The Effects of Age of Onset on VOT in L2 Acquisition and L1 Attrition

A Study of the Speech Production and Perception of Advanced Spanish-Swedish Bilinguals

Katrin Stölten

Centre for Research on Bilingualism
Department of Swedish Language and Multilingualism
Stockholm University
Abstract

This thesis explores the role of age in second language (L2) acquisition and first language (L1) attrition. The focus is on Voice Onset Time (VOT) in the production and categorical perception of word-initial L1 and L2 stops in highly advanced L1 Spanish learners of L2 Swedish. Using as the point of departure a maturational constraints perspective and the Critical Period Hypothesis (CPH), Study I examines the impact of age of onset (AO) of L2 acquisition on the production of L2 Swedish voiceless stops. The results show that there are AO effects even in the speech of highly advanced L2 learners and that the incidence of nativelike L2 learners is considerably lower than earlier assumed. However, conclusions like these are only possible when speaking rate is accounted for, thereby highlighting the importance of speaking rate effects on VOT as a measure of nativelikness. Like Study I, Study II reveals age effects on the same L2 learners’ categorical perceptions of L2 Swedish stops. Moreover, after combining the results with the data from Study I, the incidence of nativelike behavior drops remarkably with no late L2 learner performing within the range of native-speaker production and perception. The results suggest that L2 acquisition of phonetic/phonological aspects is especially sensitive to AO effects. It is concluded that theories on maturational constraints, including the CPH, cannot be refuted on the basis of the present data. Study III concerns the same participants’ production and perception of L1 Spanish stops. Age of reduced contact (ARC) is identified as an important predictor for L1 attrition and retention of voiceless stop production, although not of stop perception. This discrepancy is related to different activation thresholds as proposed by the Activation Threshold Hypothesis (ATH). It is further suggested that early bilinguals are more dependent on high-frequency L1 use than late bilinguals when compensating for age effects, but only in production.

Keywords: second language acquisition, first language attrition, maturational constraints, critical period, voice onset time, stop production, categorical perception, activation threshold hypothesis, L1 use, highly proficient L2 learners, L1 Spanish, L2 Swedish.
Acknowledgements

There are numerous people who, in one way or another, have contributed to the realization of this thesis, and I would like to take this opportunity to thank you all! However, I wish to express my gratitude to several persons in particular who have been especially important to me throughout the entire process.

First and foremost, I would like to thank my supervisor Niclas Abrahamsson for his support, invaluable competence and professional advice. I am especially thankful to you, Niclas, for encouraging me to continue with my work after a four-year break with the PhD program. I am grateful to Niclas Abrahamsson, Kenneth Hyltenstam and Emanuel Bylund who, as project leaders and/or project colleagues, as well as co-authors, have provided valuable comments and recommendations on uncountable versions of my work, thereby improving my scientific thinking and writing. I also would like to express my gratitude to my secondary supervisor, Ulla Sundberg, for her support and constructive feedback and guidance, especially on phonetic aspects and methodological issues of the thesis.

Within the framework of the two research projects I have been engaged with, I would like to thank all participants for their patience while conducting the time-consuming and demanding testing sessions that the study required. I would like to thank Alejandra Donoso, Johan Roos and Gunilla Thunberg for recruitment of participants and data collection. Special thanks go to Alejandra Donoso and Anna Ericsson for lending me their voices for the preparation of the perception stimuli. I am also grateful to Hassan Djamshidpey for technical assistance during recordings in the anechoic chamber, Olle Engstrand for suggestions and discussions on the perception tests, and Hartmut Traunmüller, Francisco Lacerda and Guillermo Montero Melis for statistical advice. I would also like to thank Lamont Antieau for checking and improving my English writing.

To my former and present colleagues at the Centre for Research on Bilingualism I would like to extend a warm collective thanks for all the endless conversations we have had, no matter whether private or work-related, and many of which included a god deal of laughter. I am thankful for the opportunity I have been given to be a part of such a stimulating, warm and friendly work environment.

I would like to extend my warmest thanks to my friends and family, especially to my parents, for always being there for me and for supporting me in all my decisions through the years. Finally, my deepest thanks to you, Björn, for always believing in me. Without you by my side I would probably not be where I am today!

Katrin Stölten
Stockholm, October 2013
The present thesis is based on the following studies:


The Effects of Age of Onset on VOT in L2 Acquisition and L1 Attrition

A Study of the Speech Production and Perception of Advanced Spanish-Swedish Bilinguals

Introduction

This thesis is based on three studies that examine the role of age in second language (L2) acquisition and first language (L1) attrition. The focus is on one specific phonetic aspect, namely, Voice Onset Time (VOT), in the production and perception of word-initial L1 and L2 stop consonants in highly advanced L1 Spanish speakers of L2 Swedish. Based on a maturational view and the Critical Period Hypothesis (CPH) for L2 acquisition, Study I aims at exploring the impact of age of onset (AO) of L2 acquisition on the production of L2 Swedish stops and underscores the importance of speaking rate effects on VOT as a measure of nativelikeness. Study II also deals with the CPH, but this time by analyzing AO effects on the categorical perception of L2 Swedish stops. By combining the results with the data from Study I, Study II suggests that the phonetic/phonological L2 acquisition of VOT in stop production and perception may be especially sensitive to AO effects. By investigating the role of age of reduced contact (ARC) for the production and perception of L1 Spanish stops, Study III reveals age as an important predictor for L1 attrition, but only for stop production. The study makes an attempt to explain the discrepancy found between L1 stop production and perception ability by referring to differences in activation thresholds according to the Activation Threshold Hypothesis (ATH) and the amount of L1 use.

This introduction aims to provide a theoretical background to the individual studies. Section I gives an overview of research on age effects in L2 acquisition, thereby presenting definitions and theoretical issues that are especially relevant for Study I and Study II. In section 2, an introduction is given to language attrition research, and some of the most important variables that are of relevance are reviewed. Among these, three factors are presented that are central to Study III: ARC, the amount of L1 use, and the ATH. Section 3 provides a comprehensive description of VOT and reviews the methodology generally used in research studies comprising measurements on stop production and categorical perception tests on the basis of VOT. A theoretical background is given in section 4 to the application of VOT to research on L2 acquisition and L1 attrition. In section 5, a short presentation of the two projects from which the data were taken is followed by a summary of Study I–III. Finally, section 6 makes some general conclusions that can be drawn from the individual studies.
1. Theoretical background to age effects in L2 acquisition

Drawing from an extensive body of research on infant speech perception, it has become a well-established fact that newborns are innately capable of perceptually discriminating a wide range of phonetic contrasts used phonologically in natural languages (e.g., Eilers & Minifie, 1975; Eilers, Wilson & Moore, 1977; Eimas & Miller, 1980; Eimas, Siqueland, Jusczyk & Vigorito, 1971; Jusczyk & Thompson, 1978; Kuhl, Tsao & Liu, 2003; Lasky, Syrdal-Lasky & Klein, 1975; Swoboda, Morse & Leavitt, 1976; Trehub, 1973, 1976). During the first year of life, however, a reorganization of this perceptual ability takes place as a function of increasing specific language exposure (Kuhl, 1987, 2001; Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992; Morse, 1978; Walley, 2005; Werker, Gilbert, Humphrey & Tees, 1981; Werker & Tees, 1984). In other words, it has become a widely accepted view that “early development entails a shift from a language-general to a language-specific pattern of perception” (Walley, 2005, p. 449), in that newborns innately equipped with sensitivities to perceive a variety of phonetic contrasts already during their first months of life become perceptually attuned to the sound system of what will become their native language. By doing so, infants experience an early reorganization of perceptual abilities that makes them increasingly insensitive to non-native contrasts.

Within the field of language acquisition, one central issue has concerned the Critical Period Hypothesis (CPH). In his 1967 book entitled Biological Foundations of Language, Eric Lenneberg proposes the existence of a timeframe of heightened sensitivity to language acquisition that spans from infancy to the onset of puberty (approximately age 12). Lenneberg (1967) emphasizes the role of biological maturation by relating the termination of the critical period to the completion of cerebral lateralization for language to the left hemisphere and to “a loss of adaptability and inability for reorganization in the brain (…)” (p. 179). Support for this hypothesis derives from reports of (fortunately) very rare incidences in which individuals under tragic circumstances have been completely isolated from L1 linguistic input from birth or very early childhood (for reviews of, e.g., the cases of “Genie” and “Chelsea”, see Hyltenstam & Abrahamsson, 2003a; Newport, 1990; with reference to Curtiss 1977, 1989). When provided with linguistic stimuli later in life (i.e. after the onset of puberty), these individuals never reached normal command of their first language.

However, other proposals have been put forward regarding the upper limit of such a biologically defined critical or sensitive period for first language acquisition. Among them, Ruben (1997) suggests that there may not be one, but several critical periods (i.e. for phonology, morphosyntax and semantics; for similar

---

1 The notion of a sensitive period sometimes found in the literature constitutes a weaker/softer formulation of the CPH that implies that the ability to achieve nativelike L2 proficiency does not change abruptly at a certain age; instead, sensitivity is reduced over a span of years. However, these two terms are often used interchangeably (see Hyltenstam & Abrahamsson, 2003a).
suggestions, see Seliger, 1978) of which “the earliest specialization is phonological, with a limiting of phoneme discrimination by the end of the first year” (p. 117). Ruben’s (1997) suggestion of a time span of heightened phonetic sensitivity coming to an end already at the age of 12 months thus fits properly into the pattern observed in the early development of L1 phonetic categories.

Given the empirical evidence from research on L1 acquisition, it has been a heavily debated issue whether such theories of one or several critical periods might also apply to second language acquisition (SLA). In accordance with Lenneberg’s maturational view, some studies propose the crucial age for successful L2 attainment to be around puberty or adolescence (for morphology and syntax, see, e.g., Long, 1990; Patkowski, 1980; for phonology, see, e.g., Scovel, 1969, 1988). Other researchers, however, suggest much lower ages (e.g., six or seven years) to constitute the upper limit for nativelike L2 proficiency (for morphology and syntax, see, e.g., Hyltenstam, 1992; Johnson & Newport, 1989; for phonology, see, e.g., Flege, Munro & MacKay, 1995a; for overviews on critical periods in L2 acquisition, see Long, 1990, 1993, 2005). Despite these differences, studies have consistently reported on overall negative correlations between age of onset (AO) of L2 acquisition and any measure of L2 ultimate attainment indicating that childhood acquirers generally outperform L2 learners with a late start (e.g., Abrahamsson, 2012; Abrahamsson & Hyltenstam, 2009; Asher & García, 1969; DeKeyser, 2000; Flege, 1999; Johnson & Newport, 1989; Munro & Mann, 2005; Oyama, 1976; Patkowski, 1990; Seliger, Krashen & Ladefoged, 1975; see also DeKeyser & Larson-Hall, 2005; Hyltenstam & Abrahamsson, 2003a; Long, 1990, 1993, 2005, for overviews). Findings like these have frequently been interpreted in relation to Lenneberg’s (1967) view of decreasing brain plasticity/flexibility and other neurological correlates such as the development and maturation of cortical neurons (e.g., Pulvermüller & Schumann, 1994; Uylings, 2006).

