Julia Pakai

Examining Voice Onset Time in Hungarian-English Bilinguals

Applications in Forensic Speech Science

Submitted in partial fulfilment of the degree of MA in Forensic Phonetics at the Department of Language and Linguistic Science, University of York

September 2022

Supervisor: Professor Paul Foulkes

Word count: 13,636

Table of Contents 1

Acknowledgements

I would like to thank my supervisor, Prof Paul Foulkes, for all his help and support.

Furthermore, I would like to thank Dr Eleanor Chodroff for helping me with the technical difficulties that occurred during forced alignment, and Dr Vincent Hughes for helping me with the calculation of likelihood ratios.

I would also like to thank my family for supporting me through this degree, and especially my mum, for relentlessly looking for Hungarians who I could interview.

Of course, I would also like to thank all participants for taking the time to be interviewed.

Finally, I would like to thank York JSoc, without whom I would not have been able to get through this year.

Abstract

Purpose: This dissertation examines the role of Voice Onset Time in Hungarian-English bilinguals, with special attention paid to applications in forensic phonetics. The research was conducted in light of an extensive literature review. It is important, as no such studies have been carried out on Hungarian-English bilinguals so far; and therefore, the role of bilinguals in forensic phonetics remains largely unexamined. Furthermore, due to the spreading of English as a *lingua franca*, and foreign accent, including disguise, becoming a more relevant topic in the field of forensic phonetics, this dissertation aimed to fill this gap in the research.

Hypotheses: The dissertation had three hypotheses: 1) Bilingual English VOT is different to monolingual English VOT. 2) People who have spent time in an English-speaking country will perform better than those who have not. 3) Bilingual data leads to better system performance in a likelihood ratio framework.

Method: An empirical, quantitative, study which used corpora and interviews for gathering data was conducted. Monolingual corpora were obtained from other researchers and institutions, while bilingual data was collected by the researcher. In the case of bilingual data, $n = 4{,}876$ observations were made and analysed. Statistical analysis was carried out in R. Likelihood ratio system performance was tested in MATLAB.

Results: The analysis of the first two hypotheses showed significant results only for selected phonemes. Therefore, no robust results were obtained. The third hypothesis was falsified, as the likelihood ratio system, when comparing bilinguals and using monolinguals as a reference population, performed worse than in the case of the monolingual data being examined on its own. Nevertheless, this still yielded some important conclusions regarding forensic phonetics, i.e. that any conclusions in speaker profiling or comparison should be made with caution.

Conclusions: It was concluded that, while no significant results could be obtained concerning the first two hypotheses, the research should be repeated, preferably as a longitudinal study. Furthermore, more research is needed on the likelihood ratio framework to examine how it can be improved to accommodate bilinguals.

Keywords: *Voice onset time, bilingualism, forensic phonetics, likelihood ratio*

List of Abbreviations

#CV	Word-initial, pre-vowel
AmE	American English
AOA	Age of Onset of Acquisition
ASR	Automatic Speaker Recognition
BEA	Beszélt Nyelvi Adatbázis
C_{llr}	Log-likelihood ratio cost
CP	Critical Period
CPH	Critical Period Hypothesis
ECH	Educated Colloquial Hungarian
EER	Equal Error Rate
EU	European Union
FSS	Forensic Speech Science
G2P	Grapheme-to-Phoneme
HS	Heritage Speakers
L1	First language
L2	Second language
L3	Third language
LADO	Language Analysis for the Determination of Origin
LR	Likelihood Ratio
MFA	Montreal Forced Aligner
MLAT	Modern Language Aptitude Test
MTA	Magyar Tudományos Akadémia (Hungarian Academy of Sciences)
MVKD	Multivariate Kernel Density
SSBE	Standard Southern British English
VOT	Voice Onset Time

List of Figures

List of Tables

1. Introduction

Egy magyar hazajön Amerikából és megkérdez valakit:
A Hungarian comes home from America and asks someone:

Hogyan kell írni?
How do you spell it?

Törölköző vagy türülköző?[1]
Towel or towel?

Eddig mennyi van leírva?
How much do you have written down so far?

TH

Versions of the above joke are common among Hungarians and were told by a lot of comedians, probably most famously by Géza Hofi (Magyar vagyok, 2010). It alludes to the fact that when Hungarians return from spending time in an English-speaking country, instead of saying /t/, they say /t^h/, aspirating their voiceless plosives. This demonstrates that Hungarians are aware of voicing distinctions between Hungarian and English, and that plosives in their first language (L1) are affected by time spent in an English-speaking environment.

Bilingualism is on the rise, especially with more people speaking English as their second language (L2) due to English being the *lingua franca* (Crystal, 2003). Therefore, bilingualism is also becoming more relevant in the field of forensic phonetics, where the presence or absence of bilingualism can necessitate further methodological considerations, or even be a distinguishing feature of a person in speaker comparison or profiling. As some people achieve nativelike pronunciation, this becomes more difficult. Hence, this study investigates the potential of Voice Onset Time (VOT) to be used as a speaker-discriminant in speaker comparison cases, or as a cue to bilingualism in speaker profiling. Further, the study can also be of use in the field of bilingualism studies, potentially uncovering valuable insights into VOT variation.

Abramson & Lisker define VOT as "the interval between the release of a stop and the onset of phonation" (1973: 1). It is used as a phonological voicing cue in many languages (Foulkes et al., 2010a; Gósy, 2001; Ladefoged, 2003, Lisker & Abramson, 1967). Since VOT in Hungarian and English are very different, examining it in the context of bilingualism and forensics is especially useful, because VOT in Hungarian-English bilinguals has received little

[1] Both words mean towel but are spelled and pronounced slightly differently due to regional variation in Hungary. Both are correct pronunciations, but the correct spelling is the former.

attention so far. The study examines if VOT is more variable in the L2 of bilinguals, and what factors play a role in potentially acquiring nativelike VOT. It also goes into potential insights of L2 to L1 influence. Furthermore, it provides insights to the field of forensic phonetics regarding VOT in bilinguals.

The dissertation is organised as follows: The first part deals with the background, where the underlying topics of the study are laid out, including bilingualism, language acquisition, VOT, and forensic linguistics. Special attention is paid to Forensic Speech Science (FSS) and the Likelihood Ratio (LR) framework. Next, the present study, the research questions, and the hypotheses are laid out. After that, the methodology of the research regarding design, data collection, and analysis is outlined. The section thereafter is concerned with the results of the study which are discussed in light of current literature on the topic afterwards. In the remaining part, proposals for further research are made, and finally, a conclusion is drawn.

2. Background and Literature Review

First, a general overview of bilingualism and second language acquisition, including relevant influencing factors, is presented. Second, VOT is explained in more detail, including language-specific information. Third, the field of forensic linguistics is presented, most importantly, FSS and the LR framework. Overall, research relevant to the hypotheses posed is reviewed.

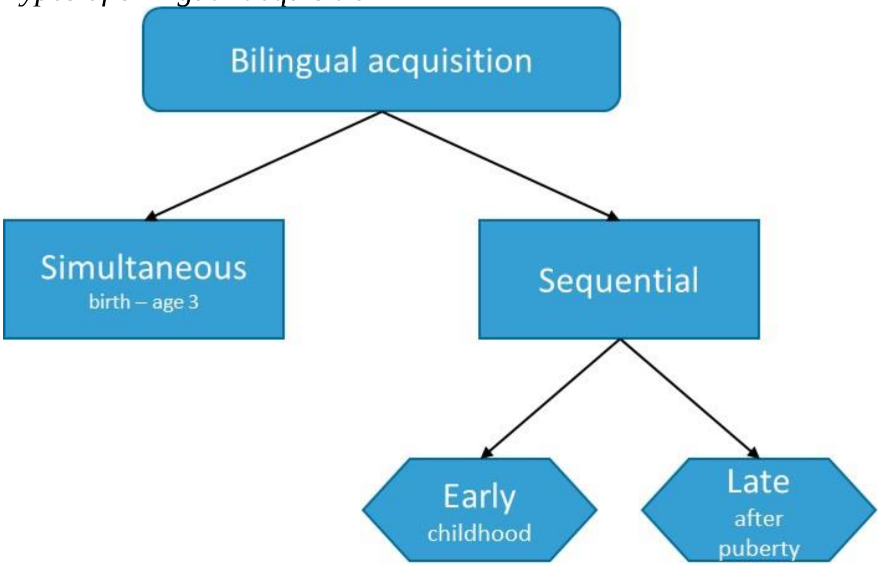
2.1 Bilingualism and Second Language Acquisition

The Oxford English Dictionary Online defines bilingualism as the "[a]bility to speak two languages; the habitual use of two languages colloquially" (OED Online, n.d.). Already, the dictionary definition is not precise, as it leaves room for multiple interpretations. On the two extremes this means that everyone is bilingual as everyone knows a couple of words in other languages or that only people who are equally highly proficient in all their languages are bilingual (Edwards, 2006; Romaine, 1989). Grosjean defines a bilingual as someone who interacts with the world in two languages, i.e. uses two languages in their everyday lives (1982; 1998; also Baker & Wright, 2021; Montrul, 2008; Wei, 2008; Yavas & Byers, 2015). The definition of bilingualism can also include someone who speaks more than two languages (Baker & Jones, 1988; Grosjean, 1982), which is the definition this paper uses. Discussing the theoretical foundations of bilingualism is important, as it is one of the many methodological considerations of the dissertation.

Types of bilingualism

Acquisition of two languages can happen either simultaneously, or sequentially (Baker & Wright, 2021; Jank & Kovács, 2020). If the two languages are acquired simultaneously, it is referred to as having two L1s rather than an L1 and an L2 (Benmamoun, 2013; Meisel, 2006; Montrul, 2008). Simultaneous acquisition is often cut-off at the age of three (Grosjean, 1982; Lakshmanan, 2009; Paradis, 2007; Romaine, 1989), while some place it as early as the first month after birth (Deuchar & Quay, 2000). Sequential acquisition takes place, if the foundations for the L1 are laid before the L2 is starting to be acquired (Taeschner, 1983). Additionally, sequential acquisition is separated into early and late sequential acquisition, with early acquisition occurring until the onset of puberty and late sequential acquisition after that (De Houwer, 1990; Lakshmanan, 2009: Yeni-Komshian et al., 2000), though some place it as early as the age of eight (Taeschner, 1983). Figure 1 shows types of bilingual acquisition.

Figure 1
Types of bilingual acquisition



Source: Adapted from Montrul (2008: 18)

Sequential bilinguals seem to be more likely to have an accent than simultaneous bilinguals (de Leeuw et al., 2010; Thordardottir, 2017), and early sequential bilinguals seem to be less likely to have an accent than late sequential bilinguals (Flege, 1999; Stölten et al., 2014; Unsworth, 2013; Wrembel, 2015). While early L2 learners may seem nativelike, they might still exhibit differences in the acoustic domain compared to native speakers (Lakshmanan, 2009). At the same time, acquiring a language simultaneously does not guarantee that the pronunciation will be nativelike as there are several other factors that play a role in the development of one's phonology (Hopp & Schmid, 2013).

A bilingual can either be balanced, i.e., speak both languages equally well, or non-balanced (Grosjean, 1982; Jank & Kovács, 2020); some say that balanced bilinguals process information differently than non-balanced ones (Albert & Obler, 1978). Non-balanced bilinguals, who are usually more nativelike in their stronger language, are more common than balanced ones (Montrul, 2008), as most bilinguals do not use their languages for the same tasks (Grosjean & Soares, 1986). It is important to mention that balanced bilinguals are not necessarily at equally highly proficient levels in their languages but may just be equally average in them (Baker, 1988), and they are also quite rare (Grosjean, 1998; Montrul, 2008). A bilingual may have different levels of proficiency in different language skills, e.g. someone may only be a receptive bilingual in one language (Bialystok, 2001). Table 1 shows a breakdown of the language skills.

Table 1
Language skills

	Oracy	**Literacy**
Receptive skills	Listening	Reading
Productive Skills	Speaking	Writing

Source: Adapted from Baker & Wright (2021: 7)

Language skills are often evaluated by using tests such as IELTS or TOEFL, and quantified by use of a scale like the Common European Framework of Reference for Languages (Baker & Wright, 2021; Bialystok, 2001; Council of Europe, n.d.). However, these are not enough, or are inappropriate to use in the case of bilinguals (Grosjean, 2008). Instead, language use surveys should be used (Baker & Wright, 2021). Often, fluency is used as an indicator of proficiency, although this is problematic as fluency may vary by the person's style, personality or even physical or mental health; therefore, fluency should only be used as an indicator for proficiency with L1 fluency as a baseline (De Jong et al., 2013; Gut, 2009).

The differences between oracy and literacy (Baker, 1988; Bialystok, 2001), and reception and production (Baker & Jones, 1998; Baker & Wright, 2021; Romaine, 1989) also have to be considered when talking about bilingualism, as the bilingual may not have the same level of proficiency in all of them (Romaine, 1989). This study is only concerned with the oracy-production interface.

Additionally, it is important to address the issue of Heritage Speakers (HS). They are non-balanced bilinguals (Scontras et al., 2015), who Benmamoun et al. define as "bilingual speakers of an ethnic or immigrant minority language, whose first language often does not reach native-like attainment in adulthood" (2013: 129). Nevertheless, often, the pronunciation of HS is still nativelike, compared to other parts of their grammar, e.g. syntax or morphology, where they seem more like an L2 learner (Benmamoun et al., 2013; Haman et al., 2017; Isurin & Ivanova-Sullivan, 2008; Montrul, 2013). They are usually naturalistic, not instructed, learners (Montrul, 2010). It is unclear whether they have only incompletely acquired their parents' L1 or if they undergo attrition as they experience a language shift (Kupisch & Rothman, 2016; Potowski, 2013; Rothman, 2009; Schmid, 2011; Scontras et al., 2015), i.e. their dominant language is replaced by the dominant L2, in which they are nativelike (Benmamoun et al., 2013; Bolonyai, 1998; Kupisch et al., 2014). Their overall proficiency usually lies somewhere between L1 and L2 speakers (Kupisch et al., 2014). Nonetheless, some HS are only receptive bilinguals (Benmamoun et al., 2013).

Problems in bilingualism

A whole community can be bilingual due to border shifts, colonisation, or mass-migration, or an individual may be bilingual due to their circumstances (Baker & Wright, 2021; Bartha, 1996; Edwards, 2006; Nemes, 2022; Ürmösné Simon, 2014). Child bilinguals are usually the children of first-generation immigrants (Montrul, 2008). As of 1996, it was estimated that more than half of the world's population is bi- or multilingual (Bartha, 1996), and globally, bilingualism is very widespread and even the norm, mostly due to the spreading of English (Lleó, 2016).

Since bilingualism usually occurs out of necessity and is oftentimes associated with lower-class minorities, and is therefore viewed negatively (Baetens Beardsmore, 1986; Romaine, 1989; Ürmösné Simon, 2014). Lower-income families from these backgrounds are also more likely to have deficits in their language skills (Qi et al., 2006). Some even see bilingualism as a threat to "language purity" (Romaine, 1989: 285). Yet, parents still think that bilingualism will affect their children negatively, either socially or cognitively (Baetens Beardsmore, 1986; Baker, 1988; Baker & Wright, 2021; Wei, 2008), despite there being no evidence of any negative cognitive effects of bilingualism (Ürmösné Simon, 2014), and some even argue that bilingualism influences cognition positively (Bialystok, 2009). However, with globalisation and the spreading of bilingualism (Baetens Beardsmore, 1986; Bhatia & Ritchie, 2013; Jank & Kovács, 2020), the topic is more relevant than ever.

In the case of Hungary, the joining of Hungary to the European Union (EU) in 2004 resulted in more need for bilingualism in the country, but even today, the least number of people speak a second language in Hungary compared to other EU countries (Ürmösné Simon, 2014). Specifically, Bartha showed that the Hungarian of Hungarians in America is generally perceived as bad in Hungary (2002).

Language modes

Bilinguals also have different language modes, i.e. what language is activated in their brain, depending on who they are speaking to (Grosjean,1998; 2001; Grosjean & Soares, 1986). This can especially vary when talking to bilinguals as there are more possibilities for expression given their shared languages (Grosjean, 2001). The language mode can lead to code-switching and -mixing, which is only appropriate to be used in certain settings (Bhatia & Ritchie, 1996; Grosjean, 2001; 2008; Grosjean & Miller, 1994; Muysken, 2000; 2006). This suggests that the activation of the L1 at the same time as the L2 may affect L2 pronunciation (Flege et al., 1997; Goldrick et al., 2014).

Bilingualism, language learning, and nativelikeness

Usually, the goal of language learning is the mastery of the target language (Bhela, 1999), but some people do not care about nativelike acquisition and are only concerned with the functionality of the L2 (Neufeld, 1977). Oftentimes, the mastery of pronunciation is seen as the mastery of the L2 (Earle & Arthur, 2017).

Generally, successful ultimate attainment is rare, especially in adult learners (Abrahamsson & Hyltenstam, 2008; Birdsong, 1999); nevertheless, it is still possible (Bongaerts, 1999; Lahmann et al., 2017; Moyer, 1999). Neufeld suggests that exceptional learners may be exceptional in not just language learning but may be a distinct population with non-normal neuro-function (2001). It is relevant that even child learners might not reach ultimate attainment, depending on various factors outlined later (Abrahamsson & Hyltenstam, 2009). Furthermore, pronunciation is the most difficult area to appear nativelike in (Montrul, 2008).

Bilinguals can have an accent or be near-native or nativelike, which is equivalent to non-nativeness that is not perceived by native speakers of the language (Abrahamsson & Hyltenstam, 2008). However, nativelikeness might vary in different contexts (Gnevsheva, 2017). Furthermore, regarding global accent judgements, prosody seems to have the strongest effect in favour of rating someone as non-native (Anderson-Hsieh et al., 1992).

An accent usually occurs when L1 targets influence L2 targets (Archibald, 2009; Bergmann et al., 2016; Costa et al., 2017; Foulkes, 2006; Grosjean, 2010; Gut, 2009). Archibald states that early acquisition seems to counteract this, and L1 targets do not influence L2 targets as much (2009). Additionally, adults that overheard the target language as a child seem to exhibit less accent than ones that did not (Au et al., 2002). The fact that there is no such thing as perfect production, as speech is dynamic, must be taken into account as well when considering accent (Tabain, 2013). Continued use of L1 can cause a stronger accent according to Gut (2009), some even say that it is impossible to reach nativelike pronunciation if L1 is kept at a high proficiency (Thompson, 1991). Taeschner shows that someone may still be a bilingual, even if they have an accent or don’t have nativelike competence in other areas (1983).

Adank et al. posit that longer exposure to an accent leads to imitation (2010). This would mean that L2 speakers spending more time with native speakers might have a better accent than those that do not, as accommodation is based on exposure (Babel, 2012). This is especially true in L1 acquisition, as L1 acquisition is essentially just imitation (Babel et al., 2014). This can lead to phonetic convergence the more time people spend with each other (Pardo, 2006; Pardo et al., 2010; 2012).

The language of bilinguals tends to be more variable in their L2 than that of native speakers, leading to increased within-category within-speaker variation (Hambly et al., 2013; Vaughn et al, 2018; Xie & Jaeger, 2020). This variability is related to the linguistic features of the L1 and L2 (Chodroff & Baese-Berk, 2019; Vaughn et al., 2018). Furthermore, any variation may be as a function of slower L2 speaking rate (Bradlow et al., 2017).

Nevertheless, English proficiency is important in finding a job and generally being accepted into the community (Guven & Islam, 2015). Furthermore, accent may lead to sociolinguistic bias and linguistic profiling (Baugh, 2007).

Language acquisition

Second language acquisition does not necessarily mean the acquisition of a foreign language, although it can be the case (Ellis, 1994; 1997). Related languages seem to be easier to learn, while languages that are more dissimilar could result in less success in learning; this is especially true for the phonetics of a language (Antoniou et al., 2014; Aoyama et al., 2004). Although some say that this is less significant in early learners (Baker & Trofimovich, 2005).

The literature highlights two theories regarding the acquisition of L2 sounds. The first is the deficit theory, which states that if a sound is not found in the L1, the learner will not be able to apply it in the L2 (Archibald, 2009), most likely due to the L2 sound not being able to be perceived (Bailey & Haggard, 1973; Flege, 1999) or due to some distinctions between sounds that exist in the L2 not existing in the L1, they are mapped onto the same category, and therefore the learner might perceptually assimilate them (Aoyama et al., 2004; Bailey & Haggard, 1973; Brown & Copple, 2018). This is also supported by the interaction hypothesis, which states that the older a person is, the more difficult it is for them to create new phonetic categories and contrasts due to not being able to perceive these anymore (Baker et al., 2008; Bright & McGregor, 1970). The second is the equivalence theory, which proposes that due to similarity in sounds, the L2 learner merges the categories and is therefore unable to produce them accurately, while phonemes that do not exist in the native language can be produced accurately as the learner has no alternatives to replace them with (Flege, 1987; 1992; Flege et al., 1996).

Age of Onset of Acquisition and the Critical Period Hypothesis

Another factor in language acquisition is the age of onset of acquisition (AOA). There is significant negative correlation between AOA and ultimate attainment of the L2 (Abrahamsson & Hyltenstam, 2008; Flege & Bohn, 2021; Scovel, 1969), including pronunciation (Flege et al., 1996). Generally, this means that children are better and more efficient at acquiring an L2 than adult learners (Abrahamsson & Hyltenstam, 2008). Because of this, the Critical Period Hypothesis (CPH) was proposed and later popularised by Lenneberg (1967). It states that there is a Critical Period (CP) in which it is ideal to acquire a language, and outside of the CP it is not possible to reach ultimate attainment in an L2 due to neurological changes and the decrease of brain plasticity (Abrahamsson & Hyltenstam, 2008; Asher & García, 1969; Birdsong, 1999; Montrul, 2008). Although it is debated if the CPH is indeed the correct theory concerning language acquisition. Some posit that CPH is only valid for L1 acquisition (DeKeyser, 2000), others say it is only relevant in implicit language acquisition (Abrahamsson & Hyltenstam, 2008; DeKeyser, 2000), further, some argue that it is not even the primary predictor in language learning success and that language learning mechanisms have to be activated before the end of the CP, but it does not close a pre-defined window (Bialystok, 1997; Ioup et al., 1994), and others say that there are advantageous periods rather than a CP (Baker & Wright, 2021). Some argue that there are different CPs for different areas of language acquisition (Benmamoun et al., 2013; Gut, 2009), usually with the CP for syntax and semantics being the longest, ending at the mid-teens (Ruben, 1997), and the one for phonetics and phonology being the shortest, as phonetic perception, which is relevant to production, starts to decline at as early as six months of age (Bialystok, 2001). Bongaerts et al. even say that only pronunciation is affected by the CPH (1997).

Exposure

Frequency of exposure and input in the target language also plays a role in language acquisition (Armon-Lotem & Ohana, 2017; Carroll, 2017; Jarvis & Pavlenko, 2008), as well as in language maintenance (Benmamoun et al., 2013). Furthermore, the source of the input also needs to be considered, as learners learning from a non-native speaker might set wrong targets (de Leeuw et al., 2010; Flege & Bohn, 2021; Flege et al., 1996; Place & Hoff, 2015). Parents can generally employ a plethora of different ways of creating a bilingual environment, of which the most known, and probably most effective, is the one person-one language strategy (Appel & Muysken, 1987; De Houwer, 2007; Flege & Bohn, 2021; Nemes, 2022). This is favourable as the child will be able to approximate native targets as they are only receiving input from native speakers. In conjunction with this, it is not just the exposure, but also the time spent practicing by the learner that impacts acquisition (Haman et al., 2017). This means that the skills of the bilingual vary as a function of exposure (Hoff & Core, 2013). In addition, the different types of exposure (parents, classroom learning, community, etc.) may also influence the language acquisition of the bilingual (Romaine, 1989; Gut, 2009; Mayr & Montanari, 2015).

