languages

MDPI

Article

Longitudinal Developments in Bilingual Second Language Acquisition and First Language Attrition of Speech: The Case of Arnold Schwarzenegger

Lisa Korndser * and Ineke Mennen

Department of English Studies, University of Graz, 8010 Graz, Austria; ineke.mennen@uni-graz.at
* Correspondence: lisa.kornder@uni-graz.at

Abstract: The purpose of this investigation was to trace first (L1) and second language (L2) segmental speech development in the Austrian German–English late bilingual Arnold Schwarzenegger over a period of 40 years, which makes it the first study to examine a bilingual's speech development over several decades in both their languages. To this end, acoustic measurements of voice onset time (VOT) durations of word-initial plosives (Study 1) and formant frequencies of the first and second formant of Austrian German and English monophthongs (Study 2) were conducted using speech samples collected from broadcast interviews. The results of Study 1 showed a merging of Schwarzenegger's German and English voiceless plosives in his late productions as manifested in a significant lengthening of VOT duration in his German plosives, and a shortening of VOT duration in his English plosives, closer to L1 production norms. Similar findings were evidenced in Study 2, revealing that some of Schwarzenegger's L1 and L2 vowel categories had moved closer together in the course of L2 immersion. These findings suggest that both a bilingual's first and second language accent is likely to develop and reorganize over time due to dynamic interactions between the first and second language system.

Keywords: first language attrition; second language acquisition; sequential bilingualism; voice onset time; vowel formants; speech development; English; (Austrian) German; phonetics

check for updates

Citation: Kornder, Lisa, and Ineke Mennen. 2021. Longitudinal Developments in Bilingual Second Language Acquisition and First Language Attrition of Speech: The Case of Arnold Schwarzenegger. *Languages* 6: 61. https://doi.org/10.3390/languages6020061

Academic Editors: Robert Mayr, Jonathan Morris, Juana M. Liceras and Raquel Fernández Fuertes

Received: 25 February 2021
Accepted: 19 March 2021
Published: 25 March 2021

1. Introduction

Learning a second language (L2) late in life often entails that speakers retain a noticeable foreign accent in their L2 resulting from first language (L1) influences on the late-acquired L2 phonetic/phonological system (Flege 1980, 1981; Flege et al. 1996; MacKay et al. 2001; Scovel 1969). Examining L1 influences on pronunciation abilities in the L2 has a long tradition in second language and bilingualism research: Against the background of theories evolving around maturational constraints, such as the critical period hypothesis (Lenneberg 1967; Penfield and Roberts 1959), a prevailing view was that an individual's native language system—once fully mature—was not likely to be affected and modified by a late-acquired L2 system (Bylund 2009; Flege et al. 2003; Scovel 1969). In the past decades, however, research came to acknowledge the bi-directional nature of interactions between a bilingual's language systems (e.g., Flege 1987; Mennen 2004; Sancier and Fowler 1997). This change in orientation resulted from the observation that also a mature native system might be affected by L2-induced changes (e.g., Flege 1987). L1 modifications and the resulting decline of linguistic abilities in one's native language observed in late L2 learners who are being long-term immersed in an L2 environment is commonly referred to as *L1 attrition* (Köpke and Schmid 2004). Interest in gaining a better understanding of whether it is possible for healthy individuals to "unlearn" their L1 as a result of L2-learning experience developed in the early 1980s (see Köpke and Schmid 2004, for an overview). Since then, a considerable amount of research has been conducted to explore attrition effects on native pronunciation abilities, which provides evidence for the permeability of late bilinguals' L1 pronunciation system with regard to segmental (Bergmann et al. 2016; de Leeuw et al. 2017; de Leeuw 2019; Dmitrieva et al. 2010; Mayr et al. 2012; Stoehr et al. 2017; Ulbrich and Ordin 2014) and prosodic (de Leeuw et al. 2012b; de Leeuw 2019; Mennen and Chousi 2018) features.

The present study aims to further extend the growing body of research on L1 attrition and L2 acquisition and to gain a more profound understanding of the relationship between L2 acquisition and L1 attrition in an immersion context. To this end, long-term changes in the L1 Austrian German and the L2 English of the late consecutive bilingual Arnold Schwarzenegger were examined by conducting acoustic analyses of voice onset time (VOT) and vowel formants based on spontaneous speech samples. To date, this case study is the first which does not exclusively focus on identifying longitudinal pronunciation changes in *either* the L1 (e.g., de Leeuw 2019) *or* the L2 (e.g., Saito et al. 2019), or examining L1–L2 interactions over a relatively short period of time (e.g., Chang 2012; Sancier and Fowler 1997) but which traces the trajectory of segmental speech development in both the L1 and a late-acquired L2 over a period of 40 years.

1.1. Bi-Directional L1–L2 Influences: An Integrated View of Bilingual Speech Development

Early investigations into second language speech acquisition started from the premise that the "steady state" of a biologically mature L1 pronunciation system is not likely to be disrupted by a late-acquired L2 system (Lado 1957; Lenneberg 1967). This resulted in a rather biased focus on identifying L1 influences on the L2 system and neglected potential interactions between a speaker's languages. Furthermore, a seemingly straightforward definition of *the bilingual* was proposed according to which a bilingual's language systems are representative of two monolingual systems, which resulted in the prevailing view that only those individuals who showed native and balanced proficiency in both languages were to be considered "real" bilinguals (Bloomfield 1933; Thiery 1978). Moving away from this static point of view, Grosjean (1989, 1997) proposed a holistic approach to bilingualism, which reflected bilingual reality more appropriately. He argued that a bilingual's linguistic configuration is per se different from that of a monolingual due to a dynamic interaction between the language systems and, hence, defining bilingual speakers against the background of monolingual competence is rather misleading. With the development of this integrated view, bilingualism and L2 acquisition research shifted perspectives and started to acknowledge bi-directional L1–L2 influences as an inherent characteristic of speech development in second language and bilingual speakers (Flege 1987; Odlin 1989, 2006; Sharwood Smith and Kellerman 1986).

Support for this holistic view and the notion of bi-directionality comes from the Speech Learning Model (SLM) proposed by Flege (1995; see Flege and Bohn 2020, for a revised version of the SLM). According to the SLM, bilinguals' sound systems are not isolated from each other, but exist in a shared phonetic space, which accounts for a mutual interaction between a speaker's language systems (Flege 1995; Flege and Bohn 2020). In this shared space, L1–L2 interactions are determined by different mechanisms which can not only lead to non-native L2 productions, but might also result in a reorganization of L1 categories. A substantial amount of empirical research on bilingual speech development has supported one of the SLM's main predictions that similar L1 and L2 sounds would be difficult to produce authentically due to assimilatory effects (Baker and Trofimovich 2005; Flege 1991; Flege et al. 1996; Flege and Hillenbrand 1984; Thornburgh and Ryalls 1998). That is, late bilinguals may fail to identify fine phonetic differences between acoustically related L1 and L2 sounds and establish a merged L1–L2 category which differs from the respective monolingual categories. Flege and Hillenbrand (1984), for example, found that both L1 French late learners of L2 English, and L1 English late learners of L2 French produced French /t/with VOT values that considerably exceeded monolingual French short-lag VOT values, but were still too short for long-lag English categories. That is, they had established merged phonetic categories due to L1–L2 assimilation. Dissimilatory effects, by contrast, have been mainly observed among early bilinguals who manage to establish distinct L1 and L2 categories for closely related sounds, but may overshoot the monolingual targets in both languages in an attempt to maintain contrast (Flege et al. 2003; Flege and Eefting 1988; Mack 1990). A recently revised version of the SLM (Flege and Bohn 2020) stresses the need to focus on individual differences between L2 learners when it comes to mastering an L2 sound system. In fact, inter- as well as intra-subject variability have been widely documented even in rather homogenous speaker groups examined in cross-sectional studies (e.g., Bongaerts et al. 2000; Major 1992; Mennen 2004; Moyer 1999). These studies show that some speakers are well able to attain native proficiency in their L2 while others clearly fall outside the native range, and thus challenge yet again a static view on bilingual speech acquisition.

While research conducted within the framework of the SLM focuses on highly proficient bilinguals with long-term L2 experience (see Flege 1995), bi-directional influences between speakers' linguistic systems and L1 modifications resulting from this interaction have also been observed to occur at a very early stage of L2 acquisition. Chang (2012, 2013), for instance, identified a *phonetic drift* of L1 English plosive and vowel categories towards L2 Korean categories in speakers learning Korean in an instructional setting. As none of the speakers had prior experience with Korean, the L1 changes did not result from long-term L2 exposure and usage, but from recent L2 experience and the novelty of L2 input. Evidence for a drift of L1 categories towards L2 norms in inexperienced individuals also comes from Kartushina et al. (2016), who trained L1 French speakers on the production of Russian and Danish vowels in 1-h training sessions. Similar to the subjects in Chang (2012, 2013), the speakers in this study experienced modifications of their L1 vowel categories in the direction of L2 norms despite lacking previous long-term L2 experience. Kartushina and Martin (2019) showed that the vowel systems of Basque-Spanish bilinguals had moved closer to English norms after completing a two-week study abroad English program. Four months after the program, however, an acoustic re-analysis of the speakers' L1 and L2 vowels revealed that the vowel categories had drifted back to native norms (see Chang 2019), which suggests that L1 modifications occurring in the initial stage of L2 learning are only temporary and are likely to be reversed due to changes in language use and linguistic environment.

Further evidence for the impact of recent L2 experience on L1 speech production and for the reversibility of L1 changes as a function of linguistic environment is provided by Sancier and Fowler (1997; see also Tobin et al. 2017). In an 11–months case study, they examined VOT produced by an L1 Brazilian-Portuguese advanced speaker of L2 English, who travelled between Brazil and the US at monthly intervals. Sancier and Fowler describe a *gestural drift* of VOT in both languages towards the most recently experienced language, i.e., VOT durations were longer and thus more English-like after returning from the US, and shorter and more Portuguese-like after staying in Brazil.

While the studies outlined above explored L2-induced changes in the L1 pronunciation of speakers who were at the onset of L2 learning (Chang 2012, 2013, 2019; Kartushina et al. 2016) or experienced regular changes in their linguistic environment (Sancier and Fowler 1997; Tobin et al. 2017), other studies examined L2-induced changes in the L1 pronunciation of experienced late bilinguals being permanently and long-term immersed in an L2 environment. Research into phonetic and phonological attrition has documented modifications in bilinguals' L1 segmental (de Leeuw 2019; Dmitrieva et al. 2010; Mayr et al. 2012; Stoehr et al. 2017) and prosodic (de Leeuw et al. 2012b; Mennen and Chousi 2018) productions in the direction of L2 norms. In addition, research provides evidence for listeners' perceptual sensitivity to divergences from L1 norms, that is, listeners have been shown to judge bilingual attrited speech as sounding less native compared to monolingual non-attrited speech (Bergmann et al. 2016; de Leeuw et al. 2010; Hopp and Schmid 2013; Mayr et al. 2020; Schmid and Hopp 2014). Research examining potential changes in long-term L2-immersed bilinguals' realizations of plosives and vowels—the two sound classes investigated in the present study—will be discussed in more detail in Section 1.2.

Taken together, these studies show that speakers may experience a restructuring of certain L1 features in the direction of the L2 as a consequence of long-term L2 learning experience and L2 immersion. However, just as bilinguals master L2 pronunciation with varying degrees of success (see Bongaerts et al. 2000; Major 1992; Mennen 2004; Moyer 1999), the extent to which native abilities in L1 pronunciation decline might also differ among individuals (see Bergmann et al. 2016; de Leeuw et al. 2017; Major 1992; Mayr et al. 2012). Furthermore, previous research shows that not all features of L1 pronunciation—within the same individual and the same sound category—undergo attrition (e.g., Bergmann et al. 2016; Mayr et al. 2012; Stoehr et al. 2017).

Notions of an end state of L2 learning (see Birdsong 2009, for a discussion) and of a steady state of the L1 (Lado 1957; Lenneberg 1967) can hardly be reconciled with the empirical findings outlined above, which document a reorganization of both the L1 and the L2 system as a result of bi-directional interaction processes, and show that L2 acquisition and L1 maintenance are determined by inter- and intra-speaker variability. Such observations offer convincing evidence that bilingual speech development is highly *dynamic* and characterized by a vivid interaction between a speaker's linguistic systems. The notion of dynamic development also lies at the core of dynamic systems theory (DST), a theoretical approach to language development (de Bot and Larsen-Freeman 2011; Verspoor et al. 2008). According to DST, which essentially supports an integrated view of bilingualism (Grosjean 1989), language development progresses in a non-linear, often unpredictable manner, and is determined by an intricate interplay of system-internal and external factors (de Bot et al. 2007). Longitudinal research, following pronunciation changes in both L2 acquisition and L1 attrition and identifying potential interactions between the two, allows us to gain a better understanding of the dynamics of language development over time.

