LISA KORNDER

ARNOLD SCHWARZENEGGER NOW AND THEN

A longitudinal case study of Schwarzenegger's English and German pronunciation

Graz University
Library Publishing

UNI GRAZ

Lisa Kornder

Arnold Schwarzenegger Now and Then

GEWI AUSGEZEICHNET : ABSCHLUSSARBEITEN

Band 4

Herausgegeben von der Geisteswissenschaftlichen Fakultät der Universität Graz

LISA KORNDER

ARNOLD SCHWARZENEGGER NOW AND THEN

A longitudinal case study of Schwarzenegger's English and German pronunciation

Graz University Library Publishing

Gedruckt mit freundlicher Unterstützung durch:
Geisteswissenschaftliche Fakultät der Universität Graz

Zitiervorschlag:
Lisa Kornder, Arnold Schwarzenegger Now and Then. A Longitudinal case study of Schwarzenegger's English and German pronunciation. Graz 2022.

Graz University Library Publishing
Universitätsplatz 3a
8010 Graz
https://library-publishing.uni-graz.at

Grafische Grundkonzeption: Roman Klug, Presse und Kommunikation, Universität Graz

Coverbild: Photo by Pawel Czerwinski on unsplash.com
Satz: Lisa Kornder
Typografie: Source Serif Pro und Roboto

ISBN 978-3-903374-10-2
DOI https://doi.org/10.25364/978-3-903374-10-2

Table of Contents

1 General introduction and literature review 8
1.1. General introduction 8
1.2. Terminology 14
1.3. Early developments in SLA research 16
1.3.1 Contrastive analysis 16
1.3.2 Maturational constraints 20
1.3.3 Interlanguage and fossilization 22
1.3.4 Shifting perspectives 26
1.4. A current perspective in SLA and bilingualism research: First language attrition 28
1.4.1 Contextualizing first language attrition 28
1.4.2 First language attrition: Permanent loss or temporary change? 35
1.4.3 When attrition should (not) be called 'attrition': Some terminological considerations 37
1.5. Dynamic systems theory 41
1.5.1 Language development from a dynamic systems perspective 41
1.5.2 Language attrition and dynamic systems theory 43
1.6. Models of second language speech production and perception 44
1.6.1 The Speech Learning Model 45
1.6.2 The Perceptual Assimilation Model 48
1.7. Interim conclusion 49
1.8. Social and cultural dimensions of the phenomenon of foreign accent 51
2 Study I: Arnold Schwarzenegger's production of plosives in his L1 and L2 54
2.1. Introduction to Study I 54
2.2. Voice Onset Time 55
2.2.1 Voice onset time in English and Austrian German 59
2.2.2 L2 acquisition of VOT 63
2.2.3 L1 attrition of VOT 67

2.2.4 Other acoustic parameters of plosives 71
2.3. The study 73
2.3.1. Aims and objectives 73
2.3.2. Methodology 74
2.3.3 Benefits and drawbacks of working with broadcast recordings 86
2.3.4 Results 89
2.3.5 Discussion of Study I 98
3 Study II: Arnold Schwarzenegger's production of vowels in his L1 and L2 105
3.1. Introduction to Study II 105
3.1.1 American English vowels 105
3.1.2 Austrian German vowels 108
3.1.3 L2 acquisition of vowels 111
3.1.4 L1 attrition of vowels 116
3.2. The study 119
3.2.1 Aims and objectives 119
3.2.2 Materials and procedure 120
3.2.3 Results 124
3.2.4 Discussion of Study II 135
4 Study III: Perceived nativeness, intelligibility, and comprehensibility of L1 speech 142
4.1. Introduction to Study III 142
4.1.1 Perceived nativeness of L1 speech 143
4.1.2 Comprehensibility and intelligibility 147
4.1.3 Methodological aspects I: Assessing global accent 148
4.1.4 Methodological aspects II: Assessing intelligibility and comprehensibility 154
4.2. The study 158
4.2.1 Aims and objectives 158
4.2.2 Methodology 158
4.2.3 Results 167

4.2.4 Discussion of Study III 174
5 Limitations and overall discussion 179
5.1. Limitations 179
5.2. Overall discussion 181
5.3. Final remarks and outlook 186
6 Literature 188
7 Appendices 232
Appendix A 232
Appendix B 238
Appendix C 240
Appendix D 242
Appendix E 244
Appendix F 246
Appendix G 248
Appendix H 254
Appendix I 256
Appendix J 258
Appendix K 259

1 General introduction and literature review

1.1. General introduction

Arnold Alois Schwarzenegger (AS henceforth) was born into an Austrian family in 1947 in Thal, a small village near Graz (Austria). At the age of 21, AS left his home country and moved to the United States to advance his career as a professional bodybuilder. Within a short period of time, he achieved his goal and became one of *the* bodybuilding legends of the 1970s. Besides that, he is well-known for being an actor and a politician, as he was elected Governor of California in 2003 (Schwarzenegger 2012). Schwarzenegger sees himself as the embodiment of the American Dream (Naumburger Tageblatt 2015) and, in fact, he is arguably one of the most famous Austrian-American celebrities of the past decades.

Against the background of his language learning history, AS can be described as a late consecutive bilingual who acquired his second language (L2, American English) in early adulthood when his first language (L1, Austrian German) had already been fully developed. Besides learning the L2 through everyday communication with native speakers, AS attended English as a Second Language classes at a community college in California shortly after he had moved to the United States (Schwarzenegger 2012). Despite the fact that he had already been learning English as a foreign language in school in Austria, he struggled with becoming fluent in the L2, and admitted that acquiring English pronunciation was the most challenging task of his language learning endeavors (ibid.). Even after more than 50 years of living in an English-speaking country, AS still maintains a detectable Austrian German accent when speaking English, something which receives much attention from the public (e.g., Daily Mail UK 2015; Gersemann 2009; Wan 2017). His pronunciation does, however, not only receive attention because of his L1-accented English, but it is also frequently discussed whether he has forgotten or even 'unlearned' his native language (e.g., Guten Morgen Österreich 2012a, 2012b; Quora 2017; Stummer 2011), which he is said to speak with an English accent.

Schwarzenegger's apparent difficulty to acquire English pronunciation is by no means exceptional. In fact, the ability to attain an authentic, native-like accent is often considered the most difficult aspect of L2 acquisition (e.g., Bongaerts et al. 1997; Hyltenstam and Abrahamsson 2006; Piske, MacKay, and Flege 2001). Even when L2 learners have acquired an extended vocabulary and excellent command of L2 grammar, their speech is often characterized by a detectable foreign accent, especially when the L2 has been acquired late in life (e.g., Flege 2003; Piske and MacKay 1999; Piske, MacKay, and Flege 2001; Scovel 1969, 1988). Scovel (1988, 247) described this apparently special status of L2 pronunciation acquisition as the 'Joseph Conrad phenomenon': Joseph Conrad, an early 20th-century novelist, moved

from his home country Poland to the UK at the age of 18. He acquired native-like proficiency in L2 writing after a short period of time, yet he was not able to get rid of his strong Polish accent (see also Jean-Aubry 1957). Over the past decades, second language acquisition (SLA) research has investigated the causes of this phenomenon and has provided a number of possible explanations. Scovel (1988), among others, argued that biological mechanisms are ultimately responsible for foreign accented speech (see also Lenneberg 1967; Penfield and Roberts 1959). With regard to this, it was assumed that attaining native-like proficiency in a second language is only possible within a limited time span, which usually ends at around puberty; after that, acquiring an L2 in a native-like fashion is no longer possible. The influence of biological constraints on second language acquisition in general and on L2 acquisition of pronunciation in particular continues to be a controversial topic among SLA researchers (see Hyltenstam and Abrahamsson 2000, for a discussion), and many scholars propose that the existence of foreign accents cannot be exclusively explained against the background of maturational constraints (e.g., Bialystok, Hakuta, and Wiley 2003; Hyltenstam and Abrahamsson 2000; Moyer 1999, 2013; Singleton 2005, 2007; Snow and Hoefnagel-Höhle 1978). Other factors which have been shown to impact L2 speech acquisition to some extent are amount and quality of L1/L2 input and use (e.g., Flege, Frieda, and Nozawa 1997; Flege and Liu 2001; Moyer 2009), formal instruction (e.g., Saito 2012; Saito and Lyster 2012; Thomson and Derwing 2014) and phonetic training (e.g., Aliaga-García and Mora 2009; Thomson 2011), motivation (e.g., Moyer 1999; Nagle 2018), and (phonetic) aptitude (e.g., Baker Smemoe and Haslam 2013; Jilka et al. 2009; Jilka 2009).[1]

Alongside the factors previously mentioned, the role of the first language and its influence on L2 speech acquisition has been extensively discussed from the onset of SLA research in the second half of the 20th century (e.g., Lado 1957; Moulton 1962a, 1962b; Odlin 1989, 2006; Weinreich [1953] 1974). Traditionally, it was assumed that the L1 provides the perceptual framework for second language learning and thus constrains and impedes an adult L2 learner's ability to acquire nonnative sound categories (Moyer 2013). More current investigations into cross-linguistic influences on speech production, however, found that not only the first language impacts L2 speech learning but that also L2 learning experience can influence L1 pronunciation features (de Leeuw, Mennen, and Scobbie 2012a, 2012b; Dmitrieva, Jongman, and Sereno 2010; Flege 1987b; Mayr, Price, and Mennen 2012; Mennen 2004), which might result in modifications of L1 phonetic categories, even to the extent that speakers are perceived to have a nonnative accent in their L1 (Bergmann et al. 2016; de Leeuw, Schmid, and Mennen 2010; Hopp and Schmid 2013; Mayr,

1 For comprehensive overviews of factors influencing L2 pronunciation abilities see Derwing and Munro (2015), Piske, MacKay, and Flege (2001), and Trofimovich, Kennedy, and Foote (2015).

Sánchez, and Mennen 2020; Schmid and Hopp 2014). As mentioned above, AS's L1 pronunciation receives much public attention as it is often claimed that he has – to some extent – unlearned his first language. This phenomenon is referred to as first language loss or first language attrition, which can be defined as a temporal or permanent decline of first language linguistic abilities in a late bilingual who is being immersed in an L2-speaking or L2-dominant environment and thus has increased contact with the L2 (Köpke 2004; Köpke and Schmid 2004; Seliger and Vago 1991).[2] The findings of studies exploring attrition effects indicate that even a mature and fully developed L1 system is sensitive to change due to L2 acquisition experience and changes in the linguistic environment (see e.g., de Leeuw, Opitz, and Lubinska 2013). Being immersed in an L2-speaking environment might be beneficial when it comes to attaining native-like proficiency in L2 pronunciation, but simultaneously "the associated reduced L1 use may cause linguistic abilities in the L1 to deteriorate" (Stoehr et al. 2017, 484). Despite the fact that there is a growing body of research examining aspects of L1 phonetic/phonological attrition in speakers with different language backgrounds (e.g., Bergmann et al. 2016; de Leeuw 2009, 2019; de Leeuw, Mennen, and Scobbie 2012a, 2012b; Mayr, Price, and Mennen 2012; Opitz 2011), there is still a lack of longitudinal investigations into L1 attrition phenomena. However, if language acquisition, including language maintenance and language loss, is considered a dynamic and ever-changing process, long-term investigations into speech and language development are indispensable. This view is in line with a Dynamic Systems Theory (DST) approach to second language learning which considers language acquisition as a constantly changing process which is influenced by a complex interplay between internal and external variables (de Bot, Lowie, and Verspoor 2007; Larsen-Freeman 1997, 2000). The relevance of DST in the context of L2 acquisition and L1 attrition will be discussed in Section 1.5.

Furthermore, research into the *relationship* between L2 speech acquisition and L1 phonetic attrition has not provided conclusive findings yet. Some argue that an advanced L2 proficiency may cause L1 pronunciation abilities to decline (Flege 1987b), while others provide evidence for changes in L1 pronunciation experienced by learners at the onset of L2 learning (Chang 2012, 2013; Dmitrieva, Jongman, and Sereno 2020; Kartushina, Frauenfelder, and Golestani 2016). Consequently, one of the objectives of this study is to find out whether there is a direct relationship between an increase in L2 pronunciation proficiency on the one hand, and a potential decline of L1 pronunciation abilities, on the other hand.

2 Note that some scholars suggest a broader definition of first language attrition (see e.g., Schmid and Köpke 2017a, 2017b). Aspects related to the use and definition of the term 'attrition' will be outlined in more detail in Section 1.4.3.

The present longitudinal case study aimed to examine AS's L1 and L2 speech production over a period of approximately 40 years, i.e., from the 1970s up to the 2010s. In order to arrive at a thorough characterization of AS's first and second language speech development, and to gain a profound understanding of potential interactions between processes of acquisition and attrition in a late consecutive bilingual, the following research questions were addressed:

Study I: AS's productions of L1 and L2 plosives

RQ 1.1: To what extent has AS's L2 realization of voice onset time (VOT) contrast changed when comparing three stages in time (early, mid, and late)? Does he produce a significant difference between phonologically voiced and voiceless plosives in terms of VOT for each of the stages, i.e., has he gradually become more native-like since he moved to the United States?

RQ 1.2: To what extent has AS's L1 realization of VOT contrast changed when comparing two stages in time (early and late)? Does he produce a significant difference between phonologically voiced and voiceless plosives in terms of VOT for each of the stages, i.e., have his L1 VOT categories moved away from native production norms and at the same time moved closer to L2 production norms since he moved to the United States?

Study II: AS's productions of L1 and L2 vowels

RQ 2.1: To what extent has AS's L2 production of monophthongs changed when comparing three stages in time (early, mid, and late)? Does he produce L2 vowels with formant frequencies (F1 and F2) which resemble those of monolingual American English speakers, i.e., has he gradually become more native-like since he moved to the United States?

RQ 2.2: To what extent has AS's L1 production of monophthongs changed when comparing two stages in time (early and late)? Does he produce L1 vowels with formant frequencies (F1 and F2) which more closely resemble those of monolingual American English speakers, i.e., have his L1 vowel categories moved away from native production norms and at the same time moved closer to L2 production norms since he moved to the United States?

Study III: Perceived nativeness, comprehensibility and intelligibility of AS's L1 speech

RQ 3.1: Is AS perceived to sound less native in his late L1 German pronunciation than in his early L1 pronunciation compared to monolingual Austrian German speakers?

RQ 3.2: To what extent does the listeners' linguistic background (monolingual vs. bilingual) affect their perception of L1 nativeness in German-English bilingual and German monolingual speakers?

RQ 3.3: To what extent is the perceived nativeness of L1 German speech related to the intelligibility and comprehensibility of L1 speech, i.e., does a lower degree of perceived nativeness impede the comprehensibility and intelligibility of L1 pronunciation, and vice versa?

In order to gain insight into the extent to which AS's L1 and L2 pronunciation has changed since he moved to an English-speaking environment and to arrive at an understanding of the relationship between L2 acquisition and L1 attrition, specific features of AS's segmental speech production, namely plosives and vowels, were investigated in two acoustic-phonetic studies (Study I and Study II). For this purpose, speech data were collected from various interviews with AS which were conducted in German or English and were broadcast on different U.S. and Austrian/German TV and radio channels. The German and English speech corpora were divided into *early* (1977–1989) and *late* (2010–2017/18) speech samples. An additional set of samples was included to represent AS's English pronunciation in the 1990s and early 2000s (*mid* samples).

With regard to AS's plosive production, the focus was on the realization of voice onset time (VOT) contrast in word-initial prevocalic plosives (Study I, Chapter 2). So far, VOT has been most frequently studied in English, which distinguishes between long-lag and short-lag VOT: Voiceless plosives are produced with long-lag VOT in the range of 30 to 100ms, i.e., they are aspirated. English phonologically voiced plosives, on the other hand, are realized with short-lag positive VOT ranging from 0 to 25ms, or with voicing lead (Docherty 1992; Lisker and Abramson 1964). The same VOT-distinction applies to Standard German German, which exhibits a relatively consistent VOT contrast between short-lag versus long-lag plosives (Braunschweiler 1997; Jessen 1998; Taylor 1975). Speakers of Austrian German varieties, however, have been observed to neutralize VOT contrast in word-initial bilabial and alveolar plosives in conversational speech by producing both phonologically voiced and voiceless targets with short-lag VOT (Hödl 2017, 2019; Moosmüller, Schmid, and Brandstätter 2015; Wiesinger 1996, 2014). Based on these cross-linguistic differences between Austrian German and English, the present study aims to determine if and to what extent AS's production of VOT in word-initial L1 and L2 plosives has changed since he moved to an English-speaking environment, i.e., if his L2 VOT productions have moved closer towards native English norms in the course of gaining more L2-learning experience, and if his L1 VOT productions have moved further away from L1 Austrian German norms as a result of increased L2 exposure over

time. The observation that AS's L1 VOT values have drifted towards the norms of L2 VOT values would suggest that his L1 plosives are affected by phonetic attrition. In order to find out to what extent his L1 and L2 pronunciation has shifted closer to or deviated from native speaker norms, AS's early and late segmental speech productions were compared to monolingual productions in both languages.

Study II (Chapter 3) focuses on examining AS's L1 and L2 monophthongal vowel space by comparing his early and late productions. The aim of this comparison is to identify potential changes and modifications in his vowel space and to address the question of whether his productions of L2 vowels have moved towards native English production norms, and whether his L1 vowel productions can be observed to have moved away from native Austrian German norms. As there are no acoustic studies available so far which investigate acoustic characteristics of vowel production in the Austrian German regional variety spoken in Thal, AS's place of birth, an age-matched monolingual control speaker from Thal was recorded for comparison purposes (see Section 2.3). Since AS moved to California at the age of 21, it can be assumed that his English is influenced by the English variety spoken in California, which primarily differs from General American and other American English varieties in terms of the quality and duration of some vowels (Ladefoged 2005). Therefore, the Californian English vowel inventory (see e.g., Hagiwara 1997) provides the basis for the investigation of changes in AS's L2 vowel productions. Modifications of his L1 and L2 vowel systems were described and compared by conducting vowel formant frequency measurements of the first and second formant (F1 and F2).

Alongside conducting acoustic-phonetic analyses of VOT and formant frequencies, the present investigation assessed perceived nativeness, intelligibility and comprehensibility of AS's L1 pronunciation (Study III, Chapter 4). For this purpose, two listener groups from Graz ($n = 20$ each), who differed in terms of their linguistic background (monolingual vs. bilingual), were invited to rate a set of 29 speech samples according to perceived nativeness, intelligibility and comprehensibility. The speech corpus included short stretches of speech produced by two Austrian German-English bilinguals, including AS, and seven monolingual Austrian German speakers from Thal, who functioned as control speakers. To represent AS's pronunciation, speech samples from two centuries were included, namely from the 1970s (early samples) and from the 2010s (late samples). The main objective of comparing early and late speech samples was to examine if AS's overall L1 accent has changed since he moved to an L2-speaking country, i.e., whether his late L1 pronunciation is perceived to sound less native compared to his early L1 pronunciation. The nativeness assessment in the present investigation is based on the rating instrument used by Moyer (1999), who studied perceived accentedness in L2 German speech. The dimension of intelligibility, i.e., a listener's understanding of an utterance (Munro and Derwing 1995a), was examined by means of listener transcripts.

Perceived comprehensibility of AS's L1 speech, i.e., how easy or difficult it is for a listener to understand an utterance (ibid.), was also tested by means of scalar ratings (see Section 4.2).

Based on the findings of the acoustical analyses and the listener ratings outlined above, the primary goal of the present investigation is to broaden the scope of existing research on L2 attainment and L1 attrition. The findings will provide insights into bidirectional L1–L2 influences operating in a late consecutive bilingual's language systems over an extended period of time. In addition, the study focuses on describing and gaining a better understanding of the dynamic processes involved in L2 speech acquisition and L1 maintenance/decline. Finally, the outcome of the study is not only expected to contribute to existing research on L2 acquisition and L1 attrition, but it will also bridge a gap between the scientific research community and the public sphere: As AS's L1 and L2 accent is of interest to a broader public audience, findings which contribute to explaining this phenomenon from a scientific perspective will also offer a valuable source for the public, including radio, television, and other media.

1.2. Terminology

First language (L1) and second language (L2)

Second language acquisition refers to the acquisition of any other language acquired subsequent to the first language (Ellis 1997; Foster-Cohen 2001; Ortega 2013), and often also includes heritage language learning (see Gass and Selinker 2008). Some scholars argue, however, that using 'L2' as a cover term for all languages a speaker acquires subsequent to the L1 is misleading as such terminology is assumed to neglect the chronology of acquisition (see Wunder 2011). Hence, the more neutral term 'nonnative language' is often used to refer to a second, third, or fourth language acquired in addition to the first (native) language, which is especially relevant in the context of multilingualism. As the present thesis does not focus on speakers with more than two languages, L1 and native language, and L2 and nonnative language are used interchangeably.

Second language acquisition (SLA) vs. L2 acquisition

A further distinction needs to be made between the terms 'SLA' and 'L2 acquisition'. While L2 acquisition primarily describes the process of learning a language subsequent to the first language, the acronym SLA is often used to refer to the scientific field of research which examines the processes involved in L2 acquisition (Gass and Selinker 2008). Despite the fact that this terminological distinction is not always made in SLA literature, it will be maintained in the present thesis for the sake of clarity.

(Foreign) language learning vs. (second) language acquisition
Particularly in the 1980s, a distinction was frequently made between language *learning* and language *acquisition*, depending on the context in which a second/foreign language is learned or acquired. With regard to this, Krashen (1981) stated that language acquisition is very similar to the process of native language acquisition in children, i.e., it describes an unconscious process which is primarily determined by interactions with native speakers and takes place in a naturalistic environment. Language learning, by contrast, was described as a conscious process which is characterized by error correction and explicit teaching of language rules in a formal instructional context (Krashen 1981). In more current SLA literature, the learning-acquisition-distinction is not always maintained due to the observation that often language acquisition/learning takes place through both naturalistic acquisition and instructed learning (see Foster-Cohen 2001). In the present thesis, second language acquisition and second language learning will be used interchangeably, both referring to the natural and/or instructed acquisition of an L2.

Austrian German
As stated by Hobel and Vollmann (2015, 5), "[t]he linguistic varieties in Austria range along a continuum between dialect and standard". Hence, Austrian German does not represent a uniform variety of German, but is used as an umbrella term for different varieties and dialects which are spoken in Austria, including both Standard Austrian German (SAG)[3] and non-standard regional and social varieties (see Wiesinger 1990, 2014). SAG, as described by Moosmüller, Schmid, and Brandstätter (2015), is oriented towards Standard German German norms and belongs to the (Middle/South-Middle) Bavarian dialect region.[4] SAG is spoken in large urban centers, such as Salzburg and Vienna, and is usually associated with speakers belonging to an educated and affluent social class (Moosmüller 2011). In the present thesis, SAG will be used in the sense of a normative variety of Austrian German, which is understood and accepted by a large number of people within the language community (see Auer 1995; Laver 1994; Moosmüller 1995).

As mentioned in the previous section, the subject (AS) examined in the present thesis was born and raised in Thal, a small municipality near Graz-City, the capital of the federal state of Styria. While Moosmüller (2011) and Moosmüller, Schmid,

3 Although the term 'standard language' is frequently used in the literature, especially from a (socio-)linguistic point of view, it should be noted that the notion if a standard language is challenged by some scholars. For an in-depth discussion of ideological, empirical and theoretical aspects which need to be taken into consideration when defining what a 'standard language' is, see Deppermann, Kleiner, and Knöbl (2011).

4 Middle and Central Bavarian dialects are spoken in the south of Germany and in Austria with the exception of the Austrian federal state of Vorarlberg, which belongs to the Alemannic dialect region (see e.g., Kleber 2018; Wiesinger 1996, 2014).

and Brandstätter (2015) describe SAG being predominantly spoken in large urban centers, Schuller et al. (2019, 2379) argue that "[u]rban centers such as Graz differ from rural regions [in that] neither 'dialect' nor 'standard' is spoken". Instead, speakers living in urban regions use a mixture of both dialect and more standard-oriented features of Austrian German. With regard to this, Schuller et al. (2019) characterize the Austrian German variety spoken in Graz as 'Urban Styrian' in contrast to more rural Styrian dialects (see also Wiesinger 1967). For the sake of convenience, AS's L1 variety will be referred to as 'Austrian German' in the context of the present thesis.

1.3. Early developments in SLA research

SLA is a relatively young field of research which emerged in the mid-20th century.[5] This emergence, according to Ellis (1997), can be explained by globalization processes which led to the need to communicate across national borders. In order to be part of and participate in this global communicative network, people needed to learn other languages in addition to their first language. As a result, an interest emerged to examine, describe and explain language acquisition processes as well as to identify factors which might be involved in these processes. In 1982, Rutherford summarized the core interests of SLA by stating that "[w]e wish to know what is acquired, how it is acquired and when it is acquired. But were we to have these answers even to these questions, we would still want to know why" (1982, 85). As this statement suggests, SLA is a complex field of research which adopts a multidisciplinary perspective including different scientific disciplines, such as first language acquisition, linguistics, psychology, sociolinguistics, and education (see Larsen-Freeman 2000; Ortega 2013. Early SLA research primarily focused on determining and describing potential areas of difficulty for L2 learners which were assumed to result from differences between the native and the target language. The systematic description of these differences aimed at facilitating foreign language learning and teaching, as specified in the Contrastive Analysis approach to second language acquisition, which will be further outlined below.

1.3.1 Contrastive analysis

As mentioned in the introduction, the traditional understanding of language transfer[6] held that the L1 was not prone to being affected by interference processes

5 For a detailed overview of the (historical) development of Applied Linguistics and SLA, see de Bot (2015) and McCarthy (2001).

6 In the literature, the notion of 'language transfer' might also be referred to by using terms such as 'language interference' (e.g., Weinreich [1953] 1974) or 'cross-linguistic influence' (e.g., Sharwood Smith and Kellerman 1986). The terminology used to describe the phenomenon of language transfer

while a learner's L2 system was likely to be influenced by the L1 system. This view developed in the 1950s and 1960s against the background of the Contrastive Analysis (CA) approach which has its roots in behaviorist learning theory (Bloomfield 1933; Skinner 1957, 1974) and which served as a basis for predicting transfer processes. The basic assumption of behaviorist learning theory was that learning takes place through imitation (i.e., copying a particular stimulus) and reinforcement (i.e., appropriate behavior is rewarded, inappropriate behavior is punished), which were assumed to trigger the formation of habits (Skinner 1974). Habits which had been established for the first language were claimed to interfere with the second language system and hence impeded the successful learning of a new language (see Mitchell, Myles, and Marsden 2019, for an overview). The origins of CA can be traced back to Fries (1945) and Lado (1957), who described language transfer as being unidirectional in nature, i.e., taking place from the L1 to the L2, but not vice versa. In addition, L1 transfer processes were considered to be highly similar in all L2 learners, regardless of the languages involved (Weinreich [1953] 1974). Based on these assumptions, systematic comparisons of the L1 and L2 were made in order to reveal areas of potential difficulty learners might face when acquiring a second language. Moulton (1962b), for instance, compared the phonetic and phonological systems of English and German and provided a classification of pronunciation errors, which should serve as a basis for teaching pronunciation in classroom settings. As pointed out by Lado (1957), such systematic L1–L2 comparisons were essential for language teaching and the design of teaching materials. Also Fries (1945) argued that "[t]he most effective materials are those that are based upon a scientific description of the language to be learned, carefully compared with a parallel description of the native language of the learner" (quoted in Dulay, Burt, and Krashen 1982, 98). Consequently, CA developed out of the pedagogical demand for effective foreign language teaching.

In terms of L2 speech production, proponents of CA argued that the more different L1 and L2 speech sounds are, the more difficult the authentic production of the L2 sounds (Lado 1957; Weinreich [1953] 1974). Furthermore, it was assumed that foreign language phonemes which are non-existent in the first language will be replaced with an L1 phoneme that is closest to the L2 phoneme, if the L2 learner fails to identify the L2 phoneme correctly (ibid.). Based on this, the nature of transfer was described being either positive or negative: Positive transfer occurs when an L1 structure has a corresponding structure in the L2 system, which results in authentic L2 productions and thus facilitates L2 acquisition. Negative transfer (i.e.,

or language interaction depends on "how researchers perceive its scope and its positive or negative effects" (Ecke 2004, 336). In the present work, the terms 'language transfer' and 'cross-linguistic influence' will be used to refer to interactions between the L1 and L2 linguistic systems.

interference), on the other hand, inhibits successful L2 acquisition because a specific L1 structure does not have a corresponding structure in the L2. Consequently, the L2 target structure is replaced with an L1 structure, resulting in incorrect L2 productions (Carroll 1968; Odlin 2006). Hence, interference from the first language was considered the primary obstacle to successful L2 (speech) acquisition (see Flege 1995, 235).

The investigation of transfer processes, however, began well before the emergence of CA. As early as 1939, Trubetzkoy argued that the inability to produce L2 sounds correctly was the result of first language interference. He described the L1 phonological system as a phonological 'sieve' which acts as a filter for L2 speech sounds and, according to this theory, incorrect speech productions result from dissimilarities between L1 and L2 sounds (Trubetzkoy [1939] 1977). A similar line of argumentation was advanced by Wode (1980), who suggested that an L2 phonological system is acquired through the 'grid' of the learner's first language. Thus, an L2 element which has a counterpart in the L1 system will be replaced by the most similar element; dissimilar sounds, by contrast, "are not substituted by the learner's L1 elements but they undergo autonomous developments similar to [...] those which the respective elements undergo in L1 acquisition" (Wode 1980, 129).

Despite its popularity in the 1960s, the CA approach to second language acquisition became a much-debated issue in the early 1970s, which led to a gradually changing view on language transfer processes. In its original formulation, the strong version of CA suggested that *all* learner errors could be predicted on the basis of L1–L2 comparisons (Lado 1957). According to Wardhaugh (1974), one of the critics of CA, this strong interpretation of CA was not tenable in practice because it failed to provide reliable and testable predictions. A number of studies supported Wardhaugh's argument and presented results contradicting the assumption that all learner errors could be predicted systematically. Richards (1971), for instance, reviewed several studies examining L2 learners of English with different L1 backgrounds. He explored the nature of learner errors observed in the studies and found that the majority of these errors could not be classified as interlingual errors, i.e., errors resulting from first language interference. Instead, certain errors were found to occur regardless of the learners' first language and were thus assumed to "reflect the learner's competence at a particular stage, and illustrate some of the general characteristics of language acquisition" (Richards 1971, 3). These errors, which Richards classified as *intralingual* and *developmental* errors, include, for example, overgeneralizations or false/incomplete applications of L2 rules. Similar findings were provided by Brooks (1960) and Dulay and Burt (1973, 1974a), who identified several types of errors which did not result from L1-interferences; in fact, the minority of errors could actually be traced back to transfer from the first language system (Dulay and Burt 1973).

In addition to different types of learner errors which were not taken into account by CA, Eckman (1977, 1996, 2003, 2008) stated that CA failed to allow for predictions to be made concerning the directionality of difficulty. He proposed that not all differences between a first and a second language will automatically entail production difficulties and, consequently, not all L2 elements will be equally difficult to acquire for learners (see Eckman 2008). Based on this, he called for an extension of CA by taking the directionality of difficulty into consideration (see Eckman 1977), which rests on the assumption that when two speakers with different L1 backgrounds learn the language of the other speaker, the speaker who has to learn a more marked L2 structure will face greater difficulty than the speaker who acquires an unmarked structure.[7]

Alongside the observation that language transfer processes might be affected by markedness relationships, it was generally questioned whether learning difficulty on the one hand, and the relative difference between two language systems on the other hand, can be equated. Odlin (1989), for example, points out that differences between an L1 and L2 do not inevitably impede successful L2 productions, as originally stated by CA. In fact, studies conducted within the framework of Flege's Speech Learning Model (Flege 1987a, 1987b, 1995, 2003, 2007) did not provide evidence for a strong interpretation of CA as they showed that differences between the L1 and L2 at a phonological level might facilitate L2 learning. These observations laid the basis for current views on cross-linguistic transfer, which will be discussed in Sections 1.4 and 1.6.

The fact that the validity of behaviorist learning theories was questioned added further to the recurrent criticism on CA (see Fisiak 1981; Major 2008). As a consequence, an alternative approach to identifying potential difficulties L2 learners might be confronted with developed in the late 1970s, known as Error Analysis (EA). EA was concerned with the systematic description of learner errors without exclusively focusing on errors resulting from L1 transfer. Corder (1967, 168), one of the pioneers of EA, disagreed with behaviorist learning theories by stating that not all learner errors[8] can be explained by "interference from the habits of the mother-tongue". Rather, errors observed in L2 learners' productions were interpreted to reflect innate strategies applied during the acquisition process. According to

7 Based on the notion of marked vs. unmarked linguistic features, Eckman (2003, 2008) formulated the Markedness Differential Hypothesis (MDH), according to which marked linguistic features are more difficult to acquire than unmarked features. Whether a linguistic structure is classified as marked or unmarked depends on its distribution within and across languages, i.e., unmarked structures occur more frequently than marked ones. In the context of L2 acquisition, the MDH holds that unmarked L1 structures are more likely to be transferred to the L2 compared to marked L1 features.

8 Corder (1967, 167) distinguishes mistakes from errors in that mistakes are "errors of performance" while errors are the result of systematic divergences from the target language, i.e., "errors of competence".

Corder, studying and analyzing these errors would provide language teachers with a sound basis for making language teaching more effective.

While CA failed to provide testable predictions, research conducted within the framework of EA seemed to offer actual empirical data (e.g., Ghadessy 1980; Richards 1971). However, also EA soon found itself under critical investigation. Schachter and Celce-Murcia (1977), for instance, discuss several insufficiencies of EA, such as the fallacy of analyzing learner errors in isolation by extracting them from the context of speaking, or the difficulty of providing a distinct classification system for errors (see also Schachter 1974; Strevens 1969). Similarly, Hammarberg (1979) voiced reservations against EA and argued that it failed to take non-errors into account which are as essential as actual errors when it comes to the realization of pedagogical objectives, such as language teaching and syllabus design.

In their concluding remark about the drawbacks of early approaches to explaining second language acquisition processes, Schachter and Celce-Murcia (1977, 449f.) state that "if anything was learned from the so-called rise and fall of CA, it should be that one single view of the language learning process [...] will not account for the diverse phenomena that exist". However, despite the fact that there has been an ongoing debate about the validity of CA and EA, some SLA researchers argue that CA still bears some plausibility from current research perspectives in Applied Linguistics. König and Gast (2007, 3), for example, point out that for translation purposes as well as in the fields of bilingualism and foreign language teaching, comparing languages and running contrastive analyses might indeed provide useful first insights into differences and commonalities between the languages under investigation. In fact, their systematic analysis and comparison of German and English phonology, morphology, and syntax can serve as a useful starting point for conducting in-depth investigations of L2 acquisition processes and thus broaden the perspective on the complex phenomenon of second language acquisition.

1.3.2 Maturational constraints

Alongside examining language transfer and interference processes, early SLA research concentrated on exploring the relationship between L2 acquisition and age-related factors, which has been a major topic of debate in SLA ever since (e.g., Long 1990; Singleton 2007; Singleton and Lengyel 1995; Singleton and Leśniewska 2021; Vanhove 2013). One of the earliest explanations of age-related influences on the acquisition of a second language were formulated in the critical period hypothesis (CPH), which held that language acquisition is predominantly – if not exclusively – controlled by biological mechanisms (Lenneberg 1967; Penfield and Roberts 1959).

Penfield and Roberts (1959) originally proposed the existence of a CPH in the context of first language acquisition. They stated that first language acquisition

processes are essentially limited due to a loss of brain plasticity which was assumed to commence at the age of around nine. After this age, the successful (native-like) acquisition of a language was considered to be no longer possible. This idea was complemented by Lenneberg (1967) who argued that biological constraints also ultimately inhibit the acquisition of a second language if it is acquired subsequent to the supposed end of a critical period, which he set at around puberty. Lenneberg's definition of a CPH is based on specific neurological processes associated with the lateralization of language functions in the brain. Once these lateralization processes are completed, the capacity to acquire a second language with native proficiency is no longer possible.

Early formulations of the CPH soon became a target of criticism, which led to modifications of its initial claims. Krashen (1973), for instance, provided evidence for a much earlier completion of lateralization, i.e., well before puberty, and thus challenged Lenneberg's arguments concerning the ultimately restrictive nature of lateralization. Scovel (1988) argued that Penfield and Roberts (1959) had failed to provide a comprehensive explanation of what they precisely meant by 'brain plasticity' and which specific neurological processes were involved in language acquisition. In addition, Scovel (1988) pointed out that not all areas of language learning are equally affected by a critical period as originally argued by Penfield and Roberts; instead, he proposed that "a critical period is defensible only for phonological learning and cannot be expanded to include the learning of other linguistic skills, such as the acquisition of new words or new grammatical patterns" (Scovel 1988, 59). He derived this assumption from the observation that phonological learning is the only area of language acquisition in which physical learning is directly involved, that is, the positioning and coordination of articulators required for speech sound production. Since learning these motoric abilities was assumed to be no longer possible once lateralization is completed, speakers who acquired an L2 late in life were thought to maintain a detectable foreign accent in their second language.

Early interpretations of the CPH outlined above have significantly influenced subsequent discussions on the role of age in L2 (speech) learning. Among many others, Bongaerts (1999; see also Bongaerts, Mennen, and van der Slik 2000) tested the CPH's prediction that native-like acquisition of second language speech is generally not possible if it is acquired after puberty. His study included highly successful native Dutch learners of English, who had acquired the L2 in adulthood. Foreign accent ratings revealed that some speakers were indeed perceived as native speakers of English, which was interpreted as evidence against any strong interpretation of the CPH and against the existence of an "absolute biological barrier" (Bongaerts 1999, 154) when it comes to the acquisition of native L2 pronunciation. Other factors besides age need to be taken into consideration in order to account for the existence of exceptionally successful L2 learners, such as a high level of motivation, access to

a sufficient amount of L2 input, and intensive L2 perception and production training (see Bongaerts 1999). Also Moyer (1999) stresses the importance of additional factors which are likely to interact with the age factor. She examined the impact of motivation, instruction, phonological training, and self-perception of productive accuracy in native English learners of German who acquired the L2 after puberty. Her conclusions suggest that despite the fact that these factors were not shown to override the age factor, they still significantly contributed to the level of L2 pronunciation proficiency. However, although numerous studies have investigated the relative influence of predictor variables on L2 pronunciation abilities and their interaction with age-related constraints (e.g., Guion 2003; Moyer 2009; Thomson 2011; Thomson and Derwing 2014), it remains unclear to date to what extent extra-linguistic variables affect successful L2 pronunciation and whether at least some of these factors play an even more essential role than age-related constraints.

Despite its popularity in the early 1950s, the validity of the CPH has been questioned by many SLA researchers, which resulted in several re-formulations and adaptions of its initial claims (see Birdsong 2014; Singleton 2007, for discussions). While some researchers still suggest that the existence of a critical period – despite several inconsistencies with regard to its end and beginning (e.g., Vanhove 2013) – is ultimately responsible for foreign accented speech in late L2 learners, others stress the importance of additional variables which interact with the age factor, as outlined above.

1.3.3 Interlanguage and fossilization

As can be concluded from the previous discussion, early developments in SLA suggest that explaining language acquisition processes solely on the basis of language transfer and maturational constraints does not satisfactorily grasp the complexity of these processes. One of the drawbacks of behaviorist learning theories and CA was that they only accounted for what could be directly observed on the 'surface' of L2 production, that is, the language input a learner receives and their output in response to the input (Ellis 1997, 32). *Internal* strategies and mechanisms applied by language learners in their attempts to acquire a second language were, however, not addressed. These insufficiencies of behaviorist learning theories led to the development of a new theoretical framework which aimed to explain language acquisition processes from a mentalist theory perspective.

The basis of mentalist approaches to SLA is the proposed existence of a Language Acquisition Device (LAD), described as an innate predisposition unique to humans which enables them to acquire language (Chomsky 1965, 1967). The concept of the LAD was first introduced by Chomsky in the 1960s and later served as a

basis for Selinker's (1979) formulation of what he termed 'interlanguage'.[9] Interlanguage, according to Selinker, is a unique learner-internal linguistic system which contains learners' abstract rules and hypotheses they establish when processing the target language. This provisional mental grammar constitutes a separate linguistic system which is different from a learner's L1 system and the linguistic system of the L2, but is likely to be influenced by both. While behaviorist learning theory emphasized the relevance of language input in the process of language acquisition, the concept of interlanguage holds that language acquisition processes are primarily determined by the LAD; input is necessary only in order to activate the LAD (e.g., Ellis 1997). With regard to this, the interlanguage system is described as being permeable, i.e., it might undergo changes and modifications due to input from the outside (language input) or due to internal mechanisms. For instance, learner errors, such as overgeneralizations or errors resulting from negative L1 transfer (see Section 3.1) which can be observed on the 'surface' are considered being symptomatic of a learner's internal processing strategies (Selinker 1979). This view is in line with Corder's (1967) argument that different types of L2 learner errors reflect the implementation of innate learning strategies. It also suggests that interlanguage systems are transitional, which means that the set of L2 rules stored in a learner's interlanguage system are constantly changing and expanding, and as a result, the L2 grammar becomes increasingly complex. Based on this, interlanguage can be described as a "dynamic system moving in the direction of the target language" (Lennon 2008, 55).

In the context of an evolving mentalist approach to second language acquisition, increasing attention was paid to investigating learners' interlanguage phonology, referred to as 'second language phonology' in current SLA (see e.g., Major 2008). Flege (1980), for example, assessed L2 segmental speech production in adult learners in order to determine whether the speakers' linguistic output reflects interlanguage phonology or whether this output is the result of L1–L2 interference processes. He investigated native Saudi Arabian speakers' realization of voicing contrast in L2 English plosives. The analysis of the target sounds revealed that more experienced Arabic speakers managed to approximate English phonetic norms for voiced and voiceless stops, while less experienced speakers' realization of the stop voicing contrast was found to be closer to Arabic norms. Based on these findings,

9 Since the introduction of the concept of interlanguage, many other terms have been used to refer to the same or similar underlying phenomena, such as 'approximative systems' (Nemser 1971), 'idiosyncratic dialects' (Corder 1971), or 'transitional competence' (Corder 1967). In addition, the term 'developing system' is frequently used, which acknowledges the fact that a learner's internal language system is constantly changing and developing as a result of their attempts to approach an L2 system similar to that of a native speaker (Wei 2009). In current SLA research, the preferred term is 'learner language' (Saville-Troike 2012).

Flege (1980) concluded that the learners' linguistic output is that of the speakers' interlanguage system rather than reflecting transfer processes from the L1. In addition, he emphasized that learners' interlanguage systems might develop at different rates, with some being faster in establishing phonetic categories for L2 speech sounds while others require more time to implement the phonetic norms of a new language.

One of the mechanisms which are assumed to influence and control the development of interlanguage systems is *fossilization* (Selinker 1979), which is frequently described as the 'incompleteness' (Schachter 1990), 'non-progression' (Han 2004b) or 'stagnation' (Nakuma 2005) of L2 acquisition. Earlier discussions of interlanguage phenomena viewed fossilization as the major property that distinguishes first from second language acquisition (e.g., Bley-Vroman 1990; Hyltenstam 1988; Schachter 1988): While L1 acquisition in typically developing children (i.e., children who do not suffer from any speech or language impairment) is usually successful in that they attain native mastery of their L1, an adult learner's L2 interlanguage system, on the other hand, is assumed to reach a point where no further progression or improvement in learning can be observed. Due to this 'end state' of L2 acquisition, late L2 learners typically acquire the L2 with varying levels of proficiency but are not likely to reach native L2 proficiency (Birdsong 1992). Selinker and Lamendella (1978, 187) define fossilization as a

> *permanent cessation of IL [interlanguage] learning before the learner has attained TL [target language] norms at all levels of linguistic structure and in all discourse domains in spite of the learner's positive ability, opportunity, and motivation to learn and acculturate into target society.*

As Selinker (1979) argues, fossilization is the reason why only a small percentage of L2 learners is actually able to achieve native or near-native competence in the second language. He exemplifies what fossilized phonological structures are by referring to typical errors observed in adult L2 learners: Native French learners of English, for example, might maintain the French uvular /R/ in their L2 English interlanguage, and native English learners of French maintain the English retroflex /ɹ/ in their L2 interlanguage.

One question that is repeatedly raised in the context of fossilization is whether it is a *global* phenomenon, i.e., whether it affects the entire interlanguage system, or a *local* phenomenon, applying to specific features of L2 learning only (e.g., Han 2004b; Larsen-Freeman 2005). Evidence seems to suggest that fossilization is selective in that it only affects certain linguistic features of the interlanguage system while other features are successfully acquired (see e.g., Long 2003; MacWhinney 1999). Tracing potential fossilization in an L2 learner's interlanguage system requires empirical longitudinal research (see Selinker and Han 2001, for a discussion)

and, in fact, most studies focusing on fossilization of L2 structures present long-term data (e.g., Han 1998; Jarvis and Pavlenko 2000; Lardiere 1998a, 1998b, 2000; Schumann 1978). However, studies examining fossilized linguistic structures have not presented entirely conclusive findings so far. Schumann (1978), for example, conducted a case study of an adult L1 Spanish speaker who learned English as an L2. The subject's use of different grammatical items was examined over a period of ten months, starting at the very onset of L2 learning, i.e., at a point where the learner had hardly any knowledge of the L2. The analysis of grammatical items revealed that the subject showed little linguistic development within the ten-months-period as he was observed to use "a reduced and simplified version of English" (Schumann 1978, 65). According to Han (2004a), this study is often viewed as the first to provide evidence for fossilization in an adult second language learner. However, interpreting Schumann's findings as evidence for fossilization mechanisms is rather misleading since it traces the subject's linguistic development over a period of ten months only, starting from the very onset of L2 acquisition. Therefore, it is not possible to determine to what extent the L2 has changed and perhaps even approached native production norms in the further course of language development. Identifying a lack of improvement within the first ten months of language acquisition does therefore not allow for definite conclusions to be drawn about whether certain linguistic features of the L2 learner's interlanguage system have fossilized or whether these features might have improved and changed at a later stage in the language acquisition process.

Another study worth noting in this context was conducted by Long (1997, see also 2003). In a 16-year-longitudinal investigation, he examined the linguistic development of an L1 Japanese adult speaker who had been living in Hawaii for 37 years at the time the study was conducted. The subject's productions of different grammatical L2 (Hawaii Creole English) structures were tested in spontaneous conversations at five points in time. It was found that while some productions were characterized by persistent errors, others showed a considerable amount of variation, which Long (2003, 510) interpreted as evidence against fossilization; he concluded:

> *What is most noticeable about Ayako's speech is that while performing far short of nativelike levels, with pervasive and persistent errors despite ample opportunity to acquire the target language, and so constituting an apparently perfect candidate for a fossilization claim, her interlanguage exhibits extensive amounts of variation, both synchronic and diachronic. Some of this variability may turn out to be systematic, but much of it appears not to be.*

Also Lardiere's research (1998a, 1998b, 2000) is frequently discussed in the context of fossilization in adult second language learners. Lardiere (1998a, 1998b) examined verbal agreement morphology in an adult L1 Chinese learner of L2 English. Similar

to the subject in Long (1997), the speaker had been long-term immersed in an L2 environment at the time of study. Recordings of natural productions were taken at different points in time. Lardiere's results indicate that the subject's morphological marking on verbs had fossilized over time, but not all grammatical features examined were observed to be affected by fossilization. This allows for the conclusion that some features of the interlanguage system are more likely to be affected by fossilization than others, and thus also supports the view illustrated above that fossilization does not operate on a global level but might affect only specific aspects of the interlanguage system.

As the previous discussion reveals, the concepts of interlanguage and fossilization contrast with earlier approaches to second language acquisition insofar as it can be reasonably assumed that L2 learning is not exclusively governed by transfer from the L1 system, but by the development of an L1- and L2-independent learner-internal system, containing abstract linguistic rules a learner establishes for the L2. Therefore, the interlanguage system is not static but develops and changes over time until it reaches an 'end state' of L2 proficiency.

As mentioned above, the vast majority of L2 learners is not likely to master the L2 in a native-like fashion as their interlanguage system is assumed to fossilize at a certain stage of learning. With regard to this, Scovel (1988) proposed that fossilization of adult learners' interlanguage phonology is ultimately responsible for their inability to produce accent-free L2 speech. However, research shows that a learner's interlanguage phonology can, in fact, attain a (near-)native state at least in some speakers (Bongaerts 1999; Bongaerts, Mennen, and van der Slik 2000). Selinker (1979) explains this phenomenon by referring to Lenneberg's (1967) description of a 'latent language structure', i.e., a genetically determined structure which is activated in children in order to acquire language. While Lenneberg states that the activation of this structure is eventually inhibited by maturational constraints, Selinker (1979, 58) argues that adults can be successful in acquiring a second language with native proficiency if they are able to "reactivate[ ] the latent language structure".

If and to what extent the concept of fossilization and its notion of an 'end state' of L2 proficiency is compatible with current theories and findings within the field of L2 acquisition will be further discussed in Section 1.7.

1.3.4 Shifting perspectives

With the development of the interlanguage concept a shift took place from investigating L2 acquisition phenomena from a teaching perspective, as it was the case in the context of CA, towards adopting a learning-oriented perspective. In order to arrive at a more thorough understanding of learner-internal mechanisms involved in

L2 learning processes, Selinker (1979) proposed analyzing learner language with regard to three systems: (1) a speaker's L1, (2) their interlanguage system, and (3) the L2 system. Empirical studies conducted along this line of inquiry in the 1970s increasingly focused on defining universal stages of second language acquisition in child and adult learners. In this context, it was explored if these stages are similar to the stages learners go through during native language acquisition (e.g., Butterworth and Hatch 1978; Dulay and Burt 1973) .

Dulay and Burt (1973), for instance, explored the acquisition of L2 English syntax in native Spanish children. Results of an error analysis revealed that only a small percentage of errors could be classified as interference errors while the majority were developmental errors, that is, L2 errors which are similar to those errors observed when monolingual children learn their first language. Dulay and Burt interpreted these findings to suggest that the acquisition of an L2 proceeds along the same 'natural route' as L1 acquisition does. Wode (1978) presented similar results in his assessment of phonological and syntactic data from L2 German and L2 English learners. He argues, however, that "L2 developmental sequences are not, on the surface of them, exactly like the respective L1 sequences" (Wode 1978, 115) as had been suggested by previous research (e.g., Corder 1967; Dulay and Burt 1973; Hatch and Wagner-Gough 1975). Instead, he proposed that L1 and L2 acquisition operate on the basis of the same set of principles, but that these principles might result in different surface forms in the respective language.

While universal principles and patterns of language acquisition were primarily investigated in young children, Butterworth and Hatch (1978) explored L2 acquisition in an adolescent speaker, acknowledging the fact that there are differences between children and adults in terms of language acquisition. In a case study of an L1 Spanish learner of English they found that the predominant strategy used by the L2 learner was simplification of L2 input and output. First language interference also played a role, but it was observed in a few instances only. Despite the fact that they assessed linguistic abilities in one speaker only, Butterworth and Hatch (1978) concluded that also adult language learners are very similar in their order of acquisition regardless of their L1 background.

Most of the studies outlined above focus on the acquisition of grammatical and syntactic L2 structures by child or adult learners while the development of interlanguage phonology did not receive as much attention in the early 1980s. However, although the development of L2 phonology is assumed to progress slower than the development of syntactic, semantic and lexical skills (Flege 1980), [10] L2

10 Flege (1980) states that one reason for this relatively slower progression in the development of L2 phonology might be that speech sound production involves a physiological component, i.e., the positioning and coordination of articulators, alongside with cognitive abilities.

phonological learning bears at least some similarities to native speech acquisition. These include, for instance, cluster reduction (e.g., Hansen Edwards 2004) or the segmentation of the speech stream (e.g., Ohala 2008). According to Ohala (2008, 35), the similarities between L1 and L2 phonological learning "suggest that the occurrence of such phenomena are not the explicit domain of either L1 or L2 phonological acquisition and that the reason behind these similarities may lie in universal tendencies".

Speaking of 'universal' tendencies when it comes to language acquisition processes and the development of interlanguage phonology conveys the impression that the interlanguage system is the same in every language learner at a particular stage of L2 acquisition. However, given that inter- and intra-learner variability is one of the most striking characteristics of L2 speech production and perception, the conclusion that "not two learners learn an L2 in exactly the same way" (Ellis 1991, 10) seems valid. The source of variability observed in learners' phonological data is determined by a number of different factors: leaner-external factors, such as quality and quantity of input (e.g., Flege, Frieda, and Nozawa 1997; Flege and Liu 2001; Moyer 1999), opportunities for language use, as well as learner-internal factors, among them motivation (e.g., Gardner 2010; Gardner and Lambert 1972; Gardner and MacIntyre 1993; Moyer 1999) and phonetic aptitude (e.g., Hinton 2012; Jilka et al. 2009; Jilka 2009). Language acquisition processes have also been shown to be influenced by linguistic variables, such as prosodic and segmental context, word shape, lexical category, and the presence or absence of orthography (see Colantoni, Steele, and Escudero 2015, 18). Hence, current SLA research acknowledges that there are stages of acquisition which are similar in L2 learners, but the transition from one stage to another might differ among learners due to the influence of some of the factors previously mentioned.

1.4. A current perspective in SLA and bilingualism research: First language attrition

1.4.1 Contextualizing first language attrition

As previously outlined, early observations concerning second language (speech) acquisition were predominantly discussed within the framework of L1 interference processes and maturational constraints (see e.g., Flege 2003; Piske, MacKay, and Flege 2001, for discussions). As the first language was assumed to reach a 'steady state' due to biological maturation mechanisms, it was claimed that the L1 system would remain unaffected by L2 learning experience, while the successful acquisition of an L2 was impeded by interferences from the first language if it was acquired

later in life (e.g., Lado 1957; Lenneberg 1967; Weinreich [1953] 1974). Thus, traditional SLA and bilingualism research seemed to provide a straightforward definition of the roles the native and the second language assumed in the context of L2 acquisition. However, with the development of the wholistic view of bilingualism (Grosjean 1989),[11] it was acknowledged that a bilingual's two (or more) language systems do not exist in isolation but constantly influence each other. Resulting from this changing perspective, it was advocated that "[linguists] will no longer examine one of the bilingual's languages without examining the other" (Grosjean 1989, 13). In the same vein, Schmid and Köpke (2007) argued that in order to be able to fully grasp the complexity of language development in bilingual speakers, it is not only necessary to explore how the L1 system influences additional language systems, but also to examine the extent to which the L1 system is affected by an evolving L2 system. Several studies which were conducted in the past decades support this view and challenge the claim that the first language is not prone to being affected by transfer processes (e.g., Flege 1987a, 1987b; Flege and Eefting 1988; Flege and Hillenbrand 1984; Flege, Schirru, and MacKay 2003; Major 1992; Mora, Keidel, and Flege 2015; Sancier and Fowler 1997). The findings of these investigations provide convincing evidence for bidirectional influences operating between the L1 and L2 system and, hence, support the observation that "there is no such thing as an end state" (De Bot, Lowie, and Verspoor 2007, 19) when it comes to both the L1 and the L2 system. Consequently, the L1 – similar to the L2 – can undergo changes and modifications in late second language learners due to L2 learning experience and long-term immersion in an L2-speaking context. This phenomenon is frequently referred to as 'first language attrition' and will be discussed in this section, focusing primarily on L1 *phonetic* attrition. Section 1.4.3 outlines some terminological and conceptual considerations concerning the use of the term 'first language attrition'.

One of the earliest investigations into the phenomenon of bidirectional influences was conducted by Flege (1987b). He examined VOT in plosives produced by native American English late learners of L2 French and native French late learners of L2 American English. His findings revealed that the subjects' production of the alveolar plosive /t/ was intermediate to the VOT norms of the respective target language. Flege explained this with reference to the Speech Learning Model (see Section 1.6.1), which is based on the prediction that similar L1 and L2 speech sounds

11 Originating from the wholistic view of bilingualism, several other theoretical frameworks have been developed, such as the integrated view of bilingual development (e.g., Schmid and Köpke 2007) and the notion of multicompetence (Cook 2003). The common ground of these concepts is that they aim to account for bidirectionality and mutual L1–L2 influence as inherent characteristics of a bilingual's linguistic systems. This view developed in opposition to the monolingual view of bilingualism which maintained that the bilingual is "the sum of two complete or incomplete monolinguals" (Grosjean 1989, 3) whose two language systems exist in isolation and are essentially independent of each other.

are assimilated due to the mechanism of equivalence classification (Flege 1987a, 1991, 1995; Flege and Hillenbrand 1984). One of the core findings of Flege's study was that the speakers' L1 plosives had moved away from native production norms, i.e., they were no longer produced within native (monolingual) ranges.

These and other more recent findings (e.g., Bergmann et al. 2016; de Leeuw 2019; Dmitrieva, Jongman, and Sereno 2010; Mayr, Price, and Mennen 2012; Mennen 2004) suggest that even a mature L1 system does not remain unaffected by L2 learning experience and is susceptible to change. As mentioned above, this phenomenon is commonly referred to as *first language attrition*, defined as non-pathological, non-age related changes in the fully developed L1 system which may result in a decline of proficiency in one's first language (Köpke and Schmid 2004). At the onset of attrition studies in the 1980s, language attrition, without exclusively referring to attrition processes in the *first* language, was described as "a special case of variation in the acquisition and use of a language or languages" (Andersen 1982, 86). Depending on the language which is 'lost' and the linguistic environment in which language loss occurs, van Els (1986) introduced four different types of attrition, as illustrated in Figure 1. Current attrition research predominantly focuses on Type II, L1 attrition in an L2 environment resulting from L2 learning experience, which is also the focus of the present thesis. However, L1 attrition does not have to be exclusively determined by the acquisition of a second language. According to Sharwood Smith and van Buren (1991), another reason for L1 attrition to occur is what they refer to as the 'desert island situation', that is, a situation in which speakers are being completely isolated from their first language – in reality, this is, however, rather unlikely to occur (Köpke 2004). In addition, L1 attrition processes have been observed to affect speakers' lexical and grammatical proficiency even without speakers having any knowledge of the L2, merely by living in an L2 environment. Baladzhaeva and Laufer (2018) examined L1 Russian speakers who migrated to Israel without acquiring and having contact to L2 Hebrew, i.e., they were solely interacting in their L1 within their immigrant community. Their study revealed that the monolingual immigrants exhibited the same amount of (or even more) L1 lexical and grammatical attrition compared to bilingual immigrants despite not learning or using the L2. Baladzhaeva and Laufer (2018) argue that the monolingual speakers experienced an indirect influence from L1-attrited Russian immigrants with whom they communicated within their language community, a phenomenon they refer to as 'second-hand attrition'.

In addition to dialect loss (Type I), the loss of the L1 in an L1 speaking environment might also occur as a result of aging and/or pathological conditions, or it might refer to the extinction of minority languages (Lambert and Freed 1982; van Els 1986). In more recent discussions on attrition phenomena a distinction is typically made between a decrease in L1 linguistic abilities observed in an individual (=

attrition) as opposed to language loss at a community/society level, which is referred to as 'language shift' (De Bot 1998, 345–46). Second language attrition (Type III) relates to the loss of a foreign or second language due to, for instance, returning to the L1 environment or due to a lack of contact with the second language. Attrition Type IV, 'language reversion', describes a loss of language which occurs rather late in life after migrating to a country in which a different language is spoken.

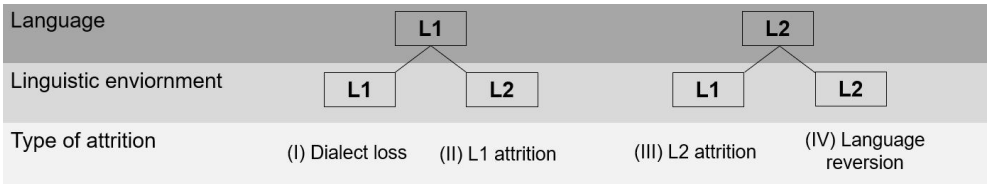


Figure 1: Four types of language attrition.

Current research into L2-induced attrition considers bidirectional influences an integral part of the dynamics of language learning processes, which is reflected in the view that "linguistic skills can grow and decline, and [...] accordingly, language acquisition and language attrition are equally relevant outcomes of developmental processes" (De Bot and Larsen-Freeman 2011, 6). While the *growth* of L2 linguistic skills has been thoroughly investigated since the emergence of SLA and bilingualism research in the second half of the 20th century, the *decline* of linguistic skills in the first language as a consequence of L2 learning experience is a relatively young field of interest (see Ecke 2004, 321–22; Schmid 2016). L1 attrition research in adult bilinguals (i.e., late consecutive bilinguals) has concentrated on language decline at all linguistic levels, such as morphology (e.g., Altenberg 1991; Keijzer 2010), syntax (e.g., Schmid 2002; Tsimpli et al. 2004; Yağmur 1997; Yağmur, de Bot, and Korzilius 1999), morphosyntax (e.g., Bergmann, Sprenger, and Schmid 2015; de Bot and Clyne 1994), the lexicon (e.g., Ammerlaan 1996; Schmid and Jarvis 2014; Yilmaz and Schmid 2013), and semantics (e.g., Jaspaert and Kroon 1992; Pavlenko 2004; Pavlenko and Malt 2011). Despite the fact that the number of studies examining attrition phenomena in the phonetic domain is still rather limited, phonetic/phonological attrition has been documented in different settings and with regard to different segmental (e.g., Bergmann et al. 2016; de Leeuw 2019; de Leeuw, Tusha, and Schmid 2017; Dmitrieva, Jongman, and Sereno 2010; Major 1992; Mayr, Price, and Mennen 2012; Stoehr et al. 2017) and prosodic features (e.g., de Leeuw 2019; de Leeuw, Mennen, and Scobbie 2012b; Mennen and Chousi 2018), as well as on the level of global foreign accent (e.g., Bergmann et al. 2016; de Leeuw, Schmid, and

Mennen 2010; Hopp and Schmid 2013; Mayr, Sánchez, and Mennen 2020; Sancier and Fowler 1997).

A frequently cited study in the context of L1 phonetic attrition was conducted by Major (1992), who examined VOT of voiceless plosives in L1 American English adult speakers with L2 Portuguese. The overall findings revealed that all subjects exhibited L1 attrition[12] to some extent, i.e., their productions of L1 English stops had moved towards the VOT norms of Portuguese stops (shorter VOT) and vice versa. Major (1992) argues that a lack of L1 use cannot be considered a significant factor triggering the attrition of L1 VOT categories in this study as all participants used English on a daily basis. Instead, he states that L1 attrition may be correlated with L2 proficiency as all of his participants were highly experienced L2 speakers. However, it needs to be taken into consideration that Major did not include control subjects with less L2 experience to test and compare potential effects of L2 proficiency on L1 decline or maintenance. In addition, a relatively small number of speakers ($N = 5$) was examined.

Phonetic attrition was also investigated by Sancier and Fowler (1997) in a longitudinal case study.[13] They examined VOT in voiceless stops produced by a native Brazilian-Portuguese female speaker who had begun learning L2 American English in late adolescence. The focus of this study was on the question of whether the fact that the speaker regularly travelled between the L1 (Brazil) and the L2 (US) environment would influence VOT productions in both languages. VOT measurements of Portuguese and English /p/ and /t/ showed that VOT was shorter and thus more Portuguese-like in both languages after the speaker had been staying in Brazil for a couple of months, and VOT was significantly longer and thus more English-like in both languages after the speaker had been staying in the US, being immersed in the L2. Sancier and Fowler (1997) take these findings as evidence for a 'gestural drift' of phonetic categories towards the norms of the most recently experienced language.

In a more recent study, Mayr, Price, and Mennen (2012) examined phonetic attrition in the speech of a late consecutive Dutch-English bilingual who emigrated to an L2 (English) speaking environment in early adulthood. Her speech was compared to that of her monozygotic twin sister who was also a late consecutive Dutch-English bilingual but remained in the L1 (Dutch) speaking environment. The study was based on the prediction that any systematic differences between the twins' L1 segmental productions can be explained by the fact that both lived in different linguistic environments. The results of an acoustic analyses of VOT and vowel

12 Major (1992) uses the term 'language loss' to refer to what is described as 'L1 attrition' in more recent research.

13 This study is frequently cited in the context of L1 (phonetic) attrition research. However, as will be further discussed in Section 1.4.3, the subjects investigated in Sancier and Fowler (1997) cannot be described as 'attriters' based on the definition of attrition applied in the present work.

formants indeed revealed differences between the twin sisters in some areas of pronunciation and thus confirmed the hypothesis that long-term L2 immersion affects L1 accent. However, systematic differences were not discovered in all areas of pronunciation: The twin living in the L2-speaking environment was, for instance, observed to produce the Dutch voiced plosives /b/ and /d/ in a native-like fashion. In addition, she produced English /b/ and /d/ with considerable prevoicing (voicing lead) typical of her L1, which indicates L1 transfer to the L2. Based on these findings, Mayr, Price, and Mennen (2012) suggest that some areas of pronunciation might be more sensitive to attrition than others.[14]

While VOT is perhaps the most widely studied feature when it comes to phonetic attrition affecting plosive production, also other consonants have been examined, such as the lateral approximant /l/ (Bergmann et al. 2016; de Leeuw, Mennen, and Scobbie 2012a) and post-vocalic /r/ (Ulbrich and Ordin 2014). Ulbrich and Ordin (2014) investigated the vocalization and realization of post-vocalic /r/ in two groups of subjects who spoke a non-rhotic variety of German as their L1, and rhotic Irish English or non-rhotic British English as their L2. Those bilinguals who had been exposed to a rhotic variety of English for an extended period of time showed signs of rhoticity in their realization of word-final /r/ in their L1, which suggests an L2-influence on L1 production.

Other studies focus on perceived global foreign accent in the L1 pronunciation of different groups of L2-immersed bilinguals (Bergmann et al. 2016; de Leeuw, Schmid, and Mennen 2010; Hopp and Schmid 2013; Mayr, Sánchez, and Mennen 2020; Schmid and Hopp 2014). De Leeuw, Schmid, and Mennen (2010), for example, explored global foreign accent in the L1 (German) of adult consecutive bilinguals living in a migrant context (Anglophone Canada or the Netherlands) with the aim to find out if the bilingual speakers' L1 pronunciation is perceived to sound less native compared to the German pronunciation of monolingual German speakers, assessed by monolingual German listeners. The foreign accent ratings obtained for the bilingual speakers were indeed significantly higher than the ratings obtained for the German monolingual controls, which suggests that the bilinguals were more likely to be perceived as foreign accented in their first language compared to monolingual control speakers. It was, however, not investigated which specific segmental and prosodic features contributed to higher foreign accent ratings.

14 It is generally acknowledged that attrition effects are selective, that is, not all linguistic (phonetic) features are equally affected by attrition processes (e.g., Köpke 2004). With regard to this, Andersen (1982) argued that language attrition proceeds along a continuum ranging from an early and complete decline to full maintenance of specific language features. However, the question of which specific features are more prone to attrition than others has not been investigated in greater detail so far.

As outlined above, most attrition studies focus on segmental productions and the perception of global foreign accent, while studies investigating attrition related to prosodic features are rare. One study worth noting in this context was conducted by de Leeuw, Mennen, and Scobbie (2012b), who examined L1 attrition in the prosody of L1 German speakers who had moved to Anglophone Canada in late adolescence or adulthood. The focus of the study was on the intonational feature of prenuclear alignment with the objective to find out whether and to what extend relative tonal alignment patterns were different in the L1 of the experimental (bilingual) group compared to a monolingual control group. The overall findings suggest that, despite a considerable amount of individual variation, some of the bilingual speakers in fact showed clear signs of L1 attrition with regard to the intonational feature under investigation, i.e., their L1 alignment of the prenuclear rise occurred earlier – as typical of L2 (British) English – compared to the control group. Hence, their productions had shifted away from monolingual German norms and fell closer within the L2 English range.

Taken together, empirical research into phonetic attrition suggests that being immersed in an L2-speaking country might be beneficial when it comes to attaining native-like pronunciation abilities in the L2, but at the same time a reduction in L1 use, which often goes along with using the L2 as a primary means of communication in an L2 setting, might lead to a decline of linguistic abilities in the L1 (see Stoehr et al. 2017, 484). However, despite the fact that language loss is frequently associated with a reduced use of the first language, there is only little evidence so far that the amount of L1 use can be considered a reliable predictive factor when it comes to attrition effects on the L1 system (see e.g., Cherciov 2012; Jaspaert and Kroon 1989; Major 1992; Schmid 2007). Schmid (2007), for example, investigated two groups of L1 German speakers being immersed in an L2 English and L2 Dutch environment, respectively. The aim of the study was to assess to what extend the amount of L1 use affects the bilinguals' verbal fluency and lexical range in their first language. The findings did not reveal a significant effect of everyday L1 use on the speakers' proficiency levels in the L1 and, based on that, Schmid (2007, 150) concluded that the "quality of [language] contact might be more important than quantity". Similar observations were made by Mayr, Price, and Mennen (2012) who examined phonetic attrition in L1 Dutch L2 English monozygotic twin sisters (see above). As both subjects used their L1 on a regular basis, the changes which were observed in the L2-immersed twin sister's L1 system could not be attributed to a lack of L1 use, but were explained with regard to fact that the sisters were living in different linguistic environments. Earlier observations concerning attrition considered a combination of both restricted language use and a change in linguistic behavior and/or environment being responsible for attrition processes to be triggered (e.g., Andersen 1982, 87). Current attrition research suggests that multiple linguistic

and non-linguistic factors influence L1 attrition, such as language learning motivation, length of residence in an L2-speaking environment, quality and quantity of L1 use and input, and age at the time of immigration (e.g., Schmid 2011, 2012; Schmid and Köpke 2007).

1.4.2 First language attrition: Permanent loss or temporary change?

Before discussing the use of the term 'L1 attrition' and the definition of L1 (phonetic) attrition applied in the present work in the subsequent section, it will be addressed if L2-induced changes observed in a late bilingual's L1 linguistic system are permanent and, thus, irreversible, or if such changes represent a temporary and reversible decline of L1 proficiency. In the context of L1 attrition, the permanency issue has been discussed by de Leeuw, Mennen, and Scobbie (2012a), who argue that a distinction needs to be made between changes relating to a speaker's L1 *competence* and modifications observed in their L1 *performance*.[15] While attrition affecting a speaker's competence, i.e., their underlying linguistic knowledge, would account for a permanent loss of linguistic abilities, changes occurring on the level of performance do not necessarily imply that the speaker is not able to reactivate linguistic knowledge on the level of competence.

Previous research suggests that attrition effects are likely to be reversible and, thus, affect a speaker's performance only. Stolberg and Münch (2010), for example, conducted a longitudinal study examining lexical-semantic, syntactic and morphological structures produced by a German-English bilingual, who had been long-term immersed in an L2 English-speaking country. The data were obtained from a range of spoken conversations between the bilingual and two native German speakers over a period of four years. The results revealed that the most significant changes within the four years of increased L1 exposure occurred in terms of the subjects' ability to re-access certain lexical items in her L1, that is, she showed fewer hesitations and word searches in the later recordings. Based on these findings, Stolberg and Münch (2010, 30) conclude that an attriter's first language skills are "being preserved at a specific level of competence that can be reached quickly as soon as this language is used again, even when an active L2 exists". Similar observations were made by Chamorro, Sorace, and Sturt (2015), who examined the effects of recent L1 exposure on the accessibility of L1 grammatical structures in long-term L2-immersed Spanish-English bilinguals. Similar to Stolberg and Münch (2010), their findings indicate that after being re-exposed to the L1 – in this case, re-exposure meant that the bilinguals had spent a minimum of one week in the L1 country

15 The distinction between competence and performance is understood in the sense of Chomsky who defined competence as speaker's knowledge of their language and performance as "the actual use of language in concrete situations" (Chomsky 1965, 3–4).

shortly before being tested – the subjects showed an increased ability in their use of the L1 grammatical features under investigation.

Although the studies outlined above seem to argue for the non-permanency of L1 attrition effects, their findings still need to be interpreted with caution. First, it is debatable if the subjects' ability to use certain L1 structures correctly represents an actual case of *relearning* these structures, or whether they represent a temporary imitation of immediate L1 input. Second, it remains unclear if and to what extent similar observations can be made when it comes to L1 pronunciation abilities. As previously discussed, pronunciation is often attributed a special status in the context of L2 acquisition (e.g., Scovel 1988), which becomes evident when considering that most L2 learners retain a detectable foreign accent if they acquire the L2 late in life, while often showing native or near-native mastery in other linguistic areas, such as reading and writing. Considering this special status of pronunciation, it might also be more difficult to relearn L1 pronunciation features which have been 'lost', despite recent re-exposure to the L1.

Furthermore, the studies discussed above do not necessarily test the speakers' level of competence. As outlined by de Leeuw, Mennen, and Scobbie (2012a) and as discussed in the previous sections above, the majority of studies examining L1 attrition in bilingual subjects are based on cross-linguistic data, comparing bilingual with monolingual speakers in order to identify if and to what extent bilinguals may diverge from monolingual patterns in their L1. Thus, cross-sectional studies test speakers' language use (performance) at a specific point in time but no conclusions can be drawn as to whether changes on the level of underlying language competence are evident or whether attrition effects observed at one point in time might be reversed at a later point. An empirical study testing an individual's linguistic competence would require testing two groups, one including potential attriters and the other including what de Leeuw, Mennen, and Scobbie (2012a, 686) refer to as "first-time learners", i.e., learners who are learning the L1 feature which has been 'lost' by the attriter group for the first time. If subjects in the attriter group are able to re-learn the L1 feature at a faster rate than the first-time learners, it can be concluded that the changes observed in the speakers' L1 system were only superficial and thus did not affect their underlying competence. However, finding two groups with a virtually identical level of performance is very unlikely, if not impossible. Therefore, de Leeuw, Mennen, and Scobbie (2012a, 688) conclude that "the question of whether or not L1 loss in late bilingual migrants is permanent, and hence a reflection of competence rather than performance, is [...] with all likelihood unanswerable" (See also Schmid 2008).

It is due to the methodological constraints outlined above that also longitudinal research spanning a period of several decades, as has been done in the present study, does not sufficiently address the permanency issue in L1 attrition. As no

feasible practical approach to reliably test the permanency of attrited L1 features seems to be available at this point, L1 attrition is not defined in terms of a loss of competence in the present context, but as a decline of linguistic abilities observed in a speaker's first language system, which might be (partially or completely) reversed due to a change in linguistic environment and linguistic behavior (see e.g., Chang 2019; Kartushina and Martin 2019).

1.4.3 When attrition should (not) be called 'attrition': Some terminological considerations

As discussed in Section 1.4.1, using the term 'attrition' to refer to different types of language loss or language decline goes back to the 1980s (e.g., Andersen 1982; van Els 1986). In current bilingualism and SLA research, attrition is mainly used to describe and explain changes and modifications which occur in a speaker's L1 linguistic system due to influences from an L2 system. The majority of studies examining L1 (phonetic) attrition focus on late consecutive bilinguals who are proficient L2 speakers and who are being immersed in an L2 environment, i.e., in a typical immersion situation (e.g., Bergmann et al. 2016; Hopp and Schmid 2013; Mayr, Price, and Mennen 2012; Ulbrich and Ordin 2014).