Another determinant that may play a role in acquiring a language is living in a country where the target language is spoken. Age of arrival in the target country (Flege et al., 1999) and time spent there (DeKeyser, 2000; Jarvis & Pavlenko, 2008) have been examined in this context. DeKeyser found that length of residence did not play a role past the first ten years (2000), while Flege found that global foreign accent is not affected past the first year of in-country residence (1988). Flege et al. state that foreign accent increases with later time of arrival (1999). However, Johnson & Newport say that age of arrival only plays a role until puberty (1989; also Kupisch et al., 2014).

Language Aptitude

The most prominently cited factor in language acquisition is that of language aptitude. Abrahamsson & Hyltenstam note that “[a]dult L2 learners require an above-average degree of language aptitude in order to attain near-native proficiency, whereas this is not the case for child learners” (2008: 488). Language aptitude, and to an extent motivation, are the best predictors of L2 learning success in adults, which may potentially be due to how aptitude influences learning strategies (Dörnyei & Skehan, 2003; Ellis, 1997). At the same time, this does not mean that child learners have increased language aptitude because they acquired an L2 early (Harley & Hart, 1997). Moreover, the child learner's initial language aptitude does not influence acquisition (Abrahamsson & Hyltenstam, 2008). Bylund et al. suggest that aptitude is also a good predictor in the maintenance of the L1 (2010). Language aptitude is usually tested using the Modern Language Aptitude Test (MLAT) (Novoa et al., 1988; Robinson, 2013).

Motivation and Attitude

In adult language learning, motivation also may play a role (Baker, 1988; Ellis, 1994; 1997; Flege et al., 1996; Moyer, 2007). Stoehr et al. say that wanting to avoid bias may be an especially good motivator for reaching nativelike pronunciation (2017).

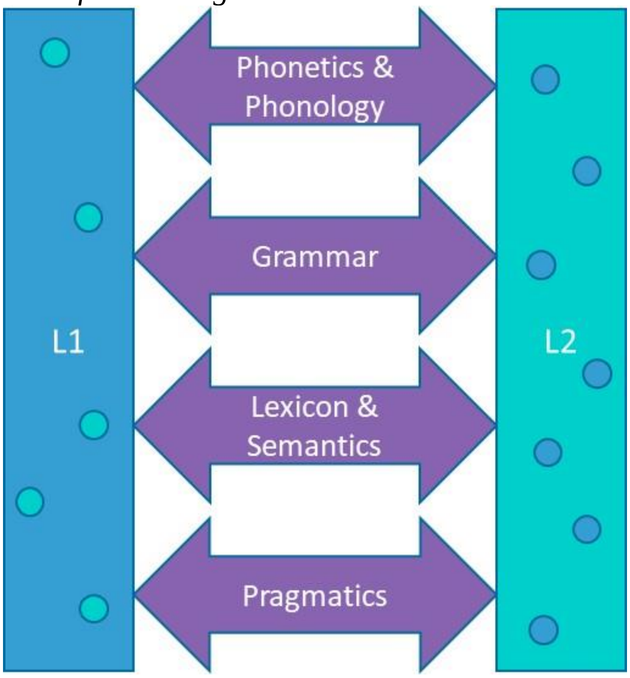
The attitude towards the target language, be that of the learner or their parents, may also play a role in becoming bilingual (Baker, 1988; Grosjean, 1982). Krashen states that motivation and language attitude are a better predictor than aptitude as they are more relevant in subconscious acquisition, which is the way a language should be learned according to the author (1981). For example, Hungarian is often seen as a less relevant language by Hungarian parents living in the West and they therefore refuse to teach it to their children (Hollstein, n.d.), taking away from them opportunities and part of their identity (Hoff & Core, 2013). Due to this, conscious efforts have to be made in Hungarian communities abroad to encourage parents to teach their children Hungarian (Hollstein, n.d.).

Language Transfer

Concerning neurological aspects, it is theorised that a bilingual either has one system (Bergmann et al., 2016), or two systems that interact with each other (Deuchar & Quay, 2000; Genesee, 1989; Holm & Dodd, 1999), this being the currently accepted theory (Hambly et al., 2013), or one system which later splits into two (De Houwer, 1990; Ingram, 1981; Genesee & Nicoladis, 2007). This suggests that a bilingual should be viewed wholistically; they are not two monolinguals in one (Cutler et al., 1989; Grosjean, 1985; 1989; 1998; 2008) or that being bilingual is the opposite of being monolingual (Lleó, 2016). As the intricacies of neurological processes are not relevant to the study, they are not discussed. Nevertheless, it is important to consider if and in how far a bilingual's systems interact with each other as it can have consequences for phonetic output. Mostly, the fact that a bilingual has two systems that interact with each other bidirectionally, i.e. L1 can influence L2, but also vice versa, is accepted (Amengual, 2021). It is debated though in how far the two systems influence each other, especially in different kinds of bilinguals. For example, Antoniou et al. posit that even L2-dominant bilinguals' L1 influences them (2011).

Language transfer also contributes to how bilinguals process language (Core & Scarpelli, 2015). It is the currently accepted view that bilinguals have two language systems that influence each other (Bergmann et al., 2016). Some say that only early bilinguals have two systems that influence each other, while in late bilinguals only the L1 influences the L2 (Baker & Trofimovich, 2005). At the same time, some argue that the systems always influence each other (Balukas & Koops, 2015; Wrembel, 2015), while others suggest that this is only the case in code-switching/-mixing (Balukas & Koops, 2015). Chang states, that even if the exposure to an L2 is brief, this can still influence the L1 (2012); if it affects pronunciation, this is called phonetic drift (Chang, 2019). Figure 2 depicts the currently accepted view of bilingual transfer processes.

Figure 2
Transfer in bilingualism



Source: Adapted from Schmid & Köpke (2007: 5)

There are two kinds of transfer in second language learning: positive and negative (Gut, 2009). Transfer takes place due to predicting L2 patterns based on the L1 (James, 1986). Positive transfer means that structures from the L1 are transferred to the L2 correctly and then used, this is especially common when the languages are similar (Corder, 1992; Ellis, 1997; Gabriel & Rusca-Ruths, 2015; Zen, 2020). Bilinguals are also generally better at learning L2 or third language (L3) phonetic contrasts and reaching targets due to positive transfer processes

(Antoniou et al., 2014; Dittmers et al., 2018; Spinu et al., 2018). Negative transfer means that structures of the L1 are transferred to the L2 incorrectly, e.g. if the languages are very dissimilar, or there is a lack of perceptual discrimination (Ellis, 1994; 1997; Flege, 1992; Jarvis & Pavlenko, 2008). Montrul suggests that negative transfer is especially problematic if these errors are not corrected and are allowed to fossilise (2008). Any further language, e.g. L3, can also influence the L1 and the L2 and vice versa, although not always the same way or the same amount, and usually a high level of L2 is required for the L2 to transfer to the L3 (Sanchez, 2015; Sanz et al., 2015; Wrembel, 2014; 2015).

Language Attrition

A further factor to consider is L1 attrition, which states that the L1 of a person can get worse over their lifetime, especially if they are bilingual (Baker & Wright, 2021; Bolonyai, 1998; Montrul, 2008). The ongoing contact with L1 speakers and being active in the community may influence attrition, as exposure can help maintain the L1 (de Leeuw, 2008; Hatoss, 2004; Schmid, 2007; 2013). Additionally, language attrition does not only take place if the L2 is proficient but also after a small amount of contact with the L2 due to L2 to L1 transfer processes (Schmid, 2012); although it is much more likely for attrition to happen when the L2 becomes the dominant language, usually as a function of length of residence in a foreign country (Schmid, 2019; Schmid & Köpke, 2007; Schmid & Yilmaz, 2018).

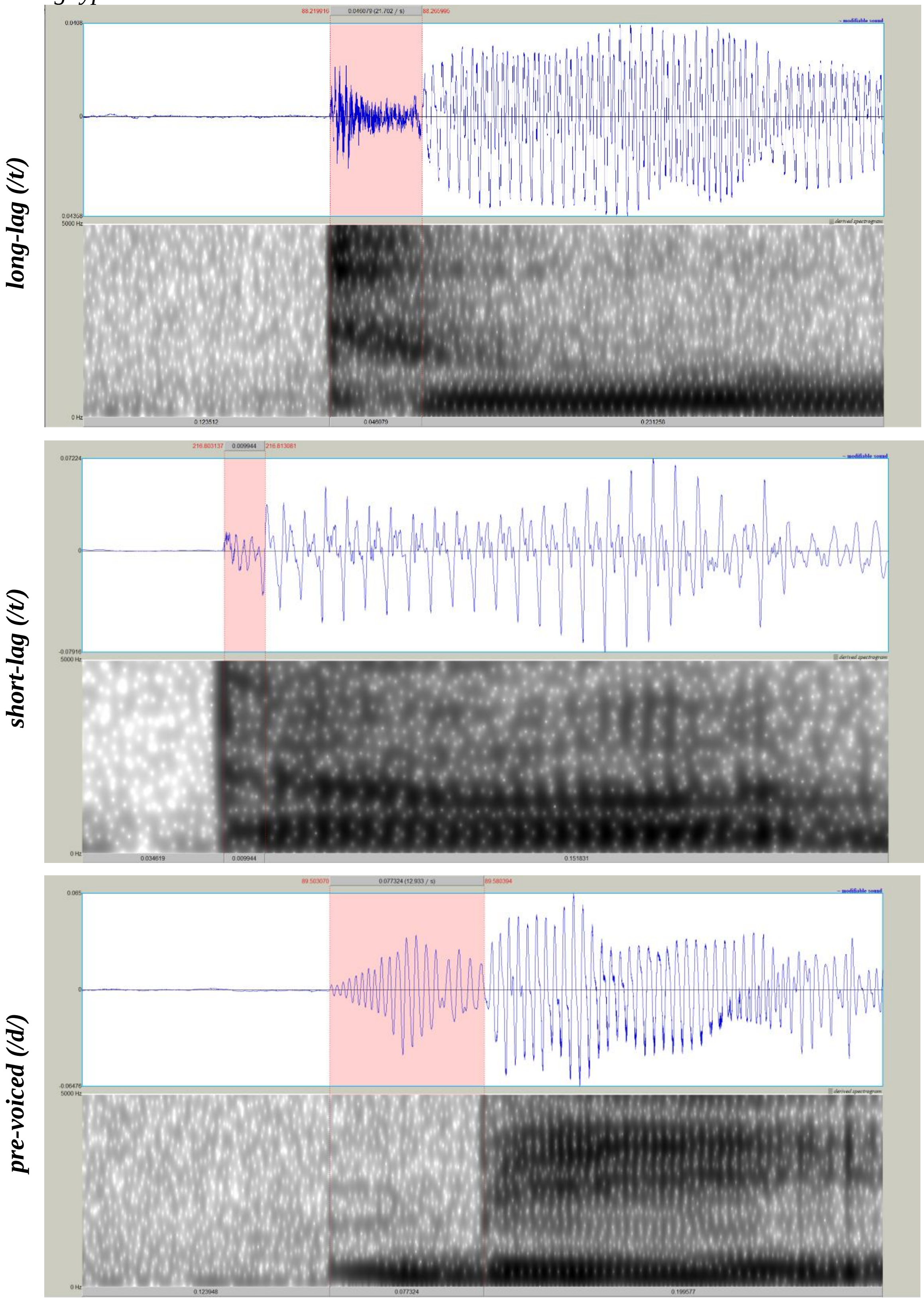
2.2 Voice Onset Time

VOT is the “primary acoustic correlate of stop voicing contrast” (Chang, 2019: 192) in many languages (Brown & Copple, 2018), including in English (Lisker & Abramson, 1964a), and in Hungarian (Bárkányi & Kiss, 2019; Gósy, 2001), but there are also other voicing cues, such as centre of gravity (Brown & Copple, 2018; Chodroff & Foulkes, forthcoming) or burst intensity (Sundara, 2005). VOT is “probably the most frequently measured acoustic component of any consonant” (Foulkes et al., 2010a: 62). Allen et al. found that VOT is speaker-specific, even after normalising for speaking rate (2003). As VOT is usually not a conscious feature that can be controlled by the speaker’s conscious efforts (Cheung & Wee, 2008), it is especially interesting to examine it in bilinguals, as it is said to be “particularly susceptible to cross-linguistic assimilation” (Bergmann et al., 2016: 73). It is suggested that perceived global accent is correlated to VOT productions (Gabriel et al., 2016; Major, 1987; Riney & Takagi, 1999; Schoonmaker-Gates, 2015; Yavas, 2002).

A plosive consists of an onset, a closure (or hold), and an offset phase; VOT is part of the hold phase (Byrd, 1993; Chodroff & Foulkes, forthcoming; Johnson, 2003). It can be divided into three categories: pre-voicing (also lead voicing), short-lag (unaspirated), and long-lag (aspirated) voicing (Abramson & Whalen, 2017), where the release of the stop is 0 (Beckman et al., 2013; Foulkes et al., 2010a). VOT is changing due to the aerodynamics of the vocal tract (Beckman et al., 2011; Cho & Ladefoged, 1999; Chodroff & Foulkes, forthcoming; Chodroff & Wilson, 2017), and the kind of distinction and the exact boundary depend on the language (Abramson & Whalen, 2017). A plosive can be produced with multiple bursts in a lot of languages (Chodroff & Foulkes, forthcoming). Regarding the acquisition of voicing, lead voicing is generally acquired later than long-lag VOT due to lead-voicing being more difficult to produce (Allen, 1985; Davis, 1995).

Hungarian speech research dates back to the 1700s (Gósy, 2000), yet the first Hungarian study on VOT was only carried out in 2001 by Gósy (compared to international research taking place from the mid-1900s onwards) (Lisker & Abramson, 1964a). Figure 3 shows the waveforms and spectrograms of the three voicing types.

Figure 3
Voicing types



Source: Own presentation

Two kinds of languages that use bimodal VOT variation can be established: true voice languages (which use pre-voicing and short-lag VOT), in this case Hungarian, and aspirating languages (which use short- and long-lag VOT), in this study English, to distinguish between voiced and voiceless categories (Cho et al., 2019; Volaitis & Miller, 1992). Table 2 presents the currently accepted features for voicing distinction in different types of languages.

Table 2

Features of plosives in aspirating and true voice languages

	Example languages	plosive	features
True voice	Hungarian, Russian, Spanish, French	/p, t, k/	$[\emptyset]$
		/b, d, g/	$[+voice]$
Aspirating	English, German	/p, t, k/	$[+spread]$
		/b, d, g/	$[\emptyset]$

Source: Own presentation based on Beckman et al. (2013: 260; with information from Bárkányi, 2018; Beckman et al., 2011; 2013; Deuchar & Quay, 2000; Jessen, 1996; Kleber, 2018; Petrova et al., 2006)

Voicing distinction is not always bimodal, it can be multimodal as well, e.g. Korean has three-way distinction (Hardcastle, 1973), while Hindi has a four-way distinction (Lisker & Abramson, 1964a). Due to this variation, VOT can be a cue to the person speaking in their L2 and may even provide superficial information about their L1 (Bóna & Auszmann, 2014).

On the one hand, English uses short-lag and long-lag VOT to distinguish between phonetically voiced and voiceless phonemes (Lisker, 1986; Lisker & Abramson, 1964a). Lisker explains this as “English /b/–/p/≠[b]–[p]” (1986). This means, that in English, the distinguishing feature is aspiration, not acoustical voicing (Lisker & Abramson, 1964b). Hungarian on the other hand uses pre-voicing and short-lag voicing in plosives (Gósy, 2001). However, some plosives may not be realised as plosives in Hungarian, this amounts to about 5% in word-initial position (Gráczi, 2011; Gráczi & Kohári, 2014; Száraz & Gráczi, 2020). This process seems to occur more often in voiced stops than in voiceless ones and is more common in males (Lavoie, 2001; Neuberger & Gráczi, 2013). As voicing is strong in Hungarian, voiced phonemes, if preceded by another word, may apply continuous voicing from the end of that word, also called VOT bleed (Gósy, 2004; 2016; Davidson, 2016; 2018). Bóna says that voiceless VOT in Hungarian moves between 9 and 96 ms (2016). The average values of VOT in Hungarian in different contexts and their standard deviations can be seen in Table 3.

Table 3
Means and standard deviations of VOT (in ms) of plosives in Hungarian in different contexts

plosive	Words in isolation (Gósy, 2001)		Spontaneous speech (Gósy, 2001)		#CV stops (Gósy & Ringen, 2009)	
	Mean	SD	Mean	SD	Mean	SD
/p/	24.64	11.95	18.51	5.92	9.7	5.4
/t/	23.30	8.18	26.59	6.47	16.0	6.7
/k/	50.17	18.71	35.31	10.78	37.6	13.4
/b/					–94.6	28.1
/d/					–95.1	29.8
/g/					–89.6	30.1

Source: Own presentation

While there is an overlap between Hungarian voiceless and English voiced stops regarding VOT, this does not mean that they are the same (Ahn, 2018; Beckman et al., 2013; Jordanidisz et al., 2015). Table 4 shows the VOT value differences between English and Hungarian.

Table 4
Means of VOT (in ms) of plosives in Hungarian and English

plosive	Words in isolation		Sentences	
	English	Hungarian	English	Hungarian
/p/	58	2	28	0
/t/	70	16	39	20
/k/	80	29	43	28
/b/	1	–90	7	–55
/d/	5	–87	9	–70
/g/	21	–58	17	–61

Source: Own presentation with information from Lisker & Abramson (1964a)

It is disputed if the Hungarian phonemes /c, ɟ/ are stops. Some argue that they are stops but they are realised as affricates in certain contexts (Siptár, 2015; Siptár & Törkenczy, 2000), while others argue that they are fully affricates and not stops (Szende, 1992). Table 5 shows the stop inventories of English and Hungarian.

Table 5
Stop inventories of English and Hungarian

	bilabial	alveolar	palatal	velar
English	p, b	t, d		k, g
Hungarian	p, b	t, d	c, ɟ	k, g

Source: Own presentation with information from The International Phonetic Association (1999: 41, 104, 10) and (Siptár, 2015)

The Hungarian phoneme /c/ only occurs in one stem-word, *tyúk* (hen) (Cseresnyési, 1992), and only makes up 0.02% of stops in Hungarian (Siptár, 2015), therefore it is not considered in the study.

Influences on VOT

First and foremost, VOT varies across categories from plosive to plosive, the further back the plosive is articulated, the longer the VOT is (Bóna, 2011; Byrd, 1993; Cho & Ladefoged, 1999). Next, factors for within-category variation are presented.

Speaking rate is the main factor mentioned when it comes to influences on VOT (Bóna, 2011). One study states that pre-voicing and long-lag VOT change as a function of speaking rate, while short-lag plosives do not (Beckman et al., 2011; Kessinger & Blumstein, 1997). Miller et al. found that the slower the speech, the longer the VOT values, and the perceptual boundary also shifts as a function of speaking rate (1986). Speaking-rate effects may vary in bilinguals, they might be similar to English if they are more proficient, but they might not exhibit any rate-effects if they are not very proficient (Flege & Schmidt, 1995).

Another important factor is that the VOT can change as a function of the length of the following vowel, where VOT is usually longer before high and tense vowels (Bóna, 2014; Chodroff & Wilson, 2017; Docherty, 1992).

Additionally, gender may also influence VOT values. In English, women tend to exhibit longer VOT values in voiceless plosives (Chodroff & Foulkes, forthcoming), while in Hungarian, men exhibit longer VOT in voiceless plosives (Chodroff & Foulkes, forthcoming; Gósy & Ringen, 2009). At the same time, pre-voicing is longer in women than in men in Hungarian (Gósy & Ringen, 2009). It is suggested that gender differences are a function of speaking rate, but this seems to not be the case (Swartz, 1992). Whereas others argue that gender only plays a role in children's speech due to biological (pulmonary) factors (Karlsson et al., 2004). While others say that gender does not play a role in VOT production at all (Morris et al., 2008).

Furthermore, the age of the speaker might also influence their VOT. In Bóna's study, Hungarian older people produced significantly longer VOTs in alveolar and bilabial stops than younger people (2011; 2014). This is most likely explained by slower articulation rate and biological changes in the vocal tract and pulmonary system (Bóna, 2011; Száraz & Gráczi, 2020). They also state that no such differences have been found in English (Bóna, 2011; 2014).

Code-switching can also affect VOT, where the speakers phonetically converge the VOTs of their two languages (Balukas & Koops, 2015; Bullock et al., 2006). As, according to the information stated in 2.1, both languages are activated in code-switching, it is important to consider the consequences of this for the purposes of the study.

Variation in VOT can be lexically conditioned, as words that occur with a higher frequency may be less accented (Baese-Berk & Goldrick, 2009; Chodroff & Wilson, 2017; Levi et al., 2017). Furthermore, differences in VOT can also be found in careful vs. spontaneous speech (Gósy, 2001), where VOT is longer in careful speech (Chodroff & Foulkes, forthcoming; Zen, 2020). According to Tar, VOT can also be affected by speech disorders (2014). Apart from this, stressed syllables have longer VOT (Lisker & Abramson, 1967), which may be a problem when comparing English and Hungarian as in Hungarian word-stress is always on the first syllable (Rounds, 2009).

Further on, studies that have been carried out on the topic of VOT in bilinguals are reviewed. Bilinguals tend to show either exaggerated or converged (compromise/hybrid) VOT values in their two languages (Bullock et al., 2006; Flege, 1987; 1991; Wrembel, 2014). Additionally, VOT in bilinguals also tends to present more variation (Chodroff & Baese-Berk, 2019).

Flege states that early L2 learners' VOT are equivalent to those of monolingual speakers (Flege, 1991). Abrahamsson & Hyltenstam found that early learners are closer to nativelike targets of VOT in their L2 than late learners (2009; also Gráczi, 2016). Early child bilinguals exhibit contrasts that are like that of their monolingual peers, but they also exhibit more within-category variation (MacLeod & Stoel-Gammon, 2005). Flege et al. found that the later the AOA, the more likely it is for VOT to differ from native targets (1996). Late learners are usually unable to acquire nativelike VOT, especially in sounds that exist in their L1, which supports the equivalence theory (Alba-Salas, 2004). Some argue that VOT can be influenced by staying in a country where the target language is spoken (Podlipsky et al., 2021), while others say that is not the case and VOT remains unaffected (van de Weijler & Kupisch, 2015).

Jordanidisz et al. found that Hungarian-English bilingual children's VOT is longer than that of monolinguals (2015). French-English bilinguals exhibit slightly less pre-voicing than monolinguals (Kehoe & Kannathasan, 2021). In French-English bilinguals, Fowler et al. found that 2L1 speakers produce shorter VOT in English than monolingual speakers, and sequential bilinguals produce even shorter VOT than 2L1 speakers (2008). Even if they seem to have a global foreign accent, 2L1 speakers show distinct VOT ranges (Lein et al., 2016). Moreover, Flege states that highly proficient L2 speakers may even see effects of L2 VOT on the VOT of their L1 (1992). In Au et al.'s study, Spanish HS do not have different VOT compared to their monolingual peers (2002). Geiss et al. support this result and say that HS are better at approximating nativelike VOT in L3 than L2 speakers (2022). Dittmers et al. state that VOT is not monolingual-like in HS (2018).