1.2. The Present Study

The present investigation aims to contribute to the expanding body of research on bilingual L2 acquisition and L1 attrition of pronunciation by exploring the segmental speech development of Arnold Schwarzenegger (AS), a late consecutive Austrian German–English bilingual who has been immersed in an L2 English environment for the past 52 years. His early (1979–1988) and late (2012–2018) L1 and L2 segmental speech development was examined based on acoustic analyses of VOT of word-initial plosives (Study 1) and formant frequency analyses of the first (F1) and second (F2) formant of stressed monophthongs (Study 2).

The temporal dimension of VOT, which has previously been described as a reliable acoustic–phonetic correlate of the voiced–voiceless distinction in pre-vocalic plosives (Abramson and Whalen 2017; Lisker and Abramson 1964), refers to the time lapse between the plosive release and the onset of vocal fold vibration (Lisker and Abramson 1964). English and Austrian German, the two languages examined in the present investigation, exhibit cross-linguistic differences with regard to their implementation of VOT contrast in word-initial pre-vocalic plosives. English is an aspirating language and distinguishes between long-lag VOT for voiceless aspirated plosives and short-lag VOT for phonologically voiced plosives (Lisker and Abramson 1964), with approximately 35 milliseconds being the threshold value for the long-lag versus short-lag distinction (e.g., Keating 1984). Some studies report the occurrence of English pre-voiced targets, that is, vocal fold vibration is maintained throughout the stop closure phase (Flege and Brown 1982; Lisker and Abramson 1964). English pre-voiced plosives have been observed to occur predominantly in controlled speech production contexts (Roach 2009), particularly in voiced environments (Docherty 1992), but overall do not appear to be produced systematically across speakers and thus occur in free variation (e.g., Flege and Brown 1982).

Unlike Standard German, which—similar to English—distinguishes between long-lag and short-lag plosives (e.g., Braunschweiler 1997; Jessen 1998), Austrian German does typically not maintain this voiced–voiceless distinction in bilabial and alveolar contexts, despite also being an aspirating language (Moosmüller et al. 2015). That is, speakers of

Austrian German varieties show a tendency to produce both phonologically voiced and voiceless bilabial and alveolar targets with short-lag VOT and, thus, neutralize the contrast between /b/and /p/, and /d/and /t/, particularly in spontaneous conversational speech (Hödl 2019; Moosmüller et al. 2015). By contrast, Austrian German velar plosives, similar to English, are produced within two distinct VOT categories, that is, long-lag VOT for voiceless and short-lag VOT for voiced velar plosives (e.g., Hödl 2019). Based on the different VOT patterns observed in English and Austrian German, Study 1 set out to explore if and to what extent AS's realization of L1 and L2 plosives has changed in the past 40 years of L2 immersion, that is, (1) whether his L2 VOT categories have moved closer to L2 norms and (2) whether his L1 categories have shifted away from L1 productions norms and, thus, have become less native.

Previous research on L2 acquisition of VOT has examined L1 speakers of voicing languages, such as French, Spanish or Dutch, acquiring an aspirating language, including English or (Standard) German, as their L2 (e.g., Simon 2009; Stoehr et al. 2017; Thornburgh and Ryalls 1998), or vice versa (e.g., Flege and Hillenbrand 1984). It has been observed that late bilinguals often fail to acquire distinct VOT categories in their L2 resulting from assimilation patterns (Flege 1991; Flege and Hillenbrand 1984) or transferring an L1 feature, such as pre-voicing, to the L2 (Mayr et al. 2012; Simon 2009; Stoehr et al. 2017). Even if bilinguals are able to produce distinct L1 and L2 VOT categories, these often do not resemble those of monolingual speakers of the respective language. Mack (1990), for instance, showed that her L1 French L2 English subject was able to maintain phonetic contrast between English and French /p t k/, but realized the target plosives with VOT values that overshot the monolingual targets in both languages.

While a considerable amount of research has explored the acquisition of late-acquired L2 VOT categories (see above), comparatively few studies so far have investigated attrition of VOT in the L1 of late L2 acquirers being long-term residents in an L2 country (Flege and Hillenbrand 1984; Major 1992; Mayr et al. 2012; Stoehr et al. 2017; Sučková 2020). Overall, findings reveal that VOT is indeed sensitive to be affected by modifications due to L2 learning experience; these modifications, however, have been shown to be more prevalent in voiceless plosives while voiced targets seem to be less likely to undergo attrition (Mayr et al. 2012; Stoehr et al. 2017). The L1 French L2 English-immersed bilinguals in Flege and Hillenbrand (1984), for example, experienced a lengthening of their L1 short-lag VOT category for /t/in the direction of English long-lag /t/. Similarly, the L1 English L2 French-immersed subjects in Flege (1987) had assimilated their L1 long-lag categories for /t/to French short-lag VOT (see also Major 1992). Mayr et al. (2012) examined VOT in the productions of monozygotic twin sisters who were both L1 Dutch speakers of L2 English, but one sister had moved to an L2-speaking country in adulthood while the other twin had remained in the L1 environment. Results showed that the VOT categories for voiceless plosives produced by the L2-immersed twin had moved closer to L2 norms, that is, she experienced a lengthening of VOT in her Dutch voiceless plosives. At the same time, her L1 voiced categories remained unaffected by L2-induced changes, i.e., she produced Dutch /b d ɡ/with consistent pre-voicing, and also pre-voiced her English voiced tokens, indicating an L1 influence on the L2.

Unlike previous investigations examining VOT, the present study does not juxtapose a voicing and an aspirating language, but compares two aspirating languages, with Austrian German featuring a neutralization of the voiced–voiceless distinction in bilabial and alveolar plosive targets (Hödl 2019; Moosmüller et al. 2015). Thus, the bilingual subject in the present study is confronted with the task of acquiring an L2 contrast which is essentially absent in his L1, at least in a bilabial and alveolar place of articulation. Successfully implementing this contrast in his L2 presupposes that he has established a distinct long-lag VOT category for all English aspirated plosives. At the same time, the observation that AS implements a VOT contrast in his late L1 bilabial and alveolar plosives would indicate that his L1 plosive categories are affected by modifications in the direction of his L2.

In order to further investigate AS's segmental speech development, an additional sound class was assessed, namely English and Austrian German stressed monophthongs. In Study 2, AS's monophthongal L1 and L2 vowel space was acoustically examined to determine if and to what extent his productions of L1 and L2 vowels have changed in the past 40 years of L2 immersion. The aim was to find out (1) whether his L2 vowel categories have moved closer to L2 production norms, and (2) whether his L1 categories have shifted away from L2 norms and, thus, have become less native. Austrian German comprises eight front vowels, i.e., /i, y, ɪ, ʏ, e, ø, ɛ, œ/, and five back vowels, i.e., /u, ʊ, o, ɔ, ɑ/. Unlike Standard German, Austrian German varieties lack the mid-central vowel /ə/, which "exists neither phonetically nor phonologically" (Moosmüller 2007, p. 52). Californian English—the variety the subject of the present study is predominantly exposed to[1]—includes five front vowels, i.e., /i, ɪ, e, ɛ, æ/, four back vowels, i.e., /u, ʊ, o, ɑ/, and the mid-central vowel /ɝ/(Hagiwara 1997; Ladefoged 2005). One characteristic of Californian English is the merging of /ɔ/and /ɑ/ in the direction of /ɑ/ (e.g., Boberg 2005).

The extent to which late L2 learners are able to successfully acquire L2 vowel categories has been investigated in a remarkable number of studies, of which just a few are listed here (e.g., Baker and Trofimovich 2005; Flege et al. 1997, 2003; Levy and Law 2010; Oh et al. 2011; Piske et al. 2002). Although it has been shown that the amount of L2 experience is positively correlated with L2 vowel production ability, even highly experienced bilinguals may fail to produce L2 vowels in a native manner (Flege et al. 1997; Levy and Law 2010). This failure to establish accurate L2 vowel categories often stems from assimilatory effects, i.e., L2 vowels are assimilated to acoustically related L1 vowels (Baker and Trofimovich 2005; Flege et al. 2003). Flege et al. (2003) have documented such assimilatory effects in a group of late Italian-English bilinguals whose productions of the L2 English vowel /eɪ/ were characterized by significantly less formant movement compared to monolingual English productions. Their inaccurate L2 productions resulted from assimilating the L2 vowel target to the acoustically related L1 Italian vowel /e/, which is typically produced with less formant movement compared to English /eɪ/. Similar observations were made by Baker and Trofimovich (2005). While their early Korean-English bilinguals had established distinct L1 and L2 vowel categories, the late bilinguals' L2 vowels were affected by acoustic features of their L1, that is, they did not manage to produce closely related L1 and L2 vowel targets within separate phonetic categories.

L2-induced articulatory changes in L1 vowel production have received little attention so far (Bergmann et al. 2016; de Leeuw 2019; Mayr et al. 2012). Research shows that late bilinguals' L1 vowels might shift towards L2 norms due to L2 learning experience; however, the extent to which a speaker's native vowel categories are affected by such modifications differs. For example, the L1 vowels produced by the Dutch-English bilingual in Mayr et al. (2012) had shifted closer to L2 production norms, which had manifested itself in an overall more open production of L1 target vowels, as typical of L2 English. Resulting from L1–L2 assimilatory effects, the speaker did not manage to maintain contrast between some of her Dutch and English vowels. These assimilatory processes were, however, not observed to affect all L1–L2 vowels as she was able, for instance, to produce distinct vowel targets for English /ɑ/ and Dutch /a/. These and other findings (Bergmann et al. 2016; de Leeuw 2019) suggest that attrition processes are selective and that not all sounds are equally sensitive to undergo modifications in the direction of the L2 system. It remains unclear, however, why some L1 features are more likely to change while others remain largely stable (Mayr et al. 2012).

Against the background of previous research into L1 attrition and L2 acquisition of plosives and vowels, the two studies conducted in the context of the present investigation aim to reveal potential changes in a bilingual's segmental productions over four decades

[1] Note that *Californian English* is commonly used to refer to different regional/local varieties and sociolects spoken in California (Eckert and Mendoza-Denton 2006). Although AS has been living in California for more than 50 years, it is likely that he was and still is exposed to multiple different English varieties and accents, not only as a result of travelling within and outside the US, but also due to contact with different native and non-native speakers of English.

and thus shed light on some of the processes affecting a speaker's accent in both his L1 and a late-acquired L2.

2. Materials and Methods

2.1. The Subject: Arnold Schwarzenegger

The subject of this case study is Arnold Alois Schwarzenegger (AS), born on July 30, 1947 in Thal, a small rural municipality near the Styrian capital city Graz in Austria. He grew up in a monolingual Austrian German[2] environment and started acquiring English as his L2 when he migrated to the US to seek a career in bodybuilding in 1968. Often described as an "embodiment of the American dream" (Allen 2011), AS did not only gain fame as a bodybuilder, but he is also well known for being an action movie actor, a successful businessman, and a politician who held the office of Governor of California for two terms (2003–2011) (Schwarzenegger 2012).

Although AS had received seven years of formal English instruction in school in Austria, his knowledge of English was rather poor when he arrived in the US at the age of 21 (Outland Baker 2006). Before becoming a permanent resident of Los Angeles, California, in the late 1960s, AS spent some time in London, where he lived with Wag and Dianne Bennett, a British couple who supported AS in the initial stages of his bodybuilding career and helped him work on his English skills (Preston 2015). At this point, however, he had not acquired English as a "functional second language" (Flege and Hillenbrand 1984, p. 710). When he arrived in the US, he struggled particularly with acquiring English pronunciation, which he described as one of the most demanding tasks he was confronted with in the first years after migration (Schwarzenegger 2012). Alongside attending English as a Foreign Language classes at Santa Monica Community College in California, AS strived to improve his English conversation skills by engaging with English-speaking friends on a regular basis (Glaister 2006). Still, his heavily accented English turned out to be problematic when he acted in his first movies in the 1970s and early 1980s, which is why he had to take training sessions with professional dialogue coaches (Miller 2012).

Despite never really losing his distinctive foreign accent, AS has grown confident speaking his L2 on a daily basis over the past 52 years living in an L2 immersion setting. In fact, English has become his dominant means of communication which he makes use of even in L1 settings. He prefers, for example, speaking English when being interviewed by German or Austrian broadcast stations and newspapers (von Uslar 2012; Ziesel 2018). In interviews, he repeatedly stated that it is much easier for him to speak English than German and that the use of his L1 is restricted predominantly to private contexts (Gala 2015; Gersemann 2009; Naumburger Tageblatt 2015). In a 2017-interview, AS reported that he rarely speaks German and, as a result, his L1 proficiency seems to have declined:

> [Interviewer:] How often do you still speak german [*sic*]? After all the years in the USA, do you feel more comfortable to speak English?
>
> [Schwarzenegger:] Not much, I am definitely more comfortable in English. Which should tell you how bad my German has gotten. (Muscle & Fitness 2017)

While some argue that AS, after living in the US for several decades, speaks English "with only a slight accent" (Ramos and Krashen 2013, p. 220), others consider him a perfect example of "embracing his Austrian English accent" (Wan 2017). In fact, his accented English pronunciation is vividly debated in online forums (e.g., Quora 2020; Reddit 2021), has led to many comic imitations (e.g., Collins 2019), and is now considered his personal trademark (Daily Mail UK 2015; Gersemann 2009). However, not only his German-accented English pronunciation attracts attention, but also the question of whether AS has "forgotten" how to speak his L1 seems to be intriguing to the general public (Jackson 2020; Quora 2017). This fascination with AS's accent suggests that people are able to perceive changes in both

[2] In this paper, *Austrian German* is used to refer to the Styrian variety of German spoken in Graz and surrounding rural regions (e.g., Wiesinger 2014).

his L1 and L2 pronunciation. The present study examines whether and to what extent such changes can be evidenced in his segmental speech development over four decades.