The existence of maturational/biological constraints as explanations to age-related differences in L2 ultimate attainment has been of much controversy in SLA research. Most of all, skepticism arises from studies that have identified individual cases of late L2 acquirers (i.e. post-puberty or adult learners) with a L2 command comparable to that of native speakers (e.g., Birdsong, 1992, 2007; Birdsong & Molis, 2001; Bongaerts, Mennen & van der Slik, 2000; Bongaerts, Planken & Schils, 1995; Bongaerts, van Summe ren, Planken & Schils, 1997; Flege, Munro & MacKay, 1996; Ioup, Boustagui, El Tigi & Moselle, 1994; Major, 1987; Moyer, 1999; van Boxtel, Bongaerts & Coppen, 2005; White & Genesee, 1996; see also Birdsong, 2006, for an overview). Based on these findings, some researchers have proposed a pattern of constant ability, suggesting that successful L2 learning is, in principle, biologically possible for any L2 learner regardless of their AO (see, e.g., Bongaerts, 1999; Bongaerts et al., 1997; see also Munro & Mann, 2005, for an overview).

However, variability exists regarding the estimation of the incidence of nativelike attainment among late L2 learners. Suggested rates range from very high
Effects of Age of Onset on VOT in L2 Acquisition and L1 Attrition

percentages (above 20%) of nativelikeness (see, e.g., Birdsong, 1992; Birdsong & Molis, 2001; Montrul & Slabakova, 2003; White & Genesee, 1996) over more moderate estimates of 5 – 20% (see, e.g., Birdsong, 2007; Bongaerts, 1999; Colantoni & Steele, 2006; Flege, Yeni-Komshian & Liu, 1999; van Boxtel et al., 2005) to very low rates, suggesting that post-puberty nativelike L2 attainment is very rare or even non-existent (e.g., Abrahamsson, 2012; Abrahamsson & Hyltenstam, 2009; Bley-Vroman, 1989; Ioup et al., 1994; Johnson & Newport, 1989; Moyer, 1999). Study I and Study II presented in this thesis can also be added to the literature viewing nativelikeness in late L2 learners as more or less unattainable.

The reason for these discrepancies could be related to the way in which the notion of nativelikeness has been used in different studies. As pointed out by Abrahamsson and Hyltenstam (2009), overestimations of the incidence of nativelikeness may be the result of overgenerous interpretations of the concept of nativelikeness. Whereas nativelikeness is sometimes measured on the basis of L2 speakers’ self-identifications by some researchers (e.g., Piller, 2002; Seliger et al., 1975), others examine nativelikeness through perceptual judgments made by native speakers of the target language (e.g., Bongaerts, 1999; Moyer, 1999; Neufeld, 2001). A third way of measuring nativelike L2 attainment is by well-controlled investigations of L2 speakers’ actual linguistic proficiency (e.g., Abrahamsson & Hyltenstam, 2008, 2009; Birdsong, 1999, 2007; Bley-Vroman, 1989; Hyltenstam & Abrahamsson, 2003b; Long, 1990; see also Study I and Study II). According to Abrahamsson and Hyltenstam (2009), only the latter operationalization of the notion of nativelikeness offers a reliable measure for capturing subtle inter-speaker differences that would allow for a distinction between near-native and truly nativelike L2 learners regarding their L2 linguistic competence.

Another problem concerns the observation that many investigations have often focused on randomly selected L2 speakers assessed through basic language tests and insufficiently sensitive analyses and elicitation techniques. The use of such methodology, the argument goes, has most probably resulted in ceiling effects and overestimations of nativelikeness among (late) L2 learners (see, e.g., Abrahamsson & Hyltenstam, 2009; Hyltenstam & Abrahamsson, 2003a, 2003b; see also Long, 1993, 2005, 2007). Moreover, even if some post-puberty or adult learners might, in fact, attain a nativelike command, it will most likely be confined to a few areas of the target language (Hyltenstam & Abrahamsson, 2000, 2003b; see also Abrahamsson & Hyltenstam, 2009). Consequently, actual nativelikeness is unlikely to be present in all relevant aspects of the language. If the goal is to challenge maturational accounts such as the CPH, investigations of nativelike L2 ultimate attainment should focus exclusively on highly advanced late L2 speakers, that is, bilinguals who give the appearance of having attained a nativelike command of their L2 (e.g., based on a perceptual judgment task). These individuals can then be examined through demanding tests that, in combination with carefully prepared elicitation methods and techniques, allow for greater in-depth analyses of a wider range of language-specific features (e.g., Abrahamsson & Hyltenstam,
As correctly pointed out by Long (1993), initial screening for potential nativelikeness is of great importance since “[t]here is no value in studying obviously non-native-like individuals intensively in order to declare them non-native-like” (p. 204). Only if late L2 learners, after such intensive evaluation, are in fact identified as having attained a proficiency level completely comparable to that of native speakers, would they most likely represent those exceptions that some researchers would refer to as evidence against maturational/biological factors constraining L2 acquisition. However, no published study has ever confirmed the existence of such exceptional adult L2 learners and thus no claim can actually be made for their status as valid counterevidence to the CPH (Hyltenstam & Abrahamsson, 2000, 2003b; see also Long, 1990, 1993).

Recent studies, among them also Study I and Study II, have shown that, even among early L2 learners (i.e. with an AO before puberty), the incidence of non-nativelike L2 attainment seems to be much more frequent than often presumed (Abrahamsson, 2012; Abrahamsson & Hyltenstam, 2008, 2009; Bylund, Abrahamsson & Hyltenstam, 2012; Flege, Birdsong, Bialystok, Mack, Sung & Tsukada, 2006; Flege et al., 1995a; Hyltenstam & Abrahamsson, 2003b; McDonald, 2000; Tsukada, Birdsong, Bialystok, Mack, Sung & Flege, 2005; Yeni-Komshian, Flege & Liu, 2000). Due to the fact that findings like these are not immediately predicted by the CPH, alternative suggestions have been put forward. Among them, a sigmoid pattern has been proposed. According to this pattern, there is a gradually changing slope of L2 proficiency that is characterized by an initial period during which nativelike L2 proficiency is generally acquired, followed by a period of increasing non-nativelikeness that finally ends in a period of maximal divergence from native-speaker proficiency (e.g., Flege & Fletcher, 1992; Flege et al., 1999). Still others prefer the pattern of an age-related linear decline instead of an abrupt change in language learning outcome at a certain age. Researchers supportive of this approach interpret the decline as a linear deterioration of general cognitive abilities in combination with social, psychological and educational circumstances throughout the entire life span (e.g., Bialystok & Hakuta, 1999; Hakuta, Bialystok & Wiley, 2003; Marinova-Todd, Marshall & Snow, 2000). Whereas this pattern is taken as evidence against the CPH and maturational explanations by some researchers (e.g., Bialystok & Hakuta, 1999), others propose that a continuous decline in the ability to acquire nativelike ultimate attainment might just as likely be the result of maturational constraints that are increasing with age, especially before the onset of puberty (Hyltenstam & Abrahamsson, 2003a).

2. Theoretical background to age effects in L1 attrition

During the last decades, researchers have been increasingly interested in the process of language attrition, a term referring to the “non-pathological decrease in proficiency in a language that had previously been acquired by an individual” (Köpke & Schmid, 2004, p. 5). As schematized in Figure 1, language attrition
Effects of Age of Onset on VOT in L2 Acquisition and L1 Attrition

should be distinguished from pathological language loss caused by neurological impairment leading to aphasia (e.g., Berko-Gleason, 1982; Bylund, 2008). Moreover, attrition implies that speakers undergo individual language loss within one generation. Language attrition should thus be clearly separated from more wide-ranging sociolinguistic contact phenomena such as language shift or death occurring across generations in bilingual communities (Köpke, 2004, with reference to de Bot, 2001).

According to the Crosslinguistic Influence (CI) framework proposed by Pavlenko (2000, 2004; see also Schmid, 2011), attrition is defined as the loss of L1 elements, resulting in the inability to produce, perceive, or recognize specific lexical items or rules in the L1. Given this definition, attrition should be differentiated from four other underlying processes taking place in the interaction between two languages, that is, (1) borrowing, (2) restructuring, (3) convergence, and (4) shift. The first process usually refers to the borrowing of lexical loanwords from the L2 into the L1. During the process of restructuring, L2 elements are incorporated in the L1, thereby leading to changes and/or substitutions in that language (e.g., the incorporation of L2 rules into L1 grammar in the case of syntactic restructuring). An often observed change taking place in the area of phonetics/phonology is the process of convergence, which implies the creation of a new sound (system) that diverges from the L1 and L2 (e.g., intermediate VOT, see section 4). Finally, the process of shift implies that L1 structures or sounds move away from the L1 norm and approximate the L2 structures or sounds (Pavlenko, 2004). However, as relevant as this separation into different processes underlying cross-linguistic influence might be, the broader definition of attrition as non-pathological decrease in L1 proficiency is preferred in this thesis, as it refers to any type of deviation from the L1 norm. By adhering to this definition, adequate comparisons and generalizations between studies can be accomplished.

Research has repeatedly demonstrated that age of reduced contact (ARC) is a highly important predictor variable for the attrition outcome (e.g., Bylund, 2009a; Köpke & Schmid, 2004). ARC is defined as the age at which an individual moves from an L1-dominant environment to an L2 setting and experiences a decline of

![Figure 1: The definition of language attrition and its differentiation from aphasia, language shift and language death (based on Schmid, 2011).](image-url)
continuous contact and input from the native language. At this point in time, the attrition process is supposed to begin.

Comparisons between adult and child attrition have shown that the linguistic system in children is affected to a higher degree than what has been reported for adults. In short, whereas attrition in late bilinguals is generally confined to obvious limitations in vocabulary retrieval (e.g., Olshtain & Barzilay, 1991) and deviances in word order (e.g., Schmid, 2002), the L1 in children is affected to a greater extent to include, for example, insufficiencies with regard to case marking (e.g., Polinsky, 1997), verb morphology (e.g., Montrul, 2002; Turian & Altenberg, 1991), and pronunciation (e.g., Yeni-Komshian et al., 2000). Findings like these have resulted in the general view that “the younger the child is when the language of her environment changes, the faster and deeper she will attrite” (Köpke & Schmid, 2004, p. 10).

Studies on both early and late bilinguals often report on general age effects on L1 proficiency (Ammerlaan, 1996; Bylund, 2009a; Hakuta & D’Andrea, 1992; Pelc, 2001; Silva-Corvalán, 1994; for pronunciation, see Yeni-Komshian et al, 2000). However, based on the data gathered by Bylund (2009a) on conceptualization patterns of L1 goal-oriented motion events in Spanish-Swedish bilinguals and by Yeni-Komshian et al. (2000) on pronunciation proficiency in native Korean speakers of L2 English, it may be suggested that a major change in attrition susceptibility takes place at the onset of puberty, or to be more precise, at the age of approximately 12. Other studies that have investigated L1 attrition in bilinguals who have moved to an L2 environment in adolescence or adulthood, however, report on non-significant effects of ARC on L1 attrition (e.g., de Leeuw, Schmid & Mennen, 2007, 2010; Scherag, Demuth, Rösler, Neville & Röder, 2004; Schmid, 2002). Taken together, these findings suggest that ARC is a predictive factor of L1 attrition/retention, especially in bilinguals who leave the L1 setting before age 12. For late bilinguals (ARC > 12), however, L1 retention remains fairly high, even after many years of residence in the L2 environment. The attrition study presented in this thesis (Study III) makes a contribution to these research findings by demonstrating age differences in that the group of Spanish-Swedish bilinguals with high ARCs (above 12) produced L1 Spanish stops within the monolingual range to a significantly higher degree than the group with low ARCs (below 12). At the same time, however, Study III adds new information to the role of ARC in L1 attrition as it identifies age effects for productive skills, but not for the perception of L1 Spanish stops.