A recent definition provided by Schmid and Köpke (2017a, 637) describes L1 attrition as "phenomena that arise in the native language of a sequential bilingual as the consequence of the co-activation of languages, cross-linguistic transfer or disuse, at *any* stage of second language (L2) development and use" (see also Schmid and Köpke 2017b). Arguing that attrition effects can be observed at any stage of L2 development suggests that even inexperienced L2 learners who are not being immersed in an L2 environment might show signs of attrition in their L1 system as a result of L2 acquisition. In fact, the findings of a study conducted by Chang (2012; see also 2013) revealed that native English speakers with no prior L2 Korean experience learning the L2 in a formal instructional context in Korea showed modifications in their L1 vowel and plosive productions. Chang (2012) concluded from these findings that recent L2 experience, even if it is only brief experience, can be considered a significant influencing factor when it comes to changes in L1 production. Findings like these challenge earlier views that an L2 which is mastered at a very basic level only should not have a significant impact on the L1 while an L2 which is mastered with high proficiency is likely to have an increased impact on the first language (see e.g., Major 1992). Another study worth noting in this context was conducted by Kartushina, Frauenfelder, and Golestani (2016), who examined mutual influences between native French and nonnative Danish/Russian vowel production during learning. The participants in this study had pre-existing experience with Danish or Russian, and yet showed changes in their native vowel system after one

hour of articulatory training of nonnative vowels. Chang's (2012, 2013) and Kartushina et al.'s (2016) findings have in common that they provide evidence for the possibility of L1 changes to occur even in inexperienced L2-speakers; hence, the L1 modifications observed in these speakers cannot be attributed to a lack of L1 use or a long-term change in linguistic environment. Rather, they suggest that "cross-language linkages are established from the *onset* of second language learning" (Chang 2012, 249; my italics), which is also implied in Schmid and Köpke's (2017a) definition cited above. The question is, however, whether the term 'attrition' is equally appropriate to describe L2-induced changes in the L1 of both highly experienced bilinguals and inexperienced L2 speakers in different linguistic settings as suggested by Schmid and Köpke (2017a). As Chang (2012, 249) argues, the L1 changes observed in his study do not represent phonetic attrition, but are the result of a 'phonetic drift' of L1 categories towards the norms of the second language, i.e., a "subtle phonological restructuring in the L1 as a consequence of L2 experience". While Schmid and Köpke (2017a) consider attrition effects being observable even at the very onset of L2 development, Chang (2012, 264) states that L1 phonetic attrition and L1 phonetic drift represent two different phenomena along the same continuum, both being results of bidirectional influences between the first and the second language at different stages of L2 acquisition:

> *[i]nstead of being symptomatic of attrition, the occurrence of phonetic drift seems to be indicative of a fluid, multifaceted quality to language development over the lifespan [...] and it may be most accurate to think of phonetic drift as one step in a continuum of cross-linguistic effects in bilinguals dependent on relative use of the L1 vs. the L2.*

Based on their broad definition of language attrition, Schmid and Köpke (2017a) explicitly challenge the notion that there is a difference between cross-language influences from the L2 to the L1 system in *early non-immersed* bilinguals and the decline of the L1 *in long-term immersed, highly proficient* bilinguals. They argue that this differentiation is "both artificial and unhelpful, as they merely represent developmental stages on the same continuum" (2017a, 640). Resulting from this claim, they conclude that every bilingual – independent of their L2 proficiency level, linguistic environment, or L1 use – must be considered an L1 attriter (see also Schmid and Köpke 2017b). However, when considering language development as a continuum, expanding the definition of language attrition to the extent that every bilingual is an attriter – no matter at which stage in the continuum he or she is – is debatable. De Leeuw (2017) challenges this definitional expansion and argues that the findings of numerous studies which investigated attrition phenomena in bilingual speakers lead to the opposite conclusion, namely that *not* every bilingual is an attriter. De Leeuw, Mennen, and Scobbie (2012b), for instance, found that attrition did not

affect features of L1 prosody of all their German-English bilinguals and their results show a considerable amount of interpersonal variation. Another study worth noting in this context was conducted by de Leeuw, Tusha, and Schmid (2017), who examined phonological attrition in native Albanian speakers with L2 English. Their investigation of the realization of velarized versus non-velarized /l/ revealed that three out of ten subjects did not show any sign of phonological attrition and two bilinguals revealed signs of minimal attrition only. On the level of global foreign accent, de Leeuw, Schmid, and Mennen (2010) found that out of 57 late sequential bilinguals (L1 German, L2 English or Dutch) only 14 were clearly perceived to sound foreign accented in their L1 while 20 participants were judged to be native German speakers. These empirical findings support an earlier observation made by Schmid (2008, 11) who states that

> *[a]ny large-scale study of attrition will uncover substantial variability between individual speakers in the degree of decomposition or maintenance of the linguistic system. Typically, some speakers preserve native-like competence over a very long attrition span, while others seem to attrite quite dramatically within a few years of emigrating.*

Against the background of a definition which describes attrition as a phenomenon which can be observed in every bilingual, the findings of the studies previously outlined are difficult to interpret.

Schmid and Köpke (2017a, 638) quite rightly state that the term 'attrition' has been used to "refer to the particular and unique process of change" which takes place in a speaker's L1 linguistic system. However, describing these changes and modifications as 'unique' and 'particular' contradicts their argument that L1 attrition is observable in all bilinguals at all stages of bilingual development. By broadening the definition of attrition to this extent, attrition is no longer a particular and specific process of change, but encompasses all influences from the second language to the L1 system (i.e., cross-linguistic influences, language co-activation, transfer, etc.) and thus turns into a rather misleading umbrella term under which a variety of different phenomena are subsumed.[16] In order to be able to investigate

16 It might be due to this broad definition of attrition that Hopp and Schmid (2013, 366) point out that "[i]t is interesting that neither Flege (1987) nor Mennen (2004) use terms such as attrition or loss." The bilingual subjects examined in Mennen (2004) can, in fact, not be considered potential attriters – at least based on the definition of attrition applied in this work – because they had not been long-term immersed in an L2-speaking country. Therefore, it is actually not surprising that Mennen (2004) does not use the term 'attrition' to refer to the changes observed in some speakers' L1 prosodic systems. Similarly, not all subjects in Flege (1987b) can be described as potential attriters: Out of the six speaker groups examined in this study, only two would qualify as potential attriters, namely late consecutive bilinguals who have been living in an L2-immersion context for an extended period of time.

bilingual development which takes place under different circumstances and in different contexts in a more differentiated manner, it is necessary to apply a narrower definition of L1 attrition as suggested by de Leeuw (2017). Therefore, the definition of attrition applied in the context of the present investigation is based on the following arguments:

(1) Bilingual language development proceeds along a continuum. The different phases or steps along this continuum are likely to result from the constant interactions and mutual influences between a speaker's language systems (e.g., de Bot 2004). In line with an integrated view of bilingualism (e.g., Grosjean 1992) and a multi-competence perspective (e.g., Cook 1991, 2003) on bilingual development, influences of the first language on the second language system and vice versa are inherent aspects of the dynamics of language development and, thus, are nothing unusual or unexpected. Therefore, cross-language influences are experienced by all bilinguals to a greater or lesser extent.
(2) When considering the opposite ends of the language development continuum, we find, on the one hand, L2 learners who start off with no L2 learning experience and who experience (superficial) modifications and changes in their L1 system due to recent and short-term L2 exposure (e.g., Chang 2012, 2013). These L1-modifications are likely to be non-permanent and be – partially (Chang 2019) or completely (Kartushina and Martin 2019) – reversed if the speakers receive, for instance, intensive L2-training for a limited period of time only and are then re-immersed in their L1 linguistic environment. On the other end of the continuum we find highly proficient L2 speakers (e.g., Hopp and Schmid 2013; Major 1992; Mayr, Price, and Mennen 2012; Ulbrich and Ordin 2014) who are being immersed in an L2-speaking environment for an extended period of time and use their L2 on a regular basis in different contexts, which may also go along with a reduced L1 use.
(3) Describing speakers at both ends of this continuum as (potential) 'attriters' ignores the very different circumstances under which language acquisition, use and development take place. *Attrition* effects resulting from L2-to-L1 influences might be observed in *late adult bilinguals who are highly proficient L2-dominant speakers or fluent bilinguals and are being immersed in an L2-speaking environment*, i.e., in a typical migration setting. Consequently, attrition will be defined as a decline in L1 proficiency experienced by bilinguals who acquired the L2 late in life after the L1 had already been fully developed and who are being immersed in an L2-speaking environment. With regard to this, *phonetic* attrition is defined as the loss of native-like pronunciation abilities in the L1 due to influences of an L2 phonetic system experienced by late consecutive bilinguals living in an

attrition situation, which might even result in speakers being perceived as having a foreign accent in their native language (e.g., Bergmann et al. 2016; de Leeuw, Schmid, and Mennen 2010; Hopp and Schmid 2013).

1.5. Dynamic systems theory

1.5.1 Language development from a dynamic systems perspective

Dynamic Systems Theory (DST), a theoretical approach to the study of development, originally developed as a branch of mathematics and has been emerging in the past two decades. It was, however, soon applied to other research areas such as biology, chemistry, and psychology (see Smith and Thelen 2003; Thelen and Smith 2006). The core assumption of DST is that all dynamic systems consist of a set of intertwined variables that change over time (e.g., de Bot and Larsen-Freeman 2011; de Bot, Lowie, and Verspoor 2007). While human developmental processes can generally be regarded as complex dynamic systems, dealing with language acquisition and development within the framework of DST is a relatively new approach in the field of SLA, which was first extensively discussed in Larsen-Freeman's paper on chaos/complexity theory (Larsen-Freeman 1997). As previously outlined, SLA is a multifaceted research area which addresses various aspects likely to influence second language acquisition processes. A DST approach to second language acquisition aims to investigate and explain these processes from a non-linear perspective based on the assumption that language acquisition – as other developmental processes – is characterized by a complex interplay of internal and external variables within a continuously changing system (de Bot, Lowie, and Verspoor 2007; Larsen-Freeman 1997, 2000). Among these variables are situational factors, linguistic input, and individual learner differences, such as language learning aptitude and motivation (see Ellis 1991, for an overview).

De Bot, Lowie, and Verspoor (2007, 19), among others, advocate the application of a DST approach in SLA research by proposing that "we can no longer work with simple cause-and-effect models in which the outcome can be predicted, but [that] we must use case studies to discover relevant sub-systems and simulate the processes". While linear cause-and-effect models of language learning and processing hold that L2 learners complete stages of acquisition which are assumed to be highly similar in every L2 learner (e.g., Dulay and Burt 1974b; Dulay, Burt, and Krashen 1982), a DST approach considers variability an inherent part of any kind of development (de Bot, Lowie, and Verspoor 2007). Variability is seen a result of the constant dynamic changes traceable in developing systems and provides evidence for the observation that learners may show both phases of regression and progression in

their linguistic development (see van Dijk, Verspoor, and Lowie 2011, 58). As de Bot, Lowie, and Verspoor (2007, 15) point out, variability is usually a major characteristic of second language learner data, even among relatively homogenous groups of learners. However, most SLA research so far has primarily "attempted to explain away seemingly FREE VARIATION by assuming full systematicity of variation in learner data" (ibid.) instead of recognizing intrapersonal and interpersonal variability as actual data which can be a valuable source for understanding the nature of second language developmental processes. A DST approach to second language learning acknowledges the existence of variation and change, and views them as integral processes and outcomes inherent to language learning.

As discussed in Section 1.3.3, the phenomenon of fossilization is often associated with terms such as 'non-progression' (Han 2004b) or 'stagnation' (Nakuma 2005) of language learning, which implies that second language learning eventually reaches a state where no further development can be observed. (How) Is it possible then to relate the view that language development is dynamic and ever-changing as proposed by DST to the assumption that there is an end state to language learning? Without specifically referring to language development, but to developmental processes in general, Thelen (1995, 77) states that development "can be envisioned as a changing landscape of preferred, but not obligatory, behavioral states with varying degrees of stability and instability, rather than a prescribed series of structurally invariant stages leading to progressive improvement." Along similar lines, Larsen-Freeman (2005) argues that since language development is not a linear and even process, fossilization might affect some features of the linguistic systems while others remain unaffected and improve further. Hence, applying a DST approach to language acquisition acknowledges that instability and continuous development on the one hand, and systematicity and stability on the other hand, are not mutually exclusive phenomena. Rather, stability and instability as well as variability are to be seen as inherent characteristics of language learning and development – a view that is also proposed in the context of the present work.

In order to examine variability and its role in second language acquisition processes from a DST perspective, researchers demand conducting longitudinal studies with a corpus of dense data (see e.g., van Dijk, Verspoor, and Lowie 2011). However, compared to cross-sectional investigations, the collection of longitudinal data is rather time-consuming and costly, which might be one reason why the number of longitudinal studies using developmental data is still limited. A solution to this problem is suggested by Thelen and Smith (1994), who state that long-term investigations do not necessarily require the collection of new data, but might also be conducted using already existing data collections and investigate them from a new, dynamic systems-oriented perspective.

Besides potential problems which go along with collecting longitudinal data, a DST approach to SLA entails further methodological issues as it is impossible to meet the challenge of grasping the *full complexity* of language learning as proposed by a DST approach. As a matter of fact, SLA research lacks applicable tools and methodology to fulfil the task of approaching L2 learning processes from an all-encompassing perspective. Therefore, the optimal approach, according to de Bot, Lowie, and Verspoor (2007, 18), "would be one in which the representation of the full complexity of the systems is linked to attempts to reduce that complexity separating highly relevant information from arguably less relevant information". It might be due to these methodological challenges that SLA and bilingualism research examining language development from a dynamic systems perspective is still relatively rare. Most work on DST in the field of SLA discusses the theoretical foundations of DST and its applications to second language acquisition studies (e.g., de Bot 2008; de Bot, Lowie, and Verspoor 2007; van Geert 2009).

1.5.2 Language attrition and dynamic systems theory

In the field of psycholinguistics, the relationship between language attrition and DST was first established by Herdina and Jessner (2002, 2013), who focused on language acquisition and decline in multilingual speakers in the context of the *Dynamic Model of Multilingualism* (Herdina and Jessner 2002).Well before the application of DST to language attrition, however, Andersen (1982, 86) established a connection between the main tenets of DST and language attrition by emphasizing the need to examine attrition phenomena "within a framework that includes all other phenomena of language acquisition and use". The common ground of discussions dealing with language attrition from a DST perspective is that language systems are dynamic in nature, i.e., they are continuously developing and changing (e.g., Herdina and Jessner 2013, 752). This suggests that even a mature and apparently fully developed L1 is sensitive to modifications due to, for instance, L2 acquisition experience and changes in the linguistic environment. Thus, a potential restructuring of linguistic skills is considered as "a normal part of language development and not as a lamentable loss of what once was" (De Bot 2004, 233). By contrast, proponents of a critical period in language acquisition (e.g., Lenneberg 1967; Scovel 1969, 1988) argued that an L1 system is not susceptible to change once it is fully developed after the end of the critical period. Therefore, any theory narrowly based on maturational constraints and on the existence of a critical period does not consider language attrition likely to affect a fully matured L1 system. A study conducted by de Leeuw, Mennen, and Scobbie (2012a) aimed to compare the predictive power of maturational constraints to DST by examining the production of the lateral phoneme /l/ in late German-English bilinguals residing in an L2-speaking country. The

results revealed that some subjects produced German /l/ with higher first formant values compared to monolingual controls, i.e., their productions of the L1 target sound were observed to have moved towards the norms of the L2. In addition, a large amount of intra- and interpersonal variation was found. De Leeuw, Mennen, and Scobbie (2012a) interpret these findings as evidence *against* maturational constraints as all participants had acquired the L2 in adulthood when the L1 had already been fully developed. Their findings are discussed in the context of DST suggesting that the variability observed within and across speakers can be "interpreted as evidence for the efficacy of additional predictor variables not specified within a theory founded on maturational constraints" (De Leeuw, Mennen, and Scobbie 2012a, 684). These findings are in line with de Leeuw et al.'s (2013) argument that extracting a single variable – such as *age* – from the system and attribute overall predictive power to this individual element "may be unproductive because the impact of one element within the system or subsystems on other elements and (sub-)systems runs the risk of being ignored" (2013, 671; see also Herdina and Jessner 2013). Cherciov (2012) further specifies this argument by stating that a DST approach acknowledges that influences on both the first and the second language might be triggered by a variety of different variables, including motivational and attitudinal factors, as well as language use and language choice.

Possible interactions of personal and environmental factors with L2 acquisition and L1 attrition processes were also part of Opitz' (2011, 2012) cross-sectional investigation. Her study set out to examine the overall linguistic development of adult German-English bilinguals who were immersed in an L2-speaking environment. Based on the findings obtained by means of extensive sociolinguistic questionnaires and (self-)ratings of language proficiency, Opitz (2012) suggests that linguistic as well as extra-linguistic factors, such as language use and input, linguistic aptitude and language attitudes, are relevant in the context of both successful second language acquisition and first language maintenance.

1.6. Models of second language speech production and perception

In the subsequent sections, two current models of second language speech learning will be outlined: the Speech Learning Model, developed by Flege and his colleagues (Flege 1987a, 1987b, 1991, 1995, 2002, 2003; Flege and Hillenbrand 1984) and the Perceptual Assimilation Model by Best et al. (Best 1994, 1995; Best and Tyler 2007).[17]

17 Another model which is often discussed in this context is the Native Language Magnet model developed by Kuhl and her colleagues (Iverson and Kuhl 2000; Kuhl 2000, 2007; Kuhl et al. 2008; Kuhl and Iverson 1995). This model is predominantly concerned with infant speech development and is based on the assumption that babies are equipped with innate mechanisms allowing them to discriminate phonetic contrasts which are non-specific to their L1 (see e.g., Kuhl 2000). After the first year of life, this ability decreases and the child turns into a language-bound listener. According to Kuhl (2000,

These models have been developed to account for differences in L2 production and perception accuracy observed in second language speakers and aim to predict and explain potential difficulties inexperienced and experienced L2 learners might face when acquiring second language speech sounds (see Strange 2007, 36). The common ground of these models is that they "consider the representations established for the L1 to be crucial for the manifestation of foreign accent in the L2" (Jilka 2009, 8). Based on this, language transfer is assumed to be more likely to occur between *similar* L1 and L2 segments, that is, accurate perception and, consequently, accurate production, are determined by cross-language similarity.

1.6.1 The Speech Learning Model

The Speech Learning Model (SLM; Flege 1987a, 1987b, 1991, 1995, 2002, 2003; Flege and Hillenbrand 1984) is perhaps one of the most frequently cited and tested models in SLA and bilingualism research (see Flege and Bohn 2021, for a revised version of the SLM, SLM-r). It provides a comprehensive framework for understanding how the phonetic categories of a bilingual "reorganize over the life span in response to the phonetic input received during naturalistic L2 learning" (Flege and Bohn 2021, 3) and is based on several postulates and hypotheses which serve as a sound basis for making testable predictions. Studies conducted within the framework of the SLM (e.g., Aoyama et al. 2004; Flege, MacKay, and Meador 1999; Flege, Munro, and Skelton 1992; Flege, Schirru, and MacKay 2003; Guion et al. 2000) focus on *advanced* L2 learners since the model's predictions are concerned with *ultimate attainment* of L2 sounds in bilingual speakers (see Flege 1995). It should be noted though that in the context of the revised SLM, the focus is no longer on an assumed 'end state' of language learning, that is, on highly experienced bilinguals; instead, the SLM-r also takes less experienced L2 speakers into account.

One of the basic assumptions of the SLM is that the ability to produce L2 sounds without a detectable foreign accent is influenced by age. Late L2 learners are assumed to be less successful when it comes to distinguishing phonetic differences between L1 and L2 sounds than learners who acquired the L2 from an early age on, i.e., the earlier an L2 is acquired, the more likely it is that L2 sounds are produced in a native-like manner. However, at stated by Flege (1995), the learning mechanisms which are responsible for the successful acquisition of the L1 phonological system are available over a speaker's lifespan and are consequently also accessible

11854), language-specific filters which have been developed during the first months of life hinder second language speech learning because "the mapping appropriate for one's primary [first] language is completely different from that required by other languages". However, as this model mainly focuses on native speech development in young children, it will not be considered in more detail here.

when learning a second language. Hence, achieving native or near-native proficiency in an L2 might be restricted by age-related factors, but is still possible due to the continuous availability of language learning mechanisms. This assumption does not only contradict any strong interpretation of a critical period hypothesis for the acquisition of L2 pronunciation (e.g., Scovel 1969), but at the same time also suggests that there might be no such thing as an 'end state' of language acquisition, as proposed by theories of fossilization (e.g., Selinker 1979; Selinker and Lamendella 1978).

Whether or not second language pronunciation is mastered in a native or near-native fashion is closely linked to the relationship between perception and production. According to the SLM, many second language production errors result from incorrect perception, that is, if differences between an L1 and an L2 speech sound are not perceived accurately, they are less likely to be produced in a native-like manner (e.g., Flege 1995, 2003).[18] Furthermore, the SLM postulates that (in-)accurate productions depend on whether a bilingual is able to establish separate phonetic categories for individual L1 and L2 speech sounds. If the speaker succeeds in forming distinct phonetic categories which resemble a monolingual's categories, L2 pronunciation is assumed to be accurate (Flege 1995, 240). In this context, phonetic categories are defined as representations of language-specific properties which are stored in a bilingual's mind. According to Flege (1995, 238), a bilingual's L1 and L2 phonetic categories coexist in a shared phonological space.[19] Within this shared phonological space, bilingual speakers aim at maintaining contrast between first and second language categories by applying different mechanisms, specified in the SLM's hypotheses.

Based on the relationship between production and perception described above, the SLM predicts that establishing a new phonetic category for an L2 speech sound which has a similar but phonetically different counterpart in the L1 will be successful if the speaker is able to perceive subtle phonetic differences between the sounds. For example, the voiceless alveolar stop /t/ exists in both English and Spanish; however, they differ phonetically insofar as Spanish /t/ is typically realized with short-lag VOT while native English /t/ is produced with long-lag VOT values (see e.g., Flege and Eefting 1988; Flege and Schmidt 1995). An accurate production of the L2 speech sound presupposes that the speaker has successfully perceived the

18 In the context of the SLM-r, the relationship between production and perception is slightly redefined: While the SLM, in its original formulation, considers accurate perception a prerequisite for production accuracy, the SLM-r proposes a bidirectional relationship between perception and production, as specified in the model's 'co-evolution hypothesis' (Flege and Bohn 2021, 29). That is, accurate perception does not necessarily have to precede accurate sound production, but the two might co-evolve and influence each other.

19 Note that in their discussion of the SLM-r, Flege and Bohn (2021) use the term 'common phonetic space'.

phonetic differences between the L2 target sound and the closely related L1 sound and, based on this, is able to establish two separate phonetic categories for the respective sounds. However, category formation for the L2 sound might be impeded by the mechanism of equivalence classification (category assimilation; see Flege 2003), that is, a speaker fails to perceive phonetic differences between two sounds and classifies them as representations of the *same* sound. Category assimilation may result in the emergence of what Flege terms 'compromise values' (e.g., Flege 1987b, 51). Referring back to the example mentioned above, this would imply that an L1 Spanish L2 English bilingual establishes a merged L1–L2 category and produces the target sound with VOT values intermediate to the monolingual norms of short-lag Spanish and long-lag English, respectively.

Another mechanism operating in bilingual speakers is category *dissimilation* (Flege 2003). In order to maintain difference between closely related L1 and L2 sounds in the shared phonological space, a bilingual's phonetic categories might be 'deflected away' from the native categories (Flege 1995, 239). This hypothesis was tested and confirmed in a case study by Mack (1990), who found that a French-English simultaneous bilingual produced the voiceless plosives /p t k/ in both languages with VOT values being considerably longer than those produced by monolingual French and English controls. Despite the fact that the speaker realized the target sounds outside the native range, the bilingual subject had been able to establish one distinct long-lag category for each of the two languages and thus maintained contrast between the L1 and L2 phonetic categories.

An earlier study conducted by Flege and Hillenbrand (1984) set out to examine the production of the voiceless alveolar plosive /t/ in native French learners of English. In line with the SLM's predictions, the findings revealed that the speakers realized /t/ with compromise VOT values in their native language, i.e., they produced values outside the short-lag VOT norm in French. Against the background of the SLM, it can be assumed that the speakers had failed to establish separated phonetic categories for French and English /t/, respectively, due to their perceived similarity. Another part of the study included native English learners of French who were tested on their production of the French vowels /y/ and /u/. The experiment revealed that the subjects failed to produce French /u/ accurately while they managed to produce /y/ in a native-like manner. These findings can also be explained in the context of the SLM: The French vowel /y/ does not have a direct counterpart in the English vowel system, which is beneficial in terms of production accuracy because the target sound is not likely to be equated with a similar L1 sound. French /u/, on the other hand, has a similar counterpart in English. Due to the mechanism of category assimilation, the bilinguals examined in Flege and Hillenbrand (1984) established a merged L1–L2 category "that subsumes the properties of phonetically

distinct L1 and L2 speech sounds that have been perceptually equated" (Flege 2003, 330).

One of the primary tenets of the SLM is the assumption that the interactions taking place between the L1 and L2 system are bidirectional in nature (Flege 1995, 241). As outlined in Section 1.3, early SLA theories maintained that interference and transfer processes are inherently unidirectional, i.e., taking place from the first to the second language, but not vice versa (e.g., Lado 1957; Weinreich [1953] 1974). Studies conducted within the framework of the SLM (e.g., Aoyama et al. 2004; Flege and Hillenbrand 1984; Guion et al. 2000; Guion, Flege, and Loftin 2000), however, suggest that a bilingual's languages mutually influence each other, which does not only lead to nonnative productions in the L2, but might also affect the production accuracy of L1 speech sounds. In addition, the SLM challenges the traditional understanding that structural *differences* between the first and the second language are ultimately responsible for production difficulties (e.g., Moulton 1962a, 1962b) by proposing that perceived *similarities* between L1 and L2 speech sounds are more likely to lead to inaccurate productions. This means that if bilingual speakers perceive two speech sounds to be dissimilar, they are more likely to create separate phonetic categories for the target sounds and thus produce them in a more native-like fashion (Flege 1995, 239). It should be noted, however, that not all difficulties in accurate sound production are perceptually motivated, but might be the result of motoric and articulatory constraints. As pointed out by Flege (1980, 1987b), adult L2 learners are likely to be less successful compared to children when it comes to adapting to new or modifying already existing articulatory patterns, that is, a speaker might be well able to identify phonetic differences between L1 and L2 sounds, but fails to produce them correctly due to an inability to position and coordinate articulators correctly.

1.6.2 The Perceptual Assimilation Model

The Perceptual Assimilation Model (PAM; Best 1994, 1995; Best and Tyler 2007) was originally developed to account for cross-language speech perception effects in infant and adult listeners (see Best 1995). While the SLM focuses on experienced and highly proficient L2 speakers, the PAM was designed to explain the perception of nonnative speech by naïve (monolingual) listeners (see Best and Tyler 2007). However, as naïve and L2-learning listeners considerably differ in terms of nonnative speech perception, the main principles of the PAM were extended to L2 speech perception, resulting in the development of the PAM-L2 (Best and Tyler 2007).

One of the basic propositions of the PAM is that while young children are capable of discriminating nonnative contrasts which do not exist in their native language, this ability is no longer available in adult speakers (see Best 1994). The

changes in perception ability, which have been shown to commence at the age of around 0;5, are said to be influenced by the development of the first language, i.e., in the course of the first months of one's life speakers become attuned to their native language, which eventually affects the perception of nonnative speech (Best 1995). With regard to this, the PAM posits that native language effects on nonnative contrasts result in "the perceptual assimilation of nonnative phones to the native phoneme categories with which those nonnative-phones share the greatest similarity in phonetic characteristics" (Best 1994, 168). The mechanisms of perceptual assimilation are said to take effect only if a nonnative phonetic contrast is perceived to be similar to a native contrast. If a listener is able to perceive differences between a nonnative segment and an L1 segment, no assimilation will take place, resulting in accurate production (Best 1994). This principle mechanism is assumed to be ultimately responsible for young children being able to master the language(s) they learn in a native fashion, while adults often speak a late-acquired L2 with a detectable foreign accent.

However, the influence of first language experience on the perception of nonnative contrasts is said to not necessarily be permanent as it can change due to perception training or extensive conversations with native speakers (Best 1994). Therefore, the likelihood of producing L2 speech sounds in a native-like fashion is enhanced by increased L2 experience, i.e., the more experienced an L2 speaker is, the more likely it is that nonnative contrasts are discriminated accurately (Best and Tyler 2007).

The main principles of the PAM and the SLM overlap to a certain extent in that, for instance, both models make predictions based on perceived similarities between the native and the nonnative language. In addition, the SLM as well as the PAM assume that perceptual learning abilities are available over the lifespan (Best and Tyler 2007). However, while the SLM is concerned with individual phonetic categories, the PAM focuses on "pairwise phonological contrasts" (Best and Bohn 2015, 44). Furthermore, the SLM aims to predict production (in)accuracy influenced by perception, and the PAM, as discussed above, focuses on cross-language speech perception.

1.7. Interim conclusion

As mentioned in the previous section, the SLM and the PAM are based on the assumption that language learning mechanisms are available across the lifespan (e.g., Best and Tyler 2007; Flege 1995), which implies that also late L2 learners are able to acquire native or near-native proficiency in the L2. In addition, both models emphasize the mutual interaction between the L1 and L2 system in the course of the language learning process and thus contradict the view that the L1 system is

stable and resistant to change once it is fully developed, as proposed in theories of contrastive analysis (e.g., Lado, 1957) and maturational constraints (e.g., Scovel 1969, 1988). Evidence for the malleability of a learners' L1 and L2 language system, which has been provided by studies conducted in the field of L2 speech acquisition (e.g., Flege, Schirru, and MacKay 2003; Flege, Takagi, and Mann 1995; Simon 2009) and L1 attrition (e.g., Bergmann et al. 2016; de Leeuw 2019; Stoehr et al. 2017), supports a dynamic view of language learning, as outlined in Section 1.5 (e.g., de Bot 2007, 2008; Larsen-Freeman 2005). In the same vein as it argues for the continuous development of the first language system, a dynamic view of language learning contradicts the notion of an 'end state' of L2 proficiency as proposed by theories of fossilization (e.g., Selinker 1979; Selinker and Lakshmanan 1992; Selinker and Lamendella 1978). While the predictions made by the SLM have been tested and confirmed in numerous empirical studies (e.g., Aoyama et al. 2004; Flege and Eefting 1987b, 1988; Guion et al. 2000; Guion, Flege, and Loftin 2000), research into fossilization (e.g., Long 1997, 2003; Schumann 1978) has not provided conclusive findings so far. One reason for the lack of reliable evidence for the occurrence of fossilization phenomena can be found in methodological problems (see Han 2005; Selinker and Han 2001). Even in longitudinal studies which trace learners' L2 development over an extended period of time (i.e., several years or even decades) it is hardly possible to make reliable statements about a learner's *final* state of L2 learning. Certain features of the learner's language system might be identified as having fossilized, but it is still possible that they further develop at a later stage in the learning process due to, for example, a change in linguistic environment, increased contact with native speakers, or specific pronunciation training. Thus, the initial claim that fossilization represents a "permanent cessation of [interlanguage] learning" (Selinker and Lamendella 1978, 187; my italics) cannot be reliably supported.

The notion of an end state is also implied in the critical period hypothesis, which – in its strong interpretation – holds that language acquisition processes are ultimately controlled and potentially constrained by biological ageing mechanisms, i.e., acquiring an L2 with native proficiency is only possible within a limited time span (e.g., Lenneberg 1967; Penfield and Roberts 1959; Scovel 1969, 1988). Research into the relationship between age and L2 pronunciation proficiency (e.g., Bongaerts et al. 1997; Bongaerts 1999; Bongaerts, Mennen, and van der Slik 2000; Flege, Yeni-Komshian, and Liu 1999; Moyer 1999) stresses the importance of age-related factors, but – at the same time – documents the existence of highly successful L2 acquisition in learners who acquired the L2 past the end of an assumed critical period. In this context, additional factors have been shown to influence L2 acquisition (see Section 1.3.2). Stressing the interplay between system-internal and system-external factors further supports a dynamics-oriented perspective on L2 acquisition.

Resulting from the previous discussion, language acquisition and linguistic development, as defined in the context of the present work, are based on the following principles:

(1) In line with a DST approach, (second) language learning is regarded as an ongoing and constantly changing process, which is characterized by a dynamic interplay between external and internal variables.
(2) As proposed by the SLM and the PAM, language learning strategies and mechanisms are available across the lifespan, which accounts for the dynamic nature of language learning. Therefore, acquiring an L2 with native or near-native proficiency is also possible for late L2 learners.
(3) Age does play an important role when it comes to L2 acquisition, yet it must be regarded in the context of additional factors which might influence successful L2 acquisition.
(4) The L1 and L2 linguistic systems mutually influence each other. Therefore, any investigation into bilingual language development necessarily requires studying both the first and the second language system.

1.8. Social and cultural dimensions of the phenomenon of foreign accent

As the previous sections revealed, second language speech acquisition is a complex field of research which has received much attention in the past decades. Several approaches and models have been developed with the aim to explain the causes of foreign accented speech in nonnative speakers (see e.g., Best 1994; Flege 1995). In addition, more recent investigations into processes of L1 attrition have shown that L2 learning experience and long-term L2-immersion can, in fact, influence L1 pronunciation abilities which may lead to speakers having a nonnative accent in their first language (see e.g., Bergmann et al. 2016; de Leeuw, Schmid, and Mennen 2010; Hopp and Schmid 2013). The discussion and juxtaposition of approaches in the field of L2 acquisition and L1 attrition leads to a range of questions which are inevitably related to the phenomenon of foreign accent: (Why) Is it necessary to be perceived as a native speaker in a second language? Which consequences might a foreign accent in both the first *and* the second language entail in different social and cultural settings? Why would speakers deliberately maintain a nonnative accent in their speech?

The following definition of foreign accent (Flege 1988, 229) serves as a starting point for discussing social, cultural and identity-related dimensions of this phenomenon:

> *Foreign accent is a phenomenological experience of listeners that is derived from detectable acoustic (and perhaps visual) differences between native and nonnative speakers in the pronunciation of sounds or other speech units. The perception of foreign accent leads to the conscious or unconscious realization by the listener that the interlocutor is not a fellow native speaker.*

While foreign accent is accepted as a widely observed phenomenon in adult L2 speakers in current bilingualism and SLA research, earlier discussions suggest that speaking a language with a nonnative accent was treated as "imperfect or defective speech" (Greene and Wells 1927, 24) which required some sort of therapy. Griffen (1980), for instance, maintained that the ultimate goal of teaching pronunciation should be to enable learners to speak the foreign language without any indication that they are nonnative speakers. Abercrombie (1949), by contrast, argued that it is more essential for L2 learners to speak the L2 in a way that is easy to understand for their interlocutors than to attain native pronunciation abilities. This view is also represented in current teaching-oriented discussions which stress the importance of being able to communicate effectively and making oneself understood in the L2, which does not necessarily presuppose a native L2 accent (see e.g., Kennedy and Trofimovich 2008).

However, as accent is commonly considered an inherent feature of communicative fluency, foreign accented speech might indeed have serious consequences for an L2 speaker in terms of intelligibility and comprehensibility. Besides being able to use the correct words in grammatically accurate sentences and choose a register appropriate to a specific communicative situation, a second language speaker's pronunciation may contribute to whether their speech is comprehensible and intelligible. If a strong foreign accent and severe pronunciation mistakes lead to an impediment of communicative fluency or even to a complete breakdown of communication, it might make speakers intentionally avoid situations in which they have to interact with native speakers. However, as Munro (2008, 197) points out, "a detailed understanding of the situations in which pronunciation errors lead to communicative breakdowns has yet to be developed". In addition, research shows that even a strong foreign accent does not *per se* entail reduced comprehensibility and intelligibility (see e.g., Munro and Derwing 1995a, 1995b) and, consequently, complete failures of communication are rarely observed, as will be further discussed in Chapter 4.

A detectable foreign accent might not only affect – at least to some extent – successful communication, but it may also lead to discrimination based on stereotypical judgments concerning a speaker's intellect and language competence. Lippi-Green (2003), who examined accent-based discrimination in English-speaking environments, points out that "[t]here is a common conception that there is a good

English, and following from that, bad English" (2003, 6). Based on this (mis-)perception, speakers of 'bad' English might be evaluated negatively because their speech does not conform to what is considered the norm of speaking. The idea that there is something like a proper way of speaking is promoted covertly, for instance, through popular media (Munro, Derwing, and Sato 2006) or by educators, politics and the general public (Trofimovich and Isaacs 2012). Negative evaluations and stereotypical judgments based on a speaker's accent can entail severe social and personal consequences for speakers, as Munro (2003) shows. He reviewed several cases of accent-based discrimination in Canada, where second language speakers were terminated or denied employment because of their foreign accent. Despite the fact, however, that accent is frequently associated with race, social status, and educational background, Munro and Derwing (2009, 486) make a valid point by stating that "accents do not cause discrimination – the fault is with intolerant, often monolingual interlocutors" who have biased ideas about foreign accents and hence might simply refuse to understand accented speech even if it is clearly understandable.

While the phenomenon of foreign accent is predominantly dealt with in the context of second language speech, as the previous discussion shows, social and personal consequences of having a nonnative accent in one's first language have received less attention so far, although the consequences associated with reduced pronunciation abilities in the L1 might be similar to what nonnative L2 speakers experience. As discussed by de Leeuw, Schmid, and Mennen (2010), the observation that L2-immersed late bilinguals might be perceived as sounding nonnative in their first language raises questions of identity and belonging in the sense that speakers might no longer be regarded as "members of the linguistically and culturally dominant group of their country of origin" (De Leeuw, Schmid, and Mennen 2010, 40). If speakers are then also not perceived as belonging to the speech community of their L2 because they retain a nonnative L2 accent, the question of identity becomes even more relevant.

In the discussion above, several negative consequences of having a foreign accent in one's first and/or second language have been illustrated. However, speakers with a detectable nonnative accent do not always have to face prejudices. Munro and Derwing (2009), for example, state that particularly European accents might be favored by listeners and speakers alike because they are often associated with social prestige and a high intellectual status. Others argue that speakers may deliberately maintain their foreign accent because it has a symbolic meaning to them or because they want to maintain their ethnic identity in an L2-dominant speech community (see e.g., Moyer 2013, 13). A famous example of a speaker who considers his nonnative L2 accent as a unique characteristic is Arnold Schwarzenegger, the subject of the present investigation, as will be presented in more detail in the subsequent chapter.

2 Study I: Arnold Schwarzenegger's production of plosives in his L1 and L2

2.1. Introduction to Study I

This study investigates the realization of voice onset time (VOT) contrast in the speech of the late consecutive bilingual Arnold Schwarzenegger (AS), who moved to an L2 English-speaking environment at the age of 21. The overall aim of this study was to trace AS's L2 segmental speech development over time and to identify potential changes and modifications in his L1 due to L2 learning experience. The speech corpus for the acoustic analysis of VOT consists of different interviews with AS which were conducted in either English or German (see Section 2.3.2, for a detailed description of the corpus).

For the purpose of the present study, the specific L1 and L2 varieties spoken by AS need to be considered in more detail. AS acquired his L2 primarily through everyday interactions with native American English speakers in his home state California, where he has been living since he moved to the United States in the 1960s. Therefore, it can be assumed that his English pronunciation is predominantly influenced by Californian English,[20] which – on the segmental level – primarily differs from General American English and other American English varieties in terms of the quality and duration of some vowels (Ladefoged 2005). When it comes to consonant contrasts, however, there are hardly any differences between individual American English varieties. Consequently, in the context of the present study, AS's L2 variety can be broadly described as American English, potentially with some Californian English influences on his vowel production, which will be considered in more detail in Study II (see Chapter 3).

AS's L1 is an Austrian German variety spoken in Thal, a municipality which is located four kilometers west of Graz-City in Austria. According to Wiesinger (1967), this local dialect can be classified as west-Styrian. Except for Wiesinger's structural description of Styrian dialect regions and of some phonetic and lexical features of the dialects spoken in these regions, there are no investigations providing an in-depth analysis of phonetic and phonological features of Styrian dialects, specifically of the regional variety spoken in Thal. Most studies examining pronunciation characteristics of Austrian German focus on *Standard* Austrian German[21] (e.g.,

20 It is, of course, likely that AS has been exposed to other English varieties through, for instance, the media and due to travelling within the United States and to other English-speaking countries. It should be noted that 'Californian English' is commonly used as an umbrella term for different regional varieties and sociolects which are spoken in the state of California (see e.g., Eckert and Mendoza-Denton 2006).

21 As pointed out by Kleber (2018, 469), standard pronunciation forms of Austrian German become

Brandstätter, Kaseß, and Moosmüller 2015; Brandstätter and Moosmüller 2015; Harrington, Hoole, and Reubold 2012; Moosmüller 2007; Moosmüller and Ringen 2004), which is strongly oriented towards Standard German German and is predominantly spoken in urban centers (Moosmüller 1991, 2011; Moosmüller, Schmid, and Brandstätter 2015). Despite the fact that AS grew up speaking a regional variety of Austrian German which is likely to deviate from Standard Austrian German norms in terms of pronunciation (and other linguistic) characteristics, it will be referred to as 'Austrian German' in this investigation for the sake of convenience. In order to be able to compare AS's production of L1 plosives to native speaker productions, normative speech data were collected from a 75-year-old monolingual Austrian German control speaker from Thal (see Section 2.3.2.3).

As the focus of this study is on AS's production of VOT in both of his languages, the subsequent sections provide an overview of aspects related to the phonetic feature of VOT in English and (Austrian) German. In addition, previous research on VOT in the context of L2 acquisition and L1 attrition will be considered in more detail.

2.2. Voice Onset Time

Voice Onset Time (VOT) is one of the most frequently investigated phonetic components of stop consonants (e.g., Abramson and Whalen 2017; Baken and Orlikoff 2000; Docherty 1992). According to Lisker and Abramson (1964, 388), this temporal acoustic parameter "serves to separate the stop categories of a number of languages in which both the number and phonetic characteristics of such categories are said to differ". VOT is defined as the time interval between the release of the plosive, which is assigned the value '0', to the onset of voicing (i.e., the onset of vocal fold vibration) of the following vowel (Lisker and Abramson 1964, 422). The time interval between the plosive release and the onset of the following vowel is assigned a positive value (see Docherty 1992). It has been shown that listeners are very sensitive to even small variations in VOT and use it as a cue to determine whether a stop is voiced or voiceless (e.g., Carney 1977; Miller and Volaitis 1989). Phonetically, the phonological distinction between 'voiced' and 'voiceless' plosives can be implemented differently as illustrated in Figure 2.1. Languages such as French, Spanish,

increasingly common among young speakers "as a result of convergence of regional varieties with the standard language". This shift towards using standard pronunciation features is considerably influenced by language contact, for instance via German media (Muhr 2003; Stuart-Smith et al. 2013) and due to an increasing number of German nationals living in Austria (192,426 in January 2019; Statistik Austria 2019). The subject of the present study grew up in Thal in the 1940s, a time when people – particularly those who lived in rural areas of Austria – had less contact with Standard German varieties compared to nowadays. This is why older speakers are more likely to maintain dialectal, non-standard pronunciation features (Kleber 2018).

and Dutch, among many others, have short positive VOT for voiceless plosives; they contrast with voiced plosives which are usually fully voiced, i.e., phonation starts well before the release of the stop and is usually maintained through the closure phase (Lisker and Abramson 1964; Simon 2009; van Alphen and Smits 2004). Consequently, in these languages the phonological distinction between voiced and voiceless plosives is reflected in the presence or absence of prevoicing (negative VOT). In German and English, by contrast, voiceless stops typically have long-lag VOT. These aspirated stops contrast with voiced plosives which do not need to be fully voiced (very short VOT or around 0) (Docherty 1992; Lisker and Abramson 1964). While most of the world's languages have a two-way contrast, such as the languages described above, only few languages, for example Thai, Eastern Armenian and Korean, have a three-way contrast (see e.g., Cho, Jun, and Ladefoged 2002; Gandour et al. 1986; Hacopian 2003), and even less languages, such as Nepali and Hindi, exhibit a four-way contrast (e.g., Cho, Whalen, and Docherty 2019; Davis 1994, 1995; Poon and Mateer 1985). However, as argued by Abramson and Whalen (2017), in languages exhibiting a four-way contrast, VOT alone might not be sufficient to distinguish the different plosive categories.[22] Therefore, additional acoustic parameters, such as fundamental frequency, may need to be included when examining plosive contrasts (see Section 2.2.4).

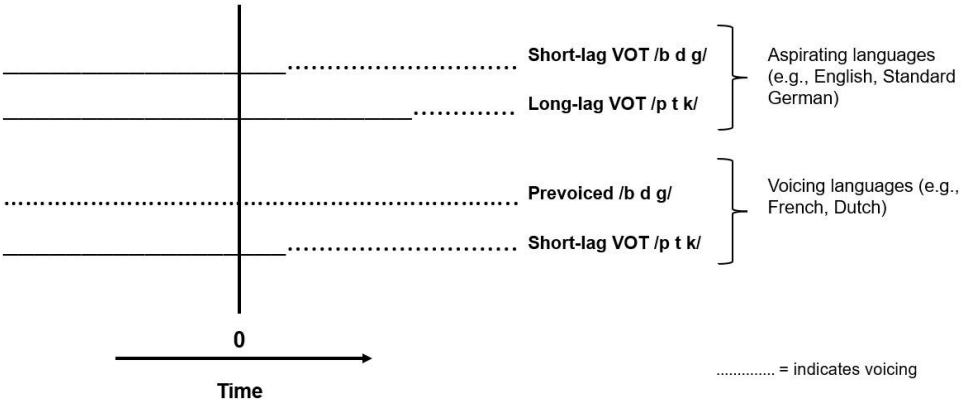


Figure 2.1: Schematic representation of different VOT categories in aspirating and voicing languages (based on Ashby and Maidment 2013, 95; Ladefoged 1975, 124).

22 Hindi, for example, displays four plosive categories, namely [b], [b^{h}], [p], and [p^{h}]. As discussed by Abramson and Whalen (2017), voiced aspirated stops, such as [b^{h}], are not distinguished from voiced stops, such as [b], in terms of VOT.

VOT duration of pre-vocalic plosives in word-initial position[23] is influenced by several linguistic factors, such as place of articulation, quality of the following vowel, stress, speaking rate, and syllable number (e.g., Baken and Orlikoff 2000; Foulkes, Docherty, and Jones 2011; Yavaş and Wildermuth 2006).[24] With regard to place of articulation, it has been shown that the further back in the oral cavity the plosive is produced, the longer the VOT duration (Braun 1996; Byrd 1993; Crystal and House 1988, 1990; Docherty 1992). The effect of place of articulation on VOT can be explained on the basis of general laws of aerodynamics: The primary reason for velar plosives having longer VOT compared to bilabial and alveolar plosives is the "relatively smaller volume of the supralaryngeal cavity in velar stops [which] causes a greater pressure" (Cho and Ladefoged 1999, 213; see also Keating, Westbury, and Stevens 1980). The increased air pressure at the point of release takes longer to drop and, consequently, longer VOT durations can be measured. The relationship between place of articulation and VOT duration is well established, i.e., the majority of studies examining VOT in different languages generally found longer VOT values for velar plosives compared to bilabial and alveolar targets (e.g., Cho and Ladefoged 1999; Morris, McCrea, and Herring 2008).

In terms of vocalic environment, some research suggests that VOT tends to be shorter before mid and low vowels, and longer before high vowels (Braun 1996; Higgins, Netsell, and Schulte 1998; Klatt 1975; Morris, McCrea, and Herring 2008; Zue 1976). Klatt (1975), for example, reported that VOT for English /p t k/ was 15% longer when occurring before high vowels than before low vowels. However, investigations into the interaction between VOT duration and vowel quality have not provided entirely conclusive findings yet. Lisker and Abramson (1967) found no significant correlations between vocalic environment and VOT, but showed that word and syllable stress have an impact on VOT duration. Their findings suggest that the effect of stress on VOT in English is most prominent in voiceless stops, that is, the degree of aspiration is higher if the target plosive occurs in stressed position while it is lower when occurring in unstressed position. In addition, they found that VOT

23 As pointed out by Docherty (1992; see also Foulkes, Docherty, and Jones 2011, 63), the acoustic dimension of VOT is not suitable for distinguishing plosives occurring in word-final position or within a consonant cluster, because plosives in these position "often lack the burst of wide-band noise in their spectra which marks the instant of the release of the stop, and which is used as a reference point for measuring voice onset time" (Docherty 1992, 16). Hence, the majority of acoustic studies examining VOT focus on plosives in word-initial position.

24 Alongside the intra-linguistic factors discussed above, also extra-linguistic factors, such as speaker sex (e.g., Li 2013; Morris, McCrea, and Herring 2008; Oh 2011; Whiteside, Henry, and Dobbin 2004; Yu, Nil, and Pang 2015), have been examined to identify potential effects on VOT. However, findings on the influence of speaker sex on VOT are inconclusive: While some studies revealed differences between male and female speakers in terms of their production of VOT duration (e.g., Oh 2011; Ryalls, Zipprer, and Baldauff 1997; Whiteside, Henry, and Dobbin 2004), others (e.g., Morris, McCrea, and Herring 2008) did not find a correlation between speaker sex and VOT.

is significantly longer in target words produced in isolation compared to the same target words produced in a sentence. The same observation was made by Braun (1996), who found that VOT is overall longer in plosives in individual words than in plosives embedded in sentences. In order to control for the possibility that syllable stress affects VOT duration in the present study, only monosyllabic and bisyllabic tokens containing a plosive in stressed onset position were selected.

Another influencing factor which needs to be considered when examining VOT is speaking rate (e.g., Kessinger and Blumstein 1997, 1998; Lisker and Abramson 1964; Miller, Green, and Reeves 1986; Theodore, Miller, and DeSteno 2009; Volaitis and Miller 1992). In their investigation of VOT in French, English and Thai stops, Kessinger and Blumstein (1997, 1998) found that if speaking rate was slowed, VOT increased in voiceless stops; however, a change in speaking rate did not affect VOT duration in voiced stops. In addition, they reported that if speaking rate increases, "the VOT differences at the voiced-voiceless boundary may be reduced, leading in some cases to overlap between the VOT categories" (Kessinger and Blumstein 1997, 144). Since the present investigation uses pre-recorded and spontaneous speech materials, controlling for speaking rate was not possible. It should be noted, though, that also in speech samples elicited by means of monitored production tasks, such as sentence readings, controlling for speaking rate can be problematic because speakers might differ in what they consider a 'slow' versus a 'fast' speaking rate (see Harrington 2010, 23).

Alongside the factors previously outlined, the number of syllables in a word might influence VOT duration. Klatt (1975) compared VOT in plosives occurring in monosyllabic and bisyllabic items embedded in carrier phrases which were produced in a reading task. He found that VOT was shorter in bisyllabic tokens than in monosyllabic tokens. In a more recent study, Yu, Nil, and Pang (2015) measured VOT in voiceless bilabial plosives in single-syllables (CV) and polysyllables (CVCVCV) produced by monolingual English-speaking children and adults. Similar to Klatt (1975), they found a significant effect of syllable number on VOT duration: All speakers produced longer mean VOT values for tokens occurring in monosyllabic tokens, and significantly shorter VOT in polysyllabic tokens.

While different linguistic factors have been shown to influence VOT duration to a greater or lesser extent, variability in VOT has also been identified as a speaker-specific phenomenon. Chodroff and Wilson (2017), for example, examined VOT duration in word-initial stops in isolated and connected speech produced by monolingual American English speakers. A considerable amount of VOT variability was evident even after controlling for factors such as speaking rate, number of syllables, and vowel context, i.e., the variability in VOT observed in this study could not be exclusively explained by linguistic factors but was shown to be the result of speaker-specific factors. Speaker-specific variation in VOT is associated with several extra-

linguistic factors, such as speaker sex (e.g., Oh 2011; Robb, Gilbert, and Lerman 2005; Swartz 1992; Whiteside and Marshall 2001), dialect (e.g., Scobbie 2005, 2006), and natural ageing processes (e.g., Ryalls, Simon, and Thomason 2004; Ryalls, Zipprer, and Baldauff 1997; Smith, Wasowicz, and Preston 1987; Torre and Barlow 2009).

2.2.1 Voice onset time in English and Austrian German

The phonetic dimension of VOT has been investigated in numerous acoustic studies on plosives in different languages (e.g., Abdelli-Beruh 2004; Chao and Chen 2008; Cho and Ladefoged 1999; Lisker and Abramson 1964; Rosner et al. 2000; Smiljanic and Bradlow 2008). So far, VOT has been most frequently examined in English (see Table 2.1), which, as described above, is an aspirating language and distinguishes between short-lag and long-lag VOT, i.e., /p t k/ are produced with long-lag VOT ranging from approximately 30 to 100ms, and phonologically voiced plosives /b d g/ are produced with short-lag positive VOT in the range of 0 to 25 ms (Docherty 1992; Lisker and Abramson 1964).

Study		**/p/**	**/b/**	**/t/**	**/d/**	**/k/**	**/g/**
Chodroff et al. (2015)	AE	51	9	60	14	54	17
Morris, McCrea, and Herring (2008)	AE	56	15	69	19	73	30
Chao and Chen (2008)	BE	62 *22–117*	11 *7–18*	73 *48–105*	22 *13–68*	86 *65–145*	24 *13–40*
Docherty (1992)	BE	42 *10–80*	15 *0–50*	64 *30–110*	21 *0–50*	62 *30–150*	27 *10–60*
Lisker and Abramson (1964)	AE	58 *20–120*	-101; 1 *-130: -20; 0–5*	70 *30–105*	-102; 5 *-155: -40; 0–25*	80 *50–135*	-88; 21 *-150: -60; 0–35*
Zue (1976)	AE	58	13	71	19	74	30

Table 2.1: Mean VOT durations (measured in milliseconds; ranges displayed in italics if available) for native British English (BE) and American English (AE) reported in previous studies.

As shown in Table 2.1, Lisker and Abramson (1964), who examined VOT in monolingual American English speakers, report two different VOT categories for the production of English phonologically voiced plosives, namely short-lag and negative VOT, the latter indicating prevoicing. However, it must be noted that Lisker and Abramson included a relatively small number of speakers ($N = 4$) in their study. In

addition, 95% of all prevoiced plosives were produced by a single participant.[25] Similarly, Flege and Brown (1982) found that only two of their monolingual American English speakers ($N = 8$) produced /b/ with prevoicing. Both studies show that the subjects were consistent in terms of producing voiced targets with prevoicing, which suggests that "speakers do not arbitrarily choose one of the two forms of timing [short-lag VOT vs. voicing lead], but have a preference for one form or the other" (Docherty 1992, 29). In addition, prevoiced English tokens are most likely to occur in controlled speech production contexts (Roach 2009), particularly in voiced environments (Docherty 1992). Based on the findings of the aforementioned studies and the studies illustrated in Table 2.1, producing voiced stops with voicing lead in American English can be considered a speaker-specific and context-dependent phenomenon, but it cannot be reliably argued that it is a general and widespread characteristic of American English VOT production.[26]

The schematic representation of VOT ranges for English plosives illustrated in Figure 2.2 shows that the VOT values reported in different studies cover quite broad ranges, which becomes particularly evident in the voiceless targets. Most VOT studies report population values without further examining individual variation in VOT production, which might be a reason for the broad ranges presented below. Theodore, Miller, and DeSteno (2009), among others (e.g., Allen, Miller, and DeSteno 2003; Chodroff et al. 2015), set out to examine speaker-specific differences in VOT duration depending on two contextual variables, namely speaking rate and place of articulation. In terms of speaking rate, findings showed that a decrease in rate entails an increase in VOT duration in voiceless plosives; VOT in voiced plosives, by contrast, was not significantly affected by speaking rate (see Kessinger and Blumstein 1997, 1998). In addition, Theodore, Miller, and DeSteno (2009) observed that the effects of speaking rate on VOT differed significantly among individual speakers. By contrast, the effect of place of articulation on VOT was not found to be speaker-specific, i.e., all subjects consistently produced the voiceless bilabial plosive with shorter VOT compared to the voiceless velar target, which is in line with previous findings (e.g., Lisker and Abramson 1964). Similarly, Docherty (1992) observed a considerable amount of inter-speaker variability in his investigation of VOT. He points out that "it does seem to be the case that the between-speaker

25 Another aspect to consider is that Lisker and Abramson (1964) do not provide further details about the participants' language backgrounds, apart from stating that all speakers were educated speakers of a standard variety of the respective language (see 1964, 388). In addition, it is not described how the speech samples were elicited, which might indeed have had an influence on VOT durations.

26 One exception is a study by Hunnicutt and Morris (2016) who examined mean VOT duration in word-initial stops produced by 13 speakers of Southern American English from Alabama and Mississippi. Their findings show that the majority of all voiced tokens (77.8%) were produced with voicing-lead, which was interpreted to represent a dialect-specific feature of Southern American English as spoken in Alabama and Mississippi.

difference in voice onset time is a reflection of different implementation strategies, as opposed to being a by-product of some other feature of the speaker concerned" (169).

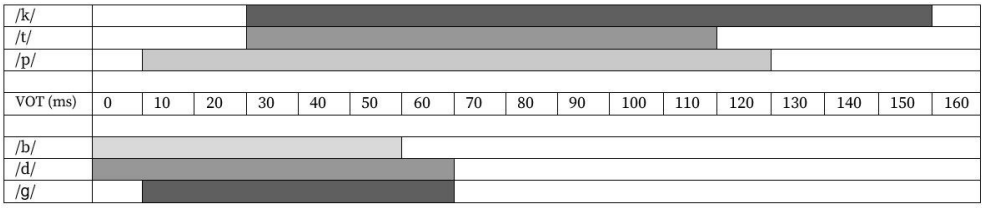


Figure 2.2: Schematic representation of VOT ranges for English /p t k/ and /b d g/ reported in the literature (Chao and Chen 2008; Docherty 1992; Lisker and Abramson 1964).

Similar to English, Standard German German (SGG) also distinguishes between long-lag (aspirated) and short-lag plosives (e.g., Braunschweiler 1997; Jessen 1998; Jessen and Ringen 2002; Kuzla and Ernestus 2011; Stock 1971).[27] While the short-lag and long-lag VOT categories are relatively stable in English and SGG, speakers of Austrian German varieties have been observed to neutralize[28] VOT contrast in bilabial and alveolar word-initial plosives by producing both voiced and voiceless targets within the short-lag VOT range (Moosmüller 2011; Moosmüller and Ringen 2004; Moosmüller, Schmid, and Brandstätter 2015). According to Wiesinger (1996, 2014), the voiced-voiceless distinction is eliminated if the stop is followed by a sonorant consonant, as in *Draht* vs. *trat* (*wire* vs. *kicked*), or *Blatt* vs. *platt* (*leaf/sheet* vs. *flat*) (see also Muhr 2001; Zeman 2009). Similarly, this neutralization of contrast has been observed to occur in bilabial or alveolar plosives followed by a vowel (e.g., *du* vs. *tu* / *you* vs. *do*; *Tank* vs. *Dank* / *tank* vs. *thank*). However, VOT contrast is usually maintained for velar plosives preceding a vowel (e.g., *Karten* vs. *Garten* / *cards* vs. *garden*). While many descriptions of Austrian German pronunciation characteristics rely on auditory-based observations and impressionistic data only (e.g., Muhr 2001; Wiesinger 1996, 2014), Moosmüller and Ringen (2004) and Moosmüller (2011) took VOT as a parameter to examine plosive production in speakers from Vienna (see Table 2.2). In a sentence reading task, they elicited target words with plosives occurring in word-initial position. The findings by Moosmüller (2011) indicated that the speakers produced a significant VOT difference between voiced and voiceless

27 Jessen (1998, 306ff.), for example, reports the following mean VOT durations for Standard German plosives: 75 ms for /p/, 7 ms for /b/, 77 ms for /t/, 18 ms for /d/, 87 ms for /k/, and 29 ms for /g/.

28 In the context of the present investigation, 'neutralization' is defined as "the elimination of phonological contrast in certain phonetic environments" (Dmitrieva, Jongman, and Sereno 2010, 483).

stop consonants: VOT for voiceless plosives ranged from 40 ms to 60 ms, and voiced plosives were produced with VOT values between 5 ms and 20 ms. Moosmüller and Ringen (2004) found a considerable amount of VOT variability in their speakers' productions of voiceless alveolar and bilabial plosives as indicated by the broad VOT ranges reported in Table 2.2. By contrast, variability for the voiced targets was relatively balanced across all places of articulation. The high variability observed in voiceless plosives with a bilabial and alveolar place of articulation, according to Moosmüller and Ringen (2004, 49), shows a tendency to neutralize VOT contrast in bilabial and alveolar plosives, that is, VOT contrast is not realized consistently.
As discussed by Moosmüller, Schmid, and Brandstätter (2015), bilabial and alveolar contrasts are most likely to be eliminated in spontaneous conversational speech styles, i.e., "bilabial and alveolar fortis [voiceless] and lenis [voiced] plosives might collapse, especially in word-initial position, so that packen 'to pack' and backen 'to bake' become homophonous: [b̥ɑkŋ]" (341). This neutralization of contrast has already been described by Luick (1904, 83), who speaks of 'blurring' the distinction ('Verwischung der Grenzen') between voiced and voiceless plosives in Austrian German. In a more recent investigation, Hödl (2019; see also 2017) examined the realization of VOT contrast in conversational and read speech produced by a group of Styrian speakers (N = 33). As displayed in Table 2.2, VOT data obtained from the spontaneously produced speech samples differed considerably compared to those obtained from read samples: While a tendency to neutralize VOT contrast in bilabial and alveolar plosives was observed in the speakers' spontaneous speech, there was no overlap between long-lag and short-lag VOT categories in the read speech, i.e., speakers realized a pronounced VOT-based distinction between voiced and voiceless plosive targets. By contrast, the results provided by Grassegger (19881988; see also 1996) do not indicate a difference in VOT measured in read speech and in speech produced on the basis of a picture naming task. As pointed out by Hödl (2019, 17), the picture naming task included in the study was used to elicit non-scripted but carefully controlled speech and, thus, the VOT measurements obtained from these samples cannot be considered representative of spontaneously produced, more naturally occurring speech.

Study		/p/	/b/	/t/	/d/	/k/	/g/
Grassegger (1988)	Read speech	29	17	39	21	87	28
	Picture naming	27	18	37	22	85	30
Moosmüller and Ringen (2004)	Read speech: Mid vowel context	29 *6–60*	11 *5–16*	33 *14–49*	14 *8–13*	64 *30–73*	25 *14–37*
	Read speech: High vowel context	52 *17–108*	13 *5–18*	61 *19–110*	26 *14–41*	99 *80–129*	31 *12–58*
Moosmüller (2011)	Read speech	37	14	46	18	80	27
Hödl (2019)	Read speech	46	10	59	14	78	23
	Spontaneous speech	22 *1–178*	11 *1–17*	29 *3–143*	19 *3–88*	63 *12–181*	24 *4–70*

Table 2.2: Mean VOT durations (in milliseconds; ranges displayed in italics if available) for Standard Austrian German speakers from Vienna (Moosmüller 2011; Moosmüller and Ringen 2004) and for Styrian speakers (Grassegger 1988; Hödl 2019)

As the previous discussion reveals, (American) English and Austrian German differ when it comes to the implementation of a VOT contrast: While English displays a relatively stable contrast between short-lag (sometimes prevoiced) versus long-lag VOT categories, this contrast tends to be neutralized in Austrian German bilabial and alveolar tokens produced in spontaneous speech contexts. These cross-linguistic differences in VOT between English and Austrian German provide the basis for the present study. In fact, examining VOT duration is quite popular in SLA and bilingualism research because data are usually easy to obtain and VOT measurements can be done rapidly and rather accurately (see Flege 2017, 2018). The subsequent sections provide an overview of research into L2 acquisition and L1 attrition of VOT.

2.2.2 L2 acquisition of VOT

Previous studies examining VOT production in different L2s have mainly focused on the question of whether bilinguals and second language learners acquire separate VOT categories for their two languages and to what extent bilingual VOT

patterns resemble those of monolingual speakers. While there are only few studies investigating VOT in the context of L1 attrition, as will be discussed in the subsequent section, there is a substantial amount of research examining L2 acquisition of VOT.

Table 2.3 displays a selection of studies which examined the acquisition of VOT in late learners of L2 English with different first languages, all of which implement VOT contrast differently compared to English (except for Stoehr et al. 2017 who investigated L2 learners of German, and the subjects in Flege 1987b and Flege and Hillenbrand 1984 who spoke English as their L1). As outlined in the previous section, English is an aspirating language, i.e., voiceless plosives are produced with long-lag VOT values. They contrast with voiced plosives which are usually not fully voiced but are produced within the short-lag VOT range. In voicing languages, such as French, Dutch and Spanish, voiceless plosives are produced with short-lag VOT and their voiced counterparts are produced with voicing lead, i.e., they are fully voiced (Lisker and Abramson 1964). Consequently, native speakers of voicing languages who learn English as an L2 have to acquire aspiration for English voiceless plosives and the short-lag VOT category for English voiced stops. Based on these cross-linguistic differences between English and different voicing languages, the studies displayed in Table 2.3 aimed at examining to what extent L2 learners of English are able to acquire distinct VOT categories for their second language. The focus here is on late L2 learners, i.e., individuals who started acquiring English late in life after the first language had already been fully developed, which reflects the order of language acquisition experienced by the subject investigated in the present study. Other VOT studies, not featured in Table 2.3, include early sequential bilinguals who acquired English as an L2 in early childhood (e.g., Caramazza et al. 1973; Flege and Eefting 1987b; Mack 1990; MacLeod, Stoehl-Gammon, and Wassnik 2009; McCarthy et al. 2014), or simultaneous bilinguals who acquired both languages from birth (e.g., Fabiano-Smith and Bunta 2012; Kehoe, Lleó, and Rakow 2004; Lein, Kupisch, and van de Weijer 2016).

Study	Languages	N of subjects	Plosives tested	Elicitation method
Flege and Hillenbrand (1984)	L1: English L2: French	14	/t/	Reading: target words in carrier phrase + story telling
Flege (1987b)	L1: English L2: French	14	/t/	Reading: target words in carrier phrase
Flege and Eefting (1987a)	L1: Dutch L2: English	50	/t d/	Reading: target words in carrier phrase
Flege (1991)	L1: Spanish L2: English	10	/t/	Reading: target words in carrier phrase
Flege, Munro, and MacKay (1996)	L1: Italian L2: English	126	/p t/	Reading: target words in carrier phrase
Thornburgh and Ryalls (1998)	L1: Spanish L2: English	16	/p t k b d g/	Reading: isolated words
MacKay et al. (2001)	L1: Italian L2: English	72	/b/	Delayed repetition task
Simon (2009)	L1: Dutch L2: English	16	/p t k b d g/	Reading: isolated words + spontaneous conversation
Stoehr et al. (2017)	L1: Dutch L2: German	18	/p t k b d g/[29]	Picture-naming

Table 2.3: Overview of studies on L2 acquisition of VOT in late bilinguals (in chronological order).[30]

A review of the studies presented in Table 2.3 reveals that bilingual speakers, whose languages differ in the way VOT contrast is implemented, often fail to produce L2 targets within native ranges. Flege and Hillenbrand (1984) and Flege (1987a), for instance, found that their English-French late consecutive bilinguals produced L2 French /t/ with considerably longer (i.e., more English-like) VOT values compared to monolingual French controls, that is, they realized values which were too short for English and too long for French. This observation can be explained with reference to the Speech Learning Model which will be discussed in Section 1.6.1: Due to the mechanism of *equivalence classification*[31] (e.g., Flege 1987a, 1995), the bilingual

29 Since word-initial /g/ is not part of the Dutch consonant inventory, it was not included in this study for either language.

30 As mentioned in the introduction to this section, research focusing on the acquisition of VOT in bilingual speakers and second language learners is quite extensive. The overview presented here is not intended to give a review of all studies of VOT in bilingual speakers, but rather includes a representative selection of studies as a basis for discussing VOT acquisition patterns which can be observed in bilingual speakers.

31 Flege and Hillenbrand (1984) refer to the phenomenon of equivalence classification as "interlingual

speakers in Flege and Hillenbrand (1984) and Flege (1987) failed to identify phonetic differences between English /t/ and French /t/ and, thus, produced the L2 target sound with VOT values intermediate between English and French monolingual norms.

The same observation was made by Flege and Eefting (1987a), who examined L1 speakers of Dutch learning L2 English: Despite the fact that the L1 Dutch speakers produced English /t/ with longer VOT values compared to their production of Dutch /t/, they did not realize the English target plosive within the monolingual English long-lag range. That is, they had not acquired the long-lag VOT category for English aspirated voiceless plosives and produced VOT values intermediate to L1 Dutch and L2 English. These findings are also reflected in Flege (1991), who observed 'compromise' – nonnative-like – VOT values in the productions of L2 English /t/ by Spanish-English late bilinguals. By contrast, the *early* L2 learner group (i.e., early sequential bilinguals) in this study did not differ from monolingual controls in their production of the target plosive, that is, they were able to establish different phonetic categories for the plosives in their two languages.

While the studies discussed above focus exclusively on the production of the voiceless alveolar plosive /t/, other studies examine the realization of VOT *contrast* in voiced and voiceless targets. Thornburgh and Ryalls (1998), for example, found that the majority of their native Spanish learners of L2 English were able to produce English /p t k/ within the long-lag monolingual range, but their productions of English voiced stops were characterized by consistent prevoicing, as typical of their L1 Spanish. This transfer of prevoicing from the L1 to the L2 was also observed in studies examining native Dutch learners of English (e.g., Mayr, Price, and Mennen 2012; Simon 2009; Stoehr et al. 2017).

Overall, the findings of the studies discussed above support Flege and Hillenbrand's (1984, 717) argument that "adult learners of a foreign language will never succeed in producing L2 stops with complete accuracy when stops in their native language differ substantially in VOT from those in L2".[32] This assumption is in line

identification" (717). However, with the development of the Speech Learning Model (e.g., Flege 1995), the term 'equivalence classification' became more commonly used.

32 Despite the fact that the majority of individuals who learn an L2 in adulthood have been shown to produce L2 features in a nonnative fashion, Flege and Hillenbrand's (1984, 717) claim that late L2 learners "will *never* succeed" (my italics) when it comes to attaining native pronunciation abilities in the L2 might be too strong. In fact, research has documented the occurrence of exceptionally successful late L2 learners (e.g., Bongaerts 1999; Bongaerts, Mennen, and van der Slik 2000; Moyer 1999). For example, some of the late learners of L2 Dutch in Bongaerts, Mennen, and van der Slik (2000) were perceived to sound (near-)native in their L2 by a group of monolingual Dutch listeners, which led to the conclusion that "it is not impossible for post-critical period [i.e., late] learners to achieve a nativelike accent in a non-primary language, in spite of the alleged biological barriers" (Bongaerts, Mennen, and van der Slik 2000, 305). In the studies cited above, (non-)nativeness of L2 learners was examined on the basis of foreign accent ratings, i.e., the focus was on perceived global

with the predictions made by the Speech Learning Model (Flege 1991, 1995, 2002, 2003): The majority of late L2 learners fail to identify phonetic differences between closely related L1 and L2 sounds and perceive them as being representations of the same phonetic category, which is likely to result in nonnative L2 productions. For L2 English learners who speak a voicing language (such as Dutch or Spanish) as their L1 this might lead to the production of compromise values which are intermediate to L1 and L2 norms (e.g., Flege and Hillenbrand 1984), or to transferring an L1 feature – such as prevoicing – to the production of the corresponding sounds in the L2 (e.g., Mayr, Price, and Mennen 2012; Thornburgh and Ryalls 1998). A second possibility alongside the merging (i.e., assimilation) of L1 and L2 phonetic categories is category *dissimilation,* or *polarization* (Keating 1984), which serves to enhance acoustic differences between L1 and L2 categories (see Flege and Eefting 1987a, 198). Flege and Eefting (1988) examined VOT in Spanish-English early and late bilinguals' productions of Spanish /p t k/. They observed that the early bilinguals produced the L1 Spanish target plosives with considerably *shorter* VOT values compared to monolingual control speakers. That is, the early bilingual subjects – unlike the late bilinguals – had established a new phonetic category for the L2 sounds, but 'overshot' the L1 target in order to maintain contrast between the L1 and L2. This dissimilatory effect is usually observable in early bilinguals, who are "more likely to develop new L2 categories and, so, are more likely to exhibit dissimilation" (Oh et al. 2011, 157). By contrast, late bilinguals are less likely to dissimilate phonetic categories, but – as outlined above – often fail to establish separate categories for closely related L1 and L2 sounds and produce *merged* categories instead.

2.2.3 L1 attrition of VOT

While there is extensive research on the acquisition of L2 VOT by bilingual speakers, as discussed above, there are only few studies which examine VOT in the context of L1 phonetic attrition.

Table 2.4 presents an overview of attrition studies examining VOT in different languages. In line with the definition of attrition discussed in Section 1.4.3, the studies included in this overview focus on *late consecutive bilinguals who have been long-term immersed in an L2-speaking environment*. This reflects the order of acquisition and the immersion context experienced by AS, the subject examined in the present study. Hence, studies including simultaneous bilinguals (as in Fowler et al. 2008) or inexperienced L2 learners receiving intensive formal instruction in the L2 (as in Chang 2012, 2013; Kartushina, Frauenfelder, and Golestani 2016) were excluded from this overview despite the fact that they are frequently cited in the context of

accent. Acoustic investigations of specific segmental and prosodic features to identify potential divergences from native speaker norms were not included.

L1 phonetic attrition. Similarly, the case study conducted by Sancier and Fowler (1997; see Tobin, Nam, and Fowler 2017, for a replication of this study) is not represented here. They examined VOT in an L1 Brazilian-Portuguese speaker with L2 English, who regularly travelled between the L1 and the L2 environment. The aim of this study was to find out whether the regular change in linguistic environment influenced the bilingual's productions of VOT in voiceless stops. As the subject in Sancier and Fowler' study regularly travelled between two linguistic environments and did not consistently stay in an L2-immersion context for an extended period of time, this study is not included in the overview.

Study	**Languages**	**N of subjects**	**Plosives tested**	**Elicitation method**
Flege and Hillenbrand (1984)	L1: English L2: French	7[33]	/t/	Reading: target words in carrier phrase + story telling
Flege (1987b)	L1: English/French L2: French/English	14[34]	/t/	Reading: target words in carrier phrase
Major (1992)	L1: English L2: Portuguese	5	/p t k/	Formal: word/sentence reading; Informal: conversation
Mayr, Price, and Mennen (2012)	L1: Dutch L2: English	2	/p t k b d ɡ/	Reading: target words in carrier phrase
Stoehr et al. (2017)	L1: German L2: Dutch	23[35]	/p t k b d/	Picture-naming

Table 2.4: Overview of studies on L1 attrition of VOT in late bilinguals (in chronological order).

Overall, the findings of the studies presented in Table 2.4 show that VOT in different L1s might be affected by processes of attrition. However, L1 voiced and voiceless stops have not been found to be influenced by L2 learning experience to the same extent; rather, voiceless stops seem to be especially sensitive to attrition while voiced stops remain largely unaffected (Mayr, Price, and Mennen 2012; Stoehr et

33 Flege and Hillenbrand (1984) examined VOT in three groups of speakers, but only one of these groups can be described as a potential attriter group, namely L1 French learners of L2 English (late consecutive bilinguals) living in an L2-speaking environment.

34 Flege (1987b) included six groups of speakers, only two of which be considered as potential attriters, namely L1 English late learners of L2 French living in France, and L1 French late learners of L2 English living in the US.

35 An additional group consisted of L1 Dutch L2 German bilinguals living in the Netherlands. As these speakers did not live an L2-immersion setting, they are not included in this overview.

al. 2017). L1 speakers of languages with voiceless aspirated plosives such as German and English have been observed to shorten VOT in voiceless targets in their L1 when speaking French, Dutch or Portuguese as their L2, and thus assimilate VOT in their L1 to the norms of the L2 (Flege 1987b; Major 1992; Stoehr et al. 2017). Flege (1987b), for example, found that the L1 English L2 French bilinguals in his study produced L1 English /t/ with considerably shorter and thus more French-like VOT compared to monolingual English controls. Although Flege (1987b) does not describe this phenomenon as 'attrition', the changes in L1 VOT categories for voiceless targets clearly show an effect of L2 learning experience on production abilities in the L1 and support Flege's (1987b, 47) hypothesis that late L2 acquisition leads to a – at least partial – restructuring of learners' L1 and L2 phonetic space.

