE]**
Voiced Stop Consonants	/b/	1	15
	/d/	5	21
	/g/	21	27
Voiceless Stop Consonants	/p/	58	42
	/t/	70	64
	/k/	80	62

Table 6: Mean measurements of VOT in American English and British English

As can be seen in the table, there is a general pattern that can be identified across the dialects. Both American English and British English have short lag values for voiced stop consonants and thus all fall into the 'unaspirated' category established by Cho and Ladefoged (1999), whilst all the voiceless stop consonants are located in the upper end of the 'slightly aspirated' region. In spite of the fact that both dialects fall within the same area, British English has higher VOT measurements for voiced stops and lower VOT measurements for voiceless stops than American English. As all of the participants for this thesis speak a variation of British English, it is expected that their VOT values for English will bear resemblance to Docherty's (1992) results, and thus it is apt to use his findings as a comparison for this thesis.

2.11. Bilingualism

The nature of VOT values belonging to a certain language is not as straightforward when numerous languages are spoken within the same vicinity and by a single speaker. As briefly mentioned in Section 2.5., Batley is home to various ethnic communities which includes the Gujarati, Punjabi, English and Polish. Therefore, it is undoubtedly the case that the town is a bilingual setting in which there is inevitable contact.

2.11.1. Language contact and second language acquisition

Despite each ethnic group in Batley having their own sense of solidarity with others who share the same heritage background, there is no doubt that the geographical setting has created a physical community to which everyone, regardless of culture, belongs. As a result, language contact in the form of *koineization* often arises, as new dialects or *koines* become the product of years of community interaction[15]. Mufwene (2001) aptly describes this condition of contact as a 'feature pool' in which various features of several languages are brought together. Speakers can then obtain features from this pool, in a somewhat assorted fashion, in order to eventually form a new variety or language.

Scholars in the field of language contact in relation to second language acquisition have differed between language transfer and language interference. The former appears to be 'more controlled and often socially loaded occurrence' (Newlin-Åukowicz, 2014) of shift of a language feature from a speaker's first language (L1) to their second language (L2) (Sankoff, 2002). The latter, however, is more uncontrolled and unintentional and typically reflects a speaker's process of acquisition of a second language. In many cases, a speaker's late acquisition of an L2 means that they continuously speak the L2 with a foreign accent, both in its production and perception. Work in this field has argued that interference from an L1 does not usually exist in the speech of successive generations (Thomason & Kaufman, 1988), however these generations may choose to intentionally retain and transfer elements of interference rather than losing it.

However, interference or transfer from an L1 to an L2 is not the only possibility that has been proposed by scholars. As mentioned in the Introduction, there are reports of bidirectionality where an L2 language influences an L1 or both influence each other simultaneously. According to Flege's Speech Learning Model (1995), "phonetic categories established in childhood for L1 sounds evolve over the life span to reflect the properties of all L1 or L2 phones identified as a realization of each category" (Flege, 1995, p. 239). In other words, the phonetic subsystem of an L1 is not in a rigid nor fossilised state, but rather is constantly changing and is contingent on the new languages that are acquired. However, if the L2 becomes the dominant language, it may become free of all interference from the L1 and consequently lead to the attrition of the first language. The final, albeit rare, possibility is that the L1 and L2 languages do not interfere at all. The two languages do not share the

[15] Though the formation of new languages is common, it is unlikely to occur in a British setting in which English is still used in the wider domain.

same phonetic space, and thus each language can be compared to the speech of monolingual speakers.

2.11.2. Language contact and VOT

A bilingual environment may sometimes affect only a few particular aspects of a language. Researchers have investigated the influence of voice onset time by bilingual speakers in the past few decades. Flege (1987) examined French-English bilingual speakers residing in their non-heritage country. He found that native French speakers living in the USA articulated the voiceless stop /t/ with a longer, more 'English-like' (Flege, 1987, p.47) VOT than French monolinguals, and the American speakers residing in France produced the same plosive with a French-like VOT. Likewise, Major (1992) discovered that his native American-English subjects living in Brazil and who learnt Portuguese as a second language began to shorten down their VOT, i.e. more Portuguese-like, in their English speech. Khattab (2000) reinforces this as she found that her Arabic-English bilingual children used the characteristic English VOT for both English and Arabic stops, the latter of which suggests that such VOT influence occurs at any age. All three of these studies strengthen the theoretical concept that an L2 influences an L1, particularly upon migration to a country in which the L2 is the dominant language, although this is not the case in Khattab's (2000) study.

With regard to languages from the Indian subcontinent, McCarthy et al. (2003) discovered that Bengali immigrants who moved to the UK in their early 20s articulated English stops with a Bengali-like VOT, whereas those Bengalis born in the UK had the same production of stops as British-English speakers. The British-born speakers correlate with the aforementioned concept of an L2 interfering with an L1, however those Bengali speakers who migrated to the UK provide support than an L1 can and does interfere with an L2, particularly when the L2 is acquired late. Heselwood and McChrystal (1999) reported that British-born Punjabi speakers living in Bradford used voiceless stops like those speakers who had learnt Punjabi in Pakistan. However, the voiced Punjabi stops were uttered in the same way as the English voiced stops. Heselwood and McChrystal (1999) put this down to the influence English (the L2) has on Punjabi (the L1) as well as the broad linguistic changes that are occurring in the Bradford Punjabi community.

The majority of the literature, therefore, suggests that a speaker's VOT value of their L1 can be influenced by other languages spoken in the community in which they live, however there is still the possibility that the L1 can interfere the L2.

2.12. Summary

The literature in this chapter has provided a clear report regarding the linguistic background of English and Gujarati as well as how the two languages have geographically and historically come into contact. In particular, the acoustic information for each of the two languages has been discussed. Studies undertaken by Lisker and Abramson (1964), Docherty (1992) and

Rami et al. (1999) indicate that English has a high positive VOT (long lag) for voiceless stops and low positive VOT (short lag) for voiced stops, whereas Gujarati has low positive VOT (short lag) for voiceless stops and negative VOT (voicing lead) for voiced stops. The literature has also displayed that VOT can be affected by other languages. Using information from this chapter, the next chapter will consider the methods that were used in order to test these aforementioned ideas further.

3. METHODOLOGY

This chapter sets out the methods used in this project. This will begin in Section 3.1, which will discuss the criteria that was applied for recruiting participants. The focus will then move onto the sampling process (Section 3.2), the data that was elicited (Section 3.3) and how it was collected (Section 3.4), the recording technicalities (Section 3.5), the impact of the fieldworker (Section 3.6), and lastly the methods used to examine the data (Section 3.7).

3.1. Participants

Participants in this experiment are defined based on linguistic and demographic qualities, which include: language background, sex, age and region. Each of these will be explained in the subsequent subsections. There were in total twelve participants that were chosen for this study. Prior to gathering these participants, ethical approval was obtained from the School of Music, Humanities and Media's Research Ethics Committee.

3.1.1. Language Background

All participants were required to be of Gujarati ethnicity, which is defined here as their parents/grandparents being born in Gujarat, India[16]. Participants were also expected to have native-like fluency in Gujarati. Although the term 'native' is rather subjective, the general requirement was that the participants could all understand Gujarati when it was spoken to them as well as sustain a conversation in the language[17].

Additionally, it was necessary that all the participants had competency in speaking and understanding English. This aspect of the criteria was easily met by the majority of the participants, due to being born in the UK and therefore receiving maximum exposure to the language. However, it was rather problematic obtaining first-generation participants (those born in India) who could speak English, due to the fact that their exposure to English only started upon their arrival to the UK, which was nearly always past the critical age of acquisition (Lenneberg, 1967). As a result of this, the standard of English proficiency was lowered for the first-generation, which allowed me to obtain speakers who had basic abilities in the language as opposed to the second and third generation of speakers who had maximum proficiency.

[16] Note that whilst selecting participants it was not essential that the subjects considered themselves to be of Gujarati ethnicity. This is because the term *ethnicity* is rather subjective, and there may be many people whose parents/grandparents are from Gujarat but they do not deem themselves to be of Gujarati ethnicity. Therefore, it appeared to be more fitting to categorise Section 3.3.1. as 'language background'.

[17] Here, 'sustain a conversation' refers to a fluent conversation with a native speaker of Gujarati (born in Gujarat).

It is also worth noting, that despite not being a requirement, many of the participants were fluent in Urdu, mainly due to its links to the Islamic religion[18].

3.1.2. Sex

Variables were restricted to only focus on female productions, and results will only be discussed in relation to females and not males. The reason for this was not only to control the investigation but also because of the researcher's access to female speakers within her circle of acquaintances.

3.1.3. Age

In order to assess the differences in this apparent-time research (see: Labov, 1963; Eckert, 1997, Bailey et al., 1991), three age categories were formed and in each of which were four participants. The three categories were:

- 60+ years old
- 39-46 years old
- 18-25 years old

These age groups were purposely chosen due to the fact that they are loosely based on three generations. Before discussing each age group, it needs to be noted that scholars disagree with who is classified as 'first generation'. In the age brackets above, the members of the 60+ group migrated to the UK in their early-mid teenage years with their parents. According to the Oxford English Dictionary, the term *generation* refers to '...designating a member of the first (or second, etc.) generation of a family to do something or live somewhere; spec. designating a **naturalized immigrant or a descendant of immigrant parents**…' (OED, 2009). This suggests that those who originally migrated and have now passed on as well as those in the 60+ age bracket above can *both* be considered to be first generation. Scholars have attempted to eliminate this ambiguity by coining the term '1.5 generation' (Rumbaut, 2004), which refers to those people who migrate during their early teens, i.e. the 60+ age group. However, in order to avoid confusion, the 60+ age bracket will be referred to as the first generation, the 39-46 age bracket will be referred to as the second generation, and the 18-25 age bracket will be referred to as the third generation.

With regards to Romaine's (1995) typology of bilingualism, all three generations parallel Type 6 of Romaine's model, named 'Mixed Languages' (1995, p. 185). In this category, the parents are bilingual speakers (of English and Gujarati) and code-mix both languages within the home. However, it is likely that the first generation, due to their late

[18] Urdu is written in the Perso-Arabic script, which originates from the script used in the Quran, the holy book for Muslims. All of the participants in this study identified as Muslim.

arrival to the UK, also correspond with Romaine's (1995) Type 3: 'non-dominant home language without community support' (1995, p.184). Typically, these speakers are usually migrants who do not speak the dominant language of the community in which they reside. It is likely that despite being able to speak basic English, the first generation use their native tongue and communicate with their children in only Gujarati

The selection of these three age brackets is extremely similar to the age range that was chosen for the pilot study. The span of the age groups for both the first generation (60+) and the third generation (18-25) have remained identical for this thesis, and representative of their respective generations. However, the age bracket for the second generation in the pilot study was fixed as 33-40 years. Thus, there was a sizeable twenty years between the first and second generation and merely eight years between the second and third. This research has amended the arrangement of this bracket so as to position the second generation exactly in the middle of the age continuum, thus there are exactly fourteen years between each category. Consequently, the new gaps between each age category are broad, yet are not too wide that a complete generation is overlooked.

3.1.4. Region

The participant pool was further narrowed by the region in which the participants resided. All of the participants lived exclusively in the town of Batley (See Section 2.5 for more information), and have done so their entire lives. This is with the exception of the first generation that moved from Gujarat. The reason for choosing such a town is not only due to its proximity to the researcher, but also because Batley conveniently has no major ethnic community from outside of the Indian subcontinent. This works well in our favour for an accurate gathering of results due to the fact the prospects of recognising the influence of Gujarati and English upon each other is much higher than it would be in major cities such as London in which the linguistic diversity is so vast that it becomes problematic to ascribing changes to one certain language.

3.1.5. Naming

With the purpose of maintaining participant anonymity, the names of all the participants were changed to names that are used within the British-Muslim community. Each of the twelve participants were ordered by age and subsequently given names alphabetically from A to L.

3.2. Sampling

Using the above criteria, a total of twelve participants were gathered. The sampling process was a combination of convenience sampling as well as random sampling. As I, as the

researcher, am of Gujarati ethnicity, the first point of contact that was explored was family and friends within the immediate vicinity[19]. Following on from this, the process had a snowball effect as some of the participants put forward their family members/friends, providing that they met the set criteria.

Figure 5 below exemplifies the distribution of participants within Batley according to their postcodes. The three colours illustrate each of the three generations that are considered in this thesis.

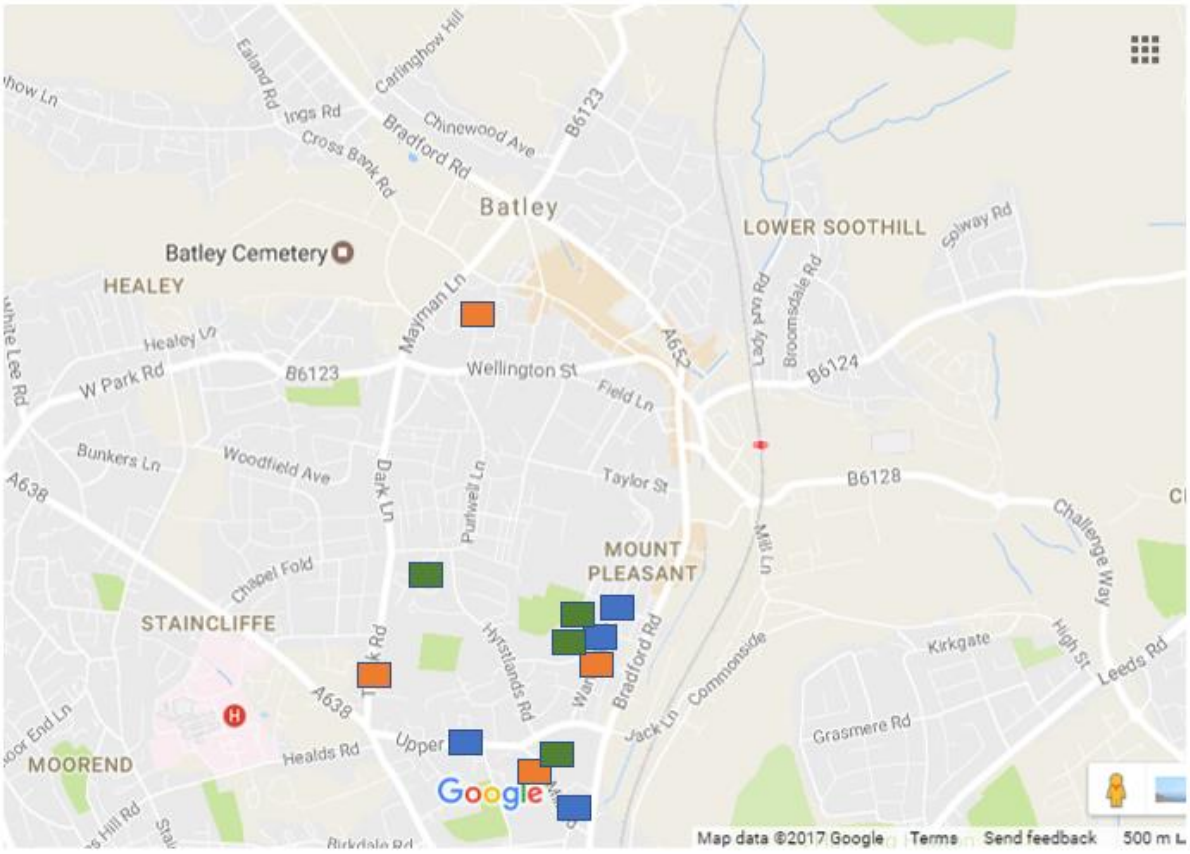

Fig. 5: A map illustrating the distribution of participants in Batley. Blue squares signify the first generation, orange squares signify the second generation and orange squares signify the third generation (Adapted from Google Maps, 2016).

3.3. Data

This section details the data collected to elicit stop consonants for analysis.

Unlike the pilot study, which due to the size of the experiment could consider only the alveolar stops /t/ and /d/, this thesis investigates all of the stops that are present in English.

[19] The same family and friends that were involved in the pilot study were not recruited again for this thesis.

This includes all of the English voiceless and voiced bilabial, alveolar and velar stops: /p b t d k g/.

In order to obtain accurate VOT measurements, four word-initial[20] target stimuli were selected per stop sound. Thus, as there are six stop sounds in English, a total of twenty-four words were chosen to be presented to the participants. Taking inspiration from Esposito and Khan (2012), all of the target stimuli were derived from ordinary contexts, such as food items or objects within the home, and belonged to no specific semantic field. This was purposely the case so that the familiarity with the words would discourage the participants from using a formal register.

Below in Table 7 is the list of the twenty-four words that were displayed to the participants.

	Word 1	**Word 2**	**Word 3**	**Word 4**
/p/	pen	paper	pyjamas	picture
/b/	biscuit	bag	banana	bottle
/t/	table	towel	train	tomato
/d/	dog	doctor	dress	door
/k/	kitchen	cat	carpet	key
/g/	garden	gate	goat	grapes

Table 7: A table of all twenty-four words shown to the participants.

To parallel many previous studies (see Lisker & Abramson, 1967; Adams, 1987; Khan, 2012), the method used to gather data was by placing the target stimuli within sentences[21]. Three sentences were concocted for each word, and each sentence was written on a flashcard[22]. Each of these cards was individually presented to the subjects who were simply required to read the word out loud. The sentences were all on average three to five words long so as to accommodate for the first generation's basic abilities in speaking English.

The positioning of the target stimuli differed in each of the three sentences. In the first sentence, the target stimuli were placed at the beginning of the sentence as the first word. In order to form grammatical sentences, all of the nouns were pluralised so that they could follow a Subject-Verb-Complement structure. Examples include: *'Towels are soft' and 'Gardens are green'* (See Appendix 8.1 for a list of all of the sentences). The second set of sentences for each of the target stimuli required that the words were positioned at the midpoint of the sentence. Every target stimulus in this set of sentences was preceded by a word ending in a voiceless plosive, for example *'Tha**t** dog is big'* and *'The to**p** garden is clean'.* With regards to the last set of sentences, the target stimuli were placed at the end of the sentence and were all preceded by a word ending in a voiceless fricative, for example *'I*

[20] Word-medial and word-final stop sounds were not investigated in this thesis.
[21] These sentences, although are grammatical, were not all semantically appropriate.
[22] The material used for these flashcards were thick card in order to minimise unnecessary background noise, which would have been the case if thin paper was used.

*like bo**th** pictures' and 'Get yourself paper'.* The purpose of creating three sets in each of which the target stimuli were positioned at various points within the sentence was to create a comprehensive understanding of the impacts of structure on Voice Onset Time.

This procedure of collecting data through sentences is a developed technique from that which was employed in the pilot study. In this preliminary experiment, the participants were asked to use the target stimuli in a sentence of their own formation. However, this proved to be rather problematic as this entrustment of power to the speakers caused them to feel self-conscious[23]. Furthermore, several of the voiced word-initial tokens collected from the pilot study had to be discounted due to the fact that the VOT of the stop sound in the target stimuli could not be measured as there was continued voicing from the previous word. Therefore, for the purpose of thesis, it was simpler to provide the participants with a pre-produced sentence. This allowed me to control the position of the target stimuli in the sentence as well as to arrange it in such a way so as to avoid this problem of continued voicing, hence the reason why the target stimuli are preceded with voiceless plosives and/or fricatives. Moreover, this pre-produced sentence also operated as a control as it meant that all the speakers articulated the same sample.

3.3.1. Total number of data tokens

In sum, every participant was provided with twenty-four words that were positioned into sentences. As each word was placed within three sentences, a total of seventy-two cards were presented to each participant, which provided me with a plentiful seventy-two tokens of VOT per speaker. As twelve participants were involved in this experiment, the overall sum of tokens that were collected for acoustic analysis was eight hundred and sixty-four. However, twenty tokens had to be removed due to the difficulty of measuring the VOT in such words[24], leaving a total of eight hundred and forty-four tokens for analysis. This extensive figure was substantial enough that any conclusions drawn from the results were from statistically significant grounds rather than unsystematic, inadequate data.

3.4. Data collection process

The process of collecting the above tokens of data included a formal review of the ethical aspects of the study as well as giving careful instructions to the participants.

At least three days prior to the recording procedure, participants were given a consent form (see Appendix 8.5) and an information sheet (see Appendix 8.3), the latter of

[23]The participants in the pilot study who felt self-conscious needed tremendous prompting. This was not only an inefficient use of time but this self-consciousness may have also impacted the VOT measurements in their speech.

[24] The main difficulty which resulted in the removal of twenty tokens was continued voicing from the word preceding the target word. Although the data was carefully chosen to avoid this problem, it could not account for the idiolect of each participant.

which explained the purpose of the thesis to the participants. However, in order to prevent the subjects from feeling self-conscious, the information sheet stated that the aim of the research project was to investigate whether there was a correlation between bilingualism and eyesight. By modifying the focus of the thesis from pronunciation to vision, the participants were less likely to alter their accent, thus decreasing the chances of a production that is atypical of their everyday style of speech. All participants were provided with a minimum of three days to decide whether they wanted to partake in the study. In order to uphold ethical practices, the participants were, upon the conclusion of the recording procedure, informed about the real objectives of the thesis and thus were given the more accurate information sheet (see Appendix 8.4). They were then provided with at least 7 days to decide if they still wanted to participate. If they chose to remain in the study, their signed consent form was collected from them thereafter.

During the recording procedure, the participants were shown one flashcard at a time. No limit was set for the time that each flashcard was held up to the participant as setting a limit may have increased the participants' speaking tempo, thus resulting in a decreased VOT. Having said this, no participant took more than ten seconds to read out each of the sentence flashcards. After each card, participants were provided with a ten second pause before the show of the next card. This short interval between each stimulus was to ensure the speakers were comfortable with the pace. This recording procedure approximately took twenty-five minutes per participant.

3.5. Recording

All of the data that was collected was recorded so that it could later be analysed acoustically. Acoustic analysis was chosen to be the focus of the study in order to consider productions that may be difficult to capture auditorily.

3.5.1. Recording Equipment

All of the tasks carried out with the speakers were taped using a Zoom H2n Voice Recorder that was positioned between 30cm and 40cm from the mouth. The recordings were then transferred onto a computer via a USB cable that directly attached to the recorder. It is worth noting that a smartphone could have been utilised instead as the quality of which almost matches that of the recorder and is sufficient enough for acoustic analysis (see De Decker and Nyez, 2011). However, a recorder was preferred due the precise nature of measuring voice onset time, which thus necessitates a high signal-to-noise ratio.

3.5.2. Recording Location

All of the recordings took place in either my own home or the homes of the participants. Although the recordings were for acoustic measurements and thus the quality needed to be clear, a laboratory environment was purposely not chosen. Despite there being no ambient noise in such a setting, there was the increased risk of the participants altering their speech to a formal register. Hence, it was more important to ensure the participants felt comfortable than to attain high quality, yet unnatural, data. In order to increase the quality of the data within the home, the recording took place in the quietest room as well as the room with the fewest hard surfaces, due to the fact that hard surfaces "reflect sound and thus compromise the clarity of the speech signal" (Podesva & Sharma, 2013, p.176). Furthermore, the recorder was placed no further than two feet away from every participant so as to guarantee that the distance between the participant and the recorder were equal in all of the recordings.