As far as VOT shift is concerned, it can occur inter-generationally with later generations having less nativelike VOT (Hrycyna et al., 2011; Nagy & Kochetov, 2013). Furthermore, productions of VOT can merge due to language attrition, especially in late learners (de Leeuw, 2019).

Chodroff & Foulkes argue that as VOT has high intra- and inter-speaker variation, it does not seem to be ideal for use in forensics (forthcoming), but Coe states that it is a good parameter, especially since it is robust against any kind of transmission problems (2012). Nonetheless, Earnshaw found that it is less good in spontaneous speech (2014).

2.3 Forensic Linguistics

Forensic linguistics concerns the "use of linguistic expert evidence in legal proceedings" (Eades, 2011: 313).

This section starts with a general overview of the forensic fields that the present research could be applied to. Afterwards, the LR framework and its uses in forensic speech science are presented.

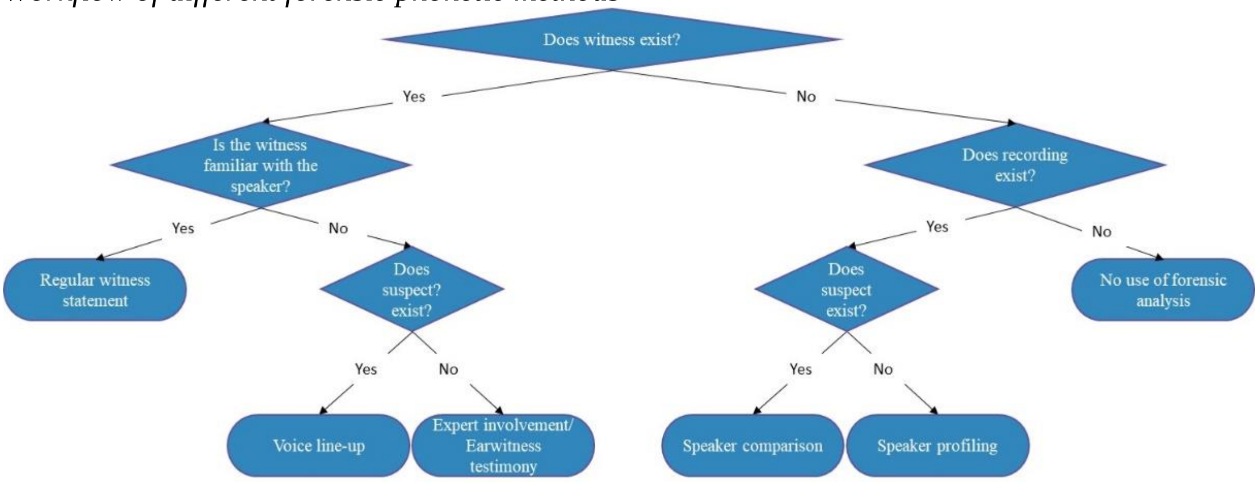
2.3.1 Forensic Speech Science

FSS deals with anything that concerns voice (and to an extent audio) as evidence. Voice evidence may be used if no other evidence is available, or to support other evidence, but it should not be a major (or only) part of the evidence (Bull & Clifford, 1984). This includes, but is not limited to, speaker comparison, speaker profiling, transcription, and voice line-ups (Baldwin & French, 1990; Coulthard & Johnson, 2007; Eades, 2011; Foulkes & French, 2001; 2012). As the field is an application of sociophonetics, it is essential that sociophonetic research is carried out with the utmost care to ensure its uses in forensics are valid (Hughes & Wormald, 2020; Rose, 2002).

Sociolinguistic speaker characteristics, such as dialect, foreign accent, age, gender, even medical conditions, are used in the field of FSS to draw conclusions about the offender (Jessen, 2007; Rock, 2011; Schilling & Marsters, 2015; Watt, 2010). According to Nolan, the individuality of the speaker should be considered in any forensic examination, as speech is variable (1997). Therefore, the best possible case is that the intra-speaker variation is low, and the inter-speaker variation is high (Rose, 2002; Rose & Morrison, 2009). The study of bilingualism in the field of forensics is that it is a source of sociophonetic variation, which can be useful (Foulkes et al., 2010b).

Figure 4 shows the different types of forensic phonetic casework and the situations they are applied in.

Figure 4
Workflow of different forensic phonetic methods



Source: Pakai, 2021

Speaker comparison is applied when a recording of the offender exists and there is also a recording of the suspect available, while speaker profiling is employed when no suspect has been identified yet (Jessen, 2008; 2010). In speaker profiling, the offender is classified into linguistic categories based on the recording, while in speaker comparison the two recordings are compared (Jessen, 2021; Rose, 2002; Schilling & Marsters, 2015). Additionally, accent identification may also be used in the field of Language Analysis for the Determination of Origin (LADO) (Bobda et al.,1999; Schilling & Marsters, 2015).

In the United Kingdom, speaker comparison was traditionally carried out on an auditory basis, but now also includes acoustic investigation (Coulthard & Johnson, 2007; French, 1994; Nolan, 1991). In the United States, "voiceprints" were used as a means of forensic speaker, but as there is no such thing, this tradition has been largely replaced by an acoustic or joint auditory-acoustic examination method (Baldwin & French, 1990; Coulthard & Johnson, 2007; Foulkes & French, 2001; French, 1994; Hollien, 1990; Nolan, 1994). Further, automatic methods can also be applied, such as Automatic Speaker Recognition (ASR).

A potential problem that could arise in FSS is the suspect having a hybrid accent, or them using an accent as disguise, may that be their own accent or another (Baldwin & French, 1990; Clark & Foulkes, 2007; Coulthard & Johnson, 2007; Dumas, 1990; Rogers, 1998; Yarmey, 2012). Therefore, before any forensic examination method is used, disguise must be ruled out (Künzel, 1994). Near-native and nativelike speakers could cause problems as their non-nativeness may not always be observable, leading to wrong reference populations and potentially wrong suspects (French & Stevens, 2013). If a foreign accent can be detected, it might still be difficult to identify the L1 of the perpetrator (Jessen, 2007; 2010). Furthermore, when using ASR with a language mismatch, it is essential to normalise for the language (Akbacak & Hansen, 2007). Additionally, a foreign accent may affect speaker profiling of ear-witness testimony due to prejudice (Baugh, 2007).

2.3.2 The Likelihood Ratio Framework

The LR framework initially gained popularity due to DNA testing in forensic science (Lucy, 2005; Morrison, 2010). Nowadays, when giving evidence is concerned, it is also applied to the subfield of forensic phonetics. According to Rose & Morrison it is the "logically and legally correct framework for the evaluation of forensic comparison evidence" (2009: 143). The Association of Forensic Science Providers states that conclusions of support for either the prosecution or the defence should be expressed based on the LR (2009). In the LR framework, not only the evidence itself is considered, but also the typicality of the given evidence in the appropriate population (Lucy, 2005). It is part of the formula used to reach the posterior odds in Bayes' theorem (Lucy, 2005; Morrison, 2011b).
The equation for the LR is shown in (1) (French & Stevens, 2013: 194).

$$LR = \frac{P(E|H_{SS})}{P(E|H_{DS})} \tag{1}$$

"The central principle [of the LR] is that the weight of the evidence is determined by the ratio of the probability of the evidence, given the prosecution alternative, to the probability of the evidence, given the defense alternative" (Evett et al., 1993: 499); i.e. the value of the evidence is measured by the LR (Champod & Meuwly, 2000). The numerator is made up of the probability of observing the evidence in the suspect, while the denominator is made up of the probability of observing the same in the appropriately selected population (Lucy, 2005). It is important, that the LR framework cannot only be applied to technical linguistic evidence but more generally to linguistic features as well (Rose, 2006).

It is of further relevance, that the members of a court, including the jury, may not be able to understand what the LR means, and it should therefore be adequately explained (Champod & Evett, 2000; Fenton et al., 2014; Lynch & McNally, 2003; Redmayne, 2001). In England, the Regina v. T. (2010) Court of Appeals case demonstrated that the LR framework was misunderstood by the court, which makes it all the more important to explain it to the court to leave no room for misunderstandings (Morrison, 2012).

For easier understanding of laypeople, the LR results can be log-scaled to avoid the awkwardness of the LR scales. The LR can also be expressed verbally, as shown in Table 6.

Table 6
Verbal LR expressions

LR	Log LR	Verbal LR	
> 10,000	***> 4***	Very strong evidence to support…	…the same-speaker hypothesis (H_{SS})
1,000 to 10,000	***3 to4***	Strong evidence to support…	
100 to 1,000	***2 to 3***	Moderately strong evidence to support…	
10 to 100	***1 to 2***	Moderate evidence to support…	
1 to 10	***0 to 1***	Limited evidence to support…	
0.1 to 1	***0 to –1***	Limited evidence to support…	…the different-speaker hypothesis (H_{DS})
0.01 to 0.1	***–1 to –2***	Moderate evidence to support…	
0.001 to 0.01	***–2 to –3***	Moderately strong evidence to support…	
0.0001 to 0.001	***–3 to –4***	Strong evidence to support…	
< 0.0001	***< –4***	Very strong evidence to support…	

Source: Adapted from Rose (2002: 61–62)

As the LR also considers the probability of an observation occurring in the appropriate population, it is important to define the correct reference populations (French & Harrison, 2007; Gold, 2014; Gold & Hughes, 2014; Hughes, 2011; Rose & Morrison, 2009). This is especially relevant in FSS, as it may not be known what exact sociolinguistic community the suspect is from, and therefore creating an appropriate reference population may prove difficult, and mismatched populations may produce incorrect LR results (Hughes & Foulkes, 2015). Additionally, the sample size of the selected reference population also needs to be considered as this has effects on the LR too (Hughes, 2014; 2017; Ishihara & Kinoshita, 2008); Hughes suggests at least 20 speakers to reach stable LR outcomes (2017).

However, the use of the Bayesian approach in the legal system can be problematic, because it could invoke the prosecutor’s fallacy, which means that the probability of the hypothesis given the evidence is interchanged with the probability of the evidence given the hypothesis (French & Harrison, 2007; Lucy, 2005).

It also has to be considered for the purpose of the study that, generally, the LR framework might be problematic due to language-mismatch (Lo, 2021). This is especially relevant in the case of bilingualism.

3. Present Study

The study is an empirical study using quantitative methods to examine VOT in Hungarian-English bilinguals. This section starts with the research aims which are followed by the research questions. At the end, the hypotheses are presented.

3.1 Research Aims and Questions

The study aims to find any differences in VOT of bilingual Hungarian-English speakers compared to monolingual speakers of the respective languages. It also aims to find reasons for this. Furthermore, the present study is created in a way that potential applications in forensic linguistics can be found and examined.

The research questions are as follows: To what extent L1 influences L2 VOT? Does VOT change with time spent in an L2-speaking country? It is expected that participants residing in English-speaking countries will be closer to native targets. Furthermore, do other factors influence English VOT in bilinguals? In the context of FSS, it is especially interesting to examine if L2 VOT could lead to better LR system performance for same-speaker recognition, and if, therefore, L2 VOT could be a better speaker-discriminant than L1 VOT.

3.2 Hypotheses

The hypotheses are as follows:

H_1: Bilingual productions of English have different VOT compared to monolinguals.

H_2: People who have spent time in an English-speaking country perform better at reaching target VOT values.

H_3: The bilingual data will lead to better system performance in an LR framework.

4. Methodology

The study was restricted to the examination of word-initial, pre-vowel (#CV) sequences, that were preceded by at least 40 ms of silence, as analytic problems may arise in other positions, e.g. due to coarticulation (Foulkes et al., 2010a).

Manual VOT measurements were taken at zero-crossings from the waveform, as this showed the burst and the onset of periodicity the best (Foulkes et al., 2010a; Ladefoged, 2003; Francis et al., 2003). When there were multiple bursts (Chodroff & Foulkes, forthcoming; Gráczi & Kohári, 2014; Lavoie, 2001), VOT was measured from the first burst, up to the onset of voicing (Foulkes et al., 2010a). VOT could not be measured if VOT bleed took place (Deuchar & Clark, 1996). For the data evaluated by the author, VOT was measured to the closest millisecond.

This section deals with the methodology. The data is introduced, then the analysis methodology is laid out, including the calculation of the LRs.

4.1 Data

First, the Hungarian monolingual corpus and collection of its data is outlined. Second, the English monolingual data is introduced. Finally, the collection of the bilingual data is described.

4.1.1 Hungarian Monolingual Data

For the purpose of monolingual Hungarian data, the Beszélt Nyelvi Adatbázis (BEA corpus) by the Hungarian Academy of Sciences (MTA) was used (Mihajlik et al., 2022). The corpus was provided by the MTA after registering online and filling in an ethics form.

The BEA Corpus

The BEA corpus was started to be developed in 2007, with first recordings taking place in 2008 (Gósy, 2008). 114 participants were interviewed by the MTA to create an audio corpus of spoken Hungarian. The corpus contains multiple speech styles: repetition of sentences, reading (of sentences and passages), retelling of an article and a story, and conversation – both with only the interviewer and with an additional conversation partner (Gósy, 2008; Gósy & Gyarmathy, 2017). All participants are monolingual Hungarian speakers from Budapest, with no dialect, speaking Educated Colloquial Hungarian (ECH) (Gósy, 2008). The corpus was provided in folders with separate sociolinguistic interviews in separate folders. The recordings were cut into approximately 1 s to 10 s segments. The files were in WAV format. Furthermore, the speech of the interviewer and potential conversation partners were also left included in the folders. This led to some issues in the processing of the data which is further explained in the next sections. Transcripts of the segments were provided as TXT files, in which foreign words were transcribed using Hungarian orthography (Gósy, 2008), e.g. *box* transcribed as <*boksz*>. The BEA corpus was also later set up to be used as an ASR benchmark, which is the dataset that was provided to the author (Mihajlik et al., 2022).

Data Processing

28 participants were selected from the BEA corpus for the purpose of analysis, based on age and gender, to create an appropriate sample of the Hungarian population.

First, the recordings were listened to so any recordings of the interviewer and the third interview partner could be excluded. Any recordings that did not have a plosive in #CV position were also excluded.

The selected recordings were aligned using the Montreal Forced Aligner (MFA) (McAuliffe et al., 2017). An acoustic model for Hungarian by Ahn & Chodroff was used (2022). A Grapheme-to-Phoneme (G2P) model was created from the existing Ahn & Chodroff dictionary (2022), which was then used to supplement the provided dictionary with words from the new data so that alignment would be of good quality.

For the voiceless plosives, AutoVOT (Keshet et al., 2014) was used and extracted using a Praat-script (Boersma & Weenink, 2022). The windows were extended by 30 ms on both sides (Chodroff et al., 2015). For the purpose of using AutoVOT, files were converted to 16,000 Hz sampling rate using SoX (SoX team, 2015). For the voiced plosives, Dr. VOT (Shrem et al., 2019) was used. First, the aligned recordings were segmented into audio files of the separate words using a Praat-script by Lennes (2002; Boersma & Weenink, 2022). The recordings were then renamed to include speech style and the participant's number. Any recordings not starting with /p, t, k, b, d, g/ were removed from the folder to be processed.

The VOT measurements were then checked to see if they are correct under strict conditions. In the case of the voiceless plosives, measurements of $VOT_p \leq 5\,ms$ or $VOT_p > 25\,ms$, $VOT_t < 10\,ms$ or $VOT_t > 35\,ms$, and $VOT_k < 15\,ms$ or $VOT_k > 45\,ms$ were manually checked and corrected if necessary. In the case of /p/, 5 ms were included because it was observed that this may indicate an error with measurement. For voiced plosives any measurements of $VOT_b < -120\ ms$ or $VOT_b > -40\ ms$, $VOT_d < -130\ ms$ or $VOT_d > -50\ ms$, $VOT_{\text{g}} < -140\ ms$ or $VOT_{\text{g}} > -60\ ms$, and $VOT_{\text{ɟ}} < -150\ ms$ or $VOT_{\text{ɟ}} > -70\ ms$ were manually checked, including random spot-checks for anything falling into the accepted ranges. This was done to ensure that the quality of the data is reliable, despite Dr. VOT not always being fully reliable in the case of negative VOT.

A total of $n = 13{,}555$ measurements were taken from the BEA database, of which $n = 2{,}183$ observations were excluded for various reasons, such as no stop release, or pre-voicing and lag-voicing occurring in the same instance. This led to $n = 11{,}372$ remaining observations that are used for the purpose of the study (Appendix A).

4.1.2 English Monolingual Data

Two sets of monolingual English data are used for the study. The first one was collected by Coe (2012) from the DyViS corpus of Standard Southern British English (SSBE) speakers (Nolan et al., 2009) and it contains measurements of /p, t, k/ for 20 speakers. It includes $\text{n} = 1{,}132$ observations. The second one was collected by Chodroff & Wilson (2018) from the Mixer 6 corpus (Brandschain et al., 2010) and contains data of all English plosives from 180 speakers of American English (AmE). It contains $\text{n} = 96{,}357$ observations. Both corpora only contain read speech.

4.1.3 Bilingual Data

Bilingual data was collected by the researcher between 22/05/2022 and 11/07/2022. A pilot study was carried out with two participants to determine the ideal length of the sociolinguistic interview. The recordings from the pilot study are not included in the analysis.

Data collection

Due to the ongoing Covid-19 pandemic and the participants living in different counties or different parts of the United Kingdom, collecting data in person was not possible, except for with a few participants. The methodology for collecting the bilingual data remotely was adapted from Zhao & Chodroff (2022). For participants with an Android phone, the app ASR (NLL, 2021) was used, while participants with an Apple device were asked to use AVR X – Voice Recorder (Newkline Co., Ltd., 2022). The recordings were asked to be set to WAV, 44,100 Hz and 128 kbps (Zhao & Chodroff, 2022). The same settings were used for in-person data collection.

Data collection took place over Zoom. A total of $n = 21$ participants were interviewed. At the beginning of the interview, the participants were given a participant number and were asked to fill in a survey (Appendix B) in which they gave consent to data collection. Furthermore, they were asked to provide additional information about themselves, including questions which confirmed whether or not they were native speakers of Hungarian and spoke English at a sufficient level. However, the considerations laid out in 2.1 make it difficult to determine where the line should be drawn between monolinguals and bilinguals (Bialystok, 2001). Originally, only nativelike speakers were meant to be interviewed, but due to lack of participants, this was only partially possible. This affected the further study, as the study was conceived to be mostly undertaken by nativelike speakers. The fact that no two bilinguals are the same, makes the choice of the correct methods for the present study increasingly problematic (Costa et al., 2017; de Bruin, 2019). The participants were also asked to provide information about where they have lived and the languages they speak. Further information was collected regarding their age, gender, and education. The use of a test such as the MLAT was not possible to test further factors such as language aptitude, as acquiring these tests is difficult. The results of the survey can be found in Appendix C.

During the interview, all participants were asked to read sentences and three passages in Hungarian and summarise them to elicit controlled spontaneous speech (Gósy, 2008). The sentences, with a couple of additions, and the passages were adopted from the BEA corpus to make the collected data more comparable. Furthermore, they were also asked to answer two to four questions to elicit spontaneous speech. The same process took place for the collection of the English data, but only one, longer, text was read. The English text was adapted from Shaw et al. (2018). The elicitation passages and potential questions can be found in Appendix D. All participants spoke for a total of approximately 30 minutes, 15 in Hungarian, and 15 in English.

Due to several reasons, such as insufficient English or Hungarian knowledge, which did not correspond with some of the survey-answers, background noise, or poor audio quality, $n = 11$ participants had to be excluded from the study. This occurred potentially due to the self-assessment in the survey leading to misjudgement of one's own competence, as self-awareness may affect self-assessment (Baker & Wright, 2021). A further three participants were not analysed as the limited time allowed for the dissertation made any further analysis difficult. As a result, 30% ($n = 7$) of the participants' data was evaluated.

As far as the limitations of the data collection are concerned, the following points must be mentioned. Since the participants were aware that they were being recorded, and in the case of the bilingual data may have even recorded themselves, the issue of the observer's paradox is raised; this might mean that the participants were more careful in their pronunciations, despite the study aiming to elicit naturalistic speech (Locke, 1983; Xie & Jaeger, 2020). Additionally, the interviews show the speech of a participant at a given day and time and might not reflect their overall competence (White & Genesee, 1996). Further, the language modes of the bilingual need to be considered. Due to the interviewer being bilingual as well, a bilingual environment was created, which may have influenced the participants' pronunciation. If the study is replicated, monolingual speakers of each language should be used as interviewers to avoid such a problem.

Data processing

Where necessary – this was the case for in-person recordings – the interviewer was removed from the recordings so only the participant would be audible. The recordings were edited using Audacity (Audacity Team, 2012) to remove the interviewer and to segment the recordings into different speech styles and questions.

A Python (Python Core Team, 2015a) script (Appendix E) was created to carry out the transcription of the data. For this purpose, the Google Speech API (Google LLC, 2022), Pydub (Robert, 2011), and the library Speech Recognition (Zhang, 2017) were used. Any transcription created was also checked by listening to the audio file and reading the transcription at the same time. The transcription was created in a way that anglicisms in Hungarian were transcribed with Hungarian orthography (Gósy, 2008) to make forced alignment easier. Punctuation marks and numeric characters had to be removed from the automatic transcriptions. There may also have been further perception problems in transcription (Vaux et al., 2006).

Forced alignment for Hungarian was carried out by using the MFA (McAuliffe et al., 2017) with an acoustic model by Ahn & Chodroff (2022). Due to compatibility issues with Windows, this was carried out using Linux and a different Python version (Python Core Team, 2015b). A new dictionary was created using the dictionary and a G2P model, which was created for the use on the BEA corpus and amended by new words from the bilingual data. The forced alignment was then manually checked and adjusted where necessary. 14 of 197 recordings had to be aligned manually for Hungarian.

The alignment for English was carried out using the english_uk_mfa G2P model, with which the english_mfa dictionary was supplemented. The english_mfa acoustic model was used for the creation of the G2P and the alignment. 4 of 108 recordings had to be aligned manually.

Since it was expected that there would be more variation within VOT, including categorical variation between negative and positive VOT, only AutoVOT was used for the processing of bilingual data. To use AutoVOT, the sampling rate of the files was converted to 16,000 Hz sampling rate using SoX (SoX team, 2015). The windows in AutoVOT were extended by 10 ms on both sides for voiced and by 30 ms for voiceless plosives (Chodroff et al., 2015). The measurements were extracted using a Praat-script. Due to the expected higher variation, every measurement was manually checked. A total of n $= 4{,}876$ observations were made and checked, of which $n = 2{,}211$ were data for English, and $n = 1{,}642$ were for Hungarian (Appendix F). All stops of the selected participants were measured to avoid any bias in selecting "good" stops.

4.2 Analysis Methodology

The data was not normalised for various reasons. First, as the main application of the study would be in FSS, normalisation is not warranted. Second, any normalisation for speech rate was not possible in the BEA corpus due to the chopped-up nature of the recordings. This was also not possible in the corpus with SSBE recordings as no data regarding speech rate was provided. Third, no normalisation using the length of the following vowel was possible as the length of the vowel was not adjusted when the VOT was checked. Furthermore, no normalisation took place for gender or age, as the variation for those factors is different in English and Hungarian (as outlined in 2.1). Therefore, it would have been counter-productive to normalise for these.

Analysis was conducted using R and the R Studio integrated development environment (R Core Team, 2021). The tidyverse (Wickham et al., 2019), dplyr (Wickham et al., 2021), plyr (Wickham, 2011), ggplot2 (Wickham, 2016), and tidyr (Wickham, 2021) packages were used in addition to base R. The R script used for the dissertation can be found in Appendix G.