As with the discussion on age findings in SLA research, age effects in L1 attrition have sometimes been related to the existence of a critical or maturational period that spans from birth to the onset of puberty. It is hypothesized that attrition susceptibility gradually decreases during this time span due to increasing maturational constraints before it drops more dramatically at its terminus, that is, at the age of approximately 12 (Bylund, 2009b). In other words, when compared to the age pattern predicted for L2 acquisition, it may be suggested that “the effect
of the Critical Period is reversed with respect to language attrition” (Schmid, 2011, p. 73).

Although age is often cited as the most important variable for L1 attrition and retention, a maturational/biological account alone cannot explain all variation observed in the attrition outcome. In fact, throughout the years myriad variables have been proposed to affect L1 maintenance to various degrees (e.g., Bylund, 2008; Köpke, 2004; Köpke & Schmid, 2004; Schmid, 2011, for an overview). Among them, length of residence (LOR), which is a factor sometimes mentioned as a possible predictor of L2 proficiency (Flege & Fletcher, 1992; Purcell & Suter, 1980; for an overview, see Piske, MacKay & Flege, 2001), is often assumed to play an important role in L1 attrition, in that the longer the time since arrival to the L2 environment, the harder it is for the individual to maintain access to the L1. This assumption, however, is generally not supported by empirical findings (e.g., Bylund, 2009a; Bylund, Abrahamsson & Hyltenstam, 2010; de Leeuw, Mennen & Scobie, 2011; Hopp & Schmid, 2011; Jaspaert & Kroon, 1989; Tsimpli, Sorace, Heycock & Filiaci, 2004). In the study by Yeni-Komshian et al. (2000), the LOR of Korean-English bilinguals in the US did not have any significant effect on either L1 or L2 pronunciation scores. Non-significant correlations between LOR and L1 proficiency are also reported for L1 conceptualization patterns and grammatical intuition measured by a grammaticality judgment test (GJT) in Spanish-Swedish bilinguals (Bylund, 2009a; Bylund et al., 2010). This also holds true for the production and perception of L1 Spanish stops analyzed in Study III, which, like Yeni-Komshian et al. (2000) and the studies by Bylund (2009a) and Bylund et al. (2010), investigated L1 mastery in individuals who had been living in the L2 environment for more than one decade. Hutz (2004), however, examined attrition effects longitudinally on several linguistic levels (e.g., lexical borrowing, semantic transfer and word order) in the writings of a German immigrant to the US. The lexicon was found to be severely affected by the time since immigration, especially during the first 10 to 15 years, whereas morphological and syntactic structures appeared to be more resistant to attrition. In reference to unpublished data by Waas (1993) that reported on considerable attrition effects in German migrants who had been living in Australia for less than ten years, de Bot and Clyne (1994) arrived at the conclusion that LOR can play some role in L1 maintenance up to ten years after the break with the L1 environment. L1 linguistic skills not affected during this time period, however, remain more or less stable, even after many years in the L2 setting (e.g., Ammerlaan, 1996; de Bot & Clyne, 1994; Gürel, 2004a; Schmid, 2002). Moreover, as pointed out by de Bot, Gommans and Rossing (1991), LOR may only have an effect when contact with the L1 is low.

Bylund et al. (2010) point out that language learning aptitude, known to play an important role for L2 ultimate attainment (DeKeyser, 2000; Abrahamsson &

---

2 Language learning aptitude is referred to as an “innate, relatively fixed, talent to acquire and process language structure” (Bylund et al., 2010, p. 447) and can be divided into four major
Hyltenstam, 2008) and identified as a predictor variable in L2 attrition (Lambert, 1982; Oxford, 1982), “has a compensatory function in situations of reduced L1 contact, in that the speaker’s degree of aptitude to a certain extent regulates his/her dependency on L1 contact to achieve and maintain L1 proficiency” (Bylund et al., 2010, p. 459). In other words, language aptitude seems to play a significant part in L1 attrition, especially in situations where the linguistic input from the L1 is reduced. This statement automatically leads us to the next variable often considered to counteract L1 attrition, that is, the amount of L1 contact/use.

A variety of studies have reported on positive effects of L1 contact on L1 maintenance, thereby giving support to the general view that the higher the amount of L1 contact, the higher the L1 proficiency and vice versa (e.g., Bylund et al., 2010; de Bot et al., 1991; Hakuta & D’ Andrea, 1992; de Leeuw et al., 2007, 2010; Lubínska, 2011; Schmid, 2002; Yeni-Komshian et al., 2000). However, other researchers have not been able to find a significant relation between these variables (e.g., Guion, Flege & Loftin, 2000; Hopp & Schmid, 2011; Schmid, 2007), and Jaspaert and Kroon (1989) even report on a negative correlation, indicating that a higher contact with the L1 can result in lower mastery of L1 skills. These discrepancies may have been caused by the fact that studies have used different research methods that can lower the possibility of adequate comparisons between them. For example, whereas de Bot et al. (1991) operationalize L1 contact/use as a dichotomous variable (i.e. many L1 contacts vs. few L1 contacts), Bylund et al. (2010) makes use of self-reported percentages of daily L1 use. In Study III, L1 use is defined by self-reported percentages that were divided into high-level and low-level L1 use. Moreover, information about L1 contact and use is usually gathered through subjective estimations of individuals that may have been influenced by language attitudes (see, e.g., Köpke, 2004; Köpke & Schmid, 2004). Schmid (2002), for example, showed that emotional trauma caused by exceptional circumstances, such as ethnic persecution, can generate negative attitudes toward the L1, thereby facilitating attrition of that language. In addition, characteristics in L1 input per se may have influenced the research outcomes. For example, in situations where there is L1 contact between bilingual speakers within the L2-dominant environment, the L1 input may already have been influenced by the L2, leading Köpke (2004) to the assertion that “not only quantity of input should be considered, but also quality” (p. 20). This assumption may be further exemplified by de Leeuw et al. (2010), who reported on perceived global foreign accent in the L1 speech of 34 German immigrants in Anglophone Canada (ARC components usually incorporated in tests batteries on language aptitude: (1) phonetic/phonemic coding ability, or the ability to identify speech sounds and to make associations between sounds and symbols; (2) grammatical sensitivity, that is, the capability to identify the grammatical functions of constituents in a sentence; (3) rote learning ability, which refers to the capacity to learn and remember new words rapidly and efficiently, thereby making associations between lexical forms and meaning; and (4) inductive learning ability, that is, the ability to infer the grammatical rules of a set of a previously unknown language material (Carroll, 1981).
10 Effects of Age of Onset on VOT in L2 Acquisition and L1 Attrition

14–40; LOR, 9–54) and 23 German immigrants in the Dutch Netherlands (ARC 16–51; LOR 16–58). Accent ratings revealed that the quantity and quality of the participants’ contact with L1 German were better predictors of perceived foreign accent than LOR and ARC. L1 contact was operationalized in terms of L1 use in the context of communicative settings (1) in which little codemixing between L1 and L2 was expected to occur, and (2) in which codemixing was more common. It was found that only L1 use with little codemixing had a significant impact on perceived foreign accent. In a study by Schmid and Dusseldorp (2010), L1 use at work was found to have a more preventive effect against language attrition than informal language use with family and friends. Taken together, these findings suggest that L1 use for professional purposes in which codemixing between L1 and L2 is expected to occur less frequently (e.g., at work) is a more powerful predictor variable for L1 maintenance than informal L1 use in which codemixing is expected to occur more frequently (e.g., with families and friends).

Continuous contact with the L1 has been found to be especially important for L1 retention in individuals who experience a change in linguistic environment early in life. Empirical evidence derives from studies on international adoptees. Ventureyra, Pallier and Yoo (2004) investigated perceptual discrimination of the Korean phonemes /k/, /p/ and /s/ in 18 individuals who had been adopted from Korea by French-speaking families in France, Switzerland and Belgium at the ages of 3–9 years. The control groups for this study contained 12 native Koreans and 12 native speakers of French. Results showed that the adoptees did not have any advantage over native French speakers regarding the perception of Korean voiceless consonants, and the researchers concluded that the L1 is prone to severe attrition, or is even lost completely at the expense of L2 acquisition, if continuous contact with the L1 is severed effectively (see also Isurin, 2000; Nicoladis & Gra- bois, 2002; Pallier, Dehaene, Poline, LeBihan, Argenti, Depoux & Mehler, 2003). Moreover, the findings imply that L1 linguistic domains such as phonology that are commonly believed to be relatively resistant to attrition due to early acquisition may be lost (e.g., Ventureyra et al., 2004; see also Schmid & Köpke, 2007).

Hyltenstam, Bylund, Abrahamsson and Park (2009) studied 21 Korean adoptees in Sweden, but in contrast to Ventureyra et al. (2004) the participants took part in relearning activities as adults (i.e. Korean language classes at university level). The native Koreans’ age of adoption varied between 1 and 10 years, and their LOR in Sweden was 22 years on average. Prior to their participation in Korean language classes, all participants had not received exposure to Korean since adoption (i.e. for an average of 22 years), and they did not report having any knowledge of Korean. In order to find out whether there were any L1 remnants in the adoptees’ language competence, they were tested by using a GJT and a discrimination task of the Korean voiceless stops /p t k/. Test results were compared to three native Koreans and to a group of eleven advanced Swedish learners of Korean who were enrolled in the same university course as the adoptees. Whereas the L2 learners exhibited significantly higher scores on the GJT than the adoptees,
the results from the stop discrimination task revealed that as many as 7 out of the 21 adoptees performed better than the highest-scoring native Swedish learner of Korean. Moreover, the individuals with the highest perception scores also showed the highest ARC scores, leading the researchers to the conclusion that intensively re-learning the L1 (i.e., re-establishing intense L1 contact and use) facilitates reactivation of L1 remnants and that the degree of reactivation is related to ARC. Also, the study by Oh, Jun, Knightly and Au (2003) showed that childhood overhearers of Korean performed better than novice learners on (re)learning L1 phonetic features. Similar to Hyltenstam et al. (2009), this advantage was found for phoneme perception. However, Au, Knightly, Jun and Oh (2002) showed that childhood overhearers of Spanish did not have any advantage regarding the production of higher linguistic structures (i.e., morphosyntax). Taken together, these results suggest that (1) if L1 remnants do exist, they are most probably found in phonology, and that (2) early L1 exposure may have long-term benefits for the reactivation of L1 remnants, especially in perception. Furthermore, these conclusions are in line with the framework of savings paradigm drawn from cognitive psychology on language attrition (e.g., de Bot & Stoessel, 2000) according to which information once learned is not lost, but rather, becomes inaccessible if disused and may be retrieved or reactivated even after a long period of time (e.g., de Bot, Martens & Stoessel, 2004; Hyltenstam et al., 2009).

The finding that early bilinguals, as opposed to late bilinguals, are more dependent on advantageous circumstances such as high-frequency daily L1 use is also confirmed by Study III. The fact that a compensatory function of L1 use in early bilinguals was found for productive skills only, calls for another explanation that takes into account differences between production and perception abilities. Such explanation is offered by the Activation Threshold Hypothesis (ATH) proposed by Paradis (e.g., 1993, 2004, 2007). According to the ATH, linguistic items have activation thresholds that vary due to recency and frequency of use. Based on the assumption that the L1 and L2 linguistic systems in bilinguals are in competition (see, e.g., the Competition Model by MacWhinney, 2005), the ATH proposes that the more particular L2 items are used, the more their activation thresholds will be lowered and thus become faster and more easily activated. At the same time, the thresholds are raised for competing L1 linguistic items that are no longer used. Consequently, different activation threshold levels arise with regard to a speaker’s L1 and L2, making it easier for the bilingual to gain access to L2 items but increasingly more difficult to (re)activate L1 items (see, e.g., Gürel, 2004b; Köpke, 2004; Paradis, 1997, 2007). Against this background, it seems reasonable to assume that the more frequent and extensive the L1 is used, the higher the probability that the L1 will be maintained.