Similar observations were made by Stoehr et al. (2017), who examined VOT in L1 German L2 Dutch-immersed bilinguals. VOT durations measured in the speakers' L1 voiceless plosives were significantly shorter, and thus more Dutch-like, compared to monolingual German speakers. The speakers' L1 voiced plosives, by contrast, were not observed to be affected by attrition as they were not consistently prevoiced as typical of L2 Dutch. 32% of the bilinguals' L1 productions of /b/ and /d/ were, in fact, characterized by prevoicing and a significantly higher percentage of prevoiced tokens (65%) was observed in the bilinguals' productions of voiced plosives in Dutch. However, no significant differences in terms of the percentage of prevoicing could be observed when comparing the bilingual group and the monolingual German group, i.e., also the monolingual speakers produced prevoiced tokens of /b/ and /d/. Therefore, the occurrence of prevoicing in the L1 German of the bilingual group was not interpreted as evidence for an L2-influence on the L1. Stoehr et al. (2017, 505) argue that despite the observation that the subjects were capable of producing prevoiced tokens in their L1, "[t]hey may [...] not necessarily be able to control the required muscular activities to a similar extent as native speakers of a prevoicing language". Prevoicing has also been observed to occur in English plosives produced by monolingual speakers (e.g., Docherty 1992; Lisker and Abramson 1964), but, as discussed in Section 2.2.1, it cannot be considered a systematic and widespread feature of English VOT production in the same vein as prevoicing in German does not seem to occur systematically.

Major (1992) also examined a group of bilinguals speaking an aspirating language (American English) as their L1 and a voicing language (Brazilian Portuguese) as their L2. He tested the subjects' productions of /p t k/ elicited in two different contexts (formal vs. casual/informal). Overall, Major's (1992) results provide evidence for deviations from the native English norm in the L1 of all subjects, i.e., they produced L1 voiceless plosives with shorter and thus more Portuguese-like VOT compared to monolingual controls. However, the subjects differed in terms of the degree of deviation from native norms. Two out of five speakers showed fewer

significant deviations from native English norms compared to the other subjects. Three speakers exhibited shorter VOT values in their L1 voiceless plosives produced in an informal context (conversational speech) when being compared to monolingual English speakers. This suggests a drift of the L1 (long-lag) VOT category towards the L2 (short-lag) category. One of the three speakers was not only observed to differ significantly from monolingual production norms, but her VOT values for English voiceless plosives were also significantly shorter than the values measured for the other subjects. In addition, this speaker scored within the native range in terms of her L2 productions of voiceless plosives, i.e., she produced the L2 targets consistently with native short-lag VOT. By contrast, the remaining four speakers were observed to produce L2 plosives within the long-lag VOT range as typical of their L1 English. While this study, similar to the studies discussed above, provides evidence for voiceless plosives being affected by attrition processes, it did not test whether or not similar effects could be observed in terms of the speakers' L1 voiced plosives. At the same time, Major's findings suggest that not all individuals are equally affected by attrition, that is, some bilinguals are able to maintain native production abilities in their L1 while others show clear divergences from native production norms in their first language.

While the studies previously discussed examine VOT in L1 speakers of aspirating languages, other studies focus on investigating VOT in L1 speakers of voicing languages who speak an aspirating language as their L2 (Flege and Hillenbrand 1984; Mayr, Price, and Mennen 2012). Overall, the subjects examined in these studies were found to produce voiceless stops in their L1 with longer VOT values compared to monolingual speakers, i.e., their L1 short-lag VOT categories moved in the direction of the long-lag VOT categories of their L2. Mayr, Price, and Mennen (2012), for example, observed that their bilingual speaker produced L1 Dutch voiceless plosives with considerably longer VOT values compared to her non-immersed bilingual twin sister, which suggests a drift of L1 categories towards the norms of the L2 English. By contrast, her L1 voiced plosives were consistently prevoiced as typical of native Dutch, i.e., no attrition effects could be observed. Interestingly, however, the subject's English voiced plosives were characterized by voicing lead (prevoicing), which was even longer in her L2 compared to her L1. One possible explanation for transferring the L1 feature of prevoicing to the L2 and overshooting the L1 target is, according to Mayr, Price, and Mennen (2012, 692), that the subject might have "acquired separate representations for voiced categories in the two languages and is attempting to differentiate them thus, exhibiting a polarization effect".

In an earlier study, Flege and Hillenbrand (1984) examined VOT duration in the alveolar plosive /t/ in monosyllables produced by French-English bilinguals. The bilinguals in this study were observed to produce L1 French /t/ with considerably longer VOT compared to monolingual French speakers and, thus, were observed to

have drifted away from native production norms. These findings support Flege and Hillenbrand's (1984, 716) hypothesis that "the perceptual target for an L1 phone – and eventually the motor plan used to realize it – changes as the foreign language learner is exposed to that phone's acoustically different counterpart in L2".

Taken together, research seems to provide evidence for a shift of L1 VOT categories in voiceless plosives towards the norms of the L2 in late adult bilinguals, which results in either a shortening of VOT towards the short-lag category of the L2 (e.g., French, Dutch) or a lengthening of VOT duration towards the long-lag category for L2 voiceless plosives (e.g. ,English, Standard German). In addition, the findings provided by Mayr, Price, and Mennen (2012) and Stoehr et al. (2017) indicate that L1 voiced plosives are less likely to be affected by L2-influences.

2.2.4 Other acoustic parameters of plosives

As illustrated in the previous sections, VOT is one of the most frequently investigated acoustic components of word-initial stops and it is often described as the most reliable cue to measure voicing contrasts in plosives (e.g., Foulkes, Docherty, and Jones 2011; Lisker 1975, 1978). However, when examining plosive voicing contrasts other acoustic cues might also play a role, which will be briefly outlined in this section.

Alongside VOT, fundamental frequency (f0) of the vowel following the plosive has been identified as an intrinsic cue to distinguish voicing in plosives. It has been observed that vowels tend to have higher f0 after phonologically voiceless plosives, and lower f0 after phonologically voiced plosives (e.g., Coetzee et al. 2018; Haggard, Ambler, and Callow 1970; Hombert, Ohala, and Ewan 1979; Shultz, Francis, and Llanos 2012). This difference in f0 depending on whether the preceding plosive is voiced or voiceless can approach up to 20 Hertz (Hz) (House and Fairbanks 1953; Lehiste and Peterson 1961). The influence of the plosive on f0 has been shown to be largest at the onset of the vowel (Coetzee et al. 2018), but f0 has also been measured at its peak (e.g., Lehiste and Peterson 1961; Löfqvist 1975), and across the entire vowel (e.g., House and Fairbanks 1953). In the present investigation, intrinsic f0 was not included as an acoustic cue, which is mainly due to the fact that the validity of f0 as a voicing cue in authentic non-prompted speech has not been reliably established so far (see Abramson and Whalen 2017, 79). In addition, Chang (2012, 253) argues that "[t]he f0 difference between English stop types has consistently been shown to be relatively subtle". This is also reflected in Coetzee et al.'s (2018, 187) overview of studies examining post-plosive f0. The f0 difference between preceding voiced and voiceless plosives for English ranges from 6 to 16 Hz, which confirms Chang's (2012) observation that English shows a relatively small f0 difference between voiced and voiceless plosives. Shultz, Francis, and Llanos (2012) arrive at a

similar conclusion: They examined the relationship between VOT and onset f0 in American English in both production and perception. In terms of production, their findings show that the majority of speakers used distinct VOT categories to produce a voicing contrast while the difference in onset f0 was not shown to be significant. Also in terms of perception, the study revealed that onset f0 was not used as a primary perceptual cue to the voiced-voiceless distinction in American English.

Another acoustic phenomenon which has been identified as a cue to the voiced-voiceless distinction in plosives is the intensity of the plosive burst. Alveolar stops have been shown to have the highest energy level (> 4000 Hz), i.e., the most intense burst, while bilabials usually have a very low frequency burst. Also, voiced plosives generally show a lower burst intensity compared to their voiceless counterparts (Baken and Orlikoff 2000; Dorman, Studdert-Kennedy, and Raphael 1977; Zue 1976). The frequency of the burst is commonly included as a phonetic component when VOT alone is not sufficient to examine voicing in plosives. According to Ladefoged and Maddieson (1997), this is, for example, the case in languages which exhibit a four-way plosive contrast, such as Hindi and Marathi, or the Indo-Aryan language Gujarati (see Rami et al. 1999). For languages with a two-way contrast, such as English and German, VOT alone serves as a reliable cue to examine voicing in plosives, at least in word-initial position (see Kleber 2018, 483). Therefore, burst frequency was not examined in the present study.

The duration of the closure interval of the plosive is the third parameter which serves as a cue to the voicing contrast in plosives. While there is extensive research on VOT in English and other languages (see Section 2.2.1), closure duration is a feature which has been less frequently investigated. So far, it has been shown that /p/ has a longer closure duration compared to /t/ and /k/ (Byrd 1993; Zue 1976). Yao (2007), for example, measured a closure duration of 69.5 ms for /p/, 48.9 ms for /t/, and 54.9 ms for /k/ in spontaneous American English speech; these findings were in line with previous results (e.g., Byrd 1993), showing that place of articulation has an effect on the duration of the closure interval. Alongside other voicing parameters, Moosmüller and Ringen (2004) examined closure duration in pre-vocalic word-initial and intervocalic word-medial plosives in Standard Austrian German speakers from Vienna. They found a significant difference in closure duration between voiced and voiceless plosives in word-medial position, i.e., closure duration was longer for voiceless plosives compared to their voiced counterparts. By contrast, their findings did not show a significant difference in closure duration in word-initial plosives, which is in line with previous studies on different varieties of German (e.g., Jessen 1998). While Moosmüller and Ringen (2004, 54) conclude that "closure duration is the only parameter which still contrasts fortis and lenis plosives" in word-medial position in Standard Austrian German, this parameter does not seem to play a significant role when it comes to distinguishing the voiced-voiceless

opposition in word-initial Austrian German plosives. As the present investigation focuses on Austrian German and English plosives occurring in word-initial position, closure duration was also not examined.

As can be concluded from the previous discussion, other acoustic cues besides VOT might be of interest when it comes to examining plosive voicing contrasts in different languages. However, VOT has been identified as the most stable voicing parameter and can be described "as a variable that, in a sense, summarizes a very complex and extremely important aspect of articulator-laryngeal coordination" (Baken and Orlikoff 2000, 274). This is why the focus of the present investigation is on examining characteristics of AS's VOT production in his two languages.

2.3. The study

2.3.1. Aims and objectives

As outlined in the previous sections, VOT has been examined in the context of both acquisition of L2 pronunciation and L1 phonetic attrition. In terms of L2 acquisition it has been shown that the majority of late L2 learners fail to authentically produce L2 VOT if their L1 differs from the L2 in terms of VOT categories (e.g., short-lag vs. long-lag / voicing lead vs. short-lag). Instead of producing L2 plosives with native VOT values, it has been observed that late L2 learners are likely to produce the target sounds with compromise VOT values which are intermediate between the L1 and L2 norms, or to transfer a specific L1 feature (such as prevoicing) to the L2.

The few studies which have been conducted to examine VOT in the context of L1 attrition show that VOT is a phonetic feature which might be affected by processes of attrition as a result of L2 learning experience and long-term immersion in an L2 environment. Attrition effects were mainly observed in terms of L1 voiceless plosives, which have been shown to have shifted closer to L2 production norms in late consecutive bilinguals. However, as previously mentioned, there is still a lack of research in the field of L1 phonetic attrition, especially in terms of longitudinal studies tracing speech development in late consecutive bilinguals over an extended period of time. The present study aims to fill this gap by examining L1 and L2 development of pronunciation in a late consecutive bilingual over a period of approximately 40 years.[36] In order to identify instances of L1 attrition as well as to trace AS's L2 speech development, this study set out to examine his realization of VOT

36 Due to the limited availability of broadcast recordings, it was not possible to trace AS's speech development right from the beginning of his L2 learning process, i.e., from the time he moved to the United States in 1968. The earliest speech samples available were recorded in 1978, ten years after he had moved to the L2-speaking environment.

contrast in word-initial plosives in Austrian German and American English. With regard to this, the following research questions were addressed:

RQ 1.1: To what extent has AS's L2 realization of VOT contrast changed when comparing three stages in time (early, mid, and late)? Does he produce a significant difference between phonologically voiced and voiceless plosives in terms of VOT for each of the stages, i.e., has he gradually become more native-like since he moved to the United States?

RQ 1.2: To what extent has AS's L1 realization of VOT contrast changed when comparing two stages in time (early and late)? Does he produce a significant difference between phonologically voiced and voiceless plosives in terms of VOT for each of the stages, i.e., have his L1 VOT categories moved away from native production norms and at the same time moved closer to L2 production norms since he moved to the United States?

2.3.2. Methodology

2.3.2.1 The subject: Arnold Schwarzenegger

The subject of the present case study is Arnold Schwarzenegger, a – at the time of investigation – 72-year-old native Austrian German speaker who moved to the United States at the age of 21 in 1968, so he had been living in an L2-speaking country for approximately 50 years at the time the study was conducted. AS can be described as a late consecutive bilingual, who acquired the L2 in adulthood when his first language had already been fully developed (Wei 2007).

Born in 1947, shortly after the end of World War II, AS grew up under rather poor circumstances in Thal, a small village near Graz in Austria.[37] His father Gustav Schwarzenegger was a local police officer, and his mother Irene took care of Arnold and his older brother Meinhard. At the age of 15, AS discovered his passion and talent for weight training, which laid the foundation for his future career as a bodybuilder. In 1966, after completing military service in the Austrian Army, AS decided to leave his home country to pursue his "mission [...] to become the world champion in bodybuilding" (Schwarzenegger 2012, 37). Before moving to the United States at the age of 21, AS spent some time in Munich and London where he met influential people from the bodybuilding business. These acquaintances helped him to finally move to the US, where he started a career as a professional bodybuilder and won the highest and most recognized awards in the business within a short period of time. After retiring from active participation in national and international

37 This paragraph is based on Schwarzenegger (2012) if not quoted otherwise.

bodybuilding competitions, AS advanced his career as a Hollywood actor, businessman, and author, and started engaging in politics. In 2003, he was elected 38th Governor of California, an office he held until 2011. Due to his tremendously successful career and his many talents, the USC Schwarzenegger Institute (2013) described AS as being "one of the most recognized individuals on the planet".

Prior to migrating to the US, AS had been learning English as a foreign language in an Austrian school for approximately three years (Absolute Radio 2014), so it can be assumed that he had at least some basic knowledge of English, but he had not acquired English as a "functional second language" (Flege and Hillenbrand 1984, 710) before moving to an L2-speaking country. Alongside acquiring the L2 through everyday communication with native English speakers, AS enrolled in English as a Second Language classes at Santa Monica College in California shortly after his arrival in the US, where had been receiving intensive instruction in English for several months. However, as mentioned in Chapter 1, AS struggled with acquiring English pronunciation and admitted that "[b]ecoming fluent in English was still the hardest thing on my to-do-list [...] Pronunciations were especially dangerous" (Schwarzenegger 2012, 135).

One question which needs to be addressed in the context of the present study is whether and to what extent AS uses his first language and how frequent he has contact with native Austrian German speakers. It is well-known that AS regularly visits Austria on different occasions (e.g., Die Presse 2011, 2018; Jungwirth 2018; Vienna.at 2019) and that he is also a welcome guest in Germany (e.g., Clamann 2017; Gerstenberger 2019; Köndgen 2019). As he is a native German speaker, one might assume that he speaks German when being interviewed by Austrian or German broadcasting stations. Interestingly, however, he prefers giving interviews in English and not in his native language (e.g., Gala 2015; Gersemann 2009; Hitradio Ö3 2015; von Uslar 2012; Ziesel 2018). In an interview with a German newspaper (Naumburger Tageblatt 2015), AS was asked about his L2-preference when giving interviews and he stated that he speaks German primarily in his private environment while English has become the language he prefers speaking in public. In addition, he reports that he feels more comfortable when speaking English and that it is easier for him to express himself in English (Gala 2015; Gersemann 2009; von Uslar 2012). It seems that AS has attributed his two languages to different domains, with German being primarily restricted to the private domain. In fact, using their languages for different purposes and in different contexts is frequently observed among bilingual speakers (Fishman 2007; Grosjean 1982; Wei 2007). However, the extent to which AS uses German when interacting with friends and family is hard to determine as there are no (self-)reports documenting his L1 use in more detail. For the purpose of this study, it will be assumed that AS's primary means of communication is his L2, which he predominantly uses in professional contexts and

when interacting with English-speaking friends. The use of his first language is likely to be restricted to the private domain when interacting with German-speaking friends and family. Thus, it can be argued that he is fully immersed in the L2, presumably with limited L1 use.[38]

AS is a famous example of an L2 speaker who regards his nonnative accent as a part of his identity. In interviews, he has repeatedly stated that his distinctive Austrian German accent has become his trademark and that his fans expect him to speak with an accent (Daily Mail UK 2015; Gersemann 2009). In an interview with Daily Mail UK (2015), he even claims that he is well able to speak 'perfect' English but deliberately maintains his accent, which makes it easy to identify him as a nonnative English speaker. If it was the case that AS consciously modifies his pronunciation to make it sound more foreign, it would hardly be possible to draw any reliable conclusions from an investigation into his L2 – and possibly also his L1 – speech development. However, a speaker's ability to consciously maintain or get rid of a nonnative accent is limited due to the fact that some aspects of pronunciation cannot be controlled for by the speaker. Alongside, for example, age-related constraints and the availability of sufficient L2 input, pronunciation involves motoric abilities, which are likely to prevent a speaker from deliberately modifying and adapting their speech patterns (Munro and Derwing 2009). Therefore, as Munro and Derwing (2009, 486) conclude, "[i]t is wrong to interpret accent as an intentional expression of identity if the speaker has no control over these non-volitional features".

In an interview with the Hindustan Times in 2014, AS was asked whether he has ever tried to lose his foreign accent. In his response it becomes clear that putting much effort into losing his accent was never a priority for him:

> *No, I never learned or tried very hard to lose it, I always felt like the most important thing is that you're understandable but I did not really work on losing the accent. I worked on it sometimes to have different accents depends on what the movie is but not to lose it now, I think it's part of the Austrian charm. (AS_E_2014_2)*[39]

In an earlier interview with the American talk show host David Letterman in 1985, AS was asked whether his nonnative L2 accent had an impact on his career as a

38 The notion of *full L2 immersion* is based on Stoehr et al. (2017, 504), who state that being fully immersed in an L2 environment is "inherently tied with L2 use in a variety of contexts and also with numerous speakers, whereas it largely limits L1 use to conversations within the family". Full L2 immersion is distinguished from L2 immersion at home which restricts L2 use to interactions within the family while the L1 is predominantly used for communication outside the home.

39 An overview of the video and audio files the present investigation is based on, including the corresponding file designations, can be found in Appendix A.

bodybuilder. Similar to the 2014-interview with the Hindustan Times, AS states that while he thinks that it is essential to be understood by his interlocutors, he does not consider losing his Austrian accent important:

> *You always have to work on it to be totally understandable [...] it's not important to lose it, I think that some people think it's attractive. (AS_E_1985_1)*

Based on AS's self-reports and on the observation that certain pronunciation features are not controllable and cannot easily be manipulated by a speaker (see Munro and Derwing 2009), it will be argued that AS's L1 and L2 speech examined in this study reflects his authentic pronunciation. In addition, the present study focuses on interviews presenting (semi-)spontaneous rather than scripted speech. In a rather unpredictable and even stressful interview situation it would be quite difficult (if not impossible) to consistently maintain a specific dialect or pronunciation.

Another question, which arises from the discussion above, is whether and to what extent AS can be considered a *bilingual*. As pointed out by Gass and Selinker (2008, 25), "SL researchers reserve the use of the term [bilingual] for only those that are truly [...] the equivalent of native speakers of two languages". This narrow and strict interpretation of bilingualism is – as recent discussions in the fields of SLA and bilingualism research show (see e.g., Grosjean 2010) – rather outdated as it defines bilingualism solely on the basis of (native) language proficiency with regard to all linguistics skills (i.e., speaking and engaging in spoken interactions, reading, and writing). Against the background of this narrow view on bilingualism, AS would rather not be considered a 'true' bilingual given that he retains a detectable foreign accent when speaking English and, thus, does not have native speaker proficiency in both languages. From a more current perspective, however, bilingualism is regarded as a continuum (e.g., Valdés 2001), which takes different dimensions of bilingualism into consideration, such as the age of acquisition, the level of proficiency in both languages, and the context and domain of language use (Wei 2007). This contradicts the concept of *the* bilingual and acknowledges the existence of different *types* of bilinguals, which are likely to show different levels of proficiency with regard to different linguistic skills in their languages and may restrict the use of their respective language to different domains and contexts (e.g., using one language as a primary means of communication within the family and among friends while using the other language in professional contexts). Based on this wholistic view of bilingualism, AS can indeed be considered a bilingual or, to be more precise, a late consecutive bilingual, who started acquiring the L2 in adulthood and who is well able to *use* both of his languages in different communicative situations despite varying levels of proficiency.

2.3.2.2 Materials and procedure

The present investigation of AS's production of word-initial pre-vocalic plosives in his L1 and L2 is based on an acoustic-phonetic analysis of VOT. Speech data were taken from interviews with AS which were conducted in either German or English and were broadcast on different Austrian and/or German and U.S. television and radio channels. The main source for the speech materials was the online video platform *YouTube* (see Appendix A, for an overview of the individual sound files). As the speech materials are freely accessible and part of the public domain, using the sound files for scientific purposes did not entail copyright issues.[40]

After converting the individual recordings into WAV-files and setting the sampling frequency uniformly to 44.1 kHz, each file was cut into several shorter files using PRAAT (Boersma and Weenink 2018); thus, a total of 656 individual sound files was included in the speech corpus. These files contained sequences of AS's speech only; interviewer sequences as well as particularly noisy sequences and those affected by speaker overlaps were not included in the sound samples used for the analysis. Table 2.5 depicts the total durations of English and German sound files according to year. The column *Duration** shows the durations of the sequences which were identified as being usable for an acoustic analysis. As the quality of the sound files largely varied, some files had to be cut to a few seconds only. Thus, sound files with a total duration of 281'02" minutes were included to represented AS's L2 pronunciation, and approximately 19'59" minutes to represent his L1 speech. The German and English speech corpora used for this analysis were divided into *early* recordings, including samples from the late 1970s and early 1980s, and *late* samples representing AS's pronunciation from 2012 until 2017/2018. An additional corpus containing English speech samples from the *mid*-1990s and early 2000s was included as well. A corresponding corpus for AS's German pronunciation could not be assembled because there were no recordings available covering the same time span.

40 According to the Austrian Copyright Act (Urheberrechtsgesetz § 42(2)), duplicating and storing works for scientific purposes is permitted as long as no commercial use is intended: "Jedermann darf von einem Werk einzelne Vervielfältigungstücke auf anderen als den in Abs. 1 genannten Trägern zum eigenen Gebrauch zu Zwecken der Forschung herstellen, soweit dies zur Verfolgung nicht kommerzieller Zwecke gerechtfertigt ist." (UrhG § 42(2))

	English			**German**		
	Year	Total duration	Duration*	Year	Total duration	Duration*
Early	1979	38'26"	22'29"	1982	3'59"	2'24"
	1984	10'39	4'06"	1985	5'58"	2'05"
	1985	12'35"	4'12"	1986	7'42"	3'10"
	1986	04'58"	3'23"			
	1987	04'29"	2'56"			
	1988	18'32"	14'00"			
Middle	1994	7'43"	5'39"			
	1999	48'45"	21'41"			
	2002	74'19"	13'33"			
	2003	20'16"	9'58"			
Late	2012	47'17"	29'22"	2012	4'54"	3'12"
	2013	82'54"	45'52"	2013	10'09"	2'16"
	2014	46'31"	19'23"	2017	6'37"	4'59"
	2015	12'00"	8'01"	2018	21'15"	2'33"
	2016	87'54"	67'25"			
	2017	81'35"	47'42"			
		598'53"	*Total used* 281'02"		60'34"	*Total used* 19'59"

Table 2.5: Total durations of the recordings and durations of sound samples included in the analysis (*Duration) according to year and stage.**

In a next step, each sound file was transcribed orthographically using OCTRA (Pömp and Draxler 2017, version 1.2.7), a web-application for the orthographic transcription of audio files provided by the Bavarian Archive for Speech Signals (BAS) at the University of Munich (Kisler, Reichel, and Schiel 2017; Schiel 1999). Each orthographic transcript is a verbatim transcription of what AS said during the interviews and is based on either American English or German spelling conventions. Repetitions, hesitations, and pauses were included in the transcriptions following the OCTRA transcription guidelines. The individual audio files in combination with the orthographic transcripts underwent automatic phonetic and phonological transcription using WebMaus Basic (Kisler, Reichel, and Schiel 2017), an automatic labelling and segmentation tool for spontaneous, non-prompted speech also provided by the BAS WebServices. Segmentations and labelling were hand-corrected where necessary.

VOT was then labelled manually using PRAAT (Boersma and Weenink 2018). Labelling was based on a combination of both oscillogram and wideband spectrogram displays, which has previously been described as providing more precise measurements than exclusively relying on one or the other (e.g., Abramson and Whalen

2017; Ladefoged 2003; Petrosino et al. 1993). Overall, the database contained 10,782 tokens including a pre-vocalic plosive in word-initial position. From this overall number of tokens, monosyllabic (mo) and bisyllabic (bi) words (content words including proper names, and function words with empathic stress) containing a word-initial plosive followed by a stressed vowel were extracted and selected for analysis (n_{mo} = 2,391; n_{bi} = 1,068; n_{total} = 3,459; see Figure 2.6), i.e., around 30% of the overall number of tokens was used for analysis. All other tokens containing an initial plosive were not included.[41] VOT was then measured from the release of the plosive, indicated by a peak in the waveform and an abrupt spectral change in the spectrogram representing the transient noise burst (see Foulkes, Docherty, and Jones 2011, 64). The plosive release was treated as one acoustic event rather than a separation of transient and frication as suggested by some authors (e.g., Stevens 1993). According to Abramson and Whalen (2017, 83), "[f]or VOT, the separation of transient and frication holds no theoretical importance, so, the two together are classified as the release". The first zero-crossing at the onset of waveform periodicity was determined as the second acoustic landmark, indicating the start of the vowel (see Figure 2.3). On the spectrogram, the onset of periodicity was defined as the occurrence of the first regular vertical striation.[42] For some tokens it was not possible to clearly identify the release of the plosive or the onset of periodicity due to noise, mumbled speech or speech produced with laughter; these tokens were excluded from the analysis (N = 42). After labelling VOT according to the conventions outlined above, VOT durations for the individual tokens were extracted using a PRAAT script.

41 That is, tokens consisting of more than two syllables, and tokens with an initial plosive followed by an unstressed vowel or a consonant (consonant clusters).

42 Different acoustic landmarks have previously been suggested to determine the onset of voicing, such as f0, F1, F2, F3, or – as in the present study – the onset of waveform periodicity, which has been shown to provide the most accurate landmark to determine voicing onset (see Foulkes, Docherty, and Jones 2011; Francis, Ciocca, and Yu 2003, for discussions)

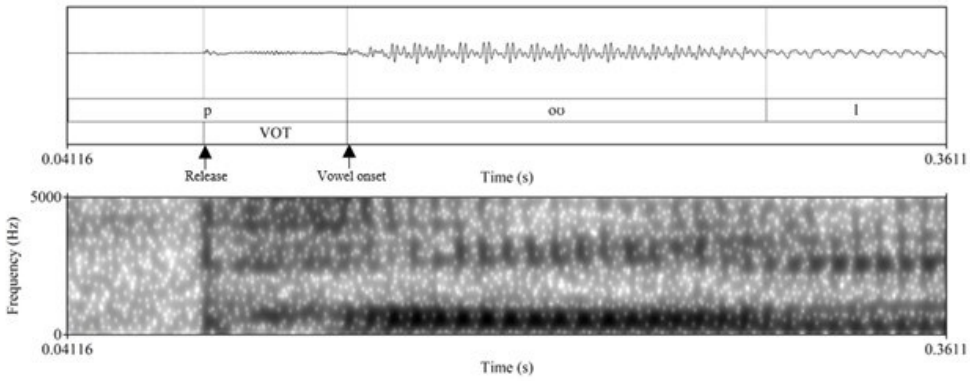


Figure 2.3: Acoustic landmarks (plosive release and vowel onset) used for VOT labelling on the waveform (top) and spectrogram (bottom) display. VOT measured for /p/ in this example is approximately 52 ms.

In some cases, velar plosives were released with a double burst (see Figure 2.4). Double or even multiple bursts are common in velar plosives due to anatomical and aerodynamic factors (see Keating, Westbury, and Stevens 1980, for a discussion) while they occur less frequently or are even completely absent in bilabial and alveolar stops (Fischer-Jørgensen 1954; Olive, Greenwood, and Coleman 1993). There is no consensus among scholars as to whether the first or the second velar burst should be taken as the first acoustic landmark for VOT measurements (see Foulkes, Docherty, and Jones 2011; Gráczi and Kohári 2014). Some suggest measuring from the burst with the highest intensity (see e.g., Fuchs 2005, for German; Lousada, Jesus, and Hall 2010, for Portuguese), others argue that measurements should be taken from the first velar burst, regardless of its intensity (see e.g., Davis 1995). Whenever double bursts occurred during the labelling process in the present investigation, the first burst was determined as the first landmark for VOT measurements, as illustrated in Figure 2.4.

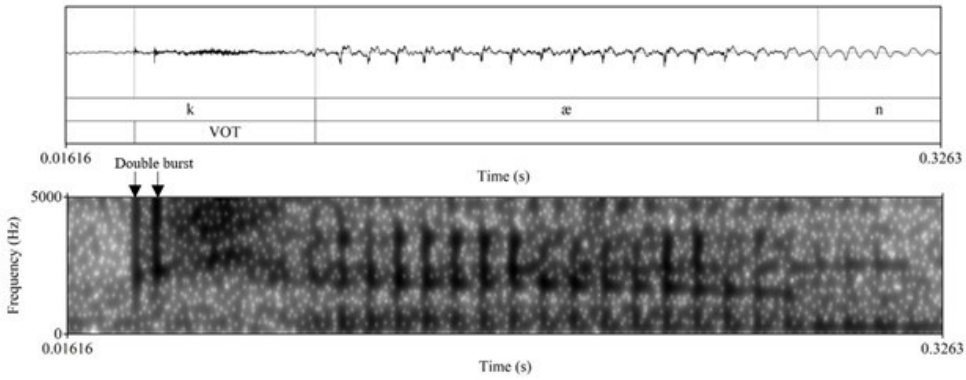


Figure 2.4: Example of a double burst in /k/ in the target word *can*. VOT duration measured from the first burst in this example is approximately 74 ms.

Prevoicing, i.e., vocal fold vibration being maintained throughout the hold phase, was observed in very few instances only. AS's prevoiced two voiced alveolar plosives in his late German, and two voiced bilabial plosives in his early and late English, respectively. Prevoicing was also identified in two voiced alveolar plosives in his late English. The onset of prevoicing was determined as the point at which vocal fold vibration could be identified (see Figure 2.5). The end of prevoicing was identified as the onset of the release of the plosive, indicated by a peak in the waveform and a spectral change in the spectrogram (see van Alphen and Smits 2004). As prevoiced plosives occurred very rarely in the present data set, it can be argued that prevoicing does not represent a systematic feature of AS's English and German pronunciation. Therefore, prevoiced tokens were excluded from further analysis.

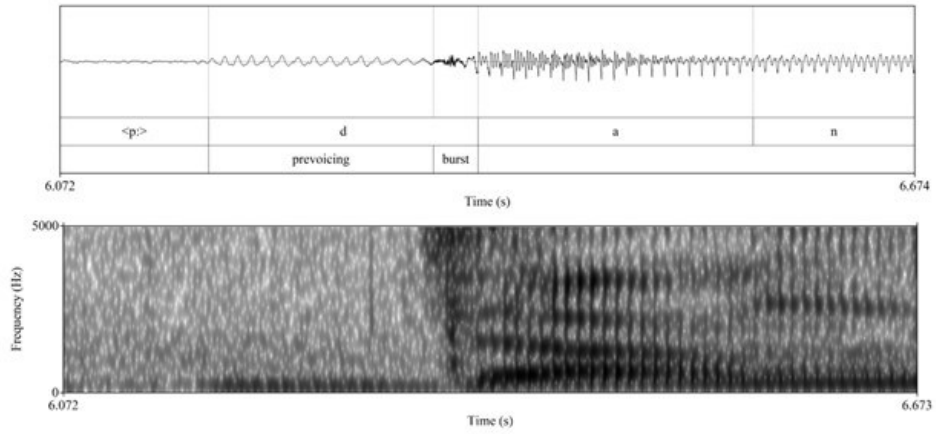
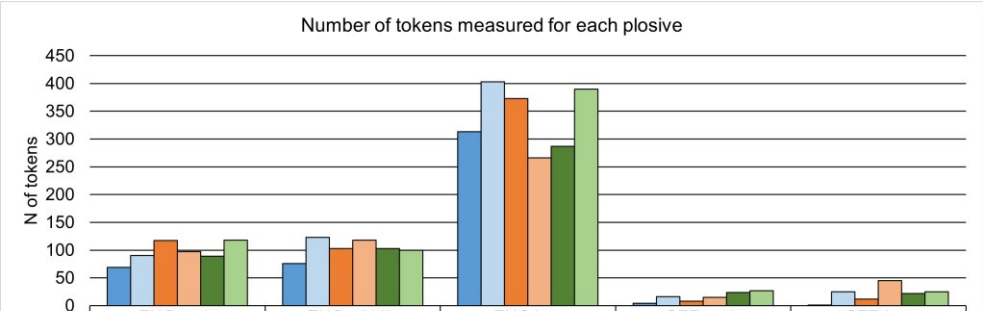


Figure 2.5: Acoustic landmarks for labelling prevoiced tokens on the waveform (top) and spectrogram (bottom) display. Prevoicing measured for German /d/ in this example is approximately –153 ms. The plosive is preceded by a pause (<p:>) of approximately 0.7 seconds.

Figure 2.6 depicts the number of tokens included in the analysis to represent AS's English and German production of plosives in the early and the late stage. As can be seen, the number of tokens largely varied for the two languages: The majority of tokens were included for his late English pronunciation, while only very few tokens could be identified to represent his German pronunciation. This is due to the fact that the present study is based on a corpus of non-prompted, pre-recorded speech, which makes it almost impossible to control for certain factors which are normally taken into consideration when conducting acoustic-phonetic analyses. Advantages and drawbacks which go along with using pre-recorded speech materials will be discussed in Section 2.3.3.



	ENG early	ENG middle	ENG late	GER early	GER late
/p/	69	76	313	4	1
/b/	90	123	403	16	25
/t/	117	103	373	8	12
/d/	97	118	266	15	45
/k/	89	103	287	24	22
/g/	118	100	390	27	25

Figure 2.6: Number of tokens (*N* = 3,459) measured for each plosive in AS's early, middle, and late English, and early and late German pronunciation.

As previously mentioned, VOT was measured in word-initial plosives followed by a vowel in monosyllabic and bisyllabic items. As discussed in Section 2.2.1, the number of syllables in a word (e.g., Klatt 1975; Yu, Nil, and Pang 2015) and the quality of the vowel following the plosive (e.g., Braun 1996; Chao and Chen 2008; Morris, McCrea, and Herring 2008), among other factors, might influence VOT duration. In order to control for influences on VOT duration induced by syllable number and vowel quality, most phonetic studies examine VOT in monosyllabic items containing a word-initial plosive followed by a pre-defined set of vowels (e.g., Chang 2012; Flege and Hillenbrand 1984; Flege and Port 1981; Mayr, Price, and Mennen 2012; Simon 2009; Thornburgh and Ryalls 1998). In order to determine potential influences of syllable number (monosyllabic vs. bisyllabic) and vowel quality (low, mid, high) on VOT in the present study, AS's VOTs depending on syllable number and vowel in English and German was plotted (see Appendix B) and visually inspected prior to conducting the main analysis presented in Section 2.3.4. In terms of AS's English plosives, only minor differences (in most cases less than four milliseconds) were found when comparing mean VOT durations in monosyllabic and bisyllabic tokens. In German, syllable number seemed to affect voiceless velar plosives only, with overall longer mean VOT values measured for /k/ in bisyllables compared to monosyllables. This observation contradicts previous research which shows that VOT duration tends to be shorter in bisyllabic than in monosyllabic tokens (e.g., Klatt 1975; Yu, Nil, and Pang 2015). Also the influence of the vowel on VOT duration in AS's speech does not seem to be systematic: In English /p/ and /t/ followed by low

vowels, VOT tends to be slightly longer compared to the same plosives followed by a high or a mid-vowel. Again, this observation does not confirm previous investigations, reporting shorter VOT values before mid and low vowels, and longer VOT values before high vowels (e.g., Braun 1996; Chao and Chen 2008; Higgins, Netsell, and Schulte 1998; Morris, McCrea, and Herring 2008). In AS's German, longer VOT values can be observed in plosives occurring before high vowels, but only at an alveolar and velar place of articulation (which is presumably due to the rather small number of observations for German /p/). Overall, as this informal visual inspection reveals, the effects of syllable number and vowel on AS's VOT seem to be rather inconsistent, i.e., no systematic correlations could be identified. Therefore, no further (statistical) analyses were conducted. In the statistical tests (linear mixed models) described in Section 2.3.4, *word* was included as a random factor in each model to control for potential variability caused by this factor.

2.3.2.3 Collection of normative data for Austrian German

In order to compare AS's L1 productions of plosives (Study I) and vowels (Study II) to monolingual Austrian German data, a 75-year-old male speaker from Thal, AS's place of birth, was recorded.[43] The speaker was recruited via personal contacts in Thal and was paid 20 Euro for participation. He was born and raised in Thal and has never lived somewhere else. Before being recorded, the participant was asked to fill in a questionnaire to collect information about his language background and use. Based on his answers and a short informal conversation which was held before the actual recording session took place, the participant can be described as a monolingual Austrian German speaker who does not use English or any other foreign language on a regular basis. In fact, he stated that he is hardly able to speak and understand English or any other language despite having learned English and French as foreign languages in school. In addition, he reported using what he referred to as "the typical Styrian dialect spoken by the people in Thal", which he considered to be distinctively different from the more standard-oriented variety which especially younger people living in Graz City speak. The participant reported ormal hearing and did not suffer from any speech impairment.

The recording session took place at the participant's home[44] in a quiet room and was conducted by a female native Austrian German speaker from Graz, who

43 Originally, five speakers from Thal were recruited to be recorded. However, due to the restrictions which have been implemented in connection with the spread of the Covid-19 virus in 2020, it was not possible to collect these data as planned. The recording of the speaker described in this section had been made shortly before the Covid-19 regulations were introduced. Reference data obtained from a single speaker add, of course, limitations to the present investigation, which will be further discussed in Chapter 5.

44 The reason for recording the participant in his home instead of using a professional recording studio was the age of the participant. As the subject was relatively old, recording him in an unfamiliar

received detailed instructions concerning the experimental procedure in advance. After the participant had signed an informed consent form, the experimenter started an approximately ten-minute conversation to create a relaxed and informal atmosphere. After that, the participant received instructions of what he was required to do while being recorded. He was shown four pictures, one after another, and was asked to describe in his own words and at his own pace what he saw on the respective picture (picture description task). A break could be taken any time. The next picture was shown when the participant had named all words included in the test corpus, which will be described further below. In order to make this task more conversational, the experimenter asked questions in between, either relating to one of the pictures, or to a different topic which came up while describing the picture. As the test words (see Appendix C) were relatively simple and easy to display in pictures, the participant did not struggle with identifying a specific word. The complete recording session lasted approximately 50 minutes.

Based on the picture description task, two sets of German test words were elicited (see Appendix C): One set contained monosyllabic and bisyllabic target words with pre-vocalic plosives, which were distinguished in terms of place of articulation and voicing. The second set included monosyllabic and bisyllabic words containing a vowel[45] in stressed onset position. Thus, a total of 52 plosive tokens and 58 vowel tokens was elicited. Due to the semi-spontaneous nature of the elicitation task, it was not possible to provide an equal number of tokens for each plosive and vowel. Labelling and measurements of VOT durations and formant frequencies (F1 and F2) were conducted in PRAAT (Boersma and Weenink 2018) following to the criteria specified in Sections 2.3.2.2 (Study I) and 3.2.2 (Study II), respectively.

2.3.3 Benefits and drawbacks of working with broadcast recordings

The collection of experimental speech data for phonetic analyses is usually based on a tightly controlled design of materials, speaking style, and recording setup (see e.g., Cieri 2011; Harrington 2010; Mertins 2016). In order to ensure a proper recording setup, experiments are often conducted in a sound-proof recording studio with professional recording devices. Thus, effects of background noise can be reduced and the recordings are usually of good quality. Another factor which is typically taken into consideration is speaking style, which has been shown to have an impact on different aspects of speech production (see Harrington 2010, 23). Speaking style refers to whether speech is produced spontaneously or on the basis of an elicitation

surrounding was assumed to cause a stressful and perhaps even intimidating situation, which should be avoided.

45 The vowels collected are the same which are included in the analysis of AS's vowel space and will be described in more detail in Chapter 3.

task, and whether the target items are embedded in carrier phrases or occur in citation form. In addition, speaking style is determined by speaking rate[46] and discourse context (DiCanio et al. 2015). Experimental phonetic research makes use of different tasks to elicit speech data and to control for the context of speech production, such as picture naming tasks (e.g., Harada 2003; Stoehr et al. 2017), film retelling tasks (e.g., de Leeuw, Schmid, and Mennen 2010; Keijzer 2004), or reading tasks containing single target items (e.g., Chang 2011; de Leeuw, Mennen, and Scobbie 2012a). On the basis of these production tasks it is possible to elicit a pre-defined set of speech data containing an exact number of tokens occurring in a specific phonetic context.

As the present study is based on a corpus of pre-recorded speech samples it was not possible to control for all the factors previously outlined. First of all, the individual recordings differed in terms of length and sound quality. The earliest samples were recorded in the 1970s and thus have a rather poor recording quality. In addition, all recordings used for this study are interviews including a minimum of two speakers, which often leads to speaker overlaps and background noise caused, for example, by an audience applauding or laughing, or several people speaking at the same time. This is why not all data were equally usable. One of the most obvious drawbacks was that it was not possible to control for word (or syllable) length and word frequency, which resulted in an unequal number of tokens representing AS's productions of plosives and vowels in his L1 and L2.

However, using pre-recorded speech material which covers a wide time span (i.e., several decades) has also major benefits. First of all, it allows conducting longitudinal analyses. Usually, collecting longitudinal data is rather time-consuming and quite costly if subjects are paid for their participation in an experiment. In addition, long-term studies always have to rely on the continuous availability of participants over an extended period of time (see e.g., Di Paolo and Yaeger-Dror 2011; Duff and Anderson 2015). These factors do not have to be taken into consideration when using broadcast recordings, which are freely available and thus offer a valuable source for longitudinal speech data. In fact, a number of studies have examined different linguistic variables in individual speakers on the basis of broadcast recordings. Harrington and his colleagues (Harrington 2006; Harrington, Palethorpe, and Watson 2000a, 2000b), for example, investigated long-term changes in Queen Elizabeth's monophthongal vowel space by conducting acoustic analyses of the Queen's Christmas broadcasts. Similar studies were conducted by Reubold and Harrington (2015, 2017) and Reubold, Harrington, and Kleber (2010) who assessed

46 As pointed out by Harrington (2010, 23), testing the effect of speaking rate on the production of specific segmental and suprasegmental (prosodic) features can be problematic since "[s]ome speakers may not vary their rate a great deal in changing from 'slow' to 'fast' and one person's slow speech may be similar to another subject's fast rate".

changes in F1 and f0 in the native German speech of the newsreader Dagmar Berghoff and in the native English speech of the British-American radio commentator Alistair Cooke, using broadcast recordings.

Conducting longitudinal research is especially relevant in the context of examining both L2 and L1 speech development because it allows investigating and characterizing the relationship between L2 acquisition and L1 phonetic attrition over an extended period of time. As outlined in Chapter 1, most of the studies on L1 attrition present cross-sectional rather than longitudinal data. Studies which are often cited as longitudinal attrition studies investigate L2 learners over a rather limited period of time, from a few weeks (e.g., Chang 2012, 2013) to just under a year (e.g., Sancier and Fowler 1997). A long-term investigation of L1 attrition effects was carried out by de Leeuw (2019), who explored segmental and prosodic features in the native German speech of the former professional tennis player Stefanie Graf. Acoustic analyses were based on data collected from different TV and radio interviews, which represented Graf's L1 speech development over more than four decades.[47]

Alongside the possibility to conduct longitudinal research, another benefit of working with broadcast recordings is the authenticity of speech. As Di Paolo and Yaeger-Dror (2011) state, using audio data which were recorded in a sound booth and were elicited, for example, by means of a reading task, might not be representative of an individual's authentic and spontaneous speech. A speaker might be affected by the experimental situation, including the location (e.g., a recording studio), the presence of the experimenter(s), and the 'unnatural' situation of being recorded while reading words or sentences from a list. In fact, reading individual words from a list is often described as triggering "the most hyperarticulated form of pronunciation" (Kleber 2018, 470) while spontaneously produced speech "reflects overall abilities the best, allowing especially representative impressions of [...] prosodic patterns and segmental realizations" (Jilka et al. 2008, 247). In formal experimental situations, dialect speakers in particular have been observed not to rely on dialectal pronunciation features but to adapt more standard-oriented pronunciation patterns (e.g., Schmid, Moosmüller, and Kaseß 2015, 4). In addition, it might be the case that the participants notice the aim of the experiment and adapt their pronunciation accordingly. This was, for instance, observed by Flege (2018) who pointed out that some of the subjects in Flege and Eefting (1987a) had in fact recognized the aim of study and thus exaggerated their productions of English

47 Although de Leeuw (2019) found modifications in Graf's L1 speech presumably resulting from L2 acquisition experience, the findings were not compared to potential changes in the speaker's L2. Despite the fact that this study provides insights into the plasticity of the native language system in a late consecutive bilingual speaker, it does not examine the relationship between L2 acquisition of speech and L1 phonetic attrition – a gap the present study aims to fill.

plosives. Consequently, VOT values measured in their plosive productions were not representative of their natural pronunciation.

Of course, also in an interview situation there are several factors which can be assumed to influence an individual's way of speaking, such as the presence of one or more interviewers, other guests and/or a studio audience, as well as the purpose and context of the interview. In addition, an individual's potential nervousness and anxiety of being interviewed in a public setting and/or having to speak in their nonnative language might have an impact on their pronunciation. This is why the materials used in this study are considered representing semi-natural/authentic rather than completely natural speech.

As mentioned above, speaking style is a variable which is almost always systematically controlled for in phonetic experiments, and it also plays a role when it comes to selecting broadcast recordings. For the purpose of this study, only TV and radio interviews were included in the corpus which focus on similar topics, that is, AS's career in the bodybuilding and film business, his childhood and upbringing in Austria, his political engagement as Governor of California, and his organization of environmental campaigns. Political speeches and inaugural addresses were excluded because they are usually representative of a rather formal speech style as they are often scripted and rehearsed.

2.3.4 Results

The aim of Study I was to identify and describe potential changes in AS's realization of VOT contrast in word-initial pre-vocalic plosives in his L1, Austrian German, and his L2, English. To this end, speech data were collected to represent AS's pronunciation over a period of 40 years, i.e., ten years after moving to the United States up to 2017/2018. His English productions were analyzed across three stages, namely early (1970s/1980s), mid (1990s/early 2000s), and late (2010s), while his German productions were compared across the early and the late stage.[48] VOT was measured in a total of 3,459 monosyllabic and bisyllabic target words containing a word-initial plosive followed by a stressed vowel. This section first gives an overview of the descriptive statistics of AS's VOT production in voiced and voiceless plosives representing his early and late English and German pronunciation and presents the statistical models used for analysis. In a next step, AS's L1 and L2 plosives will be compared to monolingual American English and monolingual Austrian German norms.

48 As mentioned in Section 2.3.2, data representing AS's German pronunciation in the mid stage were not available.

	ENG early				**ENG mid**			
Plosive	*Min*	*Max*	*Mean (SD)*	*Median*	*Min*	*Max*	*Mean (SD)*	*Median*
/p/	11.72	131.55	54.05 (23.33)	52.47	23.15	135.86	61.36 (25.05)	55.34
/b/	4.95	31.29	15.98 (5.9)	16.4	5.25	24.46	13.81 (3.93)	13.46
/t/	18.97	134.06	67.58 (22.89)	66.33	19.68	124.69	56.4 (20.13)	53.42
/d/	12.51	39.71	24.01 (6.02)	23.72	11.1	36.9	22.17 (5.74)	22.67
/k/	33.72	115.97	69.8 (18.04)	68.96	41.81	120.8	72.75 (17.27)	71.64
/g/	14.46	51.03	29.35 (7.42)	28.49	14.67	46.28	29.75 (7.3)	28.87
	ENG late				**GER early**			
Plosive	*Min*	*Max*	*Mean (SD)*	*Median*	*Min*	*Max*	*Mean (SD)*	*Median*
/p/	8.23	149.67	48.22 (24.36)	45.46	9.46	24.6	17.62 (7.27)	18.21
/b/	1.66	28.5	13.32 (4.7)	12.49	8.15	27.5	16.49 (4.64)	16.07
/t/	9.67	132.56	50.07 (20.46)	47.42	15.77	44.67	27.46 (11.4)	24.96
/d/	8.66	46.66	20.83 (7.09)	20.06	12.64	34.85	26.13 (6.29)	25.54
/k/	30.43	143.74	68.33 (19.37)	64.53	43.07	102.31	70.78 (16.68)	67.97
/g/	13.67	61.13	31.13 (7.62)	30.85	16.3	35.65	27.70 (4.89)	27.3
	GER late							
Plosive	*Min*	*Max*	*Mean (SD)*	*Median*				
/p/	*	*	*	*				
/b/	4.08	23.43	11.16 (4.85)	11.27				
/t/	14.17	109.42	44.67 (29.54)	34.08				
/d/	8.31	37.91	22.71 (7.62)	22.76				
/k/	44.67	129.24	73.06 (22.15)	72.56				
/g/	18.32	46.13	29.84 (6.28)	28.22				

Table 2.6: Descriptive statistics of VOT durations (in milliseconds) in English (ENG) and German (GER) plosives according to stage (early, mid, and late). *Note: For late German /p/ only one token was included (VOT = 13.59 ms).

(I) Descriptive statistics: AS's English VOT

Figure 2.7 depicts VOT durations measured in AS's English plosives across three stages (early, mid, and late). An inspection of the boxplots shows that, overall, AS's voiced plosives fall within a short-lag VOT range, with slightly longer VOT durations measured in his voiced velar plosives compared to the voiced bilabial and alveolar targets. When comparing the interquartile ranges for voiced and voiceless tokens across all three stages it can be seen that AS's voiced plosives cover a smaller VOT range compared to the voiceless targets, which are characterized by a considerable amount of variability. In terms of AS's production of voiceless plosives, it can also be observed that /p/ and /t/, in particular, tend to be slightly less aspirated in the late stage compared to the early stage.

Overall, Figure 2.7 indicates that AS realizes a VOT contrast for voiced and voiceless targets in his L2 across all three stages. This is particularly surprising in terms of his early L2 productions where one might have expected an L1-influence on his L2 English, i.e., a tendency to neutralize VOT contrast and produce both voiced and voiceless targets within the short-lag VOT range, as typical of his L1 German.

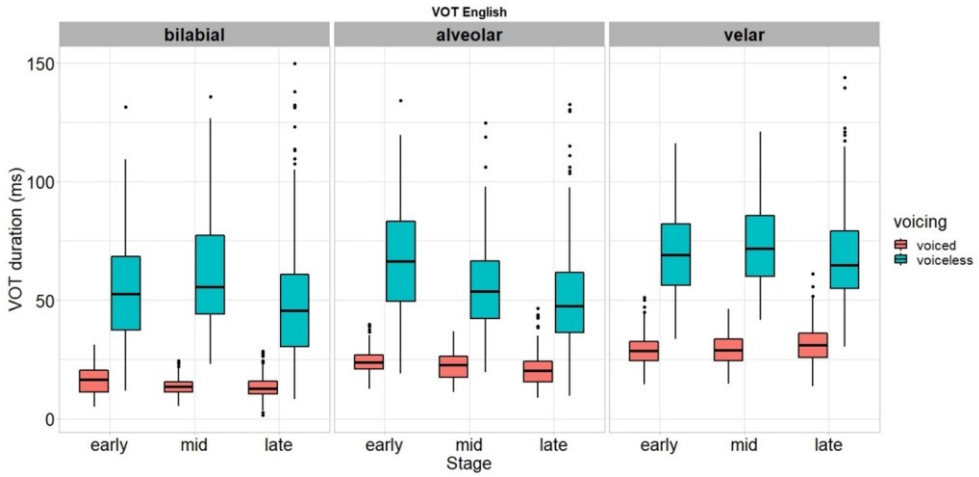


Figure 2.7: VOT durations of AS's English plosives according to stage (early, mid, late) and place of articulation (bilabial, alveolar, velar).

(II) Descriptive statistics: AS's German VOT

Figure 2.8 shows the VOT durations measured in AS's Austrian German plosives across two stages (early and late). As expected, AS realizes a clear contrast between /g/ and /k/ in both stages, with his voiceless productions falling within the long-lag

VOT range and his voiced productions covering the short-lag VOT range. It can be seen, however, that his late productions of /k/ are characterized by a larger amount of variability, ranging from a minimum VOT duration of 44.67 ms up to a maximum of 129.24 ms. In terms of his early bilabial and alveolar productions, a tendency to neutralize VOT contrast can be observed, i.e., he produces both phonologically voiced and voiceless targets within the short-lag VOT range. Despite some overlap between the interquartile ranges of voiced and voiceless alveolar plosives in the late stage, a lengthening of VOT duration in AS's voiceless alveolar plosives can be observed. This suggests a shift from neutralizing the alveolar contrast in the early stage towards realizing an L1 voiced-voiceless contrast in the late stage.

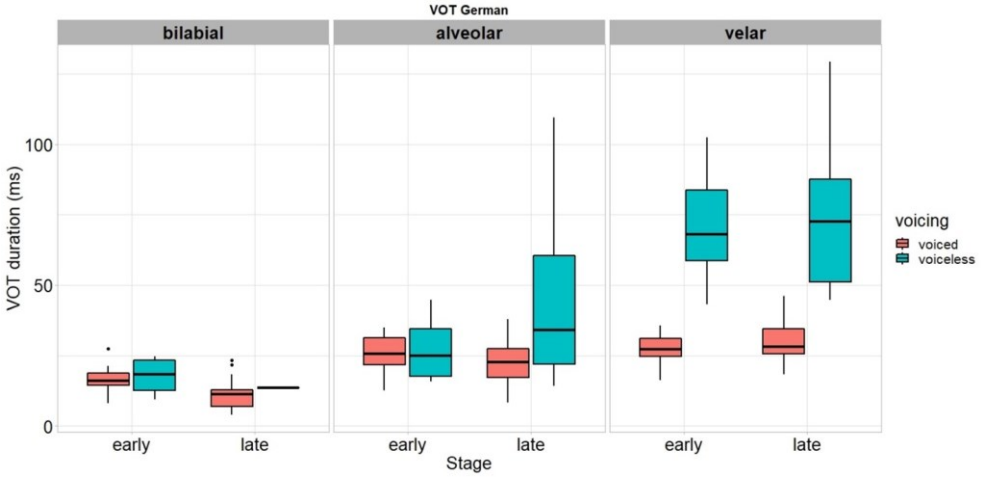


Figure 2.8: VOT durations of AS's German plosives according to stage (early, late) and place of articulation (bilabial, alveolar, velar).

(III) Descriptive statistics: Cross-linguistic comparison between AS's English and German VOT

Figure 2.9 allows for a direct comparison of VOT durations in AS's L1 German and L2 English plosives across two stages (early and late). The most obvious differences in VOT can be observed in terms of the bilabial and alveolar targets when comparing the two stages in both languages: Despite a wider distribution and thus a larger amount of variability which can be observed in AS's realizations of the English voiceless targets, a tendency towards realizing a VOT contrast even in the early stage of L2 learning can be observed. His early German bilabial and alveolar productions, on the other hand, are characterized by a tendency towards neutralizing

VOT contrast as typical of native Austrian German speech. His late German productions, however, suggest a change of VOT of L1 German /t/ in the direction of L2 English, i.e., he now produces the L1 voiceless alveolar targets predominantly in the long-lag VOT range as typical of L2 English.

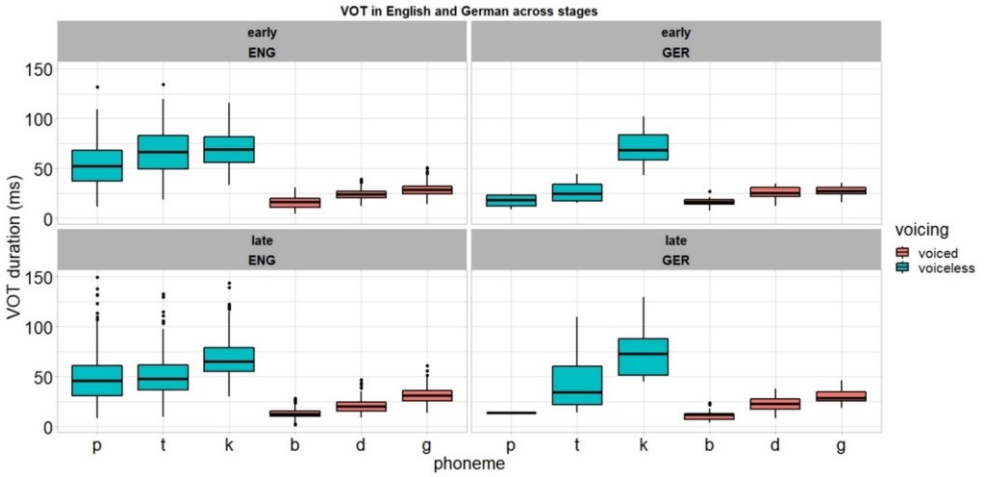


Figure 2.9: VOT durations of AS's English (ENG) and German (GER) plosives according to stage (early, late).

(IV) Statistical analysis

A series of linear mixed-effects regression analyses was performed in R (R Core Team 2020, version 3.6.3), using the *lmerTest* package (Kuznetsova, Brockhoff, and Christensen 2017) including both the *anova* method and the step function, which "performs backward elimination of non-significant effects – both random and fixed" (Kuznetsova, Brockhoff, and Christensen 2017, 1). The linear mixed models were fitted by restricted maximum likelihood (REML) *t*-tests using Satterthwaite's approximations to estimate degrees of freedom. *Post hoc* testing using the *emmeans* package (Lenth 2020)was used to identify pairwise effects. An alpha level of .05 was used throughout testing.

The first linear model regression was conducted to examine potential changes in AS's realization of VOT contrast in his L2 English across three stages (RQ 1.1). In the model, *VOT duration* was included as the dependent variable and *stage*, *phoneme*, and an interaction between the two as fixed effects. As random effects, by-*word* random intercepts and by-*stage* random slopes were entered (duration ~stage*phoneme + (stage|word)). The second regression analysis aimed to examine changes in

AS's realization of VOT contrast in his L1 German (RQ 1.2). The same model specifications outlined above were applied, using German data only. The third linear model regression was conducted to compare AS's early and late productions of plosives in his two languages. The same model specifications applied in the first and second analysis were used with *language* (*lang*) as an additional fixed effect (duration ~stage*phoneme*lang + (stage+lang|word)).

AS's English VOT

The analysis of AS's VOT duration in English plosives showed a significant effect of *stage* (F[2,97] = 23.2, p < .001) and *phoneme* (F[5,328] = 296, p < .001) on VOT duration, as well as an interaction between *stage* and *phoneme* (F[10,97.3] = 8.6, p < .001). Pairwise comparisons using *post hoc* Tukey tests revealed that AS realizes a voiced-voiceless contrast for all places of articulation in all three stages. In terms of AS's production of voiced plosives, no significant differences in VOT duration could be identified when comparing the early and the mid stage (p = .748), the early and the late stage (p = .761), and the mid and the late stage (p = .974), i.e., he consistently produced voiced targets across all places of articulation and across all three stages within the short-lag VOT range, as visible in Figure 2.7. In addition, no significant effects could be observed for the voiceless velar plosives when comparing the early and the mid stage (p = 1), the early and the late stage (p = .996), and the mid and the late stage (p = 1), i.e., the voiceless velar targets were consistently produced with long-lag VOT values. While the analysis did not show significant effects for AS's production of the voiceless bilabial plosives in the early vs. late stage (p = .072), and the early vs. mid stage (p = .999), a significant effect was found for voiceless bilabial plosives in the late and mid stage (t(104.8) = -5.21, p < .0001), showing that, as displayed in Figure 2.7, AS's voiceless bilabial plosives are less aspirated in the late stage compared to the mid stage. Significant effects were also found for the voiceless alveolar plosives in the early and the late stage (t(63.5) = 9.14, p < .0001), and the early and the mid stage (t(94.9) = 4.86, p < .001), while no effects were observed for the voiceless alveolar plosives in the late and the mid stage (p = .279).

AS's German VOT

The second analysis, which focused on VOT duration in AS's Austrian German plosives, showed a significant effect of *phoneme* on VOT duration (F[5,115.4] = 75.6, p < .001), as well as an interaction between *stage* and *phoneme* (F[5,194.6] = 2.4, p < .05). No significant effect was found for *stage* (F[1,179.5] = 0.2, p = .67). Pairwise comparisons using *post hoc* Tukey test showed that, as expected, AS realizes a significant VOT contrast between /g/ and /k/ in the early (t(139) = -9.75, p < .0001) and in the late stage (t(153) = -11.48, p < .0001). In terms of AS's realization of the bilabial contrast in the early stage, the analysis showed that he did not produce a significant

difference between /b/ and /p/ (p = 1.). Similarly, no significant VOT differences were observed in his realization of the alveolar contrast in the early stage (p = .998). In the late stage, however, AS realizes a significant contrast between /d/ and /t/ (t(144)= -4.51, p < .001), i.e., /d/ and /t/ are not neutralized any longer in the late stage, as an inspection of Figure 2.8 suggests.

Cross-linguistic comparison between AS's English and German VOT

The third analysis, which compared AS's VOT in both languages, showed significant interactions between *stage* and *language* (F[1,328.9] = 4.34, p < .05), *phoneme* and *language* (F[5,255.3] = 12.3, p < .001), and a three-way interaction between *stage*, *phoneme* and *language* (F[5,330.6] = 4.6 , p < .001). *Post hoc* Tukey tests showed a significant difference between AS's English and German in terms of his production of /p/ in the early stage (t(195.1)= 5.96, p < .0001), i.e., his early German /p/ falls within the short-lag VOT range, while his early English /p/ is produced within the long-lag VOT range (see Figure 2.9). Similarly, a significant difference was observed between his English and German production of /t/ in the early stage (t(117.3)= 8.15, p < .0001), which confirms that he produces voiceless alveolar plosives in German with short-lag VOT values while his productions of the English counterparts fall within the long-lag VOT range. In the late stage, however, no significant differences between AS's German and English productions of both the voiceless bilabial and the voiceless alveolar tokens (p = 1, and p = 0.947, respectively) could be observed. That is, he produces these target plosives in both of his languages with long-lag VOT values.

Comparison with monolingual American English VOT values

As one objective of the present investigation was to find out whether AS's production of plosives in his L2 has become more native-like since he moved to an L2-speaking environment, the mean VOT durations obtained for AS's early, middle, and late English pronunciations were compared to American English native speaker VOT values reported in Morris, McCrea, and Herring (2008)[49] (see Figure 2.10). Mean VOT durations measured in his early and late German productions were compared to the same monolingual American English VOT values in order to determine if a shift of his late L1 VOT categories towards native L2 norms can be observed (see Figure 2.11). In addition, AS's German VOT was compared to native

49 Although the study by Lisker and Abramson (1964) is one of the most frequently cited studies when it comes to VOT characteristics in different languages, among them American English, the present investigation is based on the VOT values for American English reported in Morris, McCrea, and Herring (2008). Unlike Lisker and Abramson (1964), who examined a relatively small number of speakers (N = 4), Morris et al. (2008) investigated VOT in word-initial plosives in a larger group of native American English speakers (N = 40). Thus, their results were considered more useful for a comparison. However, using VOT values reported in the literature for comparison adds some limitations to the present study, which will be discussed in Chapter 5.

Austrian German VOT data obtained from an age-matched monolingual speaker from Thal (see Section 2.3.2) to identify to what extent AS's L1 productions have moved away from native production norms.

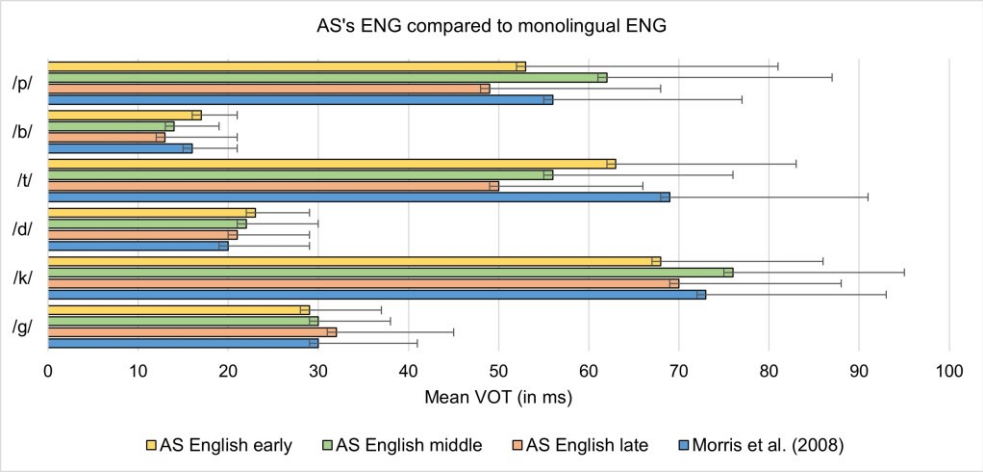


Figure 2.10: Mean VOT durations (in milliseconds) in AS's early, middle, and late English pronunciation compared to monolingual (American) English values reported in Morris, McCrea, and Herring (2008). Error bars indicate +/-1 SD.

As can be seen in Figure 2.10, AS's mean VOT values in his early, late and middle English plosives are within the native American English range, with overall longer VOT durations for the voiceless plosives and shorter VOT for the voiced counterparts. It should be noted, however, that Figure 2.10 (and Figure 2.11) reports mean VOT durations.[50] Only considering the means might create the impression that AS consistently produces a clear VOT contrast in his early, middle and late English, i.e., remains with the short-lag range for voiced plosives and within the long-lag range for voiceless targets. However, as previously discussed and as indicated by the high standard deviations in Figure 2.10, AS's production of VOT and his realization of VOT contrasts is characterized by a considerable amount of overlap between short-lag and long-lag VOT categories, which shows that he neutralizes contrast in some tokens. In addition, when comparing AS's productions to native speaker productions, a shortening of VOT in his bilabial and alveolar voiceless plosives can be identified in the late stage, while his early productions are characterized by a larger

50 Due to the fact that Morris et al. (2008) do not report ranges (maximum and minimum VOT values) or medians, it was not possible to display the variation in the data by means of a boxplot.

degree of aspiration, as typically observed among native (American) English speakers.

Figure 2.11 displays AS's early and late German VOT compared to monolingual American English values (Morris et al., 2008) and monolingual Austrian German values. When comparing his productions of the alveolar contrast in his early and late German to native English VOT values, a tendency towards a lengthening of VOT duration in the voiceless alveolar targets can be observed in the late stage, which indicates a shift towards native English long-lag VOT values. At the same time, his late productions of /t/ have moved away from native Austrian German short-lag VOT (M = 17.98, SD = 6.58). However, as can be seen in Figure 2.11, the mean VOT duration measured for his late German productions of the voiceless alveolar target is still shorter (M = 45, SD = 29) than the native mean VOT durations reported in the literature (M = 69, SD = 16, in Morris et al., 2008). As mentioned earlier, there were only very few tokens which could be included to represent AS's L1 production of the bilabial contrast. Therefore, it was not possible to determine if his realization of the bilabial contrast has also moved in the direction of L2 American English norms.

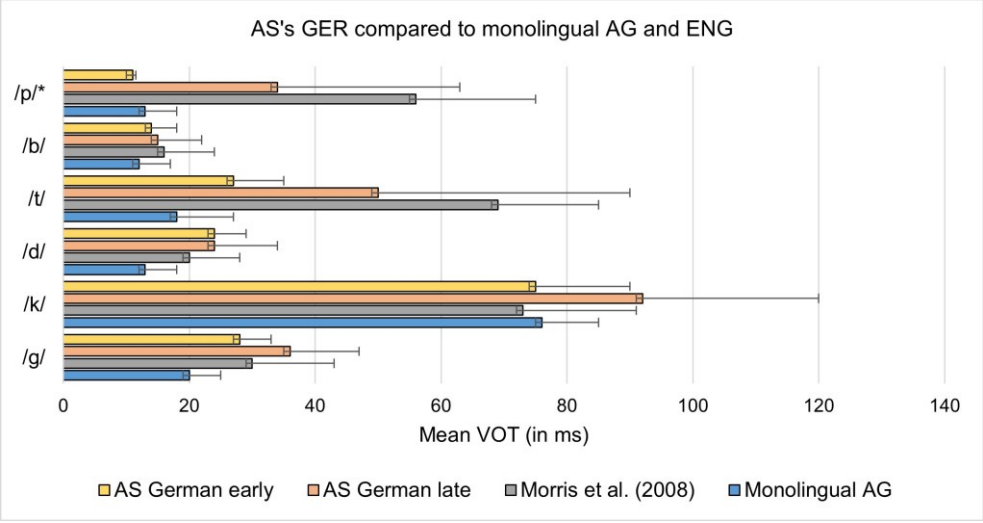


Figure 2.11: Mean VOT durations (in milliseconds) in AS's early and late German pronunciation compared to monolingual (American) English values, reported in Morris, McCrea, and Herring (2008), and monolingual Austrian German (AG) values. Error bars indicate +/-1 SD.*_Note_: For AS's late German /p/ only one token was included (VOT = 13.59 ms).

2.3.5 Discussion of Study I

The present study examined the production of word-initial plosives in the L1 (Austrian German) and L2 (American English) of the late consecutive bilingual Arnold Schwarzenegger. Based on a speech corpus of broadcast interviews, the aim of Study I was to determine if and to what extent AS's L1 and L2 production of plosives has changed over a period of 40 years, i.e., ten years after his arrival in the United States up to the year 2018. His realization of voicing contrast in German and English plosives was examined by means of VOT measurements. In the following, the results of this investigation will be discussed in the light of L1 phonetic attrition and L2 acquisition as well as in the context of potential biological ageing effects.

L2 acquisition of VOT

The first research question (RQ 1.1) addressed in the present investigation aimed to find out to what extent AS's realization of L2 VOT contrast has changed when comparing three stages in time (early, mid, and late), i.e., whether he produces a significant difference between voiced and voiceless L2 plosives in terms of VOT for each of the stages and thus has become more native-like in his L2 in the course of L2 immersion.

The analysis of VOT durations in AS's L2 English plosives revealed that, across all three stages (early, middle, late), his VOT values fell – at least partly – within the native range, i.e., he realized some of the voiceless targets within the long-lag (aspirated) VOT range and the voiced counterparts within the short-lag range, as typical of English plosives (e.g., Chao and Chen 2008; Lisker and Abramson 1964). When considering his *early* English pronunciation, the observation that he realizes a voiced-voiceless distinction for some tokens is surprising given the fact that native Austrian German speakers tend to neutralize VOT contrast in bilabial and alveolar plosives in spontaneous speech and produce *both* phonological voiced /b, d/ and voiceless /p, t/ within the short-lag VOT range (Hödl 2019; Moosmüller and Ringen 2004; Moosmüller, Schmid, and Brandstätter 2015). Due to L1-influences on L2 speech production, it might have been expected that AS transferred the L1 feature of neutralizing VOT contrast to his L2, which would have resulted in nonnative productions of L2 /p/ and /t/, especially in the early stage. In fact, L1-influences on L2 production abilities have not only been documented in relatively inexperienced, non-immersed L2 learners (e.g., Flege 1980; MacKay et al. 2001), but also in highly proficient L2 speakers who have been long-term immersed in an L2 environment. Mayr, Price, and Mennen (2012), for example, found that their L1 Dutch L2 English subject produced voiced plosives in her L2 English with a considerable amount of prevoicing, as typical of Dutch (see Section 2.2.3). The L1 French L2 English subjects in Flege (1987b), who had been long-term immersed in an L2 country, produced L2

English /t/ with significantly shorter (and thus more French-like) mean VOT compared to monolingual English control speakers, which suggests an L1-influence on L2 pronunciation. What might be indicative of an L1-influence on L2 VOT production in the present study is the considerable amount of variability observed in AS's productions of L2 voiceless plosives across all three stages. This variability shows that he is *inconsistent* in distinguishing L2 voiced and voiceless plosives, i.e., he does not realize a distinct and native-like contrast for all English tokens but neutralizes contrast in some bilabial and alveolar tokens as typical of his L1 Austrian German. Previous research into the acquisition and realization of VOT categories also provides evidence for variable production patterns in bilingual speakers (e.g., Flege 1991; Hazan and Boulakia 1993), and further suggests that the degree of variability tends to be larger in (late) bilinguals than in monolingual speakers (MacLeod, Stoehl-Gammon, and Wassnik 2009), which, as stated above, might be a result of L1-influences on the L2 system.

Another aspect which might account for the considerable amount of variability observed in AS's L2 voiceless plosives are potential articulatory constraints, which have been shown to affect the acquisition of aspiration for English voiceless plosives in particular. Previous research examining the acquisition of VOT categories in English-speaking children shows that children need longer until they are able to produce voiceless plosives than voiced plosives in their native language (e.g., Kewley-Port and Preston 1974; Lowenstein and Nittrouer 2008; Macken and Barton 1980; Preston et al. 1968; Yu, Nil, and Pang 2015). According to Yu, Nil, and Pang (2015, 153), the reason for this is that voiceless targets produced with aspiration "require fine temporal coordination to delay the onset of laryngeal vibration relative to oral closure release" (see also Whiteside and Marshall 2001). While most research examining the relatively higher degree of difficulty involved in the acquisition of voiceless aspirated stops has been conducted in the field of child language acquisition (see studies cited above), Stoehr et al. (2017) also observed this phenomenon in their late Dutch-German bilinguals (see Section 2.2.2). The L1 Dutch L2 German subjects in this study failed to produce German voiceless plosives with aspiration and produced de-aspirated plosives instead. Similar observations were made in the present investigation: As discussed above, AS's productions of English voiceless targets are characterized by a high amount of VOT variability, with some plosives falling in the short-lag VOT category (as typical of L1 Austrian German) while others are produced within the long-lag (aspirated) VOT category (as typical of L2 English). This inability to realize a consistent VOT contrast for all English tokens might be partly related to articulatory constraints which may impede a speaker's ability to consistently produce voiceless (aspirated) plosives.