Monolingual data and bilingual data were first processed separately, then they were compared to each other. Significance testing on VOT values was only carried out on /p, t, k/, while for /b, d, g/ the realisations of them having negative or positive VOT were tested. This was due to /b, d, g/ having too much of a split between positive and negative values, which would have yielded unmeaningful results if analysed the same way as /p, t, k/.

Any values that were not expected for Hungarian in the monolingual data ($/b, d, \text{g}/ > 0$ and $/p, t, k/ < 0$) were excluded from the data. In the bilingual data, they were split into two different datasets for /b, d, g, ɟ/ depending on if they had positive or negative VOT. /p, t, k/ with negative VOT were excluded. Any mean values had to be calculated separately for negative and positive VOTs, as doing otherwise would have rendered the results unmeaningful. /ɟ/ was only analysed for Hungarian, as it does not exist in English.

The data was tested for normality using the Shapiro-Wilk normality test, which returned that the data was not normally distributed (at $p < 0.05$). Therefore, the assumptions of ANOVA were not met, and a one-way ANOVA could not be performed. Instead, the Kruskal-Wallis one-way analysis of variance was used to assess if results are significantly different under different conditions. Where the Kruskal-Wallis test could not be performed, the Kolmogorov-Smirnov test was used instead. The significance threshold was set at $p < 0.05$.

Percentages for the realisations of the voiced plosives (negative or positive VOT) in Hungarian and English were also calculated for the bilingual data.

4.3 Likelihood Ratios

In the field of forensic phonetics, multiple approaches can be used to calculate LRs, e.g. Gaussian mixture models or Multivariate Kernel Density (MVKD) (Morrison, 2011a). Calculations for this study were carried out using MVKD.

For the calculation of LRs, the data had to be restructured (Appendix H). The LRs and the tests for system performance, by using Equal Error Rates (EER) and Log-likelihood-ratio cost (C_{llr}), were calculated using MATLAB (The MathWorks Inc., 2010) using three scripts (MVKD_loop_CrossVal.m, Calibration_CrossVal.m, and MVKD_loop.m) provided by Hughes (2013a; 2013b; 2013c). The EER is the intersection between the false acceptance rate and the false rejection rate.

Due to the limited amount of data available, both read and spontaneous speech data were used at the same time.

The monolingual LRs were calculated using cross validation, while the bilingual data was calculated by using the monolingual data as the reference population. To perform same-speaker comparisons, the data was divided into two sessions by splitting each participant's recordings in half. Additionally, bilingual data for each language was also calculated using cross validation using itself as a reference population. The same was done for bilingual data with language-mismatch.

LR for English were only ran on the SSBE dataset, as the AmE dataset contained too much data to be processed.

5. Results

The results are presented in this part. In the following section, the descriptive statistics are presented. After that, the results of the data analysis are outlined, including both sets of monolingual data and the bilingual data. Finally, the results for the research questions and the LR system test are laid out.

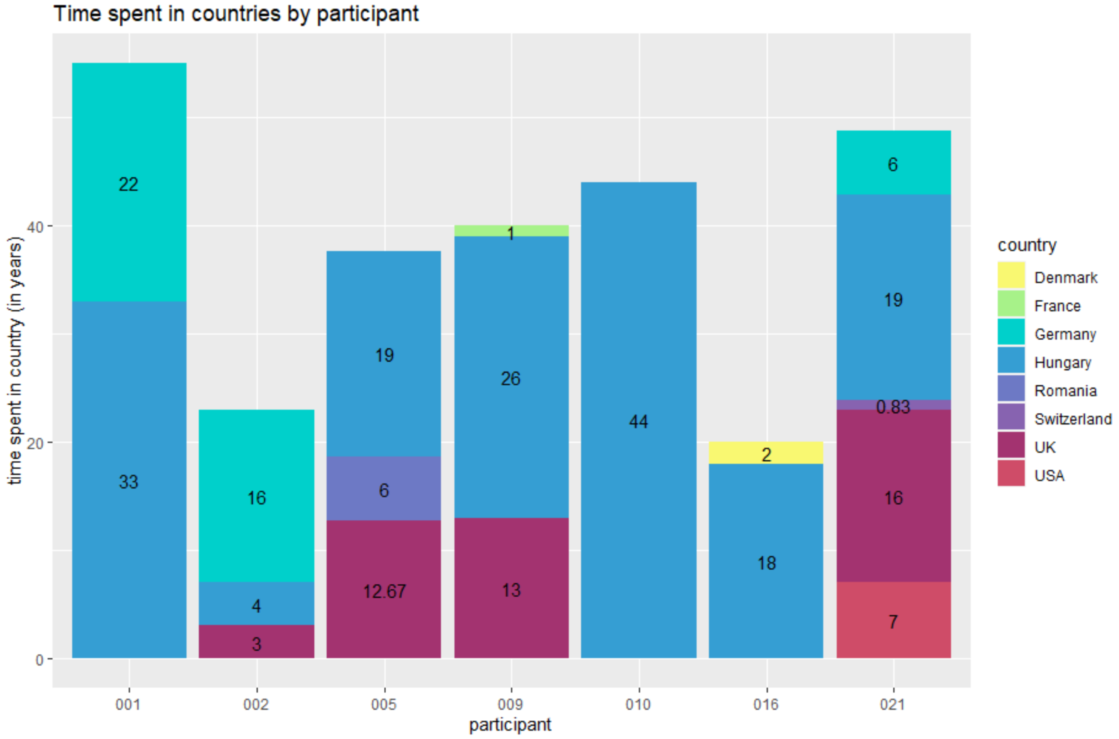
5.1 Descriptive Statistics

14 participants of each sex were selected from the BEA corpus for the analysis. The selected participants ranged between 20 and 76 years of age. There is no further information available about the participants of the BEA corpus.

A total of $n = 7$ bilingual participants were analysed from the data collected by the researcher. Of the participants, $n = 5$ were female, and $n = 2$ male. They were between 20 to 55 years old. The participants were born in Hungary, $n = 3$ in Budapest, $n = 1$ in Győr, $n = 1$ in Szőny, $n = 1$ in Tatabánya, and $n = 1$ in Miercurea Ciuc in the Hungarian-speaking part of Romania. None of them displayed any audible accent and they were all deemed speakers of ECH. Table 5 shows the years spent in different countries by the study participants.

Figure 5

Time spent in different countries by participants (in years)



Source: Own presentation

5.2 Data Analysis

The results of the analysed data are presented in this section. The monolingual Hungarian and English data are laid out first, after which the results of the analysis of the bilingual data are shown.

VOT in monolingual Hungarian speakers

Table 7 shows the means and standard deviations of the Hungarian speakers from the BEA corpus.

Table 7

VOT data (in ms) for monolingual Hungarian stops

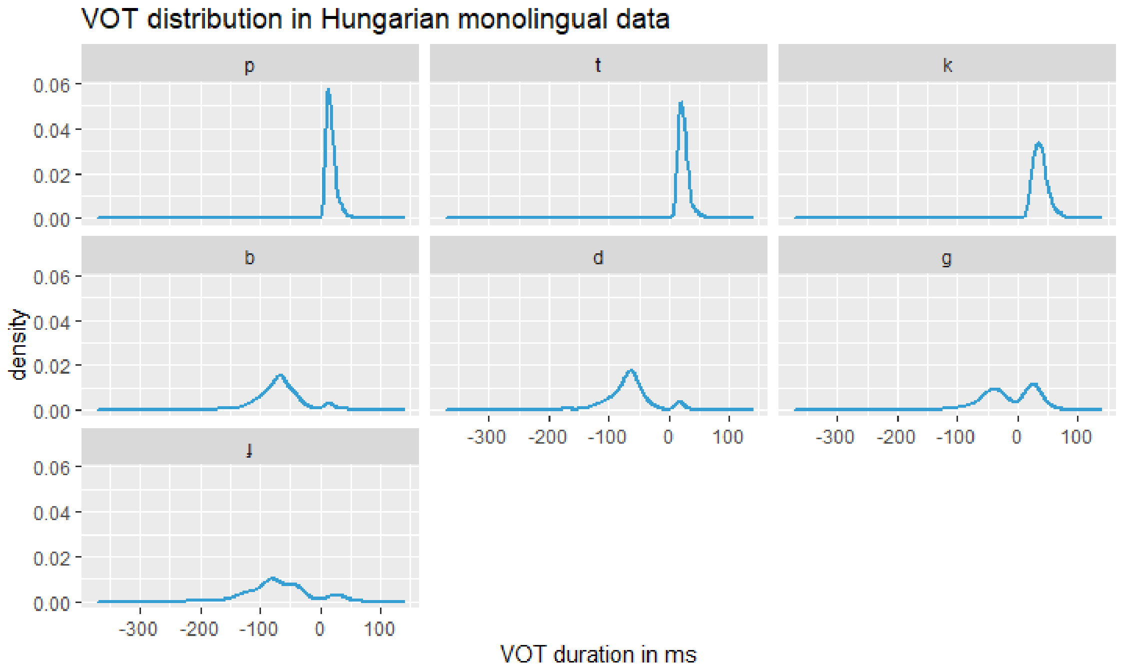
	mean	standard deviation
/p/	18.3	3.89
/t/	24.7	2.99
/k/	37.9	4.50
/b/	–74.7	8.68
/d/	–75.4	14.5
/g/	–47.5	15.1
/ɟ/	–79.2	17.0

Source: Own presentation

Figure 6 shows the distribution curves of VOT values in the Hungarian monolingual data.

Figure 6

VOT distribution in Hungarian monolingual data

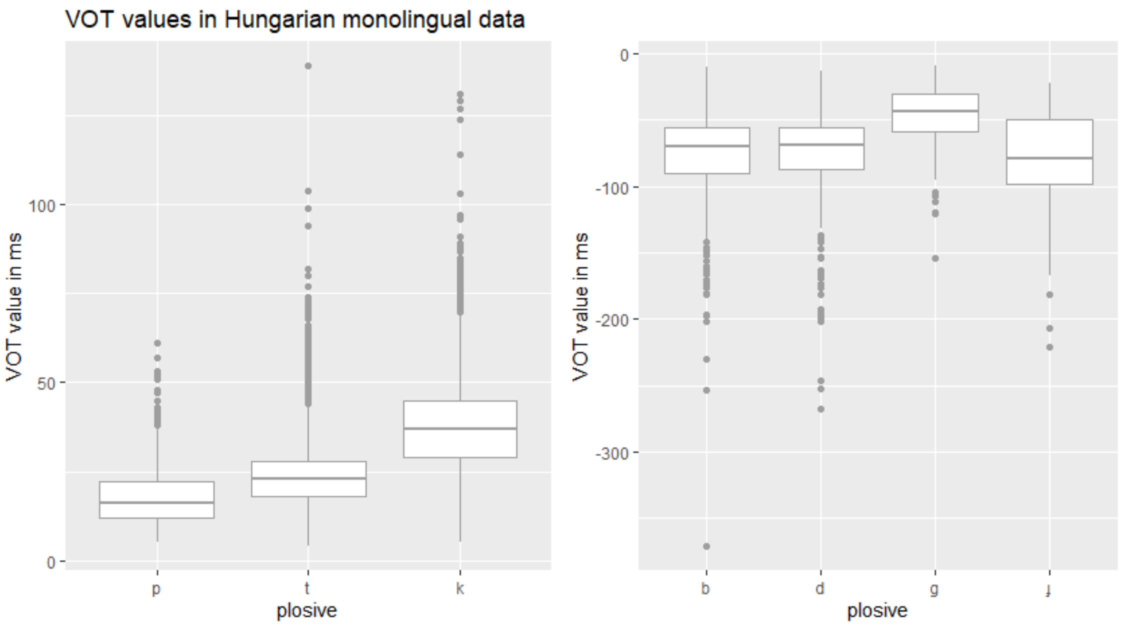


Source: Own presentation

Figure 7 shows boxplots of the VOT values in the monolingual Hungarian data.

Figure 7

Boxplot of VOT values in Hungarian monolingual data



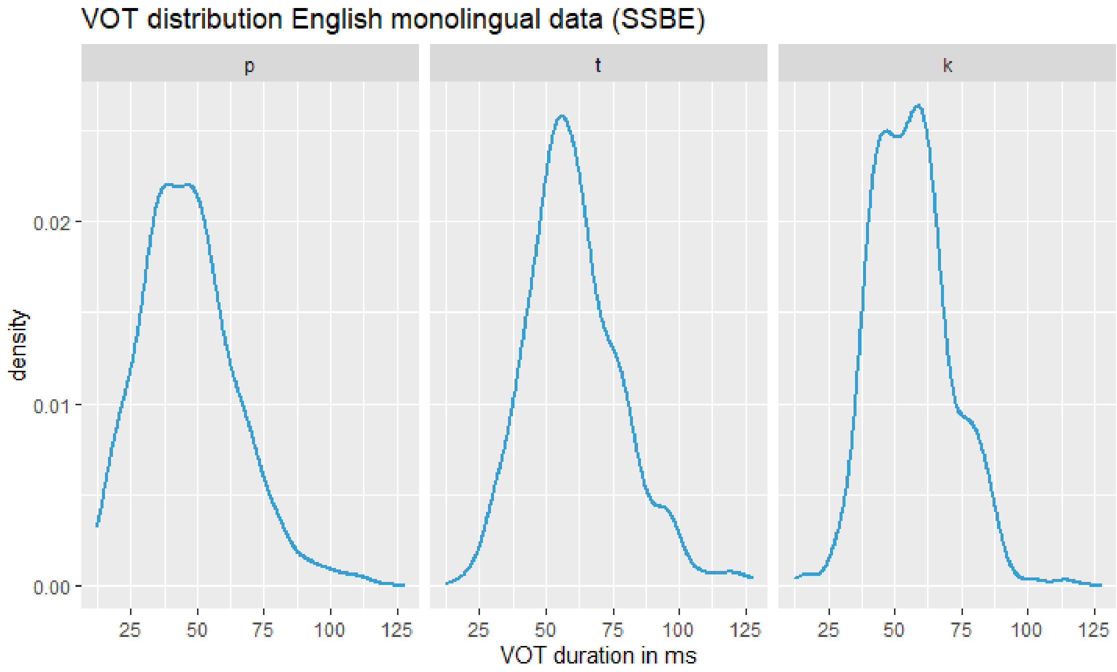
Source: Own presentation

VOT in monolingual English speakers

Figures 8 and 9 show the distributions of VOT values in the two sets of monolingual English data.

Figure 8

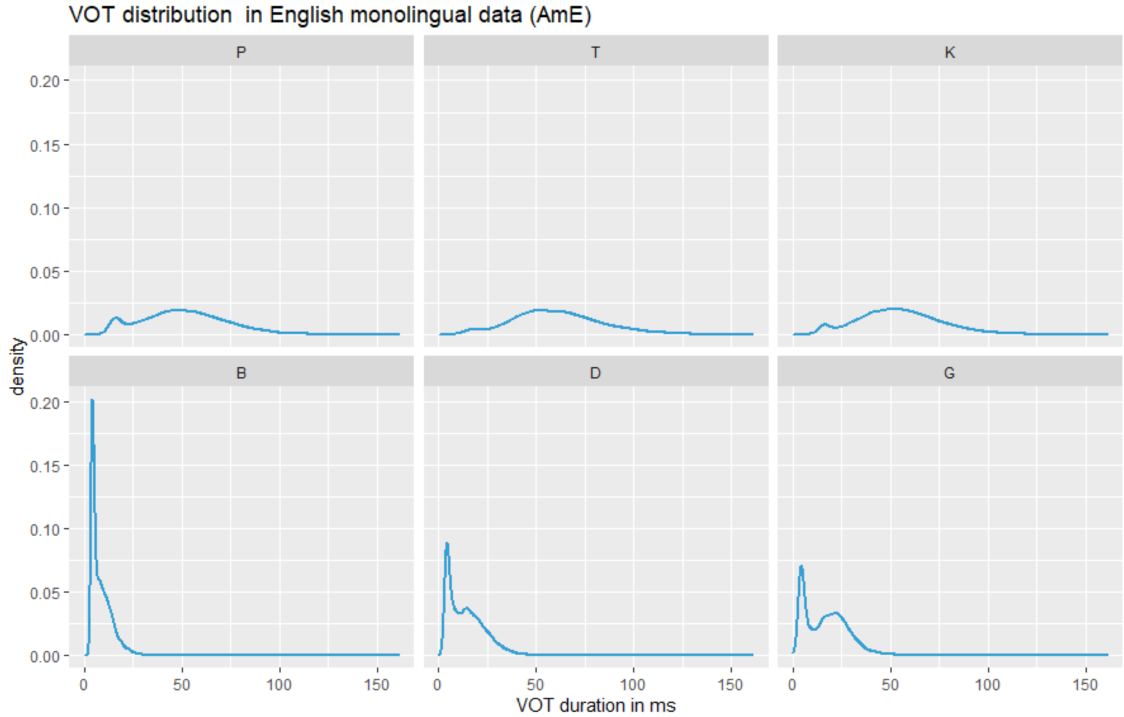
Distribution of VOT values in English monolingual data (SSBE)



Source: Own presentation

Figure 9

Distribution of VOT values in English monolingual data (AmE)



Source: Own presentation

Table 8 shows the VOT means and standard deviations for the SSBE and the AmE monolingual English datasets.

Table 8

VOT data (in ms) for both sets of monolingual English data

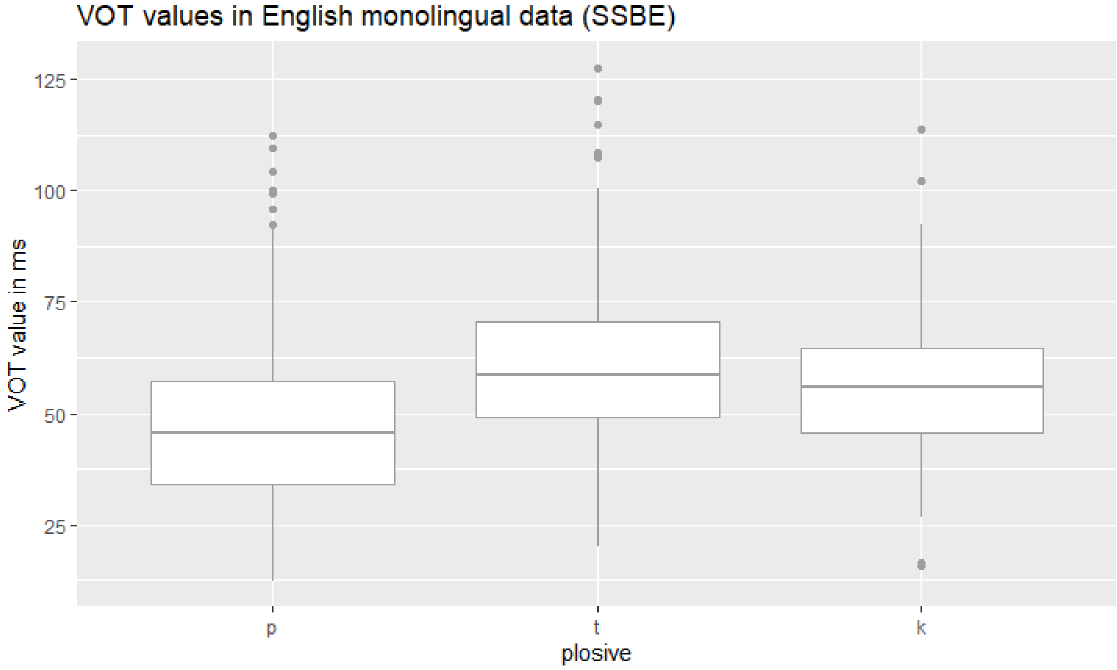
	SSBE		**AmE**	
	mean	**standard deviation**	**mean**	**standard deviation**
/p/	46.8	10.8	50.6	9.41
/t/	61.0	9.32	61.3	9.44
/k/	56.3	7.81	54.4	7.63
/b/			8.41	1.57
/d/			13.9	2.73
/g/			16.8	3.19

Source: Own presentation

Figures 10 and 11 show boxplots of the VOT values in both sets of monolingual English data.