3.6. The impact of the fieldworker on the data

As all of the participants were from within my social circle (family and friends), the environment was inevitably relaxed and upbeat. Consequently, the participants felt comfortable with my presence, which encouraged them to maintain a casual register in their speech. As I am also of Gujarati ethnicity and a fluent speaker of the language, the participants deemed me as an 'insider'. Therefore, there was no awkwardness with regards to being recorded or speaking in English (particularly for the first generation for whom English is not their first language). Entering a community in which I am already an affiliate meant that I was required to convert from being an 'ordinary community member' to a 'fieldworker'. In order to present this transition, all of the recording procedures maximised the level of formality by means of interview-like seating arrangements. However, as they still viewed me as an insider, it was hoped that this arrangement did not affect the participants' speaking register.

With regards to minimising the possibility of the participants accommodating their speech style to mine, the recording procedure was deliberately designed to require the least amount of vocal input from my role as the fieldworker. Hence, the entire procedure involved flashcards.

3.7. Acoustic Measurements

All of the recordings that were obtained were afterwards analysed in Praat in order to find the voice onset time of each of the word-initial target stimuli. Voice onset time is defined as "the time interval between the burst that marks the release of the stop closure and the onset of the quasi-periodicity that reflects laryngeal vibration" (Lisker and Abramson, 1964, p. 422).

3.7.1. Waveforms

In order to understand how VOT can be measured in Praat, we must first be able to recognise the different types of waveforms. Waveforms of speech sounds can be categorised into four groups: quiescent, transient, periodic and random, each of which are illustrated below in Figures 6 to 9. A quiescent waveform is simply a pause of speech or the closure of the vocal folds. A transient waveform is illustrated by a burst of air. A periodic waveform works by cycles that are identical and constantly repeated whilst a random waveform does not have a recurring cycle or pattern, and can therefore also be defined as aperiodic.

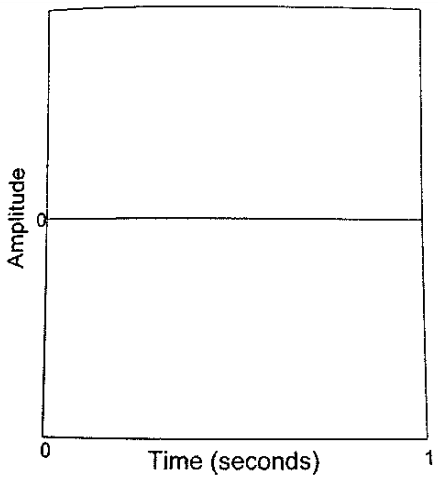

Fig. 6: Quiescent waveform (left). (Hewlett & Beck, 2006, p. 109).

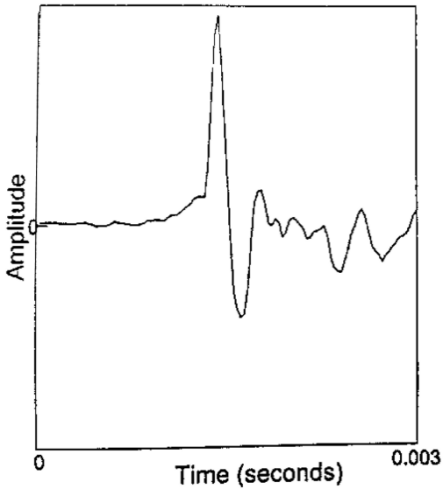

Fig. 7: Transient waveform (right). (Hewlett & Beck, 2006, p. 109).

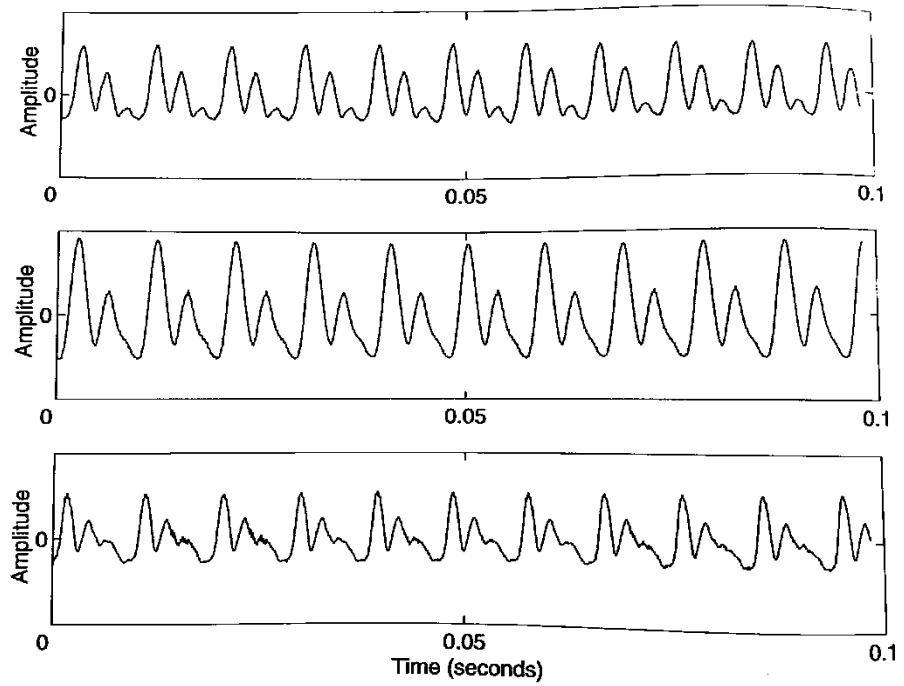

Fig. 8: Periodic waveforms of the vowel /u/ (top), the nasal consonant /n/ (middle), and the approximant /l/ (bottom). (Hewlett & Beck, 2006, p. 108).

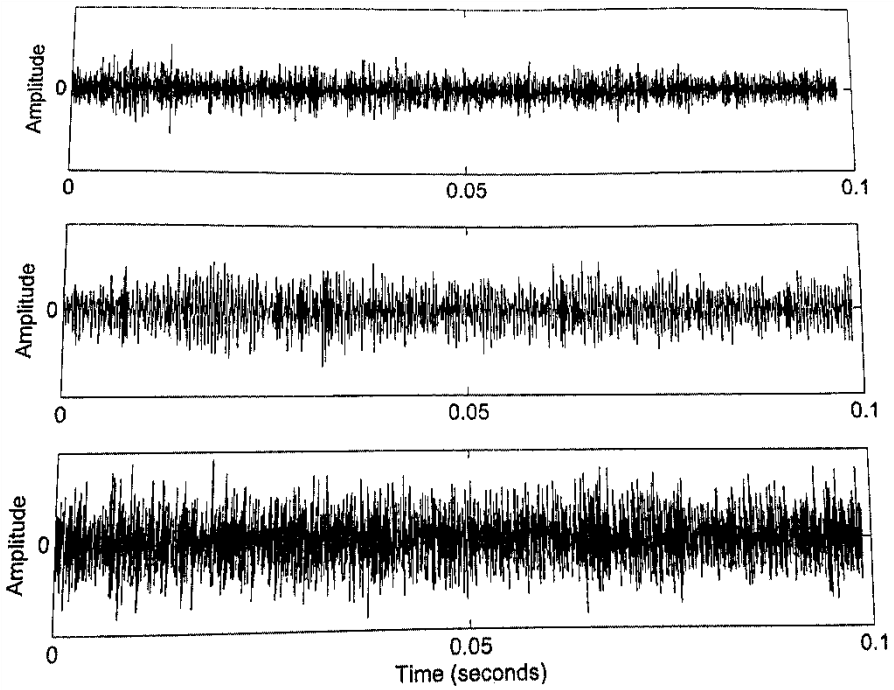

Fig. 9: Random waveforms of the fricatives /s/ (top), /ʃ/ (middle), and /f/ (bottom). (Hewlett & Beck, 2006, p. 109).

Finding the start at which voice onset time is measured is relatively uncomplicated due to the fact that it is simply at the place where the waveform changes from quiescent to transient. In other words, the speaker's articulators obstruct the airway and thus the vocal folds do not vibrate, which is represented by the flat line (Fig. 6). After this point, the air pressure that builds up during the quiescent phase is released with a burst (Fig. 7). Thus, the beginning of this burst is also the beginning of where voice onset time is measured. This is particularly visible when analysing the voiceless stops /p t k/, however there is difficulty when marking the start of VOT for voiced stops due to the continued voicing from the word preceding the target stimuli as well as due to prevoicing. Voiced stops will be discussed on page 12.

With regards to signposting the point at which voice onset time ends for a voiceless plosive, the area where the waveform transforms from transient to periodic or random is found. Distinguishing between the transient and periodic/random waveforms is rather straightforward due to the fact that a transient waveform decays very quickly and subsequently alters into a regular pattern i.e. periodic (Fig. 8).

3.7.2. Voiceless stops

To clearly demonstrate how voiceless stops were measured in Praat, a screenshot of a target stimuli being analysed on the software is provided below in Figure 10. The word being analysed is ‘tablet’ and the pink, highlighted area of the sound waves indicates the rough margins within which voice onset time is measured.

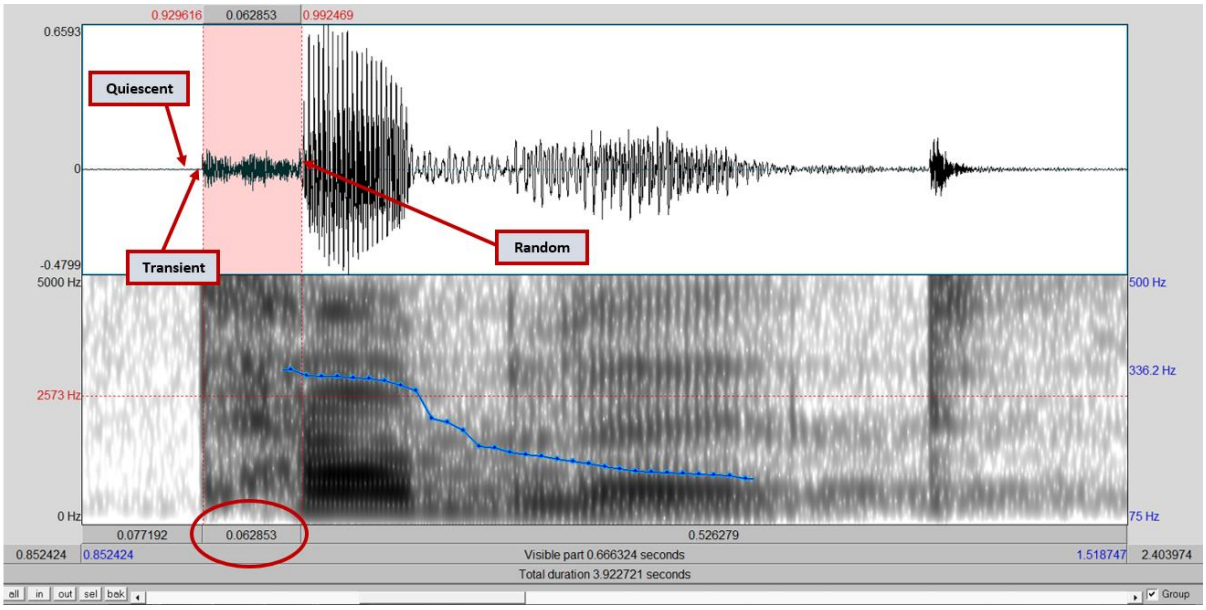

Fig. 10: A screenshot that illustrates the starting and ending point of voice onset time for the word ‘tablet’.

The highlighted portion of the sound waves in Figure 10 represents the transient waveforms, and the circled number directly below this area is the measurement of VOT in seconds. However, as VOT spans across a relatively small dimension, it is more appropriate as well as straightforward to multiply the circled number by 1000 and thus work with milliseconds. Therefore, in the example above, the VOT is 62.9ms (after rounding it to 1 decimal place).

Figure 10, however, is an approximate measurement due to the target stimuli being measured from a broad view. In order to attain more accurate results, each and every word-initial plosive was highlighted and subsequently enlarged in order to ensure that the exact points of VOT were marked. Figure 11 below is a zoomed-in version of the same target word ‘tablet’.

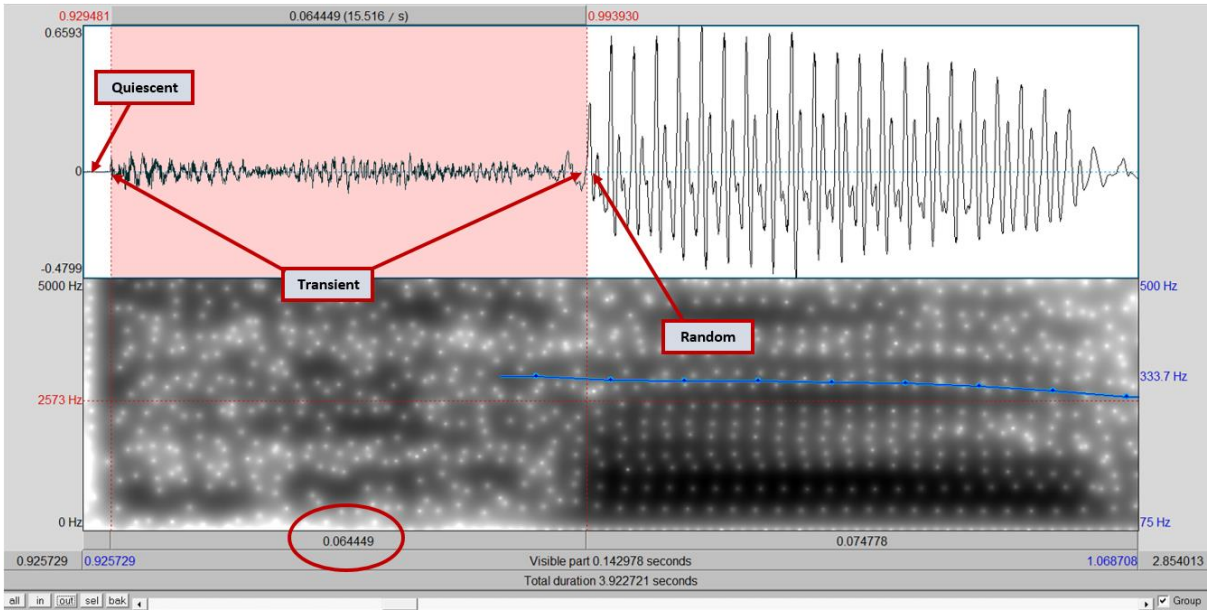

Fig. 11: A zoomed-in screenshot that illustrates the starting and ending point of voice onset time for the word 'tablet'.

A more narrow display of the target stimuli allows for more accurate markings. As can be seen in Figure 11, the VOT measurement is 64.4ms. Though this is only a 1.5ms difference from the measurements taken in Figure 10, it was extremely important to attain precise measurements so that the analysis based on the findings was as truthful as it could be.

3.7.3. Voiced stops

Voiced stops were measured and categorised into two different types: positive VOT and negative VOT. Each will be explained in the subsections below.

3.7.3.1. Positive VOT

Those voiced stops in which positive VOT was present was found using the same procedure that was utilised during the measuring of the voiceless stops. Figure 12 below is the vocalisation of the sound /d/ and the start of the vowel /ʌ/ in the word 'dustbin'. The highlighted area illustrates the transient waveforms of the plosive. As indicated by the red circle there is a short VOT of 8ms.

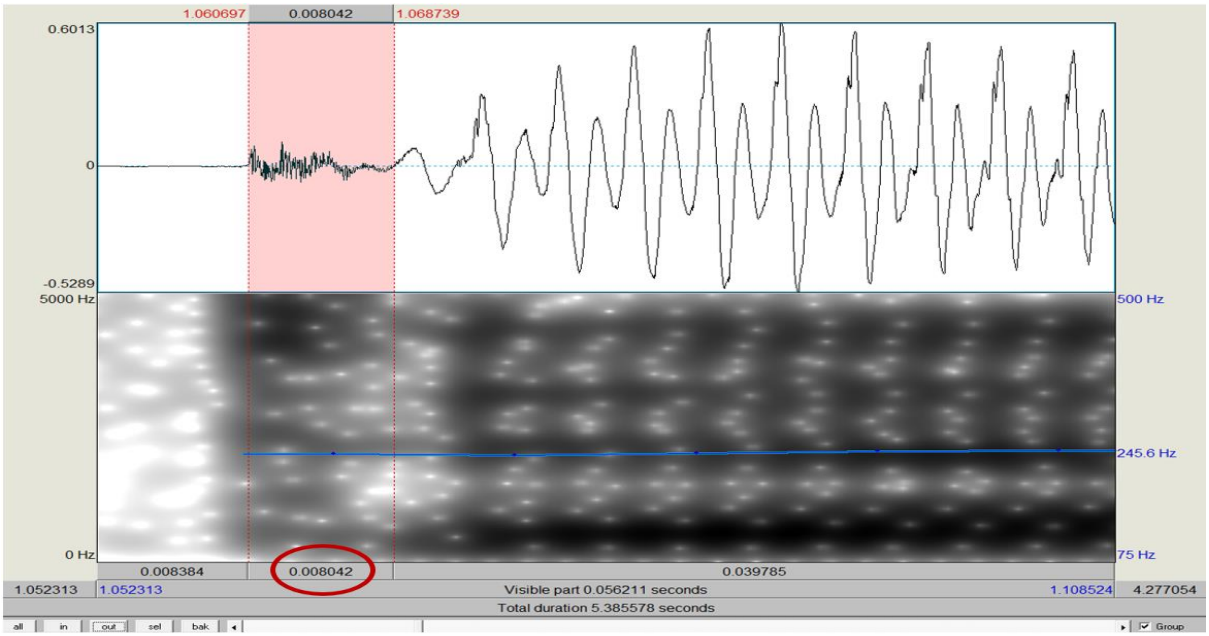

Fig. 12: A screenshot that illustrates the articulation of the stop sound /d/ said at the start of the word ‘dustbin’.

3.7.3.2. Negative VOT

Measuring voiced stops which possessed negative VOT (or prevoicing) required looking at a different reference point. Negative VOT refers to when the onset of vocal fold vibration precedes the release of the plosive. Therefore, this required me to measure the period that followed the quiescent waveforms but which came before the burst. Figure 13 below is a screenshot taken from Praat of the word 'digital'.

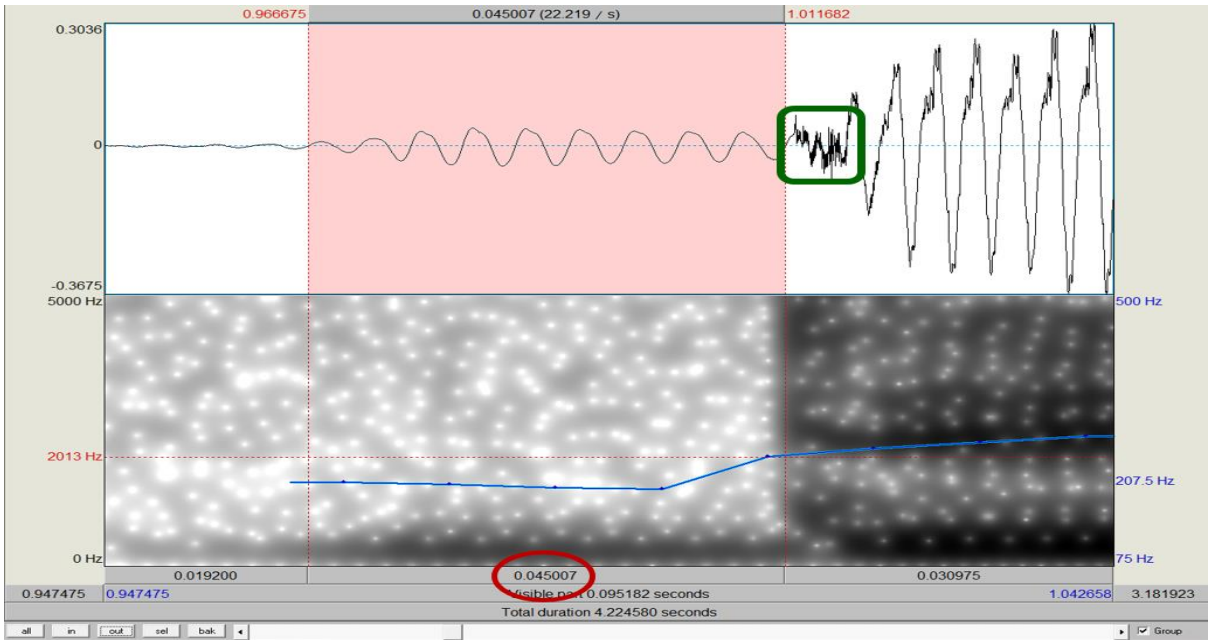

Fig. 13: A screenshot that illustrates the articulation of the stop sound /d/ said at the start of the word 'digital.

The highlighted area in Figure 13 displays an extensive amount of voicing prior to the burst shown within the green square. This onset of phonation lasts for an extensive 45ms. However, as it occurs before the release of the plosive, the VOT has a negative value and is therefore -45ms.

All of the results recorded from the participants articulating the target stimuli fit into one of the above acoustic transcriptions. In order to ensure reliability and accuracy, some of the measurements were checked by my supervisor. The results that were gathered as a result of following the methodological outlines of this chapter will be illustrated and discussed in the following chapter.

4. RESULTS

This chapter will present the results gathered from the data[25] and thus address the first half of Research Question 1 that is posited in the Introduction. The first half of the question is 'What are the word-initial VOT values for /p b t d k g/ of the data across the three generations?'

The first section will report the general differences between each of the three generations. It will subsequently examine the individual distinctions between participants. The final section of this chapter will group all the generations together and will consider the results with regards to the three positions of the target words within the utterances: utterance-initial, utterance-medial and utterance-final.

4.1. General findings

Figure 14 below is a bar chart which shows the mean VOT values for each of the three generations across all phonemes and across all three positions of the target word within the utterance. The phonemes, however, have been separated into the two groupings of voiceless and voiced. The y axis represents the three generations that are in focus for this thesis, and the x axis displays a line of measurement in the unit of milliseconds[26], stretching from -20ms to 80ms. The blue blocks of data in the chart represents the voiceless phonemes, whilst the orange data represents the voiced phonemes.

