

Age of acquisition – not bilingualism – is the primary determinant of less than nativelike L2 ultimate attainment

Emanuel Bylund^{1,2}, Kenneth Hyltenstam² and Niclas Abrahamsson²¹Stellenbosch University, South Africa and ²Centre for Research on Bilingualism, Stockholm University, Sweden

Research Article

Cite this article: Bylund E, Hyltenstam K, Abrahamsson N (2021). Age of acquisition – not bilingualism – is the primary determinant of less than nativelike L2 ultimate attainment. *Bilingualism: Language and Cognition* **24**, 18–30. <https://doi.org/10.1017/S1366728920000188>

Received: 20 June 2019
Revised: 11 February 2020
Accepted: 11 February 2020
First published online: 6 March 2020

Keywords:

critical period hypothesis; ultimate attainment; bilingualism effects; international adoptees; simultaneous bilingualism

Address for correspondence:

Niclas Abrahamsson,
E-mail: niclas.abrahamsson@biling.su.se

Abstract

It has recently been suggested that bilingualism, rather than age of acquisition, is what underlies less than nativelike attainment in childhood L2 acquisition. Currently, however, the empirical evidence in favor of or against this interpretation remains scarce. The present study sets out to fill this gap, implementing a novel factorial design in which the variables age of acquisition and bilingualism have been fully crossed. Eighty speakers of Swedish, who were either L1 monolinguals, L1 simultaneous bilinguals, L2 sequential monolinguals (international adoptees), or L2 sequential bilinguals (childhood immigrants), were tested on phonetic, grammatical, and lexical measures. The results indicate consistent effects of age of acquisition, but only limited effects of bilingualism, on ultimate attainment. These findings thus show that age of acquisition – not bilingualism – is the primary determinant of L2 ultimate attainment.

Introduction

A classic topic in the field of Second Language Acquisition (SLA), and the cognitive sciences at large, concerns the role of age of acquisition for nativelike attainment in a second language (L2). Since Lenneberg's (1967) formulation of the Critical Period Hypothesis (CPH), well over a hundred studies have sought to ferret out the effects that timing of exposure exerts on L2 acquisition, showing that those who start learning the L2 in childhood in the long run outperform those who start in adulthood. As classic a topic as age of acquisition effects is, it is also highly controversial, having instigated vigorous discussions throughout the decades. The debate has largely focused on the ultimate cause of age effects – that is, whether they are biological, experiential, socio-psychological, cognitive, etc. in nature – rather than on their actual existence.

Recently, however, the finding that individuals who acquired the L2 during childhood do not always converge fully with native speakers has called into question age of acquisition as the cause of such *near-native* (rather than fully nativelike) attainment. As an alternative explanation, it has been suggested that, rather than age of acquisition, bilingualism – in the sense of either bilingual acquisition, bilingual use, or both – accounts for the subtle non-native features in early-learner ultimate attainment, and, by inference, also the near-nativeness of exceptionally advanced adult L2 learners (e.g., Birdsong, 2018; Birdsong & Quinto-Pozos, 2018; de Leeuw, 2014; Ortega, 2010, 2013; Pfenninger & Singleton, 2017). This suggestion relates to the fact that most studies on nativelike attainment compare L2 speakers who have retained their first language (L1), and therefore are functionally bilingual, with native speakers who are functionally monolingual, thus effectively confounding age of acquisition effects with bilingualism effects.

The methodological practice of comparing bilingual L2 speakers with monolingual L1 speakers becomes particularly problematic in the light of frameworks suggesting that the linguistic behavior of bilinguals inherently differs from that of monolinguals (e.g., Cook, 1999, 2016; Flege, 1999; Grosjean, 1998), as this may ultimately render any observations on age effects inconclusive. However, despite various iterations of the notion of *bilingualism effects* on L2 ultimate attainment, few studies have actually attempted to address this question empirically. Thus, while it is indeed an intriguing possibility that bilingualism, rather than age of acquisition, underlies the subtle non-nativeness of many childhood (as well as exceptionally advanced adult) learners, this suggestion largely remains at the level of speculation due to the absence of solid empirical data.

The current study aims to address this gap, by assessing the relative impact of age of acquisition and bilingualism on L2 ultimate attainment. To achieve this, the study introduces a unique experimental design, which, in addition to an L2 bilingual group and an L1 functionally monolingual group, includes simultaneous bilinguals and international adoptees. In this design, the variables *age of acquisition at birth/after birth* vs. *mono-/bilingualism* are fully crossed.

© The Author(s), 2020. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background

The nativelikeness paradigm in CPH research

The notion that biologically scheduled changes in brain plasticity underlie child-adult differences in L2 ultimate attainment would seem to find support in research showing that non-maturational variables such as length of L2 exposure, educational level, and motivation, while important in (especially adult) L2 acquisition, only exert marginal impact compared to age of acquisition (AoA). Indeed, studies using partial correlations or regression analyses have repeatedly shown that the contributions of experiential and socio-psychological variables drop considerably (often to non-significant levels) when the AoA variable is partialled out, whereas the impact of AoA remains strong and relatively unaffected when the contributions from these variables are removed (e.g., Abrahamsson, 2012; DeKeyser, 2000; DeKeyser, Alfi-Shabtay & Ravid, 2010; DeKeyser & Larson-Hall, 2005; Granena & Long, 2013; Johnson & Newport, 1989). To this end, then, the maturation of the brain would still seem a strong explanatory candidate for AoA effects. However, despite some promising explanatory frameworks, such as the scheduled process of myelination of language-related cortical areas (e.g., Pulvermüller & Schumann, 1994) or the age-related switch from (predominantly) implicit/procedural memory to (predominantly) explicit/declarative memory in language development (e.g., Paradis, 2004, 2009; Ullman, 2004, 2015), any operationalizable neurophysiological correlates to maturation that can be closely associated with AoA are still lacking. Moreover, the AoA variable may well disguise the effects of any as yet unmeasurable (or hitherto ill-measured) non-maturational factor(s) – adult speakers' retrospectively self-assessed motivation for acquiring the L2 in childhood being just one of many examples. With the substance of the AoA variable still shrouded in darkness, the correlational approach in CPH research thus finds itself in an unfortunate deadlock.

Therefore, an alternative way of addressing the impact of maturational constraints has been to look exclusively for individual counterexamples to the hypothesis that only child learners are capable of attaining nativelike L2 proficiency and behavior. We refer to this approach as the 'nativelikeness paradigm'. The Popperian rationale behind the approach, as originally presented in detail by Long (Long, 1990, 1993, see also Long, 2007, 2013), is that, if at least *one* such individual post-critical period learner could be identified who, even after broad and detailed scrutiny, can be shown to exhibit the same linguistic knowledge and behavior as native speakers, then the CPH can be safely rejected, and the well-documented average adult disadvantage should instead be ascribed to factors other than neurobiology.

Long (1990, 1993) moreover recommended that researchers should use only linguistic tasks and structures that highly advanced learners potentially do *not* command; that the level of cognitive demand, item difficulty, and linguistic scrutiny in nativelikeness studies should be significantly higher than in studies of beginner or intermediate L2 proficiencies; and that a broad range of language abilities (rather than narrowly selected linguistic features of a limited language domain) should be scrutinized in these learners' ultimate attainment (for similar arguments and elaborations of this last point, see, e.g., Abrahamsson & Hyltenstam, 2009; DeKeyser, 2012; Granena & Long, 2013; Sorace & Robertson, 2001; Verissimo, 2018; Verissimo, Heyer, Jacob & Clahsen, 2018).

A previous project from the Stockholm lab (reported in Abrahamsson & Hyltenstam, 2008, 2009; see also Bylund, 2011;

Bylund, Abrahamsson & Hyltenstam, 2010, 2012; Hyltenstam, Bylund, Abrahamsson & Park, 2009; Stölten, Abrahamsson & Hyltenstam, 2014, 2015) aimed to follow Long's (1990, 1993) recommendations as closely as possible. The focus was set exclusively on L2 speakers who passed for native speakers in everyday oral interaction, the rationale being that there is no point in subjecting obviously non-nativelike speakers to extensive linguistic scrutiny just to declare them non-nativelike. A total of 195 candidates, who self-reported as potentially nativelike L2 speakers of Swedish (AoA 1–47 y/o), were first screened through naïve native listener judgments of their spontaneous speech. Out of these, 41 speakers were eventually selected, all of whom were perceived as native speakers by a majority of the judges (minimally 6 out of 10), and were subjected to detailed linguistic scrutiny through a challenging test battery. Thirty-one of these were early learners (AoA 1–11 y/o), and ten were late learners (AoA 13–19 y/o).

The results revealed that *every* late (seemingly nativelike) learner, and many of the early learners, were in fact *near-native* (as opposed to nativelike) when scrutinized in detail. For example, when the production *and* the categorical perception of voice onset time (VOT) were combined, for all three (i.e., bilabial, dental, and velar) places of articulation, as predicted, none of the 10 late learners fell within the native-speaker range, while, at the same time, only 16 of the 31 early learners did so (see Stölten et al., 2014). When these same learners' performance on 10 different accuracy and processing measures within various domains and modes of their L2 Swedish (phonology, morphosyntax, lexis, perception through different types of noise, etc.) was analyzed, the pattern was even clearer: again, none of the adult learners, and only a handful of the early learners, performed within the range of native-speaker controls on a majority of the measures (for details, see Abrahamsson & Hyltenstam, 2009; for similar patterns in other advanced-learner samples, see Abrahamsson, 2012; Hyltenstam, 1992; Hyltenstam & Abrahamsson, 2003a; Hyltenstam et al., 2009).¹

The findings reported in Abrahamsson and Hyltenstam (2009) were then taken as evidence that even short delays of language exposure may have minor but scientifically detectable consequences for L2 ultimate attainment, potentially indicative of the brain's decreasing capacity for nativelike language acquisition already at early AoAs. Such a conclusion is on par with accounts of atypical L1 development where small delays in L1 exposure compromise ultimate attainment, as seen both in congenitally deaf children with delayed sign-language exposure (see, e.g., Mayberry & Kluender, 2018; Morford & Mayberry, 2000) and in children with severe *otitis media* during their first year of life (e.g., Mody, Schwartz, Gravel & Ruben, 1999; Ruben, 1999) (for an overview, see Werker & Hensch, 2015).