However, different modalities may vary as to how much stimulation they need in order to reach activation, which may explain why production and perception may be affected differently by attrition. By drawing upon the savings paradigm (see above), de Bot et al. (2004) propose that “[F]or recall (i.e. the ability to pro-
duce a word), a fairly high level of activation is needed, while for recognition (i.e. the passive recognition of a word without the ability to actively produce it), a lower level will be sufficient” (p. 375). With reference to Paradis (2007), production (i.e. recall, self-activation) is more difficult to activate since the excitatory neural impulses necessary for the initiation and coordination of the articulators involved in speech must be generated from within, whereas comprehension or recognition are initialized through auditory or visual stimulation on the peripheral sensory organs (cf. the distinction between recall and recognition memory, e.g., Ward, 2010). Following this line of argumentation, it may thus be possible for a bilingual speaker to correctly perceive a particular L1 word or syntactic structure, while not being able to actually produce the same item. Results from Study III are congruent with this view, in that L1 stop perception was generally better retained than stop production with the latter being affected by the bilinguals’ ARC.

3. Voice Onset Time

Lisker and Abramson (1964) define Voice Onset Time (VOT) as “(...) the interval between the release burst of the stop and the onset of glottal vibration” (p. 389). Traditionally, voiceless stops (e.g., /p t k/) are associated with longer VOT values, whereas voiced stops (e.g., /b d ɡ/) are associated with shorter VOTs. In addition to VOT, other acoustic cues are known to contribute to the distinction between voiced and voiceless stops. Voiceless stops, for example, are characterized by high intensity of the release burst, longer duration of the occlusion phase, and delay of first formant energy in relation to the second and third formant in the following vowel (for a review, see Gelfand, 1998; Kent & Read, 1992; see also Abramson, 1977, for a discussion on VOT).³ However, VOT has long been known to be an important, salient cue for voicing in both production and perception of initial stops across languages (Abramson & Lisker, 1973; Caramazza & Yeni-Komshian, 1974; Lisker & Abramson, 1964; Williams, 1980; Zlatin, 1974).

VOT has been shown repeatedly to be affected by various factors. For example, VOT is known to increase as the place of articulation moves further back in the mouth, that is, bilabial stops are generally produced with the shortest, dental stops with intermediate, and velar stops with the longest VOT⁴ (e.g., Cho & Ladefoged, 1999; Fant, 1973; Helgason & Ringen, 2008; Krull, 1991; Lisker & Abramson, 1964; Sundberg & Lacerda, 1996; Voilaitis & Miller, 1992; Yavaş & Wildermuth, 2006). Other factors known to influence VOT are, for example, stress (e.g., Krull, 1991; Lisker & Abramson, 1967; Sundberg & Lacerda, 1999), the

³ Regarding stops occurring in word-final or medial position, an important cue for voicing appears to be the duration of the preceding vowel. In short, stops that are preceded by longer vowel durations are usually perceived as being voiced (see Gelfand, 1989, for an overview).

⁴ However, an observable overlap between places of articulation, in terms of VOT, is often seen in the data (e.g., Helgason & Ringen, 2008; Krull, 1991; Sundberg & Lacerda, 1999, for Swedish stops).
height of the following vowel (Flege, Frieda, Walley & Randazza, 1998; Klatt, 1975; Thornburgh & Ryalls, 1998; Yavaş & Wildermuth, 2006), and speaking rate (e.g., Beckman, Helgason, McMurray & Ringen, 2011; Kessinger & Blumstein, 1997; Schmidt & Flege, 1996; Volaitis & Miller, 1992; see also Study I).

On the basis of VOT, stops are roughly divided into three main phonetic categories (Lisker & Abramson, 1964). As schematized in Figure 2, stops with voicing lead or prevoicing are characterized by vocal fold vibration that starts during occlusion and are indicated by negative VOT values. In short-lag stops and long-lag stops, respectively, laryngeal pulsing is initiated slightly or considerably after the articulatory release. Short-lag and long-lag stops are collectively referred to as having voicing lag, that is, they are defined by positive VOTs. Finally, in stops where the voicing onset coincides with oral release, the VOT value becomes zero.

Languages that differentiate between voiced and voiceless stops phonologically may differ as to how they realize the voicing contrast on the phonetic level. Figure 3 displays the established view of most Germanic languages, including Swedish and English, treating aspirated long-lag stops as voiceless and voiced stops primarily as unaspirated short-lag stops (for English VOT, see, e.g., Caramazza, Yeni-Komshian, Zurif & Carbone, 1973; Deuchar & Clark, 1996, Deuchar & Quay, 2000; for measurements of Swedish stops, see, e.g., Sundberg & Lacerda, 1999). Languages like Spanish or French, however, realize the voiced-voiceless distinction by contrasting prevoiced stops with short-lag stops (e.g., Caramazza & Yeni-Komshian, 1974; Deuchar & Clark, 1996; Zampini & Green, 2001).

However, with respect to Swedish word-initial stops, the short-lag/long-lag division is not as straightforward as might be predicted from Figure 3. Helgason and Ringen (2008) demonstrated that native speakers of Swedish produced phonologically voiced stops in word-initial position more often as prevoiced rather than as short-lag stops. Findings like these (see also Fant, 1969, 1973) imply that the voiced-voiceless distinction in the production of Swedish word-initial stops may be manifested by the use of the extreme categories along the VOT continuum, that is, prevoiced versus long-lag aspirated stops, thereby increasing the perceptual
Effects of Age of Onset on VOT in L2 Acquisition and L1 Attrition

Contrast between them. However, there is still a significant difference between Swedish and Spanish regarding their phonetic realization of word-initial voiceless stops, which is the focus of the production studies incorporated in this thesis.

Cross-linguistic differences regarding the realization of the voicing contrast have frequently been demonstrated in research on stop production and perception. Regarding stop perception, listeners are usually tested on their categorical perception of synthetically generated CV-syllables in which the initial stop varies in equal steps along a continuum of VOT (e.g., Abramson & Lisker, 1973; Caracazzola et al., 1973; Flege & Eefting, 1987; Flege, Schmidt & Wharton, 1996; Gass, 1984; Mack, 1989; Williams, 1977).

Categorical perception is examined through stimulus identification and/or discrimination. In an identification test, a series of stop consonants with varying VOTs are presented to the listener, and he/she is asked to decide for each stimulus independently whether it belongs to either category A or B (e.g., /k/ or /ɡ/). Perceptual discrimination is tested by having the participant listen to stimulus pairs and then decide whether the stimuli are similar or different. Another way of investigating discrimination ability is through an ABX constellation in which the listener is introduced to three stimuli. The first two stimuli always come from different phonemic categories (e.g., /k/ or /ɡ/), and the listener is instructed to decide whether the third stimulus is similar to either stimulus A or B.

The characteristic pattern obtained from a test on categorical perception is that listeners perceptually separate the acoustic VOT continuum into distinct stop categories. As indicated by the idealized results from an identification test (Figure 4a), at a specific point along the VOT dimension an abrupt change in categorization of the VOT stimuli is taking place from voiced (e.g., /ɡ/) to voiceless (e.g., /k/). This sudden change between stop categories is referred to as a category boundary, and it is calculated by the point along the VOT continuum where the identification functions for the two stop categories overlap (Figure 4a). As displayed in Figure 4b, idealized results from a discrimination test show that whereas discrimination between stimuli from different phonemic categories is high, stimuli from within category are not easily discriminated (i.e. not greater than chance). In fact, as

---

**Figure 3:** Schematic representation of the VOT continuum, showing the relationship between the English/Swedish and Spanish stops at the phonetic and phonemic level. “0 VOT” corresponds to stop release (based on Deuchar & Clark, 1996; Zampini & Green, 2001).
comparisons between Figure 4a and Figure 4b reveal, the peak of the discrimination ability coincides with the category boundary, indicating that those VOT stimuli that are identified as belonging to different stop categories are easily discriminated, whereas those stimuli that are identified as belonging to the same voicing category are poorly discriminated. It has been found that native English listeners, for instance, perceive a change in stop category from voiced to voiceless at longer (i.e. more positive) VOTs when compared to native Spanish listeners (Abramson & Lisker, 1973; Flege & Eefting, 1986). Findings like these do not only imply that the categorization patterns are language-specific. They also indicate that patterns found in stop perception are closely related to those in production in which English voiceless stops have longer VOTs than in Spanish (see, e.g., Flege, 1991; Zampini & Green, 2001).

4. VOT in research on L2 acquisition and L1 attrition

Due to cross-linguistic variations regarding the realization of the voicing contrast in stop consonants and the fact that language-specific categories are established at an early stage in language development, VOT has emerged as an acoustic-phonetic feature well-suited for analyses on language proficiencies. Within the field of SLA research, VOT in stop production and perception has been investigated in a vast number of studies (Caramazza et al, 1973; Birdsong, 2007; Flege, 1991; Flege & Eefting, 1987; Flege & Hillenbrand, 1984; Flege & Port, 1981; Laeufer, 1996; MacKay, Flege, Piske & Schirru, 2001; Major, 1987; Obler, 1982; Schmidt & Flege, 1996; Wang & Behne, 2007; Williams, 1977, 1980). Results from these studies generally point in the same direction, namely, that late bilinguals usually
demonstrate VOTs in their L2 stop production and perception that are more deviant from native-speaker proficiency than what is the case for L2 learners with an early start to their L2 acquisition (see, e.g., Abrahamsson, 2012; Flege & Eefting, 1987; Flege, 1991; Flege et al., 1995b; MacKay et al., 2001; Schmidt & Flege, 1996; Thornburgh & Ryalls, 1998). This overall pattern is also confirmed by Study I and Study II on the production and perception of L2 Swedish stops. Moreover, the fact that these age trends were found in exclusively high-level L2 speakers and that the incidence of nativelike L2 command dropped remarkably (especially for the late L2 learners) when the data from Study I and Study II were combined suggests that aspects of phonetics/phonology may be especially sensitive to AO effects.

In contrast to the research available on VOT within the field of SLA, studies investigating phonetic/phonological attrition with regard to VOT in L1 stop production and perception in L2 speakers are more limited in number. One possible reason for this may be the existence of a general belief that phonology is less susceptible to change than higher linguistic structures due to its early acquisition. As stated by Schmid and Köpke (2007), “the phonological system of a mature L1 is probably so stable that it is impervious to L2 influence” (p. 5). However, by referring to the study by Major (1992), Schmid and Köpke (2007) noticed that some evidence of phonetic attrition exists in the speech of late bilinguals. Major (1992) examined VOT for the phonemes /p t k/ in the speech of five female native speakers of American English who as adults had moved to Brazil and who had been living in the L2 environment for 12 to 35 years. The participants had been chosen on the basis of being well-assimilated citizens of the Brazilian society with highly integrative reasons for learning L2 Portuguese as well as equally strong personal and professional motivations for wanting to maintain their L1 English. Comparisons were made between measurements of the bilinguals’ VOT productions and two control groups of native speakers of Brazilian Portuguese and American English. Results revealed that “[t]o a greater or lesser extent, all of the subjects suffered loss of native English proficiency” (Major, 1992, p. 200). Moreover, correlations were found between the bilinguals’ Portuguese VOTs and rate of L1 attrition, indicating that the shorter and thus less nativelike VOT in English, the shorter and more nativelike VOT in Portuguese. The observation that the L1 English VOT seemed to shift toward those of L2 Portuguese is explained as a result of L2 transfer. However, L2 influence was found in stops that had been elicited from a casual speaking style (i.e. informal conversation), whereas no such effect was apparent in stops from a formal speaking style (i.e. word list reading). An explanation offered by Major (1992) is that the bilingual speakers can monitor their formal speech more easily, meaning that they are able to suppress the amount of L2 influence on their L1 in formal than in casual speech.