Figure 10

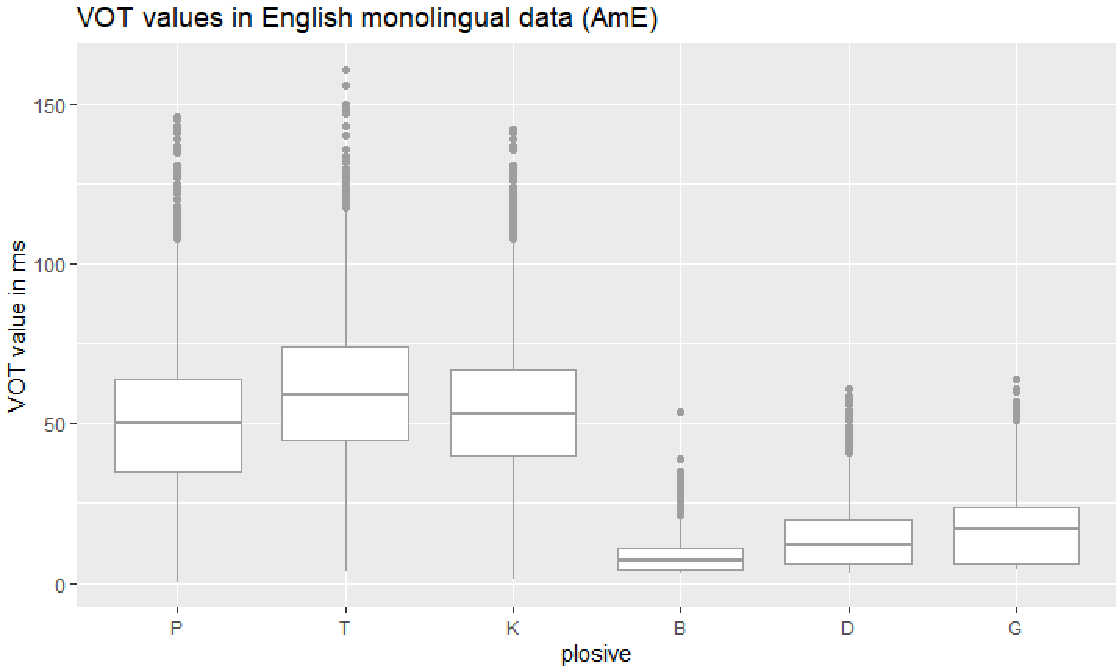
Boxplot of VOT values in English monolingual data (SSBE)



Source: Own presentation

Figure 11

Boxplot of VOT values in English monolingual data (AmE)



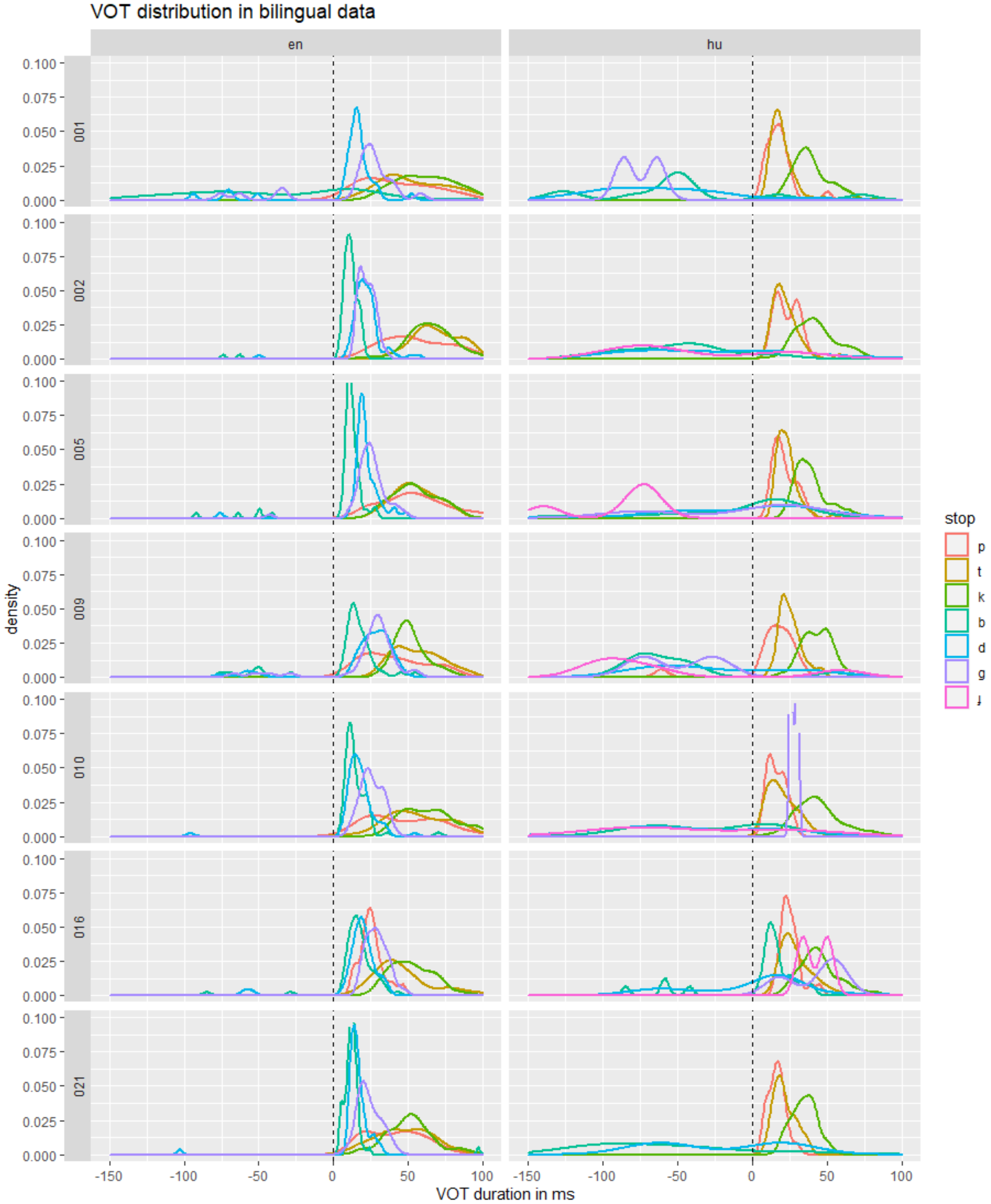
Source: Own presentation

VOT in bilingual Hungarian speakers

Overall, between $n = 250$ and $n = 483$ observations were analysed for each plosive in English. Between $n = 21$ (for ɟ, as it is only the 8th most common consonant in Hungarian (Siptár, 2015)) and $n = 701$ tokens were analysed for Hungarian bilingual speech. Figure 12 shows the distribution of VOT values in the bilingual data by participant, both for English and Hungarian.

Figure 12

VOT distribution in bilingual data



Source: Own presentation

Table 9 shows the VOT means and standard deviations of the bilingual speakers in the dataset.

Table 9

VOT data (in ms) for bilingual data

		Hungarian		English	
		mean	standard deviation	mean	standard deviation
/p/		19.9	3.51	45.0	11.3
/t/		23.3	3.25	55.8	8.67
/k/		41.2	3.61	57.8	6.57
/b/	*(pos)*	27.6	17.5	15.6	3.39
/d/	*(pos)*	27.2	7.86	20.8	4.20
/g/	*(pos)*	30.5	10.5	26.7	2.92
/ɟ/	*(pos)*	37.1	16.9		
/b/	*(neg)*	–68.2	10.7	–64.9	12.6
/d/	*(neg)*	–60.9	8.15	–76.8	21.3
/g/	*(neg)*	–57.5	12.1	–47.7	4.51
/ɟ/	*(neg)*	–74.1	18.1		

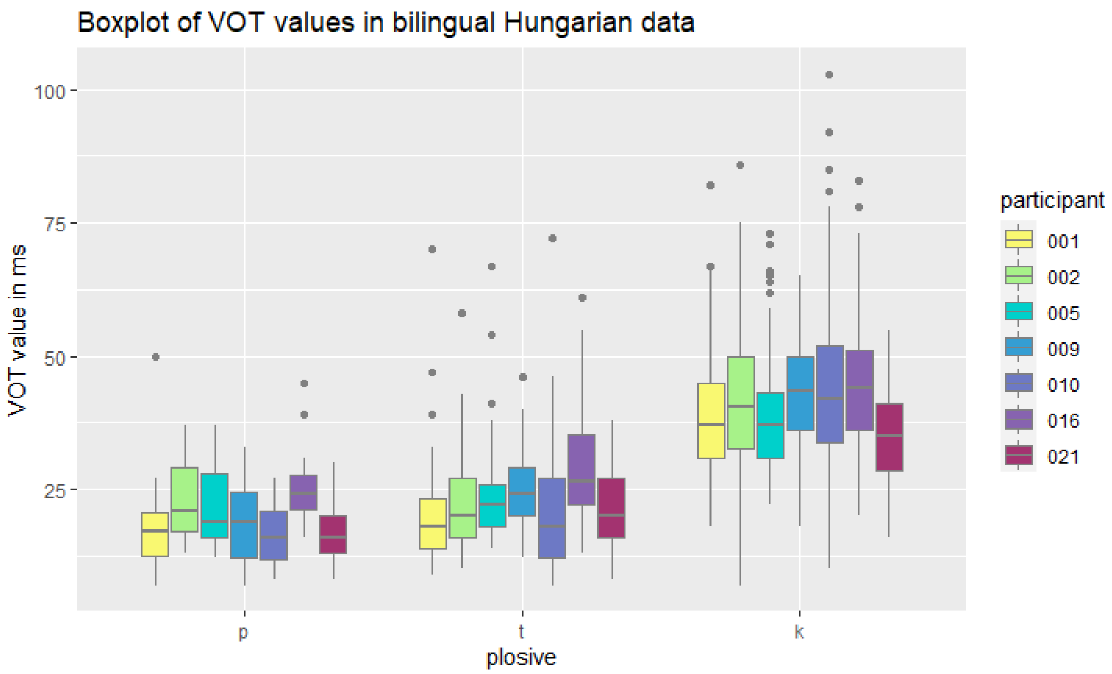
Source: Own presentation

The VOT values of the bilingual speakers are similar to those of the monolingual speakers, no mean or standard deviation stands out.

Figure 13 shows boxplots of the VOT values of voiceless plosives in Hungarian, grouped by participant.

Figure 13

Boxplot of VOT values in bilingual Hungarian data



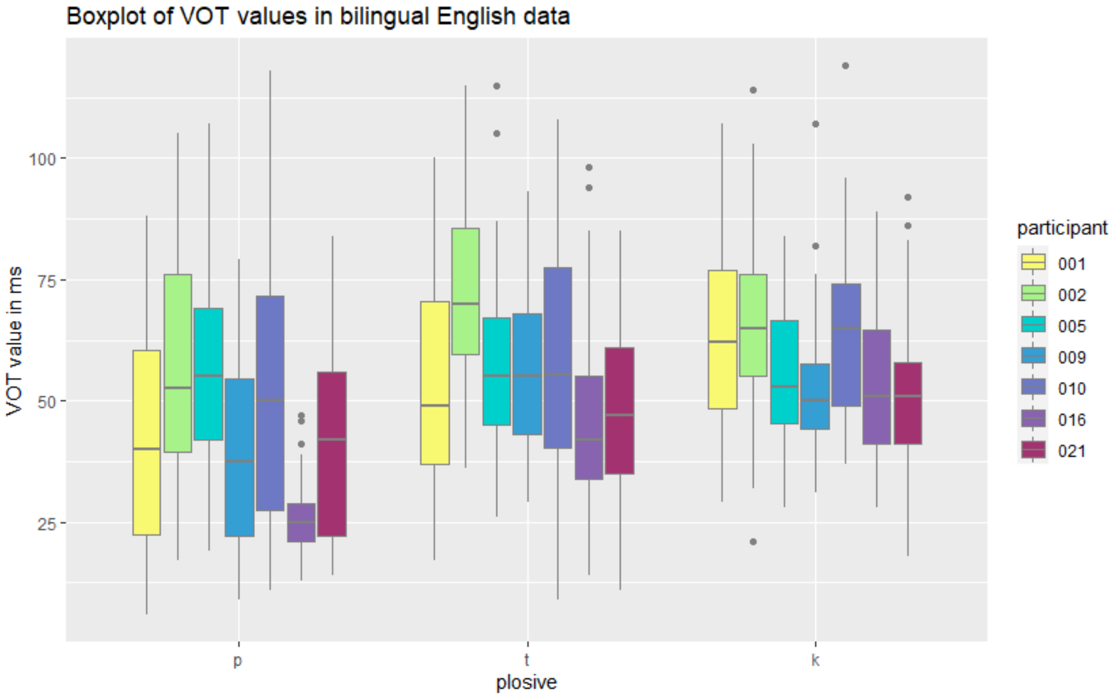
Source: Own presentation

It can be noted that no particular speaker stands out and that they all move in similar ranges of VOT for each plosive.

Figure 14 depicts boxplots of the VOT values of voiceless plosives in English, grouped by participant.

Figure 14

Boxplot of VOT values in bilingual English data



Source: Own presentation

Apart from the realisation of /p/ in participant 016, again, no realisation stands out. Nevertheless, higher variability can be seen compared to the speakers' own Hungarian data, and to both sets of monolingual English data. It is especially visible that the bilingual speakers adjust their VOT but have more variation than the monolingual English speakers.

In Table 10, the percentages of the categorical realisations of voiced plosives in bilingual speakers in Hungarian and English are shown.

Table 10

Frequency of negative and positive VOT for voiced plosives in bilingual speakers (n in brackets)

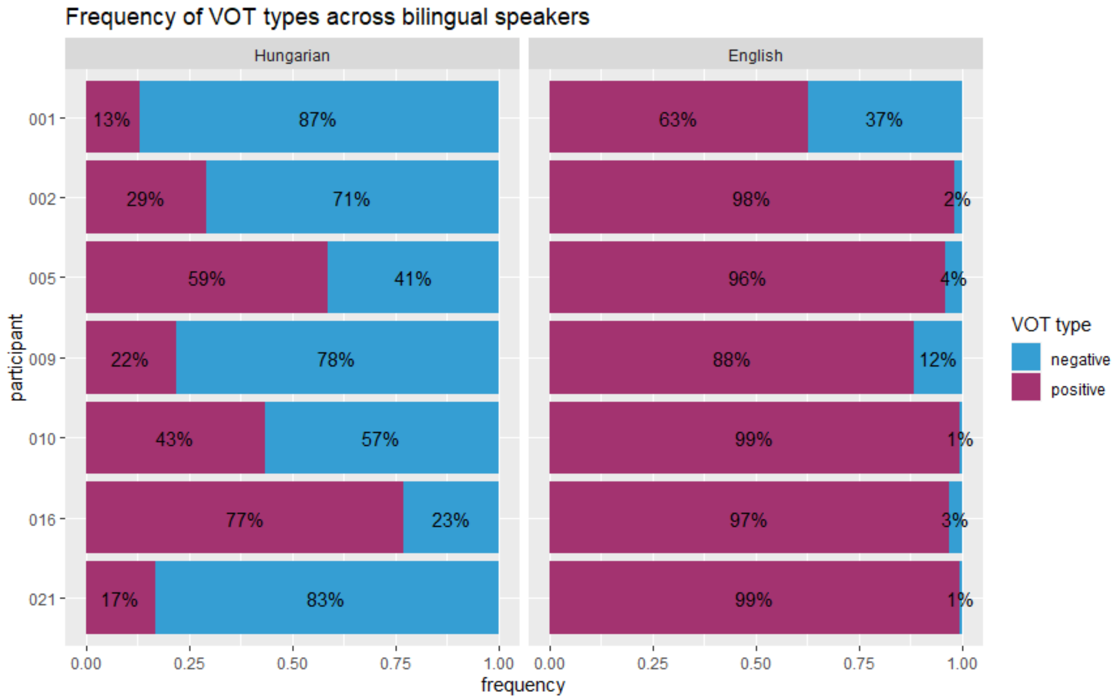
	Hungarian		**English**	
	negative VOT	**positive VOT**	**negative VOT**	**positive VOT**
001	87.1% *(27)*	12.9% *(4)*	37,3% *(47)*	62.7% *(79)*
002	70.7% *(29)*	29.3% *(12)*	1.7% *(3)*	98.3% *(173)*
005	42.2% *(19)*	57.8% *(27)*	4% *(7)*	96% *(169)*
009	78.1% *(25)*	21.8% *(7)*	11.6% *(14)*	88.4% *(107)*
010	56.7% *(21)*	43.3% *(16)*	0.6% *(1)*	99.4% *(151)*
016	22.9% *(11)*	77.1% *(37)*	3% *(4)*	97% *(126)*
021	83.3% *(10)*	16.7% *(2)*	0.6% *(1)*	99.4% *(153)*

Source: Own presentation

Figure 15 visualises the information laid out in Table 10 for ease of understanding.

Figure 15

Frequency of negative and positive VOT in bilingual speakers



Source: Own presentation

According to the literature, it would be assumed that participants with higher nativelikeness would perform better than those with an audible accent, however, this was not the case. While nativelike participants (002, 021) performed well, some non-nativelike participants (010, 016) performed well too. This means that global foreign accent is not necessarily correlated to VOT values and realisations.

5.3 Statistical Results

The Kruskal-Wallis test, which was used to test whether there was a significant difference between spontaneous and read speech, resulted in a significance of $p < 0.0005$ for all plosives. Therefore, only the read data was compared with read data, while spontaneous data was compared to spontaneous data.

A significant difference between monolingual and bilingual English speech was found in /p/ for SSBE ($p < 0.005$) and in /p, k/ for AmE ($p < 0.05$ and $p < 0.005$, respectively). A significant difference was found in Hungarian monolingual and bilingual data in spontaneous speech for /k/ ($p = 1.752e - 07$). In read speech, significant difference was only found for /p/ ($p = 7.274e - 05$).

When testing if the VOT values of people who were residing in an English-speaking country were checked against those that were not, the results were significant for /p/ at $p = 5.964e - 05$ and /t/ at $p < 0.005$, but not for /k/ at $p = 0.2217$.

5.4 Likelihood Ratio Results

When calculating system performance, the lower the EER percentage, the better the performance of the system. Concerning the C_{llr}, the closer to one it is, the worse the system performance is. A value of 1 or over indicates poor performance.

In Table 11, the EERs and C_{llr}s of the system performance test when calculating LRs are summarised.

Table 11

System performance statistics (EERs in %) in different circumstances (Monolingual = ML, Bilingual = BL; Hungarian = HU, English = ENG)

	ML HU		ML ENG		BL HU with ML HU reference population		BL ENG with ML ENG reference population		BL HU		BL ENG		BL (language-mismatch)	
	EER	C_{llr}	EER	C_{llr}	EER	C_{llr}	EER	C_{llr}	EER	C_{llr}	EER	C_{llr}	EER	C_{llr}
/p/	42.2	0.92	29.74	0.77	42.86	0.99	28.57	1.64	42.85	0.79	28.57	1.08	57.14	1.04
/t/	53.11	0.97	34.60	0.76	57.14	1.57	28.57	1.02	57.14	1.06	30.95	1.20	69.04	1.19
/k/	31.81	0.85	30.26	0.89	40.47	0.97	40.47	3.35	44.47	1.46	42.85	0.86	45.23	1.01

Source: Own presentation

6. Discussion

Due to the limited scope and timeframe for the dissertation, not all data that was collected and that was meant to be analysed could be analysed. Therefore, as the results reflect, the outcomes are limited. Nevertheless, some conclusions can be drawn which are presented in the following section.

6.1 Monolingual Results

Overall, the monolingual results conformed to the results of previous literature outlined in Tables 3 and 4, both for English and Hungarian.

6.2 Bilingual Results

The first hypothesis (H_1) was that bilingual productions of English have different VOT compared to monolinguals. This hypothesis was partially confirmed for /p/ (SSBE & AmE) and /k/ (AmE). Interestingly, the different monolingual populations have an effect on the potential difference of the populations. This may be due to Hungarians having a more British-like accent, or due to any other reason outlined in 2.1, such as language attitude or AOA. There were also significant differences found compared to monolingual Hungarian speakers in the cases of /k/ in spontaneous speech, and /p/ for read speech.

The realisations of the voiced plosives also provided some insights, especially looking at possible language transfer or even attrition. Most participants were able to keep their VOT values apart for the two languages, while participants 005, 009, and 016, were not. This is especially interesting as participants 009 and 016 have not spent any time in an English-speaking country, nevertheless, their realisations of /b, d, g/ in English are mostly short-lag. Moreover, the English productions of participant 021, who lives in the United Kingdom, were similar to those of participant 010, who has never lived in an English-speaking country. At the same time, realisations of participant 021 in Hungarian are closer to the targets than those of participant 010, despite participant 010 living in Hungary. This might indicate an individual realisation or be the product of some larger language shift, although this cannot be inferred from the data available.

The second hypothesis (H_2) assumed that people who have spent time in an English-speaking country perform better at reaching target VOT values. This was found to be the case for /p/ and /t/, but not for /k/.

Generally, it is difficult to ascertain whether the significant variation actually stemmed from the tested conditions, or if other reasons also played a role, such as AOA, aptitude, attitude, (and more outlined in 2.1), etc. Due to the vast possibilities for variation in language skills and VOT in bilinguals, it is hard to verify what the actual influence on the VOT values was. The only thing that can be said with certainty is that in some cases VOT are different between monolinguals and bilinguals and also between different groups of bilinguals. It must also be considered that the sample sizes are quite limited in the bilingual data, due to this it is difficult to draw any significant conclusions. Therefore, the results for H_1 and H_2 are inconclusive. While the alternative hypothesis cannot be accepted or rejected, the null hypothesis cannot be rejected either, since the results were inconclusive as they varied from plosive to plosive.

The third hypothesis (H_3) stated that bilingual data would lead to better system performance in an LR framework. This was not the case. Already regarding the Hungarian monolingual data, the system performance was not particularly good. The SSBE data performed marginally better. Using monolingual data as a reference population for bilinguals yielded worse results, with some C_{llr}s as high as 3, and EERs reaching over 50%, indicating poor system performance. The same was true when comparing bilingual data within a language, with English non-native data performing slightly better than the native Hungarian data. In language-mismatch conditions for bilinguals, the results also did not perform better, with EER over 65% and C_{llr} above 1 for each plosive. This conformed to the findings of previous literature. This is potentially due to the limited amount of data available but could also be due to the nature of bilingual speech; thus, the study should be repeated on a larger scale too. Therefore, the hypothesis that the LR framework performs better and that VOT in bilinguals is a better speaker-discriminant than monolingual VOT, has been falsified.

7. Applications in Forensic Speech Science

Due to the lack of conclusive results, it is difficult to suggest any real-world forensic applications of the study results.

The most useful aspect of VOT in FSS might just be the frequency specific individual realisations of VOT that are unexpected in the given language. The data collected concerning the frequency of different realisations, which were excluded from the study, might also prove useful if evaluated further.

It might be interesting to examine someone's VOT just to determine whether phonetically voiced phonemes are actually produced as such or if they are produced with short-lag VOT. This might be valuable information to determine if the speaker is a native speaker, although not in all cases. Nevertheless, it may be a criterion for exclusion in speaker comparison, although not across languages. It could also be used as a speaker profiling (or LADO) clue, although in light of the results, any such inference should be made with caution. Overall, the study proves that any speaker profiling should be treated carefully as the differences between monolinguals and bilinguals are not as clear-cut as expected.

As for the LR framework, the study has shown that bilinguals are not a better, and not even a good subject for the LR framework.

Finally, while the VOT values themselves might not be very helpful in FSS, the categorical variation in their realisations might be. Further research is needed into the topic.

8. Proposals

First and foremost, a longitudinal study is necessary to adequately cover how VOT might change as a function of language learning, exposure to the L2, and time spent in an English-speaking country. Furthermore, the study should be carried out in conjunction with testing the VOT perceptions of the participants to examine whether a similar shift as in the VOT is found in perception. Additionally, the VOT of the bilingual speakers should be perceptually tested by native and bilingual speakers to examine in how far bilingual VOTs affect perception; this was proposed originally but due to the limited scope of the dissertation it had to be discarded, although a follow-up study is warranted. In addition, a reverse study, preferably longitudinal, should be conducted with native English speakers who speak Hungarian to examine if the same effects can be observed.

The current data also warrants further examination as not all of the data could be processed. This calls for (re-)evaluation of all data, taking into account other factors that were not examined in this study as well, not just time spent in an English-speaking country, but also age of acquisition and other languages spoken. Furthermore, additional examination of HS and bilinguals who speak more than two languages is also suggested. The language attrition angle of the present study also warrants in depth examination. Finally, an interesting further study would be how bilingualism affects accent imitation, especially in nativelike speakers of English (or other languages).

9. Conclusion

The aim of the study was to examine if Hungarian-English bilinguals' VOT measurements differ from those of monolinguals, both in English and in Hungarian, what factors might influence this, and how this might affect the performance of an LR-based system. For the purpose of this, data was obtained, collected, and analysed. Unfortunately, no robust results were reached, which means that only limited applications can be proposed at the time. Nevertheless, some use-cases in the field of FSS were suggested. It must be mentioned that the hypotheses were very optimistic, therefore it is not surprising to see inconclusive or falsified results.

As outlined in the proposals, more research is necessary to determine how VOT, and the categorical realisations of plosives, might be useful in FSS and the broader fields of language acquisition and bilingualism. A longitudinal study is certainly necessary to examine how one's VOT might change when moving to an English-speaking country.

Bibliography

Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition, 30*(4), 481–509. *Cambridge Core*, doi: 10.1017/S027226310808073X

Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning, 59*, 249–306. *Wiley Online Library*, doi: 10.1111/j.1467-9922.2009.00507.x

Abramson, A. S., & Lisker, L. (1973). Voice-timing perception in Spanish word-initial stops. *Journal of Phonetics, 1*(1), 1–8. *ScienceDirect*, doi: 10.1016/S0095-4470(19)31372-5

Abramson, A. S., & Whalen, D. (2017). Voice onset time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of Phonetics, 63*, 75–86. *ScienceDirect*, doi: 10.1016/j.wocn.2017.05.002

Adank, P., Hagoort, P., & Bekkering, H. (2010). Imitation improves language comprehension. *Psychological Science, 21*(12), 1903–1909. *SAGE Journals*, doi: 10.1177/0956797610389192

Ahn, S. (2018). *The role of tongue position in voicing contrasts in cross-linguistic contexts*. [PhD dissertation, New York University]. *ProQuest*.