That brain maturation is a potential cause of childhood learners' less than nativelike L2 ultimate attainment, is not, however, an interpretation that has been embraced by everyone. Instead, results such as those above have been re-interpreted by several scholars as evidence that *bilingualism*, not maturation, is what lies behind the less than nativelike ultimate attainment of both early and exceptionally advanced late learners. This argument will be reviewed next.

¹Marinova-Todd (2003) presents data on nativelike attainment across a range of measures in adult L2 learners. This study has, however, been criticized for, among other things, uncertainties regarding the actual age of acquisition of the participants and the level of difficulty of the test instruments (for further critique, see Hyltenstam & Abrahamsson, 2003b; Long, 2007).

Monolingual bias, bilingualism effects, and the 'bi/multilingual turn' in CPH research

The status of the L2 learner's L1, and the role of cross-linguistic influence generally, has fluctuated considerably over time in SLA theory building. From having been given an absolute role under the behaviorist (pre-modern SLA) era, via a next to negligent role during the first decades of interlanguage theory development and (mainly) nativist SLA, learners' L1 and their bilingualism at large have been gradually resurrected as central components in recent (notably, connectionist/emergentist) SLA theorizing. Several modern-day cognitivist theorists would argue that the successive, age-related entrenchment of the L1 and/or the active use of two languages are the major reasons why nativelike L2 competence and behavior are not attained (e.g., Flege, 1999; Herschensohn, 2007; Pallier, 2007; Vanhove, 2013). Accordingly, the theoretical account currently gaining interpretative prerogative in the CPH debate holds that less than nativelike ultimate attainment is to be expected even in very advanced (be it early or late) L2 learners, simply because "nonmonolingual-likeness in terms of proficiency /.../ is a defining characteristic of bilingualism" (Birdsong, 2014, p. 377). In line with Grosjean's (1989) statement that the bilingual is not two monolinguals in one person, various theoretical approaches to SLA, such as the Multicompetence framework (e.g., Cook, 1991, 2003, 2016), the Competition Model (e.g., MacWhinney, 1999, 2016), the Speech Learning Model (e.g., Flege, 1999), and the Interference Hypothesis (Pallier, Dehaene, Poline, LeBihan, Argenti, Dupoux & Mehler, 2003; Ventureyra, Pallier & Yoo, 2004), all point to the inherent difference between monolingual competence and the unique linguistic competence that emerges from the existence of two language systems in one mind (for a contrasting view, see e.g., Meisel, 2008, 2017; also Montrul & Ionin, 2010).

Consequently, in view of this reasoning, various reinterpretations have been suggested for the results of the Abrahamsson and Hyltenstam (2009) study, along with general criticisms of Long's nativelikeness paradigm. On this latter point, Birdsong (2018) argues that, because of "coactivation and bidirectional effects, neither the first nor the second language of bilinguals can be expected to resemble under scrutiny that of monolinguals in either language" (p. 6), thus making it "unreasonable to hold up a standard of 'across-the-board monolingual nativelikeness' in the L2 as a criterion for falsifying the CPH" (ibid.) (see also Birdsong, 2005, 2006; Birdsong & Gertken, 2013; Birdsong & Quinto-Pozos, 2018). In a similar fashion, Vanhove (2013) holds that "the linguistic repertoires of mono- and bilinguals differ by definition and differences in the behavioural outcome will necessarily be found, if only one digs deep enough" (p. 2), and he warns us against raising the bar for highly accomplished L2 learners "to Swiftian extremes" (ibid.).² Consequently, and in line with what has been launched as "the bi/multilingual turn in SLA" (Ortega, 2010, 2013), the very comparison with monolingual speakers has been deemed theoretically misguided and it has been recommended that it should be abandoned in CPH (or, even, all SLA) research; since 'nativelike' is considered synonymous with 'monolingual-like', the expected maximal

'bilingual-like' ultimate attainment should be equivalent to what has hitherto been (mis)taken for 'near-native' proficiency, regardless of learners' AoA. Accordingly, it has been suggested by several authors (e.g., Birdsong, 2005, 2018; de Leeuw, 2014; Ortega, 2010, 2013; Cook, 1999, 2003, 2016; Muñoz & Singleton, 2011) that the comparative standard should be shifted from monolingual language proficiency to the simultaneously acquired bilingual ultimate attainment of 'crib bilinguals'. For example, de Leeuw (2014) sees the conclusions in Abrahamsson and Hyltenstam (2009) as premature, as the study potentially suffered from monolingual-speaker bias. According to her, the inclusion of an additional participant group, consisting of simultaneous bilinguals who acquired two languages from birth, would have been necessary, because

only if the /.../ simultaneous bilinguals performed according to monolingual proficiency levels, whereas the /.../ non-native speakers /.../ did not, would it have been possible to ascertain that biologically determined maturational constraints impede L2 acquisition. If, on the other hand, both simultaneous and late consecutive bilinguals performed deviantly to monolingual norms, an alternative explanation would be required. (de Leeuw, 2014: 35)

That bilingualism, rather than brain maturation, might be the best candidate for explaining any subtle differences between native and near-native ultimate attainment is indeed a theoretically intriguing hypothesis that, in our view, merits thorough empirical testing. When considering the past decades' explosion of research suggesting that bilingualism brings about cognitive advantages (in terms of divergent thinking, enhanced executive control, delayed symptoms of dementia, etc.), as well as linguistic costs (particularly in terms of a so-called bilingual lexical deficit; for overviews, see, e.g., Bialystok, 2009, 2016, 2017), the hypothesis seems well-motivated. However, the widespread reliance on this research is actually what constitutes the core problem of the current CPH debate, as the bilingualism-effects argument largely rests on indirect inferencing from non-CPH/non-ultimate attainment research. For example, when Singleton and Pfenninger (2018) assume that "[t]he reason for the slight differences between native speakers and native-like non-natives /.../ almost certainly has to do with the effects of multi-competence /.../ rather than age" (p. 260; emphasis added), they are certainly not the only ones to engage in guesswork based on research that set out to investigate something other than the relative roles of AoA and bilingualism for ultimate attainment. This is clearly problematic for a number of reasons.

To begin with, it should be noted that the bilingual cognitive advantage has been seriously challenged, both in a comprehensive meta-analysis (Lehtonen, Soveri, Laine, Järvenpää, de Bruin & Antfolk, 2018), and in a recent large-scale study (Dick, Garcia, Pruden, Thompson, Hawes, Sutherland, Riedel, Laird & Gonzalez, 2019), showing that there is no robust evidence of enhanced executive functioning in bilinguals. Secondly, and more importantly for the current argument, the majority of studies claiming to show a lexical deficit in bilinguals have actually ignored the AoA dimension or disregarded the crucial distinction between simultaneous and sequential bilingualism. Because of this, it is notoriously difficult to tell whether the lexical behavior attested in those bilingual samples is an artefact of bilingualism or L2 status. Indeed, a recent study showed that when AoA is taken carefully into account, the alleged bilingual lexical deficit turns out to predominantly be an L2 effect (Bylund, Abrahamsson, Hyltenstam & Norrman, 2019). Taken

²Cook (2016) even considers the comparison with monolinguals in Abrahamsson and Hyltenstam (2009) to be borderline racist – an accusation too outrageous to merit commenting on in this article.

together, these findings seriously undermine several assumptions on which arguments of bilingualism effects rest.

Moreover, when ultimate attainment studies have indeed included simultaneous bilinguals, results do not necessarily indicate consistent differences in proficiency or neurophysiology between monolinguals and simultaneous bilinguals (e.g., Berken, Gracco & Klein, 2017; Klein, Mok, Chen & Watkins, 2014; Reetzke, Lam, Xie, Sheng & Chandrasekaran, 2016; Verissimo et al., 2018). In those instances where different proficiency scores are indeed documented between monolinguals and simultaneous bilinguals (e.g., Hartshorne, Tenenbaum & Pinker, 2018; Sundara, Polka & Baum, 2006), it is unclear whether the language under scrutiny was the participants' dominant or non-dominant language, and whether it was the majority language or a heritage language – both of which are absolutely crucial factors to control for when performing group comparisons with monolingual majority-language speakers.

A logical extension of the bilingualism-effects argument is that the less L1 knowledge there is (and consequently, the lower the L2 speaker's degree of bilingualism), the greater the possibility of attaining nativelike/monolingual-like L2 proficiency – the extreme situation of *total L1 loss* offering the likeliest prospect for such attainment. This reasoning is captured in the Interference Hypothesis, the empirical basis of which is a series of studies (Pallier et al., 2003; Ventureyra et al., 2004; see also Pallier, 2007) showing that international adoptees seem to have completely forgotten their childhood L1, as evidenced both through behavioral tests as well as fMRI responses (however, for counter-evidence, see Choi, Cutler & Broersma, 2017b; Park, 2014; Pierce, Klein, Chen, Delcenserie & Genesee, 2014; Singh, Liederman, Mierzejewski & Barnes, 2011), while at the same time having attained (allegedly) nativelike proficiency in the L2. The conclusions drawn by these researchers were that if the L1 is lost at some point in childhood, the neural network can “reset” (Ventureyra et al., 2004: 89), which will allow for monolingual acquisition and a nativelike ultimate attainment. Conversely, the reason why some childhood L2 learners who maintain their L1 (such as immigrant children) do not attain nativelike L2 proficiency is because their L1 “acts as a filter that distorts the way in which a second language can be acquired” (Pallier et al., 2003: 160).