2.2. Speech Materials

The present investigation is based on spontaneous speech samples which were extracted from interviews with AS, collected from the online video platform YouTube. The interviews were conducted in rather informal settings, primarily including TV and radio talk shows. They covered various topics, such as AS's former career as a bodybuilder, his ongoing career in the film business, childhood memories, his political activities as Governor of California, and his commitment to environmental issues. The earliest publicly available interviews with AS were conducted in 1979, i.e., approximately ten years after he had migrated to the US. The reason for this lack of early interviews presumably is that he had not been widely recognized as a celebrity prior to the 1970s. In addition, the recordings varied in quality, with the earliest recordings made in the 1970s and 80s partly being of rather poor quality. For analysis, only those recordings were included which allowed for an exact identification of acoustic landmarks for measuring AS's plosive and vowel productions.

After downloading the online video files, the sampling frequency was set to 44.1 kHz for each file. Audio sequences which were disturbed by background noise, several speakers speaking at the same time, hesitations or disfluencies were excluded. The resulting speech corpus contained 656 audio files with a total duration of approximately 5 h. The individual recordings were then categorized according to two stages after AS's migration to the US, representing his early (1979–1988) and late (2012–2018) L1 and L2 speech. For his L2 English, a mid-stage was additionally included, containing samples from 1994 to 2003. Recordings representing his L1 German pronunciation for this stage were not available.

Each audio file was transcribed orthographically using the web-application OCTRA (Pömp and Draxler 2017). The audio files in combination with the respective orthographic transcripts were automatically segmented and labelled in WebMaus Basic, a web tool for automatic phonetic and phonological transcription of non-prompted speech (Kisler et al. 2017; Schiel 1999). The resulting segmentations were hand-corrected in Praat (Boersma and Weenink 2018).

3. Study 1: Schwarzenegger's Plosives

3.1. Analysis

From the speech corpus described above, test tokens used to examine AS's realization of VOT contrast in his L1 and L2 were selected manually in Praat (Boersma and Weenink 2018). Tokens selected for analysis contained the pre-vocalic word-initial plosives /p t k b d ɡ/ in stressed position of monosyllabic and disyllabic content words. This resulted in a total of 3459 plosive tokens which were included in the analysis, with N = 224 German and N = 3235 English tokens. The number of tokens representing AS's L2 pronunciation was considerably higher compared to the number of test tokens obtained for his L1 German. The reason for this was that fewer interviews were available in which AS used his L1 German; as previously mentioned, AS shows a preference to speak his L2 also in L1-settings (e.g., von Uslar 2012).

VOT duration was measured between two manually defined boundaries, namely the burst of the plosive, indicated by a sharp peak in the waveform and a corresponding spectral change, and the start of waveform periodicity marking the onset of the following vowel. In some tokens, the release of the plosive could not be reliably identified due to noise or mumbled speech; these tokens (N = 42) were discarded from the analysis.

3.2. Results

Study 1 aimed to identify long-term changes in AS's L1 and L2 realization of VOT contrast in word-initial stressed plosives across different stages in time. To compare VOT durations obtained for AS's L1 and L2 plosives within and across languages, we ran a mixed Anova analysis in R (R Core Team 2020, version 3.6.6), using linear mixed-effects models. A linear mixed-effects approach was considered most appropriate for the present investigation because linear mixed models allow controlling for the potential influence of random effects and manage unbalanced data more easily than, for example, repeated measures Anovas (see, e.g., Barr et al. 2013).

The first model was built to examine changes in AS's VOT in his L1 German plosives across two stages, including VOT *duration* as the dependent variable and *stage* and *phoneme* as independent variables. An interaction between *stage* and *phoneme* was included as an additional fixed effect, and *word* was included as random intercept. The same specifications were applied to the second model which aimed to examine changes in AS's VOT in his L2 English over time, using English data only. The third model, which contained *language* as an additional independent variable, was built to compare AS's VOT in his two languages over time.

The lmerTest package (Kuznetsova et al. 2017) was used in R (R Core Team 2020), including the lmerTest function to obtain *p*-values for *t*-statistics. The mixed models were REML-fitted using Satterthwaite's approximations to estimate degrees of freedom. Throughout analysis, an α-level of 0.05 was adopted. Pairwise comparisons were conducted using Tukey's HSD.

3.2.1. L1 German Plosives

Figure 1 depicts the VOT durations obtained for AS's L1 German plosives in the early and the late stage (an overview of the descriptive statistics for AS's L1 and L2 plosives is provided in Table A1, Appendix A). It can be seen that—as expected—AS realized a VOT contrast in his early and late velar plosives by producing voiceless targets with considerably longer VOT compared to the voiced counterparts. At the same time, both his voiced and voiceless bilabial and alveolar plosives were produced within a short-lag VOT range in the early stage, that is, AS neutralized contrast in his early productions of these targets. In the late stage, however, a substantial lengthening of VOT was identified in his voiceless alveolar productions while maintaining short-lag VOT in his voiced targets, which shows that he produced a voiced–voiceless distinction in the late stage. Note, however, that for his late German productions of /p/ only a small number of test tokens (N = 3) could be identified in the speech samples.

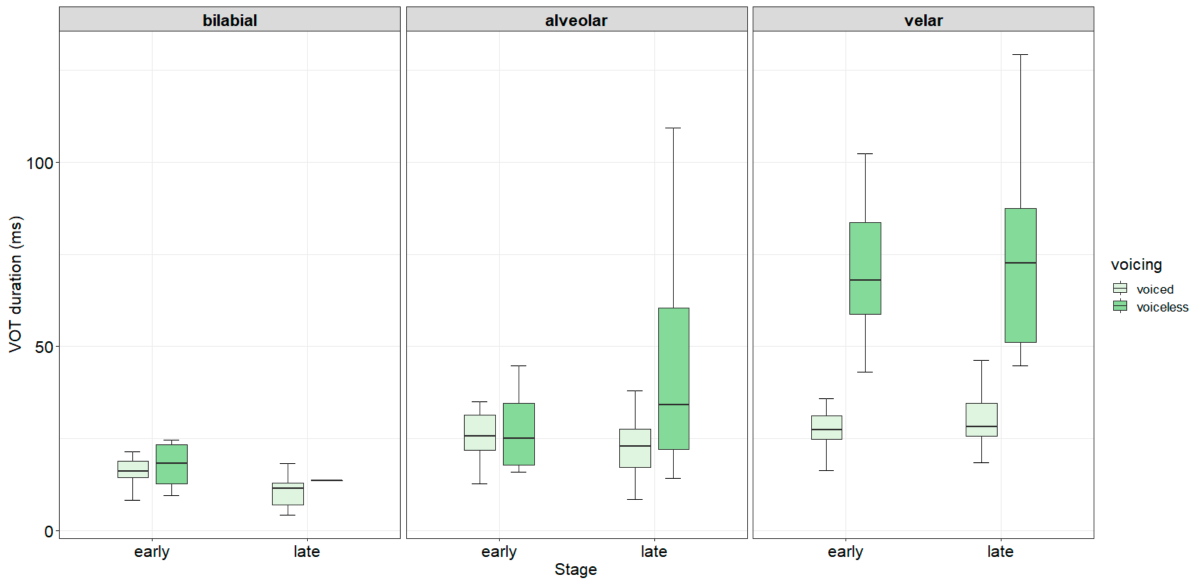

Figure 1. VOT durations (in milliseconds, ms) of AS's German plosives.

The results of the statistical analysis revealed a main effect for *phoneme* ($F[5115.4] = 75.6$, $p < 0.001$) and a significant interaction between *stage* and *phoneme* ($F[5194.6] = 2.4$, $p < 0.05$). Post hoc testing showed a significant difference in VOT duration obtained for AS's voiceless velar plosives in both the early ($t(139) = -9.75$, $p < 0.0001$) and the late ($t(153) = -11.48$, $p < 0.0001$) stage. By contrast, no significant differences in VOT duration were observed for his early bilabial and alveolar plosives, which suggests a neutralization of VOT contrast, as depicted in Figure 1. In terms of his late alveolar productions, however, the analysis revealed a significant difference in VOT duration between voiced and voiceless targets ($t(144) = -4.51$, $p < 0.001$), which indicates a shift in AS's alveolar plosives from neutralizing VOT contrast in the early stage to realizing contrast in the late stage.

3.2.2. L2 English Plosives

As depicted in Figure 2, AS showed a tendency to produce a VOT contrast in his L2 for all places of articulation and across all three stages. His voiced plosives covered a short-lag VOT range while his voiceless targets were predominantly produced with long-lag VOT values. However, the broad overlaps between VOT values obtained for his voiced and voiceless plosives suggests that AS did not consistently produce a distinct VOT contrast but was variable in his productions. Furthermore, Figure 2 suggests that his *late* voiceless bilabial and alveolar plosives are characterized by a reduced amount of aspiration—and thus a less native-like production—when comparing his mid and late voiceless bilabials, and his early and late voiceless alveolars.

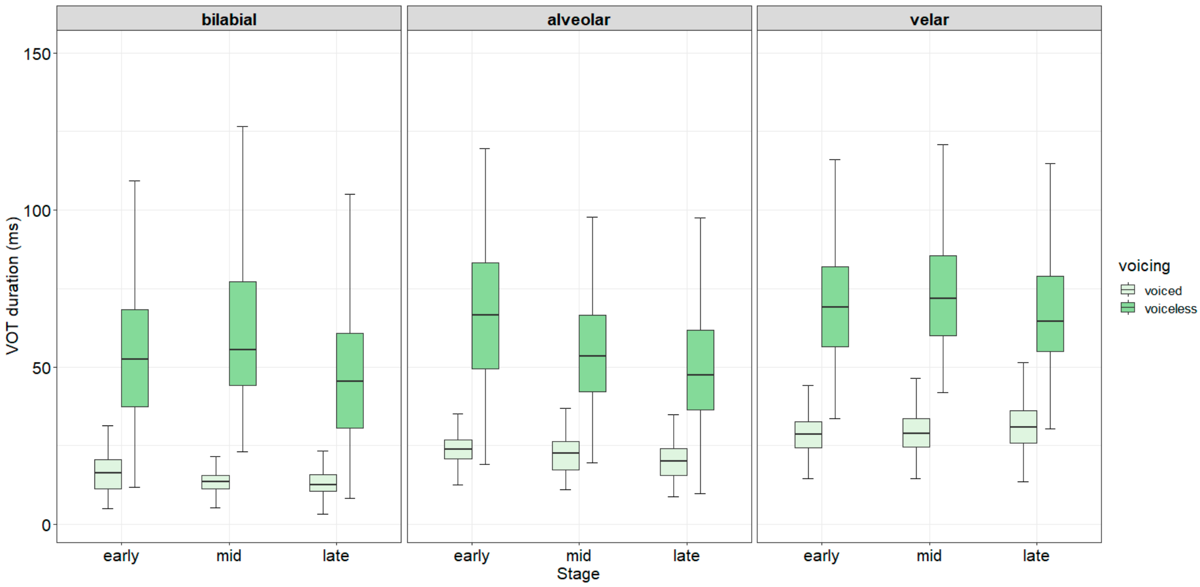

Figure 2. VOT durations (in ms) of AS's English plosives.

The analysis of AS's English VOT showed a main effect for *phoneme* ($F[5,328] = 296$, $p < 0.001$) and *stage* ($F[2,97] = 23.2$, $p < 0.001$), as well as significant interaction between *stage* and *phoneme* ($F[10,97.3] = 8.6$, $p < 0.001$). Post hoc Tukey tests revealed a significant difference in VOT duration for his late and mid voiceless bilabial plosives ($t(104.8) = -5.21$, $p < 0.0001$). Significant effects were also observed for his voiceless alveolar plosives in the early and the late ($t(63.5) = 9.14$, $p < 0.0001$) stage, and in the early and mid ($t(94.9) = 4.86$, $p < 0.001$) stage, which suggests that, as stated above, his voiceless bilabial and alveolar targets were significantly less aspirated in the late stage.

3.2.3. Comparison across Languages

Figure 3 compares AS's VOT in his L1 German and L2 English plosives in the early and the late stage. Cross-linguistic differences are most obvious in his early productions of the bilabial and alveolar targets: While he produced early English /p/ and /t/ with predominantly long-lag VOT, the same plosive targets were realized within a short-lag VOT range in his L1, which resulted in a neutralization of VOT contrast. In the late stage, he was observed to maintain a voiced–voiceless distinction for all L2 targets; at the same time, a tendency to realize a VOT contrast was also identified in his late German productions of the voiceless alveolar plosives, which were characterized by a lengthening of VOT duration in the direction of his L2.