In a study by Flege and Hillenbrand (1984), VOT was measured in the production of /t/ by native French speakers of L2 English. Similar to the study by Major (1992), the L2 learners had started to acquire their L2 in adulthood and they had
all been living in the L2 setting for 12.2 years on average. The results indicated that the speakers exhibited considerably longer and thus more English-like VOTs than monolingual French speakers so that “(...) none of the talkers in this study closely resembled monolingual native speakers of French” (Flege & Hillenbrand, 1984, p. 716). In another study, Flege (1987) reported on the VOT in the production of /t/ in a group of American English native speakers of L2 French and a group of French native speakers of L2 English. Results revealed that both learner groups tended to produce the stop /t/ in both their languages with VOTs intermediate to the VOTs observed for English and French monolinguals. Findings like these are consistent with earlier studies reporting on intermediate or compromise VOT in the production and perception of L2 and/or L1 stops (Caramazza et al., 1973; Flege, 1991; Flege & Port, 1981; Laeufer, 1996; Obler, 1982; Wang & Behne, 2007; Williams, 1977, 1980).

Observations of cross-linguistic influence in the domains of phonetics and phonology have often been interpreted in relation to the non-maturational Speech Learning Model, or SLM (Flege, 1992, 1995). According to the SLM, L1 and L2 sounds share a common phonological space that facilitates influence between them. It is predicted that L2 learners tend to classify L2 sounds that closely resemble L1 sounds as a member of an already established L1 category due to “equivalence classification” (Flege, 1992, p. 572). Only when speech sounds in the L2 differ substantially from those in the L1 are new perceptual categories established. In this view, age effects found in studies of L2 acquisition and L1 attrition are explained by age-related changes taking place in the interaction between L1 and L2 sounds. It is assumed that the L1 system, as it develops with age, exerts increasing influence on how L2 sounds are perceived and produced. Whereas late learners will continue to identify similar L2 sounds as L1 equivalents, early L2 learners will detect the acoustic differences between corresponding L1 and L2 sounds since their L1 systems are not as fixed as those of adults. Thus, early bilinguals, as opposed to late L2 learners, will eventually establish new phonetic categories for L2 sounds and produce these sounds authentically, that is, like native speakers of the L2 (Flege, 1991). Following this line of argumentation, L1 proficiency is assumed to play an important role in the perception and production of L2 sounds. Seen from the opposite perspective, it is predicted that the higher the L2 proficiency, the more the L2 will increase its impact on the L1. This mutual influence between the L1 and L2 sound system is crucial to the SLM, as it predicts that “the more a bilingual approximates the phonetic norm for an L2 speech sound, the more her production of the corresponding L1 speech sound will tend to diverge from L1 phonetic norms” (Flege, Schirru & MacKay, 2003, pp. 469–470). Against this background, opponents of biological explanations suggest that what some researchers see as age effects in L1 attrition and L2 acquisition should be interpreted as the results of bilingualism, rather than as the outcome of increasing maturational constraints (Flege, 1995, 1999; Ortega, 2010).
However, regardless of the theoretical viewpoint for explaining linguistic variation in L2 acquisition and L1 attrition, the only way of exploring the role of age is by focusing on both early and late bilingual speakers. Whereas differences in L2 proficiency have been studied extensively across a wide range of AOs (e.g., Abrahamsson, 2012; Abrahamsson & Hyltenstam, 2008, 2009; Hyltenstam & Abrahamsson, 2003b; for aspects of phonology, see Abrahamsson, 2012; Abrahamsson & Hyltenstam, 2009; Flege et al., 1995a, 1995b; Yeni-Komshian et al, 2000), systematic analyses of age effects on L1 attrition in both child and adult bilinguals are less numerous. As concerned with attrition in L1 phonetics/phonology in particular, to my knowledge there is only one study that investigates L1 and L2 pronunciation proficiencies in the speech of early and late bilinguals, namely, that by Yeni-Komshian et al. (2000). One reason for this may be related to the difficulty to appropriately analyze L1 proficiencies in bilinguals with low ARCs. To be more precise, researchers examining attrition in children generally run the risk of interpreting lack of L1 (phonetic) competence as a result of attrition whereas the limitation in L1 proficiency may in fact be a consequence of incomplete acquisition of the L1 feature. As pointed out by Bylund (2008, with reference to Montrul, 2002), attrition and incomplete acquisition are two different phenomena, although they may display similar deviational patterns. In fact, Andersen (1982) put forward his view that “[t]he researcher must (…) distinguish true attrition from a failure to acquire the language being studied” (p. 85).

Regarding the attrition study (Study III), the problem of differentiating between language attrition and incomplete acquisition needs some consideration. The study sets out to broaden the view of age effects on L1 attrition in the domain of phonetics/phonology by making detailed analyses on Spanish stop production and perception in early and late Spanish-Swedish bilinguals. As mentioned earlier, language-specific perceptual categories are formed during the first year of life. It may therefore be assumed that although the study includes participants that experienced a break with their L1 setting in early childhood, their perceptual L1 stop categories have already developed to a certain extent. Regarding stop production, research suggests that an adultlike VOT system for Spanish stop production is not established until age 4 (for review, see Macken & Ferguson, 1981). In view of these research findings, possible confusions between incomplete L1 acquisition and L1 attrition in early bilinguals (ARC < 12) are considered to be less problematic for the analysis of categorical perception than for stop production. However, interpretations of what mechanism may have caused deviances in some early Spanish-Swedish bilinguals, especially regarding the stop productions in individuals with ARCs below 5 years, should be taken with caution. By investigating both early and late bilinguals, Study III makes a contribution to the research on L1 attrition, as it reveals age-related changes in the domain of L1 phonetics/phonology, which is generally assumed to be less vulnerable to attrition.
5. Two projects – three studies

The three studies included in this thesis are based on the same group of highly advanced L1 Spanish speakers of L2 Swedish who have been participating in two major projects entitled *Age of onset and ultimate attainment in second language acquisition* (The Bank of Sweden Tercentenary Foundation, grant no. 1999-0383:01) and *First language attrition in advanced second language speakers* (Swedish Research Council, grant no. 421-2004-1975). *Study I* and *Study II* derive from the first project, whereas *Study III* has been conducted within the second project.

As the title already reveals, the central aim of the first project was to explore age effects on ultimate attainment in L2 acquisition. The project examined 41 native speakers of Spanish who had been primarily selected on the criterion that they were perceived as native speakers of their L2 Swedish by a group of native Swedish listeners (for details on participant selection, see Abrahamsson & Hyltenstam, 2009; see also *Study I* and *Study II*). The original intention was to include 60 L2 learners, evenly distributed across an AO span between 1 and 20+ years (that is, with three individuals for each AO). However, perceived nativelike command was biased towards lower AOs and it appeared to be an extremely difficult task to identify late L2 learners (AO > 12 years) who were perceived as native speakers of Swedish and who met all the background criteria (for an overview, see Abrahamsson & Hyltenstam, 2009; see also *Study I* and *Study II*). Therefore, several gaps were left in the upper half of the AO continuum.

In a four-hour session (including two breaks with fruits and refreshments), the 41 highly advanced L2 speakers were tested “across the board”, that is, they performed on some 20 instruments for speech elicitation and language testing that covered a wide range of linguistic levels and skills. The testing sessions were conducted in a sound attenuated room at Stockholm University. Apart from the tests concerning stop production and categorical perception, other test components comprised, for example, speech perception in noise, grammaticality judgments, idiomatic expressions and proverbs, and language learning aptitude. In other words, in addition to *Study I* and *Study II* other empirical studies have also derived from this project (Abrahamsson & Hyltenstam, 2008, 2009; Bylund et al., 2012; Hyltenstam et al., 2009).

Opponents of a maturational account for L2 acquisition often relate to the CPH by accentuating their findings of late L2 learners with a nativelike command and consider them valid counterexamples to the CPH and maturational constraints. However, such statements have often been based on methodological shortcomings (see section 1, as well as *Study I* and *Study II*). Therefore, systematic analyses of AO effects on the L2 learners’ performance on the different test components were

---

5 It should be noted that in the majority of cases, the AO of L2 acquisition coincided with the participants’ ages at immigration, which, in turn, coincided with their ARC.
conducted by examining the entire learner group and through analyses of 31 early (AO 1–11) and 10 late (AO 13–19) L2 learners in comparison with 15 Swedish controls. Participants were tested by a native speaker of Swedish, thereby ensuring that the L2 speakers were in a Swedish monolingual mode (cf. Grosjean, 2001). The L2 learners were separated into an early and late AO group according to the theoretically established hypothesis of a critical period for language acquisition in general (approximately age 12; Lenneberg, 1967). It is important to note, however, that the project never aimed at actually confirming any version of the CPH. This also holds true for the studies on stop production (Study I) and categorical perception (Study II), in which group comparisons were not conducted for making any statements on whether or not age 12 should be considered the fixed cut-off point of a critical period for phonetic detail. Instead, the project and the individual studies intended to evaluate various claims that late nativelike L2 speakers exist and therefore constitute evidence of no critical period or maturational constraints on language acquisition. This was done by investigating whether entirely nativelike L2 command is actually attained by highly proficient L2 learners, especially those with high AOs. Such an evaluation based on linguistic scrutiny has important implications for theories on maturational constraints and discussions on the non/existence of critical period effects in L2 acquisition.

The second project addressed the same group of advanced L2 learners, but this time their proficiency in L1 Spanish was the focus. Based on the background to attrition research (section 2; see also Study III), the overall aim of the project was to explore the impact of ARC on L1 retention/attrition. More specifically, the research question under investigation was whether L1 attrition is less likely to occur in late bilinguals than in individuals who experience a break with their L1 environment early in life, that is, before the age of approximately 12. Again, the theoretical point of departure for the project is posed by predictions made on the basis of a critical (or maturational) period. However, since six individuals (ARC 2, 4, 5, 6, and two individuals with ARC 8) resigned from participation in this project, examinations could only be conducted on 35 out of the original 41 L2 learners with bilingual competence. Moreover, this change resulted in a more irregular distribution even among lower ARCs (see Study III). Test components similar to those of the first project were adapted to Spanish, and again the participants were separated into a group of 25 early (ARC 1–11) and 10 late (ARC 13–19) bilinguals. This time, 15 native speakers of Chilean Spanish who were all Chilean citizens living in Chile served as controls. Twelve of them were tested at the Catholic University of Santiago de Chile, while three were tested during a short visit to Sweden. Data from these three controls as well as from the bilingual speakers were collected at Stockholm University. The same native speaker of Chilean Spanish led all testing sessions.

Research findings generated by the second project have been presented in, for example, a doctoral thesis (Bylund, 2008) and other empirical studies (Bylund, 2009a; Bylund et al., 2010, 2012). The present thesis, specifically Study III, makes
a contribution to this list of publications, in that it is concerned with the bilinguals’ production and categorical perception of word-initial stops in their L1 Spanish.