The problem, however, is that these studies performed no systematic linguistic assessment of the adoptees' L2 proficiency. Instead, the claim about L2 nativelikeness was based on the test administrators' impressions of the adoptees' L2 speech. Subsequent studies examining the L2 of international adoptees with proper experiments have instead found that this group exhibits the same levels of (non-nativelike) proficiency as L2 speakers who have retained their L1 (Gauthier & Genesee, 2011; Hyltenstam et al., 2009; Norrman & Bylund, 2016, see also Gauthier, Genesee & Kasparian, 2012; Pierce, Chen, Delcenserie, Genesee & Klein, 2015). Moreover, several studies have shown that international adoptees often display L1 remnants (e.g., Choi, Broersma & Cutler, 2017a; Choi et al., 2017b; Hyltenstam et al., 2009; Park, 2014; Pierce et al., 2014; Singh et al., 2011). Yet, the ideas of complete L1 loss, ‘neural resetting’, and monolingualism as prerequisites for nativelike L2 acquisition seem to be considered to be facts in the CPH debate.

Aims, design, and hypotheses of the present study

Given that the current debate on AoA and bilingualism effects in L2 acquisition is characterized by a scarcity of hard evidence, the

Table 1. The 2 × 2 factorial design of the present study, with AoA at birth/after birth vs. monolingualism/bilingualism fully crossed. (Gray cells A and D = the traditional and allegedly monolingually biased group comparison in CPH studies; white cells B and C = novel, potentially informative, but as yet largely unengaged, control groups in CPH research.)

	AoA at birth (AoA = 0 y/o)	AoA after birth (AoA = 3–8 y/o)
Monolingualism	A. Monolingual L1 speakers of Swedish (‘crib monolinguals’)	B. Sequentially monolingual L2 speakers of Swedish (‘childhood adoptees’)
Bilingualism	C. Simultaneously bilingual L1 speakers of Swedish (‘crib bilinguals’)	D. Sequentially bilingual L2 speakers of Swedish (‘childhood immigrants’)

current study set out to empirically assess the relative impact of AoA vs. bilingualism on ultimate attainment in seemingly native-like L2 speakers. To do so, we introduced a novel methodological design, in which the issue of monolingual-likeness is addressed through the addition of, first, simultaneous bilinguals as a comparison group, as advised by proponents of the bi/multilingual turn in SLA, and second, sequential monolingual L2 learners (here in the form of adult L2 speakers who were internationally adopted in early childhood), as per the notion that L1 loss increases the likelihood of nativelike L2 attainment. This yielded a 2(AoA from birth vs. after birth) × 2(monolingualism vs. bilingualism) factorial design (see Table 1), for which the following two alternative hypotheses were postulated:

1. The AoA-effects hypothesis

‘Nativelikeness’ is made possible by language exposure beginning at birth; ‘non-nativelikeness’ is the result of language exposure beginning later than birth. This hypothesis predicts a stand-alone main effect of AoA.

2. The bilingualism-effects hypothesis

‘Nativelikeness’ is made possible by monolingual language acquisition and use (and should instead be labeled ‘monolingual-likeness’); ‘non-nativelikeness’ is the result of bilingual language acquisition and use (and should instead be called ‘bilingual-likeness’). In the current design, this hypothesis would be confirmed in a stand-alone main effect of bilingualism.

In addition to these potential outcomes, alternative results may also be attested, manifested as an interaction between, or a confluence of, bilingualism and AoA.

Method

Participants

The following 80 participants took part in the study.

Monolingual L1 speakers of Swedish (n = 20)

The speakers in this group ($M_{\text{age}} = 29.8$) were ‘crib monolinguals’. They were born in Sweden to L1 Swedish parents, and had acquired Swedish from birth as their only language. They had grown up in Sweden, and used Swedish in their everyday lives for communicative purposes. These participants were recruited

Table 2. Participants' background characteristics ($n = 80$).

	L1 monolinguals ($n = 20$)	L1 bilinguals ($n = 20$)	L2 monolinguals ($n = 20$)	L2 bilinguals ($n = 20$)	Distribution	p -value
Age (M years)	29.8 (5.5)	28.8 (6.7)	33.8 (6.0)	32.2 (4.7)	$F = 3.06$	0.033/0.108 n.s. ^a
AoA (M years)	n.a.	n.a.	4.3 (1.4)	5.2 (1.8)	$t = 1.68$	0.108 n.s.
LoR (M years)	29.8 (5.6)	28.8 (6.7)	28.9 (5.7)	26.7 (5.3)	$F = 1.019$	0.389 n.s.
Education (# low)	12	12	15	12	$\chi^2 = 1.460$	0.782 n.s.
Sex (# males)	9	9	7	5	$\chi^2 = 2.347$	0.503 n.s.
Swedish use (M %)	n.a.	78.2 (15.4)	n.a.	71.0 (16.2)	$t = 1.451$	0.155 n.s.
Spanish proficiency (M % cloze test score)	n.a.	66.4 (18.5)	n.a.	65.7 (17.5)	$t = 0.20$	0.840 n.s.

Note. Standard deviations are reported in brackets.

^aThe difference in current age between the L1 bilinguals and L2 monolinguals was not significant after Bonferroni alpha correction.

through word of mouth and flyers distributed throughout Stockholm.

Simultaneously bilingual L1 speakers of Swedish and Spanish ($n = 20$)

These speakers ($M_{\text{age}} = 32.2$) were 'crib bilinguals'. They were born in Sweden to one Swedish-speaking parent and one Spanish-speaking parent. They had acquired both Swedish and Spanish from birth, and used both languages for everyday communication. These participants were recruited through newspaper advertisements.

Sequentially monolingual L2 speakers of Swedish ($n = 20$)

This group ($M_{\text{age}} = 33.7$) comprised childhood adoptees who were born in Spanish-speaking countries in Latin America and adopted to Sweden between 3 and 7 years of age ($M_{\text{age of arrival}} = 4.3$). According to self-reports, they had lost proficiency altogether in their L1 Spanish shortly after adoption, and had not engaged in relearning activities. In Sweden, they were brought up in L1 Swedish-speaking families, and consequently acquired Swedish as an L2. They used only Swedish for everyday communicative purposes. These participants were recruited through newspaper advertisements, adoption associations, adoption agencies, and social media.

Sequentially bilingual speakers of L1 Spanish and L2 Swedish ($n = 20$)

The participants in this group ($M_{\text{age}} = 28.8$) were born in Latin American countries to L1 Spanish-speaking parents and thus had acquired Spanish from birth. Together with their families, they immigrated to Sweden between the ages of 3 and 8 years ($M_{\text{age of arrival}} = 5.2$), which was when their acquisition of L2 Swedish commenced. These individuals had continued using their Spanish since arrival, and reported using both Spanish and Swedish in their everyday lives. These participants were recruited through newspaper advertisements.

As seen in Table 2, the L2-speaker groups (i.e., the 'childhood adoptees' and the 'childhood immigrants') did not differ significantly in terms of age of L2 acquisition. The bilingual groups (i.e., the 'crib bilinguals' and the 'childhood immigrants') did not differ in terms of Spanish language knowledge, as measured by their performance on a Spanish cloze test (Bylund et al., 2019), or in their everyday use of Spanish and Swedish. Through schooling in Sweden, all participants had acquired

foreign language skills in English and at least one other modern language, such as French or German. All participants spoke Swedish without any noticeable phonological, grammatical, or lexical deviations, as impressionistically judged by a linguistically trained, Swedish native-speaker research assistant. Groups were also matched in terms of education and gender.

Materials and procedure

Data was elicited on speech production and perception, morpho-syntax (accuracy and response latencies), and formulaic language, thus covering a fairly broad range of language competence and processing abilities. The linguistic instruments were identical to 7 of the 10 instruments used by Abrahamsson and Hyltenstam (2009).³

Instruments 1 and 2: production and perception of voice onset time (VOT)

The time interval between the release of a stop consonant and the onset of periodicity of the following vowel is generally referred to as *voice onset time*, or VOT. Spanish and Swedish differ as to where on the voicing continuum the voiced/voiceless categories separate: Spanish category boundaries are located at low, usually negative VOT values, whereas Swedish boundaries are found at higher, usually positive values (see, e.g., Lisker & Abramson, 1964).

In the production task (Instrument 1), the participants' reading aloud of the Swedish words *par* ('pair'), *tal* ('number'), and *kal* ('naked') was recorded. Each word was read in isolation 10 times, yielding a total of 2,398 data points (3 words \times 10 readings \times 80 participants - 2 unmeasurable tokens). Spectral analyses of the VOT of /p/, /t/, and /k/ were made in Praat (Boersma, 2002), measuring the time interval between the onset burst of the stop and the onset of vowel periodicity. Because VOT duration varies as a function of speech rate (e.g., Johnson & Wilson, 2002; Schmidt & Flege, 1996; Volaitis & Miller, 1992), VOT values in milliseconds were converted into relative VOT values, calculated as percentages of word duration (for further detail on such a procedure, see, e.g., Stölten et al., 2015). Word duration was operationalized as the interval spanning from the onset of the release burst

³The reason for not including all 10 instruments from Abrahamsson and Hyltenstam (2009) was to save time and to give room for some new phonological and lexical tests (not reported here; see instead Bylund, et al., 2019; Norrman & Bylund, 2016).

to the end of the periodicity of the final /l/ or /r/. The production task took approximately 5 minutes to complete.

The categorical perception test (Instrument 2) included the minimal pairs *par-bar* ('pair'-'bar'), *tal-dal* ('number'-'valley'), and *kal-gal* ('naked'-'crow(s)' (V_{pres})). Each word had been recorded in an anechoic chamber by a native female speaker of Swedish, and a 5-msec-step VOT continuum ranging from -60 to +90 msec was then created for all word pairs (for details on the preparation of stimuli, see Stölten et al., 2014). The stimulus items were presented in E-Prime v.2.0 (Psychology Software Tools, Inc.; Schneider, Eschman & Zuccolotto, 2002a, 2002b) through PC-350 headphones in different randomized orders for all participants. Each word was preceded by the carrier phrase *Nu hör du...* ('Now you will hear...'), and the participants' task was to indicate by pressing one of two buttons whether they heard a word beginning with a voiceless stop, /p, t, k/, or a voiced stop, /b, d, g/. The perception test took approximately 5 minutes to complete.