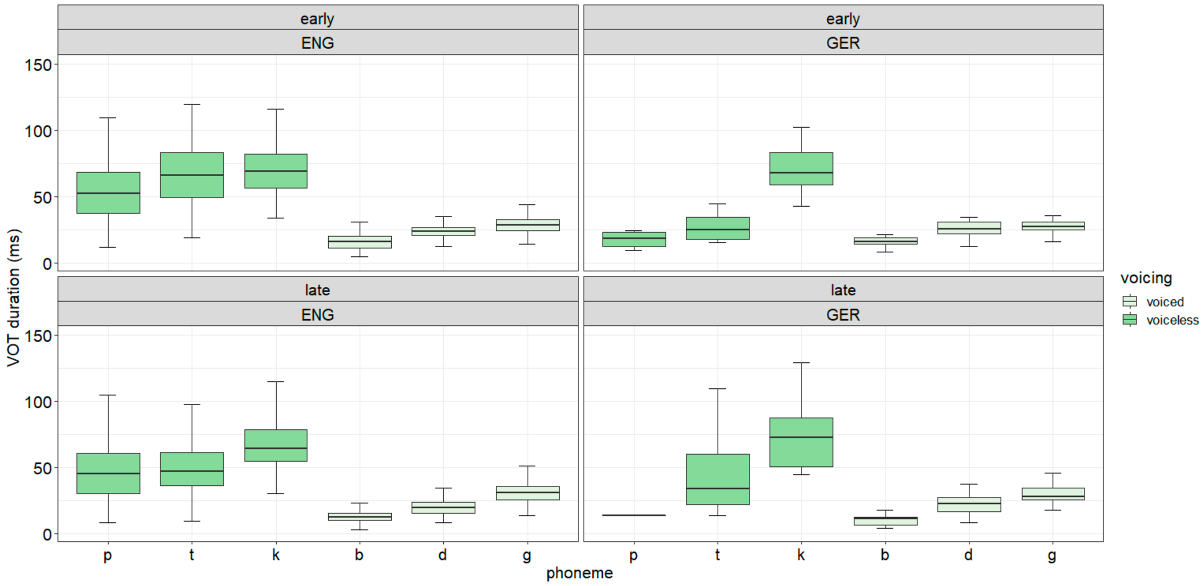

Figure 3. Comparison of VOT durations (in ms) of AS's English (ENG) and German (GER) plosives.

The statistical analysis of AS's L1 and L2 VOT durations showed significant interactions between *stage* and *language* ($F[1,328.9] = 4.34$, $p < 0.05$), *phoneme* and *language* ($F[5,255.3] = 12.3$, $p < 0.001$), as well as a three-way interaction between *stage*, *phoneme* and *language* ($F[5,330.6] = 4.6$, $p < 0.001$). Post hoc results showed a significant difference between AS's English and German early productions of /p/ ($t(195.1) = 5.96$, $p < 0.0001$) and /t/ ($t(117.3) = 8.15$, $p < 0.0001$), confirming that, as described above, he produced L1 and L2 voiceless bilabial and alveolar targets within different VOT ranges, respectively. In terms of his late productions, no significant effects were found for /p/ ($p = 1$) and /t/ ($p = 0.947$), which indicates that he realized both English and German targets within a long-lag VOT range.

3.3. Discussion

Study 1 aimed to determine if and to what extent AS's productions of L1 Austrian German and L2 English word-initial plosives have changed in the past 40 years of L2 immersion by comparing his L1 and L2 realization of VOT contrast across different stages in time.

The investigation of AS's early L1 German plosives showed a neutralization of VOT contrast in his bilabial and alveolar targets by producing both voiceless and phonologically voiced plosives with short-lag VOT, as commonly observed in Austrian German spontaneous speech (Hödl 2019; Moosmüller et al. 2015). At the same time, he maintained a distinct and native-like VOT contrast in his velar productions, with significantly longer VOT measures obtained for his voiceless velars. In his late German alveolar productions, a significant lengthening of VOT was identified, which indicates a shift of his L1 short-lag VOT categories towards English long-lag VOT. These results confirm the findings of previous studies showing that L2-immersed late bilinguals' voiceless categories are likely to be affected by attrition processes in the direction of L2 norms as a result of L2 learning experience (Flege and Hillenbrand 1984; Major 1992; Mayr et al. 2012; Stoehr et al. 2017; Sučková 2020).

While neutralizing contrast in his early L1 productions, AS showed a tendency to realize contrast in his L2 productions for all places of articulation and across all three stages. This might have been rather unexpected given that particularly in an early stage of L2 acquisition AS could have experienced L1 influences on the L2 resulting in an inability to produce English voiceless bilabial and alveolar targets with long-lag VOT, as typical of L1 Austrian German. It must be taken into consideration, however, that the recordings representing AS's early English productions were made approximately ten years after migrating to the US, i.e., he had already gained a considerable amount of L2 experience at this point. Moreover, the present investigation showed that his English voiceless plosives were characterized by considerable VOT variability, which suggests that he did not consistently maintain a distinct and native-like contrast between voiced and voiceless L2 targets. Variable L2 VOT productions are frequently observed among late bilinguals (e.g., Flege 1991; Hazan and Boulakia 1993), and can be attributed to different factors, such as diverse L2 input (de Leeuw et al. 2012a) or increased and recent L1 exposure through travelling to an L1-speaking country (de Leeuw 2019; Sancier and Fowler 1997). A further possible explanation for AS's variable L2 productions are potential articulatory constraints particularly affecting the acquisition of aspirated plosives which "require fine temporal coordination to delay the onset of laryngeal vibration relative to oral closure release" (Yu et al. 2015, p. 153). Given that also his late German productions of /t/ were observed to be variable, with some plosives falling in the short-lag VOT range and others in the long-lag range, articulatorily motivated difficulties might have impeded—at least to some extent—the production of consistently aspirated plosives.

Interestingly, AS's late English voiceless plosives were characterized by a significantly reduced amount of aspiration compared to his early and/or mid productions, where no overshooting of monolingual norms was evidenced, which suggests that his L2 productions have moved closer to L1 production norms and have thus become less native in the late stage. Although research shows that mean VOT durations for voiceless plosives typically decrease in elderly speakers as a result of physiological modifications of the vocal tract (e.g., Ryalls et al. 2004; Smith et al. 1987), such biological ageing mechanisms are not likely to have affected AS's late pronunciation of plosives given that biological effects would have resulted in a decrease of VOT in *both* languages. This is, however, not the case since the analysis of AS's late German voiceless targets revealed changes in the opposite direction, as manifested in an *increase* of VOT duration. Instead, the observed shortening of VOT in his late English voiceless plosives and an overall lengthening of VOT in his late German /t/ is indicative of a merging of L1 and L2 categories over time. This assimilation of L1 and L2 VOT categories, as identified in previous acoustic investigations of VOT (e.g., Flege 1987; Major 1992; Mayr et al. 2012), is therefore likely to be the result of cross-linguistic influences affecting pronunciation in both the L1 and the L2.

4. Study 2: Schwarzenegger's Vowels

4.1. Analysis

To examine AS's vowel space in his L1 and L2, monophthongs occurring in stressed position of monosyllabic and disyllabic content and function words were selected manually from the speech corpus described in Section 2.2 in Praat (Boersma and Weenink 2018). Table 1 depicts the English and German vowels included in the analysis (an overview of the number of tokens included for each vowel target is provided in Table A2, Appendix A). The German vowels /y/, /ʏ/, /ø/ and /œ/ were not included as not enough tokens could be identified in the German audio recordings. In terms of English /ɔ/ and /ɑ/, previous research suggests that speakers of Californian English—the variety AS is predominantly exposed to—tend to neutralize the contrast between these two vowels in the direction of /ɑ/ (e.g., Hagiwara 1997). They were, however, included as separate vowel targets in the present analysis to determine if AS produces distinct vowels. Due to the spontaneous nature of the speech analyzed in this study, vowel targets included in the analysis occurred in different consonantal contexts. Tokens preceded or followed by the approximants /w/, /l/, / ɹ/, or /j/ were excluded given that their acoustic properties are similar to those of vowels, which may impede the identification of exact measurement points (see Di Paolo et al. 2011).

Table 1. German and English vowels included in Study 2.

	German	English		German	English
Front			Back		
/i/	*bieten* "offer"	*bead*	/u/	*Mus* "pulp"	*booed*
/ɪ/	*bitten* "request"	*bid*	/ʊ/	*Bus* "bus"	*book*
/ɛ/	*Bett* "bed"	*bed*	/ʌ/	–	*bud*
/e/	*Beet* "flowerbed"	–	/ɔ/	*Motte* "moth"	*bawd*
/æ/	–	*bad*	/ɑ/	*satt* "sated"	*bod*
Central					
/ɝ/	–	*bird*			

The identification of vowel targets resulted in a total of *N* = 262 German and *N* = 2557 English vowels which were included in the analysis. In these test tokens, vowel onset and offset were marked manually in Praat (Boersma and Weenink 2018). In plosive, fricative, or affricate contexts, the first glottal striation was determined as the vowel onset, indicating the point where the spectral shape of the formants became visible. Given that the frequency of the first formant in nasal consonants is much lower compared to the F1 frequency of vowels (Ladefoged 2005), the onset and offset of vowels occurring in nasal contexts were marked at the points where the acoustic energy was rising and dropping, respectively.

After determining vowel onset and offset, the frequencies of the first and second formant were measured at the temporal mid-point of the vowel using linear predictive coding in Praat (Boersma and Weenink 2018) with a maximum formant frequency of 5000 Hertz (Hz), a window length of 0.025 s, and a pre-emphasis of 50 Hz. Burg's algorithm (Childers 1978) was used to extract formant frequencies. While some scholars suggest taking vowel formant measurements at multiple points of the vowel (Di Paolo et al. 2011), the vowel mid-point was chosen as the measurement point to reduce effects of co-articulation (e.g., Reubold et al. 2010).

4.2. Results

Study 2 set out to identify modifications in AS's L1 and L2 monophthongal vowel space across different stages in time by examining changes in F1 and F2 frequencies over time. Formant frequency measurements of F1 and F2 obtained for his German and English vowels were extracted from Praat (Boersma and Weenink 2018). As in Study 1, linear mixed-effects models were built in R (R Core Team 2020, version 3.6.6), for the same reasons outlined in Section 3.2. The first two models aimed to examine changes in F1 and F2 of AS's German and English monophthongs, including *F1* and *F2* as dependent variables, respectively, and *stage* and *phoneme* as well as an interaction between the two as fixed effects. *Word* was included as random intercept. The third model explored changes in F1 and F2 comparing AS's two languages over time and contained *language* as an additional independent variable, otherwise applying the same model specifications outlined above.

In R (R Core Team 2020), the lmerTest package (Kuznetsova et al. 2017) was used, including the lmerTest function to obtain *p*-values for *t*-statistics. The linear mixed models were REML-fitted using Satterthwaite's approximations to estimate degrees of freedom, adopting an α-level of 0.05 throughout. Tukey's HSD was used to conduct pairwise comparisons.

4.2.1. L1 German Vowels

Figure 4 depicts AS's L1 German vowel space in the early and the late stage (an overview of the descriptive statistics for AS's L1 and L2 vowels is provided in Table A2, Appendix A). An inspection of the figure shows that his late German vowels /i/, /e/, /u/ and /ɑ/ are characterized by a decrease in F1 compared to his early productions of the same targets, which indicates that these vowel targets have moved to a higher position in the late stage. The back vowel /ʊ/, by contrast, has moved to a more front position in the late stage, as manifested in considerably higher F2 values. A slight decrease in F1 and F2 can also be observed for AS's production of /ɔ/ in the late stage. Comparatively small changes are evident in AS's realizations of the front vowels /ɛ/ and /ɪ/.

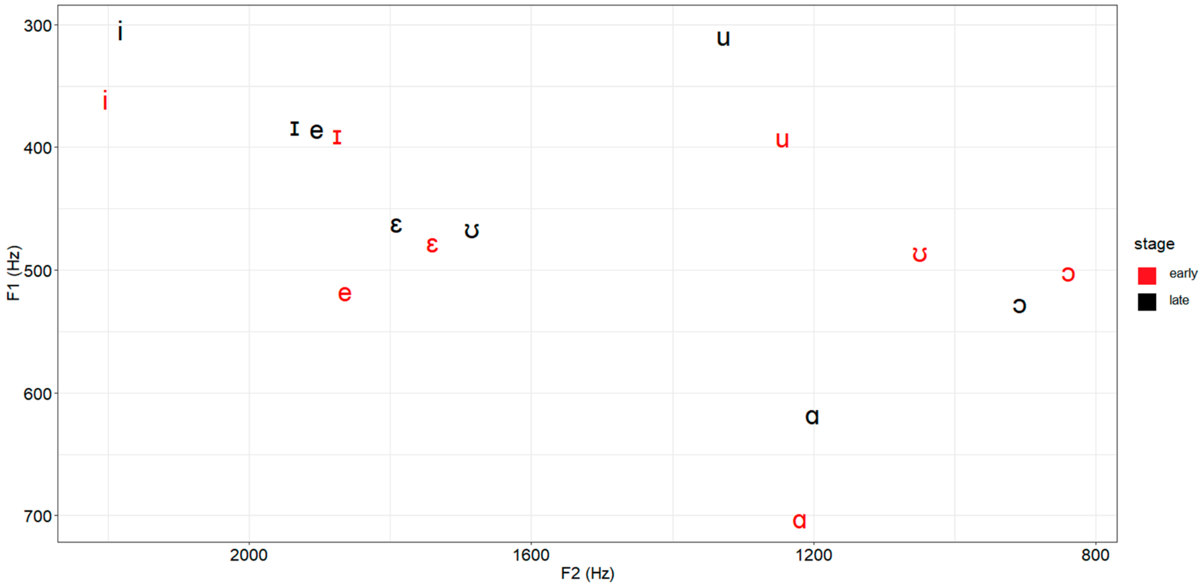

Figure 4. F1~F2 for AS's German vowels.