As illustrated in Figure 2.12,[51] an interesting development was observed in AS's productions of English /t/: When comparing early and late /t/, a significant shortening of VOT was identified in the late stage, which resulted in a reduction of aspiration in late /t/. By contrast, his early realizations of English /t/ are aspirated, i.e., they are characterized by considerably longer VOT values, which fall within the native (American) English range. One might have expected that these changes go in the opposite direction, that is, from early slight aspiration to a higher degree of aspiration in the late stage, and thus indicate an improvement of L2 production abilities in the late stage. Interestingly, however, AS's productions of English /t/ are most native (most aspirated) in the early stage and then become less aspirated in the mid and in the late stage. Similar observations were made concerning his productions of English /p/ in the mid and in the late stage: His voiceless bilabial plosives are slightly more aspirated in the mid stage while his *late* productions of English /p/ are characterized by less aspiration. Thus, as discussed in the context of the Speech Learning Model (Flege 1995, 2002, 2003; see Section 1.6.1), it can be assumed that AS has established a *merged* L1–L2 category and produces the late bilabial and alveolar targets with VOT values which are intermediate between the long-lag (aspirated) VOT category in L2 English and the short-lag (unaspirated) VOT category in L1 Austrian German. The observation that no significant changes were evident in his early vs. late productions of English /k/ were not surprising given that both English and Austrian German have (strongly) aspirated velar plosives.

51 The VOT categories displayed in Figure 2.12 are based on Cho and Ladefoged (1999, 223), who identified the following four VOT ranges for the categorization of voiceless plosives in different languages: (1) unaspirated (up to 30 ms), (2) slightly aspirated (averaging 50 ms), (3) aspirated (around 90 ms), and (4) highly aspirated (longer than 90 ms). According to Moosmüller and Ringen (2004), these ranges are also useful when it comes to classifying Austrian German voiceless plosives according to their degree of aspiration (see Figure 2.13).

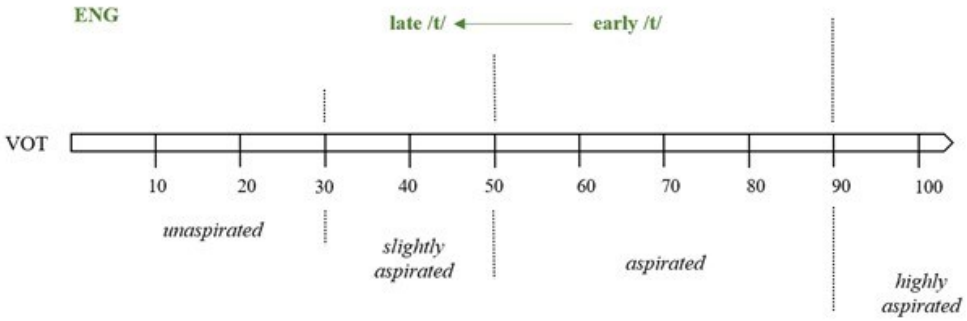


Figure 2.12: Schematic representation of VOT categories for AS's early and late English (ENG) voiceless alveolar plosives.

The observation that AS's realization of VOT contrast in his L2 English is more pronounced (i.e., more native-like) in the early stage, as manifested in a higher degree of aspiration, and has reverted to a less native-like production in the late stage, may be explained against the background of his language learning history. As discussed in Section 2.3.2, AS invested much time and effort into learning English at the very beginning of his L2 immersion by, for instance, attending English as a Second Language classes at a community college for several months and practicing with native American English speakers (Schwarzenegger 2012). Acquiring English pronunciation was one of the most challenging tasks he was confronted with after migrating to the United States (Schwarzenegger 2012), and his lack of pronunciation skills turned out to be problematic even after he had already lived in the United States for several years, when he played in his first movie *Conan the Barbarian* in 1982 (Miller 2012). In order to improve his pronunciation and acting skills, AS worked with acting and dialogue coaches in preparation for the movie (ibid.). Taken together, his endeavors to sound more native in his L2 might have resulted in an explicit awareness for American English pronunciation and, at the same time, in an improvement of certain L2 pronunciation features, such as producing aspirated plosives as typical of native English. Not much is known, however, about if and to what extent he continued working on his English pronunciation in the later years when he had already established himself as a successful actor and businessman. His success no longer depended on sounding more native, but film producers and the media began to appreciate his 'unique' way of speaking (AS_E_1985_1), which may be a reason why he did not continue working on his second language pronunciation. However, the extent to which AS worked on his English pronunciations in the early stage of L2 immersion cannot be further specified since it is only possible to refer to anecdotal evidence. Moreover, the present study is based on a speech corpus

representing AS's pronunciation over the past 40 years, i.e., ten years after he moved to the United States up to 2018. Therefore, it was not possible to examine potential changes in his L2 VOT productions in the first ten years of L2 immersion, that is, right from the onset of his L2 learning experience. It might have been the case that the most drastic changes happened within the first few years of the L2 acquisition process when, as discussed above, he received intensive L2 instruction and pronunciation training. Further research including speech data representing AS's L2 pronunciation from the very onset of L2 acquisition will be necessary to test this assumption.

L1 attrition of VOT

The second research question (RQ 1.2) addressed in this study was to what extent AS's realization of VOT contrast in his L1 German has changed when comparing two stages (early and late), i.e., whether he produces a significant difference between voiced and voiceless plosives in terms of VOT for each of the stages and has thus become less native-like in his L1, resulting from L2 learning experience.

As expected, the analysis of his *early* German plosives showed that he does not realize a significant VOT contrast for /p, b/ and /t, d/, that is, he produces both phonologically voiced and voiceless targets within the short-lag VOT range, which is a typical feature of native Austrian German conversational speech (e.g., Hödl 2019; Moosmüller, Schmid, and Brandstätter 2015). By contrast, AS realizes a significant short-lag vs. long-lag contrast for /g/ and /k/, which is also characteristic of monolingual Austrian German pronunciation.

In terms of his *late* German alveolar plosives, a lengthening of VOT duration in the direction of the L2 was observed for /t/ (from early $M = 27.46$ ms to late $M = 44.67$ ms), which might be indicative of an L2 English influence on L1 production abilities. In terms of AS's late productions of /p/, no reliable results could be obtained due to the limited number of observations; a lengthening of VOT duration in his voiceless bilabial plosives in his late German pronunciation can only be assumed. As discussed in the previous section, the most significant changes in AS's L2 development were also observed in his productions of English /t/, which are produced with less aspiration in the late stage compared to the early stage. By contrast, an opposite movement of VOT categories was observed in AS's L1 German when comparing his early (unaspirated) and late (slightly aspirated) voiceless alveolar plosives. As can be seen in Figure 2.13, his late productions of the target plosive in both of his languages fall within the same VOT range, which is most likely the result of a merging of VOT categories (Flege 2003) and, thus, indicative of a mutual interaction between the L1 and the L2 system. However, it needs to be taken into consideration that, as already discussed in the previous section, AS's early and late productions of voiceless plosives in both of his languages are characterized by a considerable

amount of variability. That is, his early and late voiceless alveolar plosives do not consistently cover the VOT ranges displayed in Figure 2.13.

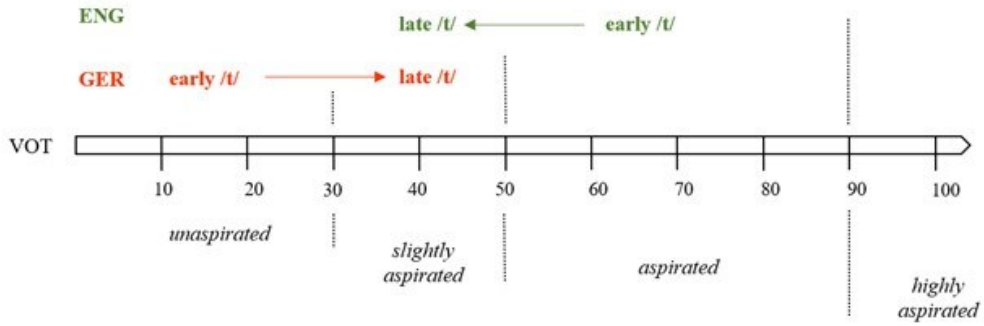


Figure 2.13: Schematic representation of VOT categories for AS's early and late English (ENG) and German (GER) voiceless alveolar plosives.

In addition to a lengthening of VOT in AS's late L1 productions of /t/, changes were also observed in his late German productions of /k/. Although no significant differences in terms of mean VOT duration could be identified when comparing early and late productions of the target plosive (early M = 70.78 ms, late M = 73.06 ms), his late German /k/ covers a broader VOT range (84.57ms) compared to the early counterpart (59.24 ms), i.e., his late German velar productions are characterized by a larger degree of variability. By contrast, his productions of German /g/ did not undergo significant changes, but remained within the short-lag VOT range.

The findings outlined above are in line with previous studies on L1 attrition of VOT, which suggest that voiceless plosives are prone to be affected by attrition processes (see Mayr, Price, and Mennen 2012; Stoehr et al. 2017). Stoehr et al. (2017), for instance, found that the L1 German L2 Dutch bilinguals in their study produced L1 voiceless plosives with considerably shorter and thus more Dutch-like VOT values. Similar observations were made by Mayr, Price, and Mennen (2012): Their L1 Dutch L2 English bilingual produced voiceless plosives in both languages with VOT values which were intermediate to native Dutch and L2 English norms while L1 voiced plosives remained unaffected by L2-influences. As stated above, these findings are also reflected in the present study since AS's productions of /t/ in his late German speech is characterized by a considerable lengthening of VOT in the direction of L2 English norms. By contrast, his L1 voiced plosives were not expected to change significantly because voiced plosives in both English and Austrian German are characterized by short-lag VOT values.

Biological ageing effects

Several studies have been conducted to examine ageing effects on VOT duration by comparing groups of older and younger monolingual speakers of different languages. Some of these studies suggest that with *increasing* age mean VOT durations of word-initial plosives *decrease* (e.g., Ryalls, Simon, and Thomason 2004; Ryalls, Zipprer, and Baldauff 1997; Smith, Wasowicz, and Preston 1987) while others show a *lengthening* of VOTs in *older* speakers (e.g., Bóna 2014).[52] Despite inconsistent results as to whether VOT durations increase or decrease with age, the studies mentioned above show that biological ageing effects on VOT seem to be most pronounced when it comes to the production of voiceless stops while their voiced counterparts seem to remain largely unaffected. Another study worth noting in this context was conducted by Petrosino et al. (1993): Unlike the studies previously mentioned, their comparison of mean VOT durations in velar stops produced by older (aged 70–87) and younger (aged 20–30) monolingual American English speakers did not reveal significant VOT differences, i.e., both age groups produced similar VOT values (see also Neiman, Klich, & Shuey, 1983). However, a considerably larger amount of VOT variability was observed in the older speaker group. Similar observations were made by Sweeting and Baken (1982), who detected a significant increase in VOT variability in their oldest speaker group (>75 years), particularly in terms of /p/. Petrosino et al. (1993) argue that the variability observed among older subjects might be due to natural anatomical and physiological changes which go along with ageing processes (see also Benjamin 1982; Smith, Wasowicz, and Preston 1987). Variability and broad overlaps of different VOT categories were also identified in the present study, which might not only be indicative of cross-linguistic influences but could also be related to natural ageing effects. However, a comparison of AS's early and late productions of /t/ and /p/ in both of his languages showed a shortening of VOT in his L2 /t/ and /p/ in the late stage (i.e., with increasing age). At the same time, his L1 productions of /t/ (and presumably /p/) were characterized by a lengthening of VOT. In the context of biological ageing effects one would expect to observe either a lengthening or a shortening of VOT in *both* languages. Therefore, the findings of the present investigation cannot be reliably attributed to age-related changes, but are rather considered being the result of bidirectional interactions between the L1 and L2 phonetic system as discussed in the previous sections.

52 Bóna (2014) observed longer VOT values for voiceless bilabial and alveolar plosives in older speakers of Hungarian, but their productions of the voiceless velar targets were characterized by considerably shorter mean VOT durations compared to the younger speakers.

3 Study II: Arnold Schwarzenegger's production of vowels in his L1 and L2

3.1. Introduction to Study II

Study II aimed to examine vowel production in the late consecutive bilingual Arnold Schwarzenegger over the past 40 years. Potential modifications and changes in AS's L1 and L2 monophthongal vowel systems were described and compared by conducting formant frequency analyses, focusing on the first and the second formant (F1, F2). The present investigation is based on the same speech corpus used in Study I (Chapter 2), that is, English and German interview sequences which have been broadcast on different television and radio channels (see Section 2.3.2, for a detailed description of the corpus).

The following sections provides an overview of the vowel systems of English, with a particular focus on Californian English, and Austrian German. In Sections 3.1.3 and 3.1.4, previous research exploring different aspects of vowel production in the context of both L1 phonetic attrition and L2 speech acquisition will be discussed in more detail.

3.1.1 American English vowels

As discussed in Chapter 2, AS's L2 variety can be broadly described as American English, with some Californian English influences given that he had spent most of his life in California at the time the present study was conducted. Of course, it can be assumed that he was and still is exposed to other (American) English varieties, but Californian English is presumably the variety he is immersed in and exposed to in his daily life. The American (Californian) English monophthongal vowel inventory contains five front vowels, /i, ɪ, e, ɛ, æ/, four back vowels, /u, ʊ, o, ɑ/, the open-mid central vowel /ʌ/, and mid-central /ɝ/ (Hagiwara 1997; Ladefoged 2005).

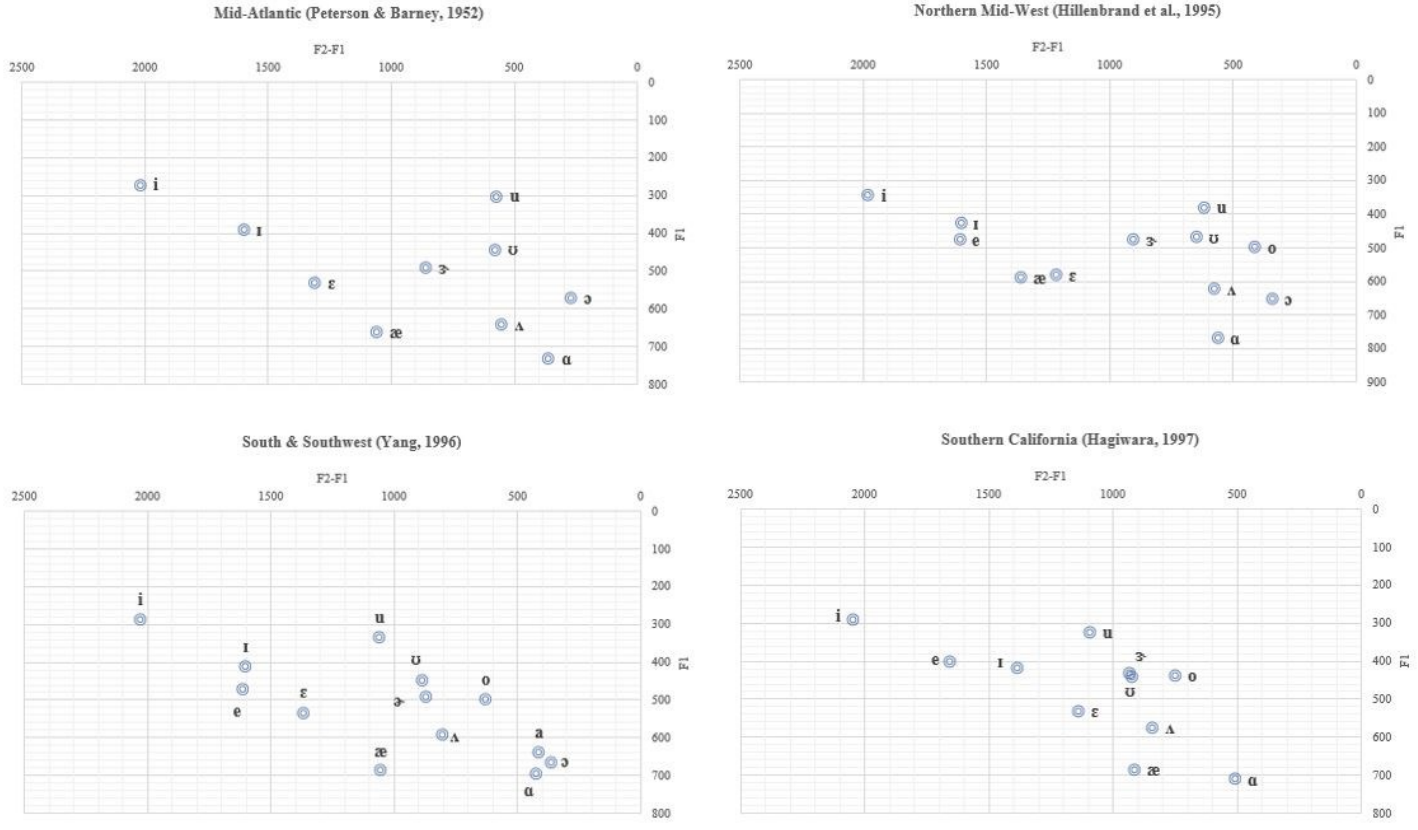


Figure 3.1: Formant frequencies (F1, F2) for male speakers measured in previous studies for the Mid-Atlantic speech area (Peterson and Barney 1952), the Northern Mid-West (Hillenbrand et al. 1995), the South and Southwest (Yang 1996), and Southern California (Hagiwara 1997).

The vowel plots in Figure 3.1 are based on formant frequency measurements (F1, F2) obtained in previous studies on American English vowels.[53] One of the earliest and perhaps most frequently cited investigations of American English vowel acoustics was conducted by Peterson and Barney (1952), who examined vowel productions in 76 male and female speakers from the Mid-Atlantic speech area.[54] It should be noted, however, that the subjects included varied in terms of their language background. Some speakers were, for instance, nonnative speakers of English and/or born outside the United States, and most of the male subjects spoke General American. Therefore, the speakers' vowel productions are not necessarily representative of the Mid-Atlantic speech area. Hillenbrand et al. (1995) replicated this study and recorded 139 speakers of the Northern Mid-West dialect, including men, women, and children. The subjects were tested on their productions of American English monophthongs in /h/–Vowel–/d/ syllables. Another replication study of Peterson and Barney (1952) was conducted by Hagiwara (1997), who analyzed formant frequencies in 15 speakers (6 male, 9 female) of Southern Californian English. Finally, Yang (1996) compared 20 American English speakers from the South and Southwest to speakers of Korean. While there are hardly any phonetic differences between individual American English varieties when it comes to consonants (Ladefoged 2005, 45), the quality of American English vowels can differ significantly due to dialectal variation as shown in the vowel plots above. Therefore, the Californian English vowel inventory (Hagiwara 1997; Ladefoged 2005) will serve as a basis for comparing AS's L2 vowels with native American English vowels in the present study.

The most striking differences between GA and Californian English can be identified in the production of the vowels in *boo* and *hood*: Californian English /u/ and /ʊ/ are characterized by a higher F2 frequency compared to GA, i.e., these vowels are produced at a more front position compared to GA (see Ladefoged 2005, 45–46). In addition, speakers of Californian English tend to realize the STRUT vowel with a slightly lower F1 and a higher F2, i.e., /ʌ/ is produced more front and more close compared to GA. According to Chang (2012, 254), the tendency to produce /u/ and /ʊ/, as well as /ʌ/, with a relatively high F2 is a result of the California Vowel Shift, which affects different parts of the Californian speech area. Another distinctive feature of the Californian Vowel Shift is the complete merging of /ɑ/ and /ɔ/ in the direction of /ɑ/, i.e., the vowels in *bod* and *bawd* have the same quality (Boberg 2005; Eckert 2004; Hinton et al. 1987; Janoff 2018). Consequently, native Californian

53 In Figure 3.1, only the results for male speakers are displayed since they are most relevant in the context of the present investigation which focuses on vowel productions in a male subject.

54 The Mid-Atlantic accent is described as a distinct American English dialect which includes the Mid-Atlantic cities of Philadelphia, Reading, Atlantic City, Wilmington, and Baltimore (Labov, Ash, and Boberg 2006, 236).

English speakers are less likely to distinguish between the two vowels in production as well as in perception, as shown in a vowel identification test conducted by Hillenbrand et al. (1995).

A further characteristic of Californian English varieties concerns the production of the high front vowel /ɪ/ in velar contexts: While RP speakers produce the same vowel in *sing* and *sin*, Californian English speakers typically produce the vowel in *sing* with the same quality as the vowel in *seen* (Ladefoged 2005, 31).

3.1.2 Austrian German vowels

The vowel inventory of (Standard) Austrian German contains nine front vowels, /i, y, ɪ, ʏ, e, ø, ɛ, œ/, and five back vowels /u, ʊ, o, ɔ, ɑ/. The majority of studies examining acoustic characteristics of Austrian German vowels focus on speakers of Standard Austrian German (SAG) from Vienna or Salzburg (Brandstätter and Moosmüller 2015; Cunha et al. 2015; Cunha, Harrington, and Hoole 2013; Moosmüller 2007). As illustrated in Figure 3.2, SAG distinguishes 14 vowels, which can be divided into rounded and unrounded monophthongs (Moosmüller 2007; Moosmüller, Schmid, and Brandstätter 2015). In Standard German German (SGG) the high vowels /i, ɪ, y, ʏ, u, ʊ/ are clearly differentiated according to vowel quality, e.g., German tense /i/ in *bieten* 'to offer' vs. German lax /ɪ/ in *bitten* 'to beg' (see Jessen et al. 1995). SAG speakers, however, show a tendency to neutralize the tense-lax-contrast in the vowel pairs /i/-/ɪ/, /y/-/ʏ/, and /u/-/ʊ/ in the direction of the tense counterpart (Brandstätter, Kaseß, and Moosmüller 2015; Cunha et al. 2015; Cunha, Harrington, and Hoole 2013; Moosmüller 2007). Brandstätter and Moosmüller (2015), for example, investigated the realization of the /i/-/ɪ/ contrast in 12 speakers of SAG from Vienna who read word lists containing the target vowels in plosive environments. Analyses of vowel duration and formant frequencies revealed a tendency towards eliminating the /i/-/ɪ/ contrast in the direction of the FLEECE-vowel, particularly in velar contexts. Similar observations were made by Cunha et al. (2015) who examined the influence of three consonantal contexts on the realization of seven tense-lax contrasts[55] in SAG and SGG read speech. They identified a greater approximation of lax vowels towards the respective tense counterpart in SAG speakers. It was also shown that the implementation of the tense-lax contrast was significantly influenced by consonantal context: Overall, the distinction between tense and lax pairs was smaller when the vowel targets occurred in velar contexts, which is in line with the results reported by Brandstätter and Moosmüller (2015). While SAG speakers showed a tendency to neutralize some tense-lax contrast, vowel quantity was

55 These were /e:/-/ɛ/, /ø:/-/œ/, /a:/-/a/, /i:/-/ɪ/, /y:/-/ʏ/, and /u:/-/ʊ/.

equally stable in both speaker groups, i.e., SAG speakers were not observed to neutralize vowel duration.

The chart displaying the Austrian German vowel inventory (Figure 3.2) is based on Moosmüller's (2007) acoustic analysis of monophthongs produced by six speakers of SAG from Vienna. SAG vowel tokens were elicited in different production tasks, including short interviews (spontaneous speech) as well as sentence and word list reading tasks (read speech). Moosmüller (2007) elicited two different sets of vowel tokens in each production task, one set containing stressed tokens and the other set unstressed tokens. For the sake of comparison with Californian English vowels (Hagiwara 1997), which were elicited in a formal reading task, the SAG vowels displayed in Figure 3.2 represent stressed vowels produced in a sentence reading task. When comparing the formant frequencies of the monophthongs which both SAG and Californian English share, the following observations can be made: The English back vowels /u/ and /ʊ/ tend to have higher F2 values, i.e., they are produced more front compared to their SAG counterparts, while the English front vowels /ɪ/, /ɛ/, and /e/ have lower F2 values, i.e., they are produced more back compared to SAG, which has higher F2 values. In terms of F1, American English /ʊ/, /ɪ/, /ɛ/, and /e/ are produced with slightly higher F1 values compared to SAG, that is, speakers of American English tend to produce these vowels more open than speakers of SAG. As can be also seen in Figure 3.2, the SAG vowel inventory lacks a mid-central vowel, which has been observed to "exist neither phonetically nor phonologically" (Moosmüller 2007, 52). The American English inventory, by contrast, includes the mid-central vowel /ɝ/ as described in the previous section. In terms of the German vowels /ɑ/ and /a/, it has been shown that SAG speakers produce hardly any differences in F1 and F2 between these two vowels (Moosmüller 2007, 115). As Moosmüller (2007) observed, the greatest degree of distinction in F1 for /ɑ/ and /a/ occurs in non-words produced in formal reading tasks. F2 frequencies, by contrast, are not discerned at all, independent of the context of production.

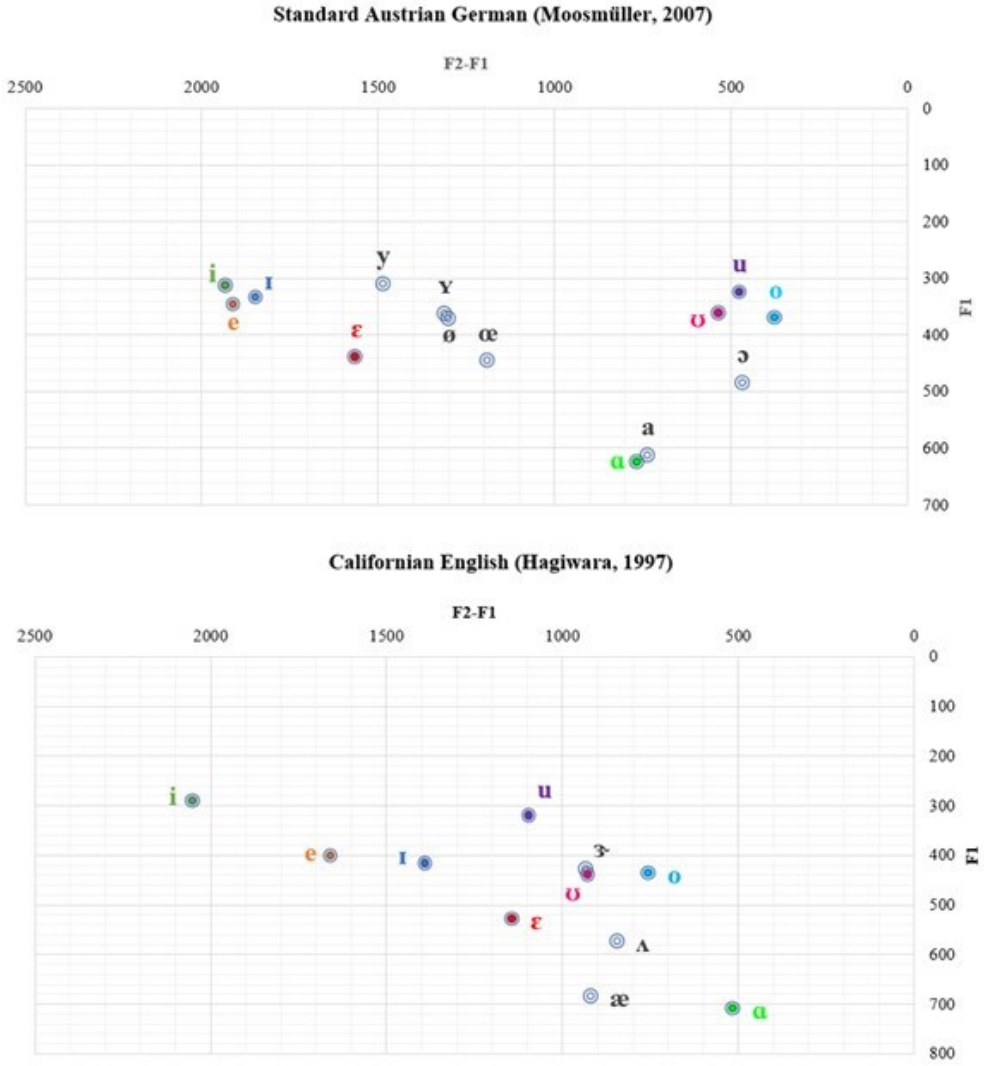


Figure 3.2: Vowel charts of Standard Austrian German vowels (Moosmüller 2007; Moosmüller, Schmid, and Brandstätter 2015) and Californian English vowels (Hagiwara 1997). Vowels which exist in both vowel inventories are displayed in the same colors.

3.1.3 L2 acquisition of vowels

A considerable amount of research has been conducted to examine the acquisition of nonnative vowels by L2 learners (see Table 3.1). Since the present study investigates the acquisition of L2 speech in a late consecutive bilingual, the studies selected for the overview provided in this section include subjects with the same order of acquisition, i.e., learners who acquired the L2 late in life after the L1 had already been fully developed. Most of the studies displayed in Table 3.1 (Baker and Trofimovich 2005; Flege, Schirru, and MacKay 2003; Oh et al. 2011; Piske et al. 2002; Tsukada et al. 2005) examined L2 acquisition of vowels in late bilinguals and compared their productions to early bilinguals in order to examine effects of age of learning and L2 experience on L2 vowel production accuracy. Studies focusing exclusively on simultaneous and early sequential bilinguals are not included in this overview (e.g., Darcy and Krüger 2012; Kehoe 2002; Khattab 2007; Mack 1989; MacLeod, Stoehl-Gammon, and Wassnik 2009), because they do not allow for conclusions to be drawn about the potential difficulties late L2 learners might encounter when it comes to acquiring L2 vowels.

Study	Languages	N of subjects	Vowels tested	Elicitation method
Flege, Bohn, and Jang (1997)	L1: German, Spanish, Mandarin, Korean L2: English	90	English /i ɪ ɛ æ/	Reading: target words in carrier phrase
Flege, MacKay, and Meador (1999)	L1: Italian L2: English	72	English /i ɪ e ɛ æ u ʊ o ʌ ɒ/	Combination of visual and auditory prompts
Piske et al. (2002)	L1: Italian L2: English	64	English /i ɪ e ɛ æ u o ʌ ɒ ʊ ɚ/	Reading: target words (non-words) in carrier phrase
Flege, Schirru, and MacKay (2003)	L1: Italian L2: English	72	English /eɪ/	Reading: target words in carrier phrase
Baker and Trofimovich (2005)	L1: Korean L2: English	20[56]	English /i u æ ɛ ʊ ɪ/; Korean /i u e ɛ ɨ/	Reading: individual words
Tsukada et al. (2005)	L1: Korean L2: English	108	English /i ɪ e ɛ æ ɑ ʌ/	Picture naming
Wang and van Heuven (2006)	L1: Mandarin or Dutch L2: English	40	Complete English vowel inventory	Reading: individual words
Levy and Law (2010)	L1: English L2: French	29	French /y œ u i a/	Repetition task
Oh et al. (2011)	L1: Japanese L2: English	32	English /i eɪ ɑ u ɪ ɛ ʌ ʊ /	Reading: individual words

Table 3.1: Overview of studies on L2 vowel acquisition (in chronological order). *Note*: Due to the substantial amount of studies examining the acquisition of nonnative vowels in bilingual speakers, the overview presented in this table includes a representative selection of studies.

Flege, Bohn, and Jang (1997) examined the effect of L2 experience on vowel production accuracy in a group of adult L2 English speakers with different L1 backgrounds and varying degrees of L2 experience.[57] The learners' L2 vowels were

56 A total of 40 subjects participated in this study, including a group of early bilinguals (n = 20) who started acquiring the L2 English in early childhood.

57 The division into experienced vs. inexperienced/relatively experienced speakers was based on the subjects' length of residence in the United States, that is, the longer the subjects had been living in the L2-speaking environment, the more experienced they were considered to be.

rated by three native English speakers according to their degree of accuracy in a forced-choice identification test. In addition, acoustic measurements of vowel duration and formant frequencies (f0, F1, F2) were conducted. In a third experiment, Flege, Bohn, and Jang (1997) tested the L2 learners' identification ability of the English /i/-/ɪ/ and /ɛ/-/æ/-continua. Overall, the findings of the vowel identification tests and the results obtained in the acoustic analyses showed that L2 experience was positively correlated with production and perception accuracy in the L2, that is, more experienced L2 learners were more likely to produce and perceive English vowels accurately compared to less experienced L2 learners. However, despite the observation that the experienced speaker group performed better in terms of vowel perception and production than the inexperienced speaker group, there were only very few experienced learners whose English vowel productions were completely native. According to Flege, Bohn, and Jang (1997, 467), attaining native production abilities in the L2 is likely to require more than approximately seven years of L2 learning experience, i.e., even the experienced English learners in Flege et al. might not have gained enough L2 experience to be able to produce L2 vowels within native norms.

The effect of experience on vowel production accuracy was also investigated by Levy and Law (2010), who tested L1 English learners of L2 French. The bilingual subjects were assigned to three different groups, depending on their amount of L2 experience (minimal, moderate, extensive).[58] Alongside investigating the effect of experience on L2 vowel production, Levy and Law (2010) aimed to find out if the consonantal context in which the vowels occurred had an effect on production accuracy. Production accuracy of the French vowels /y/ and /u/ was assessed based on native speaker judgements[59] and acoustic analyses of formant frequencies. Results showed that highly experienced L2 French learners were more accurate in their productions of French /y/ occurring in an alveolar context compared to French /u/ in the same consonantal context. However, a different trend was observed for the French target vowels occurring in a bilabial context: French /u/ was produced more accurately in all three speaker groups – regardless of L2 experience – while French /y/ – in a bilabial context – was produced inaccurately across the three groups, which shows an effect of consonantal context on L2 production accuracy. Based on these findings, Levy and Law (2010, 1304) conclude that "[t]he extensive language

58 The division into three sub-groups with minimal, moderate and extensive L2 experience was based on three factors, namely the amount of formal L2 education the subjects had received, the amount of L2 use in private and professional contexts, and the amount of L2 immersion, i.e., the time the subjects had spent in a French-speaking country.

59 Monolingual French judges (N = 9) were asked to rate vowel production accuracy by means of keywords to determine which vowel they heard in the input stimulus. In addition, the degree of perceived vowel accuracy was rated on a 9-point-Likert scale (ranging from 1 = *most foreign sounding* to 9 = *most French sounding*) by the same judges.

experience of the [highly experienced] group here suggests that reaching native-like production may not be feasible for certain L2 vowels", i.e., certain nonnative vowels might be more difficult or easier to acquire depending on the phonetic context in which they occur.

Piske et al. (2002) compared the production of L2 (Canadian) English vowels in early and late Italian-English bilinguals living in Canada. Three sub-groups were formed, depending on the speakers' amount of L1 Italian use, that is, *early-high* subjects who used their L1 on a regular basis, and *early-low* subjects who used their L1 less frequently. In the late bilingual group, all subjects made frequent use of their L1. To assess vowel production accuracy, monolingual English listeners were asked to rate the subjects' vowels on a 5-point scale, ranging from 1 = *very strong foreign accent* to 5 = *no foreign accent*. The rating results showed that, overall, the English vowels produced by the late bilinguals were perceived to be less accurate than the early subjects' vowels. In addition, the findings support the hypothesis that the amount of L1 use (i.e., high vs. low) significantly influences production accuracy of L2 vowels (see also Flege, Bohn, and Jang 1997): While the early-low subjects were not perceived to differ from monolingual English vowel production norms, some of the vowels produced by the early-high subjects received higher accent ratings, i.e., their vowels were perceived to sound less native.

Flege, Schirru, and MacKay (2003) examined the same Italian-English bilinguals as Piske et al. (2002), focusing on their production L2 English /eɪ/, which differs from Italian /e/ in that English /eɪ/ is produced with more tongue movement compared to the Italian counterpart. The findings of an acoustic analysis of formant frequencies showed that the early-low bilinguals produced the English target vowel with more formant movement compared to monolingual English speakers, i.e., they 'overshot' the monolingual target and thus maintained contrast between L1 and L2 vowel categories, despite not being able to produce the target vowel in a native-like manner. The late bilinguals, by contrast, were observed to assimilate English /eɪ/ to Italian /e/ and thus failed to establish separate phonetic categories for the L1 and L2 vowels, respectively. As discussed in Section 2.2.2, category dissimilation is most frequently observed among early bilinguals, while assimilating phonetic categories is typical of L2 learners if they are not able to perceive phonetic differences between closely related (similar) L1 and L2 sounds.

Tsukada et al. (2005) assessed L2 English vowel perception and production by Korean-English adults and children. The bilinguals' productions of L2 vowels were rated for intelligibility by a monolingual English listener. [60] Compared to

60 The English listener was asked to assign each vowel produced by the bilingual subjects to one of the American English vowels, i.e., the bilingual vowels were situated within the American English vowel inventory. In this case, *intelligibility* was defined as "how often the target vowels were heard as intended" (Tsukada et al. 2005, 287).

monolingual adult speakers, the adult bilinguals received significantly lower intelligibility ratings, while no significant differences could be observed between the monolingual and bilingual children in terms of vowel intelligibility. A comparison between the two bilingual groups revealed that the vowels produced by the bilingual children were rated as being more intelligible than the same vowels produced by the adult speakers. These findings were confirmed in an acoustic analysis: Bilingual children were not found to differ significantly from age-matched monolingual control speakers, while the adult speakers' vowels differed from native (monolingual) vowel targets.

Baker and Trofimovich (2005) examined L1 and L2 vowels produced by early and late Korean-English bilinguals, focusing on potential interactions between the native and nonnative vowel systems. Differences in L2 experience within and between the two bilingual groups were defined based on the length of residence in the United States, ranging from one to seven years in the late bilingual group, and from one to eight years in the early bilingual group. By means of acoustic analyses of f0, F1 and F2, the bilinguals' vowel productions in both languages were compared (1) to each other (bilingual comparison), and (2) to monolingual productions of English and Korean, respectively (monolingual comparison). In terms of the bilingual comparison, results indicate that late bilinguals were able to keep those English and Korean vowels apart which are dissimilar across the two languages, that is, Korean /e/ and /ɛ/, and English /æ/ and /ɛ/. By contrast, the speakers were observed to produce vowels which exist in both vowel inventories with similar acoustic properties, i.e., English /i/, /ɪ/, and /u/, /ʊ/ were found to overlap with Korean /i/ and /u/, respectively. Early bilinguals, on the other hand, had successfully established separate phonetic categories for all English and Korean vowels, regardless of whether they were similar or dissimilar. The monolingual comparison showed that the late bilinguals' vowel productions were predominantly determined by L1-influnces on the L2, while the early bilinguals' productions were observed to be affected by bidirectional influences, that is, their "renditions of at least some L2 [...] and L1 [...] vowels were 'colored' by the acoustic properties of their L1 and L2 vowels, respectively" (Baker and Trofimovich 2005, 19).

The findings outlined above are in line with the Speech Learning Model's hypothesis that dissimilarity between L1 and L2 sounds facilitates the formation of separate phonetic categories, while L1–L2 similarity is likely to impede category formation (e.g., Flege 1995, 2003). In addition, the results of the bilingual comparison in Baker and Trofimovich (2005) support the Speech Learning Model's main tenet that late L2 learners are less likely to identify differences between closely related L1 and L2 sounds and assimilate L2 to L1 categories, which results in inaccurate (nonnative) productions.

Overall, the findings of the studies discussed in this section show that accurate L2 vowel production is determined by several factors, namely age of L2 learning (early vs. late learners), amount of L2 experience and – to some extent – amount of L1 use (Piske et al. 2002), and degree of perceived cross-linguistic (dis-)similarity between L1 and L2 vowel targets. In terms of age of L2 learning, *late* learners have been shown to be less likely to produce L2 vowels with native accuracy compared to early learners, particularly when it comes to vowels which are similar in the L1 and L2 (Baker and Trofimovich 2005).

3.1.4 L1 attrition of vowels

The majority of studies investigating attrition of L1 segmental speech production have focused on examining changes in consonant production (e.g., de Leeuw, Mennen, and Scobbie 2012a; Major 1992; Stoehr et al. 2017; Ulbrich and Ordin 2014) while comparatively few attrition studies investigate potential effects of L1 attrition processes on the production of vowels. Table 3.2 includes all studies which have been conducted so far to examine L1 vowel production in long-term immersed late bilinguals.

Study	**Languages**	**N of subjects**	**Vowels tested**	**Elicitation method**
Mayr, Price, and Mennen (2012)	L1: Dutch L2: English	2	Dutch and English vowel space	Reading: target words in carrier phrase
Bergmann et al. (2016)	L1: German L2: English	20	German /a: ɛ ɔ/	Film retelling (semi-spontaneous)
De Leeuw (2019)	L1: German L2: English	1	German /i a/	(semi-)spontaneous (pre-recorded TV and radio interviews)

Table 3.2: Overview of attrition studies on L1 vowel production[61] (in chronological order).

61 Another study which is often discussed in the context of phonetic attrition of L1 vowels was conducted by Chang (2012; see also 2013), who examined L1 (American) English speakers learning L2 Korean in a six-week intensive language instruction course in Korea. As mentioned in Section 2.2.3, L1 modifications experienced by the subjects in Chang (2012) are not referred to as *attrition* effects the context of the present investigation because they have not been long-term immersed in an L2 environment, but instead learned the L2 through intensive and short-term formal instruction. Chang's (2012) findings show a *phonetic drift* of L1 English vowels towards the norms of L2 Korean vowels, which suggests that cross-linguistic effects set in already at an early stage of L2 learning and are therefore not restricted to highly experienced, proficient bilinguals. Similar observations were made by Kartushina, Frauenfelder, and Golestani (2016) who investigated L1 vowel productions in a group of native French speakers who were trained in L2 Dutch or L2 Russian vowel production in several intensive training sessions. Despite the fact that none of the subjects had prior experience with the

Mayr, Price, and Mennen (2012) conducted acoustic analyses of plosives (see Chapter 2) and vowels produced by monozygotic twin sisters with L1 Dutch and L2 English. Both twins were late consecutive bilinguals, but one sister had been immersed in an L2 environment for an extended period of time, while the other sister was living in the Netherlands. Based on formant frequency analyses, the twins' vowel spaces in their L1 and L2 were compared: The twin who stayed in the L1 speaking environment produced vowels close to native Dutch norms while the L2-immersed twin produced L1 monophthongs significantly lower (more open) compared to her sister, i.e., her vowels were characterized by higher F1 values and thus had moved closer to the L2 English counterparts. Interestingly, significantly higher F1 values were also observed in her Dutch vowels /y/ and /ʏ/, which lack a counterpart in the English inventory. In addition, she failed to produce a clear contrast between some of her vowels (e.g., Dutch and English /ɛ i ɪ ɔ u/), presumably as a result of assimilating closely related L1 and L2 vowel categories. However, not all of her vowels were influenced by assimilation processes as she managed to maintain contrast, between, for example, Dutch /a/ and English /ɑ/. This suggests that – even within the same sound category – not all sounds are equally affected by attrition.

In another study, Bergmann et al. (2016) examined changes in the L1 German vowels /a:, ɛ, ɔ/ produced by a native and a nonnative group of German-English bilinguals. The division into 'native' versus 'nonnative' attriters was based on the results of a preceding foreign accent rating study, i.e., those speakers who had received an overall low foreign accent rating were labelled as 'native' attriters, and speakers with a high foreign accent rating were assigned to the 'nonnative' attriter group. Interestingly, an acoustic analysis of L1 German vowel formants showed a significantly lower F2 in the native attriter group compared to both nonnative attriters and monolingual German controls. In terms of F2, the native attriters produced German /a:/ with lower formant values compared to monolingual speakers, i.e., the German vowel target had moved in the direction of the most similar English vowel target. Although a slight tendency towards lowering F2 could also be observed in the nonnative attriter group, this was not shown to be significant. Furthermore, no significant formant differences between both attriter groups and the monolingual controls were detected in terms of the subjects' productions of the German vowels /ɛ/ and /ɔ/. Despite the fact that Bergmann et al.'s (2016) overall findings

target L2, Kartushina, Frauenfelder, and Golestani (2016) found evidence for a phonetic drift of native vowels towards the L2 vowel targets. The results provided by Chang (2012, 2013) and Kartushina, Frauenfelder, and Golestani (2016) show that even short-term experience with an L2 can affect – at least to some degree – productions in the native language. The changes observed in the speakers' L1 systems, were, however, likely to be short-term temporary changes, resulting from what Chang (2013) refers to as a *novelty effect*, and were presumably reversed – or partially reversed (Chang 2019) – when the subjects returned to their L1 country and/or did not receive intensive L2 input any longer.

did not reveal significant differences between native and nonnative attriters in terms of vowel productions, differences were detected between the two attriter groups and the monolingual control group, particularly in terms of their productions of German /a/, which leads to the conclusion that "the sustained presence of an L2 in the phonological system of a speaker can lead to changes in the articulation of L1 speech sounds" (Bergmann et al. 2016, 84). In addition, the observation that only some of the bilinguals' L1 vowels shifted closer to L2 norms are in line with Mayr, Price, and Mennen's (2012) findings, suggesting that attrition effects operate selectively rather than affecting a speaker's entire L1 pronunciation system.

The most recent study examining potential L2 influences on L1 vowel production was conducted by de Leeuw (2019), who investigated the native German speech of Stefanie Graf, a former professional tennis player. Similar to AS, Graf can be described as a late consecutive bilingual who had permanently moved to an L2 English-speaking environment in adulthood. De Leeuw (2019) conducted formant measurements of the German vowels /i/ and /a/ to determine changes in F1 and F2 over a course of four decades, that is, the time Graf had been living in an L2 environment at the time the study was conducted. In terms of the German high front vowel /i/, her findings show a significant increase in F2 frequency, which suggests a drift towards the norms of L2 English /i/ which is characterized by higher F2 values compared to native German. In both L1 target vowels, significant changes were also found for F1, namely a decrease in /i/ and an increase in /a/. However, these changes were not assumed to reflect a movement towards L2 norms but were interpreted as being a result of natural ageing effects, which have been shown to lead to a lowering of F1 in elderly speakers (see e.g., Reubold and Harrington 2015, 2017). As mentioned in Section 2.3.3, de Leeuw (2019) did not investigate longitudinal changes in Graf's L2 English pronunciation; thus, it is not possible to establish a direct link between Graf's L2 pronunciation development and the changes observed in her first language.

Although as few as three studies have been conducted so far which examined L1 vowel productions in late consecutive bilinguals, the findings of these studies provide evidence for L1 vowels being affected by processes of attrition, presumably due to L2 learning experience. However, as stated above, vowel categories do not seem to be equally affected by attrition, which Mayr, Price, and Mennen (2012, 698) attribute to "a complex interaction between token-level and system-level shifts in bilingual vowel systems".

3.2. The study

3.2.1 Aims and objectives

As discussed in Section 3.1, a considerable amount of research has been conducted to examine L2 vowel productions in late bilinguals and L2 learners (e.g., Flege, Bohn, and Jang 1997; Flege, Schirru, and MacKay 2003; Levy and Law 2010; Piske et al. 2002) while only few studies set out to investigate vowel productions in the L1 of L2-immersed late bilinguals (Bergmann et al. 2016; de Leeuw 2019; Mayr, Price, and Mennen 2012). In terms of L2 vowel production, research shows that the successful realization of L2 vowels is influenced by several factors, including age of L2 acquisition, amount of L2 experience and use, and the degree of cross-linguistic similarity between related L1 and L2 vowel categories. With regard to this, it has been shown that late L2 learners often fail to distinguish similar L1 and L2 vowels and, consequently, tend to assimilate L2 to L1 categories, which results in nonnative productions. The few attrition studies which have been conducted so far indicate that a speaker's L1 vowel system might indeed be affected by attrition, i.e., L1 vowels have been observed to move closer towards similar L2 vowel targets and are thus assimilated to L2 phonetic categories. However, it seems that not all L1 vowels are equally affected by attrition (see Bergmann et al. 2016; Mayr, Price, and Mennen 2012). As there is still a lack of research investigating L1 attrition of vowels, it is not possible to draw any reliable conclusions at this point about which vowel targets are more likely to be affected by attrition and why some vowels are more prone to change than others.

The present study aims to expand existing research on bilingual speech development by examining L1 and L2 vowel productions in an L2-immersed late bilingual over a period of approximately 40 years. To this end, AS's productions of monophthongs in his L1 Austrian German and his L2 American English was assessed by means of formant frequency measurements. The following research questions were addressed:

RQ 2.1: To what extent has AS's L2 production of monophthongs changed when comparing three stages in time (early, mid, and late)? Does he produce L2 vowels with formant frequencies (F1 and F2) which resemble those of monolingual American English speakers, i.e., has he gradually become more native-like since he moved to the United States?

RQ 2.2: To what extent has AS's L1 production of monophthongs changed when comparing two stages in time (early and late)? Does he produce L1 vowels with formant frequencies (F1 and F2) which more closely resemble those of monolingual American English speakers, i.e., have his L1 vowel categories moved

away from native production norms and at the same time moved closer to L2 production norms since he moved to the United States?

3.2.2 Materials and procedure

Modifications in AS's L1 and L2 vowel space were examined by conducting formant frequency analyses of F1 and F2, which are correlated with vowel height and vowel backness, respectively, and are thus most relevant when it comes to distinguishing vowel quality (e.g., Harrington and Cassidy 1999, 60). The present study is based on the same speech corpus as the previous investigation of AS's plosive productions, i.e., German and English broadcast interviews. Vowel data were taken from a total of 656 individual sound files representing AS's L1 and L2 pronunciation from the late 1970s to 2017/2018 (see Chapter 2, Table 2.5). As described in Chapter 2, the individual audio files were automatically segmented and annotated using the online tool WebMaus Basic provided by the BAS WebServices (Kisler, Reichel, and Schiel 2017; Schiel 1999).

	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/	–[62]	*bird*			

Table 3.3: Target vowels with corresponding sample words for American English (see Ladefoged 2005, 46) and Austrian German (see Moosmüller, Schmid, and Brandstätter 2015, 342).

Table 3.3 shows the German and English target vowels included in the present investigation to represent AS's monophthongal vowel productions in his L1 and L2. Alongside the vowels which exist in both Austrian German and American English (i.e., /i, ɪ, ɛ, u, ʊ, ɔ, ɑ/), three additional English vowels were included: The STRUT-vowel /ʌ/, the TRAP-vowel /æ/, and the NURSE-vowel /ɝ/. For Austrian German, the

62 As stated in Section 3.1.2, Austrian German – unlike other German varieties – does not contain a mid central vowel (see Moosmüller 2007, 52).

front vowel /e/ was additionally included. The reason to include these vowels despite the fact that they are not shared by the two vowel inventories is that they frequently occur in the respective vowel systems and were therefore considered important to display AS's L1 and L2 vowel space. The German vowel inventory additionally includes the front vowels /y/, /ʏ/, /ø/, and /œ/, which were, however, excluded from the present analysis due to the observation that they rarely occur in spontaneous speech (Moosmüller 2007, 52).[63] Although it has been shown that native Californian English speakers are not likely to distinguish between the vowels /ɔ/ and /ɑ/ as a result of the Californian Vowel Shift (see Section 3.1.1), they were still included as separate vowels in the present investigation in order to identify if AS produces different vowel targets.

In the individual recordings, monosyllabic and bisyllabic tokens (function words with empathic stress and content words) which contained a stressed vowel (in bisyllabic words: stress on the first syllable) were identified by hand using PRAAT (Boersma and Weenink 2018). As the present investigation is based on a corpus of pre-recorded speech samples, it was not possible to collect all vowel data in a neutral consonantal phonetic context.[64] In order to control for effects of co-articulation on the spectral properties of the vowels examined, the preceding and following consonants were included as random factors in the statistical analysis (see Section 3.2.3). Tokens with preceding or following approximants (/w/, /j/, /l/, /r/) were entirely excluded from the analysis. These sounds are commonly categorized as semivowels (e.g., Lawson et al. 2011; Levi 2011), i.e., they share characteristics of both consonants and vowels. Therefore, as pointed out by Di Paolo, Yaeger-Dror, and Beckford Wassnik (2011, 89), selecting vowels following or preceding an approximant might make it difficult to choose exact measurement points for the spectral features of the vowel.

After identifying words containing the vowel sounds relevant for this study (see Table 3.3), vowel onset and offset were delineated in PRAAT using a combination of

63 In her investigation of SAG monophthongs produced by speakers from Vienna, Moosmüller (2007) found that out of the total number of vowel tokens analyzed in spontaneously produced speech samples, only 1.72% were /y/ and /ʏ/ tokens, while /ø/ and /œ/ occurred even less frequently with 0.38%. Given that these percentages relate to a specific speech corpus, it can, of course, not be argued that German /y/, /ʏ/, /ø/, and /œ/ occur generally less frequently in Austrian German spontaneous speech compared to other vowels. However, an inspection of the vowel data used for the present investigation also suggests that these vowels occur relatively infrequently, which is why they were not included for analysis.

64 According to Di Paolo, Yaeger-Dror, and Beckford Wassnik (2011, 88), "neutral context means that all the moving articulators are at rest as they are for /h/, which does not require any particular tongue body shape, lip protrusion, or constriction in the supralaryngeal cavity." Therefore, acoustic studies typically elicit vowel tokens in a neutral /h_d/ consonant frame (among many others, e.g., Chang 2012; Strange 2007; Wang and van Heuven 2006). Alternatively, vowels might be elicited in a plosive environment, as in Mayr, Price, and Mennen (2012) and Flege, Bohn, and Jang (1997).

both oscillogram and wideband spectrogram displays (see Figure 3.3). For vowels following a plosive, fricative, or affricate, onset was marked at the occurrence of the first glottal striation where the formants became visible on the spectrogram. Vowel offset was marked at the point where formants were no longer visible. For vowels occurring in a nasal context, onset and offset were set at the point where a rise (onset) or drop (offset) of acoustic energy[65] could be identified, which coincides with the vowel formants being visible on the spectrogram. Tokens which were affected by noise or produced with creaky voice were excluded. Also, vowel segments were not included if they were too short in duration so that it was not possible to reliably calculate formant frequencies (see e.g., Harrington, Palethorpe, and Watson 2007, 2754).

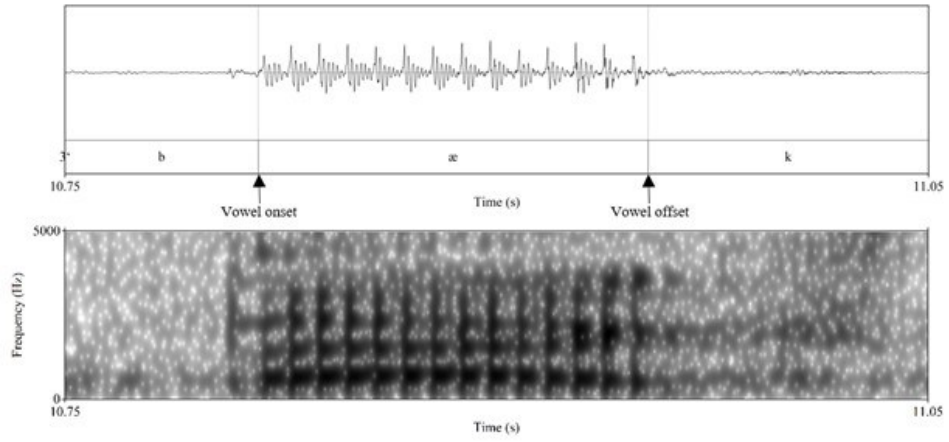


Figure 3.3: Acoustic landmarks (vowel onset and vowel offset) for vowel labelling on the waveform (top) with corresponding spectrogram (bottom). This example shows AS's late production of the English vowel /æ/ in the target word *back*.

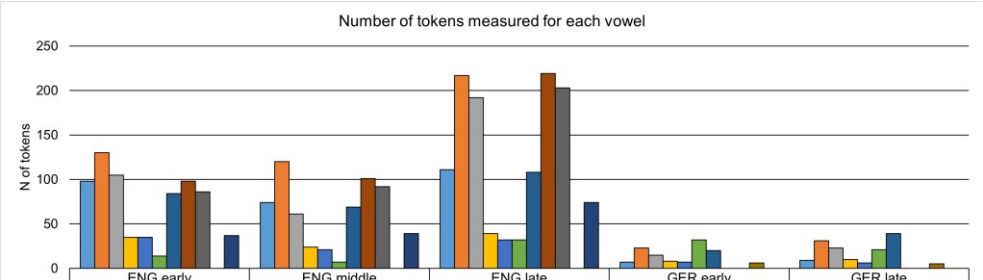
Formant frequencies were measured in PRAAT (Boersma and Weenink 2018) with a maximum formant frequency of 5000 Hz for an adult male speaker, a window length of 0.025 seconds and a pre-emphasis of 50 Hz. Based on linear predictive coding (LPC), formant frequencies were extracted using Burg's algorithm (Childers 1978; Markel and Gray 1976). Measurements of F1 and F2 were taken at the temporal mid-point of the vowel segment, i.e., halfway between onset and offset of the vowel. While it is sometimes argued that taking measurements at multiple points of the

65 Nasal consonants have a first formant with a considerably lower frequency compared to vowel sounds (see Ladefoged 2005, 55).

vowel (onset, mid-point, and offset) would allow for a more precise picture of vowel characteristics (Di Paolo, Yaeger-Dror, and Beckford Wassnik 2011, 90; see also Moosmüller 2007, 24), mid-point measurements are still most frequently used in acoustic vowel studies (e.g., de Leeuw 2019; Harrington 2006; Harrington, Palethorpe, and Watson 2000b; Mayr, Price, and Mennen 2012). Since co-articulatory effects of preceding and following consonants are only minimal at the very center of the vowel, single-point measurements also seemed most appropriate for the present study, particularly because – as mentioned before – not all vowel tokens occurred in a neutral phonetic context.

In total, 2,819 vowels (see Figure 3.4) were labelled and included in the analysis. As discussed in Section 2.3.3, due to the fact that the acoustic analysis was based on pre-recorded speech material, it was not possible to identify an equal number of vowel tokens representing AS's English and German pronunciations, which resulted in a significantly lower number of vowels obtained for his early and late German pronunciation.

Measurements of vowel duration, which are frequently conducted in addition to formant frequency analyses (e.g., Flege, Schirru, and MacKay 2003; Mayr, Price, and Mennen 2012 Flege et al., 2003; R. Mayr et al., 2012; Mok, 2011; G. E. Oh et al., 2011), were not included in the present analysis. As discussed by Paolo et al. (2010, p. 98), vowel duration is influenced by various factors, such as consonantal context and syllable number. In addition, vowel quantity is substantially affected by interactive factors, including speaking rate (see also Mok 2011, 528), i.e., in natural conversations, speakers tend to speak at varying rates, which might change even within one sentence or utterance. Therefore, as pointed out by Di Paolo, Yaeger-Dror, and Beckford Wassnik (2011, 98), "the measurement of duration from small quantities of actual conversations may not be very profitable". In order to reduce potential influences from discourse factors, duration measurements should be based on vowel tokens elicited in wordlist or short text passage readings (ibid.). In the recordings used for the present analysis, AS speaks at different rates, sometimes faster, sometimes slower – as typical of spontaneously produced conversational speech. Although vowel segments which were too short in duration were excluded from the analysis because it was not possible to provide reliable formant measurements, including an additional acoustic analysis of vowel duration was not considered useful due to the fact that the present investigation is based on conversational speech samples.



	ENG early	ENG middle	ENG late	GER early	GER late
/i/	98	74	111	7	9
/ɪ/	130	120	217	23	31
/ɛ/	105	61	192	15	23
/u/	35	24	39	8	10
/ʊ/	35	21	32	7	6
/ɔ/	14	7	32	32	21
/ɑ/	84	69	108	20	39
/æ/	98	101	219		
/ʌ/	86	92	203		
/e/				6	5
/ɜ/	37	39	74		

Figure 3.4: Number of tokens (*N* = 2,819) measured for each vowel in AS's early, middle, and late English, and early and late German pronunciation.

3.2.3 Results

The aim of Study II was to identify and describe changes in AS's monophthongal vowel space in his L1, Austrian German, and his L2, American English, by conducting formant measurements of F1 and F2. Similar to Study I, AS's English vowels were analyzed across three stages (early, mid, and late), while his German vowel productions were assessed in the early and the late stage. F1 and F2 were measured at the vowel mid-point in a total of 2,819 monosyllabic and bisyllabic tokens containing a stressed vowel. This section provides an overview of the descriptive statistics[66] of AS's vowel productions in his early and late English and German, and presents the statistical models used for analysis. The comparison of AS's L1 and L2 vowels with monolingual American English and Austrian German norms is based on an informal descriptive analysis.

(I) Descriptive statistics: AS's English vowels

Figure 3.5 shows the mean formant frequencies (in Hz) of F1 and F2 in AS's English monophthongs across the three stages. The distributions of formant frequency measurements for F1 and F2 for the individual vowel tokens are displayed in Figures 3.6 and 3.7, respectively.

66 A table showing the F1 and F2 frequency means, medians and ranges for each vowel according to language and stage can be found in Appendix D.

A visual inspection of the individual English vowels shows the following: A lowering of F1 and F2 can be observed for English /ɔ/ in the mid (F1: *M* = 615.4, *SD* = 71.73; F2: *M* = 915.4, *SD* = 78.48) and in the late (F1: *M* = 614.1, *SD* = 78.2; F2: *M* =876.5, *SD* = 73.74) stage compared to the early stage (F1: *M* = 661.5, *SD* = 69.7; F2: *M* = 1042.2, *SD* = 209.91), which indicates that the target vowel has moved towards a more close and a more back position in the late stage. Similarly, AS's early productions of English /u/ are characterized by overall higher F2 values (*M* = 1003.6, *SD* = 267.0), showing a more front production of /u/ in the early stage compared to a more back production (i.e., lower F2 values) in the mid stage (*M* = 867.8, *SD* = 246.98). However, the high standard deviations for F2 across all three stages indicate that there is a considerable amount of variation in AS's realizations of /u/. In terms of F1, only minor differences are observable, i.e., English /u/ is produced with a similar height across all three stages. A shift in F2 towards a more back position can also be observed in AS's productions of English /ʌ/ and /ɛ/ in the mid (*M* = 1218.2, *SD* = 82.09, and *M* = 1595.3, *SD* = 142.56, respectively) and in the late (*M* = 1231.2, *SD* = 88.65, and *M* = 1607.9, *SD* = 115.03, respectively) stage compared to the early stage (*M* = 1277.9, *SD* = 104.57, and *M* = 1665.2, *SD* = 188.0, respectively).

In terms of the English front vowel /ɪ/, the most obvious changes can be observed in F1, with lower F1 values for late /ɪ/ (*M* = 400.3, *SD* = 60.58) compared to mid /ɪ/ (*M* = 436.4, SD = 77.18). This indicates that AS's English /ɪ/ shifted from a more open production in the mid stage to a more close production in the late stage. For the late front vowel /i/, lower F1 values (*M* = 306.0, *SD* = 41.32) and slightly higher F2 values (*M* = 2054.8, *SD* = 117.79) can be observed compared to early and mid /i/, which indicates a more close and a more front production of the target vowel in the late stage. Finally, changes in vowel height can be observed in AS's production of the English back vowel /ʊ/, which is characterized by lower F1 values in the late stage (*M* = 384.6, *SD* = 39.41) compared to the mid stage (*M* = 418.0, *SD* = 50.3).

Overall, no uniform pattern is observable when it comes to changes in AS's monophthongal vowel space in this L2 English: While some vowels (/u/, /ɔ/, /ʌ/, and /ɛ/) seem to have moved towards a more back production in the late and/or mid stage compared to the early stage, indicated by a lowering of F2, other vowels (/ʊ/, /ɪ/, /i/, and /ɝ/) are characterized by a shift towards a more close production in the late stage compared to the early and/or mid stage, which is the result of a lowering of F1 in AS's late productions. For the vowels /æ/ and /ɑ/, only minor changes in F1 and F2 are observable, i.e., they seem to be relatively stable across all three stages.

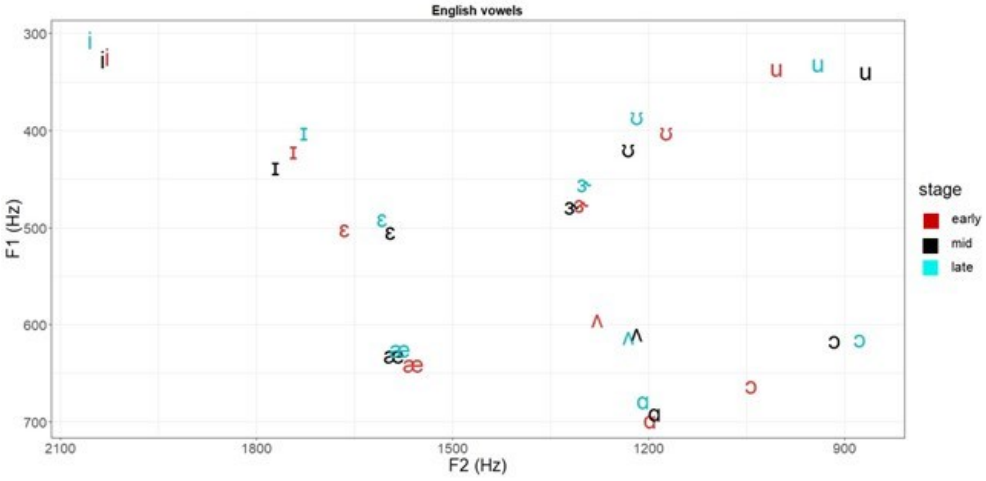


Figure 3.5: F1~F2 plot of the means (centroids) for AS's English monophthongs according to stage (early, mid, late).

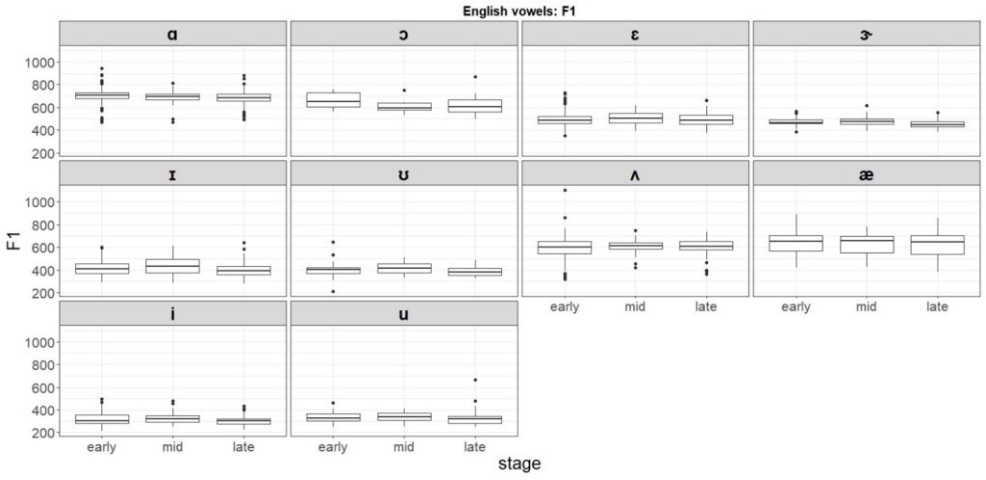


Figure 3.6: Boxplots of F1 (in Hz) in AS's English monophthongs according to stage (early, mid, late).

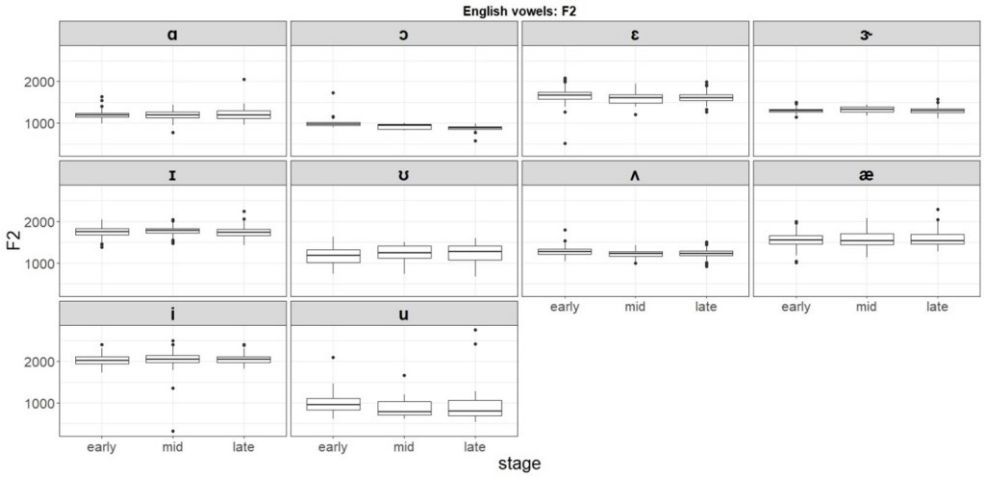


Figure 3.7: Boxplots of F2 (in Hz) in AS's English monophthongs according to stage (early, mid, late).

(II) Descriptive statistics: AS's German vowels

Figure 3.8 displays the mean formant frequencies (in Hz) of F1 and F2 in AS's German monophthongs in the early and the late stage. Figures 3.9 and 3.10 show the distributions of F1 and F2 for the individual vowel tokens across the two stages. An inspection of the figures suggests that most of AS's German vowels have changed considerably when comparing his early and late productions. The most striking changes are observable in terms of the German vowels /ʊ/ and /e/, which are characterized by an increase in F2 and a decrease in F1, respectively: F2 in his German /ʊ/ increased from *M* = 1048.4 (*SD* = 514.96) in the early stage to *M* = 1684.3 (*SD* = 129.35) in the late stage, which shows a shift towards a more front position in the late stage. The large standard deviation for early German /ʊ/, however, indicates that there is a considerably wider variation in his early compared to his late productions. While /ʊ/ has moved to a more front position in his late German, /e/ is characterized by a lowering of F1 in the late stage (*M* = 384.3, *SD* = 18.81), which shows a more close production. As can be seen in Figure 3.8, F1 for AS's late /e/ overlaps with F1 for his productions of early /ɪ/, and only minor differences in F2 can be observed between these two vowels.

His productions of German /u/ in the late stage are characterized by a considerable lowering of F1 compared to early /u/ (from early *M* = 391.3, *SD* = 27.85 to late *M* = 308.7, *SD* = 27.44), which shows a more close production of German /u/ in the late stage. At the same time, a shift towards a more front production can be seen in late German /u/, which is demonstrated by an increase in F2 (from early *M* = 1242.7,

SD = 366.13 to late *M* = 1326.7, *SD* = 227.47). However, as indicated by the large standard deviations for both his early and late productions of German /u/, his productions show a considerable amount of variability.

In AS's German /i/ a more close production, as shown by lower F1 values, can be observed in the late stage compared to the early stage (from early *M* = 360.8, *SD* = 34.69 to late *M* = 303.8, *SD* = 16.14), while only minor changes can be identified in terms of F2. A shift towards a more close position is also evident for late German /ɑ/ (from early F1 *M* = 701.6, *SD* = 76.09 to late F1 *M* = 617.1, *SD* = 76.84). Minor changes can be observed in terms of AS's productions of the German vowels /ɪ/ and /ɛ/, which have only slightly shifted towards a more close and a more back position in the late stage compared to the early stage.

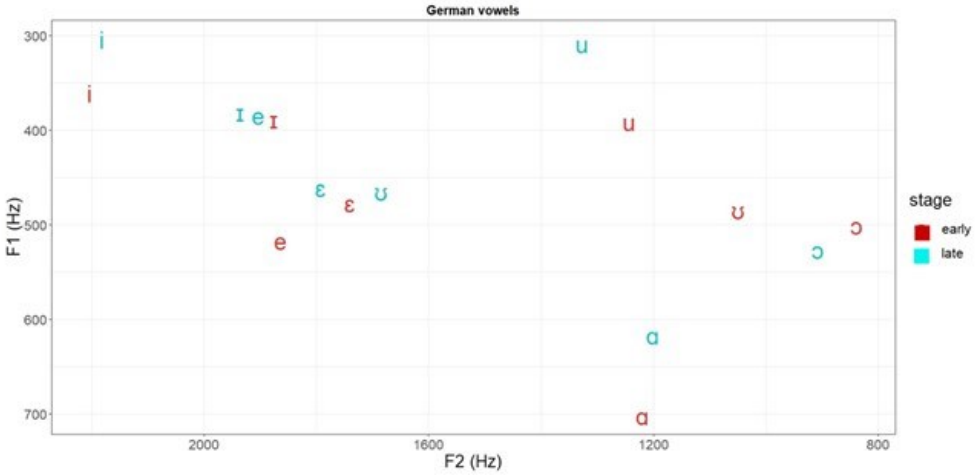


Figure 3.8: F1~F2 plot of the means (centroids) for AS's Austrian German monophthongs according to stage (early, late).

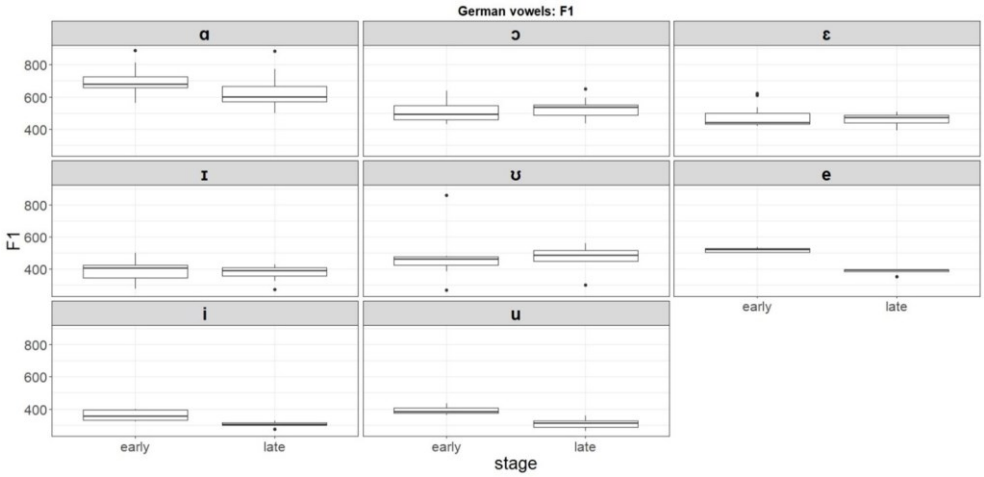


Figure 3.9: : Boxplots of F1 (in Hz) in AS's German monophthongs according to stage (early, late).

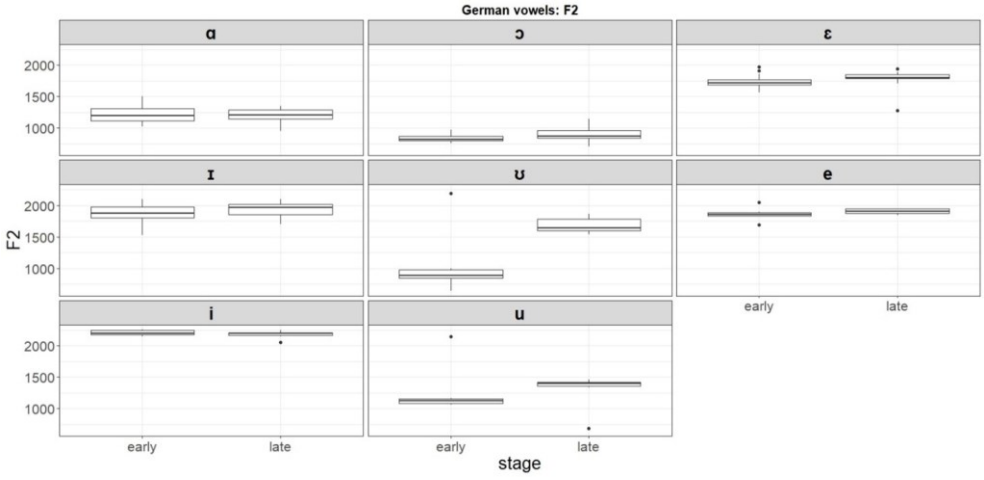


Figure 3.10: Boxplots of F2 (in Hz) in AS's German monophthongs according to stage (early, late).

(III) Descriptive statistics: Cross-linguistic comparison between AS's English and German vowels

Figure 3.11 shows a direct comparison of AS's monophthongal vowel space in L2 English and L1 German across two stages (early and late). The most obvious

changes can be identified in AS's productions of the vowels /ɑ/ and /ʊ/: While his English and German /ɑ/ in the early stage are almost identical in terms of both F1 and F2, a shift of German /ɑ/ towards English /ʌ/ can be observed in the late stage, which is reflected by a lowering of F1. Similarly, German /u/ and English /ʊ/ are very similar in the early stage, i.e., he does not seem to distinguish between the L1 and the L2 sound in terms of vowel quality. In the late stage, however, a lowering of F1 can be observed in AS's German /u/, which shows that this vowel has moved away from English /ʊ/. A tendency for L1 and L2 vowel categories to move apart in the late stage can also be identified in the English and German vowels /ɪ/ and /ɛ/: While an increase in F2 is observable in AS's late German productions of both vowels, F2 has decreased in English /ɪ/ and /ɛ/ in the late stage. Thus, his late German vowels /ɪ/ and /ɛ/ are characterized by a more front production while the same target vowels in English have moved towards a more back production.