Instruments 3 and 4: grammaticality judgment accuracy and latency

Morphosyntactic knowledge and processing ability was measured through a comprehensive and demanding, auditory grammaticality judgment test. The test consisted of 80 sentences representing four morphosyntactic features of Swedish: (1) subject-verb inversion (V2); (2) reflexive possessive pronouns; (3) placement of sentence adverbs in relative clauses; and (4) gender and number agreement. Half of the sentences were grammatically incorrect, containing one grammatical error each. All sentences contained subordinate clauses, so as to increase syntactic processing demands. The sentences had been recorded in an anechoic chamber by a female native speaker and were presented in E-Prime through PC-350 headphones in different random orders for all participants. The participants indicated whether they perceived each sentence as grammatically correct or incorrect by pressing a green or a red (respectively) button at any point during or after the sentence presentation. Along with accuracy (Instrument 3), response latencies were also recorded (Instrument 4). The test took 15–20 minutes to complete.

Instrument 5: grammatical, lexical, and semantic inferencing

A global measure of L2 Swedish proficiency was obtained through a cloze test. The cloze test technique (Taylor, 1953) mobilizes a speaker's grammatical, lexical, contextual, and pragmatic knowledge in the perception and comprehension of spoken and written language. The present test was an untimed pen-and-paper task consisting of a 300-word text where every seventh word had been replaced by a blank. The task was to fill in each of the 42 blanks with a word that would fit into the context, structurally and semantically. Responses other than those in the original text were evaluated for lexical, grammatical, and semantic appropriateness with respect to their linguistic context; encyclopedic errors or spelling errors were *not* scored as errors. The test took 15–20 minutes to complete.

Instruments 6 and 7: formulaic language

Even though L2 learners (as well as L1 learners) rely on prefabricated linguistic chunks in early language development, the idiomatic use of formulaic language has been shown to be one of the greatest difficulties for (even very advanced) L2 speakers (e.g., Erman, Forsberg Lundell & Lewis, 2018; Foster, Bolibaugh & Kotula, 2014; Granena & Long, 2013; Wray, 2005). The present

study included one test of idioms (Instrument 6) and one test of proverbs (Instrument 7). Both tests were created and run in E-Prime; they were identical in design and procedure, and included 50 items each presented on a screen (one at a time and in the same order for all participants) with a blank that was to be filled in with a missing word or phrase. Participants were given 10 seconds to complete each item, and their oral responses were recorded and later analyzed. Responses that did not correspond to the standard formulaic expression or any established variant thereof were scored as erroneous.⁴ The tests took each 7–8 minutes to complete.

Testing and data collection was performed by a male native speaker of Swedish in a sound-attenuated room individually with each participant. Normal hearing was confirmed with an OSCILLA SM910 screening audiometer, and the entire language testing session (including instructions and breaks) then lasted for approximately 2.5 hours. Participants received a remuneration of SEK 500 (approximately €50).

Statistical analyses

In the current study, AoA is defined as a categorical variable, that is, acquisition from birth (i.e., L1) versus additional language acquisition (i.e., L2), commencing in this case between 3 and 8 years of age. The study is, in other words, not designed to assess AoA as a continuous variable, because the AoA range is too narrow and only covers early childhood (a period during which pronounced differences in AoA effects are typically not attested).

Performance on the grammaticality judgement test (accuracy), the cloze test, the idioms test, and the proverbs tests was analyzed using logit mixed model regressions with response accuracy as dependent variable. AoA (i.e., at birth vs. at 3–8 years of age) and bilingualism (monolingualism vs. bilingualism) were entered as categorical fixed effects, sum coded as -1 and 1. Subject and item were added as random effects, and bilingualism, AoA, and their interaction were added as random slopes, as justified by the maximal structure that converged.

Linear mixed model regressions were conducted to analyze the performance on the VOT production test and the reaction times on the grammaticality judgment test. Again, AoA and bilingualism were entered as categorical fixed effects. Subject and item were added as random effects, and bilingualism, AoA, and their interaction were added as random slopes, as justified by the maximal model that converged. All mixed model regressions were carried out using the Lme4 package (Bates, Mächler, Bolker & Walker, 2014) in R (R Core Team, 2014).

Categorical perception of VOT was analyzed using probit (Finney, 1947), which generates estimates of the 50% crossover points of binary response curves using maximum likelihood estimation (see also Caramazza, Yeni-Komshian, Zurif & Carbone, 1973; Hazan & Boulakia, 1993). The generated probit values (one per place of articulation for each participant) were entered as dependent variable into linear models, with AoA and bilingualism as fixed effects (because there was only one data point, i.e., the probit value, per participant per articulation, random effects and random slopes were not computable).

⁴For example, the final part of the proverb *Det är inte ens fel när två träter* 'It is not one's fault when two quarrel' ('It takes two to quarrel') can also be expressed as *att två träter* 'that two quarrel' and was therefore considered as correct; however, it cannot be expressed as *när två grälar* 'when two argue', which was therefore considered an incorrect response.

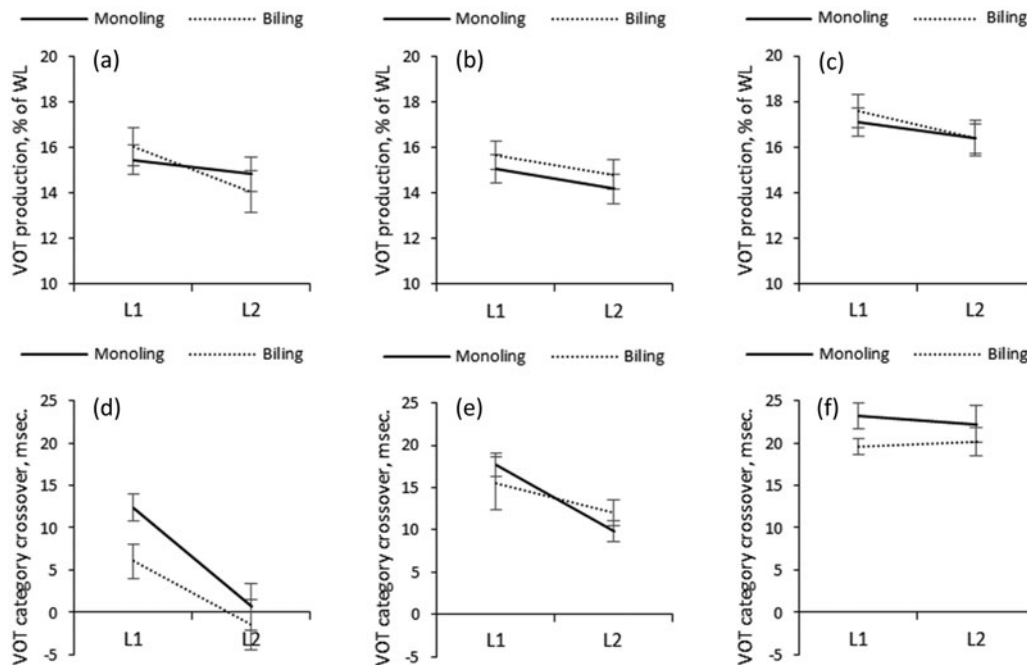


Fig. 1. (a) Mean VOT productions of /p/; (b) mean VOT productions of /t/; (c) mean VOT productions of /k/; (d) mean VOT perception of /p-b/; (e) mean VOT perception of /t-d/; and (f) mean VOT perception of /k-g/. VOT values plotted as a function of AoA at/after birth (i.e., L1/L2) and mono-/bilingualism. Error bars represent standard error of the mean. WL = Word length.

Results

Production and perception of VOT

Starting with production, the elicited VOT values of /p/ showed no significant effects of either AoA nor bilingualism, $\beta = 0.035$, $SE = 0.029$, $t = 1.192$, $p = .237$, and $\beta = -0.016$, $SE = 0.029$, $t = -0.565$, $p = .574$, respectively, and no interaction, $p = .913$. A similar result was obtained for /t/: AoA, $\beta = 0.036$, $SE = 0.025$, $t = 1.431$, $p = .157$; bilingualism, $\beta = -0.006$, $SE = 0.025$, $t = -0.255$, $p = .799$; and non-significant interaction, $p = .739$. Likewise, no significant main effects or interactions were found for /k/: AoA, $\beta = 0.024$, $SE = 0.031$, $t = 0.794$, $p = .430$; bilingualism, $\beta = -0.020$, $SE = 0.031$, $t = -0.650$, $p = .518$; and non-significant interaction, $p = .915$. All groups were thus found to produce stop intervals of similar proportions between stop release and vowel periodicity onset. These results are depicted in Figure 1a–c.

However, in terms of categorical perception, a significant main effect of AoA was found for bilabial stops, $\beta = 4.801$, $SE = 1.200$, $p < .001$, suggesting that speakers with AoA at birth were more likely to place the category boundary of /p/-/b/ towards the positive end of the voicing continuum than speakers with later AoA (see Figure 1d). This was further confirmed in planned pairwise comparisons revealing significant differences in the same direction between monolingual L1 and L2 speakers ($p < .01$) and between bilingual L1 and L2 speakers ($p = .04$). No statistically significant main effect was noted for bilingualism, $\beta = 2.112$, $SE = 1.200$, $p = .082$, nor was there any significant interaction between the predictor variables, $\beta = 1.047$, $SE = 1.200$, $p = .385$. For dental stops (Figure 1e), AoA was again found to exert a main effect on category boundary, $\beta = 2.082$, $SE = 0.663$, $p = 0.002$, with L2 speakers being prone towards negative /t/-/d/ boundaries. No main effect of bilingualism was detected, $\beta = 0.742$, $SE = 0.663$, $p = .267$, but a significant interaction between

bilingualism and AoA was found, $\beta = 1.809$, $SE = 0.633$, $p = .008$. Bonferroni post hoc tests revealed that L1 monolinguals and L1 bilinguals differed from one another ($p = .029$), but not L2 monolinguals and L2 bilinguals ($p = 1$). Likewise, monolingual L1 speakers differed significantly from monolingual L2 speakers ($p = .001$), whereas bilingual L1 speakers did not differ from bilingual L2 speakers ($p = 1$). Lastly, for velar /k/-/g/ crossover, there was no significant main effect for AoA, $\beta = 0.106$, $SE = 0.823$, $p = .898$, or for bilingualism, $\beta = 1.436$, $SE = 0.823$, $p = .09$, nor any interaction effect, $\beta = 0.404$, $SE = 0.823$, $p = .625$ (Figure 1f).⁵