5.1 Study 1: Effects of age and speaking rate on voice onset time: The production of voiceless stops by near-native L2 speakers

5.1.1 Introduction and Research Questions

For many decades now, the role of age of acquisition and the CPH have been, and still are, much debated issues in SLA research. The general observation that childhood acquirers usually outperform L2 learners with a late start to L2 acquisition has been interpreted by opponents to the CPH as a result of various socio-psychological factors rather than as caused by biological/maturational changes that have taken place in the brain. Moreover, the identification of late (i.e. adolescent or adult) learners with a seemingly nativelike command of their L2 has frequently been taken as evidence against any theory of maturational constraints, including any version of the CPH. However, such studies have been criticized for using insufficiently detailed analyses and overly simple language tests that are not sensitive enough for the crucial separation into near-native and nativelike L2 proficiency (see, e.g., Hyltenstam & Abrahamsson, 2003a). In the last two decades, a branch of SLA research has drawn attention to late L2 learners with a seemingly nativelike L2 proficiency in order to carefully scrutinize them for their actual linguistic L2 command (e.g., Abrahamsson & Hyltenstam, 2009; see also Study II). Results from these studies show that when L2 proficiencies are analyzed “across the board”, that is, through comprehensive analyses and sufficiently challenging languages tests and tasks on different linguistic levels, the incidence of nativelikeness among late learners appears to be close to non-existent.

As part of the main project described above, this study comprised 41 L1 Spanish early and late speakers of L2 Swedish who were selected on the criterion that they were perceived by native listeners as mother-tongue speakers of Swedish. The participants’ L2 proficiencies were then investigated through demanding tests and in-depth measures of a broad repertoire of their L2 Swedish skills (Abrahamsson & Hyltenstam, 2009). With respect to L2 stop production, the present study reports on the detailed analysis of VOT in all three voiceless stops that occur word-initially in Swedish (i.e. /p t k/). Based on the research background regarding age effects and discussions on the non-/existence of nativelikeness in L2 acquisition, the present study focuses on three research questions:

1. Is there a general age effect on stop production (i.e. in terms of negative correlations between AO and VOT or differences in mean VOT between early and late learners) even among L2 learners perceived as native speakers by native listeners?
2. Are there late L2 learners who produce all three voiceless stops with an average VOT within the range of native-speaker VOT?

3. Do all (or most) early L2 learners produce all three voiceless stops with an average VOT within the range of native-speaker VOT?

Most importantly, since VOT is known to be influenced by speaking rate in that VOT decreases as speaking rate increases (Beckman et al., 2011; Kessinger & Blumstein, 1997; Magloire & Green, 1999), each research question incorporates comparisons between outcomes based on absolute VOT (i.e. in milliseconds) and relative VOT (i.e. in percentages of word duration). In other words, speaking rate effects often not accounted for in studies on L2 stop production (e.g., Birdsong, 2007; MacLeod & Stoel-Gammon, 2010; Major, 1987) are controlled for in this study, thereby enabling the researcher to gather more valid data regarding age effects and actual nativelikeness in L2 acquisition or, to be more precise, in the acquisition of phonetic detail.

5.1.2 Method

The 41 Spanish early (AO 1–11) and late (AO 13–19) learners of L2 Swedish were recorded one by one, while reading aloud three isolated Swedish words ten times in a row: par /par/, tal /tal/, and kal /kal/. 15 native speakers of Swedish served as controls. The readings were digitally recorded through a KOSS R/50B headset microphone at 22.050Hz with a 16-bit resolution. VOT was measured in the word-initial stops /p t k/ and word duration was determined by using the waveform display together with synchronized wideband spectrograms. VOT was first analyzed in terms of absolute measurements (i.e. in milliseconds) before it was expressed in terms of percentages of word duration, thereby neutralizing speaking rate effects.

5.1.3 General findings

The results revealed overall negative correlations between age of onset and VOT production. However, this age effect became salient and statistically significant for all three stops ($p < .05$ for /p/, $p < .05$ for /t/, $p < .01$ for /k/) only when VOT was expressed in percentages of word duration (Research Question 1). When VOT was measured in milliseconds, a majority of both the early and late L2 learners demonstrated mean VOTs within the range of native-speaker production. The difference between the group of early and late L2 learners was not even close to being significant ($p = .224$). However, when speaking rate was taken into consideration, only a small minority of the late learners exhibited actual nativelike L2 behavior (Research Question 2), and most (but far from all) early learners performed within native-speaker range (Research Question 3). This time, the group difference between early and late L2 learners became statistically significant ($p = .047$).
5.1.4 Discussion

In SLA research, studies on randomly selected L2 speakers have frequently reported on negative correlations between AO and L2 acquisition. The present study has managed to reveal that even among L1 Spanish speakers who have been exclusively selected on the criterion that they are perceived as mother-tongue speakers of L2 Swedish, there is a statistically significant correlation between AO and the production of all three word-initial voiceless stops in Swedish (Research Question 1). The influence of AO on L2 stop production is further supported by the observation of a significant difference between the group of early and late L2 learners. However, this age effect was only found when VOT was measured in relation to word duration, thereby neutralizing the impact of speaking rate on VOT.

Comparisons in outcomes based on absolute and relative VOT further resulted in minor, but significant, readjustment regarding the distribution of early and late L2 learners who produced the Swedish stops with VOTs comparable to native-speaker stop production. After taking speaking rate into consideration, only a minority of the late, apparently nativelike L2 learners exhibited actual nativelike L2 behavior (Research Question 2) and even among low AOs, many L2 learners did not perform at a native-speaker level (Research Question 3). These findings are in accordance with reports from other linguistic tests (Abrahamsson, 2012; Abrahamsson & Hyltenstam, 2009) and suggest that actual nativelike command in both early and late L2 learners is, in fact, more infrequent, as was earlier assumed. The study comes to the conclusion that theories on maturational constraints and critical period effects on second language acquisition can still not be satisfactorily refuted. Even when the focus is on apparently nativelike late L2 learners, only a very few individuals appear to be entirely comparable to native speakers when analyzed in detail. In fact, when combined with other L2 competencies and skills (Abrahamsson & Hyltenstam, 2009; see also Study II) the incidence of actual nativelikeness among advanced late L2 learners appears to be even lower or, to be more precise, close to non-existent. Moreover, the results indicate that relative VOT, as opposed to absolute VOT, constitutes a reliable measure of nativelike L2 stop production, as it allows for the crucial distinction between near-native and actual nativelike L2 learners, which has important implications for future research on age effects and maturational constraints in L2 acquisition.

5.2 Study 2: Effects of age of learning on voice onset time: Categorical perception of Swedish stops by near-native L2 speakers

5.2.1 Introduction and Hypotheses

Study II originates from the same research background as presented in Study I. As mentioned earlier, L2 learners with an early start of L2 acquisition usually end up with a higher command of their L2 than late L2 learners. This well-established “early is better” effect on L2 ultimate attainment derives from frequent reports on
negative correlations between the learners’ AO and any measure of L2 proficiency (Abrahamsson, 2012; Abrahamsson & Hyltenstam, 2009; DeKeyser, 2000; Flege, 1999; Johnson & Newport, 1989). Findings like these are often explained with reference to maturationally conditioned changes in the brain (Pulvermüller & Schumann, 1994; Uylings, 2006) and the existence of a critical period for language learning (Lenneberg, 1967). Skeptics of biological explanations point to the existence of L2 learners who, despite high AOs, seem to reach nativelike command of the L2 (Birdsong, 2007; Bongaerts, 1999). Such findings, it is argued, are not predicted by any version of the CPH and thus are contradictory to a maturational account for L2 acquisition. However, statements like these often derive from research based on methodological shortcomings that most certainly have resulted in ceiling effects and Type II errors, or “false positives”, in that very advanced L2 learners have been erroneously classified as nativelike rather than near-native (e.g., Abrahamsson & Hyltenstam, 2009; Long, 2005).

As mentioned in Study I, a branch of SLA research has in the last two decades set out to focus on L2 learners who seem to have reached nativelike L2 proficiency or behavior. By linguistically scrutinizing their actual L2 command through challenging language tests and detailed analyses conducted “across the board”, Type II errors are avoided. Results from these studies have revealed that truly nativelike L2 command in late L2 learners is close to non-existent, and, even among early L2 learners, the incidence of non-nativelike behavior appears to be more common than generally believed (Abrahamsson & Hyltenstam, 2009; see also Study I).

As with Study I, this study is part of the same main project described earlier on ultimate L2 attainment in 41 L1 Spanish early and late learners with an exclusively high command of their L2 Swedish. Study II reports on the categorical perception of Swedish word-initial stops. Given the discussion on age and nativelikeness, Hypotheses 1 and 2 predicted that:

1. Even among L2 learners with a seemingly nativelike pronunciation, as measured by native-listener judgments, category boundary placement is affected by the L2 learners’ AO and the status as L1 or L2 speakers of Swedish, in that the native speakers will show the highest and the late L2 learners the lowest (i.e. the most non-Swedish-like) average category crossover values, with the early L2 learners exhibiting values somewhere in between.

2. Even among L2 learners with a seemingly nativelike pronunciation, as measured by native-listener judgments,
   (a) very few (if any) late L2 learners, and
   (b) most (but not all) early L2 learners
   will exhibit average category boundary values within the range of native-speaker categorization.
Since it has been suggested that aspects of phonetics/phonology seem to be most sensitive to critical period effects (Scovel, 1988), it was of interest to combine results from stop perception with previous data from the same L2 learners’ production of Swedish word-initial voiceless stops (Study I) in order to analyze in greater detail the incidence of nativelikeness when phonetic/phonological L2 command is under investigation. On the assumption that stop production and perception are especially vulnerable to the AO factor, Hypothesis 3 was formulated as follows:

3. Even among L2 learners with a seemingly nativelike pronunciation, as measured by native-listener judgments, when results from their stop perception are combined with data from their productions of word-initial voiceless stops (data from Study I), and when all three places of articulation are considered, (a) the incidence of nativelikeness will become virtually non-existent among late L2 learners, but (b) a majority of the early L2 learners will still perform within native-speaker range.

However, in order to make credible statements regarding these hypotheses, it was of high priority to first ensure the reliability and validity of the test design, including the preparation of stimuli and the execution of the listening task. Therefore, the study first investigates whether the listeners, irrespective of their status as L1 or L2 speaker of Swedish and regardless of their AO, perceive the edited natural stimuli in a categorical manner on the basis of VOT.