Ahn, E., & Chodroff, E. (2022). *VoxCommunis Corpus*. Available from https://osf.io/t957v

Akbacak, M., & Hansen, J. H. L. (2007). Language Normalization for Bilingual Speaker Recognition Systems, In *Proceedings of 2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07 IV*, 257–60. *IEEEXplore*, doi: 10.1109/ICASSP.2007.366898

Alba-Salas, J. (2004). Voice onset time and foreign accent detection: are L2 learners better than monolinguals? *Alicante Journal of English Studies, 17*, 6–58. *Universidad de Alicante*, doi: 10.14198/raei.2004.17.02

Albert, M. L., & Obler., L. K. (1978). *The bilingual brain: Neuropsychological and neurolinguistic aspects of bilingualism*. New York: Academic Press.

Allen, G. (1985). How the young French child avoids the pre-voicing problem for word-initial voiced stops. *Journal of Child Language, 12*(1), 37–46. *Cambridge Core*, doi: 10.1017/S0305000900006218

Allen, J. S., Miller, J. L., & DeSteno, D. (2003). Individual talker differences in voice-onset-time. *The Journal of the Acoustical Society of America, 113*, 544–552. *Scitation*, doi: 10.1121/1.1528172

Amengual, M. (2021). Exploring cross-linguistic effects and phonetic interactions in the context of bilingualism: Introducing the special issue. *Languages*, *6*(1), 54. *MDPI*, doi: 10.3390/languages6010054

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning*, *42*, 529–555. *Wiley Online Library*, doi: 10.1111/j.1467-1770.1992.tb01043.x

Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2011). Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching. *Journal of Phonetics, 39*(4), 558–570. *ScienceDirect,* doi: 10.1016/j.wocn.2011.03.001

Antoniou, M., Liang, E., Ettlinger, M., & Wong, P.C.M. (2014). The bilingual advantage in phonetic learning. *Bilingualism: Language and Cognition, 18*(4), 683–695. *Cambridge Core*, doi: 10.1017/S1366728914000777

Aoyama, K., Flege, J., Guion, S., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning. *Journal of Phonetics, 32*(2), 233–250. *ScienceDirect*, doi: 10.1016/S0095-4470(03)00036-6

Appel, R., & Muysken, P. (1987). *Language contact and bilingualism*. Baltimore, MD: Edward Arnold.

Archibald, J. (2009). Second Language Phonology. In W. C. Ritchie, & T. K. Bhatia (Eds.), *The new handbook of second language acquisition* (2nd ed.) (pp. 237–258). Bingley: Emerald Group.

Armon-Lotem, S., & Ohana, O. (2017). A CDI study of bilingual English-Hebrew children – frequency of exposure as a major source of variation. *International Journal of Bilingual Education and Bilingualism, 20*(2), 201–217. *Taylor & Francis Online*, doi: 10.1080/13670050.2016.1179257

Asher, J. J., & García, R. (1969). The optimal age to learn a foreign language. *The Modern Language Journal*, *53*, 334–341. *Wiley Online Library*, doi: 10.2307/323026

Association of Forensic Science Providers. (2009). Standards for the formulation of evaluative forensic science expert opinion. *Science and Justice, 49*(3): 161–164. *ScienceDirect*, doi: 10.1016/j.scijus.2009.07.004

Au, T. K., Knightly, L. M., Jun, S. A., & Oh, J. S. (2002). Overhearing a language during childhood. *Psychological Science, 13*(3), 238–243. *SAGE Journals*, doi: 10.1111/1467-9280.00444

Audacity Team. (2012). *Audacity®* [Computer program version 3.0.2]. Available from http://audacityteam.org/

Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics, 40*(1), 177–189. *ScienceDirect,* doi: 10.1016/j.wocn.2011.09.001

Babel, M., McGuire, G., Walters, S., & Nicholls, A. (2014). Novelty and social preference in phonetic accommodation. *Laboratory Phonology, 5*(1), 123–150. *De Gruyter*, doi: 10.1515/lp-2014-0006

Baese-Berk, M. M., & Goldrick, M. (2009). Mechanisms of interaction in speech production. *Language and Cognitive Processes, 24*(4), 527–554. *Taylor & Francis Online*, doi: 10.1080/01690960802299378

Baetens Beardsmore, H. (1986). Who's afraid of bilingualism? In J.-M. Dewaele, A. Housen, & W. Li (Eds.), *Bilingualism: Beyond basic principles* (pp. 10–27). Bristol: Multilingual Matters.

Bailey, P. J., & Haggard, M. P. (1973). Perception and production: some correlations on voicing of an initial stop. *Language and Speech, 16*(3), 189–195. *SAGE Journals*, doi: 10.1177/002383097301600301

Baker, C. (1988). *Key issues in bilingualism and bilingual education*. Bristol: Multilingual Matters.

Baker, C., & Jones, S. P. (1998). *Encyclopedia of bilingualism and bilingual education*. Bristol: Multilingual Matters.

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

Baker, C., & Wright, W. E. (2021). *Foundations of bilingual education and bilingualism* (7th ed.). Bristol: Multilingual Matters.

Baker, W., Trofimovich, P., Flege, J. E., Mack, M., & Halter, R. (2008). Child-adult differences in second-language phonological learning: The role of cross-language similarity. *Language and Speech, 51*(4), 316–341. *SAGE Journals*, doi: 10.1177/0023830908099068

Baldwin, J., & French, J. P. (1990). *Forensic phonetics*. London: Pinter.

Balukas, C., & Koops, C. (2015). Voice onset time in Spanish-English spontaneous code-switching. *International Journal of Bilingualism 19*(4), 423–443. *SAGE Journals*, doi: 10.1177/1367006913516035

Bárkányi, Zs. (2018). The acquisition of voicing assimilation by advanced Hungarian learners of Spanish. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics, 31*(1), 1–31. *John Benjamins e-Platform*, doi: 10.1075/resla.17016.bar

Bárkányi, Zs., & Kiss, Z. (2019). A fonetikai korrelátumok szerepe a zöngekontraszt fenntartásában – beszédprodukciós és észleléses eredmények [The role of phonetic correlates in the maintenance of voicing distinction – speech production and perception results]. *Általános nyelvészeti tanulmányok* [General linguistic studies], *31*, 57–102.

Bartha, Cs. (1996). A társadalmi kétnyelvűség típusai és főbb vizsgálati kérdései [The types of societal bilingualism and its main questions]. *Magyar Nyelvőr* [Hungarian Language Guardian], *120*, 263–282.

Bartha, Cs. (2002). Nyelvhasználat, nyelvmegtartás, nyelvcsere amerikai magyar közösségekben [Language use, language maintenance, language shift in Hungarian communities in America]. In Kovács, N., & Szarka, L. (Eds.), *Tér és terep, Az MTA kisebbségkutató intézetének évkönyve (1)* [Territory and field, Yearbook of the research institute for minority studies of the Hungarian Academy of Sciences] (pp. 111–136). Budapest: Akadémiai Kiadó.

Baugh, J. (2007). Attitudes toward variation and ear-witness testimony. In R. Bayley, & C. Lucas (Eds.), *Sociolinguistic variation: Theories, methods, and applications* (pp. 338–348). Cambridge: Cambridge University Press. *Cambridge Core*, doi: 10.1017/CBO9780511619496.018.

Beckman, J., Helgason, P., McMurray, B., & Ringen, C. (2011). Rate effects on Swedish VOT: Evidence for phonological overspecification. *Journal of Phonetics*, *39*(1), 39–49. *ScienceDirect*, doi: 10.1016/j.wocn.2010.11.001

Beckman, J., Jessen, M., & Ringen, C. (2013). Empirical evidence for laryngeal features: Aspirating vs. true voice languages. *Journal of Linguistics*, *49*(2), 259–284. *Cambridge Core*, doi: 10.1017/S0022226712000424

Benmamoun, E., Montrul, S., & Polinsky, M. (2013). Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics, 39*(3–4), 129–181. *De Gruyter*, doi: 10.1515/tl-2013-0009

Bergmann, C., Nota, A., Sprenger, S. A., & Schmid, M. S. (2016). L2 immersion causes non-native-like L1 pronunciation in German attriters. *Journal of Phonetics, 58*, 71–86. *ScienceDirect*, doi: 10.1016/j.wocn.2016.07.001

Bhatia, T. K., & Ritchie, W. C. (1996). Language mixing, universal grammar, and second language acquisition. In W. C. Ritchie, & T. K. Bhatia (Eds.), *The new handbook of second language acquisition* (2nd ed.) (pp. 592–622). Bingley: Emerald Group.

Bhatia, T. K., & Ritchie, W. C. (2013). Multilingualism and forensic linguistics. In T. K. Bhatia, & W. C. Ritchie (Eds.), *The handbook of bilingualism and multilingualism* (pp. 671–699). Oxford: Wiley-Blackwell.

Bhela, B. (1999). Native language interference in learning a second language. *International Education Journal 1*(1), 22–31.

Bialystok, E. (1997). The structure of age: In search of barriers to second-language acquisition. *Second Language Research, 13*(2), 116–137. *SAGE Journals*, doi: 10.1191/026765897677670241

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, & cognition*. Cambridge: Cambridge University Press.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12*(1), 3–11. *Cambridge Core*, doi: 10.1017/S1366728908003477

Birdsong, D. (1999). Introduction: Whys and why nots of the critical period hypothesis for second language acquisition. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 101–132). Mahwah, NJ: Erlbaum.

Bobda, S. A., Wolf, H.-G., & Peter, L. (1999). Identifying regional and national origin of English-speaking Africans seeking asylum in Germany. *The International Journal of Speech, Language and the Law 6*(2), 301–319. *Equinox*, doi: 10.1558/ijsll.v6i2.300

Boersma, P., & Weenink, D. (2022). *Praat: doing phonetics by computer* [Computer program version 6.2.18]. Available from http://www.praat.org/

Bolonyai, A. (1998). In-between languages: Language shift/maintenance in childhood bilingualism. *International Journal of Bilingualism, 2*(1), 21–43. *SAGE Journals*, doi: 10.1177/136700699800200102

Bóna, J. (2011). A [p, t, k] mássalhangzók zöngekezdési ideje idősek és fiatalok spontán beszédében és felolvasásában [Voice onset time of [p, t, k] in spontaneous and read speech of young and old adults]. *Beszédkutatás* [Speech research], *2011*, 61–72.

Bóna, J. (2014). Voice onset time and speakers' age: Data from Hungarian. *Clinical Linguistics & Phonetics, 28*(5), 366–372. *Taylor & Francis Online*, doi: 10.3109/02699206.2013.875593

Bóna, J. (2016). A zöngétlen explozívák zöngéskezdési ideje különböző életkorú felnőttek beszédében [The voice onset time of voiceless plosives in different-aged adults]. *Beszédkutatás* [Speech research], *2016,* 5–20.

Bóna, J., & Auszmann, A. (2014). Voice onset time in language acquisition: Data from Hungarian. In *Proceedings of 10th International Seminar on Speech Production*, 41–44.

Bongaerts, T. (1999). Ultimate attainment in L2 pronunciation: The case of very advanced L2 learners. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 133–159). Mahwah, NJ: Erlbaum.

Bongaerts, T., Van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition, 19*(4), 447–465. *Cambridge Core*, doi: 10.1017/S0272263197004026

Bradlow, A. R., Kim, M., & Blasingame, M. (2017). Language-independent talker-specificity in first-language and second-language speech production by bilingual talkers: L1 speaking rate predicts L2 speaking rate. *The Journal of the Acoustical Society of America, 141*(2), 886–899. *Scitation*, doi: 10.1121/1.4976044

Brandschain, L., Graff, D., Cieri, C., Walker, K., Caruso, C., & Neely, A. (2010). Greybeard Longitudinal Speech Study. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*. Valletta, Malta. European Language Resources Association (ELRA).

Bright, J. A., & McGregor, G. P. (1970). *Teaching English as a second language: Theory and techniques for the secondary stage*. London: Longman.

Brown, E. K., & Copple, M. T. (2018). Constructing two phonological systems: A phonetic analysis of /p/, /t/, /k/ among early Spanish–English bilingual speakers. *International Journal of Bilingualism, 22*(1), 51–68. *SAGE Journals*, doi: 10.1177/1367006916651983

Bull, R., & Clifford, B. R. (1984). Earwitness voice recognition accuracy. In G. L. Wells, & E. F. Loftus (Eds.), *Eyewitness testimony: Psychological Perspectives* (pp. 92–124). Cambridge: Cambridge University Press.

Bullock, B. E., Toribio, A. J., González, V., & Dalola, A. (2006). Language Dominance and Performance Outcomes in Bilingual Pronunciation. In M. Granthem O’Brien, C. Shea, & J. Archibald. (Eds.), *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA)* (pp. 9–16). Somerville, MA: Cascadilla Proceedings Project.

Bylund, E., Abrahamsson, N., & Hyltenstam, K. (2010). The role of language aptitude in first language attrition: The case of pre-pubescent attriters. *Applied Linguistics, 31*(3), 433–464. *Oxford Academic*, doi: 10.1093/applin/amp059

Byrd, D. (1993). 54,000 American stops. *UCLA Working Papers in Phonetics, 83*, 97–116.

Carroll, S. E. (2017). Explaining bilingual learning outcomes in terms of exposure and input. *Bilingualism*, *20*(1), 37–41. *Cambridge Core*, doi: 10.1017/S1366728916000511

Champod, C., & Evett, I. W. (2000). Commentary on A. P. A. Broeders (1999) 'Some observations on the use of probability scales in forensic identification', *The International Journal of Speech, Language, and the Law*, 6(2), 228–241. *Equinox*, doi: 10.1558/ijsll.v7i2.239

Champod, C., & Meuwly, D. (2000). The inference of identity in forensic speaker recognition. *Speech Communication, 31*(2), 193-203. *ScienceDirect*, doi: 10.1016/S0167-6393(99)00078-3

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

Chang, C. B. (2019). Phonetic drift. In M. S. Schmid, & B. Köpke (Eds.), *The Oxford handbook of language attrition* (pp. 190–203). Oxford: Oxford University Press.

Cheung, W. H. Y., & Wee, L.-H. (2008). Viability of VOT as a parameter for speaker identification: evidence from Hong Kong. In *Current Issues in the Unity and Diversity of Languages: Collection of the papers selected from the CIL 18 held at Korea University in Seoul, on 21-26 July, 2008*. Seoul: The Linguistic Society of Korea.

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics, 27*(2), 207–229. *ScienceDirect*, doi: 10.1006/jpho.1999.0094

Cho, T., Whalen, D. H., & Docherty, G. (2019). Voice onset time and beyond: Exploring laryngeal contrast in 19 languages. *Journal of Phonetics*, *72*, 52–65. *ScienceDirect*, doi:10.1016/j.wocn.2018.11.002

Chodroff, E., & Baese-Berk, M. (2019). Constraints on variability in the voice onset time of L2 English stop consonants. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences. International Congress of Phonetic Sciences, Melbourne, Australia 2019* (pp. 661–665). Canberra: Australasian Speech Science and Technology Association Inc.

Chodroff, E., & Foulkes, P. (forthcoming). Sociophonetic variation in stops. To appear in C. Strelluf (Ed.) *The Routledge handbook of sociophonetics*. Abingdon: Routledge.

Chodroff, E., & Wilson, C. (2017). Structure in talker-specific phonetic realization: Covariation of stop consonant VOT in American English. *Journal of Phonetics, 61*, 30–47. *ScienceDirect*, doi: 10.1016/j.wocn.2017.01.001

Chodroff, E., & Wilson, C. (2018). Predictability of stop consonant phonetics across talkers: Between-category and within-category dependencies among cues for place and voice. *Linguistics Vanguard 4* (s2). Available from https://osf.io/jt5mc/

Chodroff, E., Godfrey, J. J., Khudanpur, S., & Wilson, C. (2015). Structured variability in acoustic realization: A corpus study of voice onset time in American English stops. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Conference of Phonetic Sciences* (pp. 1–5). Glasgow: The University of Glasgow.

Clark, J., & Foulkes, P. (2007). Identification of voices in electronically disguised speech. *The International Journal of Speech, Language and the Law*, 14(2), 195–221. *Equinox*, doi: 10.1558/ijsll.v14i2.195.

Coe, S. (2012). *The effect of sample contemporaneity on the outcome of likelihood ratios for forensic voice comparison*. [Unpublished Master's Dissertation, University of York].

Corder, S. P. (1992). A role for the mother tongue. In S. Gass, & L. Selinker (Eds.), *Language transfer in language learning* (pp. 85–97). New York, NY: Newbury House Publishers.

Core, C., & Scarpelli, C. (2015). Phonological development in young bilinguals: clinical implications. *Seminars in Speech and Language*, *36*(2), 100–108. *Thieme*, doi: 10.1055/s-0035-1549105

Costa, A., Pannunzi, M., Deco, G., & Pickering, M. J. (2017). Do bilinguals automatically activate their native language when they are not using it? *Cognitive Science*, *41*(6), 1629–1644. *Wiley Online Library*, doi: 10.1111/cogs.12434

Coulthard, M., & Johnson, A. (2007). *An introduction to forensic linguistics: Language in evidence*. Abingdon: Routledge. *Taylor & Francis Online*, doi: 10.4324/9781315630311.

Council of Europe. (n.d.). *The CEFR levels*. https://www.coe.int/en/web/common-european-framework-reference-languages/level-descriptions (Accessed 06/09/2022).

Crystal, D. (2003). *English as a global language* (2nd ed.). Cambridge: Cambridge University Press.

Cseresnyési, L. (1992). An outline of Hungarian phonology. *The Journal of Intercultural Studies. Extra Series No. 2*, 79–104.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1989). Limits on bilingualism. *Nature, 340*, 229–230. *Nature*, doi: 10.1038/340229a0

Davidson, L. (2016). Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics*, *54*, 35–50. *ScienceDirect*, doi: 10.1016/j.wocn.2015.09.003

Davidson, L. (2018). Phonation and laryngeal specification in American English voiceless obstruents. *Journal of the International Phonetic Association*, *48*(3), 331–356. *Cambridge Core*, doi: 10.1017/S0025100317000330

Davis, K. (1995). Phonetic and phonological contrasts in the acquisition of voicing: Voice onset time production in Hindi and English. *Journal of Child Language*, *22*(2), 275–305. *Cambridge Core*, doi: 10.1017/S030500090000979X

de Bruin, A. (2019). Not all bilinguals are the same: A call for more detailed assessments and descriptions of bilingual experiences. *Behavioral Sciences*, *9*(3), 33. *MDPI*, doi: 10.3390/bs9030033

De Houwer, A. (1990). *The acquisition of two languages from birth: A case study*. Cambridge: Cambridge University Press.

De Houwer, A. (2007). Parental language input patterns and children’s bilingual use. *Applied Psycholinguistics*, *28*(3), 411–424. *Cambridge Core*, doi: 10.1017/S0142716407070221

De Jong, N. H., Groenhout, R., Schoonen, R., & Hulstijn, J. H. (2013). Second language fluency: Speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics*, *36*(2), 223–243. *Cambridge Core*, doi: 10.1017/S0142716413000210

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, *22*(4), 499–534. *Cambridge Core*, doi: 10.1017/S0272263100004022

de Leeuw, E. (2008). When your native language sounds foreign: A phonetic investigation into first language attrition. [Unpublished PhD dissertation, Queen Margaret University]. *Queen Margaret University Edinburgh*.

de Leeuw, E. (2019). Phonetic attrition. In M. S. Schmid, & B. Köpke (Eds.), *The Oxford Handbook of Language Attrition* (pp. 204–217). Oxford: Oxford University Press.

de Leeuw, E., Schmid, M. S., & Mennen, I. (2010). The effects of contact on native language pronunciation in a migrant context. *Bilingualism: Language and Cognition, 13*(1), 33–40. *Cambridge Core*, doi: 10.1017/S1366728909990289

Deuchar, M., & Clark, A. (1996). Early bilingual acquisition of the voicing contrast in English and Spanish. *Journal of Phonetics, 24*(3), 351–365. *ScienceDirect*, doi: 10.1006/jpho.1996.0019

Deuchar, M., & Quay, S. (2000). *Bilingual acquisition: Theoretical implications of a case study*. Oxford: Oxford University Press.

Dittmers, T., Gabriel, C., Krause, M., & Topal, S. (2018). The production of voiceless stops in multilingual learners of English, French, and Russian: Positive transfer from the heritage languages. In C. Belz, S. Mooshammer, S. Fuchs, S. Jannedy, O. Rasskazova, & M. Zygis (Eds.), *Proceedings of the 13th conference on phonetics and phonology in the German-speaking countries* (pp. 41–44). Berlin: Leibniz-Zentrum Allgemeine Sprachwissenschaft.

Docherty, G. J. (1992). *The timing of voicing in British English obstruents*. Berlin: Foris.

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty, & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589–630). Oxford: Blackwell. *Wiley Online Library*, doi: 10.1002/9780470756492.ch18

Dumas, B. K. (1990). Voice identification in a criminal law context. *American Speech, 65*(4), 341–348. *JSTOR*, doi: 10.2307/455509.

Eades, D. (2011). Forensic linguistics. In P. Hogan (Ed.), *Cambridge encyclopedia of the language sciences* (pp. 313–314). Cambridge: Cambridge University Press.

Earle, F. S., & Arthur, D. T. (2017). Native phonological processing abilities predict post-consolidation nonnative contrast learning in adults. *The Journal of the Acoustical Society of America, 142*(6), EL525–EL531. *Scitation*, doi: 10.1121/1.5013141

Earnshaw, K. (2014). Assessing the discriminatory power of /t/ and /k/ for forensic speaker comparison using a likelihood ratio approach. [Unpublished Master's dissertation, University of York].

Edwards, J. (2006). Foundations of bilingualism. In T. K. Bhatia, & W. C. Ritchie (Eds.), *The handbook of bilingualism* (pp. 7–31) Oxford: Blackwell.

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Ellis, R. (1997). *Second Language Acquisition*. Oxford: Oxford University Press.

Evett, I. W., Scranage, J. K., & Pinchin, R. (1993). An illustration of the advantages of efficient statistical methods for RFLP analysis in forensic science. *American Journal of Human Genetics, 52*(3), 498–505.

Fenton, N. E., Neil, M., & Hsu, A. (2014). Calculating and understanding the value of any type of match evidence when there are potential testing errors. *Artificial Intelligence and Law, 22*, 1–28. *SpringerLink*, doi: 10.1007/s10506-013-9147-x

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*(1), 47–65. *ScienceDirect*, doi: 10.1016/S0095-4470(19)30537-6

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

Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *The Journal of the Acoustical Society of America, 89*(1), 395–411. *Scitation*, doi: 10.1121/1.400473

Flege, J. E. (1992). Speech learning in a second language. In C. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 565–604). Timonium, MD: York Press.

Flege, J. E. (1999). Age of learning and second language speech. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 101–132). Mahwah, NJ: Erlbaum.

Flege, J. E., & Bohn, O.-S. (2021). The revised Speech Learning Model. In R. Wayland (Ed.), *Second language speech learning, Theoretical and empirical progress* (pp. 3–83). Cambridge: Cambridge University Press. *Cambridge Core*, doi: 10.1017/9781108886901.002

Flege, J. E., & Schmidt, A. (1995). Effects of speaking rate changes on native and nonnative speech production. *Phonetica, 52*(1), 41–54. *De Gruyter*, doi: 10.1159/000262028

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1996). Factors affecting the production of word-initial consonants in a second language. In R. Bayley, & D. R. Preston (Eds.), *Second language acquisition and linguistic variation*. Amsterdam: John Benjamins.

Flege, J. E., Frieda, E., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics, 25*(2), 169–186. *ScienceDirect*, doi: 10.1006/jpho.1996.0040

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

Foulkes, P. (2006). Phonological variation: A global perspective. In B. Aarts, & A. McMahon (Eds.), *The handbook of English linguistics* (pp. 625–669). *Wiley Online Library*, doi: 10.1002/9780470753002.ch27.