Grammaticality judgment accuracy and latency

Participants' performance on correctly judging the grammaticality of Swedish sentences revealed a significant main effect of AoA, $\beta = -0.538$, $SE = 0.092$, $p < .001$, but no main effect of bilingualism, $\beta = 0.011$, $SE = 0.092$, $p = .906$. However, there was also a marginally significant interaction between these two variables, $\beta = -0.741$, $SE = 0.369$, $p = .055$. As a follow-up, a series of post-hoc test (Bonferroni) was conducted. These showed no significant difference within the L1 groups (L1 monolinguals vs. L1 bilinguals, $p = 1$) nor within the L2 groups (L2 monolinguals vs. L2

⁵At the request of a reviewer, we also analysed AoA as a continuous variable in relation to test performance in the two L2 groups (i.e., sequential bilinguals and international adoptees). Because the study was not designed to this end (the AoA range being too narrow and only covering early childhood), we did not expect any significant effects. Logit mixed model regressions with continuous AoA as fixed effect indeed turned out non-significant for all measures ($ps \geq .15$). To further probe whether there was any AoA-related variation within the L2 groups, the L2 speakers were divided into two groups, with $AoA \leq 5$ and $AoA \geq 6$, respectively. Here, comparisons (logit mixed model regressions) yielded marginally significant differences or weak trends for categorical perception of velar stops ($p = .093$), GJT scores ($p = .063$), idiom scores ($p = .061$) and proverb scores ($p = .073$), with the later AoA group always obtaining less nativelike scores than the earlier AoA group. For the rest of the measures, no differences were documented ($ps \geq .20$).

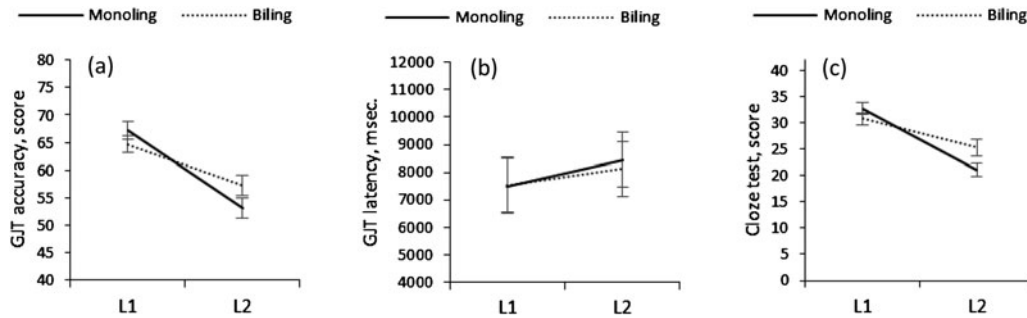


Fig. 2. (a) Mean grammaticality judgment scores; (b) mean grammaticality judgement latencies; and (c) mean cloze test scores, as a function of AoA at/after birth (i.e., L1/L2) and mono-/bilingualism. Error bars represent standard error of the mean.

bilinguals, $p = .714$). There were differences, though, within the monolingual groups, with L1 monolinguals attaining higher scores than L2 monolinguals ($p < .001$), as well as within the bilingual groups, with L1 bilinguals attaining higher scores than L2 bilinguals ($p < .019$). Lastly, the L1 monolinguals were found to outperform the L2 bilinguals ($p < .001$), and the L1 bilinguals the L2 monolinguals ($p < .01$). In other words, these results indicate a robust effect of AoA on grammatical intuition. Accuracy scores are presented in Figure 2a.

In terms of latency (log-transformed), a main effect was again documented for AoA, $\beta = 0.048$, $SE = 0.010$, $t = 4.66$, $p < .001$, showing that L1 (monolingual and bilingual) speakers exhibited overall shorter reaction times than L2 (monolingual and bilingual) speakers (corroborated in pairwise comparisons, according to which L1 speakers were significantly faster than L2 speakers, monolingual L1 vs. monolingual L2, $p < .001$; bilingual L1 vs. bilingual L2, $p = .014$). No significant main effect of bilingualism ($\beta = -0.009$, $SE = 0.010$, $t = -0.88$, $p = .377$) or interaction ($\beta = 0.011$, $SE = 0.010$, $t = 1.153$, $p = .253$) was found. Latencies are presented in Figure 2b.

Grammatical, lexical, and semantic inferencing

The cloze test scores revealed a significant main effect of AoA, $\beta = -0.678$, $SE = 0.099$, $p < .001$, but no main effect of bilingualism, $\beta = 0.068$, $SE = 0.099$, $p = .493$. However, a significant interaction was also found, $\beta = -0.251$, $SE = 0.099$, $p = .011$. Bonferroni posthoc tests revealed no differences between the L1 groups (L1 monolinguals vs. L1 bilinguals, $p = 1$) or between the L2 groups (L2 monolinguals vs. L2 bilinguals, $p = .148$). Significant differences were, however, found within the monolingual groups, with L1 monolinguals attaining higher scores than L2 monolinguals ($p < .001$) and the bilingual groups, with L1 bilinguals attaining higher scores than L2 bilinguals ($p = .025$). Finally, the L2 bilinguals were found to obtain lower scores than the L1 monolinguals ($p < 0.001$), and the L1 bilinguals higher scores than the L2 monolinguals ($p < 0.001$). Cloze test scores are depicted in Figure 2c. These comparisons thus show a significant advantage of L1 speakers over L2 speakers (irrespective of the mono-/bilingualism in either group) on this test.

Formulaic language

On the test assessing proficiency with idioms, a significant main effect of AoA was documented, $\beta = -0.467$, $SE = 0.149$, $p = .002$,

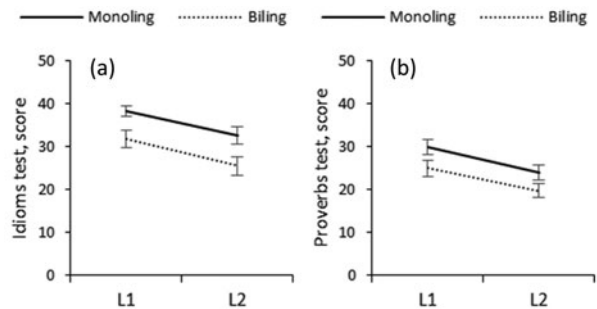


Fig. 3. Mean accuracy scores on (a) cloze test and (b) idioms test, as a function of AoA at/after birth (i.e., L1/L2) and mono-/bilingualism. Error bars represent standard error of the mean.

showing that L1 speakers in general attained higher scores than L2 speakers (the effect was consistent for both monolingual L1 vs. monolingual L2, $p = .028$, and for bilingual L1 vs. bilingual L2, $p = .026$). However, a significant main effect was also found for bilingualism, $\beta = 0.473$, $SE = 0.157$, $p = .003$ (confirmed in comparisons between monolingual and bilingual L1 speakers, $p = .013$, and between monolingual and bilingual L2 speakers, $p = .012$), suggesting that bilinguals overall scored significantly lower than monolinguals on this test. There was no significant interaction between AoA and bilingualism, $\beta = 0.065$, $SE = 0.149$, $p = .664$. Scores on the idioms test are presented in Figure 3a.

The analysis of the performance on the proverbs test revealed a significant main effect of AoA, $\beta = -0.467$, $SE = 0.149$, $p = .002$, suggesting again that, on average, L1 speakers were more proficient with the proverbs under scrutiny than L2 speakers (consistent across both monolingual L1 and L2 speakers, $p = .016$, and bilingual L1 and L2 speakers, $p = .025$). A significant main effect was also found for bilingualism, $\beta = 0.473$, $SE = 0.157$, $p = .003$, with monolingual L1 speakers in general attaining higher scores than bilingual L1 speakers ($p = .048$); for monolingual and bilingual L2 speakers this difference was only obtained at trend level ($p = .069$). No significant interaction was attested, $\beta = 0.065$, $SE = 0.149$, $p = .664$. Proverb scores are presented in Figure 3b.

Discussion

The current findings raise a number of important points for discussion, concerning not only the existence of AoA effects in ultimate attainment, but also the interpretation of AoA effects and bilingualism effects in general.

The impact of age of acquisition on ultimate attainment

The results revealed that out of the seven instruments used to assess ultimate attainment, none showed a standalone effect of bilingualism. Rather, when main effects of bilingualism were attested (for formulaic language), they always occurred in conjunction with main effects of AoA. Conversely, effects of AoA were found for six out of the seven instruments (the only exception being VOT production, where no effects of either predictor variable were documented). In those instances where an interaction between AoA and bilingualism was found (VOT perception, grammaticality judgement accuracy and cloze test performance), follow-up tests indeed confirmed consistent AoA effects and minimal bilingualism effects (if any at all). These results offer strong support for the AoA-effects hypothesis postulated above, and only limited support for the bilingualism-effects hypothesis. Moreover, the findings are consistent with previous research that has set out to examine the relative impact of AoA and bilingualism on L2 ultimate attainment using less comprehensive research designs than the current one (e.g., Bylund et al., 2012; Bylund et al., 2019; Norrman & Bylund, 2016; Veríssimo et al., 2018).

Is it possible, though, that the AoA effects attested here are in some way covert effects of bilingualism? The study has sought to disentangle these two variables in L2 speakers by including a group of international adoptees reporting to have undergone complete L1 loss. However, there is by now ample evidence showing that international adoptees may unconsciously retain some sort of L1 knowledge, even after decades of non-exposure (e.g., Choi et al., 2017a; Hyltenstam et al., 2009; Park, 2014; Pierce et al., 2014). Such knowledge is often manifested as a heightened sensitivity and/or distinct neurophysiological activation patterns to L1-phonetic contrasts in particular. It could, in other words, be argued that such L1 residual knowledge may give rise to L2 non-nativeness. It does, however, seem far-fetched that such traces would produce similar levels of bilingualism effects as would a fully functional L1. In fact, such a claim would entail that there is no proportion of bilingualism effects relative to L1 activation and proficiency, but that the mere existence of some kind of L1 knowledge, be it as a latent phonetic sensitivity or a full-fledged language, exerts an absolute effect on L2 attainment. While we have no desire to rule out the possibility that the adoptees in our study might have retained some L1 knowledge (despite self-reports to the contrary), we consider the idea of absolute bilingualism effects to be neither probable nor on par with previously reported findings on L1-L2 proficiency interactions (e.g., Bylund et al., 2012; Yeni-Komshian, Flege & Liu, 2000).