A shift towards the English counterpart can be observed when comparing AS's productions of German /i/ in the early and in the late stage: While his early German /i/ is produced more open (i.e., higher F1 values) compared to English /i/, the German target vowel is characterized by a lowering of F1 in the late stage and has thus moved closer towards the English target, which results in almost identical mean values for F1 in both languages (German late F1 M = 303.8, SD = 16.4 and English late F1 M = 306.8, SD = 41.32). In terms of the vowel /ɔ/, an approximation of the English and the German targets can be observed, which manifests itself in a lowering of F1 and F2 in English /ɔ/, and an increase in F1 and F2 in the German target.

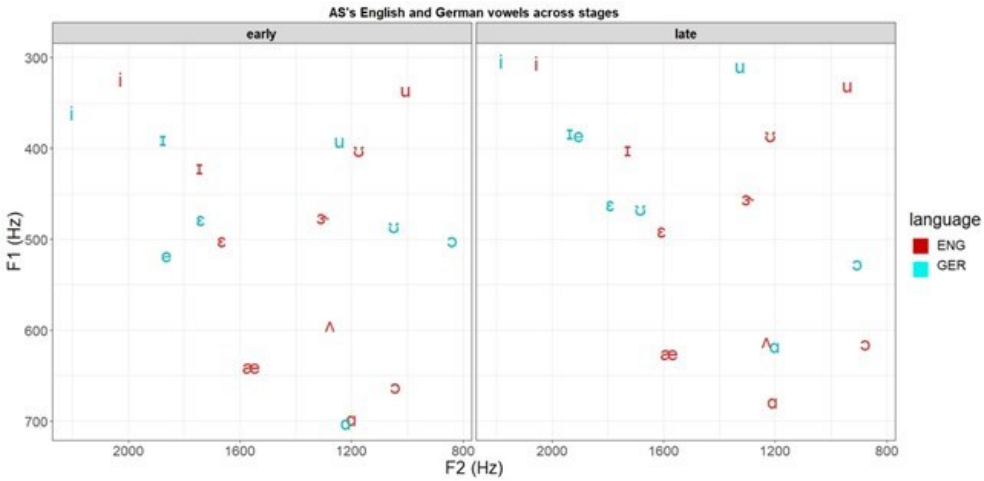


Figure 3.11: F1~F2 plot of AS's English and Austrian German monophthongs according to stage (early, mid, late).

(IV) Statistical analysis

Similar to Study I, a series of linear mixed-effects regression analyses were run in R (R Core Team, 2020, version 3.6.3), using the *lmerTest* package (Kuznetsova, Brockhoff, and Christensen 2017). The linear mixed models were fitted by REML *t*-tests using Satterthwaite's approximations to estimate degrees of freedom. *Post hoc* testing using the *emmeans* package (Lenth 2020) was used to identify pairwise effects. An alpha level of .05 was adopted throughout testing.

The first two regression models were conducted to examine changes in AS's L2 English monophthongs across three stages (RQ 2.1). In the first model, *F1* was included as the dependent variable and *stage*, *phoneme*, and an interaction between the two as fixed effects. As random effects, by-*word* random intercepts and by-*stage* random slopes were entered (F1~stage*phoneme + (stage|word)). In the second model, which also addressed Research Question 2.1, *F2* frequency was included as the dependent variable, with the same fixed and random factors specified above. The second regression analyses aimed to examine changes in AS's L1 German monophthongs across two stages (RQ 2.2). The same model specifications outlined above (with *F1* and *F2* as dependent variables, respectively) were applied, using German data only. The third linear model regression was conducted to compare AS's early and late productions of vowels in his two languages by including language (lang) as an additional fixed effect (F1/F2 ~stage*phoneme*lang + (stage+lang|word)).

AS's English vowels (F1, F2)

The analysis of F1 in AS's English vowels showed a significant effect of *stage* ($F[2,2380] = 8.5$, $p < .001$) and *phoneme* ($F[9,523.5] = 250.74$, $p < .0001$), and an interaction between *stage* and *phoneme* ($F[18,2399.2] = 1.76$, $p < .01$). Pairwise comparisons using *post hoc* Tukey tests revealed a significant difference in terms of F1 for /æ/ between the early and the late stage ($t(2358) = 2.43$, $p = .04$), i.e., AS produces English /æ/ with a slightly lower F1 frequency in the late stage compared to the early stage, as visible in Figure 3.5. A significant difference was also observed in AS's productions of English /ɪ/ in the mid and in the late stage ($t(2347) = 3.58$, $p = .001$), i.e., he produces /ɪ/ with a lower F1 frequency in the late stage compared to the mid stage. In addition, significant effects were found for his productions of English /ɔ/ in the early and in the late stage ($t(2324) = 2.55$, $p = .03$), indicating a tendency to produce the English target vowel with a slightly lower F1 frequency in the late stage than in the early stage. For the remaining vowels, no significant differences in F1 were found between the three stages.

AS's German vowels (F1, F2)

The analysis of F1 of AS's German vowels revealed a significant effect of *stage* (F[1,245.89] = 23.64, p < .0001) and *phoneme* (F[7,127.43] = 70.15, p < .0001) as well as interaction between *stage* and *phoneme* (F[7,239.55] = 3.33, p = .0021). Pairwise comparisons using *post hoc* Tukey tests showed a significant difference in F1 for /ɑ/ in the early and in the late stage (t(246) = 3.77, p < .001), indicating a lowering of F1 in AS's late productions of the target vowel as visible in Figure 3.8. Similarly, a significant lowering of F1 in AS's late compared to his early productions was found for German /e/ (t(231) = 3.47, p < .001), as can be seen in Figure 3.8. In addition, a significant effect was observed for his production of German /u/ when comparing the two stages (t(225)= 2.72, p = .007), with an overall higher F1 frequency in the early stage compared to the late stage.

In the second regression analysis, F2 frequency in AS's German vowels across the two stages was examined. Again, the analysis revealed a significant effect of *stage* (F[1,242.08] = 25.97, p < .0001) and *phoneme* (F[7,130.28] = 214.67, p < .0001), and an interaction between *stage* and *phoneme* (F[7,235.91] = 8.29, p < .0001). Pairwise comparisons using *post hoc* Tukey tests indicated significant differences in terms of F2 in the vowel /ʊ/ between the early and the late stage (t(245)= -8.03, p < .0001), as shown by an increase of F2 in the late stage. No other comparisons were significant, which was rather unexpected (see Figure 3.8), but is very likely the result of a low number of observations representing AS's productions of L1 vowels.

The second linear regression, which included *F2* of AS's English vowels as a dependent variable, also showed a significant effect of *stage* (F[2,91.2] = 3.98, p = .022) and *phoneme* (F[9,486.06] = 264, p < .001), and an interaction between *stage* and *phoneme* (F[18,87.51] = 2.64, p = .0014). Pairwise comparisons using *post hoc* Tukey tests revealed a significant difference in terms of F2 for /i/ between the early and the late stage (t(102.8) = -2.78, p = .022) and the mid and the late stage (t(102.8) = -3.0, p = .009), i.e., AS produces English /i/ with an overall higher F2 frequencies in the late stage compared to the early and the mid stage. In addition, a significant effect was found for his productions of English /ɔ/ in the early and the late stage (t(85.2) = 4.36, p < .001), i.e., his productions of the target vowel are characterized by a significantly higher F2 in the early stage. In terms of AS's productions of English /ʌ/, significant differences in F2 were also observed between the early and the late stage (t(216.4) = 2.51, p = .034), indicating a decrease of F2 frequency in the late stage.

Cross-linguistic comparison between AS's English and German vowels (F1, F2)

The analysis of F1 frequency in AS's English and German vowels showed a significant effect of *stage* (F[1,277.05] = 20.33, p < .001) and *phoneme* (F[6,420.38] = 260.04, p < .001), as well as an interaction between *phoneme* and *language* (F[6,422.54] = 7.49, p < .001) and a three-way interaction between *stage*, *phoneme* and *language*

(F[6,260.26] = 2.37, p = .03). Pairwise comparisons showed a significant difference in F1 in AS's English and German production of /ɑ/ in the late stage (t(362.7) = 5.1, p < .001), with significantly higher F1 values for his late German /ɑ/ compared to his late English /ɑ/. Significant differences in F1 were also found for English and German /ɔ/ in the early stage (t(258.6) = 5.22, p < .001) and in the late stage (t(339)= 3.92, p = .028), showing significantly higher F1 values in English /ɔ/ compared to German /ɔ/ in both stages.

In the second analysis, F2 frequency in AS's English and German vowels across the two stages was examined. The analysis revealed a significant effect of stage (F[1,1310.64] = 17.23, p < .001) and phoneme (F[6,482] = 377.39, p < .001), as well as an interaction between phoneme and language (F[6,484.6] = 7.28, p < .001) and a three-way interaction between stage, phoneme and language (F[6,1315.95] = 9.42, p < .001). *Post hoc* Tukey tests revealed significant between-language effects for AS's production of the vowels /ɛ/ and /ɪ/ in the late stage (t(574.8)= -3.91, p = .027, and t(585.1)= -4.6, p = .0017, respectively), as shown by significantly lower F2 values measured for his English productions of the target vowels. In addition, significant effects were found for F2 in English and German /ʊ/ and /u/ in the late stage (t(686.7)= -6.99, p < .001, and t(574.9)= -5.55, p < .001, respectively), with significantly higher F2 values measured for his late German /ʊ/ and /u/.

Comparison with monolingual American English and Austrian German vowels

In order to determine if AS's L2 vowels have become more native-like since he moved to an L2-speaking environment, his early and late vowels were compared to monolingual Californian English vowels reported in Hagiwara (1997). In addition, his L1 vowels were compared to monolingual Californian English vowels and monolingual Austrian German vowels (see Figure 3.12) in order to determine if his L1 vowel space has moved closer to L2 production norms and, at the same time, moved away from L1 norms. As described in Section 2.3.2, the native Austrian German reference data were collected from an age-matched monolingual Austrian German speaker living in Thal, AS's place of birth.

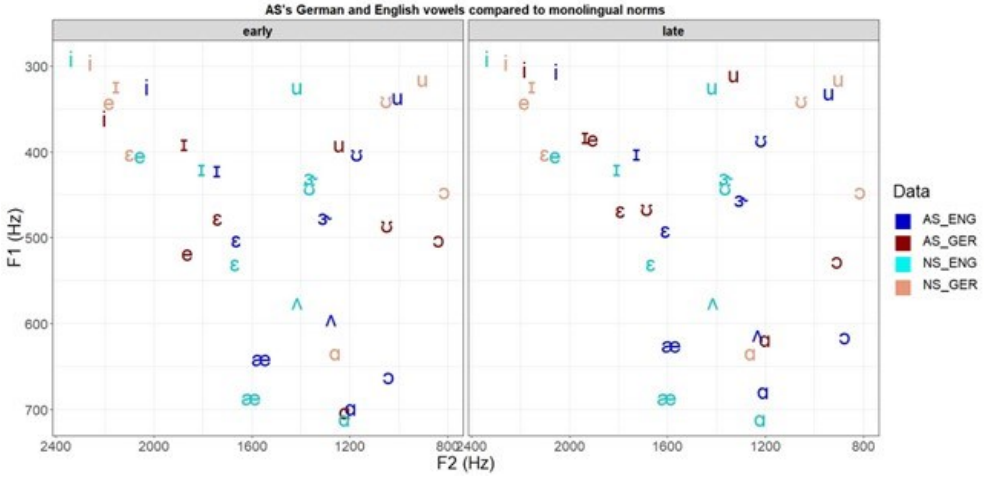


Figure 3.12: Comparison of AS's English and German vowels (AS_ENG, AS_GER) with monolingual Californian English vowels (NS_ENG) reported in Hagiwara (1997) and monolingual Austrian German vowels (NS_GER) across two stages (early, late).

As can be seen in Figure 3.12, AS's German /i/ has moved closer to native German production norms in the late stage, which is indicated by a lowering of F1 from early /i/ to late /i/. In his L2 English, AS produces the front vowel /i/ with a considerably lower F2 and a higher F1 compared to monolingual Californian English productions in both stages, i.e., his L2 target vowel is characterized by a more open and a more back production. Similar observations can be made for the English vowels /u/ and /ʌ/, which show much lower F2 values and slightly higher F1 values compared to native speaker F1 and F2 of the same vowels. At the same time, it can be seen that AS's productions of German /u/ have moved closer to L2 English norms in the late stage, indicated by a lowering of F1. Interestingly, his L2 English /u/ seems to have shifted towards native German norms in the late stage, as can be seen in a lowering of F2. When comparing AS's productions of the English target vowel /ʌ/ with native English speaker norms, it can be observed that his English /ʌ/ has moved away from native production norms in the late stage, which is shown by higher F1 values in AS's late /ʌ/ compared to lower F1 values reported for native English /ʌ/. Interesting observations can also be made in terms of AS's productions of /ɑ/: While his productions of this vowel in both English and German closely correspond to native English speaker productions in the early stage, his German /ɑ/ seems to have moved towards his English /ʌ/ in the late stage.

In terms of the back vowel /ɔ/, a comparison of AS's English and German productions with native German /ɔ/ shows a shift of AS's German /ɔ/ away from native

Austrian German production norms and closer to his English production of the target vowel in the late stage. Note that Hagiwara (1997) does not include formant frequency measurements for /ɔ/ because this vowel has merged with /ɑ/ in Californian English. However, as can be seen in Figure 3.12, AS distinguishes between English /ɔ/ and /ɑ/, with /ɑ/ being characterized by an overall higher F2 frequency compared to English /ɔ/. When it comes to AS's productions of the vowel /ɛ/, it can be observed that his early and late German productions of this vowel are characterized by lower F2 values and higher F1 values compared to native Austrian German norms, which indicates a more back and a more open production in AS's German.

AS's productions of the German front vowels /ɪ/ and /e/ are characterized by higher F1 values and lower F2 values compared to native German /ɪ/ and /e/ in both stages. In the late stage it can be seen that AS's productions of German /e/ have been raised to German /ɪ/. His productions of English /ɪ/, by contrast, closely correspond to native English speaker norms in the early stage while a lowering of F1 in the late stage indicates a more close production compared to native English norms. For the remaining vowels /æ/ and /ɝ/, which are not part of the German vowel inventory, AS's English productions in both stages are characterized by lower F1 values compared to native English production norms, i.e., he produces both vowels at a more front position compared to monolingual speakers.

3.2.4 Discussion of Study II

Study II set out to examine the monophthongal vowel space in the L1 (Austrian German) and L2 (American English) of Arnold Schwarzenegger, a late consecutive bilingual who started learning his L2 in early adulthood upon migrating to the United States. A speech corpus of broadcast interviews in English and German was used to determine if and to what extent AS's L1 and L2 monophthongs have changed over a period of approximately 40 years. His L1 and L2 vowels were examined by measuring formant frequencies of F1 and F2. In the following, the results of Study II will be discussed in the context of L2 acquisition of speech and L1 phonetic attrition. In addition, the findings will be related to potential biological ageing effects and sociophonetic factors.

L2 acquisition of vowels

The first research question (RQ 2.1) addressed in Study II aimed to find out to what extent AS's L2 production of monophthongs has changed when comparing three stages (early, mid, and late), i.e., whether he produces L2 vowels with formant frequencies (F1 and F2) which resemble those of monolingual American English speakers.

The analysis of AS's L2 vowel productions across three stages revealed significant changes in the front vowels /i/, /ɪ/ and /æ/, and in the back vowels /ɔ/ and /ʌ/. These changes are characterized by a lowering of F1 in the late stage for /ɪ/, /æ/, and /ɔ/, and by a lowering of F2 for /ɔ/ and /ʌ/. A significantly higher F2 frequency was observed for the late English vowel /i/. A comparison of AS's early and late L2 productions with monolingual Californian English vowels showed that only two vowels, namely /ɝ/ and /i/, have shifted closer to the L2 target in the late stage, indicated by a lowering of F1 in both vowels; late English /i/, however, is still produced more back compared to monolingual English /i/. His late productions of the remaining vowels (i.e., /ɛ/, /ʌ/, /æ/, /ɑ/, /u/, and /ɪ/) were observed to have moved further away from native production norms in the late stage. This shift towards less native-like productions manifests itself in a lowering of F1 in these vowels, and an additional lowering of F2 in /ɛ/, /ʌ/ and /u/. The tendency to produce less native-like vowels in the late stage is particularly evident in AS's productions of /ɑ/, which closely match the monolingual English target in the early stage but, as mentioned before, has moved towards a more close production (lowering of F1) in the late stage.

As the summary of the findings suggests, no uniform pattern can be identified when it comes to changes in AS's L2 vowel space. In the context of the Speech Learning Model (e.g., Flege 1995), which allows for predictions to be made concerning the nature of mutual interactions between the L1 and L2 sound systems, systematic differences were expected to occur between AS's productions of L2 vowels which have a *similar* counterpart in his L1 German, and those vowels which do not exist in the L1 vowel inventory (i.e., /æ/, /ʌ/, and /ɝ/). Flege, Bohn, and Jang (1997), for example, found that their L1 German late learners of L2 English failed to produce a distinction between English /ɛ/ and /æ/, presumably as a result of identifying English /æ/, which does not exist in the German vowel inventory, as an instance of German /ɛ/ or /ɛ:/. Flege and Hillenbrand (1984) showed that native English late learners of L2 French were successful in establishing a separate phonetic category for French /y/ as this vowel does not have a direct counterpart in English. By contrast, their productions of French /u/, which has a similar counterpart in English, were characterized by significantly higher F2 values compared to native French speakers, i.e., they were not able to identify phonetic differences between the L1 and the L2 sound. Also the L1 Korean L2 English subjects in Baker and Trofimovich (2005) were shown to keep those L1 and L2 vowels apart which were dissimilar in the two languages (Korean /e/ and /ɛ/, and English /æ/ and /ɛ/), while they failed to produce similar L1–L2 vowels authentically (English and Korean /i/ and /u/). Based on the findings of these studies, AS might have been predicted to move closer to native production norms from the early to the late stage in his productions of the vowels /æ/, /ʌ/, and /ɝ/, which lack a counterpart in the Austrian German vowel system and are hence less likely to be assimilated to a corresponding L1 target. However, a shift

closer towards native norms could only be observed for the mid-central vowel /ɜ˞/. The front vowel /æ/ and the back vowel /ʌ/, by contrast, have moved further away from native English norms, as indicated by a lowering of F1.

In addition, an overall improvement of AS's L2 vowels with increasing L2 experience in the late stage was expected as previous research provides evidence for a positive correlation between L2 learning experience and L2 production accuracy, i.e., production accuracy is likely to increase with increasing experience (e.g., Flege, Bohn, and Jang 1997; Levy and Law 2010). In the present investigation, however, the observed changes seem to go in the opposite direction, that is, most of AS's L2 vowels have moved even further away from native English production norms in the late stage, which, as mentioned above, is particularly evident in his production of English /ɑ/. Overall, the shift towards less native production abilities in the late stage manifests itself in a lowering of F1 in most of AS's late English vowels (with the exception of /ʌ/, which is characterized by a significant lowering of F2 and an increase of F1 in the late stage). The tendency to produce late L2 vowels with overall lower F1 frequencies might not be explained against the background of cross-linguistic (system-internal) influences, but might rather be considered a result of age-related changes. Ageing mechanisms, which lead to, for example, physiological changes in the human vocal tract, have been shown to cause a decrease in F1 frequency in vowels produced by elderly speakers (e.g., Reubold and Harrington 2015, 2017; Reubold, Harrington, and Kleber 2010).

L1 attrition of vowels

The second research question (RQ 2.2) addressed in Study II aimed to find out to what extent AS's productions of L1 monophthongs has changed when comparing two stages (early and late), i.e., whether he produces L1 vowels with formant frequencies (F1 and F2) resembling those of monolingual American English speakers and thus have become less native-like.

The analysis of AS's L1 vowel productions across two stages showed a significant decrease of F1 in his late productions of the German vowels /e/, /u/ and /ɑ/, and a significant increase of F2 in German /ʊ/ in the late stage. Interesting observations were made in terms of AS's productions of German /ɑ/. In the early stage, this vowel is nearly identical with the monolingual English target; in the late stage, however, the German target vowel has moved away from the monolingual English target, as shown by a significant decrease of F1, and has moved towards AS's production of English /ʌ/. That is, AS does not distinguish between the vowel categories for English /ʌ/ and German /ɑ/ in the late stage. At the same time, his productions of English /ʌ/ have shifted further away from monolingual English norms, in the direction of German /ɑ/. An approximation of L1 and L2 vowel categories was also observed in his early vs. late productions of /ɔ/, indicated by a significant decrease of F1 in

English, and an increase of F1 in his German. Also his German productions of the target vowel /u/ has changed significantly when comparing the two stages: While his early German /u/ is close to his English productions of /ʊ/, a significant decrease of F1 could be identified in his late German productions of /u/, which are now close to the native Californian target vowel /u/. At the same time, his productions of English /u/ have shifted towards monolingual Austrian German /u/ in the late stage.

The shift away from native (monolingual) production norms in AS's late German vowels /ɑ/, /ɔ/, and /u/, as previously outlined, might be indicative of category assimilation, that is, AS's L1 and L2 vowel categories come to resemble each other in the course of L2 learning experience (Flege 1995, 2003). It needs to be taken into consideration, however, that even in the early stage, none of his L1 vowels is identical with native Austrian German production norms. This might be a result of having already stayed in an L2- speaking environment for an extended period of time before the first speech samples used for the present investigation were recorded. In addition, not all of AS's German vowels have shifted towards L2 production norms, i.e., attrition is not evidenced uniformly across his vowel space. For instance, both his German vowels /ɛ/ and /ʊ/ have shifted towards a more front position in the late stage, further away from English production norms. The observation that not all vowels are equally affected by cross-linguistic interactions and attrition has also been made by Mayr, Price, and Mennen (2012)Mayr et al. (2012), who found evidence for cross-linguistic assimilation patterns in the L1 vowels of their L1 Dutch L2 German bilingual. However, despite an overall trend for a shift of L1 Dutch vowel categories in the direction of L2 English, not all categories were observed to be affected by cross-linguistic assimilation, such as the Dutch vowel /ɑ/ which was shown to be distinct from English /ɑ/. In the same vein, de Leeuw (2019) concluded from her investigation of Stefanie Graf's L1 German speech that L1 changes due to cross-linguistic interactions cannot necessarily be observed to affect a speaker's entire language system. Her analysis of segmental features revealed, for example, changes in Graf's productions of German /i/ and /l/ over time, while her productions of German /u/ and /r/ were not observed to have changed significantly when comparing her early and late pronunciations. Bergmann et al. (2016) further support the assumption that cross-linguistic interactions do not operate on a system-wide level: Their investigation of changes in the German vowels /a/, /ɛ/, and /ɔ/ in a group of German-English bilinguals showed that L1 German /a/ had moved towards L2 English /ɑ/, while no shift in the direction of the L2 could be observed for the vowels /ɛ/ and /ɔ/. Taken together, the findings outlined above and the findings of the present investigation seem to support the assumption that a bilingual's vowel system is characterized by "a complex interaction between token-level and system-level shifts" (Mayr, Price, and Mennen 2012, 698).

Similar to AS's English vowels, a general trend towards a decrease in F1 could also be observed in his L1 vowel productions, as discussed above. That is, whenever changes were observed in his L1 vowels, they manifested in a lowering of F1 in the late stage (with the exception of late German /ɔ/ which is characterized by an increase of F1), regardless of whether the L1 vowel moved closer to a corresponding L2 target or not. These changes in F1 in both his English and German productions might be the result of biological ageing processes, as will be discussed in the subsequent section.

Biological ageing effects

The extent to which biological ageing processes affect the production of vowels in a speaker's native language system over time has been examined in acoustic studies conducted by Reubold and Harrington (Reubold and Harrington 2015, 2017; see also Harrington 2006; Harrington, Palethorpe, and Watson 2000a, 2007; Reubold, Harrington, and Kleber 2010). These studies examined vowel changes based on longitudinal speech data of different public figures, such as Queen Elizabeth II (Reubold, Harrington, and Kleber 2010), the German newsreader Dagmar Berghoff (Reubold and Harrington 2017), and the British-American radio commentator Alistair Cooke (ibid.). Overall, findings suggest that age-related changes in vowel production mainly affect F1, that is, a decrease of F1 could be observed in older speakers.[67] As shown by Linville and Rens (2001), who compared resonance characteristics in young adults and older speakers, decreasing frequencies of F1 are likely to be the result of a lengthening of the vocal tract over a speaker's lifespan.

However, as discussed by Eichhorn et al. (2018), conclusions concerning the effect of age on vowel acoustics need to be drawn with caution since research so far has not provided entirely consistent results, which might be explained by methodological differences relating to the speech materials used and the speakers examined (Eichhorn et al. 2018, e5). In addition, most investigations (other than the studies mentioned above) into the relationship between age and vowel acoustic characteristics are based on cross-sectional comparisons of two or more groups of speakers differing in age (e.g., Debruyne and Decoster 1999; Fletcher et al. 2015; Linville and Fisher 1985; Rastatter and Jacques 1990) and, thus, do not trace ageing effects

67 In addition, the decrease in F1 in older speakers has been found to correlate with a decrease in f0 (see e.g., Reubold and Harrington 2015, 2017; Reubold, Harrington, and Kleber 2010). This correlation between F1 and f0 was observed to occur predominantly in high vowels where the frequencies of F1 and f0 are close together. In low vowels, by contrast, changes in F1 were "mainly characterized by increasing centralization at an older age" (Reubold and Harrington 2017, 129), i.e., longitudinal F1 changes were not observed to go along with changes in f0. As the present investigation did not include an acoustic analysis of fundamental frequency in AS's speech, the relationship between age-related changes in F1 and f0 will not be further discussed. However, examining this relationship based on the speech data used in the present study offers a valuable source for future research.

in the same individual or the same group of individuals over an extended period of time. In order to allow for a closer interpretation of the findings of the present study against the background of potential age-related effects, a longitudinal study conducted by Reubold and Harrington (2015), which examined age-related changes in Alistair Cooke's vowel space, was used for comparison due to two reasons: First, they examined vowel changes in a single (male) speaker[68] over a period of five decades and, second, the materials used for analysis are similar to the materials the present study is based on. Cooke's vowel productions were analyzed using broadcast recordings which covered a period of approximately 50 years (1951–2004), representing Cooke's speech from age 42 to age 95. The longitudinal changes identified in F1 (and in f0, which is not discussed here) as a function of ageing differed depending on vowel height. That is, in Cooke's phonetically high vowels F1 was found to have decreased up until his seventies, followed by an increase of F1 from his seventies until age 95. In his low vowels, an initial rise of F1 frequency was identified between the ages 42 and 61, followed by a decrease after age 61. Also in his mid-vowels, a slight decrease of F1 up until the mid-sixties was observed, which was, however, not found to be statistically significant. The outcomes of Reubold and Harrington's study are – to some extent – mirrored by the present findings. Although F1 frequencies were not averaged by the age of the speaker in the present study (as in Reubold and Harrington 2015, 19), an overall decrease of F1 could be observed when comparing AS's vowel productions in the early and in the late stage in both languages. The early stage, including recordings from 1979–1988, represents AS's pronunciation from age 32–41, and the late stage (2012–2018) represents his pronunciation from age 65–71. Consequently, he was in his late sixties/early seventies when the lowering of F1 was identified. That is, the late decrease of F1 in both his English and German vowels coincides with the decrease of F1 observed in Alistair Cooke's low vowels after the age of 61, and in his high vowels in his seventies.

It should be noted, however, that some of Alistair Cooke's vowels also seemed to have been affected by sociophonetic changes in the form of accent reversion, which means that elderly speakers might revert back to a particular accent or to accent-specific features which they used when they were younger (see Howell, Barry, and Vinson 2006, 140). Reubold and Harrington (2015) identified changes in Cooke's

68 Some previous research suggests that age-related effects on vowel production might differ depending on speaker sex (e.g., Eichhorn et al. 2018; Linville and Rens 2001; Torre and Barlow 2009). Linville and Rens (2001, 323), for example, found that the elderly female speakers in their study "demonstrate[d] more substantial lowering of spectral peaks with aging than men", which might suggest that the lengthening of the vocal tract as a consequence of biological ageing is more pronounced in women than in men. By contrast, other studies identified similar changes – predominantly in F1 – in both elderly male *and* female speakers (e.g., Debruyne and Decoster 1999; Harrington, Palethorpe, and Watson 2007) and thus do not provide evidence for an interaction with speaker sex.

BATH, LOT, THOUGHT and DRESS vowels in the period from 1950 to 1980 towards features of Received Pronunciation, a variety he had – presumably – acquired at a younger age before migrating to the United States. In AS's case it is difficult to determine what his target English looked like when he started learning the L2. As mentioned in Chapter 2, he had spent some time in the UK before migrating to California at the age of 21. Anecdotal evidence suggests that he had stayed in London for approximately 18 months in the mid-1960s (e.g., Preston 2015; The Times 2008) and, during this time, he regularly travelled between London, Austria and Germany to partake in bodybuilding competitions. Therefore, the time he was effectively exposed to Received Pronunciation or to another variety of British English was perhaps too short to reliably argue that this was the variety he acquired before moving to the United States where he adopted an American (Californian) English variety.

While AS's L1 and L2 vowels seem to be affected by an overall decrease of F1 in the late stage – perhaps as the result of biological ageing processes – no uniform patterns could be observed in terms of changes in F2 in AS's English and German productions. That is, while some of his late English vowels (/u/, /ɑ/, and /ʌ) were characterized by significantly lower F2 values in the late stage, his late /i/ was found to be produced with significantly higher F2 values. In his late German productions, by contrast, a statistically significant increase in F2 was observed for the vowel /ʊ/ only. Previous research into potential age-induced changes of F2 also present inconsistent results: While some studies report a significant decrease of F2 as a function of age (e.g., Debruyne and Decoster 1999), others found either no (e.g., (e.g., Mwangi et al. 2009; Sebastian et al. 2012) or inconsistent changes (e.g., Reubold, Harrington, and Kleber 2010; Schötz 2006).

To summarize, the findings of Study II suggest that not all phonetic features are prone to undergo attrition, i.e., attrition is not an all-or-nothing phenomenon, which is in line with previous findings (e.g., Bergmann et al. 2016; Mayr, Price, and Mennen 2012). In addition, some changes observed in a bilingual's linguistic systems might, in fact, not be directly related to system-internal interactions but are the result of extra-linguistic influences, such as biological ageing mechanisms.

4 Study III: Perceived nativeness, intelligibility, and comprehensibility of L1 speech

4.1. Introduction to Study III

As mentioned in Chapter 1, AS is often described as "one of the most recognized individuals on the planet" (USC Schwarzenegger Institute 2013) due to his outstanding career as an actor, politician, businessman, and athlete. However, not only his many talents and achievements brought him fame, but also his apparently unique way of speaking seems to be closely associated with the person(ality) of Arnold Schwarzenegger. Referring to his L1-accented English pronunciation, AS stated in an interview that when people hear him speak "they know it's Arnold [...] they don't have to see me [...] that makes me unique" (AS_E_1985_1). In German and Austrian media, it is also frequently discussed whether AS has 'unlearned' his first language (e.g., Quora 2017), which has led to many comic imitations of his apparently English-accented German pronunciation (e.g., Guten Morgen Österreich 2012a, 2012b; Maschek 2013; Stummer 2011). When Austrians or Germans hear the name *Arnold Schwarzenegger*, watch him on television or listen to him on the radio giving interviews in his native language, they might automatically expect to hear his English-accented German pronunciation.

Study III set out to answer the question if AS's L1 accent is perceived to sound nonnative by two groups of naïve (i.e., phonetically untrained) listeners who were not informed in advance that they were to listen to AS's speech. In the context of the present investigation, the phenomenon of foreign-accented speech is approached from two different perspectives: First, it is understood as a phenomenon that refers to divergences from native (monolingual) norms which can be quantified by means of acoustic-phonetic analyses – as has been done in Study I and Study II. Second, foreign accent is understood as a *perceptual* phenomenon on the part of listeners, i.e., whether and to what extent listeners perceive L1 speech to sound (non)native (see Thomson 2018), which is the focus of Study III. From this perspective, accentedness will be defined as "the degree to which the pronunciation of an utterance sounds different from an expected production pattern" (Munro, Derwing, and Morton 2006, 112). In order to assess to what extent AS's L1 speech is perceived to deviate from native pronunciation patterns, monolingual Austrian German and bilingual English-German listeners from Graz (Austria) were invited to rate a set of speech samples in a nativeness rating task.

Alongside investigating the dimension of perceived nativeness, the present study set out to examine if AS's L1 speech is perceived to be intelligible and comprehensible by the same listener groups described above. Intelligibility is defined as "the extent to which a speaker's message is actually understood" (Munro and

Derwing 1995a, 76) while the dimension of comprehensibility refers to a listener's perception of how difficult or easy a speaker's utterance is to understand, i.e., the "ease of understanding" (Saito, Trofimovich, and Isaacs 2016, 217). In the present study, intelligibility was tested by means of listener transcripts and perceived comprehensibility was assessed using a rating scale (see Section 4.2.2).

Before reviewing L1 nativeness/accent perception studies which have been previously conducted, a final remark concerning the terms 'perceived accentedness/nativeness', 'perceived intelligibility' and 'perceived comprehensibility' as used in this chapter will be made: According to Munro (2008), assessing accentedness and comprehensibility (as well as intelligibility) is only possible through direct listener responses, elicited, for example, by means of scale-ratings. Therefore, he argues, "there is no reason to use expressions like '*perceived* accentedness' or '*perceived* comprehensibility' because, in fact, there is no other kind of accentedness or comprehensibility" (Munro 2008, 200; my italics). While this argument holds true for the dimensions of comprehensibility and intelligibility, accentedness, as discussed above, can be approached from two different perspectives, namely from an analytical instrumental perspective (through acoustic-phonetic analyses) and from a perceptual, listener-oriented perspective. Therefore, the term 'perceived accentedness' will be used in this chapter to stress its perceptual dimension. It should be noted that the terms 'perceived accentedness' and 'perceived nativeness' are often used interchangeably in studies examining global foreign accent. However, perceived nativeness is not necessarily the same as perceived accentedness: Listeners who perceive a speaker as sounding nonnative might not base their judgments on perceived deviations from expected pronunciation patterns (i.e., accentedness), but other linguistic features, such as the incorrect use of grammatical structures or lexical items, might contribute to perceiving a speaker as nonnative. Therefore, a speaker's accent, though being a very salient feature when it comes to judging (non)nativeness, is not inevitably the only indicator of nativeness. In order to avoid that listeners make their judgments based on incorrect grammar or wrong word choice rather than on the speaker's pronunciation, empirical studies usually include speech samples which do not contain grammatical and/or lexical errors, as will be further discussed in Section 4.1.3. In the present investigation, the terms 'perceived nativeness' and 'perceived accentedness' will be used interchangeably, bearing in mind that both refer to the dimension of pronunciation here.

4.1.1 Perceived nativeness of L1 speech

As discussed in the previous chapters, there is only a limited number of acoustic studies which examine phonetic (e.g., de Leeuw 2019; Mayr, Price, and Mennen 2012; Stoehr et al. 2017) and phonological (e.g., de Leeuw, Tusha, and Schmid 2017;

Dmitrieva, Jongman, and Sereno 2010) attrition of L1 speech, and there are even fewer studies which assess perceived foreign accent in the L1 pronunciation of L2-immersed late consecutive bilinguals. Table 4.1 provides an overview of studies examining global foreign accent in the L1 of L2-immersed late consecutive bilinguals. As can be seen, all studies, except for Sancier and Fowler (1997), examined L1 German bilinguals being immersed in an L2 English or L2 Dutch environment.

Study	**Languages**	**N bilinguals**	**N monolingual controls**	**N raters**	**Elicitation method**	**Duration of speech materials** [69]
Sancier and Fowler (1997)[70]	L1: Brazilian-Portuguese L2: English	1	–	46	Verbal sentence translation	*not stated*
De Leeuw, Schmid, and Mennen (2010)	L1: German L2: English / Dutch	34 (L2 English) 23 (L2 Dutch)	5	19	Film retelling (semi-spontaneous)	12.6-17.7
Hopp and Schmid (2013)	L1: German L2: English / Dutch	40[71] (attriters) 40 (L2 learners)[72]	20	149	Narrative-descriptive tasks (semi-spontaneous)	10-20

Table 4.1: Overview of studies examining perceived global foreign accent in the L1 of late bilinguals (in chronological order).

De Leeuw, Schmid, and Mennen (2010) assessed global foreign accent in the L1 German of late consecutive bilinguals living in either Canada (L2 English) or the Netherlands (L2 Dutch). Monolingual German listeners rated speech samples produced by the bilingual speakers and monolingual German control speakers according to perceived nativeness in the speakers' German pronunciation. Results show that the foreign accent ratings (FARs) obtained for the L2-immersed native German speakers were significantly higher than the ratings obtained for the monolingual German

69 Per recording, in seconds.

70 Based on the definition of attrition applied in the present work (see Section 1.4.3), the female subject in Sancier and Fowler (1997) cannot be described as a potential attriter, because she was not consistently L2-immersed but changed linguistic environments on a regular basis. However, as the study by Sancier and Fowler is one of the few which do not exclusively focus on perceived L2 accent but also examine L1 accent, it is included in the present discussion.

71 Speakers were selected from the speaker group in de Leeuw, Schmid, and Mennen (2010).

72 Late L2 learners who live in an L1-speaking environment.

speakers, i.e., the bilinguals were more likely to be perceived as sounding nonnative in their L1 German compared to the control group. No differences were observed *between* the two groups of bilinguals (L2 Dutch, L2 English), which suggests that the degree of perceived accentedness was not influenced by the speakers' respective second language. In addition, de Leeuw, Schmid, and Mennen (2010) assessed the impact of several predictor variables on the outcome of the FARs, including age of arrival and length of residence in the L2 country as well as L1 contact.[73] L1 contact turned out to be the most significant predictor variable, that is, bilinguals with frequent L1 contact were less likely to be perceived as nonnative German speakers compared to those bilinguals who had less frequent contact with the L1. These findings suggest that frequent L1 contact might be beneficial when it comes to L1 maintenance.

Hopp and Schmid (2013) conducted a similar study, using some of the speakers included in de Leeuw, Schmid, and Mennen (2010). The study compared three groups of speakers in terms of perceived foreign accentedness: German-English bilinguals living in an L2 (English or Dutch) environment, L1 Dutch or English late learners of L2 German, and (quasi-)monolingual German speakers living in Germany, who served as control speakers. The late L2 German learners were compared to speakers who acquired the target language (German) from birth (L1 attriters and monolingual German controls), and both bilingual groups (late L2 learners and L1 attriters) were compared to the monolingual control group. In terms of the late L2 learners, the results obtained in the FARs showed that the majority of speakers (n = 27 out of 40) was perceived to have a foreign accent in their L2 German, while the remaining 13 speakers received low accent ratings, i.e., they were perceived to sound (near-)native in their L2. By contrast, the majority of attriters (n = 29 out of 40) did not perform differently from the German control speakers, i.e., they scored within the monolingual German range. The remaining L1 attriters (n = 11 out of 40) were perceived to sound clearly nonnative in their L1 German pronunciation. Based on these findings, Hopp and Schmid (2013, 383) conclude that late L2 acquisition does not necessarily impede native-like pronunciation in the L2, while acquiring the target language from birth "does not ensure sustained nativelikeness in speech production", as has been shown for some speakers in the L1 attriter group.

In the same vein, Bergmann et al. (2016) aimed to find out if a group of late consecutive bilinguals (L1 German, L2 English), being immersed in an L2-speaking environment, were perceived to sound foreign accented in their L1 German by a group of monolingual German listeners. Their hypothesis that the bilingual group would be perceived to sound less native compared to the monolingual group was

73 L1 contact was specified as the frequency of contact with L1 settings in which code-switching and/or code-mixing was likely (frequent L1 contact) or not likely (infrequent L1 contact) to occur.

confirmed in the FARs: Almost 40% of the bilingual speakers were rated below the native range, that is, they were perceived to sound nonnative in their L1 pronunciation. In addition, a larger amount of variation was found in the bilingual group compared to the monolingual speaker group. The effect of two variables, namely length of residence in the L2 environment and L1 use, was observed to account for this variation: The longer the bilinguals had been living in the L2 country, the more nonnative their L1 pronunciation was rated, while an increased L1 use led to lower FARs, i.e., speakers who used their L1 on a regular basis were more likely to be perceived as sounding native in the L1.

Another study which is often cited in the context of perceived L1 nativeness was conducted by Sancier and Fowler (1997). They examined whether the speech of a native Brazilian-Portuguese speaker with L2 English, who regularly travelled between the L1 (Brazil) and L2 (United States) country, was perceived to have a foreign accent in her L1 and L2, judged by monolingual listeners of the respective language. The findings of the FARs indicate that Brazilian-Portuguese listeners perceived the sentences produced right after a stay in the United States to sound more foreign-accented than those sentences produced after a stay in Brazil. By contrast, the majority of monolingual English listeners did not perceive a significant difference between the bilingual's L2 English pronunciation after staying in the United States and her pronunciation after staying in Brazil. In an acoustic analysis of the speakers' VOT, Sancier and Fowler (1997, 421) identified a "gestural drift", i.e., the bilingual's VOT categories in both languages were observed to have moved closer towards the VOT categories of the language she had been recently exposed to. As the results of the FAR show, this gestural drift was perceivable only for the Brazilian-Portuguese listeners while it was not audible to the majority of American English listeners.[74]

Taken together, the findings of the FARs discussed above show that on the level of global accent late L2-immersed bilinguals may be perceived as nonnative speakers of their L1 when being compared to monolingual control speakers. Consequently, not only specific acoustic characteristics of a bilingual's L1 segmental and prosodic productions might diverge from the native norm, but attrition – in the sense of *sounding* less native – might also be traced on a more global perceptual level. Thus, the studies presented above provide further counter-evidence to the assumption that a first language – once fully developed – is stable and resistant to changes.

74 This conclusion, however, has to be drawn with caution given that it is not clear which segmental and/or prosodic features the listeners based their nativeness judgments on, that is, it cannot be reliably stated that listeners used VOT as a (primary) cue to nativeness in this study.

4.1.2 Comprehensibility and intelligibility

Alongside examining perceived nativeness of AS's L1 German pronunciation, the present investigation focused on two further dimensions, namely comprehensibility and intelligibility of AS's L1 speech. As previously outlined, there are three studies examining global foreign accent in the L1 of bilingual speakers (Bergmann et al. 2016; de Leeuw, Schmid, and Mennen 2010; Hopp and Schmid 2013; see Table 4.1). To date, however, no investigations have been conducted which assess intelligibility and comprehensibility of potentially attrited L1 pronunciation – a gap the present study aims to fill. Given the lack of research in this area, this section will focus on studies which have been conducted in the field of second language acquisition to examine perceived intelligibility and comprehensibility of L2 pronunciation. Methodological considerations relating to the question of how to test perceived comprehensibility and intelligibility, and which speech samples to use in an empirical study will be discussed in the subsequent sections.

As defined in the introduction to this chapter, *intelligibility* refers to a listener's understanding of an utterance, while *comprehensibility* refers to the ease of understanding, i.e., how easy or difficult it is for a listener to understand an utterance (Derwing and Munro 1997; Kennedy and Trofimovich 2008; Munro and Derwing 1995a). As pointed out by Derwing and Munro (1997, 2), the terms 'comprehensibility' and 'intelligibility' have been used in different ways. Smith (1992, 76), for example, defined comprehensibility as the understanding of the actual meaning of a word or utterance, while intelligibility reflects a lower level of understanding, namely the recognition of a word or utterance. In the context of the present investigation, the two terms are used as defined by Munro and Derwing (1995a, see above), since these definitions are most widely applied in empirical studies examining perceived nativeness of L2 speech.

As outlined in Section 1.8, intelligibility, comprehensibility and accent play an essential role when it comes to communicative fluency, but they are not necessarily mutually dependent. In fact, research examining the perceived nativeness of L2 speech has shown that even a strong nonnative L2 accent does not automatically impede intelligibility and comprehensibility, i.e., heavily accented L2 speech might be still perfectly intelligible and comprehensible. Munro and Derwing (1995a), for example, asked native English listeners to rate English utterances produced by L1 Mandarin learners of L2 English and by monolingual English speakers, respectively, according to perceived accentedness, comprehensibility, and intelligibility. Despite the fact that the majority of bilingual speakers were rated being both clearly intelligible and comprehensible, a considerable number of these speakers was rated as 'heavily accented'. Similar observations were made in another study by Munro and Derwing (1995b), who tested to what extent foreign accent influences

sentence processing time. For this purpose, native English listeners were asked to rate the English pronunciation of L1 Mandarin L2 English bilinguals and native English speakers, respectively, according to perceived accentedness and comprehensibility. In addition, listeners' sentence processing times were measured by means of a sentence verification task. Results showed that English sentences produced with a Mandarin accent took the listeners longer to process than English sentences produced by native English speakers. It was also shown that English utterances which were perceived to be less comprehensible required longer processing times than sentences which were rated as being more comprehensible. While this study provides evidence for a negative correlation between processing time and comprehensibility (i.e., the more comprehensible, the less processing time), no direct correlation between processing time and degree of foreign accent could be observed. These findings indicate that "an accent – even a strong one – is by no means an inevitable barrier to communication" (Munro and Derwing 1995b, 302).

As previously mentioned, there is still a lack of research examining perceived comprehensibility and intelligibility in the L1 speech of L2-immersed late bilinguals. In order to fill this gap, the present study aims to provide first insights into the interrelation between the three perceptual dimensions of nativeness, intelligibility and comprehensibility of bilingual and monolingual (Austrian) German speech. In the subsequent sections, some methodological considerations concerning the assessment of perceived nativeness, intelligibility and comprehensibility will be discussed in detail and will be related to the aims of the present study.

4.1.3 Methodological aspects I: Assessing global accent

Global accent ratings are widely used in studies on L2 acquisition to examine if and to what extent bilinguals or second language learners are perceived to sound (non-)native in their L2. As discussed above, also in the field of L1 attrition, accent ratings have been used to assess the nativeness of L1 pronunciation of late consecutive bilinguals being immersed in an L2-speaking environment. However, the methodology to conduct such ratings – in the context of L2 acquisition and L1 attrition – has not been standardized so far, which results in differences concerning rating instruments, number and type of recordings, and the criteria for selecting speakers and listeners (see Jesney 2004). This section provides an overview of methodological aspects which should be taken into consideration when designing and conducting nativeness ratings based on previous global accent studies displayed in Table 4.2.

Study	Languages	N L2 speakers	N controls	N raters[75]	Elicitation method	Rating instrument
Tahta, Wood, and Loewenthal (1981)	Various L1s[76] L2: English	109	0	3	Reading (paragraph)	3-point-scale from *no foreign accent* (FA), to *marked accent*
Flege (1984)	L1: French L2: English	8	8	9-12[77]	Reading (single sentences)	Two-alternative-forced choice identification: *native* vs. *nonnative*
Anderson-Hsieh and Koehler (1988)	L1: Chinese L2: English	3	1	224	Reading (paragraph)	5-point-scale from heavy FA to no FA
Thompson (1991)	L1: Russian L2: English	36	10	16[78]	Reading (sentences, paragraph) + narrative task	5-point-scale from *no FA* to *heavy FA*
Bongaerts, Planken, and Schils (1995)	L1: Dutch L2: English	22	5	4	Reading (words, sentences, paragraph) + narrative task	5-point-scale from *very strong FA*: *definitively nonnative* to *no FA: definitely native*
Elliott (1995)	L1: English L2: Spanish	66	0	3	Picture description	5-point-scale from *very poor accent* to *native*
Flege, Munro, and MacKay (1995)	L1: Italian L2: English	240	24	10	Delayed repetition (sentences)	Sliding scale with endpoints no FA and *native speaker of Italian*
Munro and Derwing (1995a)	L1: Mandarin L2: English	10	2	18	Picture description	9-point-scale from *no FA* to *very strong FA*

75 'Raters' refers to the listeners who were asked to judge speech samples based on different criteria. In the present study, the terms 'raters', 'judges', and 'listeners' will be used interchangeably.

76 Among them, Arabic, Cantonese Chinese, Greek, Hindi, Spanish, and Swedish.

77 Given that Flege (1984) conducted a series of different experiments, varying numbers of judges were included, differing in terms of previous phonetic training.

78 Thompson (1991) included two different groups of judges, differing in terms of their level of L2 experience.

Derwing and Munro (1997)	Various L1s[79] L2: English	48	0	26	Picture description	9-point-scale from *no accent* to *extremely strong accent*
Riney and Flege (1998)	L1: Japanese L2: English	11	5	5	Reading (words, sentences)	9-point-scale from *strong FA* to *no FA*
Moyer (1999)	L1: English L2: German	24	4	4	Reading (word list, sentences, paragraph) + narrative task	Two-alternative-forced choice identification: *native* vs. *nonnative*, 3-point-scale to indicate degree of certainty
Jilka (2000a)	L1: English L2: German	10	10	57	Reading (sentences), repetition task + spontaneous conversation task	10-point visual analogue scale with endpoints *no FA at all* and *very strong FA*

Table 4.2: Sample studies examining perceived nativeness of L2 speech (in chronological order).

79 Among them, Cantonese, Japanese, Polish, and Spanish.

When it comes to the speech samples played to the listeners, differences can be observed in terms of the type of speech included in the samples. Moyer (1999), for example, asked her native German listeners to rate single word utterances produced by late learners of L2 German (see also Bongaerts, Planken, and Schils 1995). Other accent rating studies make use of longer stretches of read speech, such as Tahta, Wood, and Loewenthal (1981; see also Moyer 1999; Thompson 1991), who recorded 109 speakers with different L1s reading a short paragraph in their L2 English. The most frequently used type of input when it comes to the assessment of global L2 accent are single sentences produced in a reading task (e.g., Bongaerts, Planken, and Schils 1995; Flege and Fletcher 1992; Flege, Munro, and MacKay 1995; Riney and Flege 1998). Some studies use samples of spontaneous speech elicited in, for example, picture describing tasks (e.g., Elliott 1995; Munro and Derwing 1995a), free response tasks (e.g., Jilka 2000a; Moyer 1999; Thompson 1991), or film retelling tasks (e.g., Bergmann et al. 2016; de Leeuw, Schmid, and Mennen 2010, for the perception of global L1 accent). As can be seen in Table 4.2, most studies make use of reading tasks to elicit formal, non-spontaneous speech, and only few studies include (semi-)spontaneous samples. Although it is relatively easy to control for the content of speech, sentence length and phonological context when using reading tasks, speech elicited this way might not be entirely representative of a speaker's natural pronunciation. This was criticized by Long (1990, 267) who argued that using tape recordings of rehearsed and non-spontaneous speech cannot be considered indicative of a speaker's 'normal' pronunciation and of their phonological competence (see also Jilka et al. 2008). Similarly, Flege (1984, 695) pointed out that the nature of the speaking task is likely to affect the authenticity of both L2 and native speech production. The ratings conducted in the present investigation were based on non-prompted, authentic speech samples taken from different interviews with AS and monolingual Austrian German speakers (see Section 4.2.2, for a detailed description of the speakers included). This is a new approach insofar as none of the accent rating studies outlined in Table 4.2 make use of completely free speech samples produced in a non-experimental setting, but rely on semi-authentic, controlled speech instead.

A second methodological issue which has not been standardized so far is the length of stimuli presented to the judges. In the L1 accent rating studies discussed in Section 4.1.1, the duration of speech samples ranges from 10 to 23 seconds (Bergmann et al. 2016; de Leeuw, Schmid, and Mennen 2010; Hopp and Schmid 2013); similarly, the duration of free speech samples used in L2 accent rating studies typically lasts from 10 to 20 seconds (see Jesney 2004). Flege (1984), however, found that listeners actually do not need more than approximately 30 milliseconds of speech input to judge a speaker as native or nonnative. His native American English listeners were asked to detect nonnative English speech produced by a group of L1

French speakers. They were presented with speech samples which differed significantly in terms of length, ranging from the length of a short utterance to 30 milliseconds only.[80] Results showed that even when being presented with input as short as 30 milliseconds, listeners were able to make a correct judgment in 77% of the time.[81] However, as pointed out by Jesney (2004, 4), longer samples are generally considered being more informative for listeners in accent rating tasks.

In the present investigation, the speech samples presented to the listeners should not allow them to recognize AS as the speaker and thus influence their judgments. It was assumed that the longer the samples played to the listeners, the more likely they were to identify AS as the speaker. Therefore, only short samples were included (< 10 seconds), containing a single utterance or sentence. With an average duration of 5.75 seconds, the individual recordings were considered being suitable in terms of length for the listeners to make their judgments while at the same time not giving away who is speaking, as was shown in several pre-testing sessions conducted before the actual ratings. Another reason for including samples which were shorter compared to the samples used in the studies displayed in Table 4.2 was the actual feasibility and practicability of the intelligibility part of the experiment: In order to test perceived intelligibility, listeners were asked to produce a written transcript of what the speakers say in 29 different speech samples (see Section 4.2.2, for a detailed description). The intention was to avoid making the individual samples too long and thus gradually reduce the listeners' concentration or even their willingness to participate in the experiment.

Another important methodological question which needs to be addressed in the context of studying global accent is how many listeners and native versus nonnative speakers to include. The purpose of including speech samples produced by native (monolingual) control speakers in foreign accent studies on both L1 and L2 accent is the "establishment of native-speaker norms" (Jesney 2004, 6; see also Flege and Fletcher 1992). However, scholars do not seem to agree on the exact number of control speakers to include in relative proportion to the number of nonnative (bilingual) speakers. As shown in Table 4.2, the number of native controls varies from none (Elliott 1995; Tahta, Wood, and Loewenthal 1981) or one subject only (Anderson-Hsieh and Koehler 1988), to 24 control speakers in Flege, Munro, and MacKay (1995). In addition, the number of L2 speakers is not evenly distributed, ranging from three speakers in Anderson-Hsieh and Koehler (1988) to 240 in Flege, Munro,

80 In the first experiment, Flege (1984) presented the listeners with native and nonnative productions of the two short English utterances 'two little dogs' and 'two little birds'. In the second experiment, the initial CV-syllable /tu/ was extracted from each utterance and the judges were asked to rate these syllable-length stretches of speech according to perceived nativeness.

81 By comparison, 89% of the English *phrases* produced by native French speakers were identified correctly as nonnative productions.

and MacKay (1995). It can be seen, however, that generally less control speakers than L2 speakers are included, which also applies to the studies examining perceived L1 nativeness displayed in Table 4.2 (except for Bergmann et al. 2016, who included an equal number of bilingual and monolingual speakers). When deciding how many speakers to include to represent monolingual versus bilingual speech, potential range effects should be taken into consideration. As observed by Flege and Fletcher (1992, 370), the higher the number of native speech samples included, the more likely it is that listeners judge nonnative speakers as foreign-accented.

As for the number of listeners included, previous studies also largely differ: Snow and Hoefnagel-Höhle (1977), for example, asked a single listener to rate all speech samples while Anderson-Hsieh and Koehler (1988) used a total of 224 monolingual English judges to rate English speech samples produced by L1 Chinese speakers. As illustrated in Table 4.2, the number of native judges included in L1 accent rating studies also substantially differs, ranging from 19 in de Leeuw, Schmid, and Mennen (2010) up to 149 used by Hopp and Schmid (2013). Generally, it can be assumed that the more listeners are included, the more representative the findings obtained in rating tasks. For the present investigation, two groups including 20 listeners each were recruited as will be described in more detail in Section 4.2.2.

The last methodological aspect to consider concerns the rating instruments used to examine perceived nativeness/perceived accentedness. As can be seen in Table 4.2, the degree of perceived accent in L2 speech is commonly measured by means of rating scales, including 5-point scales (e.g., Anderson-Hsieh and Koehler 1988; Bongaerts, Planken, and Schils 1995), 9-point scales (e.g., Derwing and Munro 1997; Riney and Flege 1998), and 10-point scales (Jilka 2000a, 2000b). Such scales allow measuring the degree of perceived foreign accent, usually ranging from 'no foreign accent' to '(extremely) strong foreign accent'. Flege (1984) applied a different measurement tool in the form of a two-alternative-forced identification task, i.e., listeners had to decide whether the syllables they listened to were produced by a nonnative or a native speaker. Moyer (1999) added a second rating dimension to this binary judgment task (native vs. nonnative) by asking listeners to indicate how confident they were concerning their judgment on a 3-point scale ('very confident', 'fairly confident', 'not very confident'). Combining the binary judgment and the confidence level judgment then results in a six-point Likert scale, ranging from 1 *definitely native* to 6 *definitely nonnative*. This rating tool was adopted in all three accent rating studies focusing on bilinguals' first language pronunciation (Bergmann et al. 2016; de Leeuw, Schmid, and Mennen 2010; Hopp and Schmid 2013) and will also be used in the present study to evaluate if and to what extent AS's is perceived to sound (non-)native in his L1 German pronunciation.

4.1.4 Methodological aspects II: Assessing intelligibility and comprehensibility

While examining speakers' divergences from or approximations of native speaker norms by means of acoustic-phonetic analyses provides quantifiable results, assessing intelligibility and comprehensibility of L2 and L1 speech entails the question of how to operationalize these rather abstract dimensions. This section discusses some methodological aspects related to testing the degree of perceived intelligibility and comprehensibility.

As shown in Table 4.3, the most commonly applied method to assess speech intelligibility is to ask listeners to provide orthographic transcriptions of what they hear in a speech recording (Derwing and Munro 1997; Kennedy and Trofimovich 2008; Munro and Derwing 1995a; Munro, Derwing, and Morton 2006). However, earlier studies made use of different testing tools to assess how intelligible L2 speech is perceived to be: Smith and Rafiqzad (1979), for example, designed cloze-tests[82] listeners had to complete. Speaker intelligibility was then evaluated based on the number of words the listeners had identified correctly. In a more recent study, Thomson and Derwing (2016) investigated if the type of practice stimuli (real words vs. nonsense words) had a significant effect on pronunciation accuracy of English vowels produced by L2 learners of English. For this purpose, two phonetically trained listeners were asked to rate each target word by deciding whether the vowel produced in this word represents a 'good', 'poor', or 'incorrect' example of the English vowel target. As discussed by Thomson (2018), this particular rating task does not exclusively focus on assessing the dimension of intelligibility but examines perceived nativeness at the same time.

Perlmutter (1989) asked native English listeners to rate the intelligibility of L2 English speakers with different L1s on a 5-point scale, ranging from 1 being the *least intelligible* and 5 being the *most intelligible*,. In addition, raters were instructed to name the main topic of the short monologue they had listened to. With regard to examining speech intelligibility by means of a rating scale, Munro (2008, 201) argues that rating data are less reliable compared to listener transcripts "because listeners sometimes mistakenly believe that they have understood an utterance, and may therefore rate it highly, when they have not understood it well at all". As listener transcripts have been widely adopted in previous studies, this testing method was also applied in the present investigation to assess perceived intelligibility of AS's L1 speech (see Section 4.2.2).

82 First described by Taylor (1953), cloze-tests are mostly used to examine language abilities in the context of second/foreign language learning assessment. This testing format involves deleting words from a text passage or sentence and asking subjects to identify those word(s) correctly which have been removed (e.g., Bachman 1982).

As can be seen in Table 4.3, the majority of studies focus on a combination of assessing both intelligibility (through listener transcriptions) and comprehensibility (Derwing and Munro 1997; Kennedy and Trofimovich 2008; Munro and Derwing 1995a; Munro, Derwing, and Morton 2006). Comprehensibility is usually tested in scalar fashion using 9-point scales to indicate how difficult or easy an utterance is to understand for a listener. The rating results of perceived comprehensibility and accentedness in combination with the transcript scores are then used to make a statement about the interrelation between these three dimensions. As will be described in Section 4.2.2, the present study makes use of a 6-point scale, ranging from *very easy to understand* to *very difficult to understand*, to examine perceived comprehensibility. Unlike an uneven-point scale, a 6-point scale does not provide the middle option for 'neutral' or 'undecided'. The reason for adopting a forced choice measure excluding a middle option was that the middle option was considered an 'easy-way-out'-option for listeners who are undecided or who might do not want to put much thought into their ratings (see Allen and Seaman 2007; Bartram 2016).

In addition to deciding for an appropriate technique to assess perceived intelligibility and comprehensibility, also the type of speech to be rated plays an essential role. As discussed in the previous section, using speech material produced by means of reading tasks cannot necessarily be considered representative of a speaker's natural pronunciation (see Jilka et al. 2008; Long 1990). With regard to examining perceived intelligibility and comprehensibility, also Munro (2008, 202) suggests using spontaneously produced speech samples. He argues, however, that spontaneous speech often includes grammatical inaccuracies which might influence a listener's ratings, and a speaker might not necessarily produce specific target words or sounds that are supposed to be tested. As the present investigation did not aim at testing the perception of particular segments and/or prosodic features, but focused on the degree of perceived nativeness, comprehensibility and intelligibility of L1 pronunciation, the speech materials were not selected based on the sounds/words uttered by a speaker. As will be outlined in Section 4.2.2, speech samples containing grammatical and/or syntactic errors were entirely excluded from the corpus.

Study	Languages	N L2 speakers	N controls	N raters	Elicitation method	Assessment of intelligibility	Assessment of comprehensibility
Munro and Derwing (1995a)	L1: Mandarin L2: English	10	2	18	Picture description	Listener transcriptions	9-point-scale from *extremely easy to understand* to *impossible to understand*
Munro and Derwing (1995b)	L1: Mandarin L2: English	10	10	20	Reading (sentences)	–	9-point-scale from *not difficult to understand at all* to *very difficult to understand*
Derwing and Munro (1997)	L1: Cantonese, Japanese, Polish, Spanish L2: English	48	–	26	Picture description	Listener transcriptions	9-point-scale from *extremely easy to understand* to *impossible to understand*
Munro, Derwing, and Morton (2006)	L1: Cantonese, Japanese, Polish, Spanish L2: English	48	–	40	Picture description	Listener transcriptions	9-point-scale from *extremely easy to understand* to *impossible to understand*
Kennedy and Trofimovich (2008)	L1: Mandarin L2: English	12	12	24	Reading (sentences)	Listener transcriptions	9-point-scale from *very easy to understand* to *very hard to understand*
Kraut and Wulff (2013)	Various L1s[83] L2: English	24	6	78	Reading (text passage)	–	7-point-scale[84]

83 Among them, Spanish, Chinese, Japanese, and Arabic.

84 Kraut and Wulff (2013) do not describe in more detail which items were represented by the scale.

Munro, Derwing, and Saito (2013)	L1: Russian or Mandarin L2: English	31	–	3	Reading (sentences)	Identification task of English vowels in monosyllabic words	–
Saito, Trofimovich, and Isaacs (2016)	L1: Japanese L2: English	120	10	10[85]	Picture description	–	9-point-scale from *very easy to understand, no accent* to *very hard to understand, heavily accented*
Thomson and Derwing (2016)	Various L1s[86] L2: (Canadian) English	21	10	2	Reading (target items in carrier phrase)	3-alternative-forced choice identification of English vowels in monosyllabic words	–

Table 4.3: Overview of studies examining perceived intelligibility and/or comprehensibility in L2 speakers (in chronological order).[87]

85 The rater group consisted of two sub-groups, i.e., phonetically trained and untrained (naïve) listeners.

86 Among them, Chinese, Spanish, Russian, and French.

87 As mentioned in Section 4.1.2, no studies have been conducted so far which assess the dimensions of intelligibility and comprehensibility of potentially attrited L1 speech.

4.2. The study

4.2.1 Aims and objectives

As outlined in the previous sections, the perception of global accent and the dimensions of perceived comprehensibility and intelligibility have been most frequently examined in the context of the acquisition of L2 pronunciation. With regard to this, research has focused on the question of whether late bilinguals are perceived to sound (non-)native in their L2 and to what extent a strong nonnative accent might affect the comprehensibility and intelligibility of L2 speech (see overviews in Tables 4.2 and 4.3). By contrast, only three studies (Bergmann et al. 2016; de Leeuw, Schmid, and Mennen 2010; Hopp and Schmid 2013) have been conducted so far to examine perceived nativeness in the L1 of immersed late bilinguals, and to date there are no investigations which assess the potential relationship between perceived L1 accent and the comprehensibility and intelligibility of potentially accented L1 speech. The present study aimed to fill this gap by answering the following research questions on the basis of different rating tasks:

RQ 3.1: Is AS perceived to sound less native in his late L1 German pronunciation than in his early L1 pronunciation compared to monolingual Austrian German speakers?

RQ 3.2: To what extent does the listeners' linguistic background (monolingual vs. bilingual) affect their perception of L1 nativeness in German-English bilingual and German monolingual speakers?

RQ 3.3: To what extent is the perceived nativeness of L1 German speech related to the intelligibility and comprehensibility of L1 speech, i.e., does a lower degree of perceived nativeness impede the comprehensibility and intelligibility of L1 pronunciation, and vice versa?

4.2.2 Methodology

4.2.2.1 Speakers and speech samples

The samples representing AS's L1 early (1970s) and late (2017) German pronunciation (see Appendix E) were extracted from the speech corpus used in the acoustic analyses presented in Chapters 2 and 3. Overall, four samples representing AS's early German pronunciation, and four samples representing his late German pronunciation were included. The selection of speech samples was based on a set of criteria which have been applied in previous accent rating studies (see Hopp and Schmid 2013; Schmid and Hopp 2014, for overviews). The samples used in the

ratings did not include hesitations, pauses longer than one second and no or only few self-corrections. In addition, utterances with grammatical and/or lexical errors were excluded in order to avoid that listeners make their judgments on the basis of grammatically or lexically incorrect sentences[88] (see e.g., de Leeuw 2009). Also, as suggested by Hopp and Schmid (2013), samples containing borrowings from the L2 as well as place or proper names were not included. Avoiding place and proper names was especially relevant in the present investigation, because mentioning Thal (AS's place of birth), the city of Graz (which is close to Thal), and other names/places could be easily associated with the person of Arnold Schwarzenegger and thus make it easier for the listeners to identify who is speaking. The speech recordings included some moderate background noise, which was not filtered out because it was not considered disturbing.

As mention in Section 4.1.3, only short stretches of speech were selected in order to avoid that listeners identify AS as the speaker. The length of the individual samples varied from 3.21 seconds to 7.75 seconds, depending on where the sound files were cut. The recordings differed in length because each sample was supposed to represent a complete phrase or clause and it was not possible to observe clause and phrase boundaries while selecting samples of equal length at the same time. The choice of which utterances/sentences to select was based on subjective interpretation, but followed the criteria mentioned above. The average duration of 5.74 seconds was assumed to be suitable for listeners to make their judgements (see Flege 1984).

A second set of speech samples represented the speech of Frank Stronach (see Appendix E), a native Austrian-born businessman and politician, who moved to Canada at the age of 21 in 1954. Stronach was born as Franz Strohsack in a small municipality in East Styria in 1932 (Mayr 2013; Noble 2014). Similar to AS, Stronach is a late consecutive bilingual who acquired his L2 English when he migrated to an L2-speaking country. Stronach was considered an appropriate subject for the present investigation due to three main reasons: First of all, Stronach can be described as a potential L1 attriter, who started acquiring the L2 after his L1 had already been fully developed, and who has been long-term immersed in an L2-speaking environment. Second, he was exactly the same age as AS when he moved to Canada, which has been his primary place of residence ever since.[89] Third, Stronach speaks an L1

88 Trofimovich and Isaacs (2012) showed that grammatical accuracy is closely linked to perceived comprehensibility, i.e., utterances and sentences containing grammatical errors are more likely to be rated as being less comprehensible compared to sentences which contain less/no grammatical errors.

89 Similar to AS, Stronach is a public person, who runs businesses not only in Canada, but also in Austria and other countries around the world (Noble 2014). He therefore travels on a more or less regular basis to his home country and gives interviews for Austrian TV and radio channels (ARD Mittagsmagazin 2012; Menschen im Porträt TV 2018; OE24.TV 2018).

Austrian German variety which can be broadly described as East Styrian (Ernst 2004) and is thus close to AS's L1 variety. Speech samples were extracted from interviews Stronach gave for different Austrian TV and radio channels in 2018.[90] The four samples included in the ratings were selected according to the criteria discussed above and have a mean duration of 4.9 seconds (range: 3.21 to 6.51 seconds). With regard to the content, the samples were very similar to the control samples discussed below, i.e., they addressed general topics in order to avoid giving away who is speaking – although Stronach might not be as well-known as AS.

A group of seven monolingual Austrian German male speakers from Thal served as control speakers in the present study. Since AS's and Stronach's samples used in this study represent non-prompted, (semi-)spontaneous speech and were selected from a pre-recoded speech corpus, also the control speakers' samples were taken from different TV and radio recordings, i.e., all speakers produced speech in a spontaneous and non-experimental setting. They were roughly the same age as AS, ranging from 69 to 78 years, and were friends and former training partners[91] of AS, who were interviewed on different occasions, such as AS's birthdays or the opening of the *Schwarzenegger Museum* in Thal in 2011. Content-wise, the speech samples are similar, i.e., they focus on sports, weight training and different training routines without explicit or implicit references to AS's career as a bodybuilder. Some samples address more general topics, such as childhood memories or leisure activities.

	N subjects	**N samples**	**Mean word count**	**Mean sample duration (sec.)**
Bilingual group	2	12	16.42	5.86
Control group	7	17	14.23	5.62
	N_{total} subjects	N_{total} samples	Total word count	Total duration (sec.)
	9	29	440	166.62

Table 4.4: Speakers and speech samples included in the three rating tasks.

As discussed earlier, AS is well known by many Austrians since he has been a public person for over four decades and has been present in the media and in films ever since. Thus, it can be assumed that most people are familiar with his voice and his German pronunciation, which might make it easy for listeners to identify AS as the

90 Recordings representing Stronach's early pronunciation, that is, around the time he moved to Canada, were not available.

91 The speakers included are Peter Urdl, former mayor of Thal and the founder of the *Schwarzenegger Museum* in Thal, Kurt Marnul and Karl Kainrath, AS's former training partners and close friends, Peter Schickhofer, mayor of Thal from 2008 to 2015, and Alfred Gerstl, who died in 2016 at the age of 93 and is often referred to as AS's mentor who supported him at the very beginning of his career (ORF Steiermark 2016).

speaker in a listening experiment. The same is true for Frank Stronach, who is also a well-known public figure. In order to determine whether the speech materials selected for the present study were suitable and whether listeners can identify AS's or Stronach's speech, a pre-testing session was conducted. Thus, not only the materials to be used in the study, but also the experimental setup and procedure could be tested in advance. Eight randomly selected listeners participated in the ratings. Each participant was asked to listen to the individual recordings and fill out an answer sheet after each recording (see Section 4.2.2.3). After completing the session, the listeners were asked by the experimenter whether they had recognized one or more of the speakers. As none of the listeners identified AS or Stronach as the speaker, the recordings were considered suitable for the actual study.

4.2.2.2 Listeners

A total of 40 listeners participated in the ratings. The subjects were assigned to two groups (n = 20 each), depending on their language background. The first group (ML–GRA) included (quasi-)monolingual speakers of Austrian German (Styrian) from Graz while the second group (BIL–GRA) included native Austrian German speakers from Graz, who were undergraduate students of American and English studies at the University of Graz at the time the present study was conducted. All participants reported normal hearing and none of them suffered from speech impairments. Participants in both groups, which will be described in more detail below, were recruited via personal contacts in Graz and at the Department of English Studies at the University of Graz.

Participants in Group ML–GRA (10 female, 10 male), aged between 25 and 39 years (M = 31.8; see Appendix F), were (quasi-)monolingual speakers of Austrian German with some pre-existing knowledge of English. As English is a compulsory subject in Austrian schools,[92] it is nearly impossible to find participants in Austria who do not have at least some very basic knowledge of English. In order to obtain information about the participants' use of English, each participant was invited to fill in a questionnaire (see Appendix G). The assessment of overall English proficiency was based on the six competence levels provided by the Common European Framework of Reference for Languages (Council of Europe 2001; see Appendix G, for a definition of the levels) and ranged from A1 (*elementary user of English*) to B1/B2 (*independent user of English*) in the monolingual group. Figure 4.1 provides an overview of the participants' self-reported frequency of English use in different contexts (speaking, listening, reading, writing), which they had to indicate on a scale ranging

92 Austrian children start learning English as their first foreign language at the age of 10–11 in grade 5 (first year of Gymnasium).

from *very frequently* to *very rarely* or *never*.[93] As can be seen, the majority of monolingual subjects rarely or never used English, which qualifies them as "functional monolinguals" (Best and Tyler 2007, 16), who do not actively learn or use a second/foreign language. In addition, none of the participants reported having stayed in another country or another Austrian city for an extended period of time (i.e., longer than four weeks). All participants in this group were born and raised in Graz or Graz-Umgebung[94] and have been permanently living there. Thus, it could be assumed that all subjects were equally familiar with the Styrian dialect and used in on a daily basis when interacting with friends and family. Furthermore, none of the subjects in this group had received phonetic training in advance or had pre-existing phonetic knowledge.[95]

93 On the questionnaire (Appendix G), participants were asked, for example: Do you speak English in your professional environment? (yes/no), If 'yes', how often? (very rarely/rarely/frequently/very frequently/never).

94 Graz-Umgebung is a political district of Styria and includes 36 municipalities. It has the second highest number of inhabitants after Graz City (Statistik Austria 2019).

95 Previous studies on the perception of global foreign accent have included both phonetically trained listeners (e.g., Munro and Derwing 1995a; Munro, Derwing, and Saito 2013) as well as naïve, untrained listeners (e.g., Flege and Fletcher 1992; Flege, Munro, and MacKay 1995; Major 2007; Munro, Derwing, and Burgess 2003). While it might be argued that trained listeners have an advantage over untrained listeners when it comes to the detection of nonnative speech, research shows that also listeners who do not have phonetic/linguistic experience are well able to identify accented speech. Major (2007), for example, showed that phonetically naïve listeners were able to distinguish instances of foreign-accented speech from native speech in languages they were not familiar with.

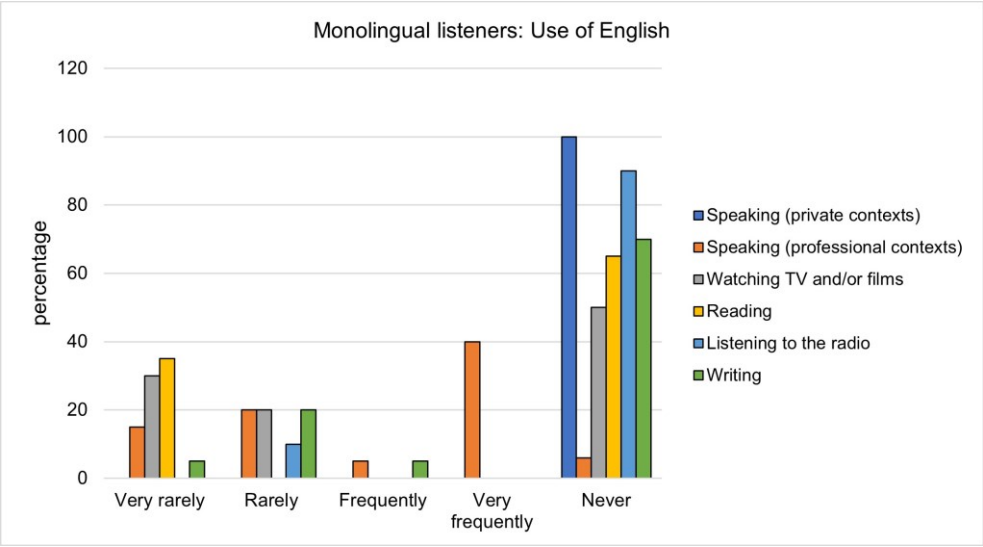


Figure 4.1: Monolingual listeners' (n = 20) self-reported use of English in different contexts.