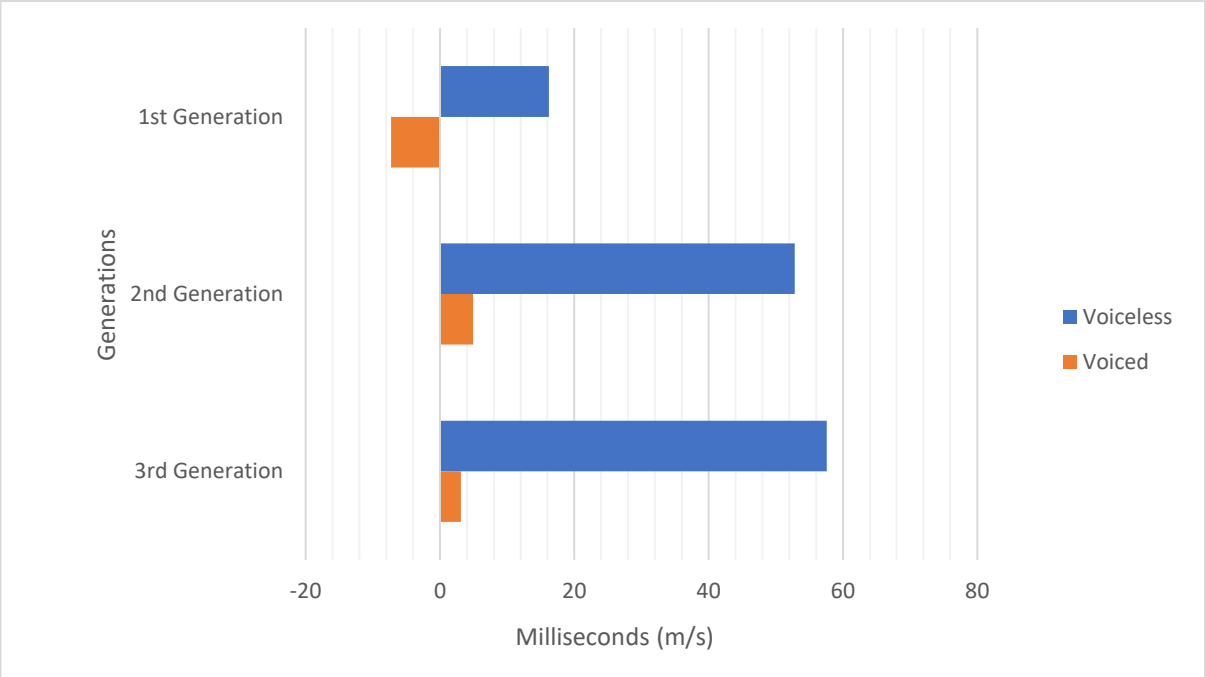

Fig. 14: The mean VOT values for the voiceless and voiced stops for each of the three generations.

[25] The data of all eight hundred and sixty-four tokens can be found in Appendix 8.2
[26] Note that this measurement using milliseconds includes negative numbers as well.

A noticeable change can be seen in Figure 14 above. With regards to the voiceless stops, the mean VOT increases as the age of the generation decreases. The first generation have a mean VOT of 16.2ms, the second generation have a mean VOT of 52.8ms and the third generation have a mean VOT of 57.6ms. There is an average difference of 36.6ms between the first and second generation, but this gap reduces between the second and third generation as there is a difference of only 4.8ms. This presents a clear progression from the Gujarati-like production of voiceless stop sounds (short lag) towards the English-like production (long lag).

The results for the voiced stops, however, do not reach broad ranges across the positive and negative VOT spectrum like the voiceless stops do. The first generation is the only age group that has a mean value of negative VOT at -7.3ms. This is most likely because *all* of the participants within this generation produced at least some of their target words with the Gujarati-like lead voicing. Yet, as Figures 18, 19 and 21 below show, members of the second and third generation also used negative VOT, though much less frequently. The second generation have a mean VOT of 4.9ms and the third generation have a mean VOT of 3.1ms, both of which are extremely close to near-zero VOT. Like the voiceless stops, there is again only a minor difference of 1.8ms between the second and third generation. However, it is the second generation that has a higher mean VOT value than the third generation.

Although Figure 14 above portrays the general overview of the findings, the results need to be considered further in order to find individual differences between each participant as well as each of the three positions of the target word. Sections 1.2, 1.3 and 1.4 will analyse such findings for each of the 12 participants.

4.2. First Generation

Figures 15, 16 and 17 below present the results for the first generation with regards to the positioning of the target word in an utterance as utterance-initial, utterance-medial and utterance-final respectively.

In all three of the figures in this section, the y axis represents each of the four members of the first generation, which follows an alphabetical order as it moves down the axis. The x axis represents the unit of measurement (milliseconds – m/s), the ranges of which vary in each chart. The mean findings for all six of the phonemes are illustrated for each participant, which first starts with the voiceless stops and subsequently the voiced stops. For visual ease, a different colour is used for each of the phonemes. These colours are consistently used for all of the figures. The key displaying such colours is situated on the right hand side of the charts.

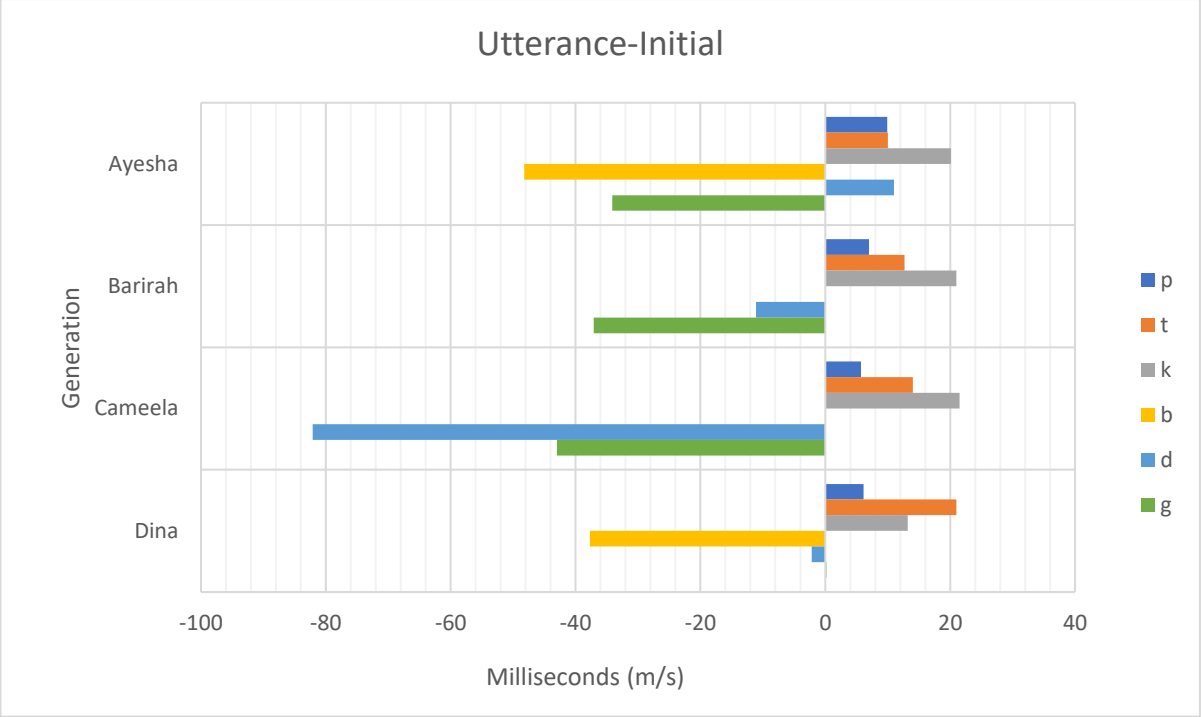

Fig. 15: The mean VOT values for each member of the first generation when the target word was in utterance-initial position.

In Figure 15 above, all of the VOT values for the voiceless plosives exceeds no higher than 21.5ms, thus all are within the short lag range. The mean VOT for all three of the voiceless stops combined is similar for all four participants. Ayesha has an average VOT of 13.3ms, Barirah has 13.6ms, Cameelah has 13.7ms and Dina 13.4ms. Therefore, no participant can be argued to have the substantially highest VOT. A common aspect between all four of the participants is that the stop sound /p/ was produced with the lowest VOT than any other voiceless plosive[27]. For three of the participants, the VOT increased as the place of articulation moved further back in the mouth. Hence, /p/ was followed with a /t/, which had a higher VOT and then /k/, which had the highest VOT. The only exception to this was Dina who articulated /t/ with a higher VOT value than /k/.

The results for the voiced plosives are much more varied, yet they do show signs of Gujarati-like VOT. Each participant produces two out of the three voiced plosives with negative VOT, therefore eight out of the twelve voiced stops (67%) are prevoiced. Ayesha has the highest negative VOT for the plosive /b/ at -48.2ms, which is followed by Dina at -37.7ms. However, Barirah and Cameelah both have zero VOT for the same plosive, placing these two participants at the extreme start of English-like VOT. All of the participants apart from Dina produce /g/ with negative VOT. With regards to /d/, Ayesha is the only participant to produce this stop sound with positive VOT and thus has the strongest English-like production for any of the voiced plosives for the first generation at utterance-initial level.

[27] In Ayesha's data, the bar for /p/ is 0.1ms shorter than the bar for /t/.

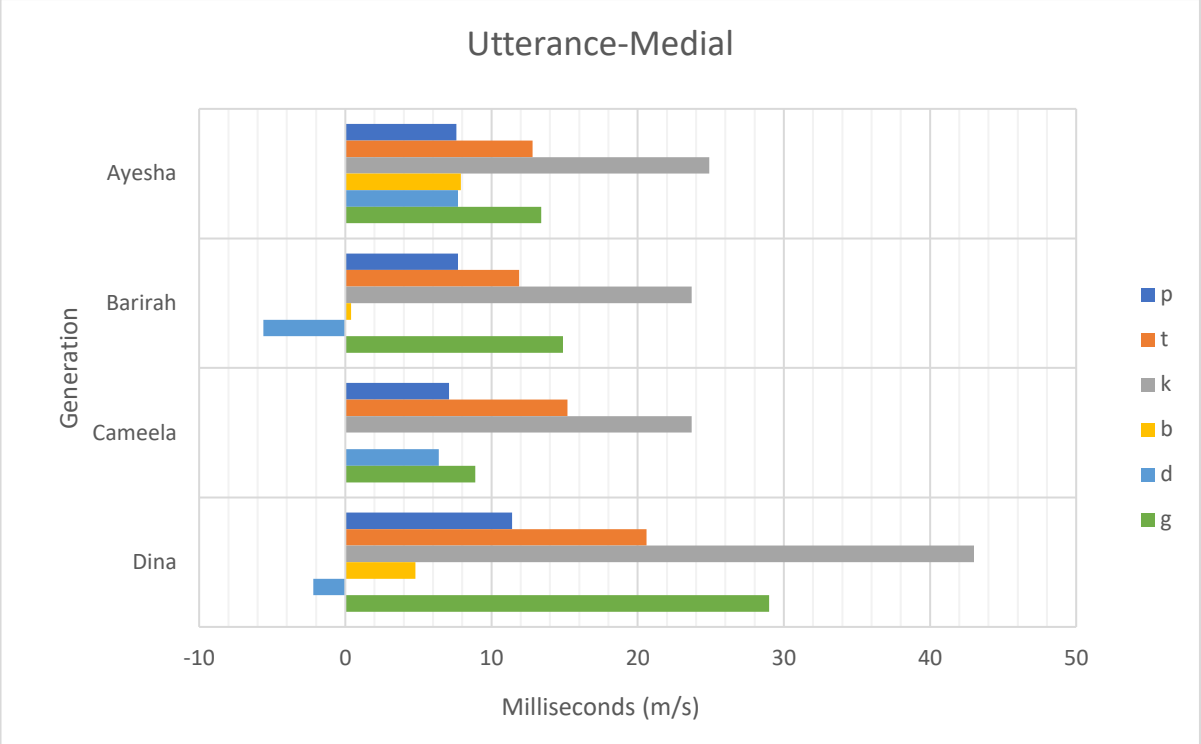

Fig. 16: The mean VOT values for each member of the first generation when the target word was in utterance-medial position.

In Figure 16 above, like Figure 15, all four of the participants produce higher VOTs for voiceless stops when the place of articulation is further back in the mouth. Thus, /k/ has the highest VOT value, which is followed by /t/ and then /p/. Ayesha, Barirah and Cameelah all have slightly increased their VOT values for when the target sound was in utterance-medial position than they did when the sound was in utterance-initial position. However, Dina's VOT values for each of the three voiceless plosives is significantly higher than its utterance-initial position as well as higher than the other participants, thus she has the highest range. The biggest rise is in her articulation of the plosive /k/, which presents an increase of nearly 30ms and is extremely similar to an English-like production.

With regards to the voiced stops, Barirah and Cameelah, who both had zero VOT for the plosive /b/ in utterance-initial position (Figure 15), continue to produce the same plosive with zero, or near zero, VOT values. However, one of the most striking observations that can be made regarding the voiced stops is that the majority of the articulations have positive VOT. Dina displays her acoustic range by producing the highest voiced sound at 29ms. However, the average VOT value for voiced stops from the other three participants is 7ms. Interestingly, all four of the participants produce /g/ with a higher VOT. This is not only higher than the voiced stops /b/ and /d/, but also than the voiceless stops /p/ and /t/[28]. This is an unusual feature of the data as voiceless stops typically have a higher VOT than voiced stops in order to mark a difference in voicing.

[28] This excludes Cameelah whose voiced stop /g/ does not exceed /t/ but does exceed the values of /p/, /b/ and /d/.

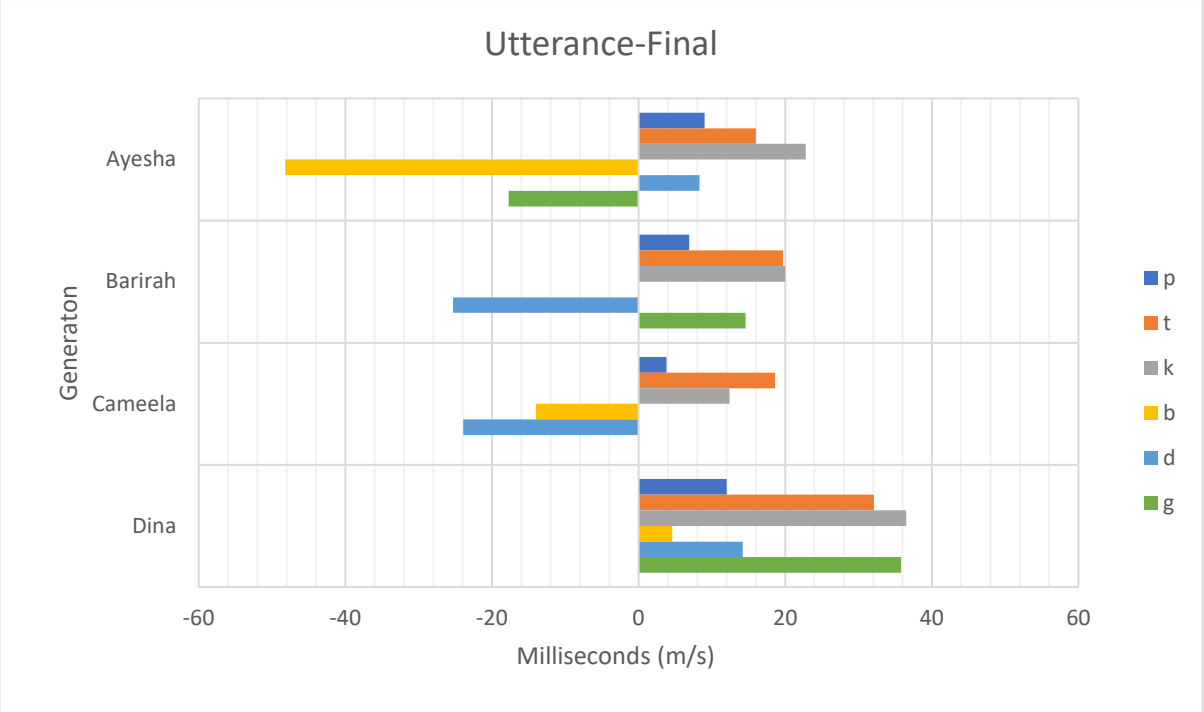

Fig. 17: The mean VOT values for each member of the first generation when the target word was in utterance-final position.

In Figure 17 above, the results for the voiceless plosives are similar to that in Figures 15 and 16. There is typically an increase in VOT values from utterance-medial position to utterance-final position. Dina produces the highest VOT values for all six of the stop sounds and all are on the positive side of the VOT continuum. Given all three of the first generation's data, it is evident that Dina has the most English-like articulations for all of her stop sounds. The other participants, however, continue to vary between positive and negative VOT. Barirah is the only participant to articulate with negative VOT in at least one of the voiced stop sounds at utterance-initial, utterance-medial as well as utterance-final place.

4.3. Second Generation

Figures 18, 19 and 20 below illustrate the mean VOT values for the second generation. In these three figures, the y axis displays the four members of the second generation, starting from E and ending with H. The x axis represents the measurement used to assess the VOT, which here is milliseconds (m/s). The same colours used to illustrate the phonemes in the figures for the first generation (Figures 15, 16 and 17) are applied here also. The key which symbolises the colours is positioned on the right-hand side of the figures.

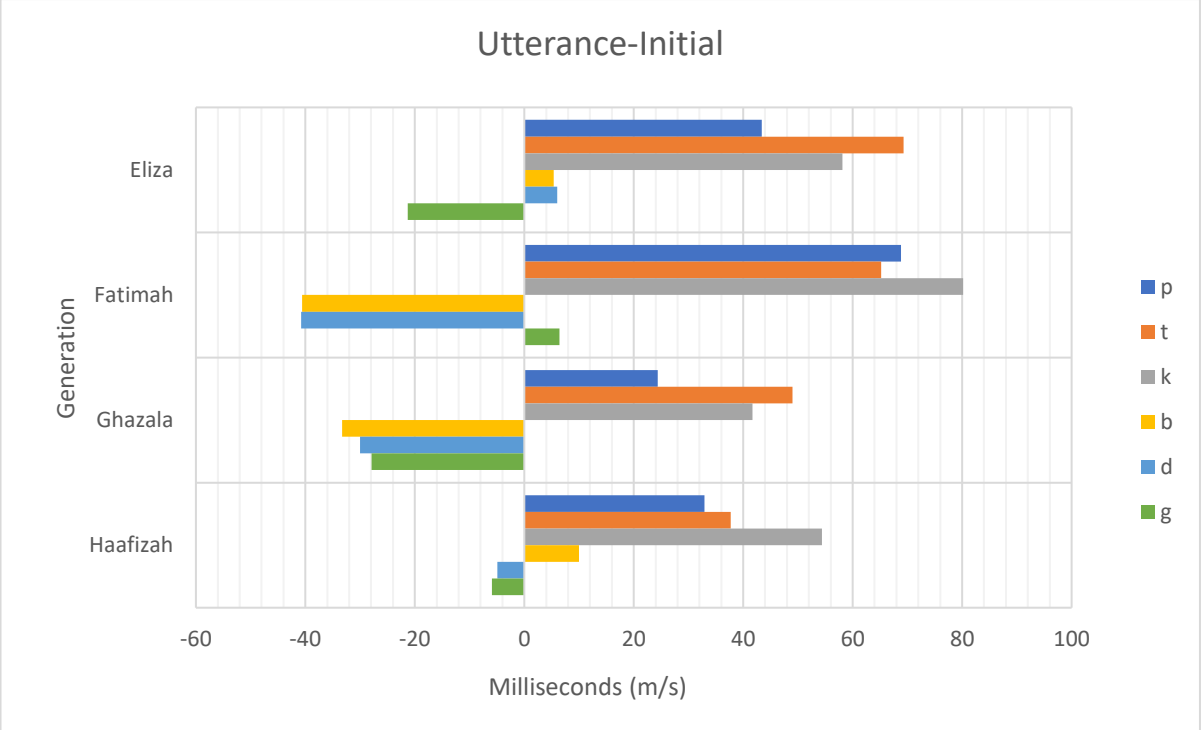

Fig. 18: The mean VOT values for each member of the second generation when the target word was in utterance-initial position.

Figure 18 above does not appear to present a strong relationship between VOT values and place of articulation. However, it does illustrate that all of the participants in the second generation pronounce voiceless plosives with a higher VOT than the majority of the first generation (Figure 15). Fatimah produces the highest VOT values for the plosives /p/ and /k/, which reaches 68.8ms and 80.2ms respectively. Ghazala produces the lowest VOT for the plosive /p/, whilst Haafizah has the lowest VOT value for /t/. These values are, nevertheless, higher than any of the first generation's productions at utterance-initial level (Figure 15). Furthermore, Ghazala and Haafizah, despite presenting a narrow range in their articulation of /p/, still show a range that extends to at least 50ms in their production of /t/ and /k/. All of the voiceless VOT values in Figure 18, therefore, fit within a range on the VOT continuum that is expected of English plosives (Cho and Ladefoged, 1999; Docherty, 1992).

Like Figure 15, Figure 18 also produces eight of the twelve voiced stops with negative VOT. This is expected of the first generation as negative VOT is an acoustic attribute of Gujarati. However, it is surprising that British-born speakers of English also produce Gujarati-like VOT values. Whilst the members of the first generation all articulate two out of three of their voiced stops with negative VOT, the second generation do not show the same consistency. Rather, it appears to be dependent on individual differences. Eliza prevoices only /g/, whereas Ghazala articulates all three of her voiced stops with negative VOT. Despite such individual differences, however, all four of the participants articulate at least one voiced stop with negative VOT, suggesting that the findings may be more than simply idiosyncratic speech patterns.

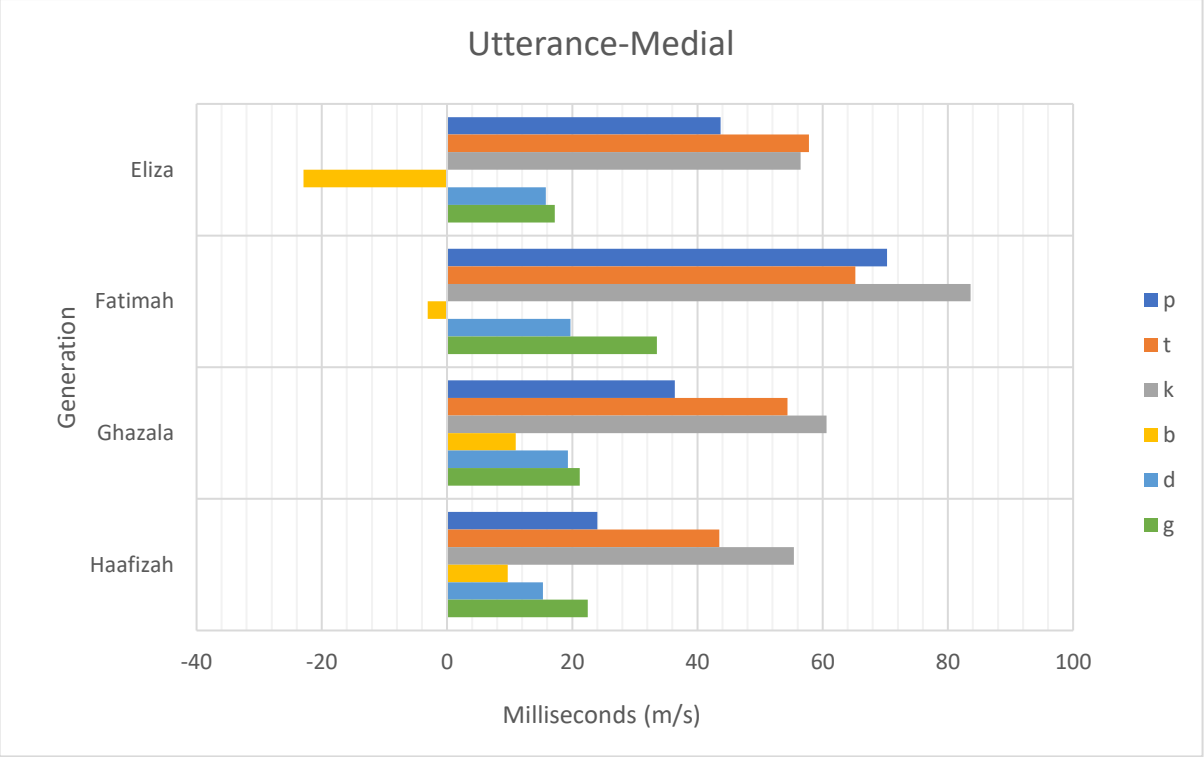

Fig. 19: The mean VOT values for each member of the second generation when the target word was in utterance-medial position.