Thus, the current findings have far-reaching consequences for the ongoing debate on bilingualism effects in L2 ultimate attainment, which to date has been characterized by a shortage of empirical evidence. The robust effects of AoA attested in the current design are orthogonal to the interpretation that bilingualism, rather than age of acquisition, gives rise to near-native (as opposed to nativelike) ultimate attainment in early learners. As such, the findings speak against, first, the idea that bilingualism *per se* automatically results in non-nativeness/non-monolingual-like linguistic behavior (for a similar point, see Meisel, 2017), and second, the notion that L1 loss brings about nativelike/monolingual-like L2 attainment (e.g., Ventureyra et al., 2004). Seeing that the current study has assessed language proficiency with the same type of instruments as several previous CPH studies (e.g., Abrahamsson & Hyltenstam, 2009; Bialystok & Miller, 1999;

Birdsong & Molis, 2001; DeKeyser, 2000; DeKeyser et al., 2010; Granena & Long, 2013; Johnson & Newport, 1989) while controlling for bilingualism effects, one could argue that the findings on age of acquisition generated here are indeed not unique to a particular experimental paradigm, but may account for – and crucially, confirm – previously reported AoA effects.

Because our instruments are identical to those in one of these previous studies, Abrahamsson and Hyltenstam (2009), we are in a unique position to re-assess the incidence of L2 nativeness in that study. As mentioned earlier, out of 41 potentially nativelike learners, 3 (with AoA ≤ 8) were in the range of native speakers on all instruments used by Abrahamsson and Hyltenstam (2009). Since the current results showed main effects of bilingualism on the tests of formulaic language (idioms and proverbs), these instruments should be considered as tapping into bilingualism effects – *in addition* to effects of AoA – and could therefore be removed from the test battery in Abrahamsson and Hyltenstam's study. This removal results in the inclusion of two additional participants (AoA 1 and 4 years) who previously did not exhibit nativelike performance because of the formulaic language tests. This changes the number of learners who performed like native speakers on the relevant instruments from three to five, corresponding to a 5% increase of the nativeness rate in the sample (from 7% to 12% of the learners). In other words, the removal of instruments sensitive (also) to bilingualism effects had but marginal effects on the original findings reported by Abrahamsson and Hyltenstam (2009). As a consequence, any suggestion that the low incidence of nativelike L2 speakers in that study was an artefact of bilingualism would find only scant support.

Implications for interpreting the ultimate cause of AoA effects

What, then, are the consequences of the current findings for interpreting the mechanisms that underlie AoA effects in L2 ultimate attainment? De Leeuw (2014, p. 35) suggests that should AoA effects be detected in a design that controls for bilingualism, such as the current one, then this can be taken as evidence for maturational constraints on L2 learning. We believe, however, that it is better not to overestimate this design when interpreting the ultimate cause of the attested effects: while the results allow us to reject the generic claim that L2 nativelike attainment is impossible due to bilingualism effects, they do not necessarily reveal the specific locus of AoA. That said, it should be emphasized that the stand-alone effects of AoA documented in the present study in no way rule out a maturational constraints-based explanation. In fact, they are consistent with such an explanation. The exact nature of maturationally induced AoA effects is, however, yet to be uncovered, concerning both the actual changes in the mechanisms of language acquisition and in the resulting learned linguistic representations, as well as the type of sensitive period (nested sensitive period with cascading effects, or independent multiple sensitive periods). Relatedly, it is necessary to ask whether the same mechanisms may really account for the whole range of behaviors studied here, or whether different explanations are needed for different linguistic behaviors (which is certainly not inconceivable, cf. Johnson, 2005).

Implications for research on bilingualism effects

While the current study is primarily concerned with the potential role of bilingualism for nativelike attainment in an L2, the

findings have important implications for the understanding of bilingualism effects on verbal behavior in general. As mentioned in the background section, there has been a tendency in some research areas (e.g., the bilingual lexical deficit literature) not to systematically factor into the study design a distinction between simultaneous bilingualism and sequential bilingualism, but to instead lump together participants with different bilingual acquisition trajectories and test them in the societally dominant language. In such a design, a large part of the bilinguals may in fact be L2 speakers, who are then compared with monolingual L1 speakers. In view of the present findings, it is clear that while bilingualism may have a certain effect on some linguistic domains (e.g., lexis), AoA exerts more consistent effects across the board (including lexis; see Bylund et al., 2019). Thus, a design that sets out to assess bilingualism effects on linguistic behavior but confounds bilingualism with L2 status runs the risk of inflating any differences between the bilinguals and the monolingual comparison group, ultimately compromising their observations on bilingualism effects. In conclusion, just as inattention to the bilingualism of L2 speakers may be problematic for assessing AoA effects, as suggested by proponents of the bi-/multilingual turn in SLA, we argue that inattention to AoA and to the fact that bilinguals may be L2 speakers is equally problematic for assessing bilingualism effects.

Conclusions

The aim of the present study was to address empirically the notion that bilingualism, rather than age of acquisition, underlies less than nativelike attainment in early (and, by inference, also exceptionally successful adult) L2 learners. In a factorial design where the variables of AoA at birth/after birth vs. monolingualism/bilingualism were fully crossed, the results from a comprehensive battery of previously used tests (see, e.g., Abrahamsson & Hyltenstam, 2009) showed minimal effects of bilingualism, but major effects of AoA. These findings were discussed in relation to the ongoing debate on AoA vs. bilingualism effects on L2 ultimate attainment, and also in terms of their implications for interpreting AoA effects and bilingualism effects in general.

There is a risk that sweeping arguments about bilingualism underlying non-nativeness in early learners may in the end backfire, as they inflate the expectations of the explanatory potential of this variable. As shown by the current study, the effects of bilingualism on L2 attainment are more limited and selective in scope than previously thought. Ultimately, it is not rhetoric, but empirical assessments, along with conceptual analyses of the notions of mono- and bilingualism (see Bylund, Hyltenstam & Abrahamsson, 2013), that will further our knowledge in this area. We choose to believe Birdsong (2018) when he ascertains that “no researchers claim that bilingualism effects alone are responsible for all divergences from monolingual-likeness in bilingualism” (p. 6). At the same time, however, we sense that the alleged negative effects of using monolingual native speakers as baseline may have been exaggerated; according to the present data and previously reported findings (e.g., Bylund et al., 2012; 2019; Hyltenstam et al., 2009; Norrman & Bylund, 2016), there may be no urgent need for making a general shift in SLA research to the use of simultaneous bilinguals as golden standard for every linguistic domain. In order to bring further clarity into this issue, we encourage future studies to not only assess different types of monolingual and bilingual L1 and L2 speakers (including simultaneous bilinguals, sequential monolinguals, and others) – but

also to factor into their designs test instruments that allow for a systematic targeting of linguistic domains and structures that exhibit different degrees of likelihood to elicit bilingualism effects (e.g., based on typological analysis). This will be crucial for understanding the differential effects of age of acquisition and bilingualism, and any potential interaction between the two, on ultimate attainment.

Acknowledgements. This research was funded by Riksbankens Jubileumsfond (The Swedish Foundation for Humanities and Social Sciences), grant no. M2005-0459 (to K.H.) and grant no. SAB16-0051:1 (to N.A.). We are grateful to João Verissimo and two anonymous reviewers for providing insightful comments on a previous version of the manuscript, to Gunnar Norrman for collecting and organizing the data, to Robyn Berghoff for statistical advice, and to Lamont Anteau for editing our English writing. Needless to say, any remaining errors are entirely our own.

References

- Abrahamsson N (2012) Age of onset and nativelike L2 ultimate attainment of morphosyntactic and phonetic intuition. *Studies in Second Language Acquisition* 34, 187–214. <https://doi.org/10.1017/S0272263112000022>
- Abrahamsson N and Hyltenstam K (2008) The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition* 30, 481–509. <https://doi.org/10.1017/S027226310808073X>
- Abrahamsson N and Hyltenstam K (2009) Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning* 59, 249–306. <https://doi.org/10.1111/j.1467-9922.2009.00507.x>
- Bates D, Mächler M, Bolker B and Walker S (2014) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67, 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Berken JA, Gracco VL and Klein D (2017) Early bilingualism, language attainment, and brain development. *Neuropsychologia* 98, 220–227. <https://doi.org/10.1016/j.neuropsychologia.2016.08.031>
- Bialystok E (2009) Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition* 12, 3–11. <https://doi.org/10.1017/S1366728908003477>
- Bialystok E (2016) The signal and the noise: Finding the pattern in human behavior. *Linguistic Approaches to Bilingualism* 6, 517–534. <https://doi.org/10.1075/lab.15040.bia>
- Bialystok E (2017) The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin* 143, 233–262. <https://doi.org/10.1037/bul0000099>
- Bialystok E and Miller B (1999) The problem of age in second-language acquisition: Influences from language, structure, and task. *Bilingualism: Language and Cognition* 2, 127–145. <https://doi.org/10.1017/S1366728999000231>
- Birdsong D (2005) Nativelikeness and non-nativelikeness in L2A research. *International Review of Applied Linguistics in Language Teaching* 43, 319–328. <https://doi.org/10.1515/iral.2005.43.4.319>
- Birdsong D (2006) Age and second language acquisition and processing: A selective overview. *Language Learning* 56, 9–49. <https://doi.org/10.1111/j.1467-9922.2006.00353.x>
- Birdsong D (2014) Dominance and age in bilingualism. *Applied Linguistics*, amu031. <https://doi.org/10.1093/applin/amu031>
- Birdsong D (2018) Plasticity, variability and age in second language acquisition and bilingualism. *Frontiers in Psychology* 9. <https://doi.org/10.3389/fpsyg.2018.00081>
- Birdsong D and Gertken LM (2013) In faint praise of folly: A critical review of native/non-native speaker comparisons, with examples from native and bilingual processing of French complex syntax. *Language, Interaction and Acquisition / Langage, Interaction et Acquisition* 4, 107–133. <https://doi.org/10.1075/lia.4.2.01bir>
- Birdsong D and Molis M (2001) On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language* 44, 235–249. <https://doi.org/10.1006/jmla.2000.2750>