Subjects in Group BIL–GRA (11 female, 9 male), aged between 19 and 25 years[96] (M = 21.65; Appendix F) and born and raised in Graz and Graz-Umgebung, have been studying English at the University of Graz for an average of 2.46 years. The majority of the subjects were enrolled in a teacher training program (n = 13), the remaining n = 7 subjects were undergrade students of English and American Studies. Besides studying English at university level, the majority of subjects reported using English on a regular basis outside university in different contexts, as depicted in Figure 4.2. Ten subjects reported that they had already stayed in an L2-speaking country for an extended period of time, ranging from 3 to 12.5 months. As all participants in the BIL–GRA Group were undergraduate students, none of them had received intensive phonetic training during the first years of their studies. However, they had gained some phonetic experience in different pronunciation and linguistics classes, which introduce students to basic concepts related to English phonetics and phonology. Thus, similar to the subjects in the ML–GRA Group described above, none of the BIL–GRA subjects could be identified as a phonetically trained listener.

96 As can be seen, the listeners in the two groups differed in terms of their mean age, with the bilingual listeners being younger (M = 21.65 years) than the monolingual listeners (M = 31.8 years), which was due to the generally limited availability of participants and due to the fact that the subjects in the bilingual group were university students. However, as pointed out by Schmid and Hopp (2014, 372), there is no evidence so far which indicates that age and/or gender differences significantly affect the perception of global accent.

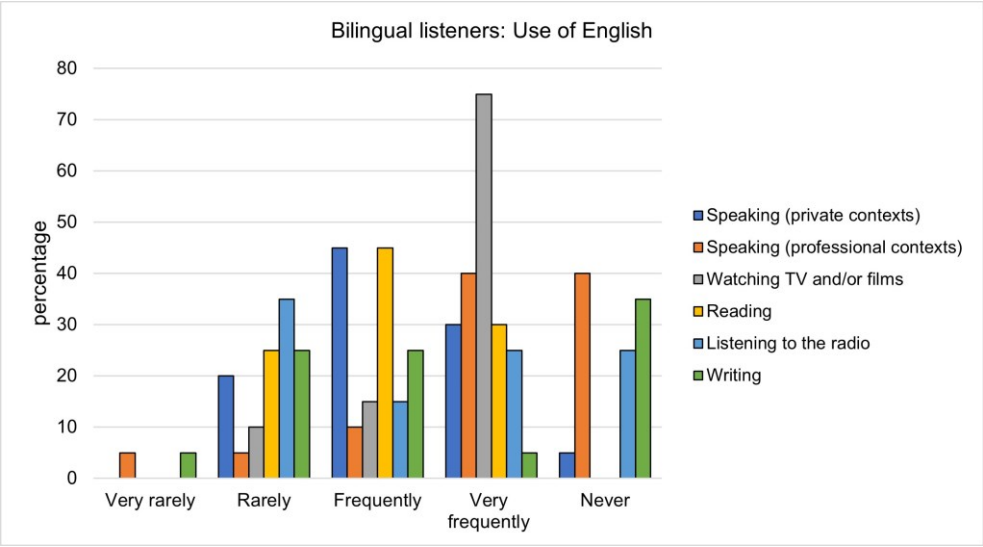


Figure 4.2: Bilingual listeners' (n = 20) self-reported use of English in different contexts.

As outlined above, participants in the two groups mainly differed in terms of their language background and linguistic experience. As pointed out by Long (1990, 267), subjects who have already experienced much linguistic diversity through, for instance, regular travelling, contract with nonnative speakers or living in another country for an extended period of time, are likely to show an increased "tolerance for and expectation of within-language variation". Consequently, linguistically experienced listeners – that is, subjects in the BIL–GRA Group – might be more hesitant to judge a speaker as nonnative compared to listeners who are predominantly exposed to their L1 variety and thus have less tolerance for variations in pronunciation (i.e., subjects in the ML–GRA Group).[97]

In terms of the listeners' L1 variety, subjects in both groups were equally familiar with the Styrian variety spoken in Graz (and Graz-Umgebung) as they were born and raised in Graz. All listeners reported being dialect speakers and the majority (n = 38 out of 40) stated that they were well able to distinguish Styrian from other Austrian dialects (the remaining two speakers chose the option *I don't know* on the questionnaire). As pointed out by de Leeuw, Schmid, and Mennen (2010, 38; see also Almberg and Husby 2000; Flege 1984; Reinisch 2005), a lack of familiarity with a

97 It should be noted, however, that due to globalization and multicultural influences, linguistic diversity is present of course also in Graz, the subjects' place of residence. Graz is the second largest city in Austria after Vienna and is located at a number of borders, which has led to an increasing contact with different nationalities and different languages in the past decades.

specific regional L1 accent might reduce listeners' ability to distinguish between foreign-accented and non-foreign-accented speech, which may lead to a misperception of an L1 variety as a non-native accent. Therefore, only listeners who reported being familiar with the Styrian variety were selected as judges in the present study.

4.2.2.3 Experimental procedure

The rating experiment was designed as an online experiment, i.e., the materials and detailed task descriptions (all presented in German) were sent to the participants via e-mail. In a first step, potential participants were asked to fill in a questionnaire (see Appendix G) in order to identify those subjects who fulfilled the listener requirements specified above. As described in the previous section, the questionnaire collected information about the subjects' language background, particularly focusing on their L1 use and L2 learning experience, and gathered information about relevant personal details, such as age, sex, and education. Out of a total of 55 subjects who were asked to complete the questionnaire, eight subjects were excluded from the study because they were either nonnative speakers of Austrian German or had lived in another Austrian city for more than four months. The remaining 47 subjects fulfilled the requirements and were assigned to one of the two groups described in the previous section. However, the responses from six listeners were excluded afterwards, because these listeners had recognized either AS or Stronach as the speaker in one or more of the recordings. One participant did not complete all three rating tasks and was therefore also excluded from analysis.

Subjects who met the participation criteria were then provided with the materials via e-mail, including a description of the overall rating procedure (see Appendix H). The materials consisted of a set of Power Point slides including the three sets of speech samples and a detailed description of each rating task. In addition, each participant received an Excel file including three answer sheets for the respective listening tasks.[98] In order to ensure that the task descriptions were clear and easy to follow, the materials were tested in advance. The eight test runs, which were conducted prior to the actual testing, showed that the participants needed approximately 35 to 45 minutes to complete all tasks. Before starting the ratings, each participant was asked to sign an informed consent sheet. The experimental part included three rating tasks, which aimed at testing (1) perceived nativeness (FAR), (2) perceived intelligibility (INTEL), and (3) perceived comprehensibility (COMP),

98 Each of the three rating sheets was included in a single Excel-file with different spreadsheets (i.e., three spreadsheets labelled "Antwortbogen_1", "Antwortbogen_2", and "Antwortbogen_3"). For each of the three rating tasks, the 29 sound files were presented via Power Point with one sound file per slide, preceded by a task description.

respectively. After each task, the listener was informed on a Power Point slide that they were allowed to take a short break and then continue with the next part.

To investigate (1) perceived nativeness, the foreign accent assessment scale introduced by Moyer (1999) was adopted, which has also been used in studies focusing on perceived nativeness of L1 pronunciation (Bergmann et al. 2016; de Leeuw, Schmid, and Mennen 2010; Hopp and Schmid 2013). Each listener was asked to make two judgments after listening to a recording (i.e., after listening to one speaker at a time): Based on a two-alternative-forced choice identification, the listener had to decide first if the speaker is a native or nonnative speaker of German. The second judgment reflected their confidence level on a three-point scale.[99] The judgments were then converted to a six-point Likert scale: 6 = *certain of nonnative speaker status*, 5 = *semi-certain of nonnative speaker status*, 4 = *uncertain of nonnative speaker status*, 3 = *uncertain of native speaker status*, 2 = *semi-certain of native speaker status*, 1 = *certain of native speaker status*.

In the second part of the rating experiment, which tested (2) perceived comprehensibility, the participants listened to the same recordings as in the first part, but in a different order. After listening to one speaker at a time, the listeners had to decide how difficult or easy the speaker was to understand on a 6-point scale, ranging from 1 = *very easy to understand* to 6 = *very difficult to understand*.[100]

The third part of the rating examined (3) perceived intelligibility. Again, the participants listened to the 29 speech samples in a different order. After each recording, they were instructed to write down word by word what they heard on a separate answer sheet (see Appendix K). They were allowed to listen to the speech sample a second time if they were not able to complete their transcription in one go.

In order to control for the possibility that listeners might have identified AS or Stronach as the speakers, each listener was asked whether there was anything in the speech samples they had noticed while listening. As mentioned above, six listeners stated that they had recognized AS and/or Stronach in one or more than one of the recordings. Although they did not identify the two speakers in all twelve recordings and they did not always seem to be entirely sure about whether it was them or not, the ratings of these participants were excluded from the analysis. None of the remaining listeners indicated that they had recognized one or more of the speakers.

---

99 On the German questionnaire (see Appendix I), each listener was asked 1) *Hat dieser Sprecher Deutsch als Muttersprache?* (*Is this speaker a native German speaker?*) with two possible responses (*Ja/Nein*, *Yes/No*); (2) *Wie sicher sind Sie sich?* (*How certain are you of this?*) with three answer options (*sicher/certain – relativ sicher/relatively certain – unsicher/uncertain*).

100 The German questionnaire (see Appendix J) included the following items: 1 = *sehr leicht verständlich* (*very easy to understand*), 2 = *leicht verständlich* (*easy to understand*), 3 = *eher leicht verständlich* (*rather easy to understand*) ; 4 = *eher schwer verständlich* (*rather difficult to understand*), 5 = *schwer verständlich* (*difficult to understand*), 6 = *sehr schwer verständlich* (*very difficult to understand*).

4.2.2.4 Analysis procedure

As specified above, each of the 40 listeners was invited to complete three rating tasks (FAR, COMP, INTEL) based on 29 speech samples, resulting in 87 ratings per listener and a total of 3,480 individual ratings (40 listeners x 3 tasks x 29 samples). The ratings where analyzed according to task and the resulting data were organized in a CSV file (Microsoft Excel 2016), which was used later for the statistical analyses (Section 4.2.3).

For the first rating task (FAR), the two judgments obtained per listener were, as described above, converted to a six-point Likert scale ranging from 1 = *certain of native speaker status* to 6 = *certain of nonnative speaker status* (de Leeuw, Schmid, and Mennen 2010; Moyer 1999). Consequently, speakers with a low accent rating were perceived to have a native or near-native L1 pronunciation, while speakers with a high accent rating were judged as nonnative or near-nonnative speakers of German. Resulting from this, a speaker who was rated '6' on the Likert scale received the highest accent rating and was perceived to sound clearly nonnative. By contrast, a speaker who received a rating of '1' was considered having a clearly native pronunciation. Also the second rating task (COMP) was based on scalar ratings, which resulted in each speaker being assigned a value between 1 (*very easy to understand*) and 6 (*very difficult to understand*) by the 40 listeners.

INTEL was assessed based on listener transcripts. Prior to conducting the ratings, the speech recordings used in Study III were transcribed by hand by the experimenter. The accuracy of each transcription was verified by two native Austrian German speakers from Graz, who were otherwise not involved in the present investigation. In line with the analysis procedure applied by Derwing and Munro (1997; see also Munro, Derwing, and Morton 2006), an INTEL score was calculated for each listener transcription defined as "the percentage of words exactly matching the original transcription" (Derwing and Munro 1997, 7). For this purpose, each listener transcript was compared to the original transcript and divergences from the original transcript were labelled as errors, including words being left out, incorrect word order and adding words which were not said by the speaker. Errors which were not counted and, thus, not included in the INTEL score calculations, were misspellings/typos and omissions of single words repeated by the speaker as typical of natural speech. The identification of errors in each transcript was conducted manually and each corrected transcript was double-checked by one of the native Austrian German speakers who verified the original transcripts.

4.2.3 Results

The main objective of Study III was to determine if AS is perceived to sound more nonnative in his late L1 German speech compared to his early L1 speech and to

what extent the degree of nativeness is perceived differently in AS's L1 speech when being compared to age-matched monolingual Austrian German control speakers from Thal, AS's place of birth (RQ 3.1). Alongside including speech samples representing AS's early and late L1 German pronunciation and the pronunciation of monolingual German control speakers, an additional set of samples was included. These samples were produced by Frank Stronach, a native Austrian-born Canadian immigrant, who – similar to AS – is an L2-immersed late consecutive bilingual and thus qualified as a control bilingual. A further aim of the present study was to examine perceived comprehensibility (COMP) and intelligibility (INTEL) of AS's early and late L1 German pronunciation compared to the monolingual control speakers in order to examine the extent to which perceived nativeness can be related to the intelligibility and comprehensibility of L1 speech (RQ 3.3). Two groups of listeners (n = 20 each), who differed in terms of their language background (monolingual vs. bilingual), rated 29 speech samples according to perceived nativeness, comprehensibility and intelligibility in three ratings tasks. Potential effects of the listeners' language background on the degree of perceived nativeness were also examined (RQ 3.2). This section gives an overview of the descriptive statistics of the results obtained in the three rating tasks and presents the statistical tests used for analysis.

(I) Nativeness ratings (FAR)

In order to assess inter-rater reliability, Cronbach's alpha coefficient was calculated. Overall inter-rater reliability was high, with α = .82. A slightly lower, but still acceptable Cronbach alpha coefficient was determined for the bilingual listener group (α = .74) compared to the monolingual listener group (α = .8). First, the statistical models used to examine the overall FAR scores obtained for the individual bilingual speakers (AS_early, AS_late and Stronach) and the monolingual Austrian German control speakers (ML_AG) will be presented. Next, the statistical analyses of rating differences between the two listener groups (BIL = bilingual vs. ML = monolingual) will be presented.

Figure 4.3 displays the FAR scores assigned to the individual bilingual speakers and to the monolingual Austrian German control group. On a six-point scale, ranging from 1 = *certain of native speaker status* to 6 = *certain of nonnative speaker status*, speakers in the monolingual control group received an average rating of 1.71 (*SD* = 0.85, *Mdn* = 1.0). AS's early L1 German pronunciation received a slightly higher average rating of 2.33 (*SD* = 1.06, *Mdn* =2.0). The speech samples produced by the two bilinguals, AS_late and Stronach, were rated with an average of 3.56 (*SD* = 1.33, *Mdn* = 3.0) and 3.83 (*SD* = 1.33, *Mdn* = 4.0), respectively.

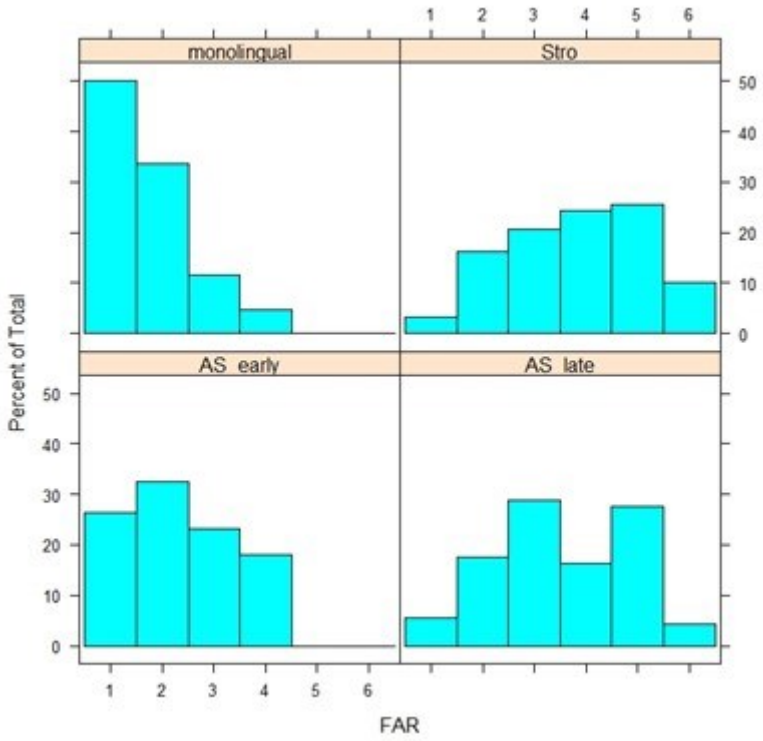


Figure 4.3: Distribution of nativeness ratings according to speakers.

Due to the fact that several responses were obtained from each listener for each speaker, two-way repeated ordinal regression analyses were performed in R (R Core Team 2020; version 4.0.1), using the *Anova.clmm* function from the *RVAide-Memoire* package (Hervé 2020) to determine main and interaction effects. *Post hoc* Tukey's tests were conducted using the *emmeans* R package (Lenth 2020). An alpha level of .05 was adopted throughout.

The first model addressed RQ 3.1, i.e., whether AS's L1 German is perceived to sound less native in his late speech (AS_late) than in his early speech (AS_early) compared to monolingual Austrian German speakers (ML_AG). In addition, it was assessed if the FAR scores assigned to the control bilingual (Stro) were significantly different from those assigned to the monolingual control speakers. In the model, *FAR* (rating score) was included as the dependent variable, and *Speakers* (four levels: *AS_early*, *AS_late*, *Stro* and *ML_AG*) as fixed effect. As a random effect, by-*Listener* random intercepts were entered (FAR~Speaker + (1|Listener)). The model showed a main effect for *Speaker*, with $\chi 2[3] = 331.47$, $p < .001$. Pairwise comparison using

post hoc Tukey tests revealed a significant difference between FARs obtained for AS_early and those obtained for AS_late (β = -1.93, *SE* = 0.21, z = -9.16, $p < .001$), i.e., AS late pronunciation was perceived to sound less native compared to his early pronunciation by the native Austrian German listeners. In addition, a significant difference in FAR scores between AS_late and the monolingual control speakers was revealed (β = 3.24, *SE* = 0.19, z = 17.32, $p < .001$), indicating that in his late German pronunciation, AS is perceived to sound less native than the monolingual control speakers. The difference between FAR scores obtained for AS_early and those obtained for the monolingual group was also shown to be significant (β = 1.3, *SE* = 0.17, z = 7.83, $p < .001$), that is, the listeners perceived his early L1 German pronunciation as being less native-like than the monolinguals' German pronunciation. As depicted in Figure 4.3, the control bilingual (Stro) received overall higher FARs compared to the monolingual speaker group; this difference was shown to be significant (β = -3.68, *SE* = 0.19, z = -19.13, $p < .001$). By contrast, the rating scores obtained for AS_late and those obtained for Stronach (see Figure 4.3) did not turn out to be significantly different (p = .13).

A second objective of the present investigation was to examine if the listeners' linguistic background (BIL vs. ML) had an effect on their perception of nativeness in AS's early and late German pronunciation (RQ 3.2). An inspection of Figure 4.4 reveals differences in the rating scores assigned to AS_late and AS_early by the monolingual and bilingual listener group, respectively: While AS_late received a mean rating of 2.94 (*SD* = 1.24, *Mdn* = 3.0) from the BIL listeners, the ML listener group assigned AS_late a higher mean rating of 4.17 (*SD* = 1.11, *Mdn* = 4.0). Similarly, the control bilingual (see Figure 4.5) was rated with a mean rating of 3.35 (*SD* = 1.36, *Mdn* = 3.0) by the BIL listeners, while the ML listeners rated him with a mean rating of 4.31 (*SD* = 1.12, *Mdn* = 4.5). By contrast, the BIL and ML listener groups rated the monolingual control speakers with 1.64 (*SD* = 0.87, *Mdn* = 1.0) and 1.78 (*SD* = 0.82, *Mdn* = 2.0), respectively.

The second model which addressed RQ 3.2. included *FAR* as the dependent variable and *Group* (two levels: *BIL* and *ML*) and *Speaker* (four levels: *AS_early, AS_late, Stro* and *ML_AG*) as well as an interaction between the two as fixed effects. As in the first model, a by-*Listener* random intercept was entered (FAR~Group + Speaker + Group:Speaker + (1|Listener)). Main effects for *Group* ($\chi 2$[1] = 12.55, $p < .001$) and *Speaker* ($\chi 2$[1] = 86.45, $p < .001$), as already identified in the first model, as well as a significant interaction between *Group* and *Speaker* ($\chi 2$[1] = 14.72, $p < .001$) were found. Pairwise comparisons using *post hoc* Tukey tests showed that the group rating differences were significant for both AS_late (β = -2.03, *SE* = 0.36, z = -5.62, $p < .001$) and Stro (β = -1.59, *SE* = 0.36, z = -4.45, $p < .001$), which indicates that BIL listeners were less likely to perceive the bilinguals as sounding nonnative in their L1 compared to the ML listeners. No significant differences were observed between

the two listener groups in terms of the ratings assigned to AS_early (p = .15) and those assigned to the monolingual control speakers (p = .05), i.e., both monolingual and bilingual listeners perceived AS_early and the monolingual speakers to sound equally native in their German pronunciation.

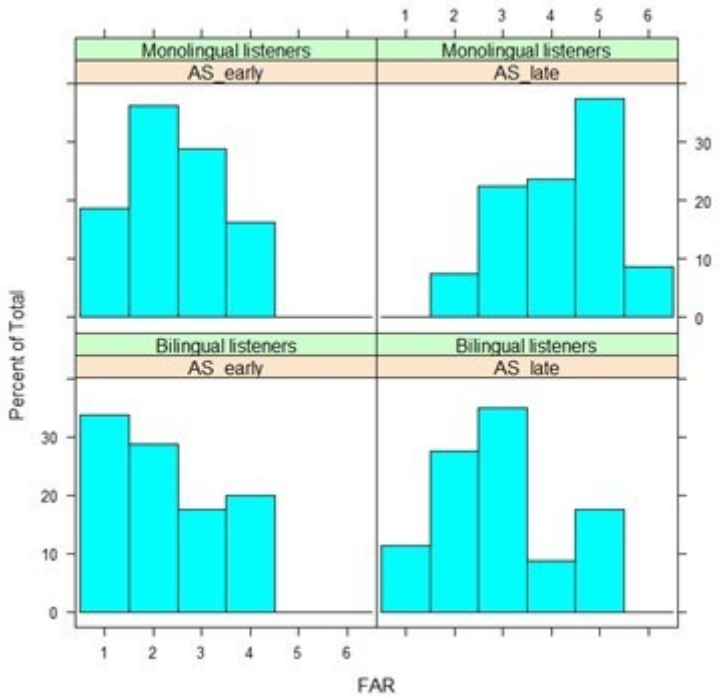


Figure 4.4: Distribution of nativeness ratings for AS_early and AS_late according to listener group.

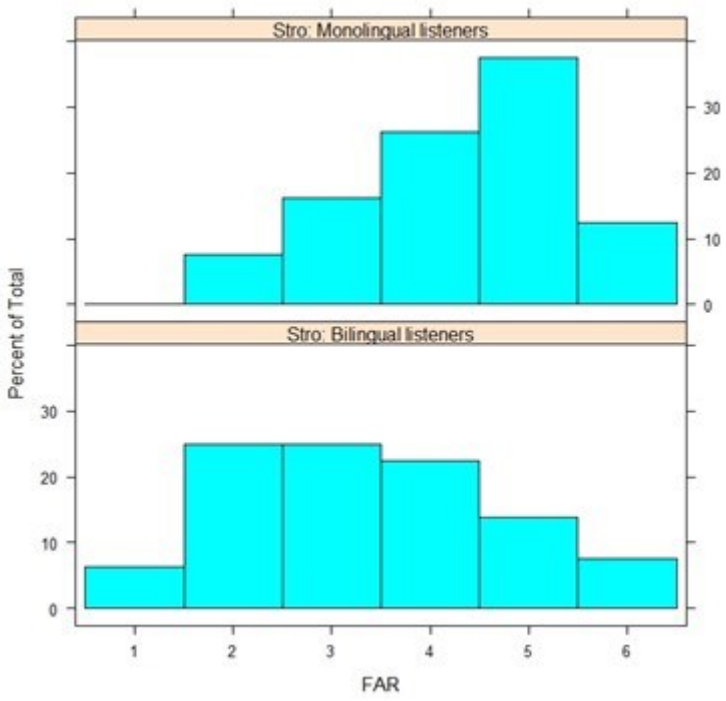


Figure 4.5: Distribution of nativeness ratings for Stronach according to listener group.

(II) Comprehensibility and intelligibility ratings (COMP, INTEL)

The results for the COMP and INTEL ratings will be first reported individually. In a next step, the results will be related to the findings of the nativeness ratings previously described in order to determine if there are any significant correlations between the degree of perceived nativeness and the degree of comprehensibility and intelligibility of L1 pronunciation (RQ 3.3). Inter-rater reliability for both INTEL and COMP was high, with a Cronbach alpha coefficient of 0.84 and 0.86, respectively.

COMP was rated on a six-point scale, ranging from 1 = *very easy to understand* to 6 = *very difficult to understand* and, as can be seen in Figure 4.6, all bilingual speakers (AS_early, AS_late, Stro) and the monolingual control group were assigned similar ratings in terms of their comprehensibility. All groups, bilinguals as well as controls, received a median rating of 2.0, i.e., all speakers were rated as being easy to understand. One monolingual speaker (ML1), however, received slightly higher COMP scores (*Mdn* = 3.0, maximum score = 5.0) compared to the remaining speakers, suggesting that this speaker was less comprehensible compared to the other speakers.

The INTEL scores (see Figure 4.7) were calculated for each speaker as the percentage of words transcribed correctly per utterance. The scores ranged from 93.1% to 99.0% (*Mdn* = 100), with only minor differences between the scores assigned to the individual speakers. In order to determine if and to what extent AS's perceived nativeness is related to the intelligibility and comprehensibility of his L1 speech, a third ordinal regression model was specified. This model included *FAR* as the dependent variable and *INTEL* scores as well as a three-way interaction between *COMP*, *Group* and *Speaker* as fixed effects. As in the previous models, a by-*Listener* random intercept was included (FAR~INTEL + COMP*Group*Speaker + (1|Listener)). As an inspection of Figures 4.6 and 4.7 already suggested, no main effects were found for INTEL (p = .29) and COMP (p = .23). This suggests that the degree of perceived nativeness was not associated with how intelligible and comprehensible the speakers were perceived to be.

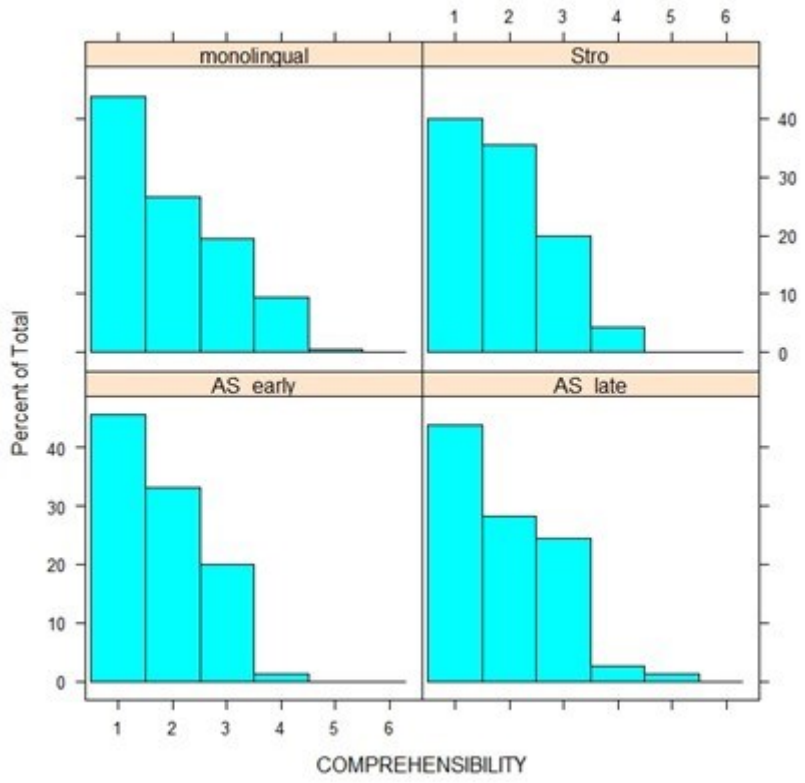


Figure 4.6: Distribution of comprehensibility ratings according to speaker.

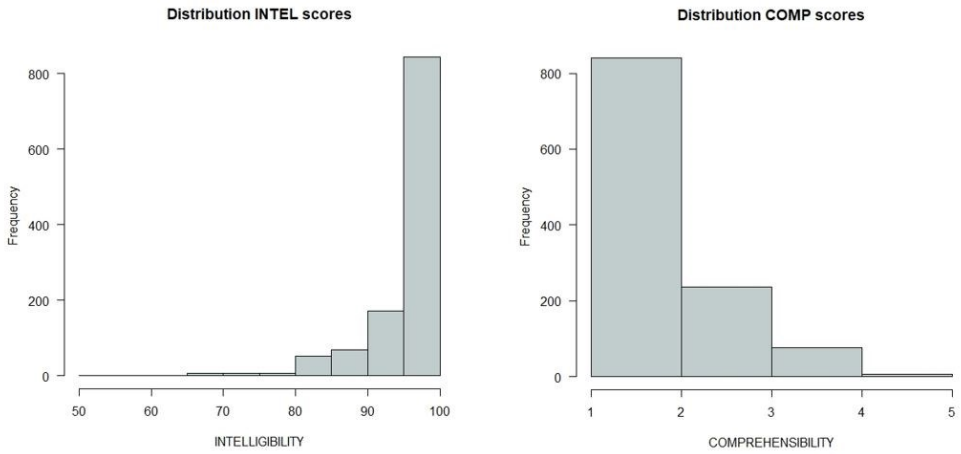


Figure 4.7: Distribution of intelligibility (100% being *completely intelligible*) and comprehensibility (1 being *very easy to understand*) scores.

4.2.4 Discussion of Study III

The aim of Study III was to determine if AS's is perceived to sound less native in his late L1 German speech than in his early L1 speech compared to monolingual Austrian German control speakers. Furthermore, the present study set out to assess if and to what extent the listeners' linguistic background (monolingual vs. bilingual) affects their perception of nativeness in AS's early and late German pronunciation. A third objective was to find out to what extent perceived nativeness is related to the intelligibility and comprehensibility of L1 speech in bilingual speakers. In this context, intelligibility was defined as the extent to which a speaker's utterance is understood, while comprehensibility was specified as the ease of understanding a speaker's utterance (Derwing and Munro 1997; Munro and Derwing 1995a).

Perceived nativeness of L1 speech

While there is abundant evidence which shows that the majority of late L2 learners maintain a detectable nonnative accent in their L2 speech (e.g., Elliott 1995; Flege 1984; Moyer 1999; Munro and Derwing 1995a; Munro, Derwing, and Morton 2006; Saito, Trofimovich, and Isaacs 2016), only little research so far has focused on the question of whether L2-immersed bilinguals may be perceived as having a nonnative accent in their first language (Bergmann et al. 2016; de Leeuw, Schmid, and Mennen 2010; Hopp and Schmid 2013). In fact, research shows that L2-immersed bilinguals are often perceived to sound less native in their L1 compared to monolingual speakers of the same language as a result of long-term exposure to the

L2 (Bergmann et al. 2016) and reduced contact with and exposure to the L1 (de Leeuw, Schmid, and Mennen 2010).

The results of the present investigation provide further evidence for the malleability of native speech in adult L2 learners as a consequence of long-term immersion in an L2 environment. AS's late speech was perceived to sound less native compared to non-immersed monolingual Austrian German speakers, who were consistently rated within the native German range. AS's early German pronunciation, by contrast, was perceived to sound more native compared to his late L1 pronunciation but, at the same time, his early speech still received higher FARs compared to the monolingual controls. This might be due to the fact that the early samples included in the present study were recorded approximately ten years after AS had moved to the United States, i.e., he had already gained a considerable amount of L2 experience, which is likely to have influenced his L1 pronunciation at an earlier stage of L2 development.

Interesting observations were made in terms of the rating differences between Austrian German monolingual listeners and German-English bilingual listeners: The monolingual listeners were more likely to perceive AS's late pronunciation as sounding nonnative compared to the bilingual listeners; by contrast, no significant differences were found between the two groups concerning the FARs assigned to AS's early German pronunciation. Similarly, the control bilingual, Frank Stronach, was rated to sound overall less native by the monolingual listeners compared to the bilingual listener group who were more likely to rate him as sounding (near-)native. These findings support Long's (1990) proposition that linguistically experienced listeners are likely to be more lenient towards accented or nonnative speech while listeners who lack linguistic experience with a second language and who are predominantly exposed to their first language can be assumed to be less tolerant when it comes to nonnative speech.

Except for the present investigation, there are no empirical studies to date which specifically test the effect of linguistic experience on nativeness ratings in a speaker's first language. Research so far has focused on examining effects of listener experience on perceived accentedness and comprehensibility in the nonnative speech of L2 learners. Kennedy and Trofimovich (2008), for instance, explored how native English listeners' experience with a particular nonnative L2 accent, in this case Mandarin-accented English, influences their perception of accentedness, intelligibility, and comprehensibility. In this context, listener experience was defined as the "extent of previous exposure to nonnative speech" (Kennedy and Trofimovich 2008, 459). Their findings showed a positive correlation between listener experience and intelligibility scores, i.e., listeners who were more experienced with nonnative L2 speech were more accurate in their intelligibility transcriptions compared to less experienced listeners. However, no significant

differences between experienced and inexperienced listeners were found in terms of comprehensibility and accentedness ratings, that is, linguistic experience did not affect listeners' comprehensibility and accent judgments. Flege and Fletcher (1992) also compared the accent ratings obtained from inexperienced and experienced listeners, *experienced* being defined as the listeners' familiarity with nonnative speech. Similar to Kennedy and Trofimovich (2008), their findings did not show a significant effect of listener experience on accent ratings. Thompson (1991), by contrast, found a clear effect of linguistic experience, which reflects the findings of the present study: Experienced raters, i.e., listeners who were fluent in an L2 and had frequent contact with nonnative speakers, judged the English speech samples produced by L1 Russian speakers of L2 English as being less accentedness compared to the inexperienced raters, who generally perceived the samples to sound more foreign-accented. The differences in the findings outlined above might be related to differences in defining and selecting experienced vs. inexperienced listeners. Kennedy and Trofimovich (2008) and Flege and Fletcher (1992) included native English subjects as *experienced* listeners who had been previously exposed to nonnative English speech but did not actively use a second language. By contrast, the experienced subjects in Thompson (1991) – similar to the subjects in the present study – were fluent speakers of an L2, as opposed to inexperienced listeners who neither spoke a second language nor had contact with nonnative speakers. Hence, it can be argued that mere exposure to nonnative speech as such does not necessarily increase a listener's tolerance for accented or nonnative speech while the active and frequent use of a second language in different communicative contexts is likely to make listeners more hesitant when it comes to judging a speaker as nonnative.

The present findings suggest that the extent to which a speaker is perceived to sound nonnative is likely to be not only affected by features of the speech itself, that is, by linguistic deviations from expected speech patterns, but also by characteristics on the part of the listener, including, for instance, their language background and linguistic experience.[101] As Reinisch (2005, 82) succinctly summarizes, "[f]oreign accent is not only the way the learners *produce* the L2, but also what the native speakers of the target language *perceive* as such" (my italics), which also holds true for L1 pronunciation as the findings of the present study suggest. Taken together, these findings add a new dimension to the question of *who* is to be considered a native speaker. From a linguistic point of view, as discussed by Davies (2003, 2004),

101 Alongside differences in linguistic experience, there are additional listener-specific factors which have been shown to influence the perception and judgment of nativeness. Munro, Derwing, and Morton (2006), for example, point out that listeners may be prejudiced against specific nonnative accents, which might make them rate a speaker as sounding less native compared to listeners who have positive attitudes towards a particular accent (e.g., Beinhoff 2013; Brennan and Brennan 1981; Kraut 2014; Kraut and Wulff 2013).

a native speaker is often regarded as being a "standard setter" and a "repository and guardian of the true language" (2004, 447). The subject of the present investigation, however, seems to have 'lost' some of what marks him as being clearly identifiable as a native speaker of his first language. Furthermore, the degree to which this 'loss' of native pronunciation abilities is perceived and rated as nonnative also depends on listener-specific characteristics as previously outlined. Hence, the question arises if AS is still considered to be a native speaker in the sense of being a 'standard setter' and a model for a 'true' language, bearing in mind that 'true' in this context is of course a rather abstract term to use, reflecting a somewhat idealized view of language. The implications which the findings of the present study and the results obtained in Studies I and II have on the notion of *the native speaker* will be further addressed in Chapter 5.

Perceived comprehensibility and intelligibility of L1 speech

Alongside examining perceived nativeness, Study III set out to investigate perceived comprehensibility and intelligibility of AS's German speech compared to the speech produced by monolingual German controls. The findings of the COMP and INTEL ratings show that the speech of all subjects, including both monolinguals and bilinguals, was consistently perceived to be clearly comprehensible and intelligible. As mentioned in Section 4.1.2, to date no research has been conducted with the aim to examine perceived comprehensibility and intelligibility in the L1 speech of L2-immersed late consecutive bilinguals. Munro (2008, 203) states that while accented L2 speech is a frequently observed phenomenon – particularly in late L2 learners – completely unintelligible L2 speech is rarely observed. The findings of the present study suggest that being unintelligible or incomprehensible in one's first language is perhaps even less common. Hence, similar to having a nonnative accent in the L2, being perceived as less native in the L1 does by no means impede communication which is determined by both intelligibility and comprehensibility. Given the clear results obtained in the COMP and INTEL ratings, one might ask whether it is possible at all that a healthy speaker, who does not suffer from any kind of speech impairments, is unintelligible and/or incomprehensible in their L1 speech. Interestingly, however, one monolingual Austrian German speaker was perceived to be slightly less comprehensible in his German speech compared to the other monolingual speakers while, at the same time, being consistently rated as sounding clearly native in the FAR. A possible explanation for this observation comes from some of the listeners' comments made in the nativeness rating task.[102] Three monolingual listeners and two bilingual listeners perceived this speaker as having a "heavy

102 As described in Section 4.2.2, listeners were invited to write down comments whenever they noticed something about a speaker, particularly about his pronunciation.

Styrian dialect" and sounding "as if he lives in the countryside".[103] As mentioned in Section 4.2.2, all listeners described themselves as dialect speakers and reported that they were familiar with the Styrian dialect, being able to distinguish it from other varieties spoken in Austria. However, with a mean age of approximately 26 years, the listeners in both groups were considerably younger than the monolingual control speakers (> 70 years). Elderly speakers are generally more likely to rely on dialectal, non-standard pronunciation features compared to younger speakers, especially if they live in rural areas of Austria and have less frequent contact with standard speakers living in urban centers (e.g., Kleber 2018, 469). Consequently, it might have been the case that the speaker's strong dialect made him slightly less comprehensible for the listeners in this study – an assumption which, of course, requires further investigation.

103 In the original comments, the listeners wrote "der Sprecher hat einen ziemlich starken Dialekt" ('the speaker has a very strong dialect') and "er klingt, als käme er irgendwo vom Land" ('he sounds as if the comes somewhere from the countryside'). One listener commented that he sounded 'typically Styrian' ("der Sprecher klingt typisch steirisch").

5 Limitations and overall discussion

5.1. Limitations

As outlined in Chapter 2, one of the risks of using pre-recorded materials for phonetic-acoustic analyses is that certain aspects are difficult to control for, such as the number of tokens extracted from the individual recordings. This resulted in a considerably lower number of plosive and vowel tokens which could be analyzed to represent AS's early and late German pronunciation compared to the number of tokens used to represent his L2 English pronunciation across three stages. One reason for this was the poor quality of some of the recordings, which made it impossible to obtain an equal or at least balanced number of observations to represent his English and German pronunciation. In addition, there are only few interviews available in which AS's actually speaks German because, as discussed in Section 2.3.2.1, he prefers giving interviews in his second language English. However, as AS has not yet retired and is still present in the media, there might be more material available in the future which can be used to examine further segmental as well as prosodic features of his first and second language pronunciation.

A further limitation of the acoustic studies is the lack of reference data to compare AS's L1 vowel and plosive productions with. Using a single control speaker, in this case a male speaker from Thal, can, of course, not be considered sufficient to establish something like a norm – if this is possible at all, as will be discussed in the subsequent section. Consequently, follow-up studies examining AS's first language speech production will need to compare his pronunciation to a larger corpus of reference data collected from male Thal-speakers.

Moreover, no reference data could be included to compare AS's early German speech production with. Ideally, these reference data would have been taken from recordings made in the 1970 and early 1980s including age-matched speakers from Thal. Such recordings were, however, not available and it might even be the case that they do not exist at all given that 40 years ago people usually did not have the technical prerequisites to produce audio and video recordings and make these recordings accessible via the internet as it is the case nowadays.

Similarly, the reference data used to compare AS's English pronunciation with are perhaps not ideal. As stated in Section 2.2, English is among the most frequently examined languages and numerous acoustic studies have been conducted to obtain data about phonetic characteristics of English. However, it is debatable if the native American English VOT and vowel data used in Studies I and II are appropriate for a comparison given that the speakers were not age-matched and that the data were obtained from both male and female speakers. Hence, in further investigations, more appropriate reference data will have to be collected, including speech

recordings from male speakers of Californian English who are approximately the same age as AS.

With regard to Study III, potential range effects need to be taken into consideration when interpreting the results of the nativeness ratings: As observed by Flege and Fletcher (1992, 370) and as mentioned in Section 4.1.3, the ratio of native and nonnative speech samples can have an effect on listener responses. That is, the larger the number of (monolingual) native speakers included, the more likely are listeners to rate speech samples produced by nonnative speakers as foreign accented (see also Schmid and Hopp 2014). Although the actual number of samples included to represent potentially attrited speech (n = 12) and monolingual speech (n = 17) was relatively balanced in the present study, the number of monolingual speakers (n = 7) still prevailed. Therefore, it cannot be ruled out that range effects influenced listener responses to some extent, i.e., listeners might have been more likely to rate the bilingual speakers as sounding nonnative because of the relative lower number of bilingual samples included. In order to control for potential range effects, a replication of Study III would ideally include more samples of bilingual L1 speech from different speakers.

One question which is often raised when conducting studies examining a single subject is the extent to which the findings of such investigations are representative of a larger population of speakers. Tracing the trajectory of first and second language speech development in an individual speaker does certainly not allow for generalizable findings; it does, however, offer a valuable starting point for further research to explore if and to what extent similar observations can be made in terms of other speakers with the same language learning history and language background. It should be noted, however, that also the generalizability of findings obtained in studies examining larger populations of speakers is debatable given that even in seemingly homogenous groups of subjects, intra- and inter-speaker variability is one of the main characteristics (e.g., Ellis 1991, 10). Hence, the possibility that speakers behave in exactly the same way is rather low, which perhaps even implies a necessity to focus on individual language development first before trying to make general statements about bilingual speech development. This is in line with Flege and Bohn's (2021, 58) call to focus more closely on *individual* learners when it comes to examining developmental processes of speech production, based on the observation that it might be difficult – or perhaps even impossible – to find learner groups which effectively differ in a single variable. It is due to this potential inter-learner variability that group data do not necessarily allow for meaningful conclusions to be drawn.

5.2. Overall discussion

The overall objective of the three studies presented in this work was to provide an in-depth longitudinal investigation of first and second language speech development in the late consecutive bilingual Arnold Schwarzenegger, who has been long-term immersed in an L2-speaking environment. To this end, his L1 German and L2 English segmental speech productions were examined in acoustic analyses focusing on the realization of VOT contrast in word-initial plosives (Study I) and on the first and second formant in his productions of monophthongs (Study II). Furthermore, it was explored if and to what extent his early and late L1 pronunciation is perceived to sound less native compared to monolingual Austrian German speakers (Study III). The three studies were based on broadcast recordings covering a period of approximately 40 years. The recordings were divided into individual stages to represent AS's early (1970s/1980s) and late (2010s) L1 and L2 pronunciation. For his L2 English, an additional set of samples was included to represent his pronunciation in the 1990s and early 2000s (mid stage).

The findings of Study I revealed a change in AS's productions of L1 German /t/ in the direction of the L2 counterpart in the late stage, which manifested itself in longer – and thus more English-like – VOT values measured in his late German /t/. At the same time, changes were observed in terms of his realization of L2 English /p/ and /t/, which were significantly less aspirated in the late stage compared to the early stage. Taken together, these findings provide evidence for bidirectional interactions operating between the speaker's L1 and L2 system, which resulted in an assimilation of L1 and L2 categories in the late stage, i.e., the first and second language targets came to resemble each other. In addition to a partial L1–L2 category assimilation of VOT, the findings of the study indicated that AS's productions are characterized by a considerable amount of variability, which became particularly evident in his L2 English. This shows that he has partly acquired and implemented an L2 contrast which does not exist in his L1 Austrian German – at least in spontaneously produced speech (Hödl 2019; Moosmüller, Schmid, and Brandstätter 2015; Wiesinger 2008). At the same time, he does not manage to produce this contrast consistently, that is, he neutralizes contrast in some plosive tokens, as typical of his L1 German.

The findings of Study II are more ambiguous when it comes to the identification of potential patterns in the changes observed. While some of AS's L1 German vowels seem to have moved closer to L2 English targets, most of his L2 vowels were observed to have moved further away from L2 production norms in the late stage. Similar to the findings of Study I, no overall improvement of L2 production abilities was shown in the late stage, but the findings point to a development which goes in the opposite direction, that is, AS seems to be less native in his late L2 English

productions, at least in terms of the sounds examined in the present investigation. Furthermore, the decrease in F1 in both his L1 and L2 vowels cannot be exclusively explained against the background of system-internal processes, but might also be the result of biological ageing mechanisms, which have been shown to predominantly affect F1 (Reubold and Harrington 2015, 2017; Reubold, Harrington, and Kleber 2010).

Study III showed that on the level of global accent, AS was perceived to sound less native in his late L1 German pronunciation compared to monolingual control speakers, which provides further evidence for the malleability of the native language system, in line with previous findings (see Bergmann et al. 2016; de Leeuw, Schmid, and Mennen 2010; Hopp and Schmid 2013). It remains unclear, however, which specific segmental and/or prosodic features of AS's pronunciation have contributed to being perceived as less native. That is, the divergences from native norms identified in the acoustic studies do not necessarily also make him being perceived as less native, but there might be additional pronunciation-related characteristics which, in combination, have influenced the listeners' judgments. This issue was, for instance, addressed by Bergmann et al. (2016) who predicted that those German-English bilinguals who were judged as sounding less native in their L1 German would also show deviances on the segmental level in their productions of specific German vowels (/aː, ɛ, ɔ/) and the lateral approximant /l/. Contradictory to this prediction, their findings did not reveal significant correlations between perceived (non-)nativeness on the one hand, and deviations on the segmental level on the other hand. It should be noted, however, that Bergmann et al. (2016) examined three vowels and one consonant only, which is of course not representative of a speaker's entire inventory. Similarly, the segmental features explored in the present investigation do not allow for predictions to be made about AS's overall pronunciation performance in both of his languages. In addition, the extent to which listeners perceive and assess system-internal divergences from an expected pattern is determined by their personal linguistic background and experience, as has been shown in Study III. The observation that *accent* is not only a phenomenon which is quantifiable by means of acoustic analyses, but is also determined by listener-specific factors does, without doubt, further contribute to the complexity and multidimensionality of accented speech.

A dynamic-systems oriented perspective

The observations that, first of all, AS's L1 segmental speech productions have to some extent moved away from native German production norms and, second, that his L1 pronunciation is perceived as sounding less native when being compared to monolingual German speakers, contradict the notion that the L1 system, once fully developed, is static and not prone to be affected by changes, as proposed in the

context of early theories on maturational constraints and their impact on language acquisition (e.g., Lenneberg 1967; Scovel 1969, 1988). In the same vein, the findings showed that also a bilingual's L2 system is not stable but likely to undergo modifications – manifested in "periods of regression and progression" (Van Dijk, Verspoor, and Lowie 2011, 58) – in the course of the acquisition process. Hence, they support a dynamic view of language learning and development as represented by the DST approach outlined in Section 1.5 (de Bot and Larsen-Freeman 2011; de Bot, Lowie, and Verspoor 2007; Larsen-Freeman 1997, 2000; van Geert 2009). This approach acknowledges the occurrence of variability and inconsistent production patterns, which could be observed in AS's segmental speech productions, as inherent characteristics of language acquisition, resulting from continuous changes in the developing language system (e.g., de Bot, Lowie, and Verspoor 2007). A further dimension of DST which is reflected by the current findings is that not all changes observed in a speaker's language systems can be traced back to system-internal bidirectional influences, but some changes might also be explained against the background of developmental processes, such as biological ageing mechanisms. This supports one of the main tenets proposed by DST, namely that there is no single factor but rather a combination of multiple internal and external factors which influence speech development. Although the three studies, particularly Study I and II, did not focus on examining additional speaker-related factors, such as level of motivation or frequency of L1 and L2 use – simply because it was not possible to collect reliable data to address these factors – they should not be completely disregarded when discussing AS's speech development from a dynamic systems point of view. Referring to anecdotal evidence, it is possible to identify a change in AS's attitude towards having a nonnative accent in his L2. While his strongly foreign-accented English pronunciation was an obstacle in the early stages of his career in the United States (Miller 2012), he soon acknowledged and appreciated his nonnative accent as part of his identity and, not least, part of his success. At the same time, his L2 became his preferred means of communication in professional contexts while the use of his L1 was and, presumably, still is restricted to the private domain (Gersemann 2009; von Uslar 2012). Alongside being immersed in an L2-speaking environment, the reduced use might be one factor which has, among others, led to the changes observed in AS's L1 plosive productions in the direction of L2 norms and, at the same time, to being perceived as less native on a global level. The observation that particularly AS's productions of voiceless plosives in his L2 has become less native-like in the late stage compared to earlier productions, as indicated by an overall decrease of aspiration, is more difficult to explain based on his language use and based on motivational factors. As discussed in Section 2.3.5, AS was highly motivated to lose or at least reduce his foreign accent in the first couple of years after migrating to the United States, since this was an essential part of starting his career

in the film business (Miller 2012; Schwarzenegger 2012). Consequently, his endeavors to work on his accent might have indeed led to an improvement of L2 pronunciation abilities in the early stage, as was observed in Study I. In the further course of L2 acquisition, it seems that AS did not invest as much time and effort into working on his pronunciation, perhaps – as mentioned above – as a consequence of starting to become recognized as a successful actor with a very unique and memorable accent (Daily Mail UK 2015). However, the extent to which external factors, such as L1 and L2 use, changing attitudes, and the motivation to be able to speak the L2 in a native manner, have influenced AS's speech development cannot be reliably determined at this point.

What further accounts for a combination of various factors and, thus, a multidimensional view on language development in the context of the present work is the finding that not only language- and speaker-specific features play a role when it comes to investigating the phenomenon of foreign accent, but that also characteristics on the part of the listener are crucial for an evaluation of potentially accented speech. In addition, this observation challenges a traditional understanding of *the native speaker* as will be discussed in the following section.

Although a dynamic systems-oriented perspective is proposed in the context of this work, it must be noted that examining a few specific features of language production only, as has been done in the present investigation, does not comply with the demand of DST that the *full complexity* of developmental processes can only be accounted for by including and examining *all factors* which potentially influence these processes. However, it is not only due to the limited scope of the present investigation, but also due to methodological issues that this requirement cannot be fulfilled, and it is questionable whether it is possible to meet it at all (see de Bot, Lowie, and Verspoor 2007). Hence, it is acknowledged that the findings of the present investigation can only partly contribute to understanding the full complexity of bilingual language development.

To be or not to be a native speaker?

As mentioned in Section 4.2.4, the findings of the three studies conducted in the context of the present longitudinal investigation lead to the need to reassess the concept of *the* native speaker. If the only requirement a native speaker has to fulfil in order to be defined as such is to have acquired a specific language from birth (e.g. , Mack 1990, 115), then the subject of the present investigation would certainly qualify as a native speaker in his first language. If, however, a native speaker is defined against a monolingual norm or a monolingual standard and is described as having complete and perfect linguistic knowledge, AS would disqualify as a native speaker not only in his first but also in his second language. He has not maintained a distinctly native accent in his first language and, at the same time, has not

achieved native pronunciation abilities in his second language; hence, he does not fulfil native-speaker requirements in the strictest sense. The observation that the perception of nativeness also depends on listener-related factors, such as their personal linguistic background, does not only further stress the ambiguity of the native speaker concept, but also raises the question of whether it is possible or useful at all to provide a one-fits-all definition of the native speaker. Any attempt to provide such a definition would have to take the perspective of the listener/interlocutor into consideration as well as the possibility that different listeners have different perceptions of who is a native speaker. Furthermore, the question arises of how a definition of nativeness can account for potential discrepancies between the findings of acoustic analyses on the one hand, and listener perceptions on the other hand. As mention above, these two dimensions do not necessarily correlate (see Bergmann et al. 2016), that is, a speaker's segmental productions might be identified as 'perfectly' native – in these sense of corresponding to the respective native target – in an acoustic analysis, but he or she might be still perceived as sounding nonnative. What or who determines nativeness in this case? From a standpoint which stresses the relevance of spoken language as the predominant means of verbal communication in real-life contexts, it is certainly the interlocutors' perceptions which determine if and to what extent a speaker is a native speaker – but yet again, not all listeners inevitably share the same perception.

While an attempt to define a native speaker seems more straight-forward when dealing with monolingual speakers, the fragility of the native speaker concept becomes more apparent when contextualizing it within bilingual acquisition and development, as the findings of the present investigations revealed. This aspect has been previously addressed by Cook (1991, 2003) and his notion of *multicompetence*, which he defines as "the knowledge of two or more languages in one mind" (Cook 2003, 2). The observation that a bilingual speaker's two (or more) languages do not exist in isolation but are inherently connected challenges the notion that a 'true' bilingual is the equivalent of two monolingual native speakers, as proposed in the context of the monolingual view of bilingualism (e.g., Bloomfield 1933; see Grosjean 1989, for a discussion).

One problematic aspect which goes along with investigating bilingual speech development is that the majority of empirical studies – the present study being no exception – take the monolingual norm as a "benchmark of true nativeness" (Rothman and Treffers-Daller 2014, 93). That is, bilingual speakers are compared with and evaluated in relation to monolingual speakers of the respective languages, who represent the 'native standard' and are thus assumed to offer reliable reference data. However, if bilingual speakers are described as having a "unique and specific linguistic configuration" (Grosjean 1989, 3), why should their language performance then be compared to the performance of monolingual speakers who have a

different linguistic configuration and are, therefore, also likely to show a different linguistic behavior?

Another dimension which should not be ignored in a discussion of the concept of the native speaker relates to social and personal consequences it might have for speakers if they are perceived as diverging from the idealized image of *the* native speaker. As illustrated in Section 1.8, a nonnative accent can lead to discrimination and negative evaluations (Lippi-Green 2003), which might even go as far as being excluded from the job market (Munro 2003) as well as losing a sense of belonging to both one's L1 and L2 speech communities. Therefore, it should be taken into consideration that each description of and attempt to explain the phenomenon of foreign accent and each re-evaluation of the native speaker concept should go beyond theoretical discussions in that it needs to be acknowledged that the extent to which individuals are judged to conform to an expected standard of nativeness entails real-life consequences for speakers engaging in real communicative situations.

Based on the previous discussion, one question remains yet to be answered: Is AS (still) a native speaker in his first language? The answer proposed in the present investigation is that he is *not a nonnative speaker*, which is in line with Davies' (2004, 434) definition *ex-negativo* "[t]o be a native speaker means not being a nonnative speaker", presumably resulting from a failure – or perhaps an inappropriateness – to provide a straight-forward definition of the native speaker, which, as the previous discussion suggests, might not even exist. An alternative answer to the aforementioned question is that AS is a native speaker to some extent, in the same vein as he is a bilingual to some extent, keeping in mind that any definition of who is (not) to be considered a bilingual inevitably also entails the question of how to define a native speaker. Perhaps, the fact that we live in a world where it is more common to speak more than one language (e.g., Aitchison 1994) implies that the native speaker as such does not exist, neither is there something as 'true' or 'perfect' nativeness as expressed in completely accent-free speech. What this illustrates is a need to move away from a static image of nativeness towards acknowledging the multidimensionality and dynamics of not only pronunciation-related features but also of language development in general.

5.3. Final remarks and outlook

As pointed out by Ellis (1994, 316), "[t]he existence of 'foreign accents' in L2 learning is so well attested that it hardly requires documenting." Undoubtedly, foreign-accented L2 speech is a phenomenon which can be frequently observed among late second language learners and has been extensively investigated in L2 studies (see Sections 2.2.2 and 3.1.3). In this respect, AS's language background and the fact that he acquired an additional language relatively late in life upon migrating to another

country, does certainly not represent a unique or unusual case of language development given that migration, bilingualism and language contact are nowadays the norm rather than being an exception. However, in order to understand and explain language development as a highly dynamic process – as proposed in the context of the present work – it is not only necessary to document and examine foreign-accented second language speech, but also to explore potential influences L2 learning experience can have on first language pronunciation abilities. The present investigation has shown that the causes and effects of accented L1 and L2 speech are so complex that they – contradictory to Ellis' statement – require both thorough documentation and investigation. Furthermore, regarding language acquisition as a developmental process *per se* entails the demand for longitudinal investigations of bilingual speech, which are still relatively rare. In this respect, the present investigation might offer an incentive for future studies to overcome the challenges which are usually linked to collecting longitudinal data and make use of what is already available in the form of, for example, broadcast recordings.

Lastly, despite the fact that foreign-accented speech is by no means an unusual phenomenon, it still entails a range of questions which have not been sufficiently answered yet and hence require further investigation: Which specific segmental and prosodic features contribute to a bilingual speaker being perceived as (non-)native in one or both of their languages? Why are some segmental and prosodic features more likely to be affected by attrition processes than others – and why are certain L2 pronunciation features easier to acquire than others? Why are some speakers more successful in maintaining native L1 production abilities while others are less successful in doing so? To what extent do extra-linguistic factors contribute to these individual differences? Which effects does it have on an individual's personal experience of being bilingual in everyday life if they are not perceived as native speakers in both of their languages?

Future studies might address and, eventually, answer these questions by more closely focusing on speakers' instances of language production in real-world contexts. Laboratory phonetic experiments have, without doubt, many advantages in terms of producing 'clean' data; the speech elicited in an experimental context is, however, not entirely representative of authentic speech. Hence, if we are interested in an individual's linguistic behavior in real communicative contexts, it seems plausible to also make use of speech data which are taken from real-world communicative situations.

6 Literature

Abdelli-Beruh, Nassima B. 2004. “The Stop Voicing Contrast in French Sentences: Contextual Sensitivity of Vowel Duration, Closure Duration, Voice Onset Time, Stop Release and Closure Voicing.” *Phonetica* 61 (4): 201–19. doi:10.1159/000084158.

Abercrombie, David. 1949. “Teaching Pronunciation.” *English Language Teaching* 3 (5): 113–22. doi:10.1093/elt/III.5.113.

Abramson, Arthur S., and Douglas H. Whalen. 2017. “Voice Onset Time (VOT) At 50: Theoretical and Practical Issues in Measuring Voicing Distinctions.” *Journal of Phonetics* 63:75–86. doi:10.1016/j.wocn.2017.05.002.

Absolute Radio. 2014. “Schwarzenegger Interview 2014 [Video File].” Accessed April 04, 2020. https://www.youtube.com/watch?v=yBpPtglHk94.

Aitchison, Jean. 1994. *Words in the Mind: An Introduction to the Mental Lexicon.* 2nd ed. Oxford, Cambridge, MA: Blackwell.

Aliaga-García, Cristina, and Joan C. Mora. 2009. “Assessing the Effects of Phonetic Training on L2 Sound Perception and Production.” In *Recent Research in Second Language Phonetics/Phonology: Perception and Production*, edited by Michael A. Watkins, Andreia S. Rauber, and Barbara O. Baptista, 2–31. Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

Allen, Elaine, and Christopher A. Seaman. 2007. “Likert Scales and Data Analyses.” Accessed June 24, 2020. http://rube.asq.org/quality-progress/2007/07/statistics/likert-scales-and-data-analyses.html.

Allen, Sean J., Joanne L. Miller, and David DeSteno. 2003. “Individual Talker Differences in Voice-Onset-Time.” *Journal of the Acoustical Society of America* 113 (1): 544–52. doi:10.1121/1.1528172.

Almberg, Jørn, and Olaf Husby. 2000. “The Relevance of Some Acoustic Parameters for the Perception of a Foreign Accent.” In *New Sounds 2000: Proceedings of the 4th International Symposium on the Acquisition of Second-Language Speech*, edited by Allan James and Jonathan Leather, 1–10. Klagenfurt, Austria: University of Klagenfurt.

Altenberg, Evelyn P. 1991. “Assessing First Language Vulnerability to Attrition.” In Seliger and Vago 1991, 189–206.

Ammerlaan, Tom. 1996. “'You Get a Bit Wobbly...' Exploring Bilingual Lexical Retrieval Processes in the Context of First Language Attrition.” PhD thesis, Katholieke Universiteit, Nijmegen, Netherlands.

Andersen, Roger W. 1982. "Determining the Linguistic Attributes of Language Attrition." In *the Loss of Language Skills,* edited by Richard D. Lambert and Barbara F. Freed, 83–118. Rowley: Newbury House Publishers.

Anderson-Hsieh, Janet, and Kenneth Koehler. 1988. "The Effect of Foreign Accent and Speaking Rate on Native Speaker Comprehension." *Language Learning* 38 (4): 561–613. doi:10.1111/j.1467-1770.1988.tb00167.x.

Aoyama, Katsura, James E. Flege, Susan G. Guion, Reiko Akahane-Yamada, and Tsuneo Yamada. 2004. "Perceived Phonetic Dissimilarity and L2 Speech Learning: The Case of Japanese /r/ and English /l/ and /r/." *Journal of Phonetics* 32 (2): 233–50. doi:10.1016/S0095-4470(03)00036-6.

ARD Mittagsmagazin. 2012. "Österreich: Milliardär Stronach Will Politik Aufmischen." Accessed May 08, 2020. https://www.youtube.com/watch?v=PhQUGc6nlHI.

Ashby, Michael, and John Maidment. 2013. *Introducing Phonetic Science.* 8th ed. Cambridge, New York: Cambridge University Press.

Auer, Peter. 1995. "Modelling Phonological Variation in German." In Werlen 1995, 9–37.

Bachman, Lyle F. 1982. "The Trait Structure of Cloze Test Scores." *TESOL Quarterly* 16 (1): 61–70. doi:10.2307/3586563.

Baken, Ronald J., and Robert F. Orlikoff. 2000. *Clinical Measurement of Speech and Voice.* 2nd ed. San Diego: Singular Publ. Group.

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): 1–27. doi:10.1177/00238309050480010101.

Baker Smemoe, Wendy, and Naomi Haslam. 2013. "The Effect of Language Learning Aptitude, Strategy Use and Learning Context on L2 Pronunciation Learning." *Applied Linguistics* 34 (4): 435–56. doi:10.1093/applin/ams066.

Baladzhaeva, Liubov, and Batia Laufer. 2018. "Is Language Attrition Possible Without Second Language Knowledge?" *IRAL: International Review of Applied Linguistics in Language Teaching* 56 (2): 103–36. doi:10.1515/iral-2016-0066.

Bartram, Dave. 2016. "Likert Scales." Accessed February 20, 2020. http://my.ilstu.edu/~eostewa/497/Likert%20topic-dane-likert.pdf.

Beinhoff, Bettina. 2013. *Perceiving Identity Through Accent: Attitudes Towards Non-Native Speakers and Their Accents in English.* Studies in Descriptive Linguistics 35. Oxford: Peter Lang.

Benjamin, Barbaranne J. 1982. “Phonological Performance in Gerontological Speech.” *Journal of Psycholinguistic Research* 11 (2): 159–67. doi:10.1007/BF01068218.

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. doi:10.1016/j.wocn.2016.07.001.

Bergmann, Christopher, Simone A. Sprenger, and Monika S. Schmid. 2015. “The Impact of Language Co-Activation on L1 and L2 Speech Fluency.” *Acta Psychologica* 161:25–35. doi:10.1016/j.actpsy.2015.07.015.

Best, Catherine T. 1994. “The Emergence of Native-Language Phonological Influences in Infants: A Perceptual Assimilation Model.” In *the Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*, edited by Judith C. Goodman and Howard C. Nusbaum, 167–224. Cambridge, MA: MIT Press.

———. 1995. “A Direct Realist View of Crosslanugage Speech Perception.” In Strange 1995, 171–204.

Best, Catherine T., and Ocke-Schwen Bohn. 2015. “Testing PAM and SLM: Perception of American English Approximants by Native German Listeners.” *Proceedings of the 6th International Symposium on the Acquisition of Second Language Speech, New Sounds 2010, Poznan, Poland*, 43–48.

Best, Catherine T., and Michael D. Tyler. 2007. “Nonnative and Second-Language Speech Perception: Commonalities and Complementarities.” In Bohn and Munro 2007, 13–34.

Bialystok, Ellen, Kenji Hakuta, and Edward Wiley. 2003. “Critical Evidence: A Test of the Critical-Period Hypothesis for Second-Language Acquisition.” *Psychological science* 14 (1): 31–38. doi:10.1111/1467-9280.01415.

Birdsong, David. 1992. “Ultimate Attainment in Second Language Acquisition.” *Language* 68:706–55. doi:10.2307/416851.

———. 2014. “The Critical Period Hypothesis for Second Language Acquisition: Tailoring the Coat of Many Colors.” In *Essential Topics in Applied Linguistics and Multilingualism, Second Language Learning and Teaching: Studies in Honor of David Singleton*, edited by Mirosław Pawlak and Larissa Aronin, 43–50. Second Language Learning and Teaching. Heidelberg, New York: Springer.

Bley-Vroman, Robert W. 1990. “The Logical Problem of Foreign Language Learning.” *Linguistic Analysis* 20 (1-2): 3–50.

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

Boberg, Charles. 2005. “The Canadian Shift in Montreal.” *Language Variation and Change* 17 (2): 133–54. doi:10.1017/S0954394505050064.

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

Bohn, Ocke-Schwen, and Murray J. Munro, eds. 2007. *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege.* Language Learning & Language Teaching 17. Amsterdam, Philadelphia: John Benjamins.

Bóna, Judit. 2014. “Voice Onset Time and Speakers' Age: Data from Hungarian.” *Clinical Linguistics & Phonetics* 28 (5): 366–72. doi:10.3109/02699206.2013.875593.

Bongaerts, Theo. 1999. “Ultimate Attainment in L2 Pronunciation: The Case of Very Advanced Late L2 Learners.” In *Second Language Acquisition and the Critical Period Hypothesis*, edited by David Birdsong, 133–59. New York: Routledge.

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 (2): 298–308. doi:10.1111/1467-9582.00069.

Bongaerts, Theo, Brigitte Planken, and Erik Schils. 1995. “Can Late Learners Attain a Native Accent in a Foreign Language? A Test of the Critical Period Hypothesis.” In *the Age Factor in Second Language Acquisition*, edited by David M. Singleton and Zsolt Lengyel, 30–50. Clevedon: Multilingual Matters.

Bongaerts, Theo, Chantal van Summeren, Brigitte Planken, and Erik Schils. 1997. “Age and Ultimate Attainment in the Pronunciation of a Foreign Language.” *Studies in Second Language Acquisition* 19 (4): 447–65. doi:10.1017/S0272263197004026.

Brandstätter, Julia, Christian H. Kaseß, and Sylvia Moosmüller. 2015. “Quality and Quantity in High Vowels in Standard Austrian German.” In Leeman et al. 2015, 79–92.

Brandstätter, Julia, and Sylvia Moosmüller. 2015. “Neutralisierung Der Hohen Ungerundeten Vokale in Der Wiener Standardsprache - a Sound Change in Progress?” In *Standarddeutsch Im 21. Jahrhundert. Theoretische Und Empirische Ansätze Mit Einem Fokus Auf Österreich*, edited by Manfred Glauninger and Alexandra N. Lenz, 185–205. Göttingen: Vandenhoeck & Ruprecht.

Braun, Angelika, ed. 1996. *Untersuchungen Zu Stimme Und Sprache. Papers on Speech and Voice.* Stuttgart: Franz Steiner Verlag.

———. 1996. “Zur Regionalen Distribution Von VOT Im Deutschen.” In Braun 1996, 19–32.

Braunschweiler, Norbert. 1997. “Integrated Cues of Voicing and Vowel Length in German: A Production Study.” *Language and Speech* 40:353–76. doi:10.1177/002383099704000403.

Brennan, Eileen M., and John S. Brennan. 1981. “Accent Scaling and Language Attitudes: Reactions to Mexican American English Speech.” *Language and Speech* 24 (3): 207–21. doi:10.1177/002383098102400301.

Brooks, Nelson. 1960. *Language and Language Learning.* 2nd ed. New York: Harcourt, Brace & World Inc.

Butterworth, Guy, and Evelyn M. Hatch. 1978. “A Spanish-Speaking Adolescent's Acquisition of English Syntax.” In Hatch 1978, 231–45.

Byrd, Dani. 1993. “54,000 American Stops.” *UCLA Working Papers in Phonetics* 83: 97–116.

Caramazza, Alfonso, Grace H. Yeni-Komshian, Edgar B. Zurif, and Elena Carbone. 1973. “The Acquisition of a New Phonological Contrast: The Case of Stop Consonants in French-English Bilinguals.” *The Journal of the Acoustical Society of America* 54 (2): 421–28. doi:10.1121/1.1913594.

Carney, Arlene E. 1977. “Noncategorical Perception of Stop Consonants Differing in VOT.” *The Journal of the Acoustical Society of America* 62 (4): 961–70. doi:10.1121/1.381590.

Carroll, John B. 1968. “Contrastive Analysis and Interference Theory.” In *Report of the Nineteenth Annual Round Table Meeting of Linguistics and Language Studies: Contrastive Analysis and Its Pedagogical Implications,* edited by James E. Alatis, 113–22. Monograph Series on Language and Linguistics 21. Washington, DC: Georgetown University Press.

Chamorro, Gloria, Antonella Sorace, and Patrick Sturt. 2015. “What Is the Source of L1 Attrition? The Effect of Recent L1 Re-Exposure on Spanish Speakers Under L1 Attrition.” *Bilingualism: Language and Cognition* 19 (3): 520–32. doi:10.1017/S1366728915000152.

Chang, Charles B. 2011. “Systemic Drift of L1 Vowels in Novice L2 Learners.” In *Proceedings of the 17th International Congress of Phonetic Sciences, 17-21 August 2011, Hong Kong,* edited by Wai-Sum Lee and Eric Zee, 428–31. Hongkong: City University of Hong Kong.

———. 2012. “Rapid and Multifaceted Effects of Second-Language Learning on First-Language Speech Production.” *Journal of Phonetics* 40 (2): 249–68. doi:10.1016/j.wocn.2011.10.007.

———. 2013. “A Novelty Effect in Phonetic Drift of the Native Language.” *Journal of Phonetics* 41 (6): 520–33. doi:10.1016/j.wocn.2013.09.006.

———. 2019. “Language Change and Linguistic Inquiry in a World of Multicompetence: Sustained Phonetic Drift and Its Implications for Behavioral Linguistic Research.” *Journal of Phonetics* 74:96–113. doi:10.1016/j.wocn.2019.03.001.

Chao, Kuan-Yi, and Li-mei Chen. 2008. “A Cross-Linguistic Study of Voice Onset Time in Stop Consonant Productions.” *Computational Linguistics and Chinese Language Processing* 13 (2): 215–31. doi:10.30019/IJCLCLP.200806.0005.

Cherciov, Mirela. 2012. “Investigating the Impact of Attitude on First Language Attrition and Second Language Acquisition from a Dynamic Systems Theory Perspective.” *International Journal of Bilingualism* 17 (6): 716–33. doi:10.1177/1367006912454622.

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

Cho, Taehong, Sun-Ah Jun, and Peter Ladefoged. 2002. “Acoustic and Aerodynamic Correlated of Korean Stops and Fricatives.” *Journal of Phonetics* 30 (2): 193–228. doi:10.1006/jpho.2001.0153.

Cho, Taehong, and Peter Ladefoged. 1999. “Variation and Universals in VOT: Evidence from 18 Languages.” *Journal of Phonetics* 27 (2): 207–29. doi:10.1006/jpho.1999.0094.

Cho, Taehong, Douglas H. Whalen, and Gerard J. Docherty. 2019. “Voice Onset Time and Beyond: Exploring Laryngeal Contrast in 19 Languages.” *Journal of Phonetics* 72:52–65. doi:10.1016/j.wocn.2018.11.002.

Chodroff, Eleanor, John Godfrey, Sanjeev Khudanpur, and Colin Wilson. 2015. “Structured Variability in Acoustic Realization: A Corpus Study of Voice Onset Time in American English Stops.” *18th International Congress of Phonetic Sciences (ICPhS), Glasgow, Scotland*.

Chodroff, Eleanor, and Colin Wilson. 2017. “Structure in Talker-Specific Phonetic Realization: Covariation of Stop Consonant VOT in American English.” *Journal of Phonetics* 61:30–47. doi:10.1016/j.wocn.2017.01.001.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax.* 3rd ed. Massachusetts Institute of Technology (Cambridge, Mass.). Research Laboratory of Electronics. Special technical report 11. Cambridge, MA: MIT Press.

———. 1967. “Recent Contributions to the Theory of Innate Ideas.” In *a Portrait of Twenty-Five Years: Boston Studies in the Philosophy of Science*, edited by Robert S. Cohen and Marx W. Wartofsky, 31–40. Dordrecht: Springer.

Cieri, Christopher. 2011. “Making a Field Recording.” In Di Paolo and Yaeger-Dror 2011, 24–35.

Clamann, Aaron. 2017. “Schwarzenegger Erfüllt Sich Kindheitstraum in Deutschland.” Accessed May 20, 2020. https://www.nrz.de/panorama/schwarzenegger-erfuellt-sich-kindheitstraum-in-deutschland-id211987215.html.

Coetzee, Andries W., Patrice S. Beddor, Kerby Shedden, Will Styler, and Daan Wissing. 2018. “Plosive Voicing in Afrikaans: Differential Cue Weighting and Tonogenesis.” *Journal of Phonetics* 66:185–216. doi:10.1016/j.wocn.2017.09.009.

Colantoni, Laura, Jeffrey Steele, and Paola Escudero. 2015. *Second Language Speech: Theory and Practice.* Cambridge: Cambridge University Press.

Cook, Vivian. 1991. “The Poverty-of-the-Stimulus Argument and Multicompetence.” *Second Language Research* 7 (2): 103–17. doi:10.1177/026765839100700203.

———. 2003. “The Changing L1 in the L2 User’s Mind.” In *Effects of the Second Language on the First,* edited by Vivian Cook, 1–18. Second Language Acquisition 3. Clevedon: Multilingual Matters.

Corder, Stephen P. 1967. “The Significance of Learner's Errors.” *IRAL: International Review of Applied Linguistics in Language Teaching* 5:161–70. doi:10.1515/iral.1967.5.1-4.161.

———. 1971. “Idiosyncratic Dialects and Error Analysis.” *International Journal of Applied Linguistics* 9 (2): 147–60. doi:10.1515/iral.1971.9.2.147.

Council of Europe. 2001. “Common European Framework of Reference for Languages, Learning, Teaching, Assessment.” Accessed March 08, 2020. https://www.europaeischer-referenzrahmen.de/sprachniveau.php.

Crystal, Thomas H., and Arthur S. House. 1988. “A Note on the Variability of Timing Control.” *Journal of Speech and Hearing Research* 31 (3): 497–502. doi:10.1044/jshr.3103.497.