Similar to Figure 16, all apart from two of the voiced stops in Figure 19 above are produced with positive VOT. Fatimah, like in the findings for the utterance-initial position (Figure 18), produces the highest VOT values for her voiceless stops as the bottom end of her range starts from no less than 65ms. Figure 19 also indicates an overall increase in VOT values from utterance-initial position to utterance-medial position. Based upon the mean values, the plosive /p/ increases by 1.2ms, and /k/ increases by 5.4ms. However, the stop /t/ decreases by 0.1ms.

With regards to the voiced stops, the velar plosive /g/ has the highest VOT value from all three voiced stops and for all participants. It is also higher than the VOT values for the plosive /g/ produced by the first generation at utterance-medial level. However, the VOT for /g/ does not exceed the VOT values for any of the voiceless stops, like it does in Figure 16 for the first generation. There also appears to be a positive relationship between VOT values and place of articulation, which has so far only been observed with voiceless plosives. It is therefore apparent that the velar position of the mouth allows participants to produce the highest VOT whereas bilabials the least.

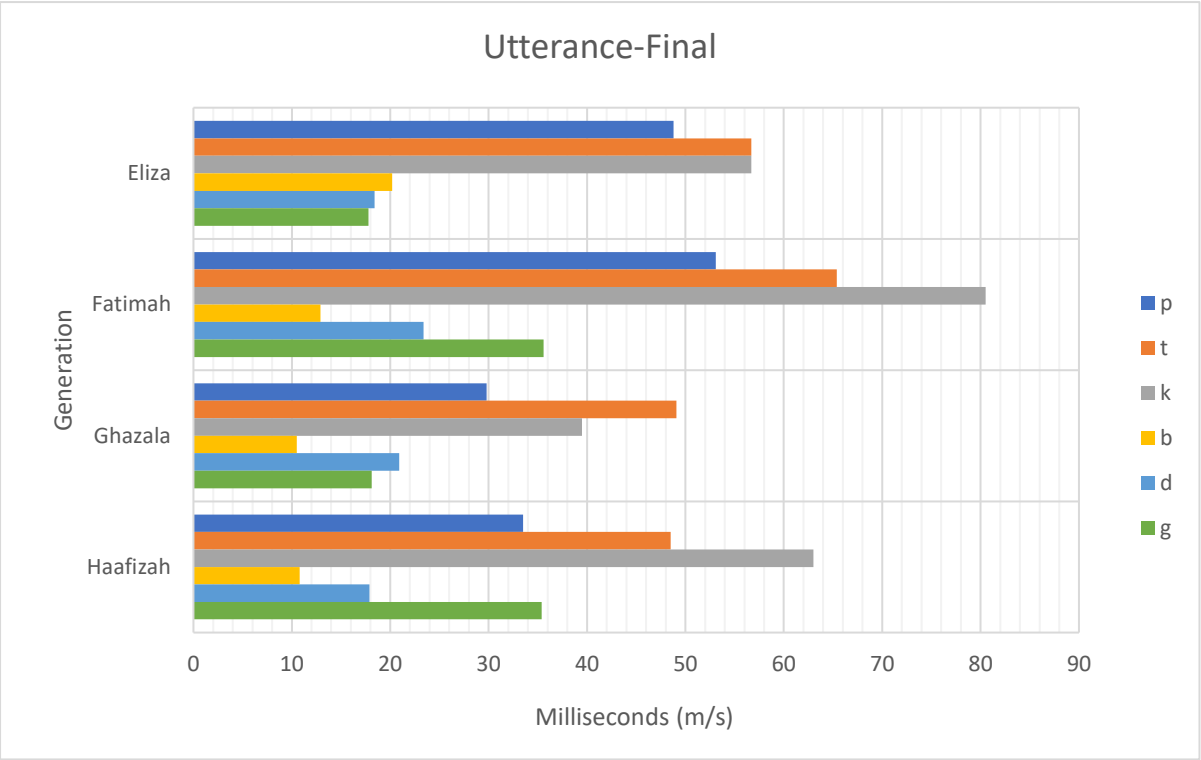

Fig. 20: The mean VOT values for each member of the second generation when the target word was in utterance-final position.

Figure 20 above illustrates that utterance-final position is the only point at which all members of the second generation produce positive VOT for both voiceless and voiced stops. It is now evident that Dina's results from the first generation is more similar to the second generation than her own respective category. This places her within the English range of VOT values, despite her being born in Gujarat and learning English as a second language.

Furthermore, Ghazala and Haafizah continue to produce /t/ with a shorter VOT than Eliza and Fatimah, as they did in Figures 18 and 19. This suggests that their acoustic range for the bilabial position does not surpass 40ms. However, we know that Ghazala's and Haafizah's acoustic capabilities stretch higher than 40ms due to the fact that their productions of /t/ and /k/ from utterance-final position as well as from utterance-medial position are between 40 and 60ms.

With regards to the voiced stops, Fatimah and Haafizah present a positive relationship as VOT values increase when the place of articulation moves further back in the mouth. Haafizah's articulation of /g/ exceeds the VOT for her production of the voiced stop /p/ by 1.9ms. Eliza is the only participant in Figure 20, as well as from all the other Figures thus far, to present a negative relationship between her voiced stops and place of articulation. Her production of /b/ has a mean VOT of 20.2ms, /d/ has a VOT of 18.4ms and /g/ 17.8ms. Her results, therefore, show an increase of VOT values when the place of articulation is closer to the front of the mouth as opposed to the back.

Overall, there appears to be very small differences between the members of the second generation at utterance-final level as the findings for all four participants are within the English-like range.

4.4. Third Generation

Figures 21, 22 and 23 below illustrate the results found for the third generation. The y axis for the three figures shows the four participants of this generation which starts at I and ends with L. Upon the x axis is the measurement of milliseconds (m/s) used to demonstrate the VOT values. The colours have remained consistent from the previous figures, and the key for which is located on the right-hand side of all the charts.

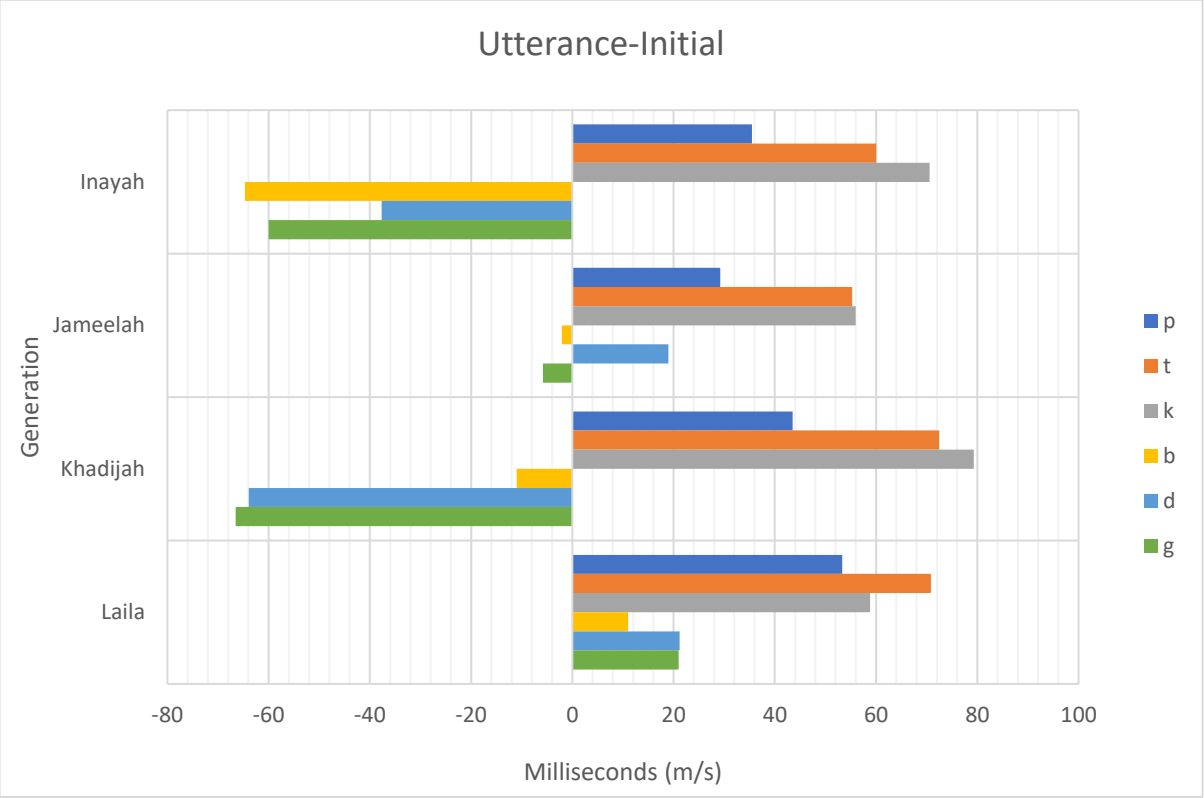

Fig. 21: The mean VOT values for each member of the third generation when the target word was in utterance-initial position.

In Figure 21 above, the findings show both positive and negative VOT. As Figure 14 in Section 4.1 states, the third generation have higher VOT values for their voiceless plosives than the second generation, which creates a considerable difference between the first and third generation. Whilst the stops /t/ and /k/ see an increase, the stop sound /p/ sees a decrease by 2ms from the second generation to the third generation. In fact, the highest VOT value for /p/, /t/ and /k/ are all produced by members of the second generation. However, the mean value makes the third generation's VOT conclusively higher than the second generation.

Figure 21 illustrates that eight out of the twelve voiced stops have negative VOT values: a feature that was shown in Figures 15 and 18 also. However, what is interesting here is that two of the four participants prevoice *all* three of their voiced stops. In the first generation, there was no participant that had negative VOT for all three plosives. However, the second generation did have one participant (Ghazala) who had negative VOT for all three plosives. Therefore, there is an increasing number of participants who prevoice per generation. Furthermore, four out of the eight negative VOT values are at least, or exceeds, -60ms. No members of the second generation produce voicing lead of this duration and neither do any members of the first generation, apart from Cameelah in her articulation of the stop sound /d/.

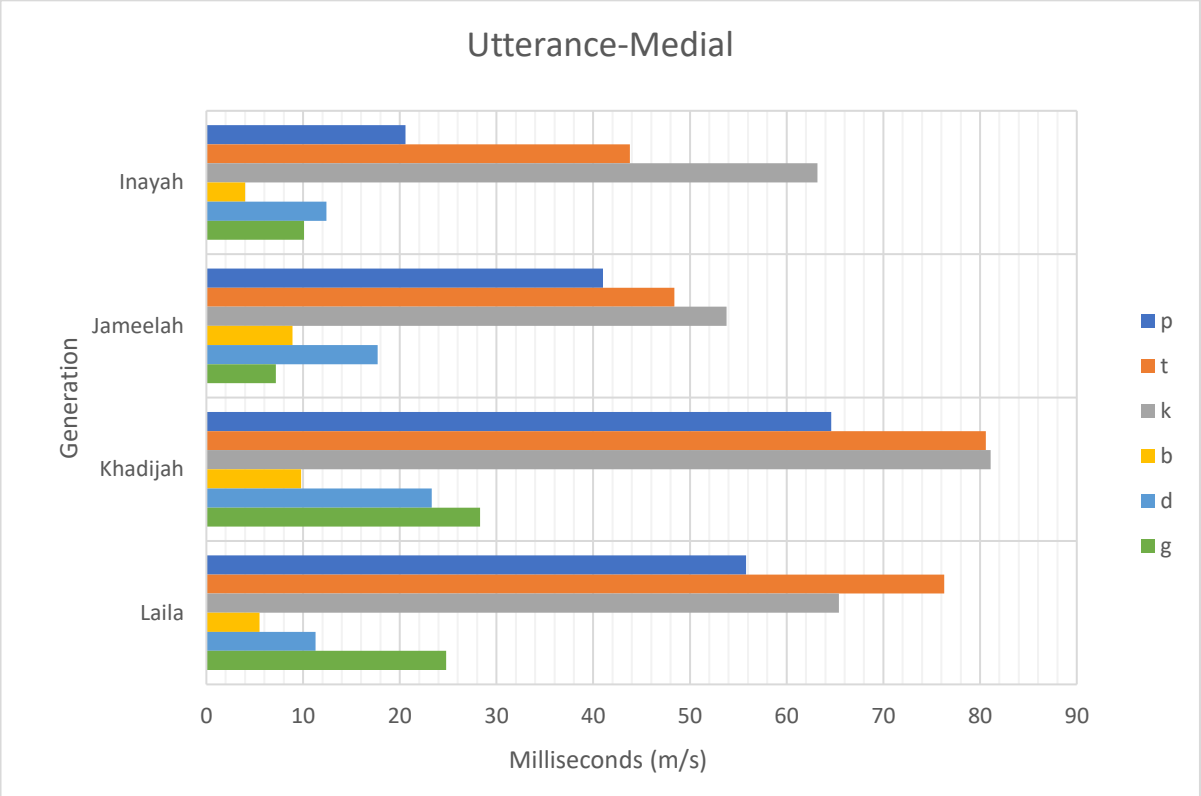

Fig. 22: The mean VOT values for each member of the third generation when the target word was in utterance-medial position.

From all the utterance-medial bar charts (see Figures 16 and 19), Figure 22 above is the only one in which all participants consistently produce both voiceless and voiced stops with positive VOT at utterance-medial level. Similar to what we have seen from the other generations, the VOT values increase from utterance-initial position to utterance-medial position for the stop sound /p/, which rises by 5.1ms. However, unlike Figures 16 and 19 in which there is an increase in VOT values in other plosives, Figure 22 above displays a reduction of 0.2ms from utterance-initial position for the stop /k/ and a 2.4ms reduction for the stop sound /t/. Khadijah once again produces /t/ and /k/ with the highest VOT as well as for /p/, which was not the case in Figure 21. Inayah's production of /p/ is the shortest VOT value from all the voiceless stops and from all four participants at utterance-medial level.

She is the only participant who presents a decrease of VOT from utterance-initial level for the bilabial plosive /p/, whilst the other three participants display considerable increases.

Two of the participants, Khadijah and Laila, present a positive relationship between VOT values and place of articulation, however Inayah and Jameelah do not present the same trend. Khadijah who has the highest negative VOT during utterance-initial position has the highest positive voice onset for all of her voiced stops than any other participant. This, along with her high range of VOT values for voiceless stops, suggests that she is at the extreme end of English-like articulation.

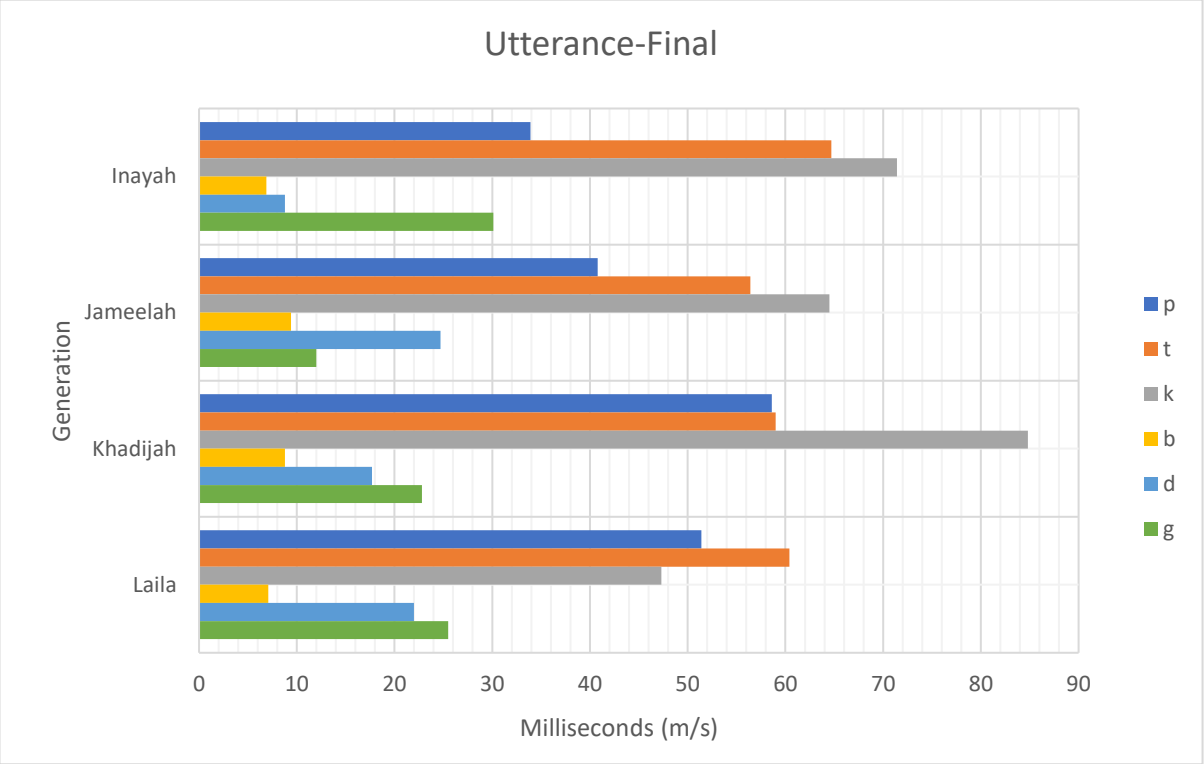

Fig. 23: The mean VOT values for each member of the third generation when the target word was in utterance-final position.

The results from utterance-final position in Figure 23 above are extremely similar to the findings from Figure 22. The plosives /p/ and /k/ see an average increase from utterance-initial and utterance-medial positions, however there is a decrease of 2.2ms in utterance-final position for /t/ than for the values in utterance-medial position. Khadijah continues to have the highest VOT values for /k/, which reaches a significant 84.8ms as well as for /p/, which has a measurement of 59ms. However, unlike Figure 22 in which Khadijah has the highest VOT values for both voiceless and voiced stop sounds, this is not the case for Figure 23 above. Inayah has the highest VOT measurement for her articulation of /g/, which is the highest value for any of the voiced stops that are produced by the third generation. However, she continues to present the lowest VOT value for her articulation of /p/.

4.5. Positioning of the target word

From Figure 24 to Figure 29 below, the participants have been grouped into their respective age category and have all been compared in relation to where the target word is placed in the sentence. Utterance-initial is presented in Section 4.5.1., utterance-medial in Section 4.5.2., and Section 4.5.3. displays the results from utterance-final position. These three subsections discuss the results for solely the voiceless stops. Section 4.5.4 discusses all of the voiced stops as a whole from the previous three sections (Figure 24 to Figure 29).

For each position of the word in the utterance, a bar chart is first presented to display the mean VOT values, which is followed by a line graph that visually displays the connections between the three generations. Figures 24 and 25 display all the results for utterance-initial, Figures 26 and 27 display the results for utterance-medial and Figures 28 and 29 display the findings for utterance-final. As opposed to Figures 14 to 23 above which present each participant separately, the subsequent charts allow us to compare the generations together with regard to the three different positions of the target words.

In Figures 24, 26 and 28 (the bar charts), the y axis displays the groupings of the three generations. The x axis displays the unit of measurement used to measure the VOT, which again is milliseconds (m/s). Every chart specifies the results for all six of the phonemes, starting with the voiceless stops and then the voiced stops.

In Figures 25, 27 and 29 (the line graphs), the y axis displays the unit of measurement (m/s) and the x axis displays the three generations. For all the Figures from Figure 24 to Figure 29, a key is provided on the right-hand side to explain the colour scheme for each stop sound.

4.5.1. Utterance-initial

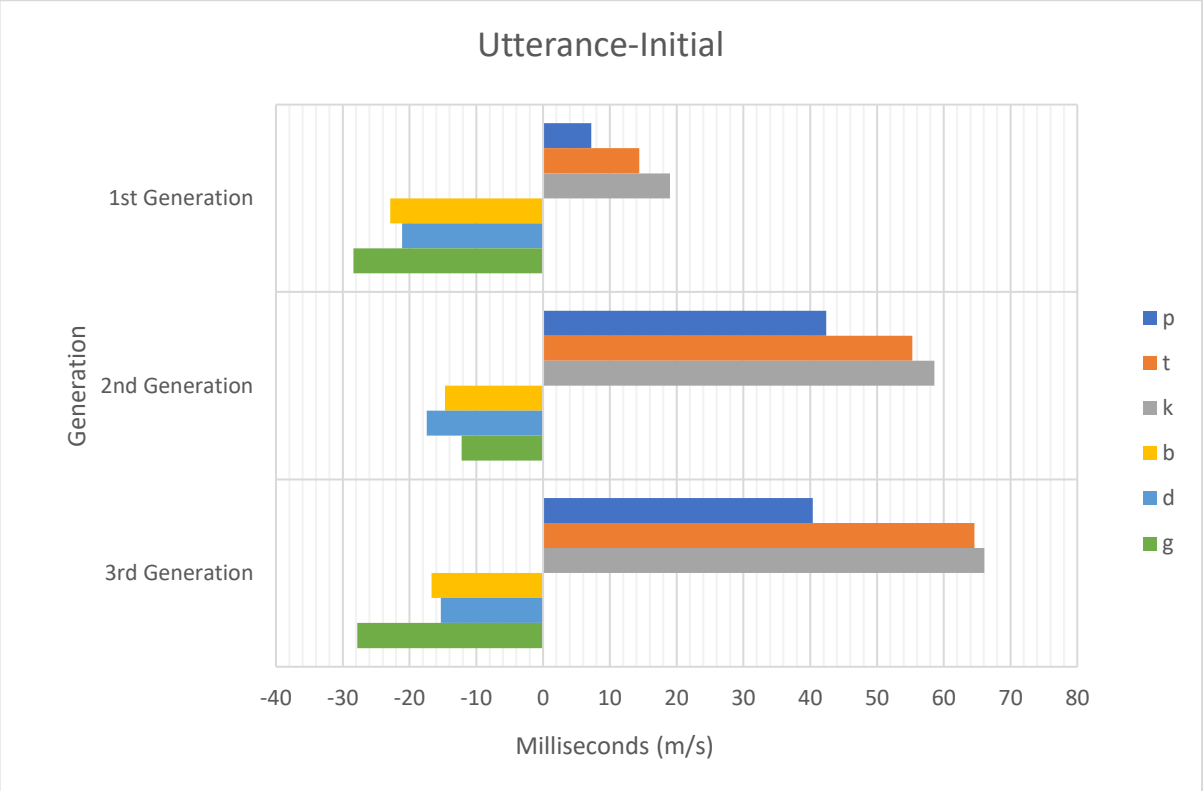

Fig. 24: A bar chart showing the mean VOT values for all articulations at utterance-initial position.