- Birdsong D and Quinto-Pozos D** (2018) Signers and speakers, age and attainment. *Bilingualism: Language and Cognition* 21, 911–912. <https://doi.org/10.1017/S1366728918000226>
- Boersma P** (2002) Praat, a system for doing phonetics by computer. *Glott International* 5. Retrieved from <https://dare.uva.nl/search?metis.record.id=200596>
- Bylund E** (2011) Segmentation and temporal structuring of events in early Spanish-Swedish bilinguals. *International Journal of Bilingualism* 15, 56–84. <https://doi.org/10.1177/1367006910379259>
- Bylund E, Abrahamsson N and Hyltenstam K** (2010) The role of language aptitude in first language attrition: The case of pre-pubescent attriters. *Applied Linguistics* 31, 443–464. <https://doi.org/10.1093/applin/amp059>
- Bylund E, Abrahamsson N and Hyltenstam K** (2012) Does first language maintenance hamper nativelikeness in a second language? *Studies in Second Language Acquisition* 34, 215–241. <https://doi.org/10.1017/S027226312000034>
- Bylund E, Abrahamsson N, Hyltenstam K and Norrman G** (2019) Revisiting the bilingual lexical deficit: The impact of age of acquisition. *Cognition* 182, 45–49. <https://doi.org/10.1016/j.cognition.2018.08.020>
- Bylund E, Hyltenstam K and Abrahamsson N** (2013) Age of acquisition effects or effects of bilingualism in second language ultimate attainment? In Granena G & Long M (eds), *Sensitive Periods, Language Aptitude, and Ultimate L2 Attainment*. Amsterdam: John Benjamins Publishing, pp. 69–103.
- Caramazza A, Yeni-Komshian GH, Zurif EB and Carbone E** (1973) The acquisition of a new phonological contrast: The case of stop consonants in French–English bilinguals. *The Journal of the Acoustical Society of America* 54, 421–428. <https://doi.org/10.1121/1.1913594>
- Choi J, Broersma M and Cutler A** (2017a) Early phonology revealed by international adoptees' birth language retention. *Proceedings of the National Academy of Sciences* 114, 7307–7312. <https://doi.org/10.1073/pnas.1706405114>
- Choi J, Cutler A and Broersma M** (2017b) Early development of abstract language knowledge: Evidence from perception–production transfer of birth-language memory. *Royal Society Open Science* 4, 160660. <https://doi.org/10.1098/rsos.160660>
- Cook V** (1991) The poverty-of-the-stimulus argument and multicompetence. *Second Language Research* 7, 103–117. <https://doi.org/10.1177/026765839100700203>
- Cook V** (1999) Going beyond the native speaker in language teaching. *TESOL Quarterly* 33, 185–209. <https://doi.org/10.2307/3587717>
- Cook V** (ed) (2003) *Effects of the Second Language on the First*. Clevedon: Multilingual Matters.
- Cook V** (2016) Premises of multi-competence. In Wei L & Cook V (eds), *The Cambridge Handbook of Linguistic Multi-Competence*. Cambridge: Cambridge University Press, pp. 1–25.
- de Leeuw E** (2014) Maturation constraints in bilingual speech. In Thomas EM & Mennen I (eds), *Advances in the Study of Bilingualism*. Bristol: Multilingual Matters, pp. 25–37. <https://doi.org/10.21832/9781783091713>
- DeKeyser R** (2000) The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22, 499–533. <https://doi.org/null>
- DeKeyser R** (2012) Age effects in second language learning. In Gass SM & Mackey A (eds), *The Routledge Handbook of Second Language Acquisition*. New York: Routledge, pp. 442–460.
- DeKeyser R, Alfi-Shabtay I and Ravid D** (2010) Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistics* 31, 413–438. <https://doi.org/10.1017/S0142716410000056>
- DeKeyser R and Larson-Hall J** (2005) What does the critical period really mean? In Kroll JF & De Groot AMB (eds), *Handbook of Bilingualism: Psycholinguistic Approaches*. New York, NY: Oxford University Press, pp. 88–108.
- Dick AS, Garcia NL, Pruden SM, Thompson WK, Hawes SW, Sutherland MT, Riedel MC, Laird AR and Gonzalez R** (2019) No evidence for a bilingual executive function advantage in the nationally representative ABCD study. *Nature Human Behaviour*, 1. <https://doi.org/10.1038/s41562-019-0609-3>
- Erman B, Forsberg Lundell F and Lewis M** (2018) Formulaic language in advanced long-residency L2 speakers. In Hyltenstam K, Bartning I & Fant L (eds), *High-Level Language Proficiency in Second Language and Multilingual Contexts*. Cambridge: Cambridge University Press, pp. 96–119. <https://doi.org/10.1017/9781316809686.005>
- Finney DJ** (1947) *Probit Analysis: A Statistical Treatment of the Sigmoid Response Curve*. Oxford, England: Macmillan.
- Fløge JE** (1999) Age of learning and second language speech. In Birdsong D (ed), *Second Language Acquisition Research. Second Language Acquisition and the Critical Period Hypothesis*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers, pp. 101–131.
- Foster P, Bolibaug C and Kotula A** (2014) Knowledge of nativelike selections in a L2. The influence of exposure, memory, age of onset, and motivation in foreign language and immersion settings. *Studies in Second Language Acquisition* 36, 101–132. <https://doi.org/10.1017/S0272263113000624>
- Gauthier K and Genesee F** (2011) Language development in internationally adopted children: A special case of early second language learning. *Child Development* 82, 887–901. <https://doi.org/10.1111/j.1467-8624.2011.01578.x>
- Gauthier K, Genesee F and Kasparian K** (2012) Acquisition of complement clitics and tense morphology in internationally adopted children acquiring French. *Bilingualism: Language and Cognition* 15, 304–319. <https://doi.org/10.1017/S1366728910000635>
- Granena G and Long MH** (2013) Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research* 29, 311–343. <https://doi.org/10.1177/0267658312461497>
- Grosjean F** (1989) Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language* 36, 3–15.
- Grosjean F** (1998) Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition* 1, 131–149. <https://doi.org/10.1017/S136672899800025X>
- Hartshorne JK, Tenenbaum JB and Pinker S** (2018) A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition* 177, 263–277. <https://doi.org/10.1016/j.cognition.2018.04.007>
- Hazan VL and Boulakia G** (1993) Perception and production of a voicing contrast by French-English bilinguals. *Language and Speech* 36, 17–38. <https://doi.org/10.1177/002383099303600102>
- Herschensohn J** (2007) *Language Development and Age*. Cambridge: Cambridge University Press.
- Hyltenstam K** (1992) Non-native features of near-native speakers: On the ultimate attainment of childhood L2 learners. In Harris RJ (ed), *Cognitive Processing in Bilinguals*. Amsterdam: North-Holland, pp. 351–368.
- Hyltenstam K and Abrahamsson N** (2003a) Age of onset and ultimate attainment in near-native speakers of Swedish. In Fraurud K & Hyltenstam K (eds), *Multilingualism in Global and Local Perspectives: Papers from the 8th Nordic Conference on Bilingualism, November 1–3, 2001*. Stockholm: Stockholm University, pp. 319–340. Retrieved from <http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-34735>
- Hyltenstam K and Abrahamsson N** (2003b) Maturation constraints in SLA. In Doughty C & Long MH (eds), *The Handbook of Second Language Acquisition*, Malden MA: Blackwell Pub, 539–588.
- Hyltenstam K, Bylund E, Abrahamsson N and Park H-S** (2009) Dominant-language replacement: The case of international adoptees. *Bilingualism: Language and Cognition* 12, 121–140. <https://doi.org/10.1017/S1366728908004008>
- Johnson CE and Wilson IL** (2002) Phonetic evidence for early language differentiation: Research issues and some preliminary data. *International Journal of Bilingualism* 6, 271–289. <https://doi.org/10.1177/13670069020060030401>
- Johnson JS and Newport EL** (1989) Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21, 60–99. [https://doi.org/10.1016/0010-0285\(89\)90003-0](https://doi.org/10.1016/0010-0285(89)90003-0)
- Johnson MH** (2005) Sensitive periods in functional brain development: Problems and prospects. *Developmental Psychobiology* 46, 287–292. <https://doi.org/10.1002/dev.20057>
- Klein D, Mok K, Chen J-K and Watkins KE** (2014) Age of language learning shapes brain structure: A cortical thickness study of bilingual and