———. 1990. “Articulation Rate and Duration of Syllables and Stress Groups in Connected Speech.” *The Journal of the Acoustical Society of America* 88 (1): 1842–48. doi:10.1121/1.399955.

Cunha, Conceição, Jonathan Harrington, and Philip Hoole. 2013. “A Physiological Analysis of the Tense/lax Vowel Contrast in Two Varieties of German.” *Proceedings of the 14th Annual Conference of the International Speech Communication Association (Interspeech), Lyon, France*, 325–28.

Cunha, Conceição, Jonathan Harrington, Sylvia Moosmüller, and Julia Brandstätter. 2015. “The Influence of Consonantal Context on the Tense-Lax Contrast in Two Satndard Varieties of German.” In Leeman et al. 2015, 65–77.

Daily Mail UK. 2015. “Arnold Schwarzenegger Reveals He Can Speak Perfect English - but Keeps Talking with an Accent Because 'Fans Expect It'.” Accessed January 25, 2021. http://www.dailymail.co.uk/tvshowbiz/article-3141778/Arnold-

Schwarzenegger-reveals-speak-perfect-English-keeps-talking-accent-fans-expect-it.html.

Darcy, Isabelle, and Franziska Krüger. 2012. "Vowel Perception and Production in Turkish Children Acquiring L2 German." *Journal of Phonetics* 40 (4): 568–81. doi:10.1016/j.wocn.2012.05.001.

Davies, Alan. 2003. *The Native Speaker: Myth and Reality.* Bilingual Education and Bilingualism 38. Clevedon: Multilingual Matters.

———. 2004. "The Native Speaker in Applied Linguistics." In *the Handbook of Applied Linguistics,* edited by Alan Davies and Catherine Elder, 431–50. Malden, MA: Blackwell.

Davis, Katharine. 1994. "Stop Voicing in Hindi." *Journal of Phonetics* 22 (2): 177–93. doi:10.1016/s0095-4470(19)30192-5.

———. 1995. "Phonetic and Phonological Contrasts in the Acquisition of Voicing: Voice Onset Time Production in Hindi and English." *Journal of Child Language* 22 (2): 275–305. doi:10.1017/S030500090000979X.

De Bot, Kees. 1998. "The Psycholinguistics of Language Loss." In *Bilingualism and Migration,* edited by Guus Extra and Ludo Verhoeven, 345–61. Studies on Language Acquisition 14. Berlin: Mouton de Gruyter.

———. 2004. "Introduction: Special Issue on Language Attrition." *International Journal of Bilingualism* 8 (3): 233–37. doi:10.1177/13670069040080030101.

———. 2007. "Dynamic Systems Theory, Lifespan Development and Language Attrition." In Köpke et al. 2007, 53–68.

———. 2008. "Introduction: Second Language Development as a Dynamic Process." *The Modern Language Journal* 92 (2): 166–78. doi:10.1111/j.1540-4781.2008.00712.x.

———. 2015. *A History of Applied Linguistics: From 1980 to the Present.* New York: Routledge.

De Bot, Kees, and Michael Clyne. 1994. "A 16-year Longitudinal Study of Language Attrition in Dutch Immigrants in Australia." *Journal of Multilingual and Multicultural Development* 15 (1): 17–28. doi:10.1080/01434632.1994.9994554.

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 Sevelopment: Methods and Techniques,* edited by Marjolijn Verspoor, Kees de Bot, and Wander Lowie, 5–24. Language Learning & Language Teaching 29. Amsterdam: John Benjamins.

De Bot, Kees, Wander Lowie, and Marjolijn Verspoor. 2007. "A Dynamic Systems Theory Approach to Second Language Acquisition." *Bilingualism: Language and Cognition* 10 (1): 7–21. doi:10.1017/S1366728906002732.

De Leeuw, Esther. 2009. "When Your Native Language Sounds Foreign: A Phonetic Investigation into First Language Attrition." PhD thesis, Queen Margaret University, Edinburgh, United Kingdom.

———. 2017. "How Phonetics and Phonology Inform L1 Attrition (Narrowly Defined) Research." *Linguistic Approaches to Bilingualism* 7 (6): 725–29. doi:10.1075/lab.00013.lee.

———. 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. doi:10.1016/j.wocn.2018.12.002.

De Leeuw, Esther, Ineke Mennen, and James M. Scobbie. 2012a. "Dynamic Systems, Maturational Constraints, and L1 Phonetic Attrition." *International Journal of Bilingualism* 17 (6): 683–700. doi:10.1177/1367006912454620.

———. 2012b. "Singing a Different Tune in Your Native Language: First Language Attrition of Prosody." *International Journal of Bilingualism* 16 (1): 101–16. doi:10.1017/S136672891100071X.

De Leeuw, Esther, Conny Opitz, and Dorota Lubinska. 2013. "Dynamics of First Language Attrition Across the Lifespan." *International Journal of Bilingualism* 17 (6): 667–74. doi:10.1177/1367006912454618.

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 (1): 33–40. doi:10.1017/S1366728909990289.

De Leeuw, Esther, Aurela Tusha, and Monika S. Schmid. 2017. "Individual Phonological Attrition in Albanian-English Late Bilinguals." *Bilingualism: Language and Cognition* 21 (2): 278–95. doi:10.1017/S1366728917000025.

Debruyne, Frans, and Wivine Decoster. 1999. "Acoustic Differences Between Sustained Vowels Perceived as Young or Old." *Logopedics Phoniatrics Vocology* 24 (1): 1–5. doi:10.1080/140154399434490.

Deppermann, Arnulf, Stefan Kleiner, and Ralf Knöbl. 2011. "'Standard Usage': Towards a Realistic Conception of Spoken Standard German." In *Language Variation – European Perspectives IV; Selected Papers from the 6th International Conference on Language Variation in Europe (ICLaVE 6), Freiburg*, edited by Göz Kaufmann, Javier C. Reina, and Peter Auer, 83–116. Studies in Language Variation 14. Amsterdam, Philadelphia: John Benjamins.

Derwing, Tracey M., and Murray J. Munro. 1997. “Accent, Intelligibility, and Comprehensibility: Evidence from Four L1s.” *Studies in Second Language Acquisition* 19 (1): 1–16. doi:10.1017/S0272263197001010.

———. 2015. *Pronunciation Fundamentals: Evidence-Based Perspectives for L2 Teaching and Research.* Amsterdam, Philadelphia: John Benjamins.

Di Paolo, Marianna, and Malcah Yaeger-Dror. 2011. “Field Methods: Gathering Data, Creating a Corpus, and Reporting Your Work.” In Di Paolo and Yaeger-Dror 2011, 7–23.

———, eds. 2011. *Sociophonetics: A Student's Guide.* New York: Routledge.

Di Paolo, Marianna, Malcah Yaeger-Dror, and Alicia Beckford Wassnik. 2011. “Analyzing Vowels.” In Di Paolo and Yaeger-Dror 2011, 87–106.

DiCanio, Christian, Hosung Nam, Jonathan D. Amith, Rey Castillo Garcia, and Douglas H. Whalen. 2015. “Vowel Variability in Elicited Versus Spontaneous Speech: Evidence from Mixtec.” *Journal of Phonetics* 48:45–59. doi:10.1016/j.wocn.2014.10.003.

Die Presse. 2011. “Arnold Schwarzenegger Auf Österreich-Tour.” Accessed May 20, 2020. https://www.diepresse.com/672167/arnold-schwarzenegger-auf-osterreich-tour.

———. 2018. “Arnold Schwarzenegger Auf Kurzbesuch in Graz.” Accessed May 20, 2020. https://www.diepresse.com/5498061/arnold-schwarzenegger-auf-kurzbesuch-in-graz.

Dmitrieva, Olga, Allard Jongman, and Joan A. Sereno. 2010. “Phonological Neutralization by Native and Non-Native Speakers: The Case of Russian Final Devoicing.” *Journal of Phonetics* 38 (3): 483–92. doi:10.1016/j.wocn.2010.06.001.

———. 2020. “The Effect of Instructed Second Language Learning on the Acoustic Properties of First Language Speech.” *Languages* 5 (4): 1–33. doi:10.3390/languages5040044.

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

Dorman, Michael F., Michael Studdert-Kennedy, and Lawrence J. Raphael. 1977. “Stop-Consonant Recognition: Release Bursts and Formant Transitions as Functionally Equivalent, Context-Dependent Cues.” *Perception & Psychophysics* 22 (2): 109–22. doi:10.3758/BF03198744.

Duff, Patsy A., and Tim Anderson. 2015. “Case Study Research.” In *the Cambridge Guide to Language Research,* edited by James D. Brown and Christine A. Coombe, 112–18. New York: Cambridge University Press.

Dulay, Heidi, and Marina Burt. 1973. “Should We Teach Children Syntax?” *Language Learning* 23 (2): 245–58. doi:10.1111/j.1467-1770.1973.tb00659.x.

———. 1974a. “Errors and Strategies in Child Second Language Acquisition.” *TESOL Quarterly* 8 (2): 129–36. doi:10.2307/3585536.

———. 1974b. “Natural Sequences in Child Second Language Acquisition.” *Language Learning* 24 (1): 37–53. doi:10.1111/j.1467-1770.1974.tb00234.x.

Dulay, Heidi, Marina Burt, and Stephen Krashen. 1982. *Language Two.* Oxford, UK: Oxford University Press.

Ecke, Peter. 2004. “Language Attrition and Theories of Forgetting: A Cross-Disciplinary Review.” *International Journal of Bilingualism* 8 (3): 321–54. doi:10.1177/13670069040080030901.

Eckert, Penelope. 2004. “Vowel Shifts in Northern California and the Detroit Suburbs.” Accessed August 01, 2019. https://web.stanford.edu/~eckert/vowels.html.

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, 139–43. Malden, MA: Blackwell.

Eckman, Fred R. 1977. “Markedness and the Contrastive Analysis Hypothesis.” *Language Learning* 27 (2): 315–30. doi:10.1111/j.1467-1770.1977.tb00124.x.

———. 1996. “A Functional-Typological Approach to Second Language Acquisition.” In *Handbook of Second Language Acquisition*, edited by William C. Ritchie and Tej K. Bhatia, 195–211. San Diego: Academic Press.

———. 2003. “Some Principles of Second Language Phonology.” *Second Language Research* 19 (3): 169–208. doi:10.1191/0267658303sr2190a.

———. 2008. “Typological Markedness and Second Language Phonology.” In Hansen Edwards and Zampini 2008, 95–115.

Eichhorn, Julie T., Raymond D. Kent, Diane Austin, and Houri K. Vorperian. 2018. “Effects of Aging on Vocal Fundamental Frequency and Vowel Formants in Men and Women.” *Journal of Voice* 32 (5): 644.e1-644.e9. doi:10.1016/j.jvoice.2017.08.003.

Elliott, Raymond. 1995. “Field Independence/Dependence, Hemispheric Specialization, and Attitude in Relation to Pronunciation Accuracy in Spanish as a Foreign Language.” *The Modern Language Journal* 79 (3): 356–71. doi:10.1111/j.1540-4781.1995.tb01112.x.

Ellis, Rod. 1991. *Understanding Second Language Acquisition.* 7th ed. Oxford, UK: Oxford University Press.

———. 1994. *The Study of Second Language Acquisition.* Oxford, UK: Oxford University Press.

———. 1997. *Second Language Acquisition.* Oxford, UK: Oxford University Press.

Ernst, Peter. 2004. "Dialektsoziologische Grenzräume in Der Oststeiermark." *Zeitschrift für Dialektologie und Linguistik* 71 (1): 3–22.

Fabiano-Smith, Leah, and Ferenc Bunta. 2012. "Voice Onset Time of Voiceless Bilabial and Velar Stops in 3-Year-Old Bilingual Children and Their Age-Matched Monolingual Peers." *Clinical Linguistics & Phonetics* 26 (2): 148–63. doi:10.3109/02699206.2011.595526.

Fischer-Jørgensen, Eli. 1954. "Acoustic Analysis of Stop Consonants." *Miscellanea Phonetica* 2: 42–49.

Fishman, Joshua A. 2007. "Who Speaks What Language to Whom and When?" In Wei 2007, 55–70.

Fisiak, Jacek, ed. 1981. *Contrastive Linguistics and the Language Teacher.* Oxford: Pergamon Press.

Flege, James E. 1980. "Phonetic Approximation in Second Language Acquisition." *Language Learning* 30 (1): 117–34. doi:10.1111/j.1467-1770.1980.tb00154.x.

———. 1984. "The Detection of French Accent by American Listeners." *The Journal of the Acoustical Society of America* 76 (3): 692–707. doi:10.1121/1.391256.

———. 1987a. "Effects of Equivalence Classification on the Production of Foreign Language Speech Sounds." In *Sound Patterns in Second Language Acquisition*, edited by Allan James and Jonathan Leather, 9–39. Dordrecht: Foris.

———. 1987b. "The Production of 'New' and 'Similar' Phones in a Foreign Language: Evidence for the Effect of Equivalence Classification." *Journal of Phonetics* 15 (1): 47–65. doi:10.1016/S0095-4470(19)30537-6.

———. 1991. "Age of Learning Affects the Authenticity of Voice-onset Time (VOT) In Stop Consonants Produced in a Second Language." *The Journal of the Acoustical Society of America* 89 (1): 395–411. doi:10.1121/1.400473.

———. 1995. "Second Language Speech Learning: Theory, Findings, and Problems." In Strange 1995, 233–72.

———. 2002. "Interactions Between the Native and Second- Language Phonetic Systems. In P. Burmeister." In *an Integrated View of Language Development: Papers in Honor of Henning Wode*, edited by Petra Burmeister, Thorsten Piske, and Andreas Rohde, 217–43. Trier: WVT Wissenschaftlicher Verlag Trier.

———. 2003. "Assessing Constraints on Second-Language Segmental Production and Perception." In *Phonetics and Phonology in Language Comprehension and*

Production: Differences and Similarities, edited by Antje S. Meyer and Niels O. Schiller, 319–55. Berlin: Mouton de Gruyter.

———. 2007. "Language Contact in Bilingualism: Phonetic System Interactions." In *Papers in Laboratory Phonology 9: Change in Phonology*, edited by Jennifer Cole and José I. Hualde, 353–82. Berlin: Mouton de Gruyter.

———. 2017. "The Cross-Language Acquisition of Stops Differing in VOT: Part 1, Historical Overview." *Phonetic Teaching and Learning Conference, University College London, August 2017*.

———. 2018. "How the Revised Speech Learning Model (SLM-R) Works for Stop Consonants." *Research Frontiers of Second Language Speech, Nanjing University of Science and Technology, China, April, 2018*.

Flege, James E., and Ocke-Schwen Bohn. 2021. "The Revised Speech Learning Model (SLM-R)." In *Second Language Speech Learning: Theoretical and Empirical Progress*, edited by Ratree Wayland, 3–83. Cambridge: Cambridge University Press.

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 (4): 437–70. doi:10.1006/jpho.1997.0052.

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 (4): 335–45. doi:10.1016/S0095-4470(19)30999-4.

Flege, James E., and Wieke Eefting. 1987a. "Cross-Language Switching in Stop Consonant Perception and Production by Dutch Speakers of English." *Speech Communication* 6 (3): 185–202. doi:10.1016/0167-6393(87)90025-2.

———. 1987b. "Production and Perception of English Stops by Native Spanish Speakers." *Journal of Phonetics* 15 (1): 67–83. doi:10.1016/S0095-4470(19)30538-8.

———. 1988. "Imitation of a VOT Continuum by Native Speakers of English and Spanish: Evidence for Phonetic Category Formation." *The Journal of the Acoustical Society of America* 83 (2): 729–40. doi:10.1121/1.396115.

Flege, James E., and Kathryn L. Fletcher. 1992. "Talker and Listener Effects on Degree of Perceived Foreign Accent." *The Journal of the Acoustical Society of America* 91 (1): 370–89. doi:10.1121/1.402780.

Flege, James E., Elaina M. Frieda, and Takeshi Nozawa. 1997. "Amount of Native-Language (L1) Use Affects the Pronunciation of an L2." *Journal of Phonetics* 25 (2): 169–86. doi:10.1006/jpho.1996.0040.

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 (3): 708–21. doi:10.1121/1.391257.

Flege, James E., and Serena Liu. 2001. “The Effect of Experience on Adults' Acquisition of a Second Language.” *Studies in Second Language Acquisition* 23 (4): 527–52. doi:10.1017/S0272263101004041.

Flege, James E., Ian R. A. MacKay, and Diane Meador. 1999. “Native Italian Speakers’ Perception and Production of English Vowels.” *The Journal of the Acoustical Society of America* 106 (5): 2973–87. doi:10.1121/1.428116.

Flege, James E., Murray J. Munro, and Ian R. A. MacKay. 1995. “Factors Affecting Strength of Perceived Foreign Accent in a Second Language.” *The Journal of the Acoustical Society of America* 97 (5): 3125–34. doi:10.1121/1.413041.

———. 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, 47–74. Amsterdam, Philadelphia: John Benjamins.

Flege, James E., Murray J. Munro, and Laurie Skelton. 1992. “Production of the Word-Final English /t/-/d/ Contrast by Native Speakers of English, Mandarin, and Spanish.” *The Journal of the Acoustical Society of America* 92 (1): 128–43. doi:10.1121/1.404278.

Flege, James E., and Robert Port. 1981. “Cross-Language Phonetic Interference: Arabic to English.” *Language and Speech* 24 (2): 125–46. doi:10.1177/002383098102400202.

Flege, James E., Carlo Schirru, and Ian R. A. MacKay. 2003. “Interaction Between the Native and Second Language Phonetic Subsystems.” *Speech Communication* 40 (4): 467–91. doi:10.1016/S0167-6393(02)00128-0.

Flege, James E., and Anna M. Schmidt. 1995. “Native Speakers of Spanish Show Rate-Dependent Processing of English Stop Consonants.” *Phonetica* 52 (2): 90–111. doi:10.1159/000262062.

Flege, James E., Naoyuki Takagi, and Virginia Mann. 1995. “Japanese Adults Can Learn to Produce English /r/ and /l/ Accurately.” *Language and Speech* 38 (1): 25–55. doi:10.1177/002383099503800102.

Flege, James E., Grace H. Yeni-Komshian, and Serena Liu. 1999. “Age Constraints on Second-Language Acquisition.” *Journal of Memory and Language* 41 (1): 78–104. doi:10.1006/jmla.1999.2638.

Fletcher, Annalise R., Megan J. McAuliffe, Kaitlin Lansford, and Juli M. Liss. 2015. “The Relationship Between Speech Segment Duration and Vowel Centralization

in a Group of Older Speakers." *The Journal of the Acoustical Society of America* 138 (4): 2132–39. doi:10.1121/1.4930563.

Foster-Cohen, Susan. 2001. "First Language Acquisition ... Second Language Acquisition: What's Hecuba to Him or He to Hecuba?'." *Second Language Research* 17 (4): 329–44. doi:10.1177/026765830101700403.

Foulkes, Paul, Gerard J. Docherty, and Mark J. Jones. 2011. "Analyzing Stops." In Di Paolo and Yaeger-Dror 2011, 58–71.

Fowler, Carol A., Valery Sramko, David Ostry, Sarah A. Rowland, and Pierre Hallé. 2008. "Cross Language Phonetic Influences on the Speech of French–English Bilinguals." *Journal of Phonetics* 36 (4): 649–63. doi:10.1016/j.wocn.2008.04.001.

Francis, Alexander L., Valter Ciocca, and Jojo M. C. Yu. 2003. "Accuracy and Variability of Acoustic Measures of Voicing Onset." *Journal of Acoustical Society of America* 133 (2): 1025–32. doi:10.1121/1.1536169.

Fries, Charles. 1945. *Teaching and Learning English as a Foreign Language.* Ann Arbor, MI: University of Michigan Press.

Fuchs, Susanne. 2005. "Articulatory Correlates of the Voicing Contrast in Alveolar Obstruent Production in German." *ZAS Papers in Linguistics* 41. http://publikationen.ub.uni-frankfurt.de/frontdoor/index/index/docId/30683.

Gala. 2015. "Der 'Terminator'-Star Ist Zurück." Accessed January 25, 2021. https://www.gala.de/stars/news/interview/arnold-schwarzenegger--der--terminator--star-ist-zurueck-20234970.html.

Gandour, Jack, Soranee H. Petty, Rochana Dardarananda, Sumalee Dechongkit, and Sunee Mukngoen. 1986. "The Acquisition of the Voicing Contrast in Thai: A Study of Voice Onset Time in Word-Initial Stop Consonants." *Journal of Child Language* 13 (3): 561–72. doi:10.1017/S0305000900006887.

Gardner, Robert C. 2010. *Motivation and Second Language Acquisition: The Socio-Educational Model.* New York: Peter Lang.

Gardner, Robert C., and Wallace E. Lambert. 1972. *Attitudes and Motivation in Second-Language Learning.* Rowley, MA: Newbury House.

Gardner, Robert C., and Peter D. MacIntyre. 1993. "A Student's Contribution to Second-Language Learning. Part II: Affective Variables." *Language Teaching* 26 (1): 1–11. doi:10.1017/S0261444800000045.

Gass, Susan M., and Larry Selinker. 2008. *Second Language Acquisition. An Introductory Course.* 3rd ed. New York, London: Routledge.

Gersemann, Olaf. 2009. "Der Mann, Der Sich Selbst Erfunden Hat." Accessed January 25, 2021. http://www.welt.de/politik/ausland/article4895817/Der-Mann-der-sich-selbst-erfunden-hat.html.

Gerstenberger, Christian. 2019. “Arnold Schwarzenegger Besucht Dortmund - Aber Nur, Wenn Die Kasse Stimmt.” Accessed May 20, 2020. https://www.ruhr-nachrichten.de/dortmund/arnold-schwarzenegger-besucht-dortmund-aber-nur-wenn-die-kasse-stimmt-1433113.html.

Ghadessy, Mohsen. 1980. “Implications of Error Analysis for Second/ Foreign Language Acquisition.” *IRAL: International Review of Applied Linguistics in Language Teaching* 18:93–104. doi:10.1515/iral.1980.18.1-4.93.

Grászi, Tekla E., and Anna Kohári. 2014. “Multiple Bursts in Hungarian Voiceless Plosives and VOT Measurements.” *Proceedings of the 10th International Seminar on Speech Production (ISSP)*.

Grassegger, Hans. 1988. *Signalphonetische Untersuchungen Zur Differenzierung Italienischer Plosive Durch Österreichische Sprecher.* Forum Phoneticum 40. Hamburg: Buske.

———. 1996. “Koartikulatorische Einflüsse Auf Die Produktion Von Anlautplosiven Bei Östserreichischen (Steirischen) Sprechern.” In Braun 1996, 12–18.

Greene, James S., and Emilie J. Wells. 1927. *The Cause and Cure of Speech Disorders: A Textbook for Students and Teachers on Stuttering, Stammering and Voice Conditions.* New York: The Macmillan Company.

Griffen, Toby D. 1980. “A Nonsegmental Approach to the Teaching of Pronunciation.” *Revue de Phonétique Appliquée* 54: 81–94.

Grosjean, François. 1982. *Life with Two Languages: An Introduction to Bilingualism.* Cambridge, Mass. Harvard University Press.

———. 1989. “Neurolinguists, Beware! The Bilingual Is Not Two Monolinguals in One Person.” *Brain and Language* 36 (1): 3–15. doi:10.1016/0093-934X(89)90048-5.

———. 1992. “Another View of Bilingualism.” *Advances in Psychology* 83:51–62. doi:10.1016/S0166-4115(08)61487-9.

———. 2010. *Bilingual: Life and Reality.* Cambridge, Mass. Harvard University Press.

Guion, Susan G. 2003. “The Vowel Systems of Quichua-Spanish Bilinguals: Age of Acquisition Effects on the Mutual Influence of the First and Second Languages.” *Phonetica* 60 (2): 98–128. doi:10.1159/000071449.

Guion, Susan G., James E. Flege, Reiko Akahane-Yamada, and Jesica C. Pruitt. 2000. “An Investigation of Current Models of Second Language Speech Perception: The Case of Japanese Adults’ Perception of English Consonants.” *The Journal of the Acoustical Society of America* 107 (5): 2711–24. doi:10.1121/1.428657.

Guion, Susan G., James E. Flege, and Jonathan D. Loftin. 2000. "The Effect of L1 Use on Pronunciation in Quichua-Spanish Bilinguals." *Journal of Phonetics* 28 (1): 27–42. doi:10.1006/jpho.2000.0104.

Guten Morgen Österreich. 2012a. "Guten Morgen Österreich Ö3 – Arnie Allein Zu Hause 'Schneemangel' [Video File]." Accessed May 25, 2020. https://www.youtube.com/watch?v=X4PtQL4sH8U.

———. 2012b. "Guten Morgen Österreich Ö3 – Arnie Allein Zu Hause 'Twilight' [Video File]." Accessed May 23, 2020. https://www.youtube.com/watch?v=bNDPRb400ZA.

Hacopian, Narineh. 2003. "A Three-Way VOT Contrast in Final Position: Data from Armenian." *Journal of the International Phonetic Association* 33 (1): 51–80. doi:10.1017/S0025100303001154.

Haggard, Mark, Stephen Ambler, and Mo Callow. 1970. "Pitch as a Voicing Cue." *The Journal of the Acoustical Society of America* 47 (2): 613–17. doi:10.1121/1.1911936.

Hagiwara, Robert. 1997. "Dialect Variation and Formant Frequency: The American English Vowels Revisited." *The Journal of the Acoustical Society of America* 102 (1): 655–58. doi:10.1121/1.419712.

Hammarberg, Björn. 1979. "The Insufficiency of Error Analysis." In *Studies in Contrastive Linguistics and Error Analysis*, edited by Dietrich Nehls, 107–17. Studies in Descriptive Linguistics 2. Heidelberg: Julius Groos.

Han, Zhaohong. 1998. "Fossilization: An Investigation into Advanced L2 Learning of a Typologically Distant Language." PhD thesis, University of London, London, United Kingdom.

———. 2004a. *Fossilization in Adult Second Language Acquisition.* Clevedon: Multilingual Matters.

———. 2004b. "Fossilization: Five Central Issues." *International Journal of Applied Linguistics* 14 (2): 212–42. doi:10.1111/j.1473-4192.2004.00060.x.

———. 2005. "Fossilization: Can Grammaticality Judgment Be a Reliable Source of Evidence?" In Han and Odlin 2005, 56–82.

Han, Zhaohong, and Terence Odlin, eds. 2005. *Studies of Fossilization in Second Language Acquisition.* Clevedon: Multilingual Matters.

Hansen Edwards, Jette G. 2004. "Developmental Sequences in the Acquisition of English L2 Codas: A Preliminary Study." *Studies in Second Language Acquisition* 26 (1): 85–124. doi:10.1017/S0272263104026142.

Hansen Edwards, Jette G., and Mary L. Zampini, eds. 2008. *Phonology and Second Language Acquisition.* Studies in Bilingualism 36. Amsterdam, Philadelphia: John Benjamins.

Harada, Tetsuo. 2003. "L2 Influence on L1 Speech in the Production of VOT." In Solé, Recasens, and Romero 2003, 1085–88.

Harrington, Jonathan. 2006. "An Acoustic Analysis of 'Happy-Tensing' in the Queen's Christmas Broadcasts." *Journal of Phonetics* 34 (4): 439–57. doi:10.1016/j.wocn.2005.08.001.

———. 2010. *Phonetic Analysis of Speech Corpora.* Malden, MA: Wiley-Blackwell.

Harrington, Jonathan, and Steve Cassidy. 1999. *Techniques in Speech Acoustics.* Dordrecht, Boston: Kluwer.

Harrington, Jonathan, Philip Hoole, and Ulrich Reubold. 2012. "A Physiological Analysis of High Front, Tense-Lax Vowel Pairs in Standard Austrian German and Standard German." *Italian Journal of Linguistics* 24 (1): 149–73.

Harrington, Jonathan, Sallyanne Palethorpe, and Catherine I. Watson. 2000a. "Does the Queen Speak the Queen's English?" *Nature* 408 (6815): 927–28. doi:10.1038/35050160.

———. 2000b. "Monophthongal Vowel Changes in Received Pronunciation: An Acoustic Analysis of the Queen's Christmas Broadcasts." *Journal of the International Phonetic Association* 30 (1-2): 63–78. doi:10.1017/S0025100300006666.

———. 2007. "Age-Related Changes in Fundamental Frequency and Formants: A Longitudinal Study of Four Speakers." *Proceedings of the 8th Annual Conference of the International Speech Communication Association (Interspeech), Antwerp, Belgium*, 2753–56.

Hatch, Evelyn M., ed. 1978. *Second Language Acquisition: A Book of Readings.* Rowley, MA: Newbury House.

Hatch, Evelyn M., and Judy Wagner-Gough. 1975. "Explaining Sequence and Variation in Second Language Acquisition." In *Papers in Second Language Acquisition. Proceedings of the 6th Annual Conference on Applied Linguistics*, edited by H. D. Brown, 39–58. Ann Arbor, MI: University of Michigan Press.

Hazan, Valerie L., and Georges Boulakia. 1993. "Perception and Production of a Voicing Contrast by French-English Bilinguals." *Language and Speech* 36 (1): 17–38. doi:10.1177/002383099303600102.

Herdina, Philip, and Ulrike Jessner. 2002. *A Dynamic Model of Multilingualism: Perspectives of Change in Psycholinguistics.* Multilingual matters 121. Clevedon: Multilingual Matters.

———. 2013. “The Implications of Language Attrition for Dynamic Systems Theory: Next Steps and Consequences.” *International Journal of Bilingualism* 17 (6): 752–56. doi:10.1177/1367006912454625.

Hervé, Maxime. 2020. “RVAideMemoire: Testing and Plotting Procedures for Biostatistics: R Package Version 0.9-77.” https://CRAN.R-project.org/package=RVAideMemoire.

Higgins, Maureen B., Ronald Netsell, and Laura Schulte. 1998. “Vowel-Related Differences in Laryngeal Articulatory and Phonatory Function.” *Journal of Speech, Language, and Hearing Research* 41 (4): 712–24. doi:10.1044/jslhr.4104.712.

Hillenbrand, James, Laura A. Getty, Michael J. Clark, and Kimberlee Wheeler. 1995. “Acoustic Characteristics of American English Vowels.” *The Journal of the Acoustical Society of America* 97 (5): 3099–3111. doi:10.1121/1.411872.

Hinton, Leanne, Birch Moonwomon, Sue Bremner, Herb Luthin, Mary van Clay, Jean Lerner, and Hazel Corcoran. 1987. “It's Not Just the Valley Girls: A Study of Californian English.” In *Proceedings of the 13th Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Grammar and Cognition*, edited by J. Aske, N. Beery, L. Michaelis, and H. Filip, 117–28. Berkeley, California: Berkeley Linguistics Society.

Hinton, Martin. 2012. “An Aptitude for Speech: The Importance of Mimicry Ability in Foreign Language Pronunciation.” In *Teaching and Researching English Accent in Native and Non-Native Speakers*, edited by Ewa Waniek-Klimczak and Linda R. Shockey, 103–11. Berlin: Springer.

Hitradio Ö3. 2015. “Arnold Schwarzenegger Im Ö3-Interview [Video File].” Accessed May 21, 2020. https://www.youtube.com/watch?v=mKiYtzWztLM.

Hobel, Bettina, and Ralf Vollmann. 2015. “Phonological Case Study of the Use of (Styrian) Dialect and Standard Language in German as a Second Language.” *Grazer Linguistische Studien* 84: 5–20.

Hödl, Petra. 2017. “Production of Voice Onset Time (VOT) In Austrian German Conversational Speech: A Pilot Study.” In *Proceedings of 8th Tutorial and Research Workshop on Experimental Linguistics, Heraklion, Crete, Greece, 19-22 June*, edited by Antonis Botinis, 41–44.

———. 2019. “Production and Perception of Voice Onset Time in Austrian German.” PhD thesis, Karl-Franzens-University, Graz, Austria.

Hombert, Jean-Marie, John J. Ohala, and William G. Ewan. 1979. “Phonetic Explanations for the Development of Tones.” *Language* 55 (1): 37–58. doi:10.2307/412518.

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 (2): 361–94. doi:10.1017/S0142716411000737.

House, Arthur S., and Grant Fairbanks. 1953. "The Influence of Consonant Environment Upon the Secondary Acoustical Characteristics of Vowels." *Journal of the Acoustical Society of America* 25:105–13. doi:10.1121/1.1906982.

Howell, Peter, William J. Barry, and David Vinson. 2006. "Strength of British English Accents in Altered Listening." *Perception & Psychophysics* 68 (1): 139–53. doi:10.3758/bf03193664.

Hunnicutt, Leigh, and Paul A. Morris. 2016. "Prevoicing and Aspiration in Southern American English." *Proceedings of the 39th Annual Penn Linguistics Conference* 22 (1): 215–24.

Hyltenstam, Kenneth. 1988. "Lexical Characteristics of Near-Native Second-Langauge Learners of Swedish." *Journal of Multilingual and Multicultural Development* 9 (1-2): 67–84. doi:10.1080/01434632.1988.9994320.

Hyltenstam, Kenneth, and Niclas Abrahamsson. 2000. "Who Can Become Native-Like in a Second Language? All, Some, or None? On the Maturational Constraints Controversy in Second Language Acquisition." *Studia Linguistica* 54 (2): 150–66. doi:10.1111/1467-9582.00056.

———. 2006. "Maturational Constraints in SLA." In *the Handbook of Second Language Acquisition*, edited by Catherine J. Doughty and Michael H. Long, 538–88. Malden, MA: Blackwell.

Iverson, Paul, and Patricia K. Kuhl. 2000. "Perceptual Magnet and Phoneme Boundary Effects in Speech Perception: Do They Arise from a Common Mechanism?" *Perception & Psychophysics* 62 (4): 874–86. doi:10.3758/bf03206929.

Janoff, Arianna. 2018. "The California Vowel Shift in Santa Barbara." *Proceedings of the 4th Annual Linguistics Conference at UGA*, 30–49. http://linguistics.uga.edu/cp2017-Janoff.

Jarvis, Scott H., and Aneta Pavlenko. 2000. "Conceptual Restructuring in Language Learning: Is There an End State?" *Paper presented at the Second Language Research Forum, Madison, WI*.

Jaspaert, Koen, and Sjaak Kroon. 1989. "Social Determinants of Language Loss." *International Journal of Applied Linguistics* 83 (1): 75–98. doi:10.1075/itl.83-84.07jas.

———. 1992. "From the Typewriter of A.L. A Case Study in Language Loss." In *Maintenance and Loss of Minority Languages*, edited by Willem Fase, Koen

Jaspaert, and Sjaak Kroon, 137–47. Studies in Bilingualism 1. Amsterdam: John Benjamins.

Jean-Aubry, Gerard. 1957. *The Sea Dreamer. A Definite Biography of Joseph Conrad.* London: Ruskin House.

Jesney, Karen. 2004. "The Use of Global Foreign Accent Rating in Studies of L2 Acquisition." *A Report Prepared for the Language Research Centre University of Calgary*. https://www.semanticscholar.org/paper/The-Use-of-Global-Foreign-Accent-Rating-in-Studies-Jesney/3e4eaa880f359be90d48eb43df7574d98630926f. Accessed January 20, 2022.

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

Jessen, Michael, Krzyzstof Marasek, Katrin Schneider, and Kathrin Claßen. 1995. "Acoustic Correlates of Word Stress and the Tense/lax Opposition in the Vowel System of German." *Proceedings of the 13th International Congress of Phonetic Sciences, Stockholm, Sweden*, 428–431.

Jessen, Michael, and Catherine Ringen. 2002. "Laryngeal Features in German." *Phonology* 19 (2): 189–218. doi:10.1017/S0952675702004311.

Jilka, Matthias. 2000a. "Testing the Contribution of Prosody to the Perception of Foreign Accent." *Proceedings of New Sounds (4th International Symposium on the Acquisition of Second Language Speech), Amsterdam*, 199–207.

———. 2000b. "The Contribution of Intonation to the Perception of Foreign Accent." PhD thesis, University of Stuttgart, Stuttgart, Germany.

———. 2009. "Talent and Proficiency in Language." In *Language Talent and Brain Activity*, edited by Grzegorz Dogil and Susanne M. Reiterer, 1–16. Trends in Applied Linguistics 1. Berlin, New York: Mouton de Gruyter.

Jilka, Matthias, Volha Anufryk, Henrike Baumotte, Natalie Lewandowski, Giuseppina Rota, and Susanne M. Reiterer. 2008. "Assessing Individual Talent in Second Language Production and Perception." *Proceedings of the 5th International Symposium on the Acquisition of Second Language Speech, New Sounds 2007, Florianópolis, Santa Catarina, Brazil*, 243–58.

———. 2009. "Assessing Phonetic Talent in Second Language Performance: The Relationship Between Perception Abilities in the Native and the Second Language and Production." *The Journal of the Acoustical Society of America* 125 (4). doi:10.1121/1.4784626.

Jungwirth, Michael. 2018. "Auf Österreich-Besuch: Warum Schwarzenegger Nicht in Die Steiermark Kommt." Accessed March 30, 2019.

https://www.kleinezeitung.at/politik/aussenpolitik/5428812/Auf-OesterreichBesuch_Warum-Schwarzenegger-nicht-in-die-Steiermark.

Kartushina, Natalia, Ulrich H. Frauenfelder, and Narly Golestani. 2016. "How and When Does the Second Language Influence the Production of Native Speech Sounds." *Language Learning* 66 (S2): 155–86. doi:10.1111/lang.12187.

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* 77:1–21. doi:10.1016/j.wocn.2019.100920.

Keating, Patricia A. 1984. "Phonetic and Phonological Representation of Stop Consonant Voicing." *Language* 60 (2): 286–319. doi:10.2307/413642.

Keating, Patricia A., John R. Westbury, and Kenneth N. Stevens. 1980. "Mechanisms of Stop-Consonant Release for Different Places of Articulation." *The Journal of the Acoustical Society of America* 67 (S1): S93. doi:10.1121/1.2018489.

Kehoe, Margaret. 2002. "Developing Vowel Systems as a Window to Bilingual Phonology." *International Journal of Bilingualism* 6 (3): 315–34. doi:10.1177/13670069020060030601.

Kehoe, Margaret, Conxita Lleó, and Martin Rakow. 2004. "Voice Onset Time in Bilingual German-Spanish Children." *Bilingualism: Language and Cognition* 7 (1): 71–88. doi:10.1017/S1366728904001282.

Keijzer, Merel. 2004. "First Language Attrition: A Cross-Linguistic Investigation of Jakobson's Regression Hypothesis." *International Journal of Bilingualism* 8 (3): 389–93. doi:10.1177/13670069040080030201.

———. 2010. "The Regression Hypothesis as a Framework for First Language Attrition." *Bilingualism: Language and Cognition* 13 (1): 9–18. doi:10.1017/S1366728909990356.

Kennedy, Sara, and Pavel Trofimovich. 2008. "Intelligibility, Comprehensibility, and Accentedness of L2 Speech: The Role of Listener Experience and Semantic Context." *The Canadian Modern Language Review/La Revue canadienne des langues vivantes* 64 (3): 459–83.

Kessinger, Rachel H., and Sheila E. Blumstein. 1997. "Effects of Speaking Rate on Voice-Onset Time in Thai, French, and English." *Journal of Phonetics* 25 (2): 143–68. doi:10.1006/jpho.1996.0039.

———. 1998. "Effects of Speaking Rate on Voice-Onset Time and Vowel Production: Some Implications for Perception Studies." *Journal of Phonetics* 26 (2): 117–28. doi:10.1006/jpho.1997.0069.

Kewley-Port, Diane K., and Malcolm S. Preston. 1974. "Early Apical Stop Production: A Voice Onset Time Analysis." *Journal of Phonetics* 2:195–210. doi:10.1016/S0095-4470(19)31270-7.

Khattab, Ghada. 2007. "Variation in Vowel Production by English-Arabic Bilinguals." *Laboratory Phonology* 9: 383–410.

Kisler, Thomas, Uwe D. Reichel, and Florian Schiel. 2017. "Multilingual Processing of Speech via Web Services." *Computer Speech and Language* 45:326–47. doi:10.1016/j.csl.2017.01.005.

Klatt, Dennis H. 1975. "Voice Onset Time, Frication, and Aspiration in Word-Initial Consonant Clusters." *Journal of Speech, Language, and Hearing Research* 18 (4): 686–706. doi:10.1044/jshr.1804.686.

Kleber, Felicitas. 2018. "VOT or Quantity: What Matters More for the Voicing Contrast in German Regional Varieties? Results from Apparent-Time Analyses." *Journal of Phonetics* 71:468–86. doi:10.1016/j.wocn.2018.10.004.

Köndgen, David. 2019. "Meet the Stars: Arnie Hautnah - Arnold Schwarzenegger in Berlin." Accessed May 25, 2020. https://www.fitnessmanagement.de/digital/arnie-in-the-city/.

König, Ekkehard, and Volker Gast. 2007. *Understanding English-German Contrasts.* Grundlagen der Anglistik und Amerikanistik 29. Berlin: Erich Schmidt Verlag.

Köpke, Barbara. 2004. "Neurolinguistic Aspects of Attrition." *Journal of Neurolinguistics* 17 (1): 3–30. doi:10.1016/S0911-6044(03)00051-4.

Köpke, Barbara, and Monika S. Schmid. 2004. "Language Attrition: The Next Phase." In Schmid et al. 2004, 1–43.

Köpke, Barbara, Monika S. Schmid, Merel Keijzer, and Susan C. Dostert, eds. 2007. *Language Attrition: Theoretical Perspectives.* Studies in Bilingualism 33. Amsterdam: John Benjamins.

Krashen, Stephen. 1973. "Lateralization, Language Learning, and the Critical Period: Some New Evidence." *Language Learning* 23 (1): 63–74. doi:10.1111/j.1467-1770.1973.tb00097.x.

———. 1981. *Second Language Acquisition and Second Language Learning.* Oxford, UK: Pergamon Press.

Kraut, Rachel. 2014. "Origins of Language Attitudes Emerge in the Perception of Foreign-Accented Speech." *Journal of Second and Multiple Language Acquisition* 2 (3): 30–38.

Kraut, Rachel, and Stefanie Wulff. 2013. "Foreign-Accented Speech Perception Ratings: A Multifactorial Case Study." *Journal of Multilingual and Multicultural Development* 34:249–63. doi:10.1080/01434632.2013.767340.

Kuhl, Patricia K. 2000. “A New View of Language Acquisition.” *Proceedings of the National Academy of Sciences of the United States (PNAS)* 97 (22): 11850–57. doi:10.1073/pnas.97.22.11850.

———. 2007. “Cracking the Speech Code: How Infants Learn Language.” *Acoustical Science and Technology* 28 (2): 71–83. doi:10.1250/ast.28.71.

Kuhl, Patricia K., Barbara T. Conboy, Sharon Coffey-Corina, Denise Padden, Maritza Rivera-Gaxiola, and Tobey Nelson. 2008. “Phonetic Learning as a Pathway to Language: New Data and Native Language Magnet Theory Expanded (NLM-E).” *Philosophical Transactions of the Royal Society* 363 (1493): 979–1000. doi:10.1098/rstb.2007.2154.

Kuhl, Patricia K., and Paul Iverson. 1995. “Linguistic Experience and the 'Perceptual Magnet Effect'.” In Strange 1995, 121–54.

Kuzla, Claudia, and Mirjam Ernestus. 2011. “Prosodic Conditioning of Phonetic Detail in German Plosives.” *Journal of Phonetics* 39 (2): 143–55. doi:10.1016/j.wocn.2011.01.001.

Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. B. Christensen. 2017. “LmerTest Package: Tests in Linear Mixed Effects Models.” *Journal of Statistical Software* 82 (13): 1–26. doi:10.18637/jss.v082.i13.

Labov, William, Sharon Ash, and Charles Boberg. 2006. *The Atlas of North American English. Phonetics, Phonology and Sound Change.* Berlin, New York: Mouton de Gruyter.

Ladefoged, Peter. 1975. *A Course in Phonetics.* New York: Hartcourt Brace Jovanovich.

———. 2003. *Phonetic Data Analysis. An Introduction to Fieldwork and Instrumental Techniques.* Oxford: Blackwell.

———. 2005. *Vowels and Consonants.* Malden, MA: Blackwell.

Ladefoged, Peter, and Ian Maddieson. 1997. *The Sounds of the World's Languages.* Phonological theory. Oxford, UK: Blackwell.

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

Lambert, Richard D., and Barbara F. Freed, eds. 1982. *The Loss of Language Skills.* Rowley: Newbury House Publishers.

Lardiere, Donna. 1998a. “Case and Tense in the 'Fossilized' Steady State.” *Second Language Research* 14 (1): 1–26. doi:10.1191/026765898674105303.

———. 1998b. “Dissociating Syntax from Morphology in a Divergent L2 End-State Grammar.” *Second Language Research* 14 (4): 359–75. doi:10.1191/026765898672500216.

———. 2000. “Mapping Features to Forms in Second Language Acquisition.” In *Second Language Acquisition and Linguistic Theory*, edited by John Archibald, 102–29. Malden, MA: Blackwell.

Larsen-Freeman, Diane. 1997. “Chaos/complexity Science and Second Language Acquisition.” *Applied Linguistics* 18 (2): 141–65. doi:10.1093/applin/18.2.141.

———. 2000. “Second Language Acquisition and Applied Linguistics.” *Annual Review of Applied Linguistics* 20:165–81. doi:10.1017/S026719050020010X.

———. 2005. “Second Language Acquisition and the Issue of Fossilization: There Is No End, and There Is No State.” In Han and Odlin 2005, 189–200.

Laver, John. 1994. *Principles of Phonetics.* Cambridge: Cambridge University Press.

Lawson, Eleanor, Jane Stuart-Smith, James M. Scobbie, Malcah Yaeger-Dror, and Margaret Maclagan. 2011. “Liquids.” In Di Paolo and Yaeger-Dror 2011, 72–86.

Leeman, Adrian, Marie-José Kolly, Stephan Schmid, and Volker Dellwo, eds. 2015. *Trends in Phonetics and Phonology: Studies from German-Speaking Europe.* Bern: Peter Lang.

Lehiste, Ilse, and Gordon E. Peterson. 1961. “Some Basic Considerations in the Analysis of Intonation.” *Journal of the Acoustical Society of America* 33:419–25. doi:10.1121/1.1908681.

Lein, Tatjana, Tanja Kupisch, and Joost van de Weijer. 2016. “Voice Onset Time and Global Foreign Accent in German–French Simultaneous Bilinguals During Adulthood.” *International Journal of Bilingualism* 20 (6): 732–49. doi:10.1177/1367006915589424.

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

Lennon, Paul. 2008. “Contrastive Analysis, Error Analysis, Interlanguage.” In *Bielefeld Introduction to Applied Linguistics*, edited by Stephan Gramley and Vivian Gramley, 51–60. Bielefeld: Aisthesis.

Lenth, Russel. 2020. “Emmeans: Estimated Marginal Means, Aka Least-Squares Means. R Package Version 1.4.7.” https://CRAN.R-project.org/package=emmeans.

Levi, Susannah V. 2011. “Glides.” In *the Blackwell Companion to Phonology: General Issues and Segmental Phonology*, edited by Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, 341–66 Vol. 1. Malden, MA: Wiley-Blackwell.

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 (3): 1290–1305. doi:10.1121/1.3466879.

Li, Fangfang. 2013. “The Effect of Speakers' Sex on Voice Onset Time in Mandarin Stops.” *The Journal of the Acoustical Society of America* 133 (2): EL142-EL147. doi:10.1121/1.4778281.

Linville, Sue E., and Hilda B. Fisher. 1985. “Acoustic Characteristics of Women’s Voices with Advancing Age.” *Journal of Gerontology* 40 (3): 324–30. doi:10.1093/geronj/40.3.324.

Linville, Sue E., and Jennifer Rens. 2001. “Vocal Tract Resonance Analysis of Aging Voice Using Long-Term Average Spectra.” *Journal of Voice* 15 (3): 323–30. doi:10.1016/S0892-1997(01)00034-0.

Lippi-Green, Rosina. 2003. *English with an Accent: Language, Ideology, and Discrimination in the United States.* London, New York: Routledge.

Lisker, Leigh. 1975. “Is It VOT or a First-Formant Transition Detector?” *The Journal of the Acoustical Society of America* 57 (6): 1547–51. doi:10.1121/1.380602.

———. 1978. “In Qualified Defense of VOT.” *Language and Speech* 21 (4): 375–83. doi:10.1177/002383097802100413.

Lisker, Leigh, and Arthur S. Abramson. 1964. “A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements.” *Word* 20 (3): 384–422. doi:10.1080/00437956.1964.11659830.

———. 1967. “Some Effects of Context on Voice Onset Time in English Stops.” *Language and Speech* 10 (1): 1–28. doi:10.1177/002383096701000101.

Löfqvist, Anders. 1975. “Intrinsic and Extrinsic F0 Variations in Swedish Tonal Accents.” *Phonetica* 31:228–47. doi:10.1159/000259671.

Long, Michael H. 1990. “Maturational Constraints on Language Development.” *Studies in Second Language Acquisition* 12 (3): 251–85. doi:10.1017/S0272263100009165.

———. 1997. “Fossilization: Rigor Mortis in Living Linguistic Systems?” *Plenary address to the EUROSLA 97 Conference, Universitat Pompeu Fabra, Barcelona, May 22-4*.

———. 2003. “Stabilization and Fossilization in Interlanguage Development.” In *the Handbook of Second Language Acquisition*, edited by Catherine J. Doughty and Michael H. Long, 487–536. Oxford, UK: Blackwell.

Lousada, Marisa, Luis M. Jesus, and Andreia Hall. 2010. “Temporal Acoustic Correlates of the Voicing Contrast in European Portuguese Stops.” *Journal of the International Phonetic Association* 40 (3): 261–75. doi:10.1017/S0025100310000186.

Lowenstein, Joanna H., and Susan Nittrouer. 2008. “Patterns of Acquisition of Native Voice Onset Time in English-Learning Children.” *The Journal of the Acoustical Society of America* 124 (2): 1180–91. doi:10.1121/1.2945118.

Luick, Karl. 1904. *Deutsche Lautlehre. Mit Besonderer Berücksichtigung Der Sprechweise Wiens Und Der Österreichischen Alpenländer.* Leipzig, Wien: Deuticke.

Mack, Molly. 1989. "Consonant and Vowel Perception and Production: Early English-French Bilinguals and English Monolinguals." *Perception & Psychophysics* 46 (2): 187–200. doi:10.3758/bf03204982.

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

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 (1): 516–28. doi:10.1121/1.1377287.

Macken, Marlys A., and David Barton. 1980. "The Acquisition of the Voicing Contrast in English: A Study of Voice Onset Time in Word-Initial Stop Consonants." *Journal of Child Language* 7 (1): 41–74. doi:10.1017/s0305000900007029.

MacLeod, Andrea A. N., Carol Stoehl-Gammon, and Alicia B. Wassnik. 2009. "Production of High Vowels in Canadian English and Canadian French: A Comparison of Early Bilingual and Monolingual Speakers." *Journal of Phonetics* 37 (4): 374–87. doi:10.1016/j.wocn.2009.07.001.

MacWhinney, Brian, ed. 1999. *The Emergence of Language.* Mahwah, NJ: Lawrence Erlbaum Associates.

Major, Roy C. 1992. "Losing English as a First Language." *The Modern Language Journal* 76 (2): 190–208. doi:10.1111/j.1540-4781.1992.tb01100.x.

———. 2007. "Identifying a Foreign Accent in an Unfamiliar Language." *Studies in Second Language Acquisition* 29 (4): 539–56. doi:10.1017/S0272263107070428.

———. 2008. "Transfer in Second Language Phonology: A Review." In Hansen Edwards and Zampini 2008, 63–94.

Markel, John D., and Augustine H. Gray. 1976. *Linear Prediction of Speech.* Berlin: Springer.

Maschek. 2013. "Maschek – Schwarzeneggers Machtergreifung in 'Willkommen Österreich' [Video File]." Accessed May 22, 2020. https://www.youtube.com/watch?v=ItiTNR91IO0.

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 (4): 687–700. doi:10.1017/S136672891100071X.

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 (4): 41. doi:10.3390/languages5040041.

Mayr, Walter. 2013. “Frank Stronach's Crusade to Transform Austria.” Accessed February 23, 2020. https://www.spiegel.de/international/europe/frank-stronach-stirs-up-austrian-politics-with-new-party-a-885810.html.

McCarthy, Kathleen M., Merle Mahon, Stuart Rosen, and Bronwen G. Evans. 2014. “Speech Perception and Production by Sequential Bilingual Children: A Longitudinal Study of Voice Onset Time Acquisition.” *Child Development* 85 (5): 1965–80. doi:10.1111/cdev.12275.

McCarthy, Michael. 2001. *Issues in Applied Linguistics.* Cambridge: Cambridge University Press.

Mennen, Ineke. 2004. “Bi-Directional Interference in the Intonation of Dutch Speakers of Greek.” *Journal of Phonetics* 32 (4): 543–63. doi:10.1016/j.wocn.2004.02.002.

Mennen, Ineke, and Denise Chousi. 2018. “Prosody in First-Generation Adult Immigrants and Second-Generation Heritage-Langauge Users: The Timing of Prenuclear Rising Accents.” *Proceedings of the 9th International Conference on Speech Prosody 2018*, 828–32. doi:10.21437/SpeechProsody.2018-167.

Menschen im Porträt TV. 2018. “Frank Stronach - Das Große Interview.” Accessed August 23, 2019. https://www.youtube.com/watch?v=v5djSCqzSrk.

Mertins, Barbara. 2016. “The Use of Experimental Methods in Linguistic Research: Advantages, Problems and Possible Pitfalls.” In *Slavic Languages in Psycholinguistics: Chances and Challenges for Empirical and Experimental Research*, edited by Tanja Anstatt, Christina Clasmeier, and Anja Gattnar, 15–33. Tübinger Beiträge zur Linguistik 554. Tübingen: Francke.

Miller, Joanne L., Kerry P. Green, and Adam Reeves. 1986. “Speaking Rate and Segments: A Look at the Relation Between Speech Production and Speech Perception for the Voicing Contrast.” *Phonetica* 43:106–15. doi:10.1159/000261764.

Miller, Joanne L., and Lydia E. Volaitis. 1989. “Effect of Speaking Rate on the Perceptual Structure of a Phonetic Category.” *Perception & Psychophysics* 46 (6): 505–12. doi:10.3758/bf03208147.

Miller, Sean. 2012. “Acting Tips from Thespian Arnold Schwarzenegger.” Accessed April 03, 2020. https://www.backstage.com/magazine/article/acting-tips-thespian-arnold-schwarzenegger-51245/.

Mitchell, Rosamond, Florence Myles, and Emma Marsden. 2019. *Second Language Learning Theories.* 4th ed. New York: Routledge.

Mok, Peggy P. 2011. "Effects of Vowel Duration and Vowel Quality on Vowel-to-Vowel Coarticulation." *Language and Speech* 54 (4): 527–45. doi:10.1177/0023830911404961.

Moosmüller, Sylvia. 1991. *Hochsprache Und Dialekt in Österreich. Sozialphonologische Untersuchungen Zu Ihrer Abgrenzung in Wien, Graz, Salzburg Und Innsbruck.* Köln & Weimar: Böhlau.

———. 1995. "Assessment and Evaluation of Dialect and Standard in Austria." In Werlen 1995, 295–316.

———. 2007. "Vowels in Standard Austrian German." Habilitation thesis, University of Vienna, Vienna, Austria.

———. 2011. "Aussprachevarianten Im Österreichischen Standarddeutsch." In *Interpersonelle Kommunikation: Analyse Und Optimierung*, edited by Ines Bose and Baldur Neuber, 193–99. Frankfurt am Main: Peter Lang.

Moosmüller, Sylvia, and Catherine Ringen. 2004. "Voice and Aspiration in Austrian German Plosives." *Folia Linguistica* XXXVIII (1-2): 43–62. doi:10.1515/flin.2004.38.1-2.43.

Moosmüller, Sylvia, Carolin Schmid, and Julia Brandstätter. 2015. "Illustrations of the IPA: Standard Austrian German." *Journal of the International Phonetic Association* 45 (3): 339–48. doi:10.1017/S0025100315000055.

Mora, Joan C., James Keidel, and James E. Flege. 2015. "Effects of Spanish Use on the Production of Catalan Vowels by Early Spanish-Catalan Bilinguals." In *the Phonetics-Phonology Interface: Representations and Methodologies*, edited by Joaquin Romero and Maria Riera, 33–53. Amsterdam: John Benjamins.

Morris, Richard J., Christopher R. McCrea, and Kaileen D. Herring. 2008. "Voice Onset Time Differences Between Adult Males and Females: Isolated Syllables." *Journal of Phonetics* 36 (2): 308–17. doi:10.1016/j.wocn.2007.06.003.

Moulton, William G. 1962a. *The Sounds of English and German.* Contrastive structure series. Chicago: University of Chicago Press.

———. 1962b. "Toward a Classification of Pronunciation Errors." *The Modern Language Journal* 46 (3): 101–9. doi:10.1111/j.1540-4781.1962.tb01773.x.

Moyer, Alene. 1999. "Ultimate Attainment in L2 Phonology: The Critical Factors of Age, Motivation, and Instruction." *Studies in Second Language Acquisition* 21 (1): 81–108. doi:10.1017/S0272263199001035.

———. 2009. "Input as a Critical Means to an End: Quantity and Quality of Experience in L2 Phonological Attainment." In *Input Matters in SLA*, edited by Martha Young-Scholten and Thorsten Piske, 159–74. Bristol, UK: Multilingual Matters.

———. 2013. *Foreign Accent: The Phenomenon of Non-Native Speech.* Cambridge: Cambridge University Press.

Muhr, Rudolf. 2001. "Varietäten Des Österreichischen Deutsch." *Revue belge de philologie et l'histoire* 79 (3): 779–803.

———. 2003. "Language Change via Satellite: The Influence of German Television Broadcasting on Austrian German." *Journal of Historical Pragmatics* 4 (1): 103–27. doi:10.1075/jhp.4.1.06muh.

Munro, Murray J. 2003. "A Primer on Accent Discrimination in the Canadian Context." *TESL Canada Journal* 20 (2): 38–51. doi:10.18806/tesl.v20i2.947.

———. 2008. "Foreign Accent and Speech Intelligibility." In Hansen Edwards and Zampini 2008, 193–218.

Munro, Murray J., and Tracey M. Derwing. 1995a. "Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners." *Language Learning* 45 (1): 73–97. doi:10.1111/j.1467-1770.1995.tb00963.x.

———. 1995b. "Processing Time, Accent and Comprehensibility in the Perception of Native and Foreign Accented Speech." *Language and Speech* 38 (3): 289–306. doi:10.1177/002383099503800305.

———. 2009. "Putting Accent in Its Place: Rethinking Obstacles in Communication." *Language Teaching* 42 (4): 476–90. doi:10.1017/S026144480800551X.

Munro, Murray J., Tracey M. Derwing, and Clifford S. Burgess. 2003. "The Detection of Foreign Accent in Backwards Speech." In Solé, Recasens, and Romero 2003, 535–38.

Munro, Murray J., Tracey M. Derwing, and Susan L. Morton. 2006. "The Mutual Intelligibility of L2 Speech." *Studies in Second Language Acquisition* 28 (1): 111–31. doi:10.1017/S0272263106060049.

Munro, Murray J., Tracey M. Derwing, and Kazuya Saito. 2013. "English L2 Vowel Acquisition over Seven Years." In *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference*, edited by John M. Levis and Kimberly LeVelle, 112–19. Ames, IA: Iowa State University.

Munro, Murray J., Tracey M. Derwing, and Kyoko Sato. 2006. "Salient Accents, Covert Attitudes: Consciousness-Raising for Pre-Service Second Language Teachers." *Prospect: An Australian Journal of TESOL* 21 (1): 67–79.

Mwangi, Samuel, Werner Spiegel, Florian Hönig, Tino Haderlein, Andreas Maier, and Elmar Nöth. 2009. "Effects of Vocal Aging on Fundamental Frequency and Formants." *Proceedings of the International Conference on Acoustics NAG/DAGA, Rotterdam*, 1761–64.

Nagle, Charles. 2018. “Motivation, Comprehensibility, and Accentedness in L2 Spanish: Investigating Motivation as a Time-Varying Predictor of Pronunciation Development.” *Modern Language Journal* 102 (1): 199–217. doi:10.1111/modl.12461.

Nakuma, Constancio K. 2005. “Researching Fossilization and Second Language (L2) Attrition: Easy Questions, Difficult Answers.” In Han and Odlin 2005, 21–34.

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

Nemser, William. 1971. “Approximative Systems of Foreign Language Learners.” *International Journal of Applied Linguistics* 9:115–23. doi:10.1515/iral.1971.9.2.115.

Noble, Kimberley. 2014. “Frank Stronach (Profile).” Accessed June 25, 2020. https://thecanadianencyclopedia.ca/en/article/frank-stronach-profile.

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

———. 2006. “Cross-Linguistic Influence.” In *the Handbook of Second Language Acquisition.,* edited by Catherine J. Doughty and Michael H. Long, 436–86. Malden, MA: Blackwell.

OE24.TV. 2018. “Fellner! Live: Frank Stronach Im Großen Interview [Video File].” Accessed June 25, 2020. https://www.youtube.com/watch?v=u2nzbP9K5VE.

Oh, Eunjin. 2011. “Effects of Speaker Gender on Voice Onset Time in Korean Stops.” *Journal of Phonetics* 39 (1): 59–67. doi:10.1016/j.wocn.2010.11.002.

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 (2): 156–67. doi:10.1016/j.wocn.2011.01.002.

Ohala, Diane K. 2008. “Phonological Acquisition in a First Language.” In Hansen Edwards and Zampini 2008, 19–39.

Olive, Joseph P., Alice Greenwood, and John Coleman. 1993. *Acoustics of American English Speech: A Dynamic Approach.* New York NY: Springer.

Opitz, Conny. 2011. “L1 Attrition and L2 Acquisition in a Second Language Environment.” PhD thesis, Trinity College Dublin, Dublin, Ireland.

———. 2012. “A Dynamic Perspective on Late Bilinguals’ Linguistic Development in an L2 Environment.” *International Journal of Bilingualism* 17 (6): 701–15. doi:10.1177/1367006912454621.

ORF Steiermark. 2016. “Schwarzenegger-Mentor Gerstl Gestorben.” Accessed July 02, 2019. https://steiermark.orf.at/v2/news/stories/2808979/.

Ortega, Lourdes. 2013. *Understanding Second Language Acquisition.* London, New York: Routledge.

Pavlenko, Aneta. 2004. “L2 Influence and L1 Attrition in Adult Bilingualism.” In Schmid et al. 2004, 47–59.

Pavlenko, Aneta, and Barbara C. Malt. 2011. “Kitchen Russian: Cross-Linguistic Differences and First-Language Object Naming by Russian-English Bilinguals.” *Bilingualism: Language and Cognition* 14 (1): 19–45. doi:10.1017/S136672891000026X.

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

Perlmutter, Marilyn. 1989. “Intelligibility Rating of L2 Speech Pre- and Postintervention.” *Perceptual and Motor Skills* 68 (2): 515–21. doi:10.2466/pms.1989.68.2.515.

Peterson, Gordon E., and Harold L. Barney. 1952. “Control Methods Used in a Study of the Vowels.” *The Journal of the Acoustical Society of America* 24 (2): 175–84. doi:10.1121/1.1906875.

Petrosino, Linda, Roger D. Colcord, Karen B. Kurcz, and Robert J. Yonker. 1993. “Voice Onset Time of Velar Stop Productions in Aged Speakers.” *Perceptual and Motor Skills* 76 (1): 83–88. doi:10.2466/pms.1993.76.1.83.

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 (1): 49–71. doi:10.1159/000056205.

Piske, Thorsten, and Ian R. A. MacKay. 1999. “Age and L1 Use Effects on Degree of Foreign Accent in English.” In *Proceedings of the XIVth International Congress of Phonetic Sciences: ICPhS 99, San Francisco, 1. - 7. August, 1999*, edited by John J. Ohala, 1433–36. Berkeley, California: University of California Department of Linguistics.

Piske, Thorsten, Ian R. A. MacKay, and James E. Flege. 2001. “Factors Affecting the Degree of Foreign Accent in an L2: A Review.” *Journal of Phonetics* 29 (2): 191–215. doi:10.1006/jpho.2001.0134.

Pömp, Julian, and Christoph Draxler. 2017. “OCTRA. A Configurable Browser-Based Editor for Orthographic Transcription.” *Proceedings of the 13th Conference*

of Phonetics and Phonology in the German Language Area (P&P13), Berlin, Germany, 145–48.

Poon, Pamela G., and Catherine A. Mateer. 1985. "A Study of VOT in Nepali Stop Consonants." *Phonetica* 42:39–47. doi:10.1159/000261736.

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

Preston, Malcolm S., Grace H. Yeni-Komshian, Rachel E. Stark, and Diane K. Port. 1968. "Developmental Studies of Voicing in Stops." *Haskins Laboratories Status Report on Speech Research, SR13/14*, 181–84.

Quora. 2017. "Can Arnold Schwarzenegger Speak German?" Accessed January 25, 2021. https://www.quora.com/Can-Arnold-Schwarzenegger-speak-German.

R Core Team. 2020. "R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria." Accessed January 02, 2021. https://www.R-project.org/.

Rami, Manish K., Joseph Kalinowski, Andrew Stuart, and Michael P. Rastatter. 1999. "Voice Onset Times and Burst Frequencies of Four Velar Stop Consonants in Gujarati." *The Journal of the Acoustical Society of America* 106 (6): 3736–58. doi:10.1121/1.428226.

Rastatter, Michael P., and Richard D. Jacques. 1990. "Formant Frequency Structure of the Aging Male and Female Vocal Tract." *Folia Phoniatrica et Logopaedica* 42 (6): 312–19. doi:10.1159/000266088.

Reinisch, Eva. 2005. "What We (Don't) Perceive as Foreign Accent." *Grazer Linguistische Studien* 64: 69–85.

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, 9–37. Amsterdam: John Benjamins.

———. 2017. "The Influence of Age on Estimating Sound Change Acoustically from Longitudinal Data." In *Panel Studies of Language Variation and Change*, edited by Suzanne E. Wagner and Isabelle Buchstaller, 129–52. London: Routledge.

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 (7-8): 638–51. doi:10.1016/j.specom.2010.02.012.

Richards, Jack C. 1971. "A Non-Contrastive Approach to Error Analysis." *ELT Journal* XXV (3): 204–19. doi:10.1093/elt/XXV.3.204.

Riney, Timothy J., and James E. Flege. 1998. "Changes over Time in Global Foreign Accent and Liquid Identifiability and Accuracy." *Studies in Second Language Acquisition* 20 (2): 213–43. doi:10.1017/S0272263198002058.

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

Robb, Michael P., Harvey Gilbert, and Jay Lerman. 2005. "Influence of Gender and Environmental Setting on Voice Onset Time." *Folia Phoniatrica et Logopaedica* 57 (3): 125–33. doi:10.1159/000084133.

Rosner, Burton S., Luis E. López-Bascuas, José E. García-Albea, and Richard P. Fahey. 2000. "Voice-Onset Times for Castilian Spanish Initial Stops." *Journal of Phonetics* 28 (2): 217–24. doi:10.1006/jpho.2000.0113.

Rothman, Jason, and Jeanine Treffers-Daller. 2014. "A Prolegomenon to the Construct of the Native Speaker: Heritage Speaker Bilinguals Are Natives Too!". *Applied Linguistics* 35 (1): 93–98. doi:10.1093/applin/amt049.

Rutherford, William. 1982. "Markedness in Second Language Acquisition." *Language Learning* 32 (1): 85–108. doi:10.1111/j.1467-1770.1982.tb00520.x.

Ryalls, Jack, Marni Simon, and Jerry Thomason. 2004. "Voice Onset Time Production in Older Caucasian- and African-Americans." *Journal of Multilingual Communication Disorders* 2 (1): 61–67. doi:10.1080/1476967031000090980.

Ryalls, Jack, Allison Zipprer, and Penelope Baldauff. 1997. "A Preliminary Investigation of the Effects of Gender and Race on Voice Onset Time." *Journal of Speech, Language, and Hearing Research* 40 (3): 642–45. doi:10.1044/jslhr.4003.642.

Saito, Kazuya. 2012. "Effects of Instruction on L2 Pronunciation Development: A Synthesis of 15 Quasi-Experimental Intervention Studies." *TESOL Quarterly* 46 (4): 842–54. doi:10.1002/tesq.67.

Saito, Kazuya, and Roy Lyster. 2012. "Effects of Form-Focused Instruction and Corrective Feedback on L2 Pronunciation Development of /ɹ/ by Japanese Learners of English." *Language Learning* 62 (2): 595–633. doi:10.1111/j.1467-9922.2011.00639.x.

Saito, Kazuya, Pavel Trofimovich, and Talia Isaacs. 2016. "Second Language Speech Production: Investigating Linguistic Correlates of Comprehensibility and Accentedness for Learners at Different Ability Levels." *Applied Psycholinguistics* 37 (2): 217–40. doi:10.1017/S0142716414000502.

Sancier, Michele L., and Carol A. Fowler. 1997. "Gestural Drift in a Bilingual Speaker of Brazilian Portuguese and English." *Journal of Phonetics* 25 (4): 421–36. doi:10.1006/jpho.1997.0051.

Saville-Troike, Muriel. 2012. *Introducing Second Language Acquisition.* 2nd ed. Cambridge: Cambridge University Press.

Schachter, Jacquelyn. 1974. "An Error in Error Analysis." *Language Learning* 24 (2): 205–14. doi:10.1111/j.1467-1770.1974.tb00502.x.

———. 1988. "Second Language Acquisition and Its Relationship to Universal Grammar." *Applied Linguistics* 9 (3): 219–35. doi:10.1093/applin/9.3.219.

———. 1990. "On the Issue of Completeness in Second Language Acquisition." *Second Language Research* 6 (2): 93–124. doi:10.1177/026765839000600201.

Schachter, Jacquelyn, and Marianne Celce-Murcia. 1977. "Some Reservations Concerning Error Analysis." *TESOL Quarterly* 11 (4): 441–51. doi:10.2307/3585740.

Schiel, Florian. 1999. "Automatic Phonetic Transcription of Non-Prompted Speech." *Proceedings of the XIVth International Congress of Phonetic Sciences, San Francisco*, 607–10. doi:10.5282/UBM/EPUB.13682.

Schmid, Carolin, Sylvia Moosmüller, and Christian H. Kaseß. 2015. "Sociophonetics of the Velarized Lateral in the Viennese Dialect." *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015). Glasgow*.

Schmid, Monika S. 2002. *First Language Attrition, Use and Maintenance: The Case of German Jews in Anglophone Countries.* Studies in Bilingualism 24. Amsterdam: John Benjamins.

———. 2007. "The Role of L1 Use for L1 Attrition." In Köpke et al. 2007, 135–53.

———. 2008. "Defining Language Attrition." *Babylonia* 2 (8): 9–12.

———. 2011. "Contact X Time: External Factors and Variability in L1 Attrition." In *Modeling Bilingualism: From Structure to Chaos. In Honor of Kees De Bot*, edited by Monika S. Schmid and Wander Lowie, 155–76. Studies in Bilingualism 43. Amsterdam, Philadelphia: John Benjamins.

———. 2012. "The Impact of Age and Exposure on Bilinugal Development in International Adoptees and Family Migrants: A Perspective from Holocaust Survivors." *Linguistic Approaches to Bilingualism* 2 (2): 177–208. doi:10.1075/lab.2.2.03sch.

———. 2016. "Research Timeline: First Language Attrition." *Language Teaching* 49 (2): 186–212. https://doi.org/10.1017/S0261444815000476.

Schmid, Monika S., and Holger Hopp. 2014. "Comparing Foreign Accent in L1 Attrition and L2 Acquisition: Range and Rater Effects." *Language Testing* 31 (3): 367–88. doi:10.1177/0265532214526175.

Schmid, Monika S., and Scott H. Jarvis. 2014. "Lexical Access and Lexical Diversity in First Language Attrition." *Bilingualism: Language and Cognition* 17 (4): 729–48. doi:10.1017/S1366728913000771.

Schmid, Monika S., and Barbara Köpke. 2007. “Bilingualism and Attrition.” In Köpke et al. 2007, 1–7.

———. 2017a. “The Relevance of First Language Attrition to Theories of Bilingual Development.” *Linguistic Approaches to Bilingualism* 7 (6): 637–67. doi:10.1075/lab.17058.sch.

———. 2017b. “When Is a Bilingual an Attriter? Response to the Commentaries.” *Linguistic Approaches to Bilingualism* 7 (6): 763–70. doi:10.1075/lab.17059.sch.

Schmid, Monika S., Barbara Köpke, Merel Keijzer, and Lina Weilemar, eds. 2004. *First Language Attrition: Interdisciplinary Perspectives on Methodological Issues.* Studies in Bilingualism 28. Amsterdam: John Benjamins.

Schötz, Susanne. 2006. “Perception, Analysis and Synthesis of Speaker Age.” PhD thesis, Lund University, Lund, Sweden.

Schuller, Björn, Anton Batliner, Christian Bergler, Florian B. Pokorny, Jarek Krajewski, Margaret Cychosz, Ralf Vollmann et al. 2019. “The INTERSPEECH 2019 Computational Paralinguistics Challenge: Styrian Dialects, Continuous Sleepiness, Baby Sounds & Orca Activity.” *Proceedings of the 20th Annual Conference of the International Speech Communication Association (Interspeech), Graz, Austria*, 2378–82.

Schumann, John F. 1978. *The Pidginization Process: A Model for Second Language Acquisition.* Rowley: Newbury House Publishers.

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

Scobbie, James M. 2005. “Intraspeaker Variation as the Long Term Outcome of Dialectally Varied Input: Speech Production Evidence for Fine-Grained Plasticity.” In *ISCA Workshop on Plasticity in Speech Perception, London*, edited by Valerie L. Hazan and Paul Iverson, 56–59.

———. 2006. “Flexibility in the Face of Incompatible English VOT Systems.” In *Laboratory Phonology 8: Varieties of Phonological Competence*, edited by Louis Goldstein, Douglas H. Whalen, and Catherine T. Best, 367–92. New Haven: Mouton de Gruyter.

Scovel, Thomas. 1969. “Foreign Accents, Language Acquisition, and Cerebral Dominance.” *Language Learning* 19 (3-4): 245–53. doi:10.1111/j.1467-1770.1969.tb00466.x.

———. 1988. *A Time to Speak: A Psycholinguistic Inquiry into the Critical Period for Human Speech.* Cambridge: Newbury House.

Sebastian, Swapna, Sonia Babu, Neethu E. Oommen, and Achamma Ballraj. 2012. "Acoustic Measurements of Geriatric Voice." *Journal of Laryngology & Voice* 2 (2): 81–84. doi:10.4103/2230-9748.106984.

Seliger, Herbert W., and Robert M. Vago, eds. 1991. *First Language Attrition.* Cambridge, MA: Cambridge University Press.

Selinker, Larry. 1979. "Interlanguage." In *Studies in Contrastive Linguistics and Error Analysis,* edited by Dietrich Nehls, 55–77. Studies in Descriptive Linguistics 2. Heidelberg: Julius Groos.

Selinker, Larry, and Zhaohong Han. 2001. "Fossilization: Moving the Concept into Empirical Longitudinal Study." In *Studies in Language Testing: Experimenting with Uncertainty. Essays in Honour of Alan Davies*, edited by Catherine Elder, Adam Brown, Elisabeth Grove, Kathryn Hill, Noriko Iwashita, Tom Lumley, Tim McNamara, and Kieran J. O'Loughlin, 276–91. Cambridge: Cambridge University Press.

Selinker, Larry, and Usha Lakshmanan. 1992. "Language Transfer and Fossilization: The Multiple Effects Principle." In *Language Transfer in Language Learning,* edited by Susan M. Gass and Larry Selinker, 97–116. Amsterdam, Philadelphia: John Benjamins.