The statistical analysis of AS's German formant frequencies over time revealed significant effects for *stage* ($F[2,2380] = 8.5$, $p < 0.001$), *phoneme* ($F[9,523.5] = 250.74$, $p < 0.0001$), and an interaction between *stage* and *phoneme* ($F[18,2399.2] = 1.76$, $p < 0.01$). Post hoc results showed a significant difference in F1 for /ɑ/ ($t(246) = 3.77$, $p < 0.001$), /e/ ($t(231) = 3.47$, $p < 0.001$), and /u/ ($t(225) = 2.72$, $p < 0.007$) in the early and the late stage, which confirms that AS's late vowels are produced more close compared to his early vowels. In terms of F2, significant differences were observed for his early and late /ʊ/ ($t(245) = -8.03$, $p < 0.0001$), with considerably higher F2 values and, thus, a more front production in the late stage.

4.2.2. L2 English Vowels

Figure 5 displays AS's vowel space in his L2 English across three stages. All of his vowels, with the exception of /u/ and /ʌ/, have moved to a higher position in the late stage, as indicated by lower F1 values in the late compared to the early and/or mid stage. Additionally, a shift in F2 is observable in his productions of /u/ and /ɔ/, with overall lower F2 frequencies and hence a more back production in the late stage. It can also be seen that AS produces distinct vowels for English /ɑ/ and /ɔ/, that is, a merging of these two categories—as typical of Californian English (e.g., Boberg 2005)—is not evident.

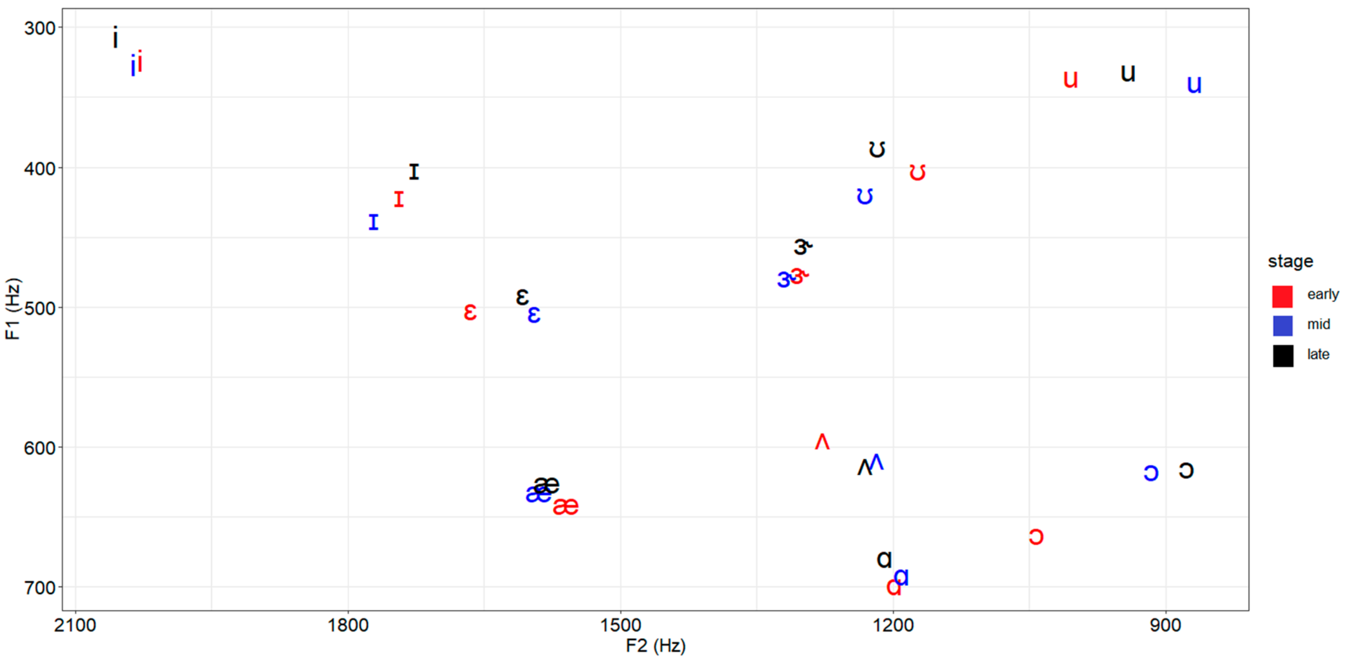

Figure 5. F1~F2 for AS's English vowels.

Results of the statistical analysis showed significant effects for *stage* ($F[2,2380] = 8.5$, $p < 0.001$), *phoneme* ($F[9,523.5] = 250.74$, $p < 0.0001$), and an interaction between *stage* and *phoneme* ($F[18,2399.2] = 1.76$, $p < 0.01$). Post hoc Tukey tests revealed a significant difference in F1 for AS's mid and late productions of the vowel /ɪ/ ($t(2347) = 3.58$, $p = 0.001$), and for his early and late production of /ɔ/ ($t(2324) = 2.55$, $p = 0.03$), with overall lower F1 values in the late stage. A significant increase of F2 in the late stage was identified for his early and late /i/ ($t(102.8) = -2.78$, $p = 0.022$), mid and late /i/ ($t(102.8) = -3.0$, $p = 0.009$), and early and late /ɔ/ ($t(85.2) = 4.36$, $p < 0.001$), indicating a shift towards a more front position. Significant effects were also found for F2 in his early and late productions of /ʌ/ ($t(216.4) = 2.51$, $p = 0.034$), i.e., his late realizations of this target vowel were characterized by a more back production as manifested in a decrease of F2.

4.2.3. Comparison across Languages

Figure 6 compares AS's L1 and L2 vowel space in the early and the late stage. Most notably, a shift of his German front vowel /i/ and German /ɔ/ closer towards the English targets can be identified in the late stage. Similarly, his German production of /ɑ/ has moved closer to his English production of /ʌ/ in the late stage, as manifested in a lowering of F1 for German /ɑ/. By contrast, German and English /ɛ/ and /ɪ/ have moved further apart in the late stage; the same is true for German /u/ and English /ʊ/, which are nearly identical in the early stage. In the late stage, however, German /u/ is characterized by considerably lower F1 values and has thus shifted away from the English target.

The statistical analysis conducted to compare F1 and F2 across languages and stages revealed a significant effect for *stage* ($F[1,277.05] = 20.33$, $p < 0.001$) and *phoneme* ($F[6,420.38] = 260.04$, $p < 0.001$), as well as an interaction between *phoneme* and *language* ($F[6,422.54] = 7.49$, $p < 0.001$), and a three-way interaction between *stage*, *phoneme* and *language* ($F[6,260.26] = 2.37$, $p = 0.03$). In the post hoc analysis, significant F1 differences between German and English were identified for AS's late productions of /ɑ/ ($t(362.7) = 5.1$, $p < 0.001$) and /ɔ/ ($t(339) = 3.92$, $p = 0.028$), as well as for early /ɔ/ ($t(258.6) = 5.22$, $p < 0.001$). Significant differences in F2 were identified for AS's late English and German realization of the target vowels /ɛ/ ($t(574.8) = -3.91$, $p = 0.027$), /ɪ/ ($t(585.1) = -4.6$, $p = 0.0017$), /u/ ($t(574.9) = -5.55$, $p < 0.001$), and /ʊ/ ($t(686.7) = -6.99$, $p < 0.001$), as manifested in overall lower F2 values identified in his late English productions of these target vowels compared to his German productions.

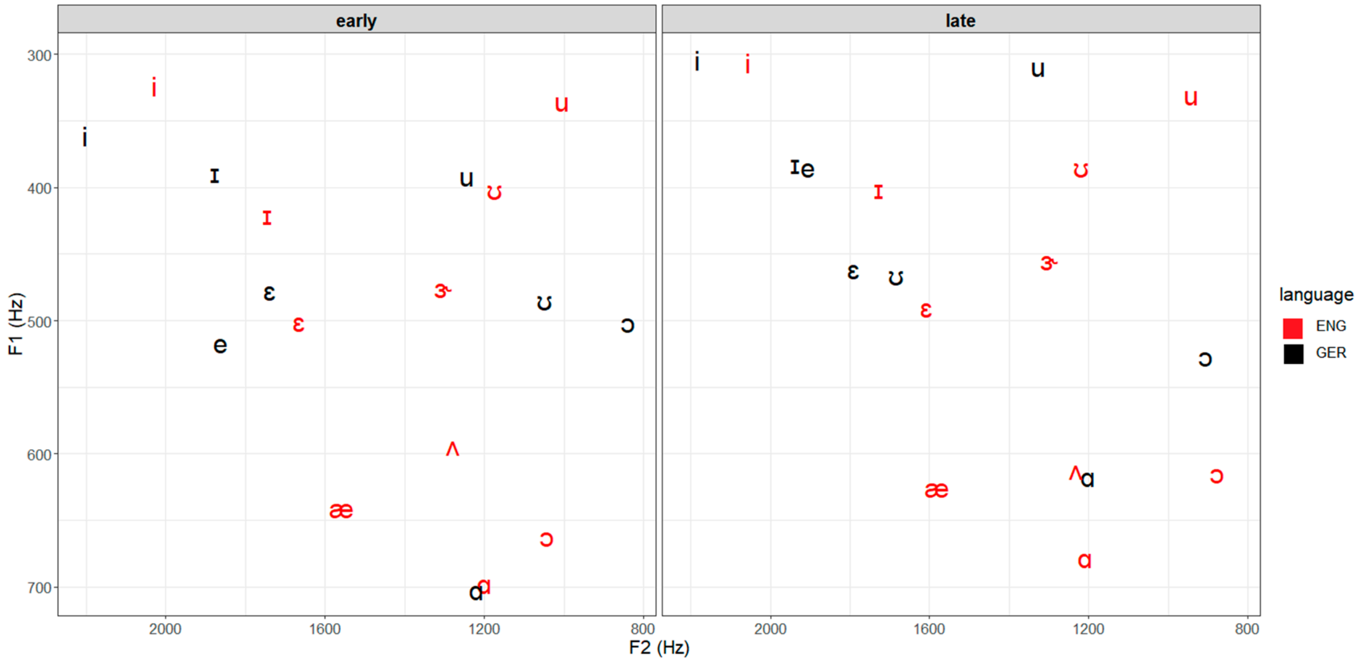

Figure 6. F1~F2 for AS's German (GER) and English (ENG) vowels.

4.3. Discussion

Study 2 set out to identify potential modifications in AS's L1 Austrian German and L2 English vowel space across three stages in time. To this end, F1 and F2 of eight German and ten English monophthongs were acoustically examined.

As the analysis revealed, two of AS's L1 vowels, i.e., /i/ and /ɑ/, were affected by a shift in the direction of the L2, that is, they came to resemble related L2 targets in the late stage, suggesting an influence of the L2 vowel system on the L1. In the case of English /ɔ/, considerable changes in both F1 and F2 were observed in the late stage, with an approximation of the L2 English target closer to the L1 German counterpart. At the same time, some of AS's L1 vowels, i.e., /u/, /ɛ/, and/ɪ/, showed changes in the opposite direction, that is, they have moved further away from English targets in the late stage, which might reflect an attempt to enhance contrast between closely related L1 and L2 vowel categories.

Overall, the investigation of AS's vowels in both of his languages revealed rather diverse modification patterns when comparing his early and late productions, showing that some of his vowels have changed considerably while others exhibited subtle modifications only. In addition, the direction of change was not uniform, with an approximation of L1 and L2 vowel categories, a dispersion of related L1 and L2 categories, and a shift of L2 vowels towards related L1 counterparts. These findings are in line with previous observations concerning the selective nature of L2-induced changes (Bergmann et al. 2016; de Leeuw 2019; Mayr et al. 2012). That is, not all speakers experience modifications in their L1 system to the same extent (Major 1992; Mennen 2004) and even within the same sound class, such modifications are not all-encompassing, as demonstrated in the present investigation. Similarly, in terms of the acquisition of L2 vowel categories, previous research suggests that speakers acquire different L2 vowels with varying degrees of success, often depending on whether an L2 vowel target has a perceptually similar counterpart in the L1 (e.g., Baker and Trofimovich 2005).