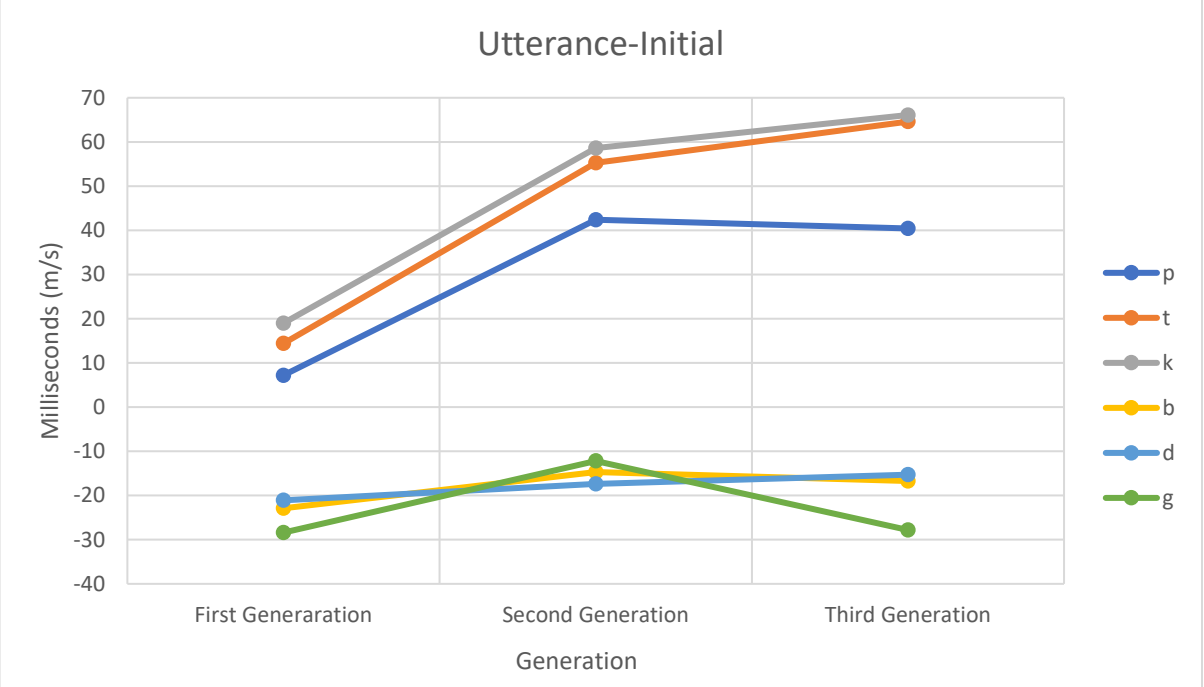

Fig. 25: A line graph showing the trend of the mean VOT values for all articulations at utterance-initial position.

A significant finding in Figures 24 and 25 is that VOT values increase as the generation becomes younger. For the second generation, the VOT values for /p/ increases from the first generation by 35.2ms, /t/ increases by 40.9ms, and /k/ sees an increase by 39.6ms. Whereas from the second generation to the third generation, the stop /t/ increases by 9.1ms (a difference of 57.5ms from the first generation), and the stop sound /k/ increases by 7.5ms (a difference of 47.1ms from the first generation). However, the bilabial stop /p/ sees a decrease by 2ms from the second generation to the third generation.

4.5.2. Utterance-medial

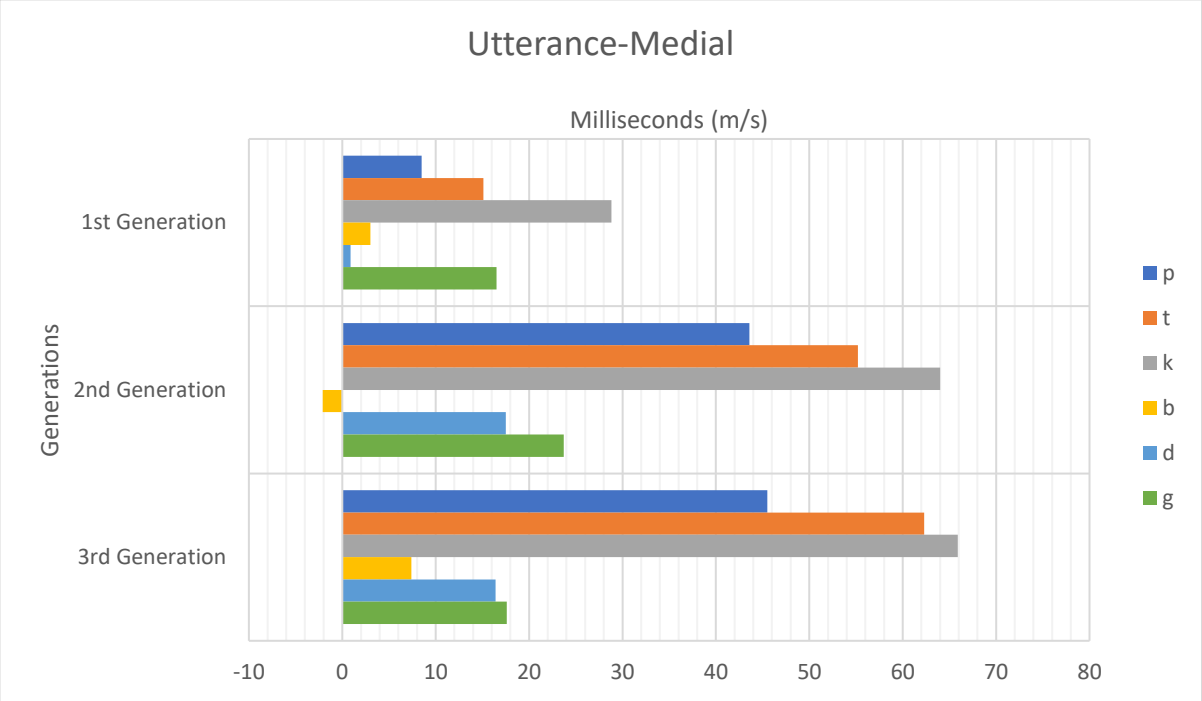

Fig. 26: A bar chart showing the mean VOT values for all articulations at utterance-medial position.

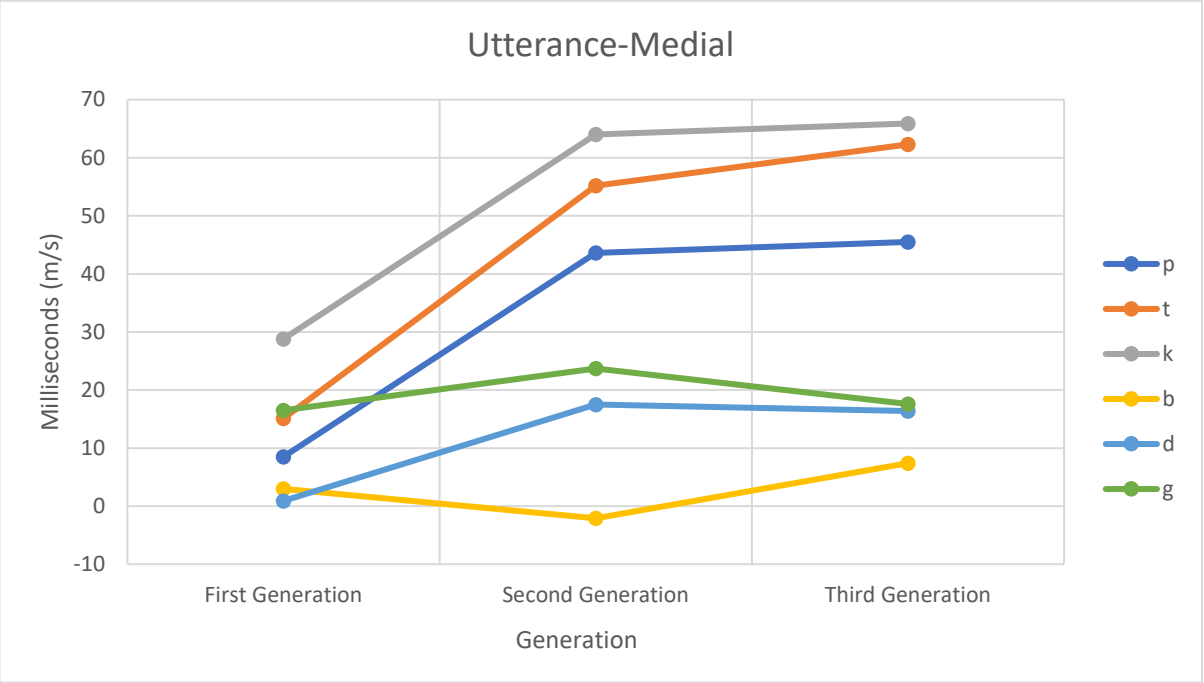

Fig. 27: A line graph showing the trend of the mean VOT values for all articulations at utterance-medial position.

Like Figures 24 and 25 for utterance-initial position, the VOT values increase across the generations. The plosive /p/ rises by 35.1ms from the first generation to the second generation, /t/ rises by 40.1ms, and /k/ rises by 35.2ms. However, from the second generation to the third generation, the stop sounds /p/ as well as /k/ increases by 1.9ms, whilst /t/ increases by a much higher 7.1ms.

4.5.3. Utterance-final

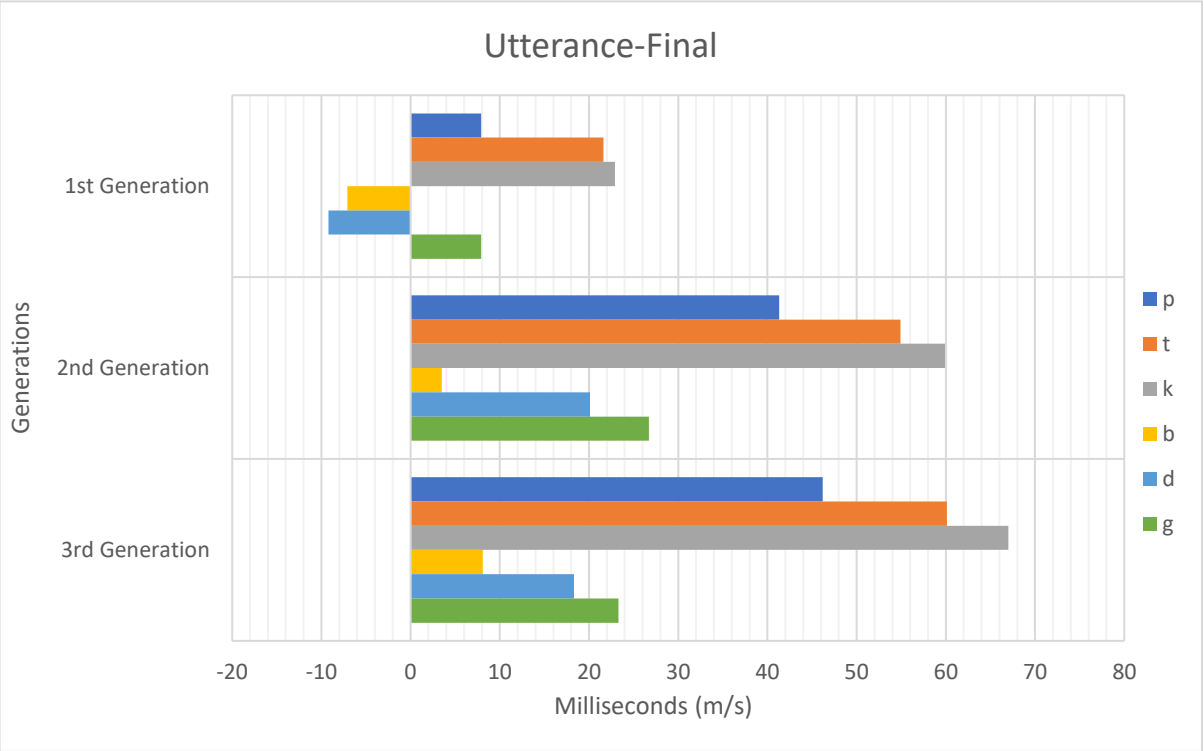

Fig. 28: A bar chart showing the mean VOT values for all articulations at utterance-final position.

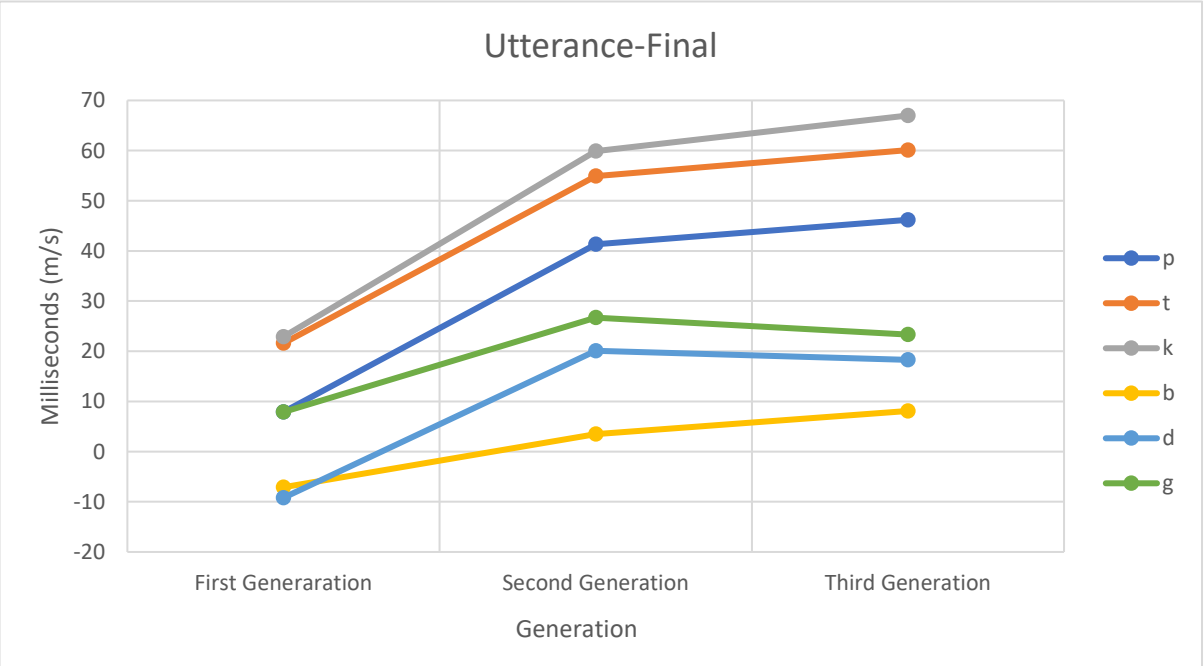

Fig. 29: A line graph showing the trend of the mean VOT values for all articulations at utterance-final position.

In Figures 28 and 29 above, the trend from the first generation to the second generation is as follows: the stop sound /p/ increases by 33.4ms, /t/ increases by 33.3ms, and /k/ increases by 33ms. This makes it apparent that utterance-final position has the most consistent and reliable rise from all three positions of the target word from the first generation to the second generation. However, from the second generation to the third generation, /p/ rises by 4.9ms, /t/ rises by 5.2ms, and /k/ rises by 7.1ms. For all three positions of the target word within the utterance (including Sections 4.5.1. and 4.5.2) as well as for all three voiceless stops, the biggest increase appears between the first and second generation, which becomes only a short rise between the second and third generation.

Furthermore, the results indicate in all six figures and for all three generations that there is a positive relationship between increased VOT values and place of articulation being moved towards the rear of the mouth. All three generations typically present an increase from utterance-initial to utterance-medial and from utterance-medial to utterance-final. As explained, the second and third generation have much higher ranges than the first generation, and therefore present much more acoustic potential.

4.5.4. Voiced stops

In contrast, an upward trend is not consistently noticeable between all three generations for the voiced stops, however it can be seen between the first generation and the second generation. This is then typically followed by a decrease from the second generation to the third generation. Suggestions of this are presented in the stop sounds /d/ and /g/ at both utterance-medial and utterance-final position. Like the voiceless stops, the biggest difference is between the first and second generation, whereas the second and third generation only present a slight variance.

Another important finding is that all three generations use negative voice onset time at utterance-initial level in Figures 24 and 25. The patterns are the most similar for the first and third generation as the velar stop /g/ has the lowest negative VOT, followed by /b/ and then /d/. However, for the second generation, the opposite pattern applies as /d/ has the lowest negative VOT, followed by /b/ and then /g/. The second generation's values for /b/ and /g/ are the smallest from all three generations. Nevertheless, all three generations displayed negative values for voiced stops which implies that, when the target word has no context before it, the stop sound is subject to negative VOT.

However, this pattern is not mirrored in Figures 25 and 26 as the majority of voiced stops are produced with positive VOT at both utterance-medial and utterance-final level, suggesting that environment is extremely important. The second generation in Figure 25 only has one stop sound that sits on the negative side of the VOT continuum as the plosive /b/ has a mean VOT of -2.1ms, largely due to Eliza's high negative value of the stop sound /b/ in Figure 19. In Figure 26, at utterance-final position, the /b/ moves to the positive side of the VOT continuum as /d/ and /g/, which have the highest positive VOT values for voiced stops from all three generations. The first generation in Figure 26, however, articulate /b/

and /d/ with negative VOT at utterance-final position. It is only the third generation that produce all voiced stops with positive VOT at both utterance-medial and utterance-final place.

4.6. Summary

Taking all three age groups and all six stop sounds into consideration, the participants display a wide range of VOT values that ranges anywhere in between -100ms and +120ms. Although this demonstrates the linguistic variation of Gujarati speakers, clear patterns can and have been identified. Such patterns provide insight into the linguistic make-up of Gujarati speakers in Batley, as well as bringing forward ideas of identity, familial relations, and social networks. Interpretations for all of the findings that have been reported in this chapter will be discussed in the following chapter.

5. DISCUSSION

This chapter will discuss and put forward various suggestions for the results presented in the previous chapter. It will do this by considering each generation in turn, starting from the first generation. The discussion will then consider the sociocultural implications of the data gathered.

5.1. Overall discussion

The overarching result is that, as the generations become younger, the VOT values change from Gujarati-like – short lag for voiceless stops and voicing lead for voiced stops – to mainly English-like – long lag for voiceless stops and short lag for voiced stops. This clearly suggests that those participants who are born in India and learnt English as a second language (L2) upon their arrival to the UK are still predominantly influenced by Gujarati, i.e. their L1. This influence is not such that the findings for the first generation are just slightly lower than the values of English, but they are also lower than the proposed Gujarati VOT values stated by Rami et al (1999). Although their short investigation only considers velar stops, their findings of -37.3ms for /g/ and 40.6ms for /k/ are still much bigger than the results for the first generation in this thesis. To understand the factors why Gujarati influences the first generation more than the other generations, we need to assess what is different across these age brackets.

The primary difference between the three generations is the age at which the participants acquired English. It has been argued by several scholars that the sequential acquisition of an L2 language that differs from a monolingual acquisition of solely the L2 is dependent on the age at which the L2 is acquired (Flege, 1991; Kang & Guion, 2006). Hence, "the L2 is learned through the filter of the L1" (Antoniou et al., 2010, p. 558). The first generation began their acquisition of English upon their arrival to the UK, which took place whilst they were between the ages of eight and fourteen years old. Although they can speak basic English, the majority of the first generation still are not fluent in the language. It is likely that when they started learning English as their L2, the phonological categories of Gujarati (their L1) were already fully formed, and interfered with the acquisition of English's phonological structures. Despite going to school and learning English academically[29], the first generation who had moved to a new land had very few established family members in

[29] Members of the first generation typically went to school until the age of 15 – the compulsory age until which they were obliged to stay in education (Education Act, 1944). Depending on the age at which they arrived in the UK, the number of years this generation went to school and acquired English varies between 1 to 8 years. Very few members went on to further education.

the in-migrant[30] community and even fewer contacts outside of their community, leading them to use Gujarati more. Consequently, the phonological categories of the L1 were continuously reinforced, making the acquisition of the L2 even more difficult.

On the other hand, the second and third generation were born in the UK and were exposed to both English and Gujarati during their childhood. They therefore established the phonological categories of each language at the same time. Researchers have stated that for simultaneous bilinguals, there is no cognitive interference or a struggle for common phonological space as both languages are acquired extremely early on and concurrently form two separate mental systems (Fromkin et al., 2013). From the outset, this theory regarding the age at which the L2 is learnt and thus the number of phonological categories the participants have appears to be a reasonable justification. This is in view of the results in Figure 14 (page 43) which present a movement from Gujarati-like to English-like VOT values across the generations. The next subsection will focus solely on the first generation and assess whether this theory or any other theory can allow us to contextualise the first generation's findings.

5.2. First generation

As explained in Section 5.1, the initial judgement from the results is that the age at which the L2 is learnt had an impact on the first generation's Gujarati-like VOTs. However, this justification presupposes that the first generation should theoretically always have Gujarati-like articulation and the British-born second and third generation should always possess English-like VOT values. Although the general overview of the results from Figure 14 presents this, the individual findings from Figures 15 to 29 are much more varied. Three members of the first generation, Ayesha, Barirah and Cameelah, present comparable results at all three positions of the target word in the utterance. However, Dina's results, particularly at utterance-medial and utterance-final level, reaches much higher VOT values than her counterparts in her generational age bracket. She is the only participant in the first generation to reach over 20ms for her articulation of /t/, over 40ms for /k/, as well as the only participant to have all positive VOT values for voiceless and voiced stops at utterance-final level (Figure 17). This latter feature is a key aspect of the second and third generation at utterance-final position to which Dina resembles, albeit with slightly lower values. Therefore, the age at which the L2 (English) is acquired cannot be the sole reason as to why the first generation have lower VOT values than the second and third generation. It is one of the many potential reasons; however, Dina's production of near English-like VOT values affirms that it is not the only possible explanation.

[30] In-migrant refers to those people arriving (in this case, to England) from another country (India). In other words, the first generation are all in-migrants.

Following brief interviews with the participants regarding their backgrounds, it was found that Dina and her husband live with their children - who are in their 40's - whilst another daughter and six grandchildren live next door. She therefore sees her children and grandchildren on a daily basis. Ayesha and Cameelah, however, live alone and see their grandchildren once a week. Barirah lives with her husband and her son who is in his 20s. All four of the participants reported to predominantly use Gujarati. However, only Dina reported to use a mixture of both Gujarati and English with her grandchildren, though Gujarati, which she described as her favoured language, was used more often. Using this information, it may be possible to explain why Dina has the most English-like VOT values. Her consistent exposure to British-born speakers of English with whom she converses on a daily basis provides Dina with the chance to use English, which has potentially influenced her own English production, and strengthened her use of long lag values in voiceless stops and short lag values for voiced stops at utterance-final level. This supports theories of the importance of the length and quality of L2 input (Flege, 1998; Flege & Liu, 2001) as well as the opportunity for L2 output (Cummins, 2000). However, we need to bear in mind that the VOT values of Dina's stops are still lower than the VOTs of the stops from the second and third generation. Although there is a present influence of English in Dina's speech which is beginning to remove prevoicing from her phonological repertoire, the fact that she still predominantly speaks Gujarati and holds an "emotional attachment to the L1" (Flege et al., 1997, p. 171), means that she maintains VOT values that have not fully assimilated to English's phonological categories.