Selinker, Larry, and John T. Lamendella. 1978. "Two Perspectives on Fossilization in Interlanguage Learning." *Interlanguage Studies Bulletin* 3 (2): 143–91.

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

Sharwood Smith, M., and Paul van Buren. 1991. "First Language Attrition and the Parameter Setting Model." In Seliger and Vago 1991, 17–30.

Shultz, Amanda A., Alexander L. Francis, and Fernando Llanos. 2012. "Differential Cue Weighting in Perception and Production of Consonant Voicing." *The Journal of the Acoustical Society of America* 132:EL95-EL101. doi:10.1121/1.4736711.

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 (3): 377–408. doi:10.1177/0267658309104580.

Singleton, David M. 2005. "The Critical Period Hypothesis: A Coat of Many Colours." *IRAL: International Review of Applied Linguistics in Language Teaching* 43 (4): 269–85. doi:10.1515/iral.2005.43.4.269.

———. 2007. "The Critical Period Hypothesis: Some Problems." *Interlingüìstica* 17: 48–56.

Singleton, David M., and Zsolt Lengyel, eds. 1995. *The Age Factor in Second Language Acquisition.* Clevedon: Multilingual Matters.

Singleton, David M., and Justyna Leśniewska. 2021. "The Critical Period Hypothesis for L2 Acquisition: An Unfalsifiable Embarrassment?" *Languages* 6 (3): 1–15. doi:10.3390/languages6030149.

Skinner, Burrhus F. 1957. *Verbal Behavior.* New York: Appleton-Century-Crofts.

———. 1974. *About Behaviorism.* New York: Knopf.

Smiljanic, Rajka, and Ann R. Bradlow. 2008. "Stability of Temporal Contrasts Across Speaking Styles in English and Croatian." *Journal of Phonetics* 36 (1): 91–113. doi:10.1016/j.wocn.2007.02.002.

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 (4): 522–29. doi:10.1044/jshr.3004.522.

Smith, Larry E. 1992. "Spread of English and Issues of Intelligibility." In *the Other Tongue: English Across Cultures,* edited by Braj B. Kachru. 2nd ed., 75–90. Urbana: University of Illinois Press.

Smith, Larry E., and Khalilullah Rafiqzad. 1979. "English for Cross-Cultural Communication: The Question of Intelligibility." *TESOL Quarterly* 13:371–80. doi:10.2307/3585884.

Smith, Linda B., and Esther Thelen. 2003. "Development as a Dynamic System." *Trends in Cognitive Sciences* 7 (8): 343–48. doi:10.1016/s1364-6613(03)00156-6.

Snow, Catherine E., and Marian Hoefnagel-Höhle. 1977. "Age Differences in the Pronunciation of Foreign Sounds." *Language and Speech* 20 (4): 357–65. doi:10.1177/002383097702000407.

———. 1978. "The Critical Period for Language Acquisition: Evidence from Second Language Learning." *Child Development* 49 (4): 1114–28. doi:10.2307/1128751.

Solé, Maria-Josep, Daniel Recasens, and Joaquin Romero, eds. 2003. *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS 15).* Barcelona, Spain: Causal Productions.

Statistik Austria. 2019. "Bevölkerung Nach Staatsangehörigkeit Und Geburtsland." Accessed December 13, 2019. http://www.statistik.at/web_de/statistiken/menschen_und_gesellschaft/bevoelkerung/bevoelkerungsstruktur/bevoelkerung_nach_staatsangehoerigkeit_geburtsland/index.html.

Stevens, Kenneth N. 1993. "Models for the Production and Acoustics of Stop Consonants." *Speech Communication* 13 (3-4): 367–75. doi:10.1016/0167-6393(93)90035-J.

Stock, Dieter. 1971. *Untersuchungen Zur Stimmhaftigkeit Hochdeutscher Phonemrealisationen.* Hamburg: Buske.

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 (4): 483–518. doi:10.1177/0267658317704261.

Stolberg, Doris, and Alexandra Münch. 2010. “'Die Muttersprache Vergisst Man Nicht' - or Do You? A Case Study in L1 Attrition and Its (Partial) Reversal.” *Bilingualism: Language and Cognition* 13 (1): 19–31. doi:10.1017/S1366728909990332.

Strange, Winifred, ed. 1995. *Speech Perception and Linguistic Experience: Issues in Cross-Language Research.* Baltimore, MD: York Press.

———. 2007. “Cross-Language Phonetic Similarity: Theoretical and Methodological Issues.” In Bohn and Munro 2007, 35–55.

Strevens, Peter. 1969. “Two Ways of Looking at Error-Analysis.” *Revised version of a paper delivered at GAL meeting, Stuttgart*. http://files.eric.ed.gov/fulltext/ED037714.pdf.

Stuart-Smith, Jane, Gwilym Pryce, Claire Timmins, and Barrie Gunter. 2013. “Television Can Also Be a Factor in Language Change: Evidence from an Urban Dialect.” *Language* 89 (3): 501–36. doi:10.1353/lan.2013.0041.

Stummer, David. 2011. “Maschek - Arnold Schwarzenegger Mit Barbara Karlich [Video File].” Accessed April 03, 2020. https://www.youtube.com/watch?v=I8oZCTUhoxs.

Swartz, Bradford L. 1992. “Gender Differences in Voice Onset Time.” *Perceptual and Motor Skills* 75 (3): 983–92. doi:10.2466/pms.1992.75.3.983.

Sweeting, Patricia M., and Ronald J. Baken. 1982. “Voice Onset Time in a Normal-Aged Population.” *Journal of Speech and Hearing Research* 25 (1): 129–34. doi:10.1044/jshr.2501.129.

Tahta, Sonia, Margaret Wood, and Kate Loewenthal. 1981. “Foreign Accents: Factors Relating to Transfer of Accent from the First Language to a Second Language.” *Language and Speech* 24 (3): 265–72. doi:10.1177/002383098102400306.

Taylor, Dennis Q. 1975. “The Inadequacy of Bipolarity and Distinctive Features: The German 'Voiced/voiceless' Consonants.” In *the Second Lacus Forum,* edited by Peter A. Reich, 107–19. Columbia, SC: Hornbeam Press.

Taylor, Wilson L. 1953. “Cloze Procedure: A New Tool for Measuring Readability.” *Journalism Quarterly* 30 (4): 415–33. doi:10.1177/107769905303000401.

The Times. 2008. “Wag Bennett: Bodybuilder Who Helped Arnold Schwarzenegger.” Accessed May 22, 2020. https://www.thetimes.co.uk/article/wag-bennett-bodybuilder-who-helped-arnold-schwarzenegger-9w53bq55tpw.

Thelen, Esther. 1995. “Time-Scale Dynamics and the Development of an Embodied Cognition.” In *Mind as Motion*, edited by Robert Port and Timothy van Gelder, 69–100. Cambridge: MIT Press.

Thelen, Esther, and Linda B. Smith. 1994. *A Dynamic Systems Approach to the Development of Cognition and Action.* Cambridge, MA: MIT Press.

———. 2006. “Dynamic Systems Theories.” In *Handbook of Child Psychology: Theoretical Models of Human Development*, edited by Richard M. Lerner and William Damon, 258–312. Hoboken, NJ: John Wiley and Sons.

Theodore, Rachel M., Joanne L. Miller, and David DeSteno. 2009. “Individual Talker Differences in Voice-Onset-Time: Contextual Influences.” *Journal of the Acoustical Society of America* 125 (6): 3974–82. doi:10.1121/1.3106131.

Thompson, Irene. 1991. “Foreign Accents Revisited: The English Pronunciation of Russian Immigrants.” *Language Learning* 41 (2): 177–204. doi:10.1111/j.1467-1770.1991.tb00683.x.

Thomson, Ron I. 2011. “Computer Assisted Pronunciation Training: Targeting Second Language Vowel Perception Improves Pronunciation.” *CALICO Journal* 28 (3): 744–65. doi:10.11139/cj.28.3.744-765.

———. 2018. “Measurements of Accentedness, Intelligibility and Comprehensibility.” In *Assessment in Second Language Pronunciation*, edited by Okim Kang and April Ginther, 11–29. New York: Routledge.

Thomson, Ron I., and Tracey M. Derwing. 2014. “The Effectiveness of L2 Pronunciation Instruction: A Narrative Review.” *Applied Linguistics* 36 (3): 326–44. doi:10.1093/applin/amu076.

———. 2016. “Is Phonemic Training Using Nonsense or Real Words More Effective?” In *Proceedings of the 7th Pronunciation in Second Language Learning and Teaching Conference*, edited by John M. Levis, Huong Le, Ivana Lucic, Evan Simpson, and Sonca Vo, 88–97. Ames, IA: Iowa State University.

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 (3): 215–29. doi:10.1016/S0021-9924(97)00053-1.

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. doi:10.1016/j.wocn.2017.05.006.

Torre, Peter, and Jessica A. Barlow. 2009. “Age-Related Changes in Acoustic Characteristics of Adult Speech.” *Journal of Communication Disorders* 45 (5): 324–33. doi:10.1016/j.jcomdis.2009.03.001.

Trofimovich, Pavel, and Talia Isaacs. 2012. "Disentangling Accent from Comprehensibility." *Bilingualism: Language and Cognition* 15 (4): 905–16. doi:10.1017/S1366728912000168.

Trofimovich, Pavel, Sara Kennedy, and Jennifer A. Foote. 2015. "Variables Affecting L2 Pronunciation Development." In *the Handbook of English Pronunciation*, edited by Marnie Reed and John M. Levis, 353–73. Malden, MA: Wiley.

Trubetzkoy, Nikolai. (1939) 1977. *Grundzüge Der Phonologie.* 6th ed. Göttingen: Vandenhoeck & Ruprecht.

Tsimpli, Ianthi, Antonella Sorace, Caroline Heycock, and Francesca Filiaci. 2004. "First Language Attrition and Syntactic Subjects: A Study of Greek and Italian Near-Native Speakers of English." *International Journal of Bilingualism* 8 (3): 257–77. doi:10.1177/13670069040080030601.

Tsukada, Kimiko, David Birdsong, Ellen Bialystok, Molly Mack, Hyekyung Sung, and James E. Flege. 2005. "A Developmental Study of English Vowel Production and Perception by Native Korean Adults and Children." *Journal of Phonetics* 33 (3): 263–90. doi:10.1016/j.wocn.2004.10.002.

Ulbrich, Christiane, and Mikhail Ordin. 2014. "Can L2-English Influence L1-German? The Case of Post-Vocalic /r/." *Journal of Phonetics* 45 (1): 26–42. doi:10.1016/j.wocn.2014.02.008.

USC Schwarzenegger Institute. 2013. "Governor Arnold Schwarzenegger, Chairman." Accessed December 18, 2021. http://schwarzeneggerinstitute.com/about-the-institute/about-the-institute-leadership.

Valdés, Guadalupe. 2001. *Learning and Not Learning English; Latino Students in American Schools.* New York: Teachers College Press.

Van Alphen, Petra M., and Roel Smits. 2004. "Acoustical and Perceptual Analysis of the Voicing Distinction in Dutch Initial Plosives: The Role of Prevoicing." *Journal of Phonetics* 32 (4): 455–91. doi:10.1016/j.wocn.2004.05.001.

Van Dijk, Marijn, Marjolijn Verspoor, and Wander Lowie. 2011. "Variability and DST." In *a Dynamic Approach to Second Language Development: Methods and Techniques*, edited by Marjolijn Verspoor, Kees de Bot, and Wander Lowie, 55–84. Language Learning & Language Teaching 29. Amsterdam: John Benjamins.

Van Els, Theo. 1986. "An Overview of European Research on Language Attrition." In *Language Attrition in Progress*, edited by Bert Weltens, Kees de Bot, and Theo van Els, 3–18. Dordrecht: Foris.

Van Geert, Paul. 2009. "A Comprehensive Dynamic Systems Theory of Language Development." In *Language Development over the Lifespan*, edited by Kees de Bot and Robert W. Schrauf, 60–104. New York, London: Routledge.

Vanhove, Jan. 2013. “The Critical Period Hypothesis in Second Language Acquisition: A Statistical Critique and a Reanalysis.” *PLoS ONE* 8 (7). doi:10.1371/journal.pone.0069172.

Vienna.at. 2019. “Arnold Schwarzenegger Auf Wien-Besuch Bei Kanzler Kurz.” Accessed May 21, 2020. https://www.vienna.at/arnold-schwarzenegger-auf-wien-besuch-bei-kanzler-kurz/6077765.

Volaitis, Lydia E., and Joanne L. Miller. 1992. “Phonetic Prototypes: Influence of Place of Articulation and Speaking Rate on the Internal Structure of Voicing Categories.” *The Journal of the Acoustical Society of America* 92:723–35. doi:10.1121/1.403997.

Von Uslar, Moritz. 2012. “'Hallo, Herr Schwarzenegger?'.” Accessed September 11, 2019. https://www.zeit.de/2012/42/Arnold-Schwarzenegger-Biografie.

Wan, Kristopher. 2017. “Accents in Hollywood: Arnold Schwarzenegger.” Accessed December 05, 2019. https://www.lucidaccent.com/post/2017/03/19/accents-in-hollywood-arnold-schwarzenegger.

Wang, Hongyan, and Vincent J. van Heuven. 2006. “Acoustical Analysis of English Vowels Produced by Chinese, Dutch and American Speakers.” *Linguistics in the Netherlands* 23 (1): 237–48. doi:10.1075/AVT.23.23WAN.

Wardhaugh, Ronald. 1974. *Topics in Applied Linguistics.* Rowley, MA: Newbury House.

Wei, Li. 2007. “Dimensions of Bilingualism.” In Wei 2007, 3–22.

———, ed. 2007. *The Bilingualism Reader.* 2nd ed. London, New York: Routledge.

Wei, Longxing. 2009. “The Composite Nature of Interlanguage as a Developing System.” *Research in Language* 7 (1): 5–30. doi:10.2478/v10015-009-0002-9.

Weinreich, Uriel. (1953) 1974. *Languages in Contact: Findings and Problems.* 6th ed. The Hague, Paris: Mouton de Gruyter.

Werlen, Iwar, ed. 1995. *Verbale Kommunikation in Der Stadt.* Tübingen: Narr.

Whiteside, Sandra P., Luisa Henry, and Rachel Dobbin. 2004. “Sex Differences in Voice Onset Time: A Developmental Study of Phonetic Context Effects in British English.” *The Journal of the Acoustical Society of America* 116 (2): 1179–83. doi:10.1121/1.1768256.

Whiteside, Sandra P., and Jeni Marshall. 2001. “Developmental Trends in Voice Onset Time: Some Evidence for Sex Differences.” *Phonetica* 58 (3): 196–210. doi:10.1159/000056199.

Wiesinger, Peter. 1967. *Mundart Und Geschichte in Der Steiermark: Ein Beitrag Zur Dialektgeographie Eines Österreichischen Bundeslandes.* Marburg/Lahn: N.G. Elwert Verlag.

———. 1990. “Standardsprache Und Mundarten in Österreich.” In *Deutsche Gegenwartssprache: Tendenzen Und Perspektiven*, edited by Gerhard Stickel, 218–32. Berlin, New York: Mouton de Gruyter.

———. 1996. “Das Österreichische Deutsch Als Eine Varietät Der Deutschen Sprache.” *Die Unterrichtspraxis/Teaching German* 29 (2): 154–64. doi:10.2307/3531825.

———. 2008. *Das Österreichische Deutsch in Gegenwart Und Geschichte.* 2nd ed. Wien: LIT.

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

Wode, Henning. 1978. “Developmental Sequences in Naturalistic L2 Acquisition.” In Hatch 1978, 101–17.

———. 1980. “Phonology in L2 Acquisition.” In *Second Language Development: Trends and Issues*, edited by Sascha W. Felix, 123–36. Tübingen: Narr.

Wunder, Eva-Maria. 2011. “Crosslinguistic Influence in Multilingual Language Acquisition: Phonology in Third or Additional Language Acquisition.” In *New Trends in Crosslinguistic Influence and Multilingualism Research*, edited by Gessica de Angelis and Jean-Marc Dewaele, 105–28. Bristol, UK: Multilingual Matters.

Yağmur, Kutlay. 1997. *First Language Attrition Among Turkish Speakers in Sydney.* Studies in Multilingualism 7. Tilburg, Netherlands: Tilburg University Press.

Yağmur, Kutlay, Kees de Bot, and Hubert Korzilius. 1999. “Language Attrition, Language Shift and Ethnolinguistic Vitality of Turkish in Australia.” *Journal of Multilingual and Multicultural Development* 20 (1): 51–69. doi:10.1080/01434639908666369.

Yang, Byunggon. 1996. “A Comparative Study of American English and Korean Vowels Produced by Male and Female Speakers.” *Journal of Phonetics* 24 (2): 245–61. doi:10.1006/JPHO.1996.0013.

Yao, Yao. 2007. “Closure Duration of VOT of Word-Initial Voiceless Plosives in English in Spontaneous Connected Speech.” *UC Berkeley Phonology Lab Annual Report* 3 (3): 183–225. doi:10.5070/P71hs7h769.

Yavaş, Mehmet, and Renée Wildermuth. 2006. “The Effects of Place of Articulation and Vowel Height in the Acquisition of English Aspirated Stops by Spanish Speakers.” *International Journal of Applied Linguistics* 44 (3). doi:10.1515/IRAL.2006.011.

Yilmaz, Gülsen, and Monika S. Schmid. 2013. “L1 Accessibility Among Turkish-Dutch Bilinguals.” *The Mental Lexicon* 7 (3): 249–74. doi:10.1075/ml.7.3.01yil.

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 (2): 152–67. doi:10.1177/0023830914522994.

Zeman, Dalibor. 2009. *Überlegungen Zur Deutschen Sprache in Österreich. Linguistische, Sprachpolitische Und Soziolinguistische Aspekte Der Österreichischen Varietät.* Hamburg: Dr. Kovač.

Ziesel, Günther. 2018. “Arnold Schwarzenegger Im Gespräch Mit Günther Ziesel [Video File].” Accessed May 21, 2020. https://www.youtube.com/watch?v=fW3YU3HFO5M.

Zue, Victor. 1976. *Acoustic Characteristics of Stop Consonants: A Controlled Study*. Bloomington, IN: Indiana University Linguistics Club.

7 Appendices

Appendix A

(I) English					
Stage	**Year**	**Video name**	**Link**	**Total duration**	**File designation**
Early	1979	Johnny Carson's Tonight Show (Part I-III)	https://www.youtube.com/watch?v=Y4OXujrjcRk (I) https://www.youtube.com/watch?v=5CUG06OOE0A (II) https://www.youtube.com/watch?v=A2AiLbpkngA (III)	05:57 (I); 04:32 (II); 06:41 (III)	AS_E_1979_1 AS_E_1979_2 AS_E_1979_3
	1979	Brian Linehan, City Lights	https://www.youtube.com/watch?v=1_Jl-k0x-lA	21:16	AS_E_1979_4
	1984	Late Night, David Letterman	https://www.youtube.com/watch?v=9H_2rs-1KZY	10:38	AS_E_1984_1
	1985	Late Night, David Letterman	https://www.youtube.com/watch?v=_ZCSiU37OOc	12:35	AS_E_1985_1
	1986	Terry Wogan, BBC1	https://www.youtube.com/watch?v=jGJqvW3mj5A	04:58	AS_E_1986_1
	1987	Good Morning America	https://www.youtube.com/watch?v=J-Qhvk4Gi5Y	04:29	AS_E_1987_1
	1988	***'Twins'-interview	https://www.youtube.com/watch?v=IYe60yfHYvw	06:14	AS_E_1988_1
	1988	Predator-interview	https://www.youtube.com/watch?v=BkX2CMCXhM8	12:17	AS_E_1988_2

Mid	1994	Interview on 'True Lies'	https://www.youtube.com/watch?v=fcfpP1fVkOw	05:23	AS_E_1994_2
	1994	Jimmy Carter, 'Junior' interview	https://www.youtube.com/watch?v=jau2rOxYNoo	02:20	AS_E_1994_1
	1999	Johnny Vaughan, Part I	https://www.youtube.com/watch?v=jhSBPC0qEfs	12:51	AS_E_1999_1
	1999	Johnny Vaughan, Part II	https://www.youtube.com/watch?v=41xHp4BIvB0	10:19	AS_E_1999_2
	1999	Johnny Vaughan, Part III	https://www.youtube.com/watch?v=JGodwIV4wq8	13:16	AS_E_1999_3
	1999	Johnny Vaughan, Part IV	https://www.youtube.com/watch?v=vBXLgjtM098	12:19	AS_E_1999_4
	2002	Exclusive interview (Raw Iron Documentary)	https://www.youtube.com/watch?v=wiXxifU5ilQ	74:00	AS_E_2002_1
	2003	**ITV F1 studio interview	https://www.youtube.com/watch?v=_kIHhXe_l1g	07:12	AS_E_2003_1
	2003	Richard & July Show	https://www.youtube.com/watch?v=bdqVizucLHQ	13:04	AS_E_2003_2
Late	2012	ABC News	https://www.youtube.com/watch?v=KIQfBhDhsPg	07:08	AS_E_2012_1
	2012	Google Play Presents	https://www.youtube.com/watch?v=FSPTkjtIHsI	29:01	AS_E_2012_6
	2012	ABC News	https://www.youtube.com/watch?v=KIQfBhDhsPg	03:28	AS_E_2012_3
	2012	NDTV Interview, India questions Arnold Schwarzenegger	https://www.youtube.com/watch?v=33KeqIgmxB0	23:34	AS_E_2012_7
	2012	Graham Norton Show Series 12 (I)	https://www.youtube.com/watch?v=WBWb2pCJ4mE	06:04	AS_E_2012_4
	2012	Graham Norton Show Series 12 (II)	https://www.youtube.com/watch?v=WBWb2pCJ4mE	03:08	AS_E_2012_5

	2013	Interview with Swedish TV	https://www.youtube.com/watch?v=gZO_eP402h4	30:33	AS_E_2013_3
	2013	Pressekonferenz The Last Stand	https://www.youtube.com/watch?v=NZkFPsR8gFs	04:56	AS_E_2013_1
	2013	10 Questions for A., Time	https://www.youtube.com/watch?v=ArV6f5kbmfE	04:39	AS_E_2013_2
	2013	JT Fox interview	https://www.youtube.com/watch?v=6WhVfdtiiEg	42:46	AS_E_2013_4
	2014	Absolute Radio	https://www.youtube.com/watch?v=yBpPtglHk94	08:35	AS_E_2014_1
	2014	Hindustan Times, Work-out Regimen	https://www.youtube.com/watch?v=QQGPWoUBzII	09:54	AS_E_2014_2
	2014	Kiss FM UK, Money Making Secrets	https://www.youtube.com/watch?v=Z7NqhRqLm8U	04:42	AS_E_2014_3
	2014	The Making of The Terminator	https://www.youtube.com/watch?v=KBF4Rxm_dlc	20:30	AS_E_2014_5
	2014	NDTV Interview, Want to do a Shankar movie	https://www.youtube.com/watch?v=itp0ppfE8o0	02:49	AS_E_2014_4
	2015	Arnold Schwarzenegger im OE3-Interview	https://www.youtube.com/watch?v=mKiYtzWztLM	02:40	AS_E_2015_1
	2015	James Corden, On 30 years of The Terminator	https://www.youtube.com/watch?v=odr2LDlixFI	03:54	AS_E_2015_3
	2015	Good Morning Britain	https://www.youtube.com/watch?v=gulu4QjRiMI	05:26	AS_E_2015_2
	2016	Graham Bensinger, Had to get out of Austria	https://www.youtube.com/watch?v=wFiNHn8sZ-A	05:32	AS_E_2016_1
	2016	Graham Bensinger, heart surgery	https://www.youtube.com/watch?v=iOOoyDXgHPs	04:47	AS_E_2016_2
	2016	Graham Bensinger, Karl Rove's bad advice	https://www.youtube.com/watch?v=NY8ijBhmUyE	07:18	AS_E_2016_3

	2016	NBC News, Politics	https://www.youtube.com/watch?v=CMbx9tW1b7I	07:50	AS_E_2016_4
	2016	Graham Bensinger, Father's negativity	https://www.youtube.com/watch?v=DJQOgp9a66I	04:57	AS_E_2016_5
	2016	Graham Bensinger, My army tank mistake	https://www.youtube.com/watch?v=Lyw6mJhb090	04:47	AS_E_2016_6
	2016	Graham Bensinger, Stealing Terminator	https://www.youtube.com/watch?v=rf6gO6jxcXA	06:46	AS_E_2016_7
	2016	Graham Bensinger, Best decision	https://www.youtube.com/watch?v=ZHL_UXfWT44	04:39	AS_E_2016_8
	2016	Graham Bensinger, Lucille Ball saved me	https://www.youtube.com/watch?v=n403JSJxNuE	05:51	AS_E_2016_9
	2016	Graham Bensinger, Real estate mogul	https://www.youtube.com/watch?v=15kLC-T-8Ew	04:20	AS_E_2016_10
	2016	Graham Bensinger, Partying with President Bush	https://www.youtube.com/watch?v=Jm87olgC_14	03:47	AS_E_2016_11
	2016	Graham Bensinger , Governor was role of life time	https://www.youtube.com/watch?v=NIUvNbSVTaE	04:04	AS_E_2016_12
	2016	Graham Bensinger, Maria Shriver didn't want me to run	https://www.youtube.com/watch?v=XFr722eNHMA	05:29	AS_E_2016_13
	2016	**Vision, goals & success	https://www.youtube.com/watch?v=K79pNxaliwk	09:59	AS_E_2016_14
	2017	Glenn Twiddle, Arnold Schwarzenegger Interview	https://www.youtube.com/watch?v=40xM0oZB_Po	24:22	AS_E_2017_6
	2017	USC, Democracy isn't a spectator sport	https://www.youtube.com/watch?v=nPDZfN4iXF8	02:12	AS_E_2017_1
	2017	CNN, Why Schwarzenegger didn't vote for Trump	https://www.youtube.com/watch?v=94tmN4UtLr0	01:31	AS_E_2017_2
	2017	**About Terminator 2	https://www.youtube.com/watch?v=T7r1lMT8Zpo	04:12	AS_E_2017_3

	2017	Arnold Schwarzenegger speaks with Guenther Ziesel	https://www.youtube.com/watch?v=ni-QbE1XrhAg	47:43	AS_E_2017_4
	2017	R20 Austrian World Summit 2017 in Vienna	https://www.youtube.com/watch?v=dH2B-5ZiFPc	19:22	AS_E_2017_5
	2018	Opie Radio, Arnold Schwarzenegger on all those crappy impressions on him	https://www.youtube.com/watch?v=j4LTcxjd9pA	01:36	AS_E_2018_1
	2018	CNN, Arnold Schwarzenegger: Politics 'sucks'	https://www.youtube.com/watch?v=OX5qbuAGOM4	07:32	AS_E_2018_2
(II) German					
Stage	**Year**	**Video name**	**Link**	**Total duration**	**File designation**
Early	1977	Arnold Schwarzenegger speaking German	https://www.youtube.com/watch?v=bkWCfuCmWRI	01:00	AS_G_1977_1
		Auf los geht's los, ARD 1982 Mensch Maier, ARD, 1986	https://www.youtube.com/watch?v=z_OaPkR-rVs	03:58	AS_G_1982_86_1
	1985	Na sowas! Arnold Schwarzenegger Interview	https://www.youtube.com/watch?v=4uFlF0ZoRIE	05:59	AS_G_1985_1
	1986	ZiB2 Interview mit Robert Hochner	https://www.youtube.com/watch?v=SDZIn9z_sak	07:42	AS_G_1986_1
	2012	**Kleine Zeitung, Interview "War echt abnormal"	https://www.youtube.com/watch?v=SfZID-D5O48	02:47	AS_G_2012_1
	2012	**Kleine Zeitung, Hautnah bei den Fans	https://www.youtube.com/watch?v=RQU3VUKeCro	02:07	AS_G_2012_2
	2013	Pressekonferenz, The Last Stand	https://www.youtube.com/watch?v=NZkFPsR8gFs	04:56	AS_G_2013_1

	2013	ORF1, Ski Opening Schladming	https://www.youtube.com/watch?v=NlQhLtmHU1A	03:38	AS_G_2013_2
	2013	Goldenes Ehrenzeichen des Landes STMK	https://www.youtube.com/watch?v=exfUZKdNhIM	01:35	AS_G_2013_3
	2017	Arnold Schwarzenegger speaks with Guenther Ziesel	https://www.youtube.com/watch?v=niQbE1XrhAg	03:10	AS_G_2017_4
	2017	R20 Austrian World Summit in Vienna	https://www.youtube.com/watch?v=dH2B-5ZiFPc	03:27	AS_G_2017_5
	2017	Auszeichnung durch Landeshauptmann	https://www.youtube.com/watch?v=bvDuXxrmZqo	03:26	AS_G_2017_6
	2017	OE24, Schwarzenegger trifft Van der Bellen	https://www.youtube.com/watch?v=8jpXjUrSeTY	04:34	AS_G_2017_7
	2018	**R20 Austrian World Summit	https://www.youtube.com/watch?v=tpqQjpCopBE	02:33	AS_G_2018_1

Table A: Sound files used for the analysis of AS's (I) English and (II) German pronunciation. *Note*: The videos marked with ** are no longer available on YouTube (date of last access: June 28th, 2020).

Appendix B

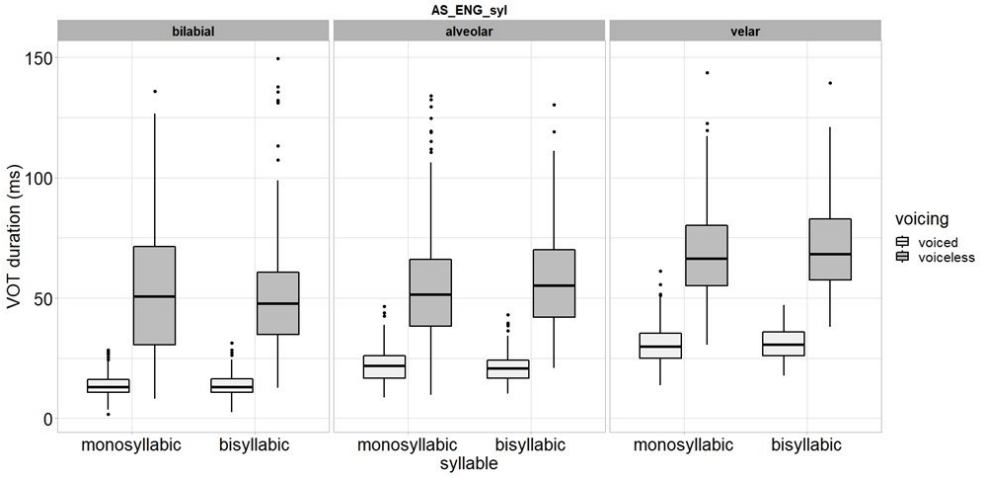


Figure B.1: VOT durations of AS's L2 English voiced and voiceless plosives according to place of articulation (bilabial, alveolar, velar) and syllable number in the test words (monosyllabic, bisyllabic).

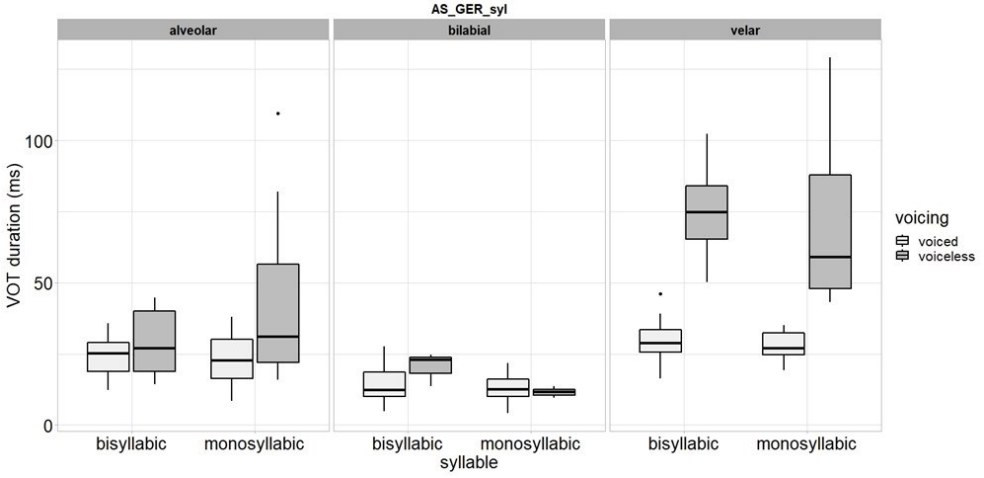


Figure B.2: VOT durations of AS's L1 German voiced and voiceless plosives according to place of articulation (bilabial, alveolar, velar) and syllable number in the test words (monosyllabic, bisyllabic).

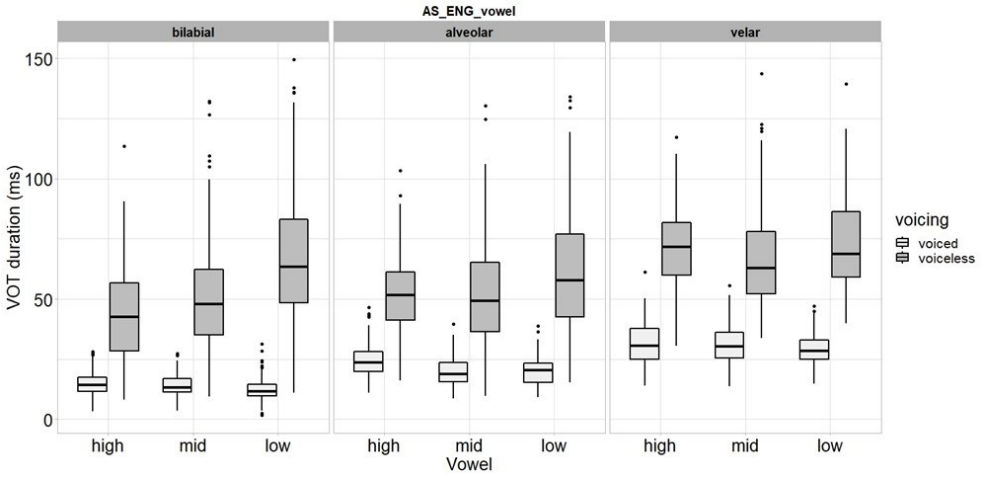


Figure B.3: VOT durations of AS's L2 English voiced and voiceless plosives according to place of articulation (bilabial, alveolar, velar) and quality of the vowel following the plosive (low, mid, high).

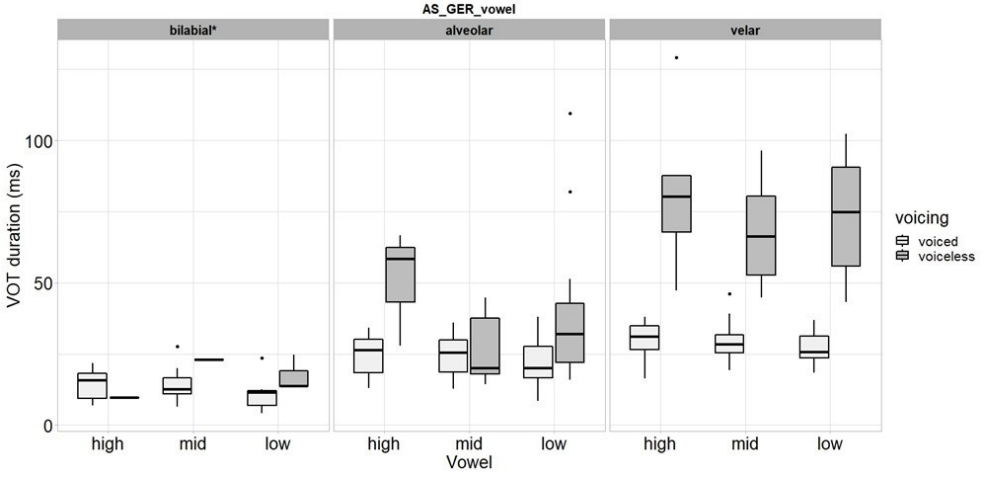


Figure B.4: VOT durations of AS's L1 German voiced and voiceless plosives according to place of articulation (bilabial, alveolar, velar) and quality of the vowel following the plosive (low, mid, high).

Appendix C

Target word	English translation	IPA transcription	Phoneme(s) tested
Baum	*tree*	/baʊm/	/b/
Sonne	*sun*	/ˈzɔnə/ - /sɔnə/	/ɔ/
Bank	*bank*	/baŋk/	/b/, /a/
Park	*park*	/paːk/	/p/, /a/
Katze	*cat*	/ˈkatsə/	/k/, /a/
Dach	*roof*	/dax/	/d/, /a/
Bus	*bus*	/bʊs/	/b/, /ʊ/
Buch	*book*	/buːx/	/b/, /u/
Tisch	*table*	/tɪʃ/	/t/, /ɪ/
Packerl	*package*	/ˈpakəl/	/p/, /a/
Wiese	*meadow*	/ˈviːzə/	/i/
Picknick	*picnic*	/ˈpɪkˌnɪk/	/p/, /ɪ/
Gans	*goose*	/gans/	/g/, /a/
Tonne	*barrel*	/ˈtɔnə/	/t/, /ɔ/
Teich	*pond*	/taɪç/	/t/
Mutter	*mother*	/mʊtɐ/	/ʊ/
Ball	*ball*	/bal/	/b/, /a/
Hund	*dog*	/hʊnt/	/ʊ/
Decke	*blanket*	/ˈdɛkə/	/d/, /ɛ/
Sieb	*sieve*	/ziːp/	/i/
Topf	*pot*	/tɔpf/	/t/, /ɔ/
Hund	*dog*	/hʊnt/	/ʊ/
Butter	*butter*	/ˈbʊtɐ/	/b/, /ʊ/
Käse	*cheese*	/ˈkɛːzə/	/k/, /ɛ/
Geld	*money*	/gɛlt/	/g/
Mus	*puree*	/muːs/	/u/
Kuh	*cow*	/kuː/	/k/, /u/
Tisch	*table*	/tɪʃ/	/t/, /ɪ/
Dieb	*thief*	/diːp/	/d/, /i/
Koffer	*suitcase*	/ˈkɔfɐ/	/k/, /ɔ/
Decke	*blanket*	/ˈdɛkə/	/d/, /ɛ/
Giebel	*gable*	/ˈgiːbl/	/g/, /i/
Gipfel	*peak*	/ˈgɪpfl/	/g/, /ɪ/
Wiese	*meadow*	/ˈviːzə/	/i/
Gans	*goose*	/gans/	/g/, /a/
Peter	*Peter*	/ˈpeːtə/	/p/
Bett	*bed*	/bɛt/	/b/, /ɛ/
Kissen	*pillow*	/ˈkɪsn/	/k/, /ɪ/
Topf	*pot*	/tɔpf/	/t/, /ɔ/
Puppe	*doll*	/ˈpʊpə/	/p/, /ʊ/
Dusche	*shower*	/ˈduːʃə/	/d/, /u/

Decke	*blanket*	/ˈdɛkə/	/d/, /ɛ/
Buch	*book*	/buːx/	/b/, /u/

Table C: Tokens collected for Austrian German reference data.

Appendix D

	ENG early			
	F1		F2	
Vowel	*Mean (SD)*	*Median*	*Mean (SD)*	*Median*
/i/	323.6 (60.24)	305.1	2028.1 (135.9)	2018.0
/ɪ/	420.2 (69.66)	410.3	1743.2 (118.3)	1752.6
/ɛ/	500.6 (77.41)	485.4	1665.2 (188.0)	1667.9
/u/	334.9 (45.38)	328.5	1003.6 (267.0)	950.8
/ʊ/	400.9 (68.98)	403.5	1172.9 (211.76)	1181.7
/ɔ/	661.5 (69.7)	651.1	1042.2 (209.91)	975.1
/ɑ/	697.2 (83.84)	707.0	1197.5 (109.31)	1182.6
/æ/	639.9 (99.28)	650.9	1559.4 (207.47)	1545.2
/ʌ/	593.1 (106.16)	601.7	1277.9 (104.57)	1265.8
/ɝ/	474.7 (43.34)	462.7	1302.5 (76.26)	1300.9
	ENG mid			
	F1		F2	
Vowel	*Mean (SD)*	*Median*	*Mean (SD)*	*Median*
/i/	326.3 (43.04)	320.0	2035.5 (252.94)	2041.8
/ɪ/	436.4 (77.18)	429.1	1771.1 (100.58)	1779.0
/ɛ/	502.7 (51.58)	503.3	1595.3 (142.56)	1599.6
/u/	338.5 (43.05)	340.3	867.8 (246.98)	778.9
/ʊ/	418.0 (50.3)	412.6	1231.2 (209.24)	1236.7
/ɔ/	615.4 (71.73)	594.9	915.4 (78.48)	936.1
/ɑ/	690.0 (55.31)	696.6	1190.3 (129.17)	1191.7
/æ/	630.6 (91.12)	659.9	1589.1 (194.24)	1534.0
/ʌ/	607.8 (52.27)	610.3	1218.2 (82.09)	1228.8
/ɝ/	477.1 (42.31)	476.3	1317.4 (74.51)	1323.9
	ENG late			
	F1		F2	
Vowel	*Mean (SD)*	*Median*	*Mean (SD)*	*Median*
/i/	306.0 (41.32)	301.9	2054.8 (117.79)	2040.1
/ɪ/	400.3 (60.58)	390.5	1727.2 (125.9)	1737.4
/ɛ/	489.8 (52.2)	485.9	1607.9 (115.03)	1598.78
/u/	330.0 (75.71)	317.9	940.6 (438.66)	799.6
/ʊ/	384.6 (39.41)	381.5	1217.6 (257.99)	1272.8
/ɔ/	614.1 (78.2)	602.3	876.5 (73.74)	882.4
/ɑ/	677.8 (73.99)	681.2	1208.2 (140.06)	1193.6
/æ/	624.5 (99.86)	648.2	1580.6 (169.3)	1535.2
/ʌ/	611.1 (57.01)	609.0	1231.2 (88.65)	1229.4
/ɝ/	454.0 (36.82)	448.1	1298.6 (85.71)	1292.07
	GER early			

	F1		F2	
Vowel	*Mean (SD)*	*Median*	*Mean (SD)*	*Median*
/i/	360.8 (34.69)	352.7	303.8 (16.14)	302.1
/ɪ/	388.7 (59.77)	404.0	381.4 (35.03)	387.0
/ɛ/	476.8 (66.88)	437.4	460.7 (30.86)	469.3
/u/	391.3 (27.85)	381.4	308.7 (27.44)	311.9
/ʊ/	484.5 (181.51)	461.2	465.2 (91.02)	483.1
/ɔ/	500.6 (55.08)	490.4	526.1 (52.32)	534.8
/ɑ/	701.6 (76.09)	679.7	617.1 (76.84)	598.5
/e/	516.6 (14.59)	518.4	384.3 (18.81)	390.9

Table D: Means, standard deviations (SD), and medians of F1 and F2 (in Hz) of AS's English (ENG) and German (GER) monophthongs according to stage (early, mid, late).

Appendix E

File name	Orthographic transcript	Duration (in sec.)
AS_1977_1_1	zwanzig Wiederholungen mit ungefähr dreihundert Pfund und das ungefähr a halbe Stund täglich	5.85
AS_1977_1_3	i[ch] hab ungefähr zweihundert Kilo hinten oben auf der Maschin[e] und i[ch] moch[mach] ungefähr zehn bis zwölf Wiederholungen	6.86
AS_1986_1_1	und i[ch] wollte irgendwie an an Mittel finden dass i[ch] das erreich früher oder später	6.74
AS_1986_1_2	und i[ch] hab dann natürlich ganz wie a Wahnsinniger angefangen mit dem Training vier fünf Stunden täglich	5.33
AS_2017_4_2	i hab heute schon i i geh jeden Tag in der Früh um sechs Uhr [] zum Radfahren ich fahr ungefähr a Stunde mit dem Rad	5.84
AS_2017_4_3	wenn ich an unsere Radtour denke ja dann denk i[ich] net[nicht] so viel an deine Leistung	7.75
AS_2017_5_1	die erste Klasse besucht haben wie wir olle[alle] Angst gehabt haben vom der Schul[e] und wie wir noch Schlitten fahrn gangen[gegangen] san[sein]	6.28
AS_2017_5_2	Eisschießen, Schifahrn und so weiter Fußball gespielt haben wir zusammen also mir ham so viel Gaude gehabt	5.8
Stro_2018_1_1_1	ich hab sehr viele Erfahrungen gesammelt über die Jahre nicht ich war in Aufsichtsräten von Spitälern	6.42
Stro_2018_2_1_1	die richtigen Eigenschaften hat dann kann sehr viel passieren	3.68
Stro_2018_2_2_1	aber ich sag auch immer noch ganz gleich wie g[e]scheit du bist es gehört auch a[ein] bissl ein Glück dazu ja	6.51
Sro_2018_2_4_1	man soll irgendetwas tun was man gerne macht	3.21
Kainr_2018_1_1	ich darf leider ka[keine] Torte ess[e]n weil sonst verlier ich meine acht Prozent Körperfett	5.41
Kainr_2016_1_1	unter andern auch seinen neunundsechzigsten Geburtstag	4.26
Kainr_2017_1	und bin deswegen brav im Studio und tu brav trainieren	4.49
Kainr_2017_2	jetzt mach ma an old school Training wie vor fuffzig[fünfzig] Joa[Jahr] und zwar unsere Lieblingsübung warn immer Klimmzüge	5.9
Urdl_2018_1_10	sei[n] bester Freund der is der is eigentlich aufgewachsen im Gitterbett z[us]ammen	5.47
Urdl_2017_1	du so schnell wie die Zeit vergeht jetzt wer[d] ma siebzig Joa[Jahr]	5.02

Urdl_2017_2	wir erinnern uns noch olle[alle] an diese Zeit wo ma[wir] eigentlich in die Volksschul[e] gangen[gegangen] san[sind]	5.7
Marnul_2017_1_5	und dort hab ich begonnen neunzehnhundertzweiundfünfzig mit Gewichtheben	4.03
Marnul_2017_1_9	aber zu diesem Zeitpunkt hots[hat's] bei uns in Österreich ja kein Eiweiß noch gegeben	5.72
Marnul_2017_2	und dann musst ungefähr a[ein] Gewicht verwenden acht bis neun Wiederholungen und bei der neunten zehnten hast du nicht mehr die Kraft des richtig auszuführen	8.17
Schick_2013_1_1	hätt net sein müssen nicht i[ch] mein andererseits ist es wirklich a[ein] jeder sagt des[das] is[t] seine Privatgeschichte	5.66
Gerstl_2011_2_1	studieren weiterarbeiten und nach oben geh[e]n	5.17
Gerstl_2011_2_2	es is[t] wunderbar deshalb weil es ist die Motivation für die Jugend	6.34
FAR_SpeakerA_1	Naja i hätt einmal gsogt[gesagt] im Durchschnitt so mit der richtigen Frequenz fürs Herz acht Minuten	6.57
FAR_SpeakerA_2	das siecht[sieht] man so oft und i[ch] hob[hab] erst in sog i amol fünf Joa[Jahr] zwa[zwei] gesehen die an Fisch rausgeholt haben	6.42
FAR_SpeakerA_3	es is[t] halt wie bei an Tier der an[einen] gewissen Jagdinstinkt hat und so is[t] halt beim Menschen auch	5.28
FAR_SpeakerB_1	mir gefallt's Fischerei Fischen tu ich schon lang und aufhörn mog[mag] i[ch] no[ch] net[nicht]	6.74

Table E: Orthographic transcripts of speech samples used for perceived nativeness, intelligibility and comprehensibility ratings.

Appendix F

Listener	Age	Sex	Highest education	Age at onset of L2 learning	Years of schooling in the L2	Level of L2 proficiency
MT_ML	34	female	Apprenticeship diploma	10	9	B1
PH_ML	34	male	Matura	9	10	A2-B1
TK_ML	33	female	University degree (Bachelor)	10	9	B1-B2
BE_ML	38	female	Matura	9	9	B1
SR_ML	27	female	Matura	10	8	B1
ST_ML	32	male	Matura	10	9	A2
HP_ML	26	male	University degree (FH)	10	8	A2-B1
AV_ML	31	female	Matura	10	7	B1
DL_ML	32	male	Apprenticeship diploma	10	9	B2
MR_ML	34	male	University degree	9	9	B1-B2
UN_ML	25	female	University degree (FH)	10	8	B1
RP_ML	32	male	University degree (Bachelor)	10	8	B1
EB_ML	25	male	Matura	9	9	A2-B1
BS_ML	29	female	Matura	10	8	A2
VK_ML	39	female	University degree (FH)	10	7	B1
JS_ML	35	female	Apprenticeship diploma	10	6	A2
LD_ML	31	female	University degree (PH)	9	9	A2-B1
EH_ML	33	male	University degree	10	9	B2
ON_ML	38	male	Apprenticeship diploma	10	8	A1
FL_ML	28	male	Matura	9	6	A2

Table F.1: Listener details for group ML–GRA (n = 20).

Listener	Age	Sex	Age at onset of L2 learning	Years of schooling in the L2	Years of L2 experience at university level	Level of L2 proficiency	Months spent in an L2 country
BD_BIL	25	male	10	11	2	C1-C2	12
HT_BIL	21	female	9	9	3	C1	-
TS_BIL	19	female	10	8	2	C1	-
WS_BIL	20	female	9	9	2	C1	2
NL_BIL	19	female	10	8	2	C1	-
BN_BIL	20	female	11	8	2	B2-C1	-
WK_BIL	24	female	10	8	3	C2	12
LM_BIL	20	female	9	10	2	C2	12.5
LC_BIL	23	female	10	9	3	C2	8
BK_BIL	22	male	9	9	2.5	C1-C2	-
NK_BIL	20	male	10	9	2.5	C2	6
LB_BIL	23	male	11	8	3	C1-C2	-
CH_BIL	21	male	11	8	2	C2	8
SV_BIL	24	female	9	9	3.5	C2	12
VK_BIL	25	male	10	9	2.5	C1-C2	
LT_BIL	20	male	10	8	2	C1	-
ML_BIL	21	male	9	9	2.5	C1	6
PM_BIL	22	male	9	9	3	B2-C1	-
JT_BIL	20	female	10	8	2	C1-C2	3
SM_BIL	24	female	10	8	3	C2	-

Table F.2: Listener details for group BIL–GRA (n = 20).

Appendix G

Frage 1	Alter in Jahren:	
Frage 2	Geburtsjahr:	
Frage 3	Geschlecht:	
Frage 4	Geburtsort, Geburtsland:	
Frage 5	Nationalität:	
Frage 6	Derzeitiger Wohnort:	

Frage 7	Haben Sie in der Vergangenheit für eine längere Zeit (länger als 4 Wochen) an einem anderen Ort (Stadt, Land) als dem oben genannten gelebt?	□ Ja Wo? _____ Wie lange? _____ □ Nein
Frage 8	Höchste abgeschlossene Ausbildung?	□ Pflichtschule □ Höherbildende Schule mit Matura □ Lehrabschluss □ Hochschulabschluss (Universität, FH)
Frage 9	Falls Hochschul**abschluss**: Welche Studienrichtung?	
Frage 10	Was ist Ihr **Beruf** bzw. Ihre derzeitige **Tätigkeit** (darunter fällt z.B. Karenz, Kinderbetreuung, Hausfrau, StudentIn, in Ausbildung, etc.)?	
Frage 11	Falls StudentIn, welche **Studienrichtung**? In welchem Semester?	Studienrichtung (falls Lehramt, bitte Fächer angeben): _____ Semester: _____
Frage 12	Haben Sie eine diagnostizierte **Sprach- oder Hörbeeinträchtigung**?	□ Ja □ Nein
Frage 13	Was ist Ihre **Muttersprache** (Erstsprache)?	

Frage 14	Sind Sie **zweisprachig** aufgewachsen?	□ Ja □ Nein		
Frage 15	Sprechen Sie noch **andere Sprachen** außer Ihrer Muttersprache(n)?	□ Ja □ Nein	Welche? _____	
Frage 16	Wenn Sie noch **andere Sprachen** außer Ihrer Muttersprache sprechen, wie schätzen Sie Ihr **Sprachniveau** (sprachliche Kompetenz) in diesen Sprachen ein? **** Bitte geben Sie hier eines der sechs Sprachniveaus (A1, A2, B1, B2, C1, C2) an, die am ENDE dieses Fragebogens aufgelistet und erläutert sind – bitte nach unten scrollen) ****	Sprache: Sprache: Sprache:	Sprachniveau: Sprachniveau: Sprachniveau:	Kommentar (optional): Kommentar (optional): Kommentar (optional):
Frage 17	Sind Ihre **Eltern** oder ist ein **Elternteil** zweisprachig aufgewachsen?	□ Ja □ Nein		
Frage 18	Würden Sie sagen, dass Sie **Dialektsprecher** oder eher **Standardsprecher** sind?	□ Dialektsprecher □ Standardsprecher		
Frage 19	Welchen **Dialekt** sprechen Sie (steirisch oder eher einen anderen österreichischen Dialekt)?	□ Steirisch □ Anderer Dialekt		
Frage 20	Welchen Dialekt sprechen Ihre **Eltern** (steirisch oder eher einen anderen österreichischen Dialekt)?	□ Steirisch □ Anderer Dialekt		
Frage 21	Würden Sie sagen, dass Sie den steirischen Dialekt gut von anderen österreichischen Dialekten unterscheiden können?	□ Ja □ Nein □ Weiß nicht		
Frage 22	Wie häufig haben Sie Kontakt zu Sprechern, die einen **anderen österreichischen Dialekt** sprechen?	□ Sehr häufig □ Eher häufig □ Eher selten		

		□ Sehr selten
Frage 23	Haben Sie **Englisch** in der Schule und/oder im Studium gelernt?	□ Ja □ Nein
	Wenn Sie Englisch in der Schule gelernt haben, wie **lange** (in Jahren)?	Jahre: _____
	Mit welchem **Alter** haben Sie begonnen, Englisch in der Schule zu lernen?	Alter: _____
Frage 24	**Sprechen** Sie im **beruflichen** Umfeld häufig Englisch?	□ Ja □ Nein
	Wenn ja, wie häufig?	□ Sehr selten (weniger als 1x pro Monat) □ Eher selten (1-2x pro Monat) □ Eher häufig (mehrmals pro Monat) □ Sehr häufig (mehrmals pro Woche)
Frage 25	**Sprechen** Sie im **privaten** Umfeld häufig Englisch und wenn ja, mit wem (Freunde, Familie, StudienkollegInnen)?	□ Ja Mit wem? _____ □ Nein
	Wenn ja, wie häufig?	□ Sehr selten (weniger als 1x pro Monat) □ Eher selten (1-2x pro Monat) □ Eher häufig (mehrmals pro Monat) □ Sehr häufig (mehrmals pro Woche)
Frage 26	**Schauen** Sie TV-Serien und/oder Filme in Englisch?	□ Ja □ Nein
	Wenn ja, wie häufig?	□ Sehr selten (weniger als 1x pro Monat) □ Eher selten (1-2x pro Monat) □ Eher häufig (mehrmals pro Monat) □ Sehr häufig (mehrmals pro Woche)
Frage 27	**Lesen** Sie Bücher und/oder Zeitungen/Magazine (online oder Print) in Englisch?	□ Ja □ Nein

	Wenn ja, wie häufig?	□ Sehr selten (weniger als 1x pro Monat) □ Eher selten (1-2x pro Monat) □ Eher häufig (mehrmals pro Monat) □ Sehr häufig (mehrmals pro Woche)
Frage 28	**Hören** Sie Radio, online Broadcasts oder Hörbücher in Englisch?	□ Ja □ Nein
	Wenn ja, wie häufig?	□ Sehr selten (weniger als 1x pro Monat) □ Eher selten (1-2x pro Monat) □ Eher häufig (mehrmals pro Monat) □ Sehr häufig (mehrmals pro Woche)
Frage 29	**Schreiben** Sie häufig Texte (E-Mails, Briefe, Kurzgeschichten, Gedichte, online Kommentare/Postings, etc.) in Englisch (unabhängig von den Schreibaufgaben, die Sie im Rahmen Ihres Studiums bekommen)?	□ Ja Welche Art von Texten? _____ □ Nein
	Wenn ja, wie häufig?	□ Sehr selten (weniger als 1x pro Monat) □ Eher selten (1-2x pro Monat) □ Eher häufig (mehrmals pro Monat) □ Sehr häufig (mehrmals pro Woche)
Frage 30	Basierend auf Ihren vorherigen Antworten, würden Sie sagen, dass für Sie **Englisch** neben Ihrer Muttersprache/Erstsprache eine Sprache ist, die Sie auch außerhalb der Universität **häufig** verwenden und in der Sie sich **gut** und **sicher** in verschiedenen kommunikativen Kontexte ausdrücken können?	□ Ja Kommentar (optional): _____ □ Nein Kommentar (optional): _____

****

Erläuterungen der Sprachniveaus (siehe Frage 16)[104]: Bitte lesen Sie sich die folgenden Erläuterungen durch und beurteilen Sie Ihre Sprachkompetenz dementsprechend.

Elementare Sprachverwendung (A1, A2)

Spracheniveau A1:
Ich kann vertraute, alltägliche Ausdrücke und ganz einfache Sätze verstehen und verwenden, die auf die Befriedigung konkreter Bedürfnisse zielen. Ich kann mich und andere vorstellen und anderen Leuten Fragen zu ihrer Person stellen - z. B. wo sie wohnen, was für Leute sie kennen oder was für Dinge sie haben - und kann auf Fragen dieser Art Antwort geben. Ich kann mich auf einfache Art verständigen, wenn die Gesprächspartnerinnen oder Gesprächspartner langsam und deutlich sprechen und bereit sind zu helfen

Sprachniveau A2:
Ich kann Sätze und häufig gebrauchte Ausdrücke verstehen, die mit Bereichen von ganz unmittelbarer Bedeutung zusammenhängen (z.B. Informationen zur Person und zur Familie, Einkaufen, Arbeit, nähere Umgebung). Ich kann mich in einfachen, routinemäßigen Situationen verständigen, in denen es um einen einfachen und direkten Austausch von Informationen über vertraute und geläufige Dinge geht. Ich kann mit einfachen Mitteln die eigene Herkunft und Ausbildung, die direkte Umgebung und Dinge im Zusammenhang mit unmittelbaren Bedürfnissen beschreiben.

Selbstständige Sprachverwendung (B1, B2)

Sprachniveau B1:
Ich kann die Hauptpunkte verstehen, wenn klare Standardsprache verwendet wird und wenn es um vertraute Dinge aus Arbeit, Schule, Freizeit usw. geht. Ich kann die meisten Situationen bewältigen, denen man auf Reisen im Sprachgebiet begegnet. Ich kann mich einfach und zusammenhängend über vertraute Themen und persönliche Interessengebiete äußern. Ich kann über Erfahrungen und Ereignisse berichten, Träume, Hoffnungen und Ziele beschreiben und zu Plänen und Ansichten kurze Begründungen oder Erklärungen geben.

Sprachniveau B2:

[104] The definition of the six competence levels provided here are based on the Common European Framework of Reference for Languages (CEFR; Council of Europe 2001).

Ich kann die Hauptinhalte komplexer Texte zu konkreten und abstrakten Themen verstehen; ich verstehe im eigenen Spezialgebiet auch Fachdiskussionen. Ich kann mich so spontan und fließend verständigen, dass ein normales Gespräch mit Muttersprachlern ohne größere Anstrengung auf beiden Seiten gut möglich ist. Ich kann mich zu einem breiten Themenspektrum klar und detailliert ausdrücken, einen Standpunkt zu einer aktuellen Frage erläutern und die Vor- und Nachteile verschiedener Möglichkeiten angeben.

Kompetente Sprachverwendung (C1, C2)

Sprachniveau C1:
Ich kann ein breites Spektrum anspruchsvoller, längerer Texte verstehen und auch implizite Bedeutungen erfassen. Ich kann mich spontan und fließend ausdrücken, ohne öfter deutlich erkennbar nach Worten suchen zu müssen. Ich kann die Sprache im gesellschaftlichen und beruflichen Leben oder in Ausbildung und Studium wirksam und flexibel gebrauchen. Ich kann mich klar, strukturiert und ausführlich zu komplexen Sachverhalten äußern und dabei verschiedene Mittel zur Textverknüpfung angemessen verwenden.

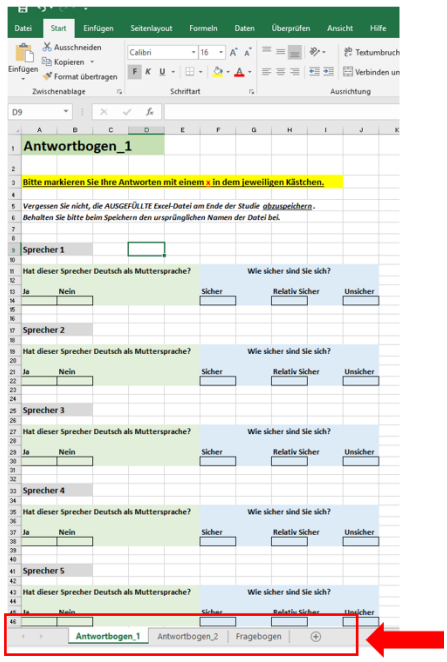
Sprachniveau C2 (nahezu muttersprachliche Sprachverwendung):
Ich kann praktisch alles, was ich lese oder höre, mühelos verstehen. Ich kann Informationen aus verschiedenen schriftlichen und mündlichen Quellen zusammenfassen und dabei Begründungen und Erklärungen in einer zusammenhängenden Darstellung wiedergeben. Ich kann mich spontan, sehr flüssig und genau ausdrücken und auch bei komplexeren Sachverhalten feinere Bedeutungsnuancen deutlich machen.

Table G: Listener questionnaire used in Study III.

Appendix H

Description of rating procedure (Study III): E-mail sent to listeners

Liebe Teilnehmerin, lieber Teilnehmer!
Vielen Dank, dass Sie sich Zeit für meine Studie nehmen, die ich im Rahmen meiner Dissertation an der Karl-Franzens-Universität am Institut für Anglistik durchführe. Die Studie besteht insgesamt aus **drei Teilen**. Ihre Aufgabe wird es sein, sich unterschiedliche **Tonaufnahmen** anzuhören und anhand verschiedener Fragestellungen zu bewerten. Um die Aufnahmen gut hören zu können, benutzen Sie bitte **unbedingt Kopfhörer** und befinden sich in einer **ruhigen Umgebung** (keine störenden Hintergrundgeräusche, wie beispielsweise Radio oder TV).
Bitte bearbeiten Sie die Studie am **PC** oder **Laptop** (kein Tablet oder Handy!).
Hier eine kurze Anleitung zur Vorgehensweise (alle weiteren Informationen finden Sie im Anhang dieser Mail):
Öffnen Sie die **zwei Dateien im Anhang dieser E-Mail** (eine Excel-Datei ‚Antwortbögen' und eine Power Point Präsentation ‚INPUT_online'). Bitte öffnen Sie beide Dateien parallel auf Ihrem Computerbildschirm – das erleichtert das Arbeiten.



Die Excel-Datei beinhaltet drei verschiedene Tabellenblätter: **Antwortbogen_1**, **Antwortbogen_2** und **Antwortbogen_3**.
Um zu dem jeweiligen Antwortbogen zu gelangen, klicken Sie auf die entsprechende **Registerkarte**, die Sie am linken unteren Rand (siehe Screenshot) finden.
Die Anleitungen zu den unterschiedlichen Teilen der Studie, die Sie in der Power Point Präsentation finden, sagen Ihnen immer ganz genau, **in welchem Tabellenblatt Sie Ihre Antworten vermerken sollen**.

Um zu starten, klicken Sie auf die **erste Folie** der Power Point Präsentation und folgen Sie der Anleitung.
Bitte beachten Sie: Die erhobenen Daten werden streng vertraulich behandelt, nicht an Dritte weitergegeben und ausschließlich für Zwecke dieser

wissenschaftlichen Studie verwendet. In Übereinstimmung mit den Richtlinien zum Datenschutz, bitte ich Sie zudem, sich das angefügte **Informationsblatt** inkl. **Datenschutzerklärung** genau durchzulesen und mir unterschrieben zukommen zu lassen.

Bitte vergessen Sie nicht, die ausgefüllten Fragebögen (d.h. die komplette Excel-Datei) nach Bearbeitung der drei Teile an ich zu retournieren.

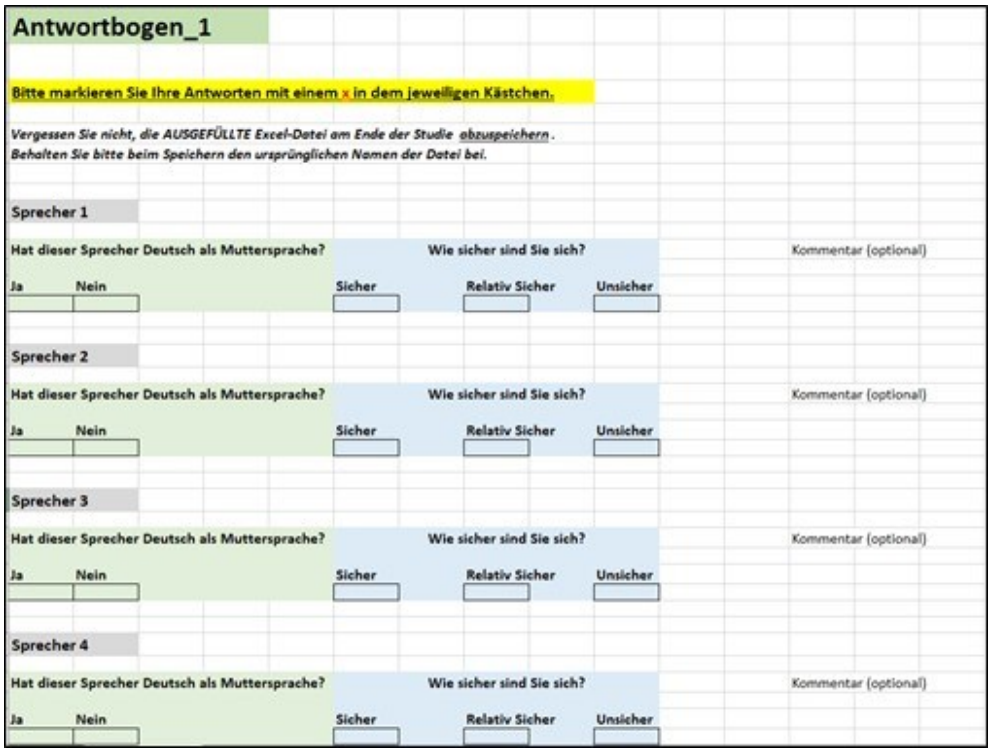
Herzlichen Dank, dass Sie mich bei meiner Studie unterstützen!

Appendix I

Sample of an Excel-rating sheet used in the nativeness rating task

Task description:

Sie werden nun nacheinander eine Reihe unterschiedlicher kurzer Tonaufnahmen hören. Es handelt sich hierbei um Aufnahmen von **Sprechern, die kurze Sätze in österreichischem Deutsch sagen.**
Hören Sie sich die jeweilige Tonaufnahme zunächst an. Vermerken Sie dann auf **Antwortbogen_1**, ob Sie glauben, dass der **Sprecher Deutsch als Muttersprache hat** (**JA** oder **NEIN**). Vermerken Sie Ihre Antwort mit einem **x** in dem jeweiligen Kästchen.
Vermerken Sie dann, **wie sicher Sie sich sind**, dass der Sprecher Deutsch als Muttersprache hat bzw. dass er Deutsch nicht als Muttersprache hat (**Sicher – Relativ Sicher – Unsicher**). Vermerken Sie auch hier Ihre Antwort mit einem **x** in dem jeweiligen Kästchen.
Sie können zudem aufschreiben, wenn Ihnen bei dem jeweiligen Sprecher etwas Besonderes auffällt (z.B. wie er etwas ausspricht, die Verwendung eines bestimmten Wortes, Betonungen, etc.).
Ihre Antworten sollten Sie nach **EINMALIGEM Anhören** der Tonaufnahme vermerken. Sie dürfen nach jeder Aufnahme eine kurze Pause einlegen, um Ihre Antworten auf dem Antwortbogen zu vermerken. Beachten Sie bitte, dass es kein Richtig oder Falsch gibt – verlassen Sie sich auf Ihre Ohren und Ihr Bauchgefühl.



Antwortbogen_1

Bitte markieren Sie Ihre Antworten mit einem x in dem jeweiligen Kästchen.

Vergessen Sie nicht, die AUSGEFÜLLTE Excel-Datei am Ende der Studie abzuspeichern.
Behalten Sie bitte beim Speichern den ursprünglichen Namen der Datei bei.

Sprecher 1

Hat dieser Sprecher Deutsch als Muttersprache?		Wie sicher sind Sie sich?			Kommentar (optional)
Ja	Nein	Sicher	Relativ Sicher	Unsicher	

Sprecher 2

Hat dieser Sprecher Deutsch als Muttersprache?		Wie sicher sind Sie sich?			Kommentar (optional)
Ja	Nein	Sicher	Relativ Sicher	Unsicher	

Sprecher 3

Hat dieser Sprecher Deutsch als Muttersprache?		Wie sicher sind Sie sich?			Kommentar (optional)
Ja	Nein	Sicher	Relativ Sicher	Unsicher	

Sprecher 4

Hat dieser Sprecher Deutsch als Muttersprache?		Wie sicher sind Sie sich?			Kommentar (optional)
Ja	Nein	Sicher	Relativ Sicher	Unsicher	

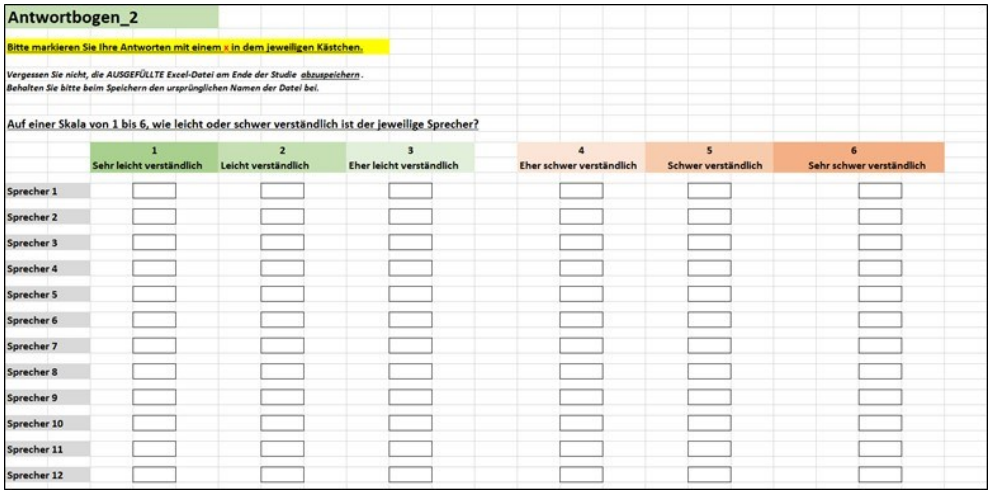
Figure I: Screenshot ‘Antwortbogen_1’.

Appendix J

Sample of an Excel-rating sheet used in the comprehensibility rating task

Task description:

Wie im vorherigen Teil werden Sie nun eine Reihe unterschiedlicher kurzer Tonaufnahmen hören.
Hören Sie sich die jeweilige Tonaufnahme zunächst an. Vermerken Sie dann auf Antwortbogen_2, wie leicht oder schwer verständlich der jeweilige Sprecher ist. Hierfür benutzen Sie eine Skala von 1 (= Sehr leicht verständlich) bis 6 (= Sehr schwer verständlich). Vermerken Sie Ihre Antwort mit einem x in dem jeweiligen Kästchen.
Ihre Antworten sollten Sie nach EINMALIGEM Anhören der Tonaufnahme vermerken. Sie dürfen nach jeder Aufnahme eine kurze Pause einlegen, um Ihre Antworten auf dem Antwortbogen zu vermerken.
Beachten Sie bitte, dass es auch hier kein Richtig oder Falsch gibt – verlassen Sie sich auf Ihre Ohren und Ihr Bauchgefühl.



Antwortbogen_2

Bitte markieren Sie Ihre Antworten mit einem x in dem jeweiligen Kästchen.

Vergessen Sie nicht, die AUSGEFÜLLTE Excel-Datei am Ende der Studie abzuspeichern.
Behalten Sie bitte beim Speichern den ursprünglichen Namen der Datei bei.

Auf einer Skala von 1 bis 6, wie leicht oder schwer verständlich ist der jeweilige Sprecher?

	1 Sehr leicht verständlich	2 Leicht verständlich	3 Eher leicht verständlich	4 Eher schwer verständlich	5 Schwer verständlich	6 Sehr schwer verständlich
Sprecher 1	☐	☐	☐	☐	☐	☐
Sprecher 2	☐	☐	☐	☐	☐	☐
Sprecher 3	☐	☐	☐	☐	☐	☐
Sprecher 4	☐	☐	☐	☐	☐	☐
Sprecher 5	☐	☐	☐	☐	☐	☐
Sprecher 6	☐	☐	☐	☐	☐	☐
Sprecher 7	☐	☐	☐	☐	☐	☐
Sprecher 8	☐	☐	☐	☐	☐	☐
Sprecher 9	☐	☐	☐	☐	☐	☐
Sprecher 10	☐	☐	☐	☐	☐	☐
Sprecher 11	☐	☐	☐	☐	☐	☐
Sprecher 12	☐	☐	☐	☐	☐	☐

Figure J: Screenshot ‘Antwortbogen_2’.

Appendix K

Sample of an Excel-rating sheet used in the intelligibility rating task

Task description:

Wie in den vorherigen Teilen werden Sie nun eine Reihe unterschiedlicher kurzer Tonaufnahmen hören.
Hören Sie sich die jeweilige Tonaufnahme zunächst an. Schreiben Sie dann **Wort für Wort auf, was der Sprecher sagt**, d.h. fertigen Sie ein Transkript an. Hierbei ist es **nicht** wichtig, auf eine korrekte Interpunktion (Beistrichsetzung etc.) zu achten.
Verwenden Sie **Antwortbogen_3**. Bitte schreiben Sie Ihr Transkript direkt hinter die jeweilige Sprecherbezeichnung in der linken Spalte.
Wenn möglich, sollten Sie Ihre Antwort nach **einmaligen Anhören** vermerken. Sie dürfen die Tonaufnahme jedoch ein zweites Mal anhören, wenn Sie nach dem ersten Anhören noch nicht alles notiert haben.
Sie können nach jeder Aufnahme eine kurze **Pause** einlegen, um Ihre Antworten auf dem Antwortbogen zu vermerken.

Antwortbogen_3

Orthographische Transkription

Schreiben Sie Wort für Wort auf, was der jeweilige Sprecher sagt.

Sprecher 1

Sprecher 2

Sprecher 3

Sprecher 4

Sprecher 5

Sprecher 6

Sprecher 7

Sprecher 8

Sprecher 9

Sprecher 10

Sprecher 11

Figure K: Screenshot 'Antwortbogen_3'.

Arnold Schwarzenegger is perhaps the most famous Austrian-American celebrity who is not only well-known for his stunning career but who has also gained considerable attention for his supposedly unique accent in both his first (L1) and second (L2) language. This book presents an investigation of Schwarzenegger's L1 German and L2 English pronunciation over a period of 40 years, using a speech corpus of broadcast interviews which illustrate his pronunciation at different stages in time. Study I and II present acoustic investigations of his segmental speech production (plosives and vowels). Study III examined if Schwarzenegger is perceived to sound native in his German pronunciation when being compared to monolingual German speakers and if listeners' linguistic background influences their judgments of L1 nativeness. The findings of the acoustic analyses and the perception study show that a mature L1 system may be altered in response to a developing L2 system and that speakers might 'lose' their status as native speakers in their first language, providing evidence for the inherently dynamic nature of speech development.