Again, one could argue that the changes observed in AS's vowel productions are related to natural ageing mechanisms, which have been shown to lead to a decrease in F1 frequency in elderly speakers (e.g., Reubold and Harrington 2015, 2017). In fact, an overall decrease in F1 was also identified in some of AS's German and English vowels in the late stage. However, as addressed in Section 3.3, changes resulting from biological ageing processes are not expected to be selective, that is, to have an effect on some vowels only. Hence, the influence of biological mechanisms does not offer a convincing explanation for the modifications observed in AS's German and English vowels. Instead, the changes in AS's vowel productions over time seem to be indicative of complex system-internal mechanisms, which result in various modification patterns. The specific reasons, however, why some vowel categories are affected differently or to a greater extent than others are still to be uncovered.

In order to gain further insight into the extent to which AS's vowels have been affected by bi-directional L1–L2 influences, it would be interesting for future research to also explore the compactness of his vowels over time to identify if the acoustic stability of his vowel targets has changed in the course of L2 immersion (see, e.g., Kartushina et al. 2016; Kartushina and Martin 2019).

5. Overall Discussion

The aim of the present investigation was to trace developments in the L1 and L2 segmental speech production in the late consecutive bilingual Arnold Schwarzenegger over a period of four decades. The findings of Study 1, focusing on AS's realization of VOT contrast in German and English plosives, revealed an assimilation of his late L1 and L2 categories for voiceless targets, which was particularly visible in his alveolar plosives. That is, his short-lag L1 productions of /t/ were characterized by significantly longer and thus more English-like VOT durations in the late stage. At the same time, his late L2 targets were produced with considerably less aspiration, which suggests a drift away from native English production norms closer to L1 short-lag norms. The findings of Study 2, exploring AS's L1 and L2 monophthongal vowel space, to some extent reflect the results of Study 1. In both studies, the changes observed cannot be reliably explained against the background of age-related biological mechanisms, and both investigations provide evidence for a merging of L1 and L2 categories in AS's late productions, which supports one of the main tenets of the Speech Learning Model that closely related L1 and L2 categories may come to resemble each other due to assimilatory effects (Flege 1995; Flege and Bohn 2020). Study 2 further showed that a bilingual's segmental productions are not necessarily equally affected by L2-induced changes (Mayr et al. 2012) and that some changes are relatively subtle (Bergmann et al. 2016; Chang 2012).

Taken together, the present findings confirm that "cross-linguistic transfer is not a one-way street" (Schmid and Köpke 2017, p. 637) in that the use and development of a late-acquired L2 system can indeed exert influence on a speaker's L1 accent, in the same way as the L1 system affects pronunciation abilities in the L2. In this respect, the findings contradict the notion of an impermeable, invariable L1 system and thus challenge a static view on bilingualism (Lado 1957; Lenneberg 1967). Instead, they support a dynamic systems-oriented approach to language development, according to which mutual L1–L2 interactions, sensitivity to system-internal and external influences, and sometimes even unpredictability are inherent characteristics of bilingual development (e.g., de Bot and Larsen-Freeman 2011). This approach considers intra-individual variability in bilingual productions as valuable evidence for the dynamic nature of linguistic development, arguing that "both free and systematic variability will be relatively high when the system is reorganizing" (Verspoor et al. 2008, p. 216). As outlined above, a restructuring and reorganization of a speaker's pronunciation system(s) might be internally motivated, that is, resulting from L1–L2 interactions over time, but also external—social, environmental, and personal—factors can trigger and shape such modification processes (de Bot et al. 2007; Verspoor et al. 2008).

While our findings are indicative of L1–L2 changes that are most likely internally driven, it might be argued that external factors may have also played a role. For instance, AS's pronunciation may have been affected by the interview situation leading to stress-induced changes to his pronunciation and a possible adaption to the accent, speaking style and other linguistic features of his interlocutors (e.g., Giles and Ogay 2007). However, we expect this influence to be minimal given that over the years AS has had ample experience conducting interviews and he had been exposed to the speech of many different interviewers, which will have reduced the adaption to an individual's pronunciation. Another disadvantage of the use of spontaneous speech samples is that it is likely to entail variations in speaking rate, which in turn can influence features of pronunciation, including VOT duration and vowel production (Kessinger and Blumstein 1998). However, controlling for speaking rate has also been found to be problematic in experimental settings given that speakers often have different perceptions of what fast and slow speech is (Harrington 2010). Another possible external factor that might have influenced the results is AS's recent and enhanced L1 exposure through travelling to his home country Austria, which has previously been shown to have an impact on a speaker's accent (Sancier and Fowler 1997). Furthermore, potential changes in his private and professional environment might have contributed to the changes and variability observed in his L1 and L2 speech production (see, e.g., Schoonmaker-Gates 2015).

Despite the limitations of using spontaneous speech data outlined above, the present investigation offers a rare insight into the longitudinal development of a late bilingual's L1 and L2 speech production over a period of 40 years. The investigation thus offers useful insights into the dynamic nature of bilingual speech development and sheds light on the complexity of L1–L2 interaction processes affecting a speaker's pronunciation. To gradually comprehend this complexity and to fully understand how and to what extent system-internal processes are intertwined with social and environmental factors, further long-term investigations focusing on additional segmental and prosodic variables in both individuals and groups of speakers will be necessary.

Author Contributions: Conceptualization, L.K. and I.M.; methodology, L.K. and I.M.; formal analysis, L.K.; writing—original draft preparation, L.K. and I.M.; writing—review and editing, I.M. and L.K.; visualization, L.K.; supervision, I.M. All authors have read and agreed to the published version of the manuscript.

Funding: Open Access Funding by the University of Graz.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Ethics Committee of the University of Graz (protocol code: GZ. 39/1/63 ex 2019/20; date of approval: 12 November 2019).

Informed Consent Statement: Participant consent was waived due to the use and analysis of publicly available speech material. Using the materials for scientific purposes did not entail copyright issues according to the Austrian Copyright Act (UrhG) §42(2).

Data Availability Statement: The data are available on request from the corresponding author. The data are not publicly available due to ongoing data analyses. Data will be made available once all analyses have been completed.

Acknowledgments: We would like to thank Robert Mayr and the two anonymous reviewers for their helpful comments and feedback. We also thank Ulrich Reubold for his valuable statistical advice, and Kerstin Endes and Matthias Wedenig for their assistance during the data labelling process.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Means, standard deviations and medians for AS's VOT durations (reported in milliseconds) obtained for his English and German plosives across different stages in time, including the number of tokens (*N*) measured for each plosive according to language and stage.

English early				**English mid**		
Plosive	*N*	*Mean (SD)*	*Median*	*N*	*Mean (SD)*	*Median*
p	69	54.05 (23.33)	52.47	76	61.36 (25.05)	55.34
b	90	15.98 (5.9)	16.4	123	13.81 (3.93)	13.46
t	117	67.58 (22.89)	66.33	103	56.4 (20.13)	53.42
d	97	24.01 (6.02)	23.72	118	22.17 (5.74)	22.67
k	89	69.8 (18.04)	68.96	103	72.75 (17.27)	71.64
ɡ	118	29.35 (7.42)	28.49	100	29.75 (7.3)	28.87
English late				**German early**		
Plosive	*N*	*Mean (SD)*	*Median*	*N*	*Mean (SD)*	*Median*
p	313	48.22 (24.36)	45.46	4	17.62 (7.27)	18.21
b	403	13.32 (4.7)	12.49	16	16.49 (4.64)	16.07
t	373	50.07 (20.46)	47.42	8	27.46 (11.4)	24.96
d	266	20.83 (7.09)	20.06	15	26.13 (6.29)	25.54
k	287	68.33 (19.37)	64.53	24	70.78 (16.68)	67.97
ɡ	390	31.13 (7.62)	30.85	27	27.70 (4.89)	27.3
German late						
Plosive	*N*	*Mean (SD)*	*Median*			
p	3	16.81 (6.55)	13.6			
b	25	11.16 (4.85)	11.27			
t	12	44.67 (29.54)	34.08			
d	45	22.71 (7.62)	22.76			
k	22	73.06 (22.15)	72.56			
ɡ	25	29.84 (6.28)	28.22			

Table A2. Means, standard deviations and medians of F1 and F2 (reported in Hertz) obtained for AS's English and German monophthongs across different stages in time, including the number of tokens (N) measured for each monophthong according to language and stage.

English early					
		F1		F2	
Vowel	*N*	*Mean (SD)*	*Median*	*Mean (SD)*	*Median*
/i/	98	323.6 (60.24)	305.1	2028.1 (135.9)	2018.0
/ɪ/	130	420.2 (69.66)	410.3	1743.2 (118.3)	1752.6
/ɛ/	105	500.6 (77.41)	485.4	1665.2 (188.0)	1667.9
/u/	35	334.9 (45.38)	328.5	1003.6 (267.0)	950.8
/ʊ/	35	400.9 (68.98)	403.5	1172.9 (211.76)	1181.7
/ɔ/	14	661.5 (69.7)	651.1	1042.2 (209.91)	975.1
/ɑ/	84	697.2 (83.84)	707.0	1197.5 (109.31)	1182.6

Table A2. *Cont.*

/æ/	98	639.9 (99.28)	650.9	1559.4 (207.47)	1545.2
/ʌ/	86	593.1 (106.16)	601.7	1277.9 (104.57)	1265.8
/ɝ/	37	474.7 (43.34)	462.7	1302.5 (76.26)	1300.9
English mid					
		F1		F2	
Vowel	*N*	*Mean (SD)*	*Median*	*Mean (SD)*	*Median*
/i/	74	326.3 (43.04)	320.0	2035.5 (252.94)	2041.8
/ɪ/	120	436.4 (77.18)	429.1	1771.1 (100.58)	1779.0
/ɛ/	61	502.7 (51.58)	503.3	1595.3 (142.56)	1599.6
/u/	24	338.5 (43.05)	340.3	867.8 (246.98)	778.9
/ʊ/	21	418.0 (50.3)	412.6	1231.2 (209.24)	1236.7
/ɔ/	7	615.4 (71.73)	594.9	915.4 (78.48)	936.1
/ɑ/	69	690.0 (55.31)	696.6	1190.3 (129.17)	1191.7
/æ/	101	630.6 (91.12)	659.9	1589.1 (194.24)	1534.0
/ʌ/	92	607.8 (52.27)	610.3	1218.2 (82.09)	1228.8
/ɝ/	39	477.1 (42.31)	476.3	1317.4 (74.51)	1323.9
English late					
		F1		F2	
Vowel	*N*	*Mean (SD)*	*Median*	*Mean (SD)*	*Median*
/i/	111	306.0 (41.32)	301.9	2054.8 (117.79)	2040.1
/ɪ/	217	400.3 (60.58)	390.5	1727.2 (125.9)	1737.4
/ɛ/	192	489.8 (52.2)	485.9	1607.9 (115.03)	1598.78
/u/	39	330.0 (75.71)	317.9	940.6 (438.66)	799.6
/ʊ/	32	384.6 (39.41)	381.5	1217.6 (257.99)	1272.8
/ɔ/	32	614.1 (78.2)	602.3	876.5 (73.74)	882.4
/ɑ/	108	677.8 (73.99)	681.2	1208.2 (140.06)	1193.6
/æ/	219	624.5 (99.86)	648.2	1580.6 (169.3)	1535.2
/ʌ/	203	611.1 (57.01)	609.0	1231.2 (88.65)	1229.4
/ɝ/	74	454.0 (36.82)	448.1	1298.6 (85.71)	1292.07
German early					
		F1		F2	
Vowel	*N*	*Mean (SD)*	*Median*	*Mean (SD)*	*Median*
/i/	7	360.8 (34.69)	352.7	2203.3 (44.15)	2198.1
/ɪ/	23	388.7 (59.77)	404.0	1875.7 (145.09)	1879.3
/ɛ/	15	476.8 (66.88)	437.4	1739.8 (107.46)	1722.6
/u/	8	391.3 (27.85)	381.4	1242.7 (366.13)	1122.3
/ʊ/	7	484.5 (181.51)	461.2	1048.4 (514.96)	880.4
/ɔ/	32	500.6 (55.08)	490.4	837.6 (52.06)	823.8
/ɑ/	20	701.6 (76.09)	679.7	1218.8 (135.32)	1192.5
/e/	6	516.6 (14.59)	518.4	1863.5 (115.1)	1852.4

Table A2. *Cont.*

German late					
		F1		F2	
Vowel	*N*	*Mean (SD)*	*Median*	*Mean (SD)*	*Median*
/i/	9	303.8 (16.14)	302.1	2182.1 (58.89)	2199.8
/ɪ/	31	381.4 (35.03)	387.0	1935.5 (110.93)	1964.2
/ɛ/	23	460.7 (30.86)	469.3	1791.9 (122.85)	1799.1
/u/	10	308.7 (27.44)	311.9	1326.7 (227.47)	1398.7
/ʊ/	6	465.2 (91.02)	483.1	1684.3 (129.35)	1645.0
/ɔ/	21	526.1 (52.32)	534.8	906.5 (118.46)	865.2
/ɑ/	39	617.1 (76.84)	598.5	1201.0 (99.85)	1209.4
/e/	5	384.3 (18.81)	390.9	1904.1 (43.14)	1908.8

References

Abramson, Arthur S., and D. H. Whalen. 2017. Voice Onset Time (VOT) At 50: Theoretical and Practical Issues in Measuring Voicing Distinctions. *Journal of Phonetics* 63: 75–86. [CrossRef]

Allen, Nick. 2011. Arnold Schwarzenegger: From Austrian Immigrant to Superstar to Politician. Available online: https://www.telegraph.co.uk/culture/tvandradio/8419751/Arnold-Schwarzenegger-from-Austrian-immigrant-to-superstar-to-politician.html (accessed on 30 September 2020).