- monolingual individuals. *Brain and Language* **131**, 20–24. <https://doi.org/10.1016/j.bandl.2013.05.014>
- Lehtonen M, Soveri A, Laine A, Järvenpää J, de Bruin A and Antfolk J** (2018) Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin* **144**, 394–425. <https://doi.org/10.1037/bul0000142>
- Lenneberg EH** (1967) *Biological Foundations of Language*. New York: Wiley.
- Long MH** (1990) Maturation constraints on language development. *Studies in Second Language Acquisition* **12**, 251–285. <https://doi.org/10.1017/S0272263100009165>
- Lisker L and Abramson A** (1964) A cross-language study of voicing in initial stops: Acoustical measurements. *Word* **20**, 384–422. <https://doi.org/10.1080/00437956.1964.11659830>
- Long MH** (1993) Second language acquisition as a function of age. In Hyltenstam K & Viberg A (eds), *Progression and Regression in Language: Sociocultural, Neuropsychological and Linguistic Perspectives*. Cambridge: Cambridge University Press, pp. 196–221.
- Long MH** (2007) *Problems in SLA*. Hillsdale, N.J.: L. Erlbaum Associates.
- Long MH** (2013) Maturation constraints on child and adult SLA. In Granena G & Long MH (eds), *Sensitive Periods, Language Aptitude, and Ultimate L2 Attainment*. Amsterdam: John Benjamins Publishing, pp. 3–42. Retrieved from <https://benjamins.com/catalog/llt.35>
- MacWhinney B** (Ed.). (1999) *The Emergence of Language*. Mahwah, N.J.: Lawrence Erlbaum.
- MacWhinney B** (2016) Entrenchment in second-language learning. In Schmid H-J (ed), *Entrenchment and the Psychology of Language Learning: How We Reorganize and Adapt Linguistic Knowledge*. Berlin: De Gruyter Mouton, pp. 343–366. <https://doi.org/10.1515/9783110341423-016>
- Marinova-Todd S** (2003) *Comprehensive Analysis of Ultimate Attainment in Adult Second Language Acquisition*. PhD Thesis. MIT, Harvard.
- Mayberry RI and Kluender R** (2018) Rethinking the critical period for language: New insights into an old question from American Sign Language. *Bilingualism: Language and Cognition* **21**, 886–905. <https://doi.org/10.1017/S1366728917000724>
- Meisel JM** (2008) The bilingual child. In Bhatia TK & Ritchie WC (eds), *The Handbook of Bilingualism*. Malden: Blackwell Publishing, pp. 90–113. <https://doi.org/10.1002/9780470756997.ch4>
- Meisel JM** (2017) Bilingual acquisition. In Fernández EM & Smith Cairns H (eds), *The Handbook of Psycholinguistics*. Hoboken, NJ: Wiley Blackwell, pp. 635–652. <https://doi.org/10.1002/9781118829516.ch28>
- Mody M, Schwartz RG, Gravel JS and Ruben RJ** (1999) Speech perception and verbal memory in children with and without histories of otitis media. *Journal of Speech, Language, and Hearing Research* **42**, 1069–1079.
- Montrul S and Ionin T** (2010) Transfer effects in the interpretation of definite articles by Spanish heritage speakers. *Bilingualism: Language and Cognition* **13**, 449–473. <https://doi.org/10.1017/S1366728910000040>
- Morford JP and Mayberry RI** (2000) A reexamination of ‘early exposure’ and its implications for language acquisition by eye. In Chamberlain C, Morford JP & Mayberry RI (eds), *Language Acquisition by Eye*. Mahwah, NJ: Lawrence Erlbaum, pp. 111–127.
- Muñoz C and Singleton D** (2011) A critical review of age-related research on L2 ultimate attainment. *Language Teaching* **44**, 1–35. <https://doi.org/10.1017/S0261444810000327>
- Norrman G and Bylund E** (2016) The irreversibility of sensitive period effects in language development: Evidence from second language acquisition in international adoptees. *Developmental Science* **19**, 513–520. <https://doi.org/10.1111/desc.12332>
- Ortega L** (2010) *The Bilingual Turn in SLA*. Presented at the American Association for Applied Linguistics Conference, Atlanta, GA, March 6–9.
- Ortega L** (2013) Ways forward for a bi/multilingual turn in SLA. In May S (ed), *The Multilingual Turn. Implications for SLA, TESOL and Bilingual Education*. London and New York: Routledge, pp. 32–52. <https://doi.org/10.4324/9780203113493-9>
- Pallier C** (2007) Critical periods in language acquisition and language attrition. In B Köpcke, Schmid M, Keijzer M & Dostert S (eds), *Language Attrition. Theoretical Perspectives*. Amsterdam: John Benjamins Publishing, pp. 135–153.
- Pallier C, Dehaene S, Poline J-B, LeBihan D, Argenti A-M, Dupoux E and Mehler J** (2003) Brain imaging of language plasticity in adopted adults: Can a second language replace the first? *Cerebral Cortex* **13**, 155–161. <https://doi.org/10.1093/cercor/13.2.155>
- Paradis M** (2004) *A Neurolinguistic Theory of Bilingualism*. Amsterdam: John Benjamins.
- Paradis M** (2009) *Declarative and Procedural Determinants of Second Languages*. Amsterdam: John Benjamins Publishing.
- Park H-S** (2015) Korean adoptees in Sweden: Have they lost their first language completely? *Applied Psycholinguistics*, 773–797. <https://doi.org/10.1017/S0142716413000507>
- Pfenninger SE and Singleton D** (2017) *Beyond Age Effects in Instructional L2 Learning. Revisiting the Age Factor*. <https://doi.org/10.21832/9781783097630>
- Pierce LJ, Chen J-K, Delcenserie A, Genesee F and Klein D** (2015) Past experience shapes ongoing neural patterns for language. *Nature Communications* **6**, 10073. <https://doi.org/10.1038/ncomms10073>
- Pierce LJ, Klein D, Chen J-K, Delcenserie A and Genesee F** (2014) Mapping the unconscious maintenance of a lost first language. *Proceedings of the National Academy of Sciences* **111**, 17314–17319. <https://doi.org/10.1073/pnas.1409411111>
- Pulvermüller F and Schumann JH** (1994) Neurobiological mechanisms of language acquisition. *Language Learning* **44**, 681–734. <https://doi.org/10.1111/j.1467-1770.1994.tb00635.x>
- R Core Team**. (2014) *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
- Reetzke R, Lam BP-W, Xie Z, Sheng L and Chandrasekaran B** (2016) Effect of simultaneous bilingualism on speech intelligibility across different masker types, modalities, and signal-to noise ratios in school-age children. *PLOS ONE* **11**, e0168048. <https://doi.org/10.1371/journal.pone.0168048>
- Rubén RJ** (1999) Persistency of an effect: Otitis media during the first year of life with nine years follow-up. *International Journal of Pediatric Otorhinolaryngology* **49**, Supplement 1, S115–S118. [https://doi.org/10.1016/S0165-5876\(99\)00145-7](https://doi.org/10.1016/S0165-5876(99)00145-7)
- Schmidt AM and Flege JE** (1996) Speaking rate effects on stops produced by Spanish and English monolinguals and Spanish/English bilinguals. *Phonetica* **53**, 162–179. <https://doi.org/10.1159/000262196>
- Schneider W, Eschman A and Zuccolotto A** (2002a) *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools, Inc.
- Schneider W, Eschman A and Zuccolotto A** (2002b) *E-Prime Reference Guide*. Pittsburgh: Psychology Software Tools, Inc.
- Singh L, Liederman J, Mierzejewski R and Barnes J** (2011) Rapid reacquisition of native phoneme contrasts after disuse: You do not always lose what you do not use. *Developmental Science* **14**, 949–959. <https://doi.org/10.1111/j.1467-7687.2011.01044.x>
- Singleton D and Pfenninger SE** (2018) L2 acquisition in childhood, adulthood and old age. Misreported and under-researched dimensions of the age factor. *Journal of Second Language Studies* **1**, 254–275. <https://doi.org/10.1075/jsls.00003.sin>
- Sorace A and Robertson D** (2001) Measuring development and ultimate attainment in non-native grammars. In Elder C (ed), *Experimenting with Uncertainty: Essays in Honour of Alan Davies*. Cambridge: Cambridge University Press, pp. 264–274.
- Stölten K, Abrahamsson N and Hyltenstam K** (2014) Effects of age of learning on voice onset time: Categorical perception of Swedish stops by near-native L2 speakers. *Language and Speech* **57**, 425–450. <https://doi.org/10.1177/0023830913508760>
- Stölten K, Abrahamsson N and Hyltenstam K** (2015) Effects of age and speaking rate on voice onset time: the production of voiceless stops by near-native L2 Speakers. *Studies in Second Language Acquisition* **37**, 71–100. <https://doi.org/10.1017/S0272263114000151>
- Sundara M, Polka L and Baum S** (2006) Production of coronal stops by simultaneous bilingual adults. *Bilingualism: Language and Cognition* **9**, 97–114. <https://doi.org/10.1017/S1366728905002403>
- Taylor WL** (1953) “Cloze Procedure”: A new tool for measuring readability. *Journalism Bulletin* **30**, 415–433. <https://doi.org/10.1177/107769905303000401>

- Ullman MT** (2004) Contributions of memory circuits to language: The declarative/procedural model. *Cognition* **92**, 231–270. <https://doi.org/10.1016/j.cognition.2003.10.008>
- Ullman MT** (2015) The Declarative/Procedural Model: A neurobiologically motivated theory of first and second language. In VanPatten B & Williams J (eds), *Theories in Second Language Acquisition. An Introduction*. New York: Routledge, pp. 135–158. <https://doi.org/10.4324/9780203628942-13>
- Vanhove J** (2013) The critical period hypothesis in second language acquisition: A statistical critique and a reanalysis. *PLoS ONE* **8**, e69172. <https://doi.org/10.1371/journal.pone.0069172>
- Ventureyra VAG, Pallier C and Yoo H-Y** (2004) The loss of first language phonetic perception in adopted Koreans. *Journal of Neurolinguistics* **17**, 79–91. [https://doi.org/10.1016/S0911-6044\(03\)00053-8](https://doi.org/10.1016/S0911-6044(03)00053-8)
- Verissimo J** (2018) Sensitive periods in both L1 and L2: Some conceptual and methodological suggestions. *Bilingualism: Language and Cognition* **21**, 932–933. <https://doi.org/10.1017/S1366728918000275>
- Verissimo J, Heyer V, Jacob G and Clahsen H** (2018) Selective effects of age of acquisition on morphological priming: Evidence for a sensitive period. *Language Acquisition* **25**, 315–326. <https://doi.org/10.1080/10489223.2017.1346104>
- Volaitis LE and Miller JL** (1992) Phonetic prototypes: Influence of place of articulation and speaking rate on the internal structure of voicing categories. *The Journal of the Acoustical Society of America* **92**, 723–735. <https://doi.org/10.1121/1.403997>
- Werker JF and Hensch TK** (2015) Critical periods in speech perception: New directions. *Annual Review of Psychology* **66**. <https://doi.org/10.1146/annurev-psych-010814-015104>
- Wray A** (2005) *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.
- Yeni-Komshian GH, Flege JE and Liu S** (2000) Pronunciation proficiency in the first and second languages of Korean–English bilinguals. *Bilingualism: Language and Cognition* **3**, 131–149.