5.2.2 Method

Prior to the test on categorical perception, three voicing continua had been created on the basis of natural Swedish word pairs with /p–b/ (i.e. par–bar), /t–d/ (i.e. tal–dal), and /k–ɡ/ (i.e. kal–gal) in word-initial position. Each voicing continuum consisted of 30 speech stimuli varying in 5ms increments from –60ms to +90ms VOT, thereby covering the VOT systems of both Spanish and Swedish. Forty-one Spanish early (AO 1–11) and late (AO 13–19) near-native speakers of L2 Swedish performed one at a time on a forced-choice identification task, in which the VOT stimuli were presented binaurally through headphones (KOSS TX/PRO). The three VOT-continua were tested separately. All stimuli were automatically randomized for each listener and embedded in a Swedish carrier phrase. Due to time restrictions, each VOT stimulus was presented only once. The control group comprised 15 native speakers of Swedish, and the experiment leader was a native speaker of Swedish.
5.2.3 General findings

All three listener groups exhibited discernible categorization patterns for all three stop continua. Differences were not significant for either participant group or for stop continuum, indicating that all listener groups irrespective of their L1/L2 status and regardless of AO perceived the three edited natural stop pairs categorically on the basis of VOT. The results further revealed overall negative correlations between category boundary placement and AO. The correlations were statistically significant for the bilabial \((p < .01)\) and the dental stop continuum \((p < .05)\), and very close to significant for the velar stops \((p = .054)\). On the group level, the native speakers exhibited the highest and the late L2 learners the lowest category boundaries with the early L2 learners in between (Hypothesis 1). This group difference was significant for the bilabial \((p < .001)\), the dental \((p = .001)\), and the velar stop continuum \((p = .001)\). Whereas only a small minority of the late learners perceived the voicing contrast in a way comparable to native-speaker categorization (Hypothesis 2a), most early L2 learners demonstrated nativelike categorization patterns (Hypothesis 2b). Again, the difference observed between the early and the late AO group was statistically significant \((p < .01)\). However, when the results were combined with data from the same L2 learners’ production of Swedish voiceless stops (see Study I, nativelike production and perception was not found among the late learners (Hypothesis 3a), but there was still a majority of the early learners that exhibited nativelike production and perception (Hypothesis 3b). This difference turned out to be statistically significant \((p < .001)\).

5.2.4 Discussion

First of all, irrespective of L1/L2 status and regardless of AO the listeners perceived all three stop contrasts categorically on the basis of VOT, thereby confirming the reliability and validity of the test design, including the preparation of VOT stimuli and the performance on the listening task.

Although statistically significant only for the bilabial and the dental stop pair, AO had an impact on all three Swedish voicing continua as indicated by significant differences between the group of early and late L2 learners. Moreover, overall group differences also revealed that category boundary placement was affected by the L2 learners’ status as L1 or L2 speakers. These findings show that AO affects L2 ultimate attainment even in highly advanced L2 learners, if their L2 command is analyzed on the basis of L2 stop perception. This outcome is also in accordance with Hypothesis 1, in that the late L2 learners, as opposed to the early learners, deviated the most from native-speaker perception. In other words, even among L2 learners with a seemingly nativelike command of the target language, the ability to establish L2 phonetic categories on the basis of VOT is limited and becomes increasingly difficult with higher AOs. Whereas only a small minority of the late learners demonstrated actual nativelike L2 command in terms of perceived category boundary (Hypothesis 2a), most, but still far from all, early L2 speakers per-
formed on native-speaker level when their L2 stop perception was analyzed in detail (Hypothesis 2b). However, when the results were combined with the data obtained from the same L2 learners’ Swedish stop production (see Study I), the incidence of nativelikeness on both measures decreased considerably. In fact, entirely nativelike command was never found among the late learners (Hypothesis 3a). The study arrives at the conclusion that earlier estimates of the incidence of nativelikeness in adult learners should be questioned (Birdsong, 2007; Bongaerts, 1999; White & Genesee, 1996). Instead, the findings support the view that complete nativelikeness in late learners is, in principle, never found when L2 linguistic proficiency is scrutinized in detail (Abrahamsson & Hyltenstam, 2009), and it is concluded that the CPH or any theory of maturational constraints on language acquisition cannot be rejected on the basis of these results. Regarding the early L2 learners, a majority of the listeners produced and perceived the Swedish stops within native-speaker range (Hypothesis 3b). However, seen from the opposite side, incidences of non-nativelikeness were observed among early L2 learners, even in those with very low AOs, indicating that an early start to L2 acquisition does not guarantee an entirely nativelike outcome. Moreover, these findings are also congruent with the view that the domain of phonetics/phonology seems to be especially vulnerable to AO effects.

5.3 Study 3: The L1 production and perception of VOT in Spanish-Swedish bilinguals: The role of age and L1 use

5.3.1 Introduction and Research Questions

Research on L1 attrition has repeatedly shown that age (or ARC) is a decisive factor for L1 retention, especially if the break with the L1 setting is located early in life. Moreover, as has been suggested by Bylund (2009a), a major change in attrition susceptibility may take place at the onset of puberty, or, more precisely, at the age of approximately 12 years. Findings like these have been related to maturational constraints and theories of a critical or maturational period (Bylund, 2009b). However, other factors apart from ARC have been suggested to have an impact on L1 attrition and retention, among them L1 use (Bylund, 2009a; Schmid, 2002; Yeni-Komshian et al., 2000) and the ATH (Paradis, 2004, 2007). Against this background, the study sets out to explore the effects of ARC, L1 use and the ATH on the production and perception of L1 Spanish stops (in terms of VOT) in Spanish-Swedish bilinguals. As mentioned above, the study is part of a larger project on 35 L1 Spanish speakers with near-native proficiency levels of L2 Swedish who have experienced a break with their L1 Spanish at varying ARCs (1–19 years). At the time of testing, all bilinguals were functional in Spanish and used the L1 on a regular basis as indicated by self-reported percentages of daily L1 use. With reference to the study by Bylund (2009a) pointing to the age of approximately 12 years to be critical for L1 retention, Research Questions 1 and 2 address the following:
1. Is there an age effect on VOT in the production of L1 Spanish stops, in that late bilinguals (ARC > 12) exhibit Spanish-like VOTs to a greater extent than early bilinguals (ARC < 12)?

2. Is there an age effect on the categorical perception of VOT of L1 Spanish stops, in that late bilinguals (ARC > 12) exhibit Spanish-like category boundaries to a greater extent than early bilinguals (ARC < 12)?

Due to differences in activation levels between production and perception (as predicted by the ATH), L1 maintenance may be expected to be greater for stop perception. This motivates the formulation of Research Question 3:

3. Do bilinguals exhibit Spanish-like VOT more often in their perception than in their production of L1 stops?

Within reference to the ATH, a higher frequency of L1 use is supposed to lower activation thresholds, thereby facilitating maintenance of L1 linguistic items. Considering the differences in activation thresholds for production and perception and the differences between early and late bilinguals addressed above, it may be expected that a higher amount of L1 use is needed for the retention of Spanish-like stop production, especially in early bilinguals, leading to the formulation of Research Question 4:

4. Does L1 use have a stronger impact on stop production than on perception, and is this effect more prominent in early than in late bilinguals?

5.3.2 Method

In the stop production task, the 25 early (ARC 1–11) and 10 late (ARC 13–19) Spanish-Swedish bilinguals were instructed to read aloud three isolated words in Spanish ten times in a row: pata /pata/, tata /tata/, and kata /kata/. The control group comprised 15 monolingual speakers of Spanish. The readings were digitally recorded through a KOSS R/50B headset microphone at 22.050Hz with a 16-bit resolution. By using wideband spectrograms synchronized with the waveform display, VOT in the word-initial voiceless stops /p t k/ as well as the duration of the first CV-syllable was measured for each speaker. In order to account for speaking rate effects, VOT was analyzed in percentages of syllable duration.

Prior to the test on categorical perception, three voicing continua were created on the basis of natural Spanish word pairs with /p–b/ (i.e. pata–bata), /t–d/ (i.e. tata–data), and /k–ɡ/ (i.e. kata–gata) in word-initial position. Each voicing continuum consisted of 30 speech stimuli varying in 5ms increments from –60ms to +90ms VOT, thereby ensuring that the VOT systems of both Spanish and Swedish
were covered. In a forced-choice identification task, in which the 35 Spanish-Swedish bilinguals were tested individually, the VOT stimuli were presented binaurally through headphones (KOSS TX/PRO) in a random order for each participant. The stimuli were presented within a Spanish carrier phrase, and all instructions were given by the same experiment leader as in the production task. The same group of 15 monolingual Spanish speakers from the stop production task also served as controls for the test on categorical perception.

5.3.3 General findings

Results from stop production showed non-significant negative correlations between ARC and VOT. However, differences between speaker groups were apparent, in that the monolingual controls produced the shortest VOTs, and the group of the early bilinguals exhibited the longest mean VOTs regardless of place of articulation. Group differences were statistically significant for the labial ($p < .001$), the dental ($p < .001$), and the velar stops ($p < .001$). As indicated by a post hoc test, differences were significant between the groups of early and the late bilinguals ($p < .05$) and between the monolingual group and the early bilinguals ($p < .001$). The control group and the late bilinguals did not differ significantly ($p > .05$) for any of the stops. Individual analyses revealed that whereas the entire group of late bilinguals exhibited mean VOTs within the range of the monolingual controls for either two or three stops, this was found for only 52% of the early bilinguals. This difference was significant ($p = .01$). In accordance with Research Question 1, the results revealed an age effect on L1 stop production, in that the late bilinguals (ARC > 12) exhibited Spanish-like VOT to a greater extent than early bilinguals (ARC < 12). Concerning stop perception, neither significant correlations nor group differences were observed. In other words, there was no age effect on the categorical perception of L1 stops in Spanish-Swedish bilinguals (Research Question 2). Comparisons between stop production and perception revealed that Spanish-like performance was generally more frequent for categorical perception than for L1 stop production ($p < .02$). However, closer analysis of the three places of stop articulation revealed that this difference was significant only for the labial stops, $p = .001$ (Research Question 3). With respect to Research Question 4, high-frequency L1 use (25–50%), as opposed to low-level frequency L1 use (0–24%), resulted in a significantly higher amount of Spanish-like VOT productions ($p < .01$). In contrast, no significant difference between high-/low-frequency L1 use was found for categorical perception ($p = .90$). Moreover, L1 use was found to have a significant effect on Spanish stop production in early bilinguals only ($p < .02$).

5.3.4 Discussion

The study revealed an effect of ARC on the production of L1 stops in Spanish-Swedish bilinguals (Research Question 1), whereas L1 stop perception remained
Effects of Age of Onset on VOT in L2 Acquisition and L1 Attrition

unaffected (Research Question 2). This suggests that the age at which a person experiences a break with the L1 environment plays an important role for L1 retention, especially where production skills are concerned. Differences observed between the group of early and late bilinguals are in accordance with previous studies reporting on changes in attrition susceptibility at the onset of puberty or, more precisely, at the age of approximately 12 (Bylund, 2009a; Yeni-Komshian et al., 2000). Moreover, findings like these seem to support the view of a reversed critical period for language attrition (Schmid, 2011). Whereas maturational constraints may account for the underlying ARC effects on L1 stop production, they alone do not satisfactorily explain why categorical perception was less affected by ARC than production. One explanation is offered by the ATH (Paradis, 2004, 2007), according to which speech perception has lower activation thresholds than production and may therefore be expected to be more easily activated and retained. Although statistically significant only for the labial stops, the overall tendency of Spanish-like performance being more common for perception than for production is congruent with the predictions made by the ATH (Research Question 3). Based on these findings, it may be suggested that changes are more likely to occur in the pronunciation of L1 sounds, whereas the underlying L1 sound system seems to be more resistant to attrition. Moreover, the assumption that perception is activated more easily and is somewhat better maintained than production may offer a contributing explanation to the finding of the former being less affected by ARC. Since the ATH proposes that activation levels are controlled by language use, it was further expected that production as opposed to perception should be more strongly influenced by L1 use. Although higher rates of Spanish-like stop productions were, in fact, related to high-frequency L1 use, this effect was only confirmed for the bilinguals with ARCs below age 12 (Research Question 4). From these findings it is concluded that early bilinguals are more dependent on advantageous factors such as extensive L1 use than late bilinguals in order to compensate for ARC effects, but only where production skills are concerned.