Foulkes, P., & French, J. P. (2001). Forensic phonetics and sociolinguistics. R. Mesthrie (Ed.), *Concise Encyclopedia of Sociolinguistics* (pp. 329–332). Amsterdam: Elsevier Press.

Foulkes, P., & French, J. P. (2012). Forensic speaker comparison: a linguistic-acoustic perspective. In P. Tiersma, & L. M. Solan (Eds.), *Oxford handbook of Language and the Law* (pp. 557–572). Oxford: Oxford University Press.

Foulkes, P., Docherty, G. J., & Jones, M. J. (2010a). Analysing stops. In M. Di Paolo, & M. Yaeger-Dror (Eds.), *Sociophonetics: A student's guide* (pp. 58–71). Abingdon: Routledge.

Foulkes, P., Scobbie, M., & Watt, D. (2010b). Sociophonetics. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (2nd ed.) (pp. 703–754). Oxford: Blackwell.

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

Francis, A. L., Ciocca, V., & Yu, J. M. C. (2003). Accuracy and variability of acoustic measures of voicing onset. *The Journal of the Acoustical Society of America, 113*(2), pp. 1025–1032. *Scitation*, doi: 10.1121/1.1536169

French, P. (1994). An overview of forensic phonetics with particular reference to speaker identification. *The International Journal of Speech, Language and the Law, 1*(2): 169–181. *Equinoxpub*, doi: 10.1558/ijsll.v1i2.169.

French, P., & Harrison, P. (2007). Position statement concerning use of impressionistic likelihood terms in forensic speaker comparison cases. *The International Journal of Speech, Language and the Law, 14*(1), 137–144. *Equinoxpub,* doi: 10.1558/ijsll.v14i1.137

French, P., & Stevens, L. (2013). Forensic speech science. In M. J. Jones, & R. Knight (Eds.), *Bloomsbury Companion to Phonetics* (pp. 183–197). London: Bloomsbury.

Gabriel, C., & Rusca-Ruths, E. (2015). Der Sprachrhythmus bei deutsch-türkischen L3-Spanischlernern: Positiver Transfer aus der Herkunftssprache? [Speech rhythm in German-Turkish L3-learners: Positive transfer from the native language?] In S. Witzigmann, & J. Rymarczyka (Eds.), *Mehrsprachigkeit als Chance. Herausforderungen und Potentiale individueller und gesellschaftlicher Mehrsprachigkeit* [Multilingualism as an opportunity: Challenges and potentials of individual and societal multilingualism] (pp. 185–204). Frankfurt a. M.: Peter Lang.

Gabriel, C., Kupisch, T., & Seoudy, J. (2016). VOT in French as a foreign language: A production and perception study with mono- and multilingual learners (German/Mandarin-Chinese). *SHS Web of Conferences, 27,* 1–17. *SHS Web Conferences,* doi: 10.1051/shsconf/20162709002

Geiss, M., Gumbsheimer, S., Lloyd-Smith, A., Schmid, S., & Kupisch, T. (2022). Voice onset time in multilingual speakers: Italian heritage speakers in Germany with L3 English. *Studies in Second Language Acquisition, 44*(2), 435–459. *Cambridge Core*, doi: 10.1017/S0272263121000280

Genesee, F. (1989). Early bilingual development: One language or two? *Journal of Child Language, 16*(1), 161–179. *Cambridge Core,* doi: 10.1017/S0305000900013490

Genesee, F., & Nicoladis, E. (2007). Bilingual first language acquisition. In E. Hoff, & M. Shatz (Eds.), *Blackwell handbook of language development* (pp. 324–342). Oxford: Blackwell.

Gnevsheva, K. (2017). Within-speaker variation in passing for a native speaker. *International Journal of Bilingualism, 21*(2), 213–227. *SAGE Journals*, doi: 10.1177/1367006915616197

Gold, E. (2014). *Calculating likelihood ratios for forensic speaker comparisons using phonetic and linguistic parameters*. [PhD dissertation, University of York]. *White Rose eTheses Online*.

Gold, E., & Hughes, V. (2014). Issues and opportunities: The application of the numerical likelihood framework to forensic speaker comparison. *Science and Justice, 54*(4), 292–299. *ScienceDirect*, doi: 10.1016/j.scijus.2014.04.003

Goldrick, M., Runnqvist, E., & Costa, A. (2014). Language switching makes pronunciation less nativelike. *Psychological Science, 25*(4), 1031–1036. *SAGE Journals*, doi: 10.1177/0956797613520014

Google LLC (2022). *Google Python Speech API* [Python API version 1.3.3]. Available from https://github.com/googleapis/python-speech

Gósy, M. (2000). On the early history of Hungarian speech research. *International Journal of Speech Technology, 3*(3), 155–164. *SpringerLink*, doi: 10.1023/A:1026511031853

Gósy, M. (2001). The VOT of the Hungarian voiceless plosives in words and in spontaneous Speech. *International Journal of Speech Technology, 4*(1), 75–85. *SpringerLink*, doi: 10.1023/A:1009608900453

Gósy, M. (2004). *Fonetika, a beszéd tudománya* [Phonetics, the science of speech]. Budapest: Osiris Kiadó.

Gósy, M. (2008). Magyar spontánbeszéd-adatbázis – BEA [Hungarian Spontaneous Speech Database – BEA]. *Beszédkutatás* [Speech research], *2008*, 194–207.

Gósy, M. (2016). Beszédhangok viselkedése a spontán beszédben [The behaviour of speech sounds in spontaneous speech]. In J. Bóna (Ed.), *Fonetikai Olvasókönyv* [Phonetics reader] (pp. 19–31). Budapest: ELTE Fonetikai Tanszék.

Gósy, M., & Gyarmathy, D. (2017). A BEA beszélt nyelvi adatbázis tízéves fejlesztése és kutatási eredményei [The ten-year development and research outcomes of the BEA spoken language database]. *Magyar Nyelvőr* [Hungarian Language Guardian], *141*(4), 70–82.

Gósy, M., & Ringen, C. O. (2009). Everything you always wanted to know about VOT in Hungarian. Presented at: *Ninth International Conference on the Structure of Hungarian*. Budapest.

Gráczi, T. E. (2011). Explozívák a zöngésségi oppozíció függvényében [Plosives in interdependence with voicing opposition]. In T. Váradi (Ed.), *V. Alkalmazott nyelvészeti doktoranduszkonferencia* [5th doctoral candidate conference of applied linguistics] (pp. 51–66). Budapest: MTA Nyelvtudományi Intézet.

Gráczi, T. E. (2016). A zöngéskezdési időről [About voice onset time]. In J. Bóna (Ed.), *Fonetikai olvasókönyv* [Phonetics reader] (pp. 61–74). Budapest: ELTE Fonetikai Tanszék.

Gráczi, T. E., & Kohári, A. (2014). Multiple bursts in Hungarian voiceless plosives and VOT measurements. Presented at: *10th International Seminar on Speech Production (ISSP)*. Cologne.

Grosjean, F. (1982). *Life with two languages*. Cambridge, MA: Harvard University Press.

Grosjean, F. (1985). The bilingual as a competent but specific speaker-hearer. *Journal of Multilingual and Multicultural Development, 6*(6), 467–477. *Taylor & Francis Online*, doi: 10.1080/01434632.1985.9994221

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

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition, 1*(2), 131–149. *Cambridge Core*, doi: 10.1017/S136672899800025X

Grosjean, F. (2001). The bilingual's language modes. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 23–48). Oxford: Blackwell.

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

Grosjean, F. (2010). *Bilingual: Life and Reality*. Cambridge, MA: Harvard University Press.

Grosjean, F., & Miller, J. L. (1994). Going in and out of languages: An example of bilingual flexibility. *Psychological Science, 5*(4), 201–209. *SAGE Journals,* doi: 10.1111/j.1467-9280.1994.tb00501.x

Grosjean, F., & Soares, C. (1986). Processing mixed language: Some preliminary findings. In J. Vaid (Ed.), *Language processing in bilinguals: Psycholinguistic and neuropsychological perspectives* (pp. 145–179). Hillsdale, NJ: Lawrence Erlbaum.

Gut, U. (2009). *Non-native speech: A corpus-based analysis of phonological and phonetic properties of L2 English and German*. Frankfurt a. M.: Peter Lang.

Guven, C., & Islam, A. (2015). Age at migration, language proficiency, and socioeconomic outcomes: Evidence from Australia. *Demography, 52*(2), 513–542. *Duke University Press*, doi: 10.1007/s13524-015-0373-6

Haman, E., Wodniecka, Z., Marecka, M., Szewczyk, J., Białecka-Pikul, M., Otwinowska, A., Mieszkowska, K., Łuniewska, M., Kołak, J., Miekisz, A., Kacprzak, A., Banasik, N., & Forys-Nogala, M. (2017). How does L1 and L2 exposure impact L1 performance in bilingual children? Evidence from Polish-English migrants to the United Kingdom. *Frontiers in Psychology, 8*, 1444. *Frontiers*, doi: 10.3389/fpsyg.2017.01444

Hambly, H., Wren, Y., McLeod, S., & Roulstone, S. (2013). The influence of bilingualism on speech production: a systematic review. *International Journal of Language & Communication Disorders, 48*(1), 1–24. *Wiley Online Library,* doi: 10.1111/j.1460-6984.2012.00178.x

Hardcastle, W. J. (1973). Some observations of the tense-lax distinction in initial stops in Korean. *Journal of Phonetics, 1*(3), 263–271. *ScienceDirect,* doi: 10.1016/S0095-4470(19)31390-7

Harley, B., & Hart, D. (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition, 1*, 379–400. *Cambridge Core,* doi: 10.1017/S0272263197003045

Hatoss, A. (2004). A magyar nyelv és kultúra Ausztráliában. Felmérés a queenslandi nyelvközösség nyelvmegtartásáról [Hungarian language and culture in Australia. Study about the language maintenance of the Hungarian speech community in Queensland]. *Hungarológia Évkönyv* [Hungarology Yearbook], *5*(1), 201-215.

Hoff, E., & Core, C. (2013). Input and language development in bilingually developing children. *Seminars in Speech and Language, 34*(4), 215–226. *Thieme,* doi: 10.1055/s-0033-1353448

Hollien, H. (1990). *The acoustic crime: The new science of forensic phonetics*. New York, NY: Plenum.

Hollstein, B. (n.d.) *Tanácsok a kétnyelvű gyerekneveléshez* [Advice for raising bilingual children]. Lingua Hungarica. https://www.lingua-hungarica.de/images/Tanacsok%20Ketnyelvu%20gyermekneveles.pdf (Accessed 27/08/2022)

Holm, A., & Dodd, B. (1999). A longitudinal study of the phonological development of two Cantonese-English bilingual children. *Applied Psycholinguistics*, *20*(3), 349–376. *Cambridge Core,* doi: 10.1017/S0142716499003021

Hopp, H., & Schmid, M. S. (2013). Perceived foreign accent in first language attrition and second language acquisition: The impact of age of acquisition and bilingualism. *Applied Psycholinguistics*, *34*(2), 361–394. *Cambridge Core,* doi: 10.1017/S0142716411000737

Hrycyna, M., Lapinskaya, N., Kochetov, A., & Nagy, N. (2011). VOT drift in 3 generations of heritage language speakers in Toronto. *Canadian Acoustics*, *39*(3), 166–167.

Hughes, V. (2011). The effect of variability on the outcome of numerical likelihood ratios for forensic voice comparison. [Master’s dissertation, University of York]. *White Rose eTheses Online*.

Hughes, V. (2013a). *MVKD_loop_CrossVal.m*. [MATLAB script].

Hughes, V. (2013b). *Calibration_CrossVal.m*. [MATLAB script].

Hughes, V. (2013c). *MVKD_loop.m*. [MATLAB script].

Hughes, V. (2014). The definition of the relevant population and the collection of reference data for likelihood ratio-based forensic voice comparison. [PhD dissertation, University of York]. *White Rose eTheses Online*.

Hughes, V. (2017). Sample size and the multivariate kernel density likelihood ratio: how many speakers are enough? *Speech Communication, 94*, 15–29. *ScienceDirect*, doi: 10.1016/j.specom.2017.08.005

Hughes, V., & Foulkes, P. (2015). The relevant population in forensic voice comparison: Effects of varying delimitations of social class and age. *Speech Communication, 66*, 218–30. *ScienceDirect,* doi: 10.1016/j.specom.2014.10.006

Hughes, V., & Wormald, J. H. (2020). Sharing innovative methods, data and knowledge across sociophonetics and forensic speech science. *Linguistics Vanguard, 6*(s1): 1–9. *De Gruyter*, doi: 10.1515/lingvan-2018-0062

Ingram, D. (1981). The emerging phonological system of an Italian-English bilingual child. *Journal of Italian Linguistics, 2,* 95–113.

Ioup, G., Boustagui, E., El Tigi, M., & Moselle, M. (1994). Reexamining the critical period hypothesis: A case study in a naturalistic environment. *Studies in Second Language Acquisition 16*(1), 73–98.

Ishihara, S., & Kinoshita, Y. (2008). How many do we need? Exploration of the population size effect on the performance of forensic speaker classification. *Proceedings of Interspeech 2008*, 1941–1944. *ISCA*, doi: 10.21437/Interspeech.2008-513

Isurin, L., & Ivanova-Sullivan, T. (2008). Lost in between: The case of Russian heritage speakers. *Heritage Language Journal*, *6*, 72–104.

James, A. (1986) Phonic transfer and phonological explanation: Some theoretical and methodological issues. In E. Kellerman, & M. Sharwood Smith (Eds.), *Crosslinguistic influence in second language acquisition* (pp. 134–149). Oxford: Pergamon.

Jank, I., & Kovács, L. (2020). *Mi az a kétnyelvűség, és miért érdekes nekünk?* [What is bilingualism and why is it interesting to us?] Nyelv és Tudomány [Language and Science]. https://m.nyest.hu/hirek/mi-az-a-ketnyelvuseg-es-miert-erdekes-nekunk (Accessed 27/08/2022).

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

Jessen, M. (1996). The relevance of phonetic reality for underlying phonological representation: the case of tense versus lax obstruents in German. In U. Kleinhenz (Ed.), *Interfaces in phonology* (pp. 294–328). Berlin: Akademie Verlag

Jessen, M. (2007). Speaker classification in forensic phonetics and acoustics. In C. Müller (Ed.), *Speaker classification I: Fundamentals, features, and methods* (pp. 180–204). Berlin: Springer. *SpringerLink*, doi: 10.1007/978-3-540-74200-5_3.

Jessen, M. (2008). Forensic phonetics. *Language and Linguistics Compass, 2*(4), 671–711. *Wiley Online Library*, doi: 10.1111/j.1749-818X.2008.00066.x

Jessen, M. (2010). The Forensic Phonetician: Forensic Speaker Identification by Experts. In M. Coulthard, & A. Johnson (Eds.), *The Routledge handbook of forensic linguistics* (pp. 378–394). Abingdon: Routledge. *Routledge Handbooks Online*, doi: 10.4324/9780203855607.ch25.

Jessen, M. (2021). Speaker profiling and forensic voice comparison: The auditory-acoustic approach. In M. Coulthard, & A. Johnson (Eds.), *The Routledge handbook of forensic linguistics* (2nd ed.) (pp. 378–394). Abingdon: Routledge.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*(1), 60–99. *ScienceDirect*, doi: 10.1016/0010-0285(89)90003-0

Johnson, K. (2003). *Acoustic and auditory phonetics* (2nd ed.). Oxford: Blackwell.

Jordanidisz Á., Auszmann A., & Bóna J. (2015) Voice onset time of the voiceless alveolar and velar stops in bilingual Hungarian-English children and their monolingual Hungarian peers. In E. Babatsouli & D. Ingram (Eds.), *Proceedings of the international symposium on monolingual and bilingual speech 2015* (pp. 105–111). Chania: Institute of Monolingual and Bilingual Speech.

Karlsson, F., Zetterholm, E., & Sullivan, K. P. H. (2004). Development of a gender difference in voice onset time. In *Proceedings of the 10th Australian International Conference on Speech Science and Technology,* 316–321. Canberra.

Kehoe, M., & Kannathasan, K. (2021). Development of voice onset time in monolingual and bilingual French-speaking children. *Lingua, 252,* 102937. *ScienceDirect,* doi: 10.1016/j.lingua.2020.102937

Keshet, J., Sonderegger, M., & Knowles, T. (2014). *AutoVOT: A tool for automatic measurement of voice onset time using discriminative structured prediction* [Computer program version 0.94]. Available from https://github.com/mlml/autovot/

Kessinger, R. H., & Blumstein, S. E. (1997). Effects of speaking rate on voice-onset time in Thai, French, and English. *Journal of Phonetics, 25*(2), 143–168. *ScienceDirect,* doi: 10.1006/jpho.1996.0039

Kleber, F. (2018). VOT or quantity: what matters more for the voicing contrast in German regional varieties? Results from apparent-time analyses. *Journal of Phonetics, 71*, 468–486. *ScienceDirect,* doi: 10.1016/j.wocn.2018.10.004

Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.

Künzel, H. J. (1994). On the problem of speaker identification by victims and witnesses. *The International Journal of Speech, Language and the Law, 1*(1): 45–57. *Equinox*, doi: 10.1558/ijsll.v1i1.45

Kupisch, T., & Rothman, J. (2016). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*, *22*(5), 564–582. *SAGE Journals*, doi: 10.1177/1367006916654355

Kupisch, T., Barton, D., Klaschik, E., Lein, T., Stangen, I., & van de Weijer, J. (2014). Foreign accent in adult simultaneous bilinguals. *The Heritage Language Journal, 11*(2), 123–150. *BRILL*, doi: 10.46538/hlj.11.2.2

Ladefoged, P. (2003). *Phonetic data analysis: An introduction to fieldwork and instrumental techniques*. Oxford: Blackwell.

Lahmann, C., Steinkrauss, R., & Schmid, M. S., (2017). Speed, breakdown, and repair: An investigation of fluency in long-term second-language speakers of English. *International Journal of Bilingualism, 21*(2), 228–242. *SAGE Journals,* doi: 10.1177%2F1367006915613162

Lakshmanan, U. (2009). Child second language acquisition. In W. C. Ritchie, & T. K. Bhatia (Eds.), *The new handbook of second language acquisition* (2nd ed.) (pp. 377–399). Bingley: Emerald Group.

Lavoie, L. (2001). *Consonant strength: Phonological patterns and phonetic manifestations*. New York, NY: Garland.

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. *SAGE Journals*, doi: 10.1177%2F1367006915589424

Lenneberg, E. (1967). *Biological foundations of language*. Chichester: John Wiley & Sons.

Lennes, M. (2002). *Save labeled intervals to wav sound files*. [Praat Script]. Available from https://github.com/lennes/spect/blob/master/scripts/save_labeled_intervals_to_wav_sound_files.praat

Levi, S. V., Winters, S. J., & Pisoni, D. (2007). Speaker-independent factors affecting the perception of foreign accent in a second language. *The Journal of the Acoustical Society of America, 121*(4), 2327–2338. *Scitation*, doi: 10.1121/1.2537345

Lisker, L. (1986). “Voicing” in English: a catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech, 29*(1). *SAGE Journals,* doi: 10.1177%2F002383098602900102

Lisker, L., & Abramson, A. S. (1964a). A cross-language study of voicing in initial stops: acoustical measurements. *Word, 20*(3): 384–423. *Taylor & Francis Online,* doi: 10.1080/00437956.1964.11659830

Lisker, L., & Abramson, A. S. (1964b). Stop categorization and voice onset time. In *Proceedings of the 5th international congress of phonetic sciences*. Münster.

Lisker, L., & Abramson, A. S. (1967). Some effects of context on voice onset time in English stops. *Language and Speech, 10*(1), 1–28. *SAGE Journals,* doi: 10.1177%2F002383096701000101

Lleó, C. (2016). Bilingualism and child phonology. In *Oxford Handbook Topics in Linguistics* (online edition). *Oxford Handbooks Online*. doi: 10.1093/oxfordhb/9780199935345.013.53

Lo, J. J. H. (2021). *Issues of Bilingualism in likelihood ratio-based forensic voice comparison*. [PhD dissertation, University of York]. *White Rose eTheses Online*.

Locke, J. L. (1983). *Phonological acquisition and change*. New York: Academic Press.

Lucy, D. (2005). *Introduction to statistics for forensic scientists*. Chichester: John Wiley & Sons.

Lynch, M., & McNally, R. (2003). ‘Science’, ‘common sense’ and DNA evidence: a legal controversy about the public understanding of science. *Public Understanding of Science*, *12*(1), 83–103. *SAGE Journals*, doi: 10.1177%2F0963662503012001246

MacLeod, A. N., & Stoel-Gammon, C. (2005). Are bilinguals different? What VOT tells us about simultaneous bilinguals. *Journal of Multilingual Communication Disorders, 3*(2), 118–127. *Taylor & Francis Online,* doi: 10.1080/14769670500066313

Magyar vagyok. (2010). *Hofi Géza - Pusszantás mindenkinek (1)*. http://www.magyarvagyok.hu/videok/3-Humor/11325-Hofi-Geza-Pusszantas-Mindenkinek-1.html (Accessed 28/08/2022)

Major, R. (1987). English voiceless stop production by speakers of Brazilian Portuguese. *Journal of Phonetics, 15*(2), 197–202. *ScienceDirect,* doi: 10.1016/S0095-4470(19)30560-1

Mayr, R., & Montanari, S. (2015). Cross-linguistic interaction in trilingual phonological development: The role of the input in the acquisition of the voicing contrast. *Journal* *of Child Language, 42*(5), 1006–1035. *Cambridge Core,* doi: 10.1017/S0305000914000592

McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M., & Sonderegger, M. (2017). Montreal Forced Aligner: Trainable Text-Speech Alignment Using Kaldi. [Computer program version 2.0.6] *Proceedings of Interspeech 2017*, 498–502. doi: 10.21437/Interspeech.2017-1386. Available from https://github.com/MontrealCorpusTools/Montreal-Forced-Aligner

Meisel, J. M. (2006). The Bilingual Child. In T. K. Bhatia, & W.C. Ritchie (Eds.), *The handbook of bilingualism* (pp. 90–113). Oxford: Blackwell.

Mihajlik, P., Balog, A., Gráczi, T. E., Kohári, A., Tarján, B., & Mády, K. (2022). BEA-Base: A benchmark for ASR of spontaneous Hungarian. *Electrical Engineering and Systems Science*. *Cornell University,* doi: 10.48550/arXiv.2202.00601

Miller, J. L., Green, K. P., & Reeves, A. (1986). Speaking rate and segments: A look at the relation between speech production and speech perception for the voicing contrast. *Phonetica, 43*(1-3), 106–115. *De Gruyter,* doi: 10.1159/000261764

Montrul, S. (2008). *Incomplete acquisition in bilingualism: Re-examining the age factor*. Amsterdam: John Benjamins.

Montrul, S. (2010). Current issues in heritage language acquisition. *Annual Review of Applied Linguistics, 30,* 3–23. *Cambridge Core,* doi: 10.1017/S0267190510000103

Montrul, S. (2013). Bilingualism and the heritage language speaker. In T. K. Bhatia, & W. C. Ritchie (Eds.), *The Handbook of Bilingualism and Multilingualism* (pp. 168–190). Oxford: John Wiley & Sons.