Dina's results are not the only evidence of VOT values that do not correspond with the hypothesised results. The other three participants from the first generation also use both Gujarati-like and English-like VOTs for voiced stops by moving across both sides of the VOT continuum. For example, Ayesha has negatives values for /b/ and /g/ and positive values for /d/ (Figure 15). Barirah and Dina both prevoice the stop /b/ at utterance-medial position in contrast to the other stops which all have positive VOT (Figure 16). The same can be seen at utterance-final position at which Ayesha, Barirah and Cameelah use both positive and negative VOT. Ayesha at utterance-medial position produces all voiceless and voiced stops with positive VOT: a feature that was earlier stated that Dina uses at utterance-final position as well as a feature that is almost constant for the second and third generation. However, Ayesha, as do Barirah and Cameelah, regularly use only Gujarati and have very little exposure or interaction to English speakers. We know that such interaction with English speakers has an impact on VOT as Dina has higher VOT values than Ayesha's as well as Barirah's and Cameelah's use of short lag voicing 60% of the time. Thus, the theories of the quality of L2 input and opportunity for L2 output proposed by Flege (1998) and Cummins (2000) still stand as Dina is still more English-like in her articulations. However, these theories do not account for why other members of the first generation still produce some voiced stops with English-like values. An explanation for this could simply be due to an unintended influence from English after having lived in England for over forty years.

Although the first generation are native speakers of Gujarati, the separation from their native country for several decades means that it is inevitable for VOT values to be less Gujarati-like and more English-like.

Nonetheless, this does not explain why the VOTs for the first generation's voiced stops remain extremely low at an average of 16.2ms when Gujarati VOT values are typically positioned much higher on the VOT continuum in other studies (Rami et al., 1999). A possible explanation for this could be the flashcards that were presented to the participants. On these flashcards were the target words and sentences written in English. Like many other Indo-Aryan languages, Gujarati marks a phonemic contrast between unaspirated and aspirated stop sounds. In a transliteration of Gujarati speech, this contrast is depicted with a /h/, e.g. /p/ (unaspirated) and /ph/ (aspirated)[31]. It is highly possible, and has been supported by other scholars also (see Rao, 1961), that when these participants read the flashcards, they assumed that the voiceless stops in English – a language that does not mark a phonemic distinction – were meant to be unaspirated, and thus articulated these stops with a lower VOT than they normally would for aspirated sounds. This could also be the reason as to why the findings from this generation are even lower than the findings from Rami et al.'s (1999) study of Gujarati velar stops[32].

5.3. Second and third generation: an overview

In contrast to the first generation, on the opposite end of the spectrum is the second and third generation who all are fluent in both English and Gujarati. The majority of these participants' VOT values for voiceless stops vary between 'slightly aspirated' and 'highly aspirated' (Cho and Ladefoged, 1999), ranging mainly from 40ms to 85ms. However, from all the thirty-six voiceless stops from the second generation, seven of the tokens are what Cho and Ladefoged (1999) would classify as 'unaspirated' due to them having a mean VOT value of 30ms. For the third generation, the number of tokens in the unaspirated range reduces to four. With regard to the voiced stops, seven out of the eight participants from these two generations use a mixture of both positive and negative VOT, constantly switching between the two at different positions of the utterance. The results for these two generations are remarkably similar. Both generations have a mean difference of 4.8ms for the voiceless stops and a 1.8ms difference for the voiced stops (Figure 14). These participants are all born in the UK and learnt Gujarati from their parents, due to the fact that such acquisition of Gujarati is "an integral part of family background (and) a sign of…local identity" (Blom and Gumperz, 2007, p. 76). At the same time, there is constant contact with English, especially from the onset of full-time education. Consequently, the second and third generation are

[31] Note that in a transliteration the diacritic /^{h}/ becomes an orthographic /h/.

[32] Although Rami et al. (1999) only studied velar stops, they found that the unaspirated /k/ had a VOT value of 40.6ms and the aspirated /k^{h}/ had 75ms, both of which are much higher than the results of this study.

raised with two identities: the inherited Gujarati identity that is acquired from birth and the English identity that is acquired from the host community.

However, the reason for the third generation having a higher mean VOT for voiceless stops than the second generation is most probably owing to how often each language is spoken within the home. The members of the second generation were asked what language(s) they were brought up with; all four participants responded with Gujarati. This is unsurprising as the second generation were raised by in-migrant parents who, although learnt to speak basic English, found it easier to use their native language in a foreign country. It was only until the second generation attended preschool that their acquisition of English began. Nonetheless, the immigrant community was still so close-knit that Gujarati remained a language used with neighbours, in local shops as well as in other general community services. On the other hand, when the third generation were asked the same question of what language(s) they were brought up with, all the participants reported a varied mixture of both English and Gujarati, which expectedly is a consequence of their parents – the second generation – being native speakers of English. As a result, it is likely that the higher usage of the host language English within the home has allowed the third generation to produce slightly higher mean VOT values than the second generation who were raised with more Gujarati than English in the household. The following two sections will focus on the second and third generation more comprehensively in an attempt to individually consider both of these age groups' findings.

5.4. Second generation

This section will be split into two subsections. Section 5.4.1. will focus on voiceless stops, and section 5.4.2. will focus on voiced stops.

5.4.1. Voiceless stops

Section 5.1 stated that the second generation's VOT values were higher than the findings for the first generation, and the third generation's voiceless stops were bigger than the values found for the second generation. However, it is in fact Fatimah from the second generation who maintains the highest VOT results for the stops /p t k/ from all twelve participants. Fatimah has the highest VOTs for /p/ and /k/ at utterance-initial and utterance-medial position (Figures 18 and 19) and the highest VOT for /t/ at utterance-final position (Figure 20) than any participant from her age category as well as the participants from the third generation. It is unexpected of a second-generation subject who was raised with only Gujarati in the household to have more English-like results than those participants from the third generation who typically have greater exposure to English, both within and outside the home.

During interviews, the second generation were asked for contextual background; it was revealed that Fatimah is the only member of the second generation who works full-time. Eliza works part-time, Ghazala is a full-time carer for her mother who lives with her and Haafizah is a housewife. These occupational differences are possibly important to consider as they may affect VOT production. Fatimah's profession in a school environment requires her to speak English for most of the day. She also reported that her daughter speaks basic Gujarati, whilst her son can only speak English. As a result, the most used language within the home is English. All three other participants reported using either an equal mixture of both languages or more Gujarati than English[33]. Consequently, this persistent use of English in Fatimah's household in her adult life is a contrast to the predominant usage of Gujarati during her childhood, which has most likely been the reason behind her extremely high VOTs. Pandit (1977) comments that the second-generation speakers of in-migrant parents living in the west renounce their native tongue in favour of the area's most dominant language, which indicates that "language shift is the norm and language maintenance an exception" (1977, p. 9). This remark, on the most part, appears to be accurate and pertinent to the findings of this thesis. The second generation, and by extension the third generation, present an extreme shift from native Gujarati to English by moving out of their heritage community and building networks with a wider range of people from various backgrounds. This is reflected in Fatimah's, as well as all of the other members of the second generation's, English-like VOTs that are dissimilar from that of the first generation.

5.4.2. Voiced stops

Despite the abovementioned marked shift from the first generation to the second generation, members of the second generation still at times articulate voiced stops with the values associated with Gujarati. For example, at utterance-initial position, Ghazala uses negative VOT for all three voiced stops, which we can perhaps assume is due to her position as a full-time carer to her monolingual-Gujarati mother who lives in the same household. This is possibly also why Ghazala has the lowest VOT values for her voiceless stops four out of a possible nine times[34]. Her living conditions of Gujarati being the dominant language of her home has reduced her VOTs, not necessarily down to the VOT values of Gujarati itself[35], but rather to a moderately low long lag/high short lag range. Such voiceless stops parallel the results from Haafizah, who has the lowest VOT values for voiceless stops for the other five times out of nine. Thus, Ghazala and Haafizah together make up all of the times when the lowest VOT for voiceless stops is used. Both of these participants do not have an occupation that requires them to have an extensive number of social networks outside of

[33] The latter was only reported by Ghazala
[34] The nine has been obtained from the fact that /p t k/ is reiterated at three positions of the utterance.
[35] Ghazala's VOT values for voiceless stops are still higher than those from the first generation.

their heritage community. This, combined with the fact that they were born into a close-knit immigrant environment from which they did not move away by means of gaining employment, has most likely allowed them to continuously strengthen the phonological variants of Gujarati. Thus, occupation, or lack of, is a possible covariate of VOT.

Ghazala's and Haafizah's results that have been discussed above imply that occupation may influence VOT production; lack of occupation equals to low VOT which theoretically should mean that having an occupation equals to higher VOT. However, the results from Eliza and Fatimah refute this theory. Eliza and Fatimah who both respectively work part-time and full-time produce voiced stops with negative VOT. Ghazala prevoices *all* three of her voiced stops at utterance-initial position and Fatimah prevoices only two (/b/ and /d/). However, both of Fatimah's stops have higher negative VOT values than Ghazala's (Figure 18). Eliza prevoices the voiced plosive /g/ at utterance-initial position which, although is not bigger than Ghazala's values, is only 6.6ms smaller. It does, however, exceed at least Haafizah's VOT values. Moreover, at utterance-medial position (Figure 19), Eliza and Fatimah both produce one voiced stop each with negative VOT, whilst Ghazala and Haafizah produce all three voiced stops with positive VOT i.e. Ghazala and Haafizah resemble the English-like production. It is unclear why the two participants who work, and who thus conveyed speak English more than Gujarati on a daily basis, present more Gujarati-like tokens than the participants who speak Gujarati more often. Going back to Pandit's (1977) observation that language shift is the norm, Eliza's and Fatimah's findings present something much more complex as there appears to be an attempt at language shift and language maintenance simultaneously. As these are assumptions made from the results of only two participants, we do not know or can say for certain that the results are not coincidental. Further research is required with more participants and with a further focus on social networks.

5.5. Third generation

This section will be split into three subsections. Section 5.5.1. will provide contextual information gathered during interviews with the third generation. Section 5.5.2 will discuss the voiceless stops and Section 5.5.3 will discuss the voiced stops.

5.5.1. Contextual background

Inayah and Laila are university students, whilst Jameelah and Khadijah work full-time. When asked what language is mainly spoken within the home, Inayah reported to only speak Gujarati, Laila responded with mostly English but some occurrences of Gujarati, whilst Jameelah and Khadijah stated an equal mixture of both Gujarati and English. Nevertheless, all four of these participants have at least some exposure to both languages.

5.5.2. Voiceless stops

All four members of the third generation produce high VOT values for voiceless stops and mostly short lag VOT for voiced stops. Using only the provided contextual information above, we could assume that Inayah may theoretically have the most Gujarati-like VOT, or at least lower values than the other participants due to her sole usage of Gujarati at home. Conversely, Laila may hypothetically have the most English-like VOT owing to her high use of English. However, what we find is that Khadijah has the highest VOTs for voiceless stops than any member of the third generation, exceeding over 80ms several times and therefore matching the expected English VOT values (Lisker and Abramson, 1984). This is the case at every position of the utterance as Khadijah has a mean VOT of 65.1ms at utterance-initial (Figure 21), 75.4ms at utterance-medial (Figure 22) and 67.5ms at utterance-final (Figure 23). At the first two positions of the utterance, Laila has the second highest VOT values. At utterance-final position, Inayah, Jameelah and Laila all bear similar findings as Khadijah maintains the highest results.

There appears to be a similarity here between Khadijah from the third generation and Fatimah from the second generation. Both participants work full-time and both display the highest VOT values for voiceless stops within their respective age brackets. However, Jameelah who also works full-time has the lowest mean VOTs at utterance-initial position and the second lowest at utterance-medial and utterance-final position. In order to assess whether there is in fact any significance of occupation in relation to VOT, both these participants were asked more about the language(s) spoken at their place of employment. Whilst Khadijah reported that she is the only Gujarati-speaking employee at her workplace and therefore solely speaks English, Jameelah reported that her peers in her workplace are mainly British-Asian and code-switch between English, Gujarati and occasionally Urdu on a daily basis. Similar to Holmes and Stubbe's (2004) study of code-switching at work, these participants in their workplace use more than one language to "signal and enact their ethnic identities, as well as to construct and reinforce good relationships with members of their own ethnic group in the workplace" (Holmes and Stubbe, 2004, p. 140). It is plausible to suggest that, as the majority of Jameelah's time is spent at the workplace in which she has the opportunity to strengthen her Gujarati identity, the VOT attributes have an influence and are somewhat lower than those participants who speak solely English in the workplace i.e. Khadijah. Therefore, these results from Khadijah and Jameelah not only reinforce that having an occupation is a covariate of VOT, but it provides support for the belief that it is the language(s) spoken within that occupation that is also important.

Inayah and Laila, however, both do not work as they are still university students. It is therefore appropriate to discuss these participants separately. It is firstly important to state that Inayah and Laila spend a large portion of their day at university, and so the amount of time spent away from home is similar to those participants who work full-time (Khadijah and Jameelah). Inayah and Laila both reported to speak English with their friends at

university. This, in addition to her high exposure to English at home, is possibly why Laila mainly has the second highest VOT values at different positions of the utterance. Laila is also the only participant out of all three generations to have positive VOT for all voiceless and voiced stops at every position of the utterance. Therefore, she is the only participant to consistently resemble a monolingual English speaker, which is likely due to the fact that her daily routine encompasses very slight proportions of her heritage language Gujarati. This is indicative of Dunkel's (1948) quote who states that "the 'less' L1 there is, the smaller will be its influence on the L2" (Dunkel, 1948). By this same token, we can also understand the contrasting argument. Inayah, who also goes to university and speaks English with her friends, does not bear the same results as Laila. Although Inayah has the second highest VOT values at utterance-final position, she has the second lowest VOTs at utterance-initial and the lowest at utterance-medial position. Moreover, she prevoices all three of her voiced stops at utterance-initial position. Whilst Laila speaks mainly English with her family, Inayah predominantly speaks in Gujarati. Thus, Dunkel's (1948) theory can be inverted: the more L1 there is, the bigger its influence will be on the L2. The reason why Inayah predominantly speaks in Gujarati at home is due to the fact that she lives with her grandparents. Not only does she have exposure to Gujarati through her parents like the other participants, she also has constant exposure from native Gujarati speakers belonging to the first generation. Therefore, Inayah's negative VOT values that are Gujarati-like have possibly been influenced by the sociological make-up of her household. Although there is an influence of English from university, her influence of native Gujarati is still present by means of her relatively low voiceless stops as well as the prevoicing in her voiced stops.

5.5.3. Voiced stops

Voiced stops have been briefly discussed in Section 5.5.1., however, has only been in regard to Inayah's results. Inayah is not the only participant who presents voiced stops on the negative side of the VOT continuum. Jameelah and Khadijah also articulate voiced stops with negative VOT at utterance-initial position (Figure 21). However, Jameelah's values do not surpass -6ms and are thus very close to near zero VOT[36]. In contrast to this, Khadijah articulates all three voiced stops with negative VOT with values exceeding over -60ms. Despite her overall mean VOT value (-47.1ms) not being higher than Inayah's mean value (-54.1), Khadijah still presents higher negative values for the stops /d/ and /g/ than Inayah. As aforementioned, Inayah has exposure to Gujarati from speakers born in Gujarat i.e. her grandparents. She therefore speaks solely Gujarati at home. However, Khadijah speaks a mixture of both languages at home and speaks only English in her workplace. Therefore, it is curious why she has the highest, most English-like VOT values for voiceless stops and the lowest, most Gujarati-like VOT values for her voiced stops. These findings precisely correspond with Fatimah's, and to an extent Eliza's result (second generation). All these

[36] Jameelah articulates /d/ with positive VOT, thus her overall mean VOT at utterance-initial position is 3.7ms.

participants speak English as their dominant language, however possess Gujarati-like negative VOTs for voiced stops. This appears to be a strong support for the single-system hypothesis and a refutation of many scholars', such as Flege et al.'s (2002), assertion that L2-dominant bilinguals suppress language interference. Rather, the findings indicate that despite being dominant speakers of English, members of the second and third generation are still affected by their non-dominant Gujarati L1. These speakers cannot fully separate the phonetic categories of their L1 from their L2, which therefore triggers the occurrence of 'phonetic bleeding' or 'incomplete suppression' (Antoniou et al., 2010, p. 560) between the two languages. It is unclear *why* these robust differences between English-like VOT for voiceless stops and Gujarati-like VOT for voiced stops is only observed in a few participants; this particular finding will be revisited in Section 5.7.1.

5.6. Sentence position of tokens

So far in this chapter, the sociolinguistic factors have been considered, however none has been said about the influence of syntax. Lisker and Abramson (1967) state in their cross-language study that "…there appeared to be no relation between our measure and the position of the stop, whether position is defined in relation to syntactic constituents or merely by reference to the number of preceding or following words or syllables in the sentence". (Lisker and Abramson, 1967, p. 20). It needs to be kept in mind that the variety of sentence types in Lisker and Abramson's (1964) study was limited, and therefore does not stand to be reliable nor irrefutable.

With regard to the findings of this study which considered three places of position in an utterance, there is no discernible pattern for the voiceless stops. However, for the voiced stops there does appear to be a connection between VOT and the positioning of the target word. At utterance-initial position, prevoicing can be seen for all voiced stops and all three generations (Figures 24 and 25) as the average is -19.6ms. At utterance-medial position, the voicing dramatically becomes more English-like[37] (Figures 26 and 27) by moving over 30ms to the positive side of the VOT continuum and thus holding an average of 11.2ms. At utterance-final position, due to the occurrence of prevoicing from the first generation, the mean VOT marginally falls to 10.2ms (Figures 28 and 29). This shows that utterance-medial followed by utterance-final position produce the most English-like articulations. It is highly possible that the English utterance that syntactically surrounds the target word triggers the speakers to articulate the word with a VOT that matches the phonological condition of the rest of the sentence. This observation bears resemblance to Grosjean's Language Mode Framework (1982, 2001, 2008) in which he argues that both languages of a bilingual are within one system, and are activated at different times. He states that certain linguistic

[37] The stop sound /b/ from the second generation is still prevoiced at utterance-medial position (-21.ms). However, it is a decrease from the result at utterance-initial position (-14.7) and therefore does become more English-like.

factors such as the context of interaction activates one language whilst temporarily deactivating the other. Thus, the unilingual English context of the utterances during the recording process possibly activated the participants' English mode whilst the Gujarati mode became dormant. This theory is reinforced by the fact that the utterance-initial position creates the most mixed results and the most prevoicing. Using the Language Mode Framework, it is conceivable that, as there is no prior syntactic influence from English at utterance-initial position, the uni-lingual English context had not yet been established, thus causing many tokens of the target word to be produced with negative VOT. However, as there are still some results at utterance-initial position with positive VOT, and at utterance-medial and final position with negative VOT, it does imply that a dormant language during a unilingual context is not entirely deactivated, and thus a bilingual can never fully function in an exclusively monolingual mode.

5.7. Sociocultural implications

Although numerous rationalisations and theories have been discussed, it is conceivable that the results may not be swayed by the participants' personal circumstances but rather by something more societal. This certain inference refers solely to the voiced stops. The voiceless stops follow the hypothesised pattern of increasing as the generations become younger. However, the findings for the voiced stops are much more complex. All three generations prevoice mainly at utterance-initial level (Figures 24 and 25). This is explicable for the in-migrant first generation who tend to prevoice at all positions, however it is not expected of the second and third generation. Moreover, those participants who prevoice in the two British-born generations primarily only do so at utterance-initial position[38]. Indeed, this could be due to the above reasoning regarding the Language Mode Framework (Grosjean, 1982, 2001, 2008), however, it could alternatively be due to a wider transformation occurring within the Gujarati-English community. It is plausible that this immigrant community, whose language is the second most spoken after English in Batley East (QPZM, 2017), have become so accustomed to the exposure to both languages within and outside of the home that it has blurred the lines of each language. Similar to Mufwene's (2001) feature pool, the phonetic sets of English and Gujarati have, through language contact, become input varieties and speakers within this community are free to pick different combinations of features from the pool to develop their output variety. Therefore, the sociocultural makeup of Batley has made it so that there is no pressure for British-born members of this community to articulate stops with native accuracy of each language, but rather it is acceptable to use both positive and/or negative VOT irrespective of which of the two languages is being spoken.

[38] This is excluding two tokens from utterance-medial position in Figure 19.

5.7.1. Identity

Building upon the previous section, it is possible that this acceptance of using both languages relates to identity that is stronger within the younger generation. Both the second and third generation prevoice at utterance-initial level. Whilst there is no single participant that articulates all of /b d g/ with negative VOT in the first generation (Figure 15), there is one participant in the second generation that does so (Figure 18), which increases to two participants in the third generation (Figure 21). Moreover, the third generation also appear to be more Gujarati-like in terms of their VOT values. Khadijah's and Inayah's articulations of voiced stops (Figure 21) are more negative than all articulations from the first[39] and second generation. More generally, for six out of the nine voiced stops across all three positions of the target word, the third generation have a higher negative VOT than the second generation, as depicted in Figure 14 and more precisely in Figures 24-29. It is confusing as to why the articulation of voiced stops becomes more accurate to the Gujarati pronunciation as the generations become younger[40]. This broad trend resembles Eckert's (1988) study in which "The progress of phonological change is…reflected in age differences within communities, with a general increase of innovative forms as one moves downward through the age continuum" (1988, p. 184). However, in this case, the 'innovative form' is not something new, but rather something old, i.e. from native Gujarati.

Additionally, it is interesting that the speakers who have the highest positive VOT values for voiceless stops have the highest negative VOT values for voiced stops. Fatimah, Khadijah and Eliza who reported to speak the most English whilst speaking very little or average amounts of Gujarati have the most Gujarati-like VOT for /b d g/. Although the earlier justification provided for this finding was the Language Mode Framework (Grosjean, 1982, 2001, 2008), it may be possible that these participants are linguistically portraying their ethnic identity. These participants were asked whether they were satisfied with the amount of Gujarati that they speak and the number of opportunities that they have for speaking it. All responded with no and all wished that they could it speak it more often. Therefore, this could be a possible indication that, as accent can be a marker of identity, such participants want to associate themselves with their native tongue and distinguish themselves in some way from their monolingual English contemporaries. Of course, this is not a complete separation from their non-Gujarati peers as these participants still have the highest, most English-like VOTs for their voiceless stops. However, for voiced stops, the Gujarati-like VOT could be a linguistic struggle against complete westernisation. Eckert (1988) reports that adolescents want to "develop an identity independent of the family

[39] This is discounting one token from Cameelah in the first generation (Figure 15).