6. Summary and general conclusions

The aim of this thesis has been to explore age effects in L2 acquisition and L1 attrition in L1 Spanish speakers with an exclusively high command of L2 Swedish. More specifically, concerned as it is with phonetic/phonological aspects of L2 acquisition and L1 attrition, this work has focused on age-related differences in the production and categorical perception of L1/L2 stops, in terms of VOT. The original contributions of the three empirical studies included in this thesis are the following:

First, Study I indicates that even in exclusively high-level L2 speakers, there is an overall age effect regarding the production of L2 stops, thereby lowering the rates of actual nativelikeness that have been assumed in previous research. Most importantly, Study I suggests that such conclusions are only justified if the
analysis is based on sufficiently detailed examinations that in the case of VOT account for possible effects of speaking rate.

Second, based on the investigation of categorical perception behavior in the same group of L2 learners as in Study I, Study II gives further support to the observation of age effects in highly advanced L2 speakers. Study II concludes that the ability to establish nativelike L2 phonetic categories on the basis of VOT is limited and becomes increasingly constrained with higher ages of onset of L2 acquisition. Furthermore, Study II accentuates the sensitivity of phonetic/phonological components to age effects. This interpretation is based on the finding that the incidence of nativelike L2 command decreases remarkably, especially among the late L2 learners (AO > 12), when nativelikeness is examined by combining the measures of stop production and perception. Based on the data gathered in Study I and Study II, it is concluded that theories of biological/maturational constraints including the CPH cannot yet be satisfactorily rejected.

Third, Study III investigates the effects of ARC on the production and categorical perception of L1 Spanish stops and identifies the age factor as an important predictor variable for the attrition outcome of production abilities. Whereas age effects on stop production are in accordance with earlier studies and may be related to maturational constraints, such explanations do not support the finding of categorical perception being less affected by attrition regardless of ARC. By referring to the ATH, Study III suggests that this discrepancy between production and perception may be affected to some degree by differences in neurologically based activation levels, leaving pronunciation skills generally more vulnerable to attrition. Finally, Study III highlights the role of L1 use as an advantageous factor in L1 attrition, in that it enables bilinguals, especially those with low ARCs, to compensate for negative age effects on the ability to retain L1 pronunciation.

In sum, this thesis has contributed to broaden the view of age effects in L1 and L2 phonetic development by focusing on one specific aspect, that is, VOT in the production and perception of word-initial stops. Study I and Study II make important contributions to the field of L2 acquisition research in that they emphasize methodological issues including the examination of advanced L2 learners, the detailed analysis of a specific L2 feature, and the combination of measures of nativelikeness. Based on this research agenda, Study I and Study II add new information to the discussion on age effects or, more specifically, on the non-/existence of nativelikeness in L2 ultimate attainment that have important implications for theory building in SLA research. Study III offers an insight into the relation between maturational and non-biological factors and their explanatory value regarding age-related differences in L1 attrition. As has become clear from this thesis, attrition studies exploring the role of age in the domain of phonetics are still scarce when compared to the work that has been done on the acquisition of L2 phonetics/phonology. It is hoped that this thesis will inspire researchers to gather more empirical data regarding both the production and perception of different
Effects of Age of Onset on VOT in L2 Acquisition and L1 Attrition

By focusing on comparisons between productive and perceptual language abilities, it is my hope that future research will arrive at more conclusive explanations and generalizations regarding possible differences in susceptibility to age effects in the attrition of production and perception of L1 skills.

Sammanfattning på svenska

Denna avhandling fokuserar på startålderns inverkan på inlärningen av ett andraspråk (L2) och förlust, s.k. attrition, av ett förstaspråk (L1) i en andraspråksmiljö. Dessa aspekter undersöks utifrån ett fonetiskt fenomen, Voice Onset Time (VOT), som analyseras i produktion och perception av ordinitiala klusiler hos avancerade andraspråksinlärare av svenska med spanska som L1. Avhandlingen baseras på tre individuella studier. Studie I och Studie II undersöker L2-inlärarnas produktion respektive kategoriska perception av svenska klusiler, medan Studie III utforskar L2-inlärarnas produktion och perception av ordinitiala klusiler i deras L1-spanska.


Under de senaste två decennierna har en forskningsgren inom andraspråksinlärningsfältet börjat rikta upp出てくるlenheten mot andraspråksinlärare som verkar

Studie I och Studie II utgör fristående delar av ett omfattande projekt med titeln ”Startålder och slutlig färdighetsnivå vid andraspråksinlärning” (finansierat av Riksbankens Jubileumsfond, 1999-0383:01) där L2-behärskningen hos 41 avancerade tidiga (SÅ 1–11) och sena (SÅ 13–19) L2-inlärare (L1 spanska, L2 svenska) har undersöks genom detaljerade analyser och omfattande lingvistiska tester. Utifrån teorin om en kritisk period för språkinlärning och diskussionen kring ålderseffekter och existensen/avsaknaden av inföddlikhet vid andraspråksinlärning undersöks i Studie I om inlärningsålder har en inverkan på VOT i svenska ordinatia klusiler (/p t k/) producerade av andraspråksinlärare med en hög språkbehärskningsnivå i L2 svenska (forskningsfråga 1). Vidare utforskas om sena inlämare som producerar svenska klusiler med inföddlika VOT (forskningsfråga 2) kan identifieras och även huruvida de flesta av de tidigare inläarna uppskattar svenska VOT-värden (forskningsfråga 3). Eftersom det är allmänt känt att VOT påverkas av talhastighet (Beckman et al., 2011; Johnson & Wilson, 2002; Magloire & Green, 1999) belyses varje forskningsfråga genom jämförelser mellan absoluta VOT-värden (i millisekunder) och relativa VOT-värden (i procent av ordduration). Resultaten visade att startålder har en negativ inverkan på VOT även bland avancerade L2-inlärare men att denna ålderseffekt endast blev tydlig (dvs. statistiskt signifikant) när talhastigheten beaktades. Talhastighetens effekter gav även upphov till mindre förmågor gällande föreläningarnas/avskaknaden av inföddlig VOT-värden bland tidiga och sena L2-inlärare. Denna justering resulterade i signifikanta gruppseffekter där endast en liten minoritet av sena inlämare uppskattade inföddlig VOT samtidigt som många av de tidiga inläarna inte uppnådde en inföddlik behärskningsnivå. Resultaten följer samma mönster som har observerats i andra studier (Abrahamsson, 2012; Abrahamsson & Hyltenstam, 2009) och tyder på att en faktiskt inföddlik språkbehärskning hos både tidiga och sena L2-inlärare är mer sällsynt än vad som tidigare antogits. Detta utfall i kombination med data från noggranna och omfattande analyser av andra lingvistiska färdigheter hos samma avancerade L2-inlärare (Abrahamsson & Hyltenstam, 2009; se även Studie II) tyder dessutom på att sena inlämare med en faktisk inföddlig L2-behärskning är i det närmaste obefintliga. En allmän slutsats som dras är att teorier om mognadsbegränsningar och specifikt hypotesen om en kritisk period för språkinlärning inte kan förkastas utifrån dessa data. Vidare dras i Studie I slutsatsen att användningen av relativa (snarare än absoluta) VOT-mätningar är helt avgörande när det gäller att särskilja nästan infödda från faktiskt inföddliga slutnivåer i andraspråket.
Parallellt med Studie I utforskas i Studie II huruvida inlärningsålder påverkar den kategoriska perceptionen av tonkontrasten i svenska klusiler (/p–b/, /t–d/, /k–ɡ/) hos samma grupp av avancerade L2-inlärare. Samtliga VOT-stimuli till perceptionstestet har skapats genom manipulationer av naturligt producerade ord i svenska. Det antogs att L2-inlärarna skulle skilja sig från kontrollgruppen på så sätt att de sena L2-inlärarna förväntades upppvisa perceptuella kategoriseringar som avvikar mera markant från den svenska kontrollgruppen än de tidiga inlärarna. Med andra ord antogs den svenska kontrollgruppen upppvisa de högsta kategori-gränserna (dvs. mest positiva VOT) och de sena inlärarna de lägsta och därmed mest avvikande kategorigränserna, medan inlärarergruppen med låg startålder (SÅ 1–11) förväntades perceriera en övergång mellan klusilkategorierna någonstans mitt emellan kategorigränsvärdena hos de sena inlärarna och den infödda kontrollgruppen (hypotes 1). Vidare antogs att endast en minoritet av de sena inlärarna skulle upppvisa svensksliknande kategorigränser (hypotes 2a) medan de flesta av de tidiga inlärarna antogs visa perceptuella kategorigränser jämförbara med den svenska kontrollgruppens (hypotes 2b). På basis av tidigare antaganden om att den kritiska perioden torde ha störst inflytande på fonetiska/fonologiska aspekter (Scovel, 1988), förväntades åldersseffekterna framträdna ännu tydligare när perceptionsresultaten kombineades med produktionsdata från Studie I – en inföddlik behärskning av både produktion och perception förutspåddes bli närmast oförliklig bland de sena inlärarna (hypotes 3a) medan en majoritet av de tidiga L2-inlärarna antogs fortfarande upppvisa en inföddlik behärskningsnivå (hypotes 3b).

Resultaten visade att deltagarna, oberoende av deras status som L1/L2-talare och oavsett startålder, perciperade alla tre klusilkontraster på ett kategoriskt sätt. Detta bekräftar en hög reliabilitet och validitet hos studiedesignen, inklusive framställandet av stimuli och genomförandet av lyssningsexperimentet. Ålderseffekter observerades, vilket ger stöd till hypotes 1 samt hypotes 2a och 2b. Sammanlagningen av produktions- och perceptionsdata resulterade i en markant minskning av inföddlik behärskning. Medan inföddlik produktion och perception inte alls kunde konstateras bland de sena L2-inlärarna (hypotes 3a) upppvisade majoriteten av inlärrarna med låg startålder fortfarande en inföddlik behärskningsnivå på båda måtten (hypotes 3b). Dessa resultat ger stöd till tidigare konstateranden att efter tillämpandet av omfattande och detaljerade lingvistiska analyser är sena L2-inlärrare med en faktiskt inföddlik behärskningsnivå mycket svåra, om inte omöjliga, att identifiera (Abrahamsson & Hyltenstam, 2009). Följaktligen ifrågasätts tidigare upp- skattningar av inföddlighet bland sena andraspråksinlärare och parallellt med Studie I dras slutsatsen att hypotesen om en kritisk period och teorier om mognadsbegränsningar vid andraspråksinlärning inte kan förkastas mot bakgrund av de erhållna resultaten. Vidare visar denna studie att icke-inföddlik klusilproduktion och -perception även förekommer hos avancerade L2-inlärrare med mycket låga startåldrar vilket tyder på att en inföddlik behärskningsnivå i ett L2 inte kan garanteras enbart utifrån att andraspråktillägnandet påbörjas tidigt i livet. Dess-
utom beskrivs utfallet vara i linje med antagandet om att fonetiska/fonologiska aspekter som VOT verkar vara speciellt känsliga för ålderseffekter.


6 Eftersom sex av de tidigare deltagarna avstod från medverkan i attritionsprojektet baserar Studie III på 25 deltagare med låg attritionsålder (1–11 år) och 10 deltagare med hög attritionsålder (13–19 år).


References


Kuhl, P. K. (1987). The special mechanisms debate in speech research: Categorization tests on animals and infants. In S. Harnad (Ed.), *Categorical


International Review of Applied Linguistics in Language Teaching, 44, 251–263.