Morris, R. J., McCrea, C. R., & Herring, K. (2008). Voice onset time differences between males and females: Isolated syllables. *Journal of Phonetics, 36*(2), 308–317. *ScienceDirect,* doi: 10.1016/j.wocn.2007.06.003

Morrison, G. S. (2010). Forensic Voice comparison. In I. Frekelton, & G. Selby (Eds.), *Expert Evidence* (Ch. 99). Sydney: Thomson Reuters.

Morrison, G. S. (2011a). A comparison of procedures for the calculation of forensic likelihood ratios from acoustic–phonetic data: multivariate kernel density (MVKD) versus Gaussian mixture model-universal background model (GMM-UBM). *Speech Communication, 53*(2), 242–256. *ScienceDirect,* doi: *10.1016/j.specom.2010.09.005*

Morrison, G. S. (2011b). Measuring the validity and reliability of forensic likelihood-ratio systems. *Science & Justice, 51* (3), 91–98. *ScienceDirect,* doi: *10.1016/j.scijus.2011.03.002*

Morrison, G. S. (2012). The likelihood ratio framework and forensic evidence in court: A response to R v T. *The International Journal of Evidence & Proof, 16*(1), 1–29. *SAGE Journals,* doi: 10.1350%2Fijep.2012.16.1.390

Moyer, A. (1999). Ultimate attainment in L2 phonology: The critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition, 21*(1), 81–108. *Cambridge Core,* doi: 10.1017/S0272263199001035

Moyer, A. (2007). Empirical considerations on the age factor in L2 phonology. *Issues in Applied Linguistics, 15*(2), 109–129. *eScholarship*, doi: 10.5070/L4152005087

Muysken, P. (2000). *Bilingual speech: A typology of code-mixing*. Cambridge: Cambridge University Press.

Muysken, P. (2006). Two linguistic systems in contact: Grammar, phonology, and lexicon. In T. K. Bhatia, & W. C. Ritchie (Eds.), *The handbook of bilingualism* (pp. 146–168). Oxford: Blackwell.

Nagy, N., & Kochetov, A. (2013). VOT across the generations: A cross-linguistic study of contact-induced change. In P. Siemund, I. Gogolin, M. Schulz, & J. Davydova (Eds.), *Multilingualism and language contact in urban areas: Acquisition - Development - Teaching - Communication* (pp. 19–38). Amsterdam: John Benjamins.

Nemes, M. (2022) The difficulties of becoming bilingual: Hungarian children in the United Kingdom. *Gyermeknevelés Tudományos Folyóirat* [Scientific Journal of Early Years Education], *10*(2), 231–251. *ELTE*, doi: 10.31074/gyntf.2022.2.231.251

Neuberger, T., & Gráczi, T. E. (2013). Az alveoláris zöngétlen explozíva variabilitása [The variability of the alveolar voiceless plosive]. *Beszédkutatás* [Speech research], *2013*, pp. 160–172.

Neufeld, G. G. (1977). Language learning ability in adults. A study on the acquisition of prosodic and articulatory features. *Working papers on bilingualism, 12*, 46–60.

Neufeld, G. G. (2001). Non-foreign-accented speech in adult second language learners: Does it exist and what does it signify? *ITL Review of Applied Linguistics*, *133*(1), 185–206. *John Benjamins e-Platform,* doi: 10.1075/ITL.133-134.01NEU

Newkline Co., Ltd. (2022). *AVR X – Voice Recorder*. [Mobile iOS software version 2.0.4]. Available from https://apps.apple.com/gb/app/avr-x-voice-recorder/id1344740222

NLL. (2021). *ASR Voice Recorder*. [Mobile Android software version 243 arm64-v8a] Available from https://play.google.com/store/apps/details?id=com.nll.asr&gl=GB

Nolan, F. (1991). Forensic phonetics. *Journal of Linguistics, 27*(2), 487–493. *Cambridge Core,* doi: 10.1017/S0022226700012755.

Nolan, F. (1994). Auditory and acoustic analysis in speaker recognition. In J. Gibbons (Ed.), *Language and the Law* (pp. 326–345). London: Longman.

Nolan, F. (1997). Speaker recognition and forensic phonetics, In W. J. Hardcastle, & J. Laver (Eds.), *The Handbook of Phonetic Sciences* (1st ed.) (pp. 744–767). Oxford: Blackwell.

Nolan, F., McDougall, K., de Jong, G., & Hudson, T. (2009). The DyViS database: style-controlled recordings of 100 homogeneous speakers for forensic phonetic research. *The International Journal of Speech, Language and the Law 16*(1), 31–57 *Equinox*, doi: 10.1558/ijsll.v16i1.31

Novoa, L., Fein, D., & Obler, L. K. (1988). Talent in foreign languages: A case study. In L. K. Obler, & D. Fein (Eds.), *The exceptional brain: Neuropsychology of talent and special abilities* (pp. 294–302). New York, NY: The Guildford Press.

OED Online. (n.d.). bilingualism, n. [online dictionary entry]. www.oed.com/view/Entry/18968. Oxford University Press. (Accessed 29/08/2022)

Pakai, J. (2021). *The perception of speech varieties – An analysis of authentic and feigned accents (applications in forensic linguistics)*. [Unpublished bachelor's dissertation, Ludwig-Maximilians-Universität München].

Paradis, J. (2007). Second language acquisition in childhood. In E. Hoff, & M. Shatz (Eds.), *Blackwell handbook of language development* (pp. 387–405). Oxford: Blackwell.

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America, 119*(4), 2382–2393. *Scitation*, doi: 10.1121/1.2178720

Pardo, J. S., Jay, I. C., & Krauss, R. M. (2010). Conversational role influences speech imitation. *Attention, Perception, & Psychophysics, 72*(8), 2254–2264. *SpringerLink*, doi: 10.3758/APP.72.8.2254

Pardo, J. S., Gibbons, R., Suppes, A., & Krauss, R. M. (2012). Phonetic convergence in college roommates. *Journal of Phonetics*, *40*(1), 190–197. *ScienceDirect*, doi: 10.1016/j.wocn.2011.10.001

Petrova, O., Plapp, R., Ringen, C., & Szentgyörgyi, S. (2006). Voice and aspiration: evidence from Russian, Hungarian, German, Swedish, and Turkish. *The Linguistic Review, 23*(1), 1–35. *De Gruyter*, doi: 10.1515/TLR.2006.001

Place, S., & Hoff, E. (2015). Effects and noneffects of input in bilingual environments on dual language skills in 2 ½-year-olds. *Bilingualism: Language and Cognition, 19*(5), 1–19. *Cambridge Core,* doi: 10.1017/S1366728915000322

Podlipsky, V. J., Simackova, S., & Chladkova, K. (2021). Phonetic drift reveals interconnected phonological representation in simultaneous bilinguals: A case study of English and Czech stop consonants. *International Journal of Bilingualism, 25*(3), 789–799. *SAGE Journals,* doi: 10.1177%2F1367006920932211

Potowski, K. (2013). Language maintenance and shift. In R. Bayley, R. Cameron, & C. Lucas (Eds.), *The Oxford handbook of sociolinguistics* (pp. 321–339). Oxford: Oxford University Press.

Python Core Team. (2015a). *Python: A dynamic, open source programming language*. Python Software Foundation. [Programming language version 3.9.7] Available from https://www.python.org/

Python Core Team. (2015b). *Python: A dynamic, open source programming language*. Python Software Foundation. [Programming language version 3.10.5] Available from https://www.python.org/

Qi, C. H., Kaiser, A. P., Milan, S., & Hancock T. (2006). Language performance of low-Income African American and European American preschool children on the PPVT–III. *Language, Speech, and Hearing Services in Schools, 37*(1), 5–16. *ASHAWire*, doi: 10.1044/0161-1461(2006/002)

R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. [Programming language version 1.4.1106]. Available from https://www.R-project.org

Redmayne, M. (2001). *Expert evidence and criminal justice*. Oxford: Oxford University Press.

Regina v. T. (2010). Court of Appeal, Criminal Division, case 2439. *Casemine* [online]. https://www.casemine.com/judgement/uk/5b46f1f02c94e0775e7ee6c4 (Accessed 29/08/2022)

Riney, T. J., & Takagi, N. (1999). Global foreign accent and voice onset time among Japanese EFL speakers. *Language Learning*, *49*(2), 275–302. *Wiley Online Library*, doi:10.1111/0023-8333.00089

Robert, J. (2011). *Pydub* [Python library version 0.25.1]. Available from https://github.com/jiaaro/pydub

Robinson, P. (2013). Aptitude and second language acquisition. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. *Wiley Online Library*, doi: 10.1002/9781405198431.wbeal0035

Rock, F. (2011). Variation and forensic linguistics. In W. Maguire, & A. McMahon (Eds.), *Analysing variation in English* (pp. 199–218). Cambridge: Cambridge University Press. *Cambridge Core*, doi: 10.1017/CBO9780511976360.010.

Rogers, H. (1998). Foreign accent in voice discrimination: A case study. *The International Journal of Speech, Language and the Law, 5*(2), 203–208. *Equinox*, doi: 10.1558/sll.1998.5.2.203

Romaine, S. (1989). *Bilingualism*. Oxford: Blackwell.

Rose, P. (2002). *Forensic speaker identification*. London: Taylor & Francis.

Rose, P. (2006). Technical forensic speaker recognition: Evaluation, types and testing of evidence. *Computer Speech and Language*, *20*(2–3): 159–191. *ScienceDirect*, doi: 10.1016/j.csl.2005.07.003.

Rose, P., & Morrison, G. S. (2009). A response to the UK Position Statement on forensic speaker comparison. *The International Journal of Speech, Language and the Law, 16*(1), 139–163. *Equinox*, doi: 10.1558/ijsll.v16i1.139

Rothman, J. (2009). Understanding the nature and outcomes of early bilingualism: romance languages as heritage languages. *International Journal of Bilingualism, 13*(2), 155–163. *SAGE Journals,* doi: 10.1177/1367006909339814

Rounds, C. (2009). *Hungarian: An Essential Grammar* (2nd ed.). Abingdon: Routledge

Ruben, R. J. (1997). A time frame of critical/sensitive periods of language development. *Acta Oto-Laryngologica*, *117*(2), 202–205. *Taylor & Francis Online,* doi:10.3109/00016489709117769

Sanchez, L. (2015). L2 activation and blending in third language acquisition: Evidence of crosslinguistic influence from the L2 in a longitudinal study on the acquisition of L3 English. *Bilingualism: Language and Cognition, 18*(2), 252–269. *Cambridge Core,* doi: 10.1017/S1366728914000091

Sanz, C., Park, H., & Lado, B. (2015). A functional approach to cross-linguistic influence in ab initio L3 acquisition. *Bilingualism: Language and Cognition*, *18*(2), 236–251. *Cambridge Core,* doi: 10.1017/S1366728914000285

Schilling, N., & Marsters, A. (2015). Unmasking identity: Speaker profiling for forensic linguistic purposes. *Annual Review of Applied Linguistics*, *35*, 195–214. *Cambridge Core*, doi: 10.1017/S0267190514000282.

Schmid, M. S. (2007). The role of L1 use for L1 attrition. In B. Köpke, M. S. Schmid, M. Keijzer, & S. Dostert (Eds.), *Language Attrition: Theoretical perspectives* (pp. 135–153). Amsterdam: John Benjamins.

Schmid, M. S. (2011). *Language Attrition*. Cambridge: Cambridge University Press.

Schmid, M. S. (2012). Attrition and Multilingualism. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. *Wiley Online Library,* doi: 10.1002/9781405198431.wbeal0060

Schmid, M. S. (2013). First language attrition. *WIREs Cognitive Science, 4*(2), 94–115. *WIREs*, doi: 10.1002/wcs.1218

Schmid, M. S. (2019). The impact of frequency of use and length of residence on L1 attrition. In M. S. Schmid, & B. Köpke (Eds.), *The Oxford handbook of language attrition* (pp. 288–303). Oxford: Oxford University Press.

Schmid, M. S., & Köpke, B. (2007). Bilingualism and attrition. In B. Köpke, M. S. Schmid, M. Keijzer, & S. Dostert (Eds.), *Language Attrition: Theoretical perspectives* (pp. 1–7). Amsterdam: John Benjamins.

Schmid, M. S., & Yilmaz, G., (2018). Predictors of language dominance: an integrated analysis of first language attrition and second language acquisition in late bilinguals. *Frontiers in Psychology, 9*(1306), 1–19. *Frontiers*, doi: 10.3389/fpsyg.2018.01306

Schoonmaker-Gates, E. (2015). On voice-onset time as a cue to foreign accent in Spanish: Native and nonnative perceptions. *Hispania, 98*(4), 779–791.

Scontras, G., Fuchs, Z., & Polinsky, M. (2015). Heritage language and linguistic theory. *Frontiers in Psychology, 6*(1545), 1–20. *Frontiers*, doi: 10.3389/fpsyg.2015.01545

Scovel, T. (1969). Foreign accent, language acquisition, and cerebral dominance. *Language Learning, 19*(3–4), 245–254. *Wiley Online Library*, doi: 10.1111/j.1467-1770.1969.tb00466.x

Shaw, J., Best, C., Docherty, G., Evans, B., Foulkes, P., Hay, J., & Mulak, K. (2018). Resilience of English vowel perception across regional accent variation. *Laboratory Phonology 9*(1), 11. doi: 10.5334/labphon.87

Shrem, Y., Goldrick, M., & Keshet, J. (2019). Dr. VOT: Measuring positive and negative voice onset time in the wild. [Computer program] *Proceedings of Interspeech 2019*, 629–633. *ISCA*, doi: 10.21437/Interspeech.2019-1735. Available from https://github.com/MLSpeech/Dr.VOT

Siptár, P. (2015). *Kis magyar fonológia: Válogatott karcolatok* [Small Hungarian Phonology]. Veszprém: Pannon Egyetemi Kiadó.

Siptár, P., & Törkenczy, M. (2000). *The phonology of Hungarian*. Oxford: Oxford University Press.

SoX team. (2015). *SoX - Sound eXchange* [Computer program version 14.4.2]. Available from https://sourceforge.net/projects/sox/

Spinu, L. E., Hwang, J., & Lohmann, R. (2018). Is there a bilingual advantage in phonetic and phonological acquisition? The initial learning of word-final coronal stop realization in a novel accent of English. *International Journal of Bilingualism, 22*(3), 350–370. *SAGE Journals*, doi: 10.1177/1367006916681080

Stoehr, A., Benders, T., Van Hell, J. G., & Fikkert, P. (2017). Second language attainment and first language attrition: The case of VOT in immersed Dutch-German late bilinguals. *Second Language Research*, *33*(4), 483–518. *SAGE Journal*, doi: 10.1177/0267658317704261

Stölten, K., Abrahamsson, N., & Hyltenstam, K. (2014). Effects of age of learning on voice onset time: Categorical perception of Swedish stops by near-native L2 speakers. *Language and Speech*, *57*(4), 425–450. *SAGE Journals*, doi: 10.1177/0023830913508760

Sundara, M. (2005). Acoustic-phonetics of coronal stops: A cross-language study of Canadian English and Canadian French. *The Journal of the Acoustical Society of America, 118*(2), 1026–1037. *Scitation*, doi:10.1121/1.1953270

Swartz, B. L. (1992). Gender difference in voice onset time. *Perceptual and Motor Skills, 75*(3), 983–992. *SAGE Journals*, doi: 10.2466/pms.1992.75.3.983

Száraz, B., & Gráczi, T. E. (2020). Explozívák realizációja fiatal és öregedő felnőttek beszédében [The realisation of explosives in young and aging adults' speech]. *Beszédtudomány* [Speech Science], *2020*, 73–98. *MTAK*, doi: 10.15775/Besztud.2020.73-98

Szende, T. (1992). Phonological representation and lenition processes. *Magyar fonetikai füzetek* [Hungarian Papers in Phonetics], *24*.

Tabain, M. (2013). Research methods in speech production. In M. J. Jones, & R. Knight (Eds.), *Bloomsbury Companion to Phonetics* (pp. 39–56). London: Bloomsbury.

Taeschner, T. (1983). *The sun is feminine: A study on language acquisition in bilingual children*. Berlin: Springer.

Tar, É. (2014). The acquisition of the voicing contrast in word-initial bilabial and alveolar stops – Atypical data from Hungarian. *Clinical Linguistics & Phonetics*, *28*(4), 269–282. *Taylor & Francis Online*, doi: 10.3109/02699206.2013.835445

The International Phonetic Association. (1999). *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press.

The MathWorks Inc. (2010). MATLAB [Computer program version 7.10.0 (R2010a)]. Natick, Massachusetts. Available from https://uk.mathworks.com/products/matlab.html

Thompson, I. (1991). Foreign accents revisited: The English pronunciation of Russian immigrants. *Language Learning*, *41*(2), 177–204. *Wiley Online Library*, doi: 10.1111/j.1467-1770.1991.tb00683.x

Thordardottir, E. (2017). Amount trumps timing in bilingual vocabulary acquisition: effects of input in simultaneous and sequential school-age bilinguals. *International Journal of Bilingualism*, *23*(1), 236–255. *SAGE Journals*, doi: 10.1177/1367006917722418

Unsworth, S. (2013). Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism: Language and Cognition, 16*(1), 86–110. *Cambridge Core,* doi: 10.1017/S1366728912000284

Ürmösné Simon, G. (2014). A bilingvizmus pszicho- és neurolingvisztikai aspektusai, határtudományai és oktatáspolitikai szerepe [The psycho- and neurolinguistic aspects of bilingualism, and its roles in border studies and educational policy]. *Magyar Rendészet* [Hungarian Policing], *14*(2), 117–125.

van de Weijer, J., & Kupisch, T. (2015). Voice onset time in heritage speakers and second-language speakers of German. In E. Babatsouli & D. Ingram (Eds.), *Proceedings of the international symposium on monolingual and bilingual speech 2015* (pp. 414–420). Chania: Institute of Monolingual and Bilingual Speech.

Vaughn, C., Baese-Berk, M., & Idemaru, K. (2018). Re-examining phonetic variability in native and non-native speech. *Phonetica*, *76*(5), 1–32. *De Gruyter*, doi: 10.1159/000487269

Vaux, B., Tucker, E., & Cooper, J. (2006). *Linguistic field methods*. Eugene, OR: Wipf & Stock.

Volaitis, L., & Miller, J. (1992). Phonetic prototypes: Influence of place of articulation and speaking rate on the internal structure of voicing categories. *The Journal of the Acoustical Society of America*, *92*, 723–735. *Scitation*, doi: 10.1121/1.403997

Watt, D. (2010). The identification of the individual through speech. In C. Llamas, & D. Watt (Eds.), *Language and Identities* (pp. 76–85). Edinburgh: Edinburgh University Press.

Wei, L. (2008). Research perspectives on bilingualism and multilingualism. In L. Wei, & M. G. Moyer (Eds.), *The Blackwell guide to research methods in bilingualism and multilingualism* (pp. 3–17). Oxford: Blackwell. *Wiley Online Library*, doi: 10.1002/9781444301120.ch1

White, L., & Genesee, F. (1996). How native is near-native? The issue of ultimate attainment in adult second language acquisition. *Second language research, 12*(3), 233–265. *SAGE Journals*, doi: 10.1177/026765839601200301

Wickham, H. (2011). The Split-Apply-Combine Strategy for Data Analysis. *Journal of Statistical Software, 40*(1), 1–29. http://www.jstatsoft.org/v40/i01/. [R package version 1.8.7] Available from https://CRAN.R-project.org/package=plyr

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. [R package version 3.3.6]. Available from https://ggplot2.tidyverse.org

Wickham, H. (2021). *tidyr: Tidy Messy Data*. [R package version 1.1.3]. Available from https://CRAN.R-project.org/package=tidyr

Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn. M., Pedersen, T. L., Miller, E., Milton Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., & Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software 4*(43). *The Journal of Open Source Software*, doi: 10.21105/joss.01686. [R package version 1.3.2]. Available from https://CRAN.R-project.org/package=tidyverse

Wickham, H., François, R., Henry, L., & Müller, K. (2021). *dplyr: A Grammar of Data Manipulation*. [R package version 1.0.6]. Available from https://CRAN.R-project.org/package=dplyr

Wrembel, M. (2014). VOT patterns in the acquisition of third language phonology. *Concordia Working Papers in Applied Linguistics*, *5*, 751–771.

Wrembel, M. (2015). Cross-linguistic influence in second vs. third language acquisition of phonology. In U. Gut, R. Fuchs, & E.-M. Wunder (Eds.), *Universal or diverse paths to English phonology* (pp. 41–70). Berlin: De Gruyter.

Xie, X., & Jaeger, T. F. (2020). Comparing non-native and native speech: Are L2 productions more variable? *The Journal of the Acoustical Society of America*, *147*(5), 3322–3347. *Scitation*, doi: doi.org/10.1121/10.0001141

Yarmey, A. D. (2012). Factors affecting lay persons' identification of speakers. In L. M. Solan, & P. M. Tiersma (Eds.), *The Oxford handbook of language and law* (pp. 547–556). Oxford: Oxford University Press. *Oxford Academic,* doi: 10.1093/oxfordhb/9780199572120.013.0040

Yavas, M. (2002). Voice onset time patterns in bilingual phonological development. In F. Windsor, M. L. Kelly, & N. Hewlett (Eds.), *Investigations in clinical phonetics and linguistics* (pp. 341–350). Mahwah, NJ: Lawrence Erlbaum.

Yavas, M., & Byers, E. (2015). Production of voiceless stops in early sequential Spanish-English bilinguals. In M. Yavas (Ed.), *Unusual productions in phonology* (pp. 242–260). New York: Psychology Press.

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

Zen, E. (2020). Role of regional language background and speech styles on the production of voice onset time (VOT) in English among Indonesian multilinguals. *Indonesian Journal of Applied Linguistics, 10*(2), 359-368. *Universitas Pendidikan* Indonesia, doi: 10.17509/ijal.v10i2.28604

Zhang, A. (2017). *Speech recognition* [Python library version 3.8]. Available from https://github.com/Uberi/speech_recognition

Zhao, L., & Chodroff, E. (2022). The ManDi Corpus: A spoken corpus of Mandarin regional dialect. *Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022)*, 1985–1990.

Appendices

Appendix A: BEA Corpus VOT Data

The VOT data from the BEA corpus can be found at:
https://github.com/julipakai/hungarian-english_VOT/blob/main/diss_mono_hu.csv

Appendix B: Survey

A pdf of the survey can be found at:
https://github.com/julipakai/hungarian-english_VOT/blob/main/survey.pdf
The survey is also available as a Qualtrics (qsf) file (including JavaScript) at:
https://github.com/julipakai/hungarian-english_VOT/blob/main/survey.qsf

Appendix C: Survey Results

The raw survey results are available at:
https://github.com/julipakai/hungarian-english_VOT/blob/main/surveydata.csv

Appendix D: Interview Material

A word document of the interview material can be found at:
https://github.com/julipakai/hungarian-english_VOT/blob/main/materials.docx

Appendix E: Python Transcription Script

The Python script for the transcription can be found at:
https://github.com/julipakai/hungarian-english_VOT/blob/main/recognition.py

Appendix F: Bilingual VOT Data

The bilingual VOT data can be found at:
https://github.com/julipakai/hungarian-english_VOT/blob/main/diss_bi_all.csv

Appendix G: R Script

The R script used for the statistics and visualisation can be found at:
https://github.com/julipakai/hungarian-english_VOT/blob/main/r_survey.R

Appendix H: Data for LR Calculations

The data for the LR calculations can be found at:
https://github.com/julipakai/hungarian-english_VOT/tree/main/lr_data