[40] We need to bear in mind that the higher negative VOT of the third generation than the second generation is largely due to two participants' – Inayah's and Khadijah's – extremely high values, whilst Jameelah's and Laila's results are positioned at near-zero or positive VOT. Further studies are needed with a greater number of participants in order to clarify whether the third generation do in fact possess more Gujarati-like voiced stops than the second generation.

structure" which is "based on their characteristics as individuals" (1988, p. 187). In this study, it appears to be that the participants support and at the same time refute Eckert's two aforementioned observations. From the small number of participants in this study, it seems that the British-born participants have an independent English identity by means of their voiceless stops, yet maintain their ethnic Gujarati identity by means of their voiced stops, thus representing a dual identity. We cannot put forward that this is representative of the entire Gujarati community. However, we can therefore surmise that there may be an active exertion, whether that is consciously or subconsciously, from those participants who are impelled to use English more than Gujarati to hold onto their Gujarati roots. However, the fact that these VOT values are even more native-like than the native speakers of Gujarati could be to overstress the contrast between the voiceless and voiced stops and thus clearly emphasize the dual identity. Why this representation of a dual identity is only marked at utterance-initial position is as of yet unclear. Further research is warranted to examine this, with more speakers and adolescents from the Gujarati community, to understand the full sociological makeup of this town and its impact on VOT production. Without a study with a greater number of participants, all of the above explanations cannot be considered to be reliable assertions.

6. CONCLUSION

This thesis set out to examine whether VOT transfer is present between Gujarati-English speakers across three generations as well as across three positions of an utterance. It also aimed to evaluate what may provide as rationalisations for the patterns that were observed. With the information presented in the previous two chapters, it is now possible to answer the research questions that were specified in the Introduction. Each research question is individually addressed in the subsections below.

6.1. Research Questions

6.1.1. What are the word-initial VOT values for /p b t d k g/ of the data across the three generations, and do such values correspond more with English or with Gujarati VOTs?

The word-initial VOT values for /p b t d k g/ have been comprehensively illustrated in Chapter 4. In brief, there appears to be short-lag VOTs for voiceless stops and a predominant use of prevoicing for voiced stops in the first generation. The second and third generation bear similar results as both age categories largely use long lag VOTs for voiceless stops and short lag VOTs for voiced stops. The third generation have slightly higher VOT values for voiceless stops than the second generation, however, the second generation have marginally higher VOT values for voiced stops. Therefore, with regard to the language with which the VOT values correspond, the findings present a split between the Gujarat-born first migrants and the British-born second and third generation. The former's results correspond more with Gujarati VOTs whilst the latter's findings more closely resemble English VOTs.

6.1.2. Are there any variances in word-initial VOT at different places in an utterance, and is there any significance of these variances?

No meaningful or substantial variances were observed for any of the voiceless stops /p t k/ across the generations, but were observed for the voiced stops /b d g/. Whilst utterance-medial and utterance-final position generally yielded English-like VOT values, the utterance-initial position appeared to be inconsistent with this. All three generations were typically observed at utterance-initial position to articulate voiced stops with negative VOT. The Discussion put forth the theory that this observation may be due to no prior lexical influences at utterance-initial position (see Grosjean, 1982, 2001, 2008), whereas there are such influences at utterance-medial and utterance-final position. However, the main significance of prevoicing at utterance-initial position is that it may possibly be a reflection of the sociocultural makeup of Batley in which the third generation are actively attempting to keep hold of Gujarati-like VOTs in order to index their ethnic identity.

6.1.3. Are there any sociolinguistic stimuli/influences that affect the transfer of VOT for any of the participants?

Aside from the aforementioned theory regarding the change occurring in the Gujarati-English community, other stimuli were also observed and discussed in the previous chapter. These include the country of birth/the age of learning the L2 (Kang & Guion, 2006), the amount of direct exposure to an L1/L2 (Flege & Liu, 2001), the opportunity for its output (Cummins, 2000), as well as the dominant language used within the home or out of the home, e.g. at work/university (see Piske et al. 2001). Some of these influences appeared to have a greater sway than others. Generally, the age of learning the L2 combined with the dominant language used in and out of the home had the most significant effects on VOT transfer. This reinforces previous literature from Flege (1991), Kang & Guion (2006) and Piske et al. (2001) who all considered an early age of acquiring the L2 is paramount to its fluency and native-like articulation. As English is more frequently used than Gujarati as the generations become younger, the implication of this study is that there is a possibility of language attrition occurring in the future. However, the second implication is that the Gujarati-like negative VOT may possibly be maintained.

6.2. Limitations and shortcomings of the study

Despite the care and caution that was taken to complete this research, there are undoubtedly limitations. It is therefore paramount that the methodological shortcomings are considered.

The main drawback of this thesis that impedes authenticity of the results is the sample size. A size of twelve participants was tolerated for the purpose of methodological ease due to the fact that such participants generated eight hundred and forty-four tokens for analysis. This, although it appears extensive on the outset, was a practical amount of data to work with. A larger sample size would expectedly mean more tokens and thus necessitate a greater amount of time to analyse. However, it would also mean a more accurate representation of the Gujarati-English community in focus to whom the results would be generalised.

Secondly, a shortcoming, that is not an inherent limitation of the study, is a lack of literature from monolingual Gujarati speakers. Lisker and Abramson's (1964), Docherty's (1992) and Rami et al.'s (1999) VOT values of English and Gujarati were specified in the Literature Review as a base rate for the results of this research. However, these scholars' findings are dated and focus only on standard dialects of the languages that they researched. It would have therefore been more relevant as well as more current to record monolingual speakers of English and Gujarati living in the same community as the bilingual participants in order to fully comprehend the similarity between the participants' VOTs and the two languages English and Gujarati. Without these recordings, it is difficult to accurately verify the extent to which the participants' productions of stops differed from their monolingual counterparts.

A further shortcoming of the data is the lack of questionnaires or a contextual interview given at the same time as the data collection recording procedure. Although participants were approached again and were asked for basic details regarding their background, this was only a brief and informal conversation. There would have been greater benefit to provide questionnaires, as this not only would have efficiently supplied the required details without having to re-contact the participants, it would have also provided the opportunity to gain participants' personal opinions regarding their affiliation to their L1/L2 and the identity that they deem themselves to have.

6.3. Areas for future research

Should this area of research concerning Gujarati-English and VOT be expanded in the future, the primary modifications should be based on the aforementioned limitations of the present study. As future research may not be constricted by the same time restrictions as this thesis, it is essential that it includes more participants as well as more tokens. This will support whether the reasonings for the results that are made in this study are in fact accurate. In particular, the results regarding the use of negative VOT by all three generations were somewhat surprising. A follow-up study with more participants and a focus on this abovementioned result will determine whether our intuitions regarding a sociocultural change is in fact occurring or not.

This thesis has been able to examine the influence Gujarati (the L1) has on English (the L2) with regards to VOT values in English speech. Whilst it has effectively managed to address this, the matter that this research has not been able to resolve is the influence English has on Gujarati. Therefore, future studies should aim to record Gujarati speech to investigate what changes, if any, are occurring in the ethnic language.

As this thesis exclusively focused on British-Gujarati females, there is a gap in the literature regarding males as well as children. Studying such groups in the future will provide a more well-rounded understanding of the Gujarati-English community in Batley as opposed to the current study which could only focus on a few socio-demographic qualities. Future studies could also try to incorporate more various styles of obtaining data before examining other dialects and other geographical communities.

More broadly, the body of the data that currently exists for Gujarati is severely limited. This thesis regarding VOT has contributed to the literature that has so far discussed only breathy voice. Therefore, any studies that are undertaken in the future regarding any dialect or feature of Gujarati in relation to any field of linguistics, be it acoustics or not, will be extremely valued.

7. REFERENCES

- Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2010). Language context elicits native-like stop voicing in early bilinguals' productions in both L1 and L2. *Journal of Phonetics*. 38 (4), 640-653.
- A little piece of India that's forever Yorkshire. (2016, May 29). *The Yorkshire Post.* Retrieved from http://www.yorkshirepost.co.uk/news/a-little-piece-of-india-that-s-forever-yorkshire-1-7928405 [Accessed 2017, June 3].
- Adams, M. R. (1987). Voice onsets and segment durations of normal speakers and beginning stutterers. *Journal of Fluency Disorders.* 12 (2), 133-139.
- Bailey, G., Wikle, T., Tillery, J., & Sand, L. (1991). The apparent time construct. *Language Variation and Change.* 3, 241-264.
- Bates, C. (2003). The hidden story of partition and its legacies. Retrieved from http://www.bbc.co.uk/history/british/modern/partition1947_01.shtml
- Bohn, O. S., & Flege, J. E. (1993). Perceptual switching in Spanish/English bilinguals. *Journal of Phonetics.* 21 (3), 267-290.
- Blom, J. & Gumperz, J.J. (2007). Social meaning in linguistic structure: code-switching in Norway. In L. Wei (Ed.) The Bilingualism Reader. (2nd ed.) (pp. 75-96). London: Routledge.
- Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America*. 54, 421-428.
- Cardona, G. (1965). A Gujarati Reference Grammar. Philadelphia: The University of Pennsylvania Press.
- Census of India. (2011). *Ministry of Statistics and Programme Implementation.* Retrieved from http://mospi.nic.in/mospi_new/upload/SYB2013/ch2.html [Accessed 2016, September 16].
- Chaudhuri, K. N. (1999). The English East India Company: The Study of an Early Joint-stock Company 1600-1640. London: Routledge.
- Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics.* 27, 207-229.
- Communities and Local Government (2009). The Indian Muslim Community in England: Understanding Muslim Ethnic Communities. London: Department for Communities and Local Government.
- Cummins, J. (2000). Language, power and pedagogy: Bilingual children in the crossfire. Clevedon, UK: Multilingual Matters.
- De Decker, P., & Nycz, J. (2011). For the record: Which digital media can be used for sociophonetic analysis? *University of Pennsylvania Working Papers in Linguistics.* 17, 51-9.

- Discoveredindia.com. (2017). Gujarat Attractions - A Memorable Replica of Glowing India. Retrieved from http://www.discoveredindia.com/gujarat/attractions/ [Accessed 2017, May 30].
- Discoveredindia.com. (2017). Regions in Gujarat. Retrieved from http://www.discoveredindia.com/gujarat/about-gujarat/regions-in-gujarat.htm [Accessed 2017, June 2].
- Docherty, G. J. (1992). The timing of voicing in British English Obstruents. Foris Publications: New York.
- Dunkel, H. (1948). Second Language Learning. Boston: Ginn and Company.
- Eckert, P. (1997). Age as a sociolinguistic variable. The handbook of sociolinguistics, ed. by Florian Coulmas, 151–67. Oxford, UK: Blackwell.
- Eckert, P. (1988). Adolescent social structure and the spread of linguistic change. *Language in Society*. 17 (2), 183-207.
- Education Act 1944 (7 & 8 Geo 6 c. 31). London: HMSO.
- Ember, M., Ember, C. R., & Skoggard, I. (2005). Encyclopaedia of Diasporas: Immigrant and refugee cultures around the world. Springer, United States of America.
- Esposito, C. M., & Khan, S. U. D. (2012). Contrastive breathiness across consonants and vowels: A comparative study of Gujarati and White Hmong. *Journal of the International Phonetic Association.* 42 (2), 123-143.
- Fellbaum, M. L. (1996). The acquisition of voiceless stops in the interlanguage of second language learners of English and Spanish. Proceedings of the 4th International Conference on Spoken Language Processing on 3rd October 1996, Philadelphia, USA. 3, 1648-1651.
- Firth, J. R. (1967). Phonetic observations on Gujarati. *Bulletin of the School of Oriental and African Studies.* 20 (1), 231-241.
- Fischer- Jørgensen, E. (1954). Acoustic Analysis of Stop Consonants. *Miscellanea Phonetica.* 42-59.
- Fischer- Jørgensen, E. (1967). Phonetic analysis of breathy (murmured) vowels. *Indian Linguistics.* 28, 71-139.
- Flege, J. E. (1987). The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics.* 15, 47-65.
- Flege, J. E. (1991). Age of learning affects the authenticity of voice onset time (VOT) in stop consonant produced in a second language. *Journal of the Acoustical Society of America. 89,* 395-411.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), Speech perception and linguistic experience: Issues in cross-language research (pp. 233–277). Baltimore: York Press.
- Flege, J. E. (1998). The role of subject and phonetic variables in L2 speech acquisition. In M. Gruber, D. Higgins, K. Olsen, & T. Wysocki (Eds.) Papers from the 34th Annual Meeting of the Chicago Linguistic Society (pp. 213-232). Volume II, The Panels. Chicago: Chicago Linguistic Society.

- Flege, J. E., & Efting, W. (1987a). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*. 15, 67–83.
- Flege, J. E., & Efting, W. (1987b). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*. 6, 185–202.
- Flege, J. E., Frieda, E. M., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics*. 25 (2), 169-186.
- Flege, J. E., & Hammond, R. M. (1982). Mimicry of non-distinctive phonetic differences between language varieties. *Studies in Second Language Acquisition.* 5 (1), 1-17.
- Flege, J. E., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition.* 23, 527-552.
- Flege, J. E., MacKay, I. R. A., & Piske, T. (2002). Assessing bilingual dominance. *Applied Psycholinguistics*. 23, 567–598.
- Flege, J. E., & Port, R. (1981). Cross-Language Phonetic Interference: Arabic to English. *Language & Speech*. 24 (2), 125-145.
- Fowler, C. A., Sramko, V., Ostry, D. J., Rowland, S. A., & Halle´, P. A. (2008). Cross language phonetic influences on the speech of French–English bilinguals. *Journal of Phonetics*. 36, 649–663.
- Fromkin, V., Rodman, R., & Hyams, N. (2013). An Introduction to Language. Boston, MA: Wadsworth Publishing.
- Generation [Def. 4]. (2009). In *Oxford English Dictionary.* Retrieved from http://www.oed.com.libaccess.hud.ac.uk/view/Entry/77521?redirectedFrom=generation#eid [Accessed 2017, July 2].
- Ghosh, S. (2013, August 14). Gujarat to get 7 new districts on Independence Day. *One India.* Retrieved from http://www.oneindia.com/2013/08/14/gujarat-to-get-7-new-districts-on-independence-day-1283210.html [Accessed 2017, June 16].
- Google Maps. (2017). Batley. Retrieved from https://www.google.co.uk/maps/place/Batley/@53.7191298,-1.6649035,13z/data=!3m1!4b1!4m5!3m4!1s0x48795fe289ab4b01:0x9d69a51635cd025d!8m2!3d53.717028!4d-1.635083 [Accessed 2016, December 28].
- Government of India. (1968). The Official Language Resolution. *Ministry of Home Affairs.* Retrieved from https://web.archive.org/web/20100318123802/http://www.rajbhasha.gov.in/dolresolutioneng.htm [Accessed 2017, May 31].
- Grimes, B. F. (1996). Ethnologue: Languages of the World *(13th Edition).* Summer Institute of Linguistics: Dallas.
- Grosjean, F. (1982). Life with two languages: An introduction to bilingualism. Cambridge: Harvard University Press.
- Grosjean, F. (2001). The bilingual's language modes. In Nicol, J. (Ed.). One Mind, Two Languages: Bilingual Language Processing (pp. 1-22). Oxford: Blackwell.
- Grosjean, F. (2008). Studying Bilinguals. Oxford: Oxford University Press.

- Harada, T. (2007). The production of voice onset time (VOT) by English-speaking children in a Japanese immersion program. IRAL, International Review of Applied Linguistics in Language Teaching, 45(4), 353-378.
- Heselwood, B., & McChrystal, L. (1999). The effect of age-group and place of L1 acquisition on the realisation of Punjabi stop consonants in Bradford: An acoustic sociophonetic study. *Leeds Working Papers in Linguistics.* 7, 49-69.
- Hewlett, N., & Beck, J. M. (2006). An introduction to the science of phonetics. Mahwah, NJ: Lawrence Erlbaum Associates.
- History of Indian Muslim Welfare Society. 50 years of excellence. (2007). *Indian Muslim Welfare Society.* Retrieved from http://217.199.187.65/imws.org.uk/wp-content/uploads/2013/10/imwshistory.pdf [Accessed 2017, June 3].
- Holmes, J. & Stubbe, M. (2004). Strategic code-switching in New Zealand workplaces: Scaffolding, solidarity and identity construction. In J. House & J. Rehbein (Eds.) *Multilingual Communication* (pp. 133-154). Philadelphia: John Benjamins Publishing Company.
- Jivraj, S., & Finney, N. (2013). Local dynamics of diversity: Evidence from the 2011 Census. *Centre on Dynamics of Ethnicity (CoDE).* Manchester: The University of Manchester.
- Jubb, S. (1860). The History of the Shoddy-trade: Its Rise, Progress and Present Position. London: Houlston and Wright.
- Kang, K., & Guion, S. G. (2006). Phonological systems in bilinguals: Age of learning effects on the stop consonant systems of Korean-English bilinguals. *The Journal of the Acoustical Society of America*. 119 (3), 1672-1683.
- Keating, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language.* 60, 286-319.
- Keating, P. A., Linker, W., & Huffman, M. (1983). Patterns in allophone distribution for voiced and voiceless stops. *Journal of Phonetics.* 11, 277-290.
- Kehoe, M. M., Lleo, C., & Rakow, M. (2004). Voice onset time in bilingual German-Spanish children. *Bilingualism.* 7 (1), 71-88.
- Khan, S. U. D. (2012). The phonetics of contrastive phonation in Gujarati. *Journal of Phonetics.* 40 (6), 780-795.
- Khattab, G. (2000). VOT production in English and Arabic bilingual and monolingual children. *Leeds Working Papers in Linguistics*. 8, 95-122.
- Krieman, J., Gerratt, B. R., & Khan, S. U. D. (2010). Effects of native language on perception of voice quality. *Journal of Phonetics.* 38 (4), 588-593.
- Labov, W. (1963). The social motivation of a sound change. *Word.* 19, 273-309.
- Lein, T., Kupisch, T., & van de Weijer, J. (2016). Voice onset time and global foreign accent in German–French simultaneous bilinguals during adulthood. *International Journal of Bilingualism.* 20(6), 732-749.
- Lenneberg, E. H. (1967). Biological foundations of language. New York: Wiley.
- Lewis, M. P. (2009). Ethnologue: Languages of the World *(16th Edition).* Summer Institute of Linguistics: Dallas. Retrieved from https://www.ethnologue.com/language/guj [Accessed 2017, March 13].

- Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word.* 20 (3), 384-422.
- Major, R. C. (1992). Losing English as a first language. *The Modern Language Journal.* 76, 190-208.
- Lisker, L., & Abramson, A. S. (1967). Some effects of context on voice onset time in English stops. *Language and Speech*. 10 (1), 1-28.
- Mattausch, J. (2001). After 'ethnicity': Migration, identity and political economy. *Immigrants & Minorities.* 20 (3), 59-74.
- Mapsofindia.com (2014). Gujarat District Map. Retrieved from http://www.mapsofindia.com/maps/gujarat/gujarat.htm [Accessed 2017, May 31].
- Mapsofindia.com. (2017). History of Gujarat. Retrieved from https://web.archive.org/web/20100526165223/http://www.mapsofindia.com/gujarat/history/ [Accessed 2016, September 21].
- McCarthy, K. M., Evans, B. G., & Mahon, M. (2013). Acquiring a second language in an immigrant community: The production of Sylheti and English stops and vowels by London-Bengali speakers. *Journal of Phonetics.* 41 (5), 344-358.
- Messina, A. (2001). The Impacts of Post-WWII Migration to Britain: Policy Constraints, Political Opportunism and the Alteration of Representational Politics. *The Review of Politics*. 63 (2), 259-285.
- Mufwene, S. S. (2001). The ecology of language evolution. Cambridge: Cambridge University Press.
- Nair, U. (1979). Gujarati phonetic reader. Mysore: Central Institute of Indian Languages.
- National Congress of Gujarati Organisations UK (2012). Gujaratis in Britain. Retrieved from http://www.ukncgo.org/gujaratis_britain.html [Accessed 2017, April 3].
- Newlin-Åukowicz, L. (2014). From interference to transfer in language contact: Variation in voice onset time. Language Variation and Change. 26, (3), 359-385.
- ONS (2013a). 2011 Census: Quick Statistics for England and Wales, March 2011. *Office for National Statistics*
- ONS (2013b). 2011 Census: Detailed Characteristics for England and Wales, March 2011. *Office for National Statistics.*
- ONS (2013c). 2011 Census: Language in England and Wales, 2011. *Office of National Statistics.*
- ONS (2013d). 2011 Census: Ethnic group, local authorities in the United Kingdom. *Office of National Statistics.*
- Pandit, P. B. (1961). Historical phonology of Gujarati vowels. *Language.* 37 (1), 54-66.
- Pandit, P. B. (1977). Language in a plural society: the case of India. Delhi: Devraj Chanana Memorial Committee.
- Patil, H. A., Madhavi, M. C., Malde. K. D., & Vachhani, B. B. (2012). Phonetic transcription of fricatives and plosives for Gujarati and Marathi languages. *International Conference on Asian Language Processing.* Hanoi, 177-180.
- Piske, T., MacKay, I. R., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics.* 29, 191-215.

- Podesva, R., & Sharma, D. (2013). Research methods in linguistics. Cambridge: Cambridge University Press.
- QPZM. (2017). Kirklees Census Demographics United Kingdom. Retrieved from http://localstats.co.uk/census-demographics/england/yorkshire-and-the-humber/kirklees. [Accessed 2017, August 1].
- Rami, M. K., Kalinowski, J., Stuart, A., Rastatter, M. P. (1999). Voice onset times and burst frequencies of four velar stop consonants in Gujarati. *The Journal of the Acoustical Society of America.* 106 (6), 3736-3738.
- Rao, S. R. (1985). Lothal. Archaeological Survey of India: New Delhi.
- Rao, K. S. N. (1961). A footnote to the Indian pronunciation of the initial /k,t,p/ and /v/ and /w/ in English. *Indian Linguistics*. 22, 160-3.
- Reid, S. (2015, June 15). The breeding ground for jihadis where even the ice cream lady wears a burka: How Dewsbury, the once great textile town of the North, has undergone a terrifying transformation. *The Daily Mail.* Retrieved from http://www.dailymail.co.uk/news/article-3125530/The-breeding-ground-jihadis-ice-cream-lady-wears-burka-great-textile-town-Dewsbury-undergone-terrible-transformation.html [Accessed 2016, December 19].
- Roach, P. (2004). British English: Received Pronunciation. *Journal of the International Phonetic Association.* 34 (2), 239-245.
- Romaine, S. (1995). Bilingualism. (2nd ed.). Oxford: Blackwell.
- Rumbaut, R. G. (2004). Ages, Life Stages, and Generational Cohorts: Decomposing the Immigrant First and Second Generations in the United States. *International Migration Review.* 38 (3), 1160-1205.
- Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics.* 25, 421-436.
- Sankoff, (2002). Linguistic outcomes of language contact. In J. K. Chambers, P. Trudgill, N. Schilling-Estes (Eds.) The Handbook of Sociolinguistics (pp. 638 – 668). Oxford: Basil Blackwell.
- The Constitution of India. (1950). Article 344(1), 351. *Eighth Schedule.* The Constitution (Ninety-Second Amendment) Act. Retrieved from http://www.constitution.org/cons/india/p17345.html [Accessed 2017, May 30].
- The Constitution of India. (1950). Article 345(17). Official Language; Language of the Union. Retrieved from http://www.constitution.org/cons/india/p17345.html [Accessed 2016, December 20].
- The Yorkshire Post. (2016, June 29). BNP posts Jo Cox 'Muslim' slur through neighbours' letterboxes as racist attack soar. Retrieved from http://www.yorkshirepost.co.uk/news/crime/bnp-posts-jo-cox-muslim-slur-through-neighbours-letterboxes-as-racist-attacks-soar-1-7988476 [Accessed 2016, December 19].
- Thompson, I. (2015). Gujarati. Retrieved from http://aboutworldlanguages.com/gujarati [Accessed 2016, December 20].
- Turner, R. L. (1921). Gujarati Phonology. *Journal of the Royal Asiatic Society of Great Britain and Ireland.* 3, 329-365.