Baker, Wendy, and Pavel Trofimovich. 2005. Interaction of Native- and Second-Language Vowel System(S) In Early and Late Bilinguals. *Language and Speech* 48: 1–27. [CrossRef]

Barr, Dale J., Roger Levy, Christoph Scheepers, and Harry J. Tily. 2013. Random Effects Structure for Confirmatory Hypothesis Testing: Keep It Maximal. *Journal of Memory and Language* 68: 255–78. [CrossRef] [PubMed]

Bergmann, Christopher, Amber Nota, Simone A. Sprenger, and Monika S. Schmid. 2016. L2 Immersion Causes Non-Native-Like L1 Pronunciation in German Attriters. *Journal of Phonetics* 58: 71–86. [CrossRef]

Birdsong, David. 2009. Age and the End State of Second Language Acquisition. In *The New Handbook of Second Language Acquisition*. Edited by William C. Ritchie and Tej K. Bhatia. Amsterdam: Elsevier, pp. 401–24.

Bloomfield, Leonard. 1933. *Language*. New York: Holt.

Boberg, Charles. 2005. The Canadian Shift in Montreal. *Language Variation and Change* 17: 133–54. [CrossRef]

Boersma, Paul, and David Weenink. 2018. Praat: Doing Phonetics by Computer (Version 6.0.37). Computer Program. Available online: http://www.fon.hum.uva.nl/praat (accessed on 14 March 2018).

Bongaerts, Theo, Susan Mennen, and Frans van der Slik. 2000. Authenticity of Pronunciation in Naturalistic Second Language Acquisition: The Case of Very Advanced Late Learners of Dutch as a Second Language. *Studia Linguistica* 54: 298–308. [CrossRef]

Braunschweiler, Norbert. 1997. Integrated Cues of Voicing and Vowel Length in German: A Production Study. *Language and Speech* 40: 353–76. [CrossRef]

Bylund, Emanuel. 2009. Maturational Constraints and First Language Attrition. *Language Learning* 59: 687–715. [CrossRef]

Chang, Charles B. 2012. Rapid and Multifaceted Effects of Second-Language Learning on First-Language Speech Production. *Journal of Phonetics* 40: 249–68. [CrossRef]

Chang, Charles B. 2013. A Novelty Effect in Phonetic Drift of the Native Language. *Journal of Phonetics* 41: 520–33. [CrossRef]

Chang, Charles B. 2019. Language Change and Linguistic Inquiry in a World of Multicompetence: Sustained Phonetic Drift and Its Implications for Behavioral Linguistic Research. *Journal of Phonetics* 74: 96–113. [CrossRef]

Childers, Donald G. 1978. *Modern Spectrum Analysis*. New York: IEEE Press.

Collins, Ben. 2019. This Viral Schwarzenegger Deepfake Isn't Just Entertaining. It's a Warning. Available online: https://www.nbcnews.com/tech/tech-news/viral-schwarzenegger-deepfake-isn-t-just-entertaining-it-s-warning-n1016851 (accessed on 25 January 2021).

Daily Mail UK. 2015. Arnold Schwarzenegger Reveals He Can Speak Perfect English—But Keeps Talking with an Accent Because 'Fans Expect It'. Available online: http://www.dailymail.co.uk/tvshowbiz/article-3141778/Arnold-Schwarzenegger-reveals-speak-perfect-English-keeps-talking-accent-fans-expect-it.html (accessed on 25 January 2021).

de Bot, Kees, and Diane Larsen-Freeman. 2011. Researching Second Language Development from a Dynamic Systems Theory Perspective. In *A Dynamic Approach to Second Language Development: Methods and Techniques*. Edited by Marjolijn Verspoor and Kees de Botand Wander Lowie. Amsterdam: John Benjamins, pp. 5–24.

de Bot, Kees, Wander Lowie, and Marjolijn Verspoor. 2007. A Dynamic Systems Theory Approach to Second Language Acquisition. *Bilingualism: Language and Cognition* 10: 7–21. [CrossRef] de Leeuw, Esther. 2019. Native Speech Plasticity in the German-English Late Bilingual Stefanie Graf: A Longitudinal Case Study over Four Decades. *Journal of Phonetics* 73: 24–39. [CrossRef]

de Leeuw, Esther, Monika S. Schmid, and Ineke Mennen. 2010. The Effects of Contact on Native Language Pronunciation in an L2 Migrant Setting. *Bilingualism: Language and Cognition* 13: 33–40. [CrossRef]

de Leeuw, Esther, Ineke Mennen, and James M. Scobbie. 2012a. Dynamic Systems, Maturational Constraints, and L1 Phonetic Attrition. *International Journal of Bilingualism* 17: 683–700. [CrossRef]

de Leeuw, Esther, Ineke Mennen, and James M. Scobbie. 2012b. Singing a Different Tune in Your Native Language: First Language Attrition of Prosody. *International Journal of Bilingualism* 16: 101–16. [CrossRef]

de Leeuw, Esther, Aurela Tusha, and Monika S. Schmid. 2017. Individual Phonological Attrition in Albanian-English Late Bilinguals. *Bilingualism: Language and Cognition* 21: 278–95. [CrossRef]

Di Paolo, Marianna, Malcah Yaeger-Dror, and Alicia Beckford Wassnik. 2011. Analyzing Vowels. In *Sociophonetics: A Student's Guide*. Edited by Marianna Di Paolo and Malcah Yaeger-Dror. New York: Routledge, pp. 87–106.

Dmitrieva, Olga, Allard Jongman, and Joan Sereno. 2010. Phonological Neutralization by Native and Non-Native Speakers: The Case of Russian Final Devoicing. *Journal of Phonetics* 38: 483–92. [CrossRef]

Docherty, Gerard J. 1992. *The Timing of Voicing in British English Obstruents*. Berlin and New York: Foris.

Eckert, Penelope, and Norma Mendoza-Denton. 2006. Getting Real in the Golden State (California). In *American Voices: How Dialects Differ from Coast to Coast*. Edited by Walt Wolfram and Ben Ward. Malden: Blackwell Publishing, pp. 139–43.

Flege, James E. 1980. Phonetic Approximation in Second Language Acquisition. *Language Learning* 30: 117–34. [CrossRef]

Flege, James E. 1981. The Phonological Basis of Foreign Accent: A Hypothesis. *TESOL Quarterly* 15: 443–55. [CrossRef]

Flege, James E. 1987. The Production of 'New' and 'Similar' Phones in a Foreign Language: Evidence for the Effect of Equivalence Classification. *Journal of Phonetics* 15: 47–65. [CrossRef]

Flege, James 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: 395–411. [CrossRef] [PubMed]

Flege, James E. 1995. Second Language Speech Learning: Theory, Findings, and Problems. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Edited by Winifred Strange. Baltimore: York Press, pp. 233–72.

Flege, James E., and Ocke-Schwen Bohn. 2020. The Revised Speech Learning Model. Unpublished Preprint. Available online: https://www.researchgate.net/publication/348265785_The_revised_Speech_Learning_Model_SLM-r (accessed on 28 September 2020).

Flege, James E., and W. S. Brown. 1982. The Voicing Contrast Between English/p/and/b/as a Function of Stress and Position-in-Utterance. *Journal of Phonetics* 10: 335–45. [CrossRef]

Flege, James E., and Wieke Eefting. 1988. Imitation of a VOT Continuum by Native Speakers of English and Spanish: Evidence for Phonetic Category Formation. *The Journal of the Acoustical Society of America* 83: 729–40. [CrossRef] [PubMed]

Flege, James E., and James Hillenbrand. 1984. Limits on Phonetic Accuracy in Foreign Language Speech Production. *The Journal of the Acoustical Society of America* 76: 708–21. [CrossRef]

Flege, James E., Murray J. Munro, and Ian R. A. MacKay. 1996. Factors Affecting the Production of Word-Initial Consonants in a Second Language. In *Second Language Acquisition and Linguistic Variation*. Edited by Robert Bayley and Dennis R. Preston. Amsterdam: John Benjamins, pp. 47–74.

Flege, James E., Ocke-Schwen Bohn, and Sunyoung Jang. 1997. Effects of Experience on Non-Native Speakers' Production and Perception of English Vowels. *Journal of Phonetics* 25: 437–70. [CrossRef]

Flege, James E., Carlo Schirru, and Ian R. A. MacKay. 2003. Interaction between the Native and Second Language Phonetic Subsystems. *Speech Communication* 40: 467–91. [CrossRef]

Gala. 2015. Der 'Terminator'-Star Ist Zurück. Available online: https://www.gala.de/stars/news/interview/arnold-schwarzenegger--der--terminator--star-ist-zurueck-20234970.html (accessed on 25 January 2021).

Gersemann, Olaf. 2009. Der Mann, Der Sich Selbst Erfunden Hat. Available online: http://www.welt.de/politik/ausland/article4895817/Der-Mann-der-sich-selbst-erfunden-hat.html (accessed on 25 January 2021).

Giles, Howard, and Tania Ogay. 2007. Communication Accommodation Theory. In *Explaining Communication: Contemporary Theories and Exemplars*. Edited by Bryan B. Whaley and Wendy Samter. Mahwah: Lawrence Erlbaum, pp. 293–310.

Glaister, Dan. 2006. California Split over Teaching: English Classes for Hispanic Students Branded as 'Return to Segregation'. Available online: https://www.theguardian.com/education/2006/aug/11/tefl.danglaister (accessed on 25 January 2021).

Grosjean, François. 1989. Neurolinguists, Beware! The Bilingual Is Not Two Monolinguals in One Person. *Brain and Language* 36: 3–15. [CrossRef]

Grosjean, François. 1997. The Bilingual Individual. *Interpreting* 2: 163–87. [CrossRef]

Hagiwara, Robert. 1997. Dialect Variation and Formant Frequency: The American English Vowels Revisited. *The Journal of the Acoustical Society of America* 102: 655–58. [CrossRef]

Harrington, Jonathan. 2010. *Phonetic Analysis of Speech Corpora*. Malden: Wiley-Blackwell.

Hazan, Valerie L., and Georges Boulakia. 1993. Perception and Production of a Voicing Contrast by French-English Bilinguals. *Language and Speech* 36: 17–38. [CrossRef]

Hödl, Petra. 2019. Production and Perception of Voice Onset Time in Austrian German. Ph.D. thesis, Karl-Franzens-University, Graz, Austria.

Hopp, Holger, and Monika S. Schmid. 2013. Perceived Foreign Accent in First Language Attrition and Second Language Acquisition: The Impact of Age of Acquisition and Bilingualism. *Applied Psycholinguistics* 34: 361–94. [CrossRef]

Jackson, Betty. 2020. Does Arnold Schwarzenegger Speak German? Available online: https://celebanswers.com/does-arnold-schwarzenegger-speak-german/ (accessed on 25 January 2021).

Jessen, Michael. 1998. *Phonetics and Phonology of Tense and Lax Obstruents in German*. Studies in Functional and Structural Linguistics 44. Amsterdam: John Benjamins.

Kartushina, Natalia, and Clara D. Martin. 2019. Third-Language Learning Affects Bilinguals' Production in Both Their Native Languages: A Longitudinal Study of Dynamic Changes in L1, L2 and L2 Vowel Production. *Journal of Phonetics* 77: 100920. [CrossRef]

Kartushina, Natalia, Alexis Hervais-Adelman, Ulrich H. Frauenfelder, and Narly Golestani. 2016. Mutual Influences Between Native and Non-Native Vowels in Production: Evidence from Short-Term Visual Articulatory Feedback Training. *Journal of Phonetics* 57: 21–39. [CrossRef]

Keating, Patricia A. 1984. Phonetic and Phonological Representation of Stop Consonant Voicing. *Language* 60: 286–319. [CrossRef]

Kessinger, Rachel H., and Sheila E. Blumstein. 1998. Effects of Speaking Rate on Voice-Onset Time and Vowel Production: Some Implications for Perception Studies. *Journal of Phonetics* 26: 117–28. [CrossRef]

Kisler, Thomas, Uwe D. Reichel, and Florian Schiel. 2017. Multilingual Processing of Speech via Web Services. *Computer Speech and Language* 45: 326–47. [CrossRef]

Köpke, Barbara, and Monika S. Schmid. 2004. Language Attrition: The Next Phase. In *First Language Attrition: Interdisciplinary Perspectives on Methodological Issues*. Edited by Monika S. Schmid, Barbara Köpke and Merel Keijzerand Lina Weilemar. Amsterdam: John Benjamins, vol. 28, pp. 1–43.

Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. B. Christensen. 2017. LmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82: 1–26. [CrossRef]

Ladefoged, Peter. 2005. *Vowels and Consonants*. Malden: Blackwell Publishing.

Lado, Robert. 1957. *Linguistics across Cultures. Applied Linguistics for Teachers*. Ann Arbor: University of Michigan Press.

Lenneberg, Eric H. 1967. *Biological Foundations of Language*. New York: John Wiley and Sons.

Levy, Erika S., and Franzo F. Law. 2010. Production of French Vowels by American-English Learners of French: Language Experience, Consonantal Context, and the Perception-Production Relationship. *The Journal of the Acoustical Society of America* 128: 1290–1305. [CrossRef]

Lisker, Leigh, and Arthur S. Abramson. 1964. A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. *Word* 20: 384–422. [CrossRef]

Mack, Molly. 1990. Phonetic Transfer in a French-English Bilingual Child. In *Language Attitudes and Language Conflict. Spracheinstellungen Und Sprachkonflikte*. Edited by Peter H. Nelde. Plurilingua IX. Bonn: Dümmler, pp. 107–24.

MacKay, Ian R. A., James E. Flege, Thorsten Piske, and Carlo Schirru. 2001. Category Restructuring During Second-Language Speech Acquisition. *The Journal of the Acoustical Society of America* 110: 516–28. [CrossRef]

Major, Roy C. 1992. Losing English as a First Language. *The Modern Language Journal* 76: 190–208. [CrossRef]

Mayr, Robert, Sacha Price, and Ineke Mennen. 2012. First Language Attrition in the Speech of Dutch–English Bilinguals: The Case of Monozygotic Twin Sisters. *Bilingualism: Language and Cognition* 15: 687–700. [CrossRef]

Mayr, Robert, David Sánchez, and Ineke Mennen. 2020. Does Teaching Your Native Language Abroad Increase L1 Attrition of Speech? The Case of Spaniards in the United Kingdom. *Languages* 5: 41. [CrossRef]

Mennen, Ineke. 2004. Bi-Directional Interference in the Intonation of Dutch Speakers of Greek. *Journal of Phonetics* 32: 543–63. [CrossRef]

Mennen, Ineke, and Denise Chousi. 2018. Prosody in First-Generation Adult Immigrants and Second-Generation Heritage-Langauge Users: The Timing of Prenuclear Rising Accents. Paper presented at 9th International Conference on Speech Prosody, Poznań, Poland, June 13–16; pp. 828–32. [CrossRef]

Miller, Sean. 2012. Acting Tips from Thespian Arnold Schwarzenegger. Available online: https://www.backstage.com/magazine/article/acting-tips-thespian-arnold-schwarzenegger-51245/ (accessed on 3 April 2020).

Moosmüller, Sylvia. 2007. Vowels in Standard Austrian German. Habilitation thesis, University of Vienna, Vienna, Austria.

Moosmüller, Sylvia, Carolin Schmid, and Julia Brandstätter. 2015. Illustrations of the IPA: Standard Austrian German. *Journal of the International Phonetic Association* 45: 339–48. [CrossRef]

Moyer, Alene. 1999. Ultimate Attainment in L2 Phonology: The Critical Factors of Age, Motivation, and Instruction. *Studies in Second Language Acquisition* 21: 81–108. [CrossRef]

Muscle & Fitness. 2017. Arnold Schwarzenegger Opens up to Fans on Live Chat. Available online: https://www.muscleandfitness.com/athletes-celebrities/news/arnold-schwarzenegger-opens-fans-live-chat/ (accessed on 25 January 2021).

Naumburger Tageblatt. 2015. Arnold Schwarzenegger Im Interview: 'Beim Rechnen Denke Ich Auf Deutsch'. Available online: https://www.naumburger-tageblatt.de/nachrichten/panorama/arnold-schwarzenegger-im-interview--beim-rechnen-denke-ich-auf-deutsch--22639912-seite3 (accessed on 24 June 2020).

Odlin, Terence. 1989. *Language Transfer: Cross-Linguistic Influence in Language Learning*. Cambridge: Cambridge University Press.

Odlin, Terence. 2006. Cross-Linguistic Influence. In *the Handbook of Second Language Acquisition*. Edited by Catherine J. Doughty and Michael H. Long. Malden: Blackwell, pp. 436–86.

Oh, Grace E., Susan Guion-Anderson, Katsura Aoyama, James E. Flege, Reiko Akahane-Yamada, and Tsuneo Yamada. 2011. A One-Year Longitudinal Study of English and Japanese Vowel Production by Japanese Adults and Children in an English-Speaking Setting. *Journal of Phonetics* 39: 156–67. [CrossRef]

Outland Baker, Barbara. 2006. *Arnold and Me: In the Shadow of the Austrian Oak*. Bloomington: Author House.

Penfield, Wilder, and Lamar Roberts. 1959. *Speech and Brain Mechanisms*. Princeton: Princeton University Press.

Piske, Thorsten, James E. Flege, Ian R. A. MacKay, and Diane Meador. 2002. The Production of English Vowels by Fluent Early and Late Italian-English Bilinguals. *Phonetica* 59: 49–71. [CrossRef]

Pömp, Julian, and Christoph Draxler. 2017. OCTRA. A Configurable Browser-Based Editor for Orthographic Transcription. Paper presented at 13th Conference of Phonetics and Phonology in the German Language Area (P&P13), Berlin, Germany, September 28–29; pp. 145–48.

Preston, John. 2015. An Austrian Hick in London: Arnie's Early Years. Available online: https://www.telegraph.co.uk/film/terminator-genisys/arnold-schwarzenegger-early-life-bodybuilding/ (accessed on 20 May 2020).

Quora. 2017. Can Arnold Schwarzenegger Speak German? Available online: https://www.quora.com/Can-Arnold-Schwarzenegger-speak-German (accessed on 25 January 2021).

Quora. 2020. Can Arnold Schwarzenegger Speak Without an Accent? Available online: https://www.quora.com/Can-Arnold-Schwarzenegger-speak-without-an-accent (accessed on 25 January 2021).

R Core Team. 2020. *R: A Language and Environment for Statistical Computing. Computer Program*. Vienna: R Foundation for Statistical Computing, Available online: https://www.r-project.org/ (accessed on 20 September 2020).

Ramos, Francisco, and Stephen Krashen. 2013. Arnold's Advantages: How Governor Schwarzenegger Acquired English Through De Facto Bilingual Education. *International Multilingual Research Journal* 7: 220–29. [CrossRef]

Reddit. 2021. Isn't It Kinda Weird That Arnold Schwarzenegger Hasn't Lost His Accent? Available online: https://www.reddit.com/r/Showerthoughts/comments/3hb6bj/isnt_it_kinda_weird_that_arnold_schwarzenegger/ (accessed on 25 January 2021).

Reubold, Ulrich, and Jonathan Harrington. 2015. Disassociating the Effects of Age from Phonetic Change: A Longitudinal Study of Formant Frequencies. In *Language Development: The Lifespan Perspective*. Edited by Annette Gerstenberg and Anja Voeste. Amsterdam: Benjamins, pp. 9–37.

Reubold, Ulrich, and Jonathan Harrington. 2017. The Influence of Age on Estimating Sound Change Acoustically from Longitudinal Data. In *Panel Studies of Language Variation and Change*. Edited by S. E. Wagner and I. Buchstaller. London: Routledge, pp. 129–52.

Reubold, Ulrich, Jonathan Harrington, and Felicitas Kleber. 2010. Vocal Aging Effects on F0 and the First Formant: A Longitudinal Analysis in Adult Speakers. *Speech Communication* 52: 638–51. [CrossRef]

Roach, Peter. 2009. *English Phonetics and Phonology: A Practical Course*, 4th ed. Cambridge: Cambridge University Press.

Ryalls, Jack, Marni Simon, and Jerry Thomason. 2004. Voice Onset Time Production in Older Caucasian- and African-Americans. *Journal of Multilingual Communication Disorders* 2: 61–67. [CrossRef]

Saito, Kazuya, Yui Suzukida, and Hui Sun. 2019. Aptitude, Experience, and Second Language Pronunciation Proficiency Development in Classroom Settings: A Longitudinal Study. *Studies in Second Language Acquisition* 41: 201–25. [CrossRef]

Sancier, Michele L., and Carol A. Fowler. 1997. Gestural Drift in a Bilingual Speaker of Brazilian Portuguese and English. *Journal of Phonetics* 25: 421–36. [CrossRef]

Schiel, Florian. 1999. Automatic Phonetic Transcription of Non-Prompted Speech. Paper presented at XIVth International Congress of Phonetic Sciences, San Francisco, CA, USA, August 1–7; pp. 607–10. [CrossRef]

Schmid, Monika S., and Holger Hopp. 2014. Comparing Foreign Accent in L1 Attrition and L2 Acquisition: Range and Rater Effects. *Language Testing* 31: 367–88. [CrossRef]

Schmid, Monika S., and Barbara Köpke. 2017. The Relevance of First Language Attrition to Theories of Bilingual Development. *Linguistic Approaches to Bilingualism* 7: 637–67. [CrossRef]

Schoonmaker-Gates, Elena. 2015. On Voice-Onset Time as a Cue to Foreign Accent in Spanish: Native and Nonnative Perceptions. *Hispania* 98: 779–91. [CrossRef]

Schwarzenegger, Arnold. 2012. *A Total Recall. My Unbelievably True Life Story*. London and New York: Simon & Schuster.

Scovel, Thomas. 1969. Foreign Accents, Language Acquisition, and Cerebral Dominance. *Language Learning* 19: 245–53. [CrossRef]

Sharwood Smith, M., and E. Kellerman. 1986. *Crosslinguistic Influence in Second Language Acquisition*. New York: Pergamon Press.

Simon, Ellen. 2009. Acquiring a New Second Language Contrast: An Analysis of the English Laryngeal System of Natives Peakers of Dutch. *Second Language Research* 25: 377–408. [CrossRef]

Smith, Bruce L., Jan Wasowicz, and Judy Preston. 1987. Temporal Characteristics of the Speech of Normal Elderly Adults. *Journal of Speech and Hearing Research* 30: 522–29. [CrossRef]

Stoehr, Antje, Titia Benders, Janet G. van Hell, and Paul Fikkert. 2017. Second Language Attainment and First Language Attrition: The Case of VOT in Immersed Dutch-German Late Bilinguals. *Second Language Research* 33: 483–518. [CrossRef]

Sučková, Magda. 2020. First Language Attrition in Voice Onset Times in Anglophone Expatriates Residing in the Czech Republic. *Brno Studies in English* 46: 47–66. [CrossRef]

Thiery, Christopher. 1978. True Bilingualism and Second-Language Learning. In *Language Interpretation and Communication*. Edited by David Gerver and H. W. Sinaiko. New York: Plenum, pp. 145–53.

Thornburgh, Dianne F., and John H. Ryalls. 1998. Voice Onset Time in Spanish-English Bilinguals: Early Versus Late Learners of English. *Journal of Communication Disorders* 31: 215–29. [CrossRef]

Tobin, Stephen J., Hosung Nam, and Carol A. Fowler. 2017. Phonetic Drift in Spanish-English Bilinguals: Experiment and a Self-Organizing Model. *Journal of Phonetics* 65: 45–59. [CrossRef]

Ulbrich, Christiane, and Mikhail Ordin. 2014. Can L2-English Influence L1-German? The Case of Post-Vocalic/r/. *Journal of Phonetics* 45: 26–42. [CrossRef]

Verspoor, Marjolijn, Wander Lowie, and Marijn van Dijk. 2008. Variability in Second Language Development from a Dynamic Systems Perspective. *The Modern Language Journal* 92: 214–31. [CrossRef]

von Uslar, Moritz. 2012. 'Hallo, Herr Schwarzenegger?'. Available online: https://www.zeit.de/2012/42/Arnold-Schwarzenegger-Biografie (accessed on 11 September 2019).

Wan, Kristopher. 2017. Accents in Hollywood: Arnold Schwarzenegger. Available online: https://www.lucidaccent.com/post/2017/03/19/accents-in-hollywood-arnold-schwarzenegger (accessed on 5 December 2019).

Wiesinger, Peter. 2014. *Das Österreichische Deutsch in Gegenwart Und Geschichte*, 3rd ed. Wien and Berlin: LIT Verlag.

Yu, Vickie Y., Luc F. de Nil, and Elizabeth W. Pang. 2015. Effects of Age, Sex and Syllable Number on Voice Onset Time: Evidence from Children's Voiceless Aspirated Stops. *Language and Speech* 58: 152–67. [CrossRef] [PubMed]

Ziesel, Günther. 2018. Arnold Schwarzenegger Im Gespräch mit Günther Ziesel. *Video File*. Available online: https://www.youtube.com/watch?v=fW3YU3HFO5M (accessed on 21 May 2020).