- Vuppala, A. K., & Bhaskararao, P. (2014). Automatic detection of breathy voiced vowels in Gujarati speech. *International Journal of Speech Technology.* 17 (2), 75-82.
- Wasey, A. (2014). 50th Report of The Commissioner for Linguistic Minorities in India (July 2012 to June 2013). *Ministry of Minority Affairs.* 28, 118-122.

8. APPENDICES

8.1. Appendix 1: Data

/p/

1. Pen

a) Pens are blue
b) That pen is new
c) Please press pen

2. Paper

a) Papers are white
b) Damp paper is bad
c) Get yourself paper

3. Pyjamas

a) Pyjamas are spotty
b) Pink pyjamas are the best
c) I like both pictures

4. Picture

a) Pictures are clear
b) I take pictures for fun
c) I like both pictures

/b/

1. Biscuit

a) Biscuits are dry
b) Burnt biscuits are no good
c) She makes biscuits

2. **Bag**
   a) Bags are nice
   b) The best bag is mine
   c) I want a waterproof bag

3. **Banana**
   a) Bananas are healthy
   b) The cheap bananas are big
   c) That is her fourth banana

4. **Bottle**
   a) Bottles are full
   b) I keep bottles at home
   c) That is a glass bottle

/t/

1. **Table**
   a) Tables are dirty
   b) The dark table is there
   c) That is a posh table

2. **Towel**
   a) Towels are soft
   b) The wet towel is white
   c) That is the chef towel
3. **Train**
   a) Trains are long
   b) The quick train has gone
   c) Where is the south train

4. **Tomato**
   a) Tomatoes are red
   b) Chop tomatoes now
   c) I like to crush tomatoes

/d/

1. **Dog**
   a) Dogs are noisy
   b) That dog is big
   c) He is a cross dog

2. **Doctor**
   a) I like doctors
   b) The quiet doctor left

3. **Dress**
   a) I like that dress
   b) The damp dress is here

4. **Door**
   a) I like doors
   b) The back door is closed

/k/

1. **Kitchen**
   a) I like her kitchen
   b) Her kitchen is clean

2. **Kebab**
   a) I like kebabs
   b) That kebab is spicy

3. **Carpet**
   a) I like that carpet
   b) His carpet is white

4. **Colour**
   a) I like this colour
   b) That colour is white

/g/

1. **Garden**
   a) I like gardens
   b) The top garden is clean

2. **Gate**
   a) I like that gate
   b) The black gate is open

3. **Goat**
   a) I like goats
   b) That goat is mine

4. **Grapes**
   a) I like grape.
   b) Take grapes with you

8.2. Appendix 2: Data results

This section presents the VOT for all eight hundred and sixty-four tokens of data. Each participant's data is displayed in a different figure.

	VOT (m/s)					
	VOICELESS			VOICED		
	Sentence-Initial	Sentence-Medial	Sentence-End	Sentence-Initial	Sentence-Medial	Sentence-End
pen	19.5	6.9	4.7			
paper	10.1	6.6	12.7			
pyjamas	4.6	7.8	10			
picture	5.2	9.1	8.6			
biscuit				-143.2	N/A	-39.9
bag				0	0	0
banana				-57	N/A	4.6
bottle				7.4	15.7	-40.5
table	11.5	12.8	18			
towel	12.1	14.2	12.4			
train	11.9	17.6	23.8			
tomato	4.5	6.6	9.7			
dog				0	4.8	N/A
doctor				2.4	8.1	0
dress				15.7	17.7	24.8
door				25.9	0	0
kitchen	10.6	13.2	25.4			
cat	34.9	33.4	31			
carpet	21.9	30.5	30.3			
key	15.2	22.5	4.5			
garden				-31.9	16.3	10.9
gate				23.3	0	29.2
goat				-84.2	19.9	-65.9
grapes				-43.4	17.2	-45

Table 8: A table showing all of the VOT results gathered for Ayesha.

	VOT (m/s)					
	VOICELESS			VOICED		
	Sentence-Initial	Sentence-Medial	Sentence-End	Sentence-Initial	Sentence-Medial	Sentence-End
pen	0	7.9	0			
paper	10.1	0	14.4			
pyjamas	14.2	13.3	6.3			
picture	3.7	9.7	7			
biscuit				0	0	0
bag				0	0	0
banana				0	0	0
bottle				0	1.7	0
table	12.2	10.1	8.1			
towel	6.3	5.5	22.3			
train	23	18.6	38.4			
tomato	9.4	13.5	10			
dog				8.9	6.5	0
doctor				-26.9	8.5	0
dress				-32.2	14.7	-101
door				5.9	-51.9	0
kitchen	10.7	25.4	22.2			
cat	21.9	25.7	20.1			
carpet	25.6	16.4	18.9			
key	24.8	27.2	18.6			
garden				-47.4	21	22.2
gate				-46.2	22.5	18.5
goat				-22.5	0	0
grapes				-32.5	15.9	17.8

Table 9: A table showing all of the VOT results gathered for Bareerah.

	VOT (m/s)					
	VOICELESS			VOICED		
	Sentence-Initial	Sentence-Medial	Sentence-End	Sentence-Initial	Sentence-Medial	Sentence-End
pen	9.2	9.4	0			
paper	4.5	9.1	7.2			
pyjamas	4.8	4.8	0			
picture	4.3	5.2	8			
biscuit				0	N/A	5.6
bag				0	0	0
banana				0	N/A	-61.7
bottle				N/A	0	0
table	12.7	13.6	19.5			
towel	7	9.3	12.4			
train	25.9	18.5	24.9			
tomato	10.2	19.4	17.6			
dog				-75.3	7.5	0
doctor				-61.5	5.6	-47.8
dress				-99.6	6.2	13.8
door				-91.9	N/A	-61.5
kitchen	21.5	23.9	10.7			
cat	16.9	28.5	12.9			
carpet	23.1	5.4	3.4			
key	24.4	37.1	22.7			
garden				0	6.8	N/A
gate				-90.2	6.9	N/A
goat				0	21.7	0
grapes				-81.7	0	N/A

Table 10: A table showing all of the VOT results gathered for Cameelah.

	VOT (m/s)					
	VOICELESS			VOICED		
	Sentence-Initial	Sentence-Medial	Sentence-End	Sentence-Initial	Sentence-Medial	Sentence-End
pen	5.2	12.4	18.1			
paper	11.1	6.9	18.4			
pyjamas	1.9	3.5	0			
picture	6	22.6	11.5			
biscuit				-162.6	0	0
bag				0	0	11.6
banana				11.8	7.8	3
bottle				0	11.3	3.9
table	15.1	13.7	21.8			
towel	12.5	12.4	8			
train	38.8	44.6	70			
tomato	16.6	11.8	28.5			
dog				2.7	10.3	N/A
doctor				3.9	6.6	18.8
dress				24.9	17.4	15.2
door				-40.2	-48.7	8.6
kitchen	7.9	41.2	38.1			
cat	9.3	26.8	36.8			
carpet	35.7	43.1	25.2			
key	0	60.8	45.7			
garden				11	22	34.9
gate				19.5	31.9	37.9
goat				-54.4	44.3	34.7
grapes				24.7	17.9	N/A

Table 11: A table showing all of the VOT results gathered for Dina.

	VOT (m/s)					
	VOICELESS			VOICED		
	Sentence-Initial	Sentence-Medial	Sentence-End	Sentence-Initial	Sentence-Medial	Sentence-End
pen	58.9	73.8	56.7			
paper	51.1	36.5	69.9			
pyjamas	36.9	36.7	40.5			
picture	26.7	27.6	28			
biscuit				6.1	2.7	0
bag				5.6	-101.8	0
banana				9.7	6	-80.6
bottle				0	1.5	0
table	52.5	55.8	58.1			
towel	57.3	64.8	57			
train	88.2	79.6	69.1			
tomato	79	30.8	42.7			
dog				9	7.7	16.2
doctor				6.7	10.9	12
dress				0	30.8	36
door				8.4	13.6	9.2
kitchen	47.7	34.7	38.2			
cat	65.9	61.1	50.4			
carpet	52.2	52.7	60.6			
key	66.4	77.6	77.5			
garden				16.2	14	19.8
gate				12	10.6	12.8
goat				-56.6	39.8	30.8
grapes				-56.8	4.3	7.6

Table 12: A table showing all of the VOT results gathered for Eliza.

	VOT (m/s)					
	VOICELESS			VOICED		
	Sentence-Initial	Sentence-Medial	Sentence-End	Sentence-Initial	Sentence-Medial	Sentence-End
pen	66.4	87.3	57.9			
paper	75	56.1	58.9			
pyjamas	69.2	59.1	44.7			
picture	64.7	78.6	51			
biscuit				-71.4	-59.9	10
bag				12.2	13.9	16.1
banana				-50.9	14.2	10.4
bottle				-52.3	19.5	15
table	50.1	54.9	63.2			
towel	53.9	64.7	54.8			
train	99.3	95.4	101			
tomato	57.5	45.7	42.7			
dog				-73.4	15	23.4
doctor				-39.1	16.8	17.3
dress				-65.4	24.6	36.9
door				14.9	22.2	15.8
kitchen	71.9	67	57.2			
cat	73.4	76.6	89.1			
carpet	77.1	80.8	79.3			
key	98.5	109.9	96.4			
garden				33.5	29.5	37.1
gate				30.2	34.6	23.1
goat				0	33.8	41.1
grapes				-38.11	36	41.2

Table 13: A table showing all of the VOT results gathered for Fatimah.

	VOT (m/s)					
	VOICELESS			VOICED		
	Sentence-Initial	Sentence-Medial	Sentence-End	Sentence-Initial	Sentence-Medial	Sentence-End
pen	36.4	39.9	47.5			
paper	31.1	41.5	30.8			
pyjamas	16.1	32.8	38.8			
picture	14	31.3	2.1			
biscuit				-51.6	15.9	8.3
bag				5	7.3	17.8
banana				-44	9.9	15.7
bottle				-42.7	N/A	0
table	44.6	56.4	36.7			
towel	53	53	50.6			
train	67.7	59.9	69			
tomato	30.7	48.1	40			
dog				-59.9	10.7	16.3
doctor				-35.4	9.5	10.4
dress				30.8	44	42.6
door				-55.4	13	14.2
kitchen	22.1	34	24.9			
cat	31.5	45.8	36.2			
carpet	28.5	61.5	44			
key	84.5	101.2	52.7			
garden				-38.1	26.4	16.2
gate				-68.5	23.8	14.6
goat				-57.9	13.3	21
grapes				53.1	N/A	20.5

Table 14: A table showing all of the VOT results gathered for Ghazala.

	VOT (m/s)					
	VOICELESS			VOICED		
	Sentence-Initial	Sentence-Medial	Sentence-End	Sentence-Initial	Sentence-Medial	Sentence-End
pen	44.5	26.7	39.2			
paper	26.9	29.4	31.6			
pyjamas	29.3	18.2	24.8			
picture	30.8	21.7	38.5			
biscuit				10.3	11.7	6.7
bag				8.9	10.2	10.9
banana				9.8	10.2	10.7
bottle				10.9	6.6	14.8
table	37.1	24.3	38.8			
towel	37.7	43.2	54			
train	53.4	84.8	78.9			
tomato	22.5	21.7	22.1			
dog				-43	11.4	15.5
doctor				12.3	11.7	11.4
dress				0	27.9	30.6
door				11.2	10	14.2
kitchen	43.2	37.5	54.4			
cat	54.8	62.8	62.4			
carpet	58.7	34.5	53.2			
key	60.8	86.6	81.9			
garden				22.7	20.5	26.9
gate				44.4	31.4	47.4
goat				-102.7	16	33.5
grapes				12.2	21.9	33.9

Table 15: A table showing all of the VOT results gathered for Haafizah.

	VOT (m/s)					
	VOICELESS			VOICED		
	Sentence-Initial	Sentence-Medial	Sentence-End	Sentence-Initial	Sentence-Medial	Sentence-End
pen	39.5	22	30.9			
paper	33.8	18.8	35.2			
pyjamas	24.8	25.2	28.7			
picture	43.8	16.3	40.7			
biscuit				-61.6	N/A	15
bag				-71.6	0	4.5
banana				-57.2	3.4	8.2
bottle				-68.3	8.5	0
table	40.7	49.4	41.6			
towel	61.4	29.3	78.8			
train	94.3	58.4	99.8			
tomato	43.5	38.2	38.6			
dog				11.9	14.7	13.4
doctor				-44.1	11.7	7.8
dress				-53.7	N/A	0
door				-64.7	10.9	13.8
kitchen	56.7	64.9	56.7			
cat	62.5	53	75.5			
carpet	63.6	49.6	68.2			
key	99.7	85.1	85.3			
garden				-39.3	16.9	20.6
gate				-74.6	27.1	27.7
goat				-87.7	31.3	20.6
grapes				-38.3	-35	52.2

Table 16: A table showing all of the VOT results gathered for Inayah.

	VOT (m/s)					
	VOICELESS			VOICED		
	Sentence-Initial	Sentence-Medial	Sentence-End	Sentence-Initial	Sentence-Medial	Sentence-End
pen	5	48.4	40.3			
paper	30.3	41.7	35.9			
pyjamas	50.7	34	46.8			
picture	30.7	39.9	40.2			
biscuit				10.8	N/A	10
bag				7	8.4	12.7
banana				5.2	8.6	0
bottle				-31.2	9.8	14.7
table	48.7	37.5	41.3			
towel	64.5	49.8	54.7			
train	68.7	51.3	75.9			
tomato	39.3	55	53.6			
dog				14.3	10.2	8.3
doctor				11.7	12	18.4
dress				42.2	36.8	55.5
door				7.9	11.9	16.5
kitchen	52.6	49.1	57.4			
cat	45.7	51.2	64.7			
carpet	58.6	50.2	60.5			
key	67.2	64.7	75.5			
garden				24.1	-31.8	N/A
gate				22.5	16.4	26.2
goat				23.8	27.8	9.7
grapes				-93.6	16.3	0

Table 17: A table showing all of the VOT results gathered for Jameelah.

	VOT (m/s)					
	VOICELESS			VOICED		
	Sentence-Initial	Sentence-Medial	Sentence-End	Sentence-Initial	Sentence-Medial	Sentence-End
pen	68.9	104.7	73.7			
paper	36.5	77.4	84			
pyjamas	26.2	43	36.7			
picture	42.5	33.3	40.1			
biscuit				7.1	16	14.4
bag				7.5	6.3	8.7
banana				9.7	9.3	3.9
bottle				-68.3	7.6	8.3
table	72.1	70.9	59.1			
towel	90.1	102.1	62			
train	95.4	104.1	66.7			
tomato	32.5	45.2	48.2			
dog				-63.4	10.5	15
doctor				-46.6	21.6	16.7
dress				-88.4	43.9	22.8
door				-57.1	17	16.1
kitchen	57.3	63.9	66.2			
cat	63.7	52.3	71.4			
carpet	92.2	101	85.7			
key	104	107.3	115.8			
garden				56.8	13.9	14.7
gate				-105.3	30.7	30.6
goat				-144	32.6	18.9
grapes				-73.6	36.1	26.9

Table 18: A table showing all of the VOT results gathered for Khadijah.

	VOT (m/s)					
	VOICELESS			VOICED		
	Sentence-Initial	Sentence-Medial	Sentence-End	Sentence-Initial	Sentence-Medial	Sentence-End
pen	56.9	88.7	67.5			
paper	63.7	60.9	73.8			
pyjamas	49.7	41.8	28.9			
picture	42.9	31.8	35.5			
biscuit				15	N/A	10.8
bag				0	0	1.1
banana				12.1	10.9	13.4
bottle				16.7	N/A	3.2
table	47.8	69	54.4			
towel	64.3	74.6	59.4			
train	108.3	92.6	97			
tomato	62.6	68.8	30.7			
dog				12.4	14.1	13.5
doctor				12.7	10.3	12.2
dress				46.2	4.5	51.2
door				13.4	16.3	10.9
kitchen	59.8	43.7	38.3			
cat	57.1	55.6	28.9			
carpet	40.2	50.9	48.4			
key	77.9	111.3	73.4			
garden				13.3	18.7	20.2
gate				17.8	19.9	30.1
goat				26.8	33.1	25.6
grapes				26.2	27.5	26

Table 19: A table showing all of the VOT results gathered for Laila.

8.3. Appendix 3: Information sheet – Fake Version

Participant Information Sheet

Research Project Title: The influence of Gujarati heritage on vision.

Name of Researcher: Safiyyah Fadia

Contact Details of Researcher: 07938815825

You are being invited to take part in a research project. Before you decide, it is important for you to understand why this research is being done and what it will involve. Please take time to read the following information and discuss it with others if you wish. Ask questions if there is anything that is not clear or if you would like more information. Take time to decide whether or not you want to take part in this study. May I take this opportunity to thank you for taking your time to read this.

1. What is the purpose of the project?
In this research project, which is for my Masters by Research degree, I am investigating if and how vision is affected by being bilingual in two languages. In this specific case, the two languages are English and Gujarati.

2. Why have I been chosen?
You have been chosen due to your bilingual ability to speak Gujarati and English, your gender as well as the fact that you fit into one of the three age brackets that I am investigating. You will be one of 12 participants that are part of this study.

3. Do I have to take part?
Participation in this study is entirely voluntary, so please do not feel obliged to take part. Refusal will involve no penalty whatsoever and you may withdraw from the study at any stage without providing me with an explanation.

4. What do I have to do?
You will be invited to take part in one recorded interview. I, the researcher, will present you with one card at a time. The first set of 24 cards will have a word written in English. The second set of 48 cards will have these same English words written within 2 English sentences per word.
You will be expected to read out all the cards that I show you. This should take no more than 20 minutes of your time, and you will feel neither pain nor discomfort during this process.

5. Are there any disadvantages to taking part?
There should be no foreseeable disadvantages to your participation. If you are unhappy or have further questions at any stage in the process, please address your concerns initially to myself if this is appropriate. Alternatively, please contact Professor M. Adkins **(m.adkins@hud.ac.uk)** at the School of Music, Humanities and Media, University of Huddersfield.

6. Will all my details be kept confidential?
All information which is collected will be strictly confidential and anonymised before the data is presented in any work, in compliance with the Data Protection Act and ethical research guidelines and principles. This will be achieved by removing your identity from the data; storing the recordings in a

private, locked folder, as well as limiting who can listen to the recordings (only myself and my tutor will have access). After the thesis is over, all of the data will be destroyed.

7. What will happen to the results of the research study?
The results of this research will be written up as part of my Master's degree. There may be a possibility that the thesis will be published; if this is the case, your identity will still remain anonymous in the publication. If you would like a copy please contact me on the phone number provided at the top of the sheet.

8. What happens to the data collected?
After the data is recorded, it will be analysed using an online software called Praat.

9. Will I be paid for participating in the research?
No payment will be given to you for your time.

10. Where will the research be conducted?
The research will be conducted in my home or, if you prefer, in your own home. The location is mainly based on where you will feel at ease.

11. Who has reviewed and approved the study, and who can be contacted for further information?
For further details, please contact myself; my contact details are written at the top of this document. You can also contact my supervisor Dr Erica Gold. Her contact details are E.Gold@hud.ac.uk.

8.4. Appendix 4: Information sheet – Real Version

Participant Information Sheet

Research Project Title: The influence of Gujarati on the VOT of English stops.

Name of Researcher: Safiyyah Fadia

Contact Details of Researcher: 07938815825

You are being invited to take part in a research project. Before you decide, it is important for you to understand why this research is being done and what it will involve. Please take time to read the following information and discuss it with others if you wish. Ask questions if there is anything that is not clear or if you would like more information. Take time to decide whether or not you want to take part in this study. May I take this opportunity to thank you for taking your time to read this.

1. What is the purpose of the project?
In this research project, which is for my Masters by Research degree, I am investigating if and how English and Gujarati can influence each other's voice onset time.

2. Why have I been chosen?
You have been chosen due to your bilingual ability to speak Gujarati and English, your gender as well as the fact that you fit into one of the three age brackets that I am investigating. You will be one of 12 participants that are part of this study.

3. Do I have to take part?
Participation in this study is entirely voluntary, so please do not feel obliged to take part. Refusal will involve no penalty whatsoever and you may withdraw from the study at any stage without providing me with an explanation.

4. What do I have to do?
You will be invited to take part in one recorded interview. I, the researcher, will present you with one card at a time. The first set of 24 cards will have a word written in English. The second set of 48 cards will have these same English words written within 2 English sentences per word.
You will be expected to read out all the cards that I show you. This should take no more than 20 minutes of your time, and you will feel neither pain nor discomfort during this process.

5. Are there any disadvantages to taking part?
There should be no foreseeable disadvantages to your participation. If you are unhappy or have further questions at any stage in the process, please address your concerns initially to myself if this is appropriate. Alternatively, please contact Professor M. Adkins **(m.adkins@hud.ac.uk)** at the School of Music, Humanities and Media, University of Huddersfield.

6. Will all my details be kept confidential?
All information which is collected will be strictly confidential and anonymised before the data is presented in any work, in compliance with the Data Protection Act and ethical research guidelines and principles. This will be achieved by removing your identity from the data; storing the recordings in a

private, locked folder, as well as limiting who can listen to the recordings (only myself and my tutor will have access). After the thesis is over, all of the data will be destroyed.

7. What will happen to the results of the research study?
The results of this research will be written up as part of my Master's degree. There may be a possibility that the thesis will be published; if this is the case, your identity will still remain anonymous in the publication. If you would like a copy please contact me on the phone number provided at the top of the sheet.

8. What happens to the data collected?
After the data is recorded, it will be analysed using an online software called Praat.

9. Will I be paid for participating in the research?
No payment will be given to you for your time.

10. Where will the research be conducted?
The research will be conducted in my home or, if you prefer, in your own home. The location is mainly based on where you will feel at ease.

11. Who has reviewed and approved the study, and who can be contacted for further information?
For further details, please contact myself; my contact details are written at the top of this document. You can also contact my supervisor Dr Erica Gold. Her contact details are E.Gold@hud.ac.uk.

8.5. Consent Sheet

Participant Consent Form (E4)

Title of Research Study: The influence of VOT within English and Gujarati speech

Name of Researcher: Safiyyah Fadia

Participant Identifier Number:

☐ I confirm that I have read and understood the participant Information sheet related to this research, and have had the opportunity to ask questions.

☐ I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason.

☐ I understand that all my responses will be anonymised.

☐ I give permission for members of the research team to have access to my anonymised responses.

☐ I agree to take part in the above study

Name of Participant:

Signature of Participant:

Date:

Name of Researcher: Safiyyah Fadia

Signature of Researcher: 

Date: