1 Age Effects on Language Acquisition, Retention and Loss

Key Hypotheses and Findings

Niclas Abrahamsson, Kenneth Hyltenstam and Emanuel Bylund

1.1 Introduction

One of the longest-standing scientific debates on language acquisition concerns the relationship between age of first exposure and language development, the big divide being between proponents and opponents of the idea of one or several critical period(s) for language. While being perhaps the most-observed language acquisition phenomenon among laypeople, ironically, children's success in second language (L2) acquisition relative to that of adults is still, fifty years after the publication of Lenneberg's (1967) *Critical Period Hypothesis* (CPH), a highly contested issue among scholars of linguistics, cognitive psychology, second language acquisition and bilingualism. Considering the boldness of a biologically underpinned hypothesis on human language acquisition that assumes maturational changes to be the cause of child–adult differences, it should have come as no surprise to Lenneberg (nor anyone else) that the CPH was to become subject to falsification, and quite intensively so. The earliest challenges were based on an entirely unrestricted interpretation of the hypothesis espoused by some researchers (e.g. Asher & Price 1967; Olson & Samuels 1973; Snow & Hoefnagel-Höhle 1977, 1978), who compared the language performance of small children, older children, adolescents and/or adults either in a 'teach and test' format in laboratory-like settings, sometimes with target structures from languages previously unknown to the participants, or for their initial language learning achievements during their first few months of residence in new language settings. Because adolescent and adult learners outperformed child learners, and because older children did better than very young children in these extremely time-constrained and sometimes artificial learning situations, the CPH was considered falsified. Needless to say, the relevance of this version of the hypothesis and the empirical data that emanated from it have been seriously questioned (see e.g.

Krashen, Long & Scarcella 1979; Long 1990; for discussions, see also Hyltenstam & Abrahamsson 2000, 2001, 2003a), and it was soon abandoned in favour of more restricted versions that made predictions about the relationship between starting age and long-term (rather than short-term) L2 achievement.

Since then, the controversy has centred not so much on the very fact that most children eventually attain higher levels of second language proficiency than most adults, but rather on the reasons for why this should be so. Do child–adult differences in language acquisition originate in maturational changes in the brain during a limited critical period in childhood, as implied by the CPH, or are they more likely to stem from non-biological yet age-dependent changes in learning conditions? Is nativelike L2 attainment possible for adult L2 learners and, if so, what would be the significance of such individuals for the CPH? Does L2 development come about through the same or fundamentally different neurocognitive systems in children and adults, allowing automatic, implicit language acquisition only during a limited period of heightened sensitivity to naturalistic language exposure in childhood? Aside from the many ideologically coloured objections over the years to 'nativelikeness' as falsification criterion,[1] one of the most theoretically challenging – and most recent – objections to the maturational constraints interpretation makes the claim that age of acquisition (being the confounding variable it undeniably is) does not conceal maturational changes per se, but rather the age-related entrenchment of the bilingual's first language (L1) or other cognitive effects – some of which are dependent and some independent of age – of being a bilingual speaker. Thus, although many would have thought that fifty years of empirical research and theoretical refinement should have sufficed to settle the maturational constraints issue, the critical period debate seems as vibrant as ever.

Rather than providing yet another comprehensive review of the CPH-related literature,[2] our aim in this chapter is to recapitulate what may be seen as the key falsification criteria and to investigate how our own empirical results relate to these. The research that we draw upon encompasses a wide range of language learning circumstances that include: second and heritage language acquisition; early and late language exposure; simultaneous and sequential bilingualism, as compared to early and late/sequential monolingualism (in the case of international adoptees);[3] near-native ultimate attainment; language learning aptitude;

[1] We will not review this debate here, but see e.g. Birdsong (2005); Birdsong & Gertken (2013); Cook (1999, 2002); Davies (2003); Dewaele (2017); and Ortega (2010, 2013), for critical discussions on nativelikeness as falsification criterion and the 'monolingual native-speaker bias'.

[2] See instead Abrahamsson (2013); Bylund (2009a); Bylund, Hyltenstam & Abrahamsson (2013); Hyltenstam (2010, 2012); Hyltenstam & Abrahamsson (2000, 2003a).

[3] Strictly speaking, international adoptees are not monolingual in the sense that they have childhood language acquisition experience with more than one language, albeit sequential. In addition, there is clear evidence today of the existence of remnants of their L1, at least for phonology (Choi, Broersma & Cutler 2017; Hyltenstam, Bylund, Abrahamsson & Park

first language attrition; and first language residual activation (again, in the case of international adoptees). These issues were investigated using innovative and comprehensive research designs comprising a strict selection of participants and native controls, a large battery of elicitation techniques with demanding language tasks covering aspects of phonology, grammar and lexical idiomaticity in both production and perception, as well as an aptitude component. The findings privilege an interpretation of age-of-acquisition effects in terms of maturational constraints rather than primarily as an effect of cross-linguistic or sociopsychological circumstances. We conclude the chapter with implications for further research directions, specifically with respect to potential bilingualism effects and a dual system approach.

1.2 The Impact of Age of Onset on L2 Ultimate Attainment

The minimal evidence of maturational constraints on language acquisition should be an observable, general age effect on learning outcomes. The overarching spirit of both the layman observation and of Lenneberg's (or any other version of the) CPH is that children somehow are 'better' language learners, not least because they – undeniably and with very few exceptions – ultimately develop into more competent L2 speakers than adults. In fact, a wealth of studies have corroborated the inverse relation between the *age of onset* (AO) of acquisition and the *ultimate attainment* (UA) of the second language, either by comparing the average UA of different AO cohorts, or through inference from AO–UA correlations (see e.g. Asher & García 1969; Bialystok & Miller 1999; DeKeyser 2000; Flege, Yeni-Komshian & Liu 1999; Granena & Long 2013; Huang 2014; Johnson & Newport 1989; Munro & Mann 2005; Oyama 1976, 1978; Patkowski 1980). Symptomatically, no study has ever reported a lacking (let alone positive) relationship between AO and UA. Although the correlation coefficient varies across studies (depending on the methodology used; for a review of research designs and correlations, see DeKeyser & Larson-Hall 2005), AO has notoriously turned out to be the independent variable that best explains the data, 'typically accounting for about 30% of the variance in ultimate attainment in L2 grammar . . . and around 50% in pronunciation' (Long 2013: 7).

2009; Park 2015; Pierce, Klein, Chen, Delcenserie & Genesee 2014). However, it has also recently been clarified that such remnants are not retrievable for language processing except after a period of reexposure and relearning (Pierce, Klein, Chen, Delcenserie & Genesee 2014). It is therefore reasonable to consider adults with such experience in childhood, but totally unaware of any memories of their birth language, as functionally monolingual, which is the reason behind our use of this terminology throughout. To claim that they are bilingual would certainly be inappropriate in any sense of the term.

To exemplify, Abrahamsson (2012) investigated 200 immigrant L1 Spanish learners of L2 Swedish for their UA of grammatical and phonetic intuition through an auditory grammaticality judgment test (GJT) and a test of categorical perception of voice onset time (VOT) of /p/ and /b/. The learners were evenly distributed over an AO span between 1 and 30 years (ca. 6–8 participants per AO interval), they were adults at the time of testing (M = 40 years, *range* = 21–63), and their length of residence (LOR) in Sweden was extensive (M = 25 years, *range* = 15–46). The results revealed statistically significant differences between the 101 early learners (AO ≤ 15) and the 99 late learners (AO ≥ 16) on both measures, and the overall AO–UA correlations r were −.60 and −.47 for GJT and VOT, respectively. Weaker yet statistically significant correlations were obtained between UA and other variables, such as LOR, current age and percentage L1 use, but these dropped considerably and to nonsignificant levels when the confounding effect of AO was partialled out,[4] while the AO–UA correlations remained at robust levels (−.52 to −.61 for GJT and −.34 to −.47 for VOT) when the effect of other variables was removed. This result is representative of studies that have used partial correlations or regression analyses: AO stands out as the most predictive variable, whereas other (affective, psychological, cognitive or experiential) variables seem to contribute significantly less (if at all) to the variance of the data (see e.g. Granena & Long 2013; Huang 2014; Johnson & Newport 1989; Stevens 1999).

Further evidence of the robust impact of AO on UA can be found in Abrahamsson & Hyltenstam (2009), which comprised 195 L2 learners of Swedish (again, with Spanish as L1 and, again, distributed over a wide AO continuum, although this time between <1 and 47 years). Audio recorded samples of informal speech from the learners (who all considered themselves to be potentially nativelike or near-native speakers of Swedish) were judged by ten native listeners. The judges were to decide for each speech sample whether it was produced by a native or a non-native speaker of Swedish, and the judgments were then transformed into scores of *perceived nativelikeness* between 0 and 10 (corresponding to the number of judges who perceived a speaker as a native speaker). Despite the speakers' self-assessed nativelikeness, the difference between learners with AO ≤ 11 and AO ≥ 12 was large (M = 7.9 and 2.5, respectively), and a strong, statistically significant overall negative correlation (r = −.72) was obtained between AO and perceived nativelikeness, explaining around 52 per cent of the variance, while other independent variables (LOR, current age, L1 use, sex) explained only about 2 to 8 per cent each, which fits well with the DeKeyser and Larson-Hall (2005) meta-analysis of

[4] Except for L1 use, which remained at a weak but statistically significant correlation coefficient (r = −.29 and −.21 for GJT and VOT, respectively).

accent studies suggesting a contribution from AO of 'at least 50%' as compared to 'only 5% or less to the total variance' (p. 93) from other variables.

In order to obtain a measure of *actual* (rather than *perceived*) nativelikeness, a subsample of 41 participants (31 with AO 1–11 and 10 with AO 13–19), all of whom had passed for native speakers with a majority (i.e. at least 6) of the 10 judges, were then subjected to a comprehensive and demanding test battery that produced 10 phonetic, grammatical and lexical UA measures, and the learners' results were then compared to the performance of 15 native control speakers on those same measures (see description in Section 1.4). As with perceived nativelikeness, there was a difference in mean scores of *scrutinized nativelikeness* (M = 6.0 *versus* 3.5 out of 10) between the early-learner and the older-learner group, as well as a statistically significant (although substantially weaker) negative correlation with AO ($r = -.38$). Similarly, detailed analysis of the production and categorical perception of Swedish VOT for these 41 near-native L2 speakers, reported in Stölten, Abrahamsson & Hyltenstam (2014, 2015), revealed significantly different group means as well as medium-strong AO–UA correlations (r between –.31 and –.47) for bilabial, dental and velar stops. One might have expected these learners' initial passing for native speaker to have severely reduced, even neutralized, the effects of AO, but this was clearly not the case – general age effects are prevalent even among L2 speakers who see themselves as potentially nativelike and who pass for natives by actual native speakers of the target language. In other words, the predictive power of AO on UA is, as we see it, indisputable and beyond compare.

1.3 The AO–UA Function: Linear or Discontinuous?

A major interpretational problem with overall correlations is that a general inverse relation between AO and UA does not constitute sufficient evidence that language acquisition is constrained by maturational changes in the brain, as the correlation itself remains uninformative for the very causes of age effects. In fact, a steady decline in UA that continues across the full AO range and even beyond the alleged maturational asymptote constitutes potential *counter*evidence to the CPH. A nonlinear (rather than a linear) AO–UA function, with discontinuities at theoretically motivated ages, is to be expected if a critical period is at work, while a continuous decline across those ages (and beyond) can be expected if non-maturational factors are the cause of the observed age effects.[5] As with critical period-induced developments in other species (such as imprinting in goslings, birdsong in sparrows and finches and vision in cats), we

[5] Linear patterns have been reported, based on large sets of self-reported census data in the United States (see e.g. Hakuta, Bialystok & Wiley 2003; Chiswick & Miller 2008; Stevens 1999). We will not deal with these studies here, due to their somewhat peculiar formatting; for critical discussions, see Long (2013) and DeKeyser (2006).

should expect a critical period for language acquisition to manifest itself through a pattern that can be visualized as a 'stretched Z' (see Birdsong, e.g. 2006; Granena & Long 2013), that is, through:

1. a relatively short-lived plateau of maturationally unconstrained development in early childhood, evidenced by uniformly high (though possibly weakly declining) levels of UA across these early AOs;
2. a period of increasingly constrained development in later childhood, even adolescence, evidenced by a distinct negative AO–UA correlation with a steeper gradient; and
3. a fully constrained language learning mechanism after the closure of the critical period, evidenced by an entirely absent or (more likely) retained but markedly weakened decline caused by factors other than maturation.

By way of illustration, let us first return to the Abrahamsson & Hyltenstam (2009) data referred to in the previous section, and the inverse relation between AO and UA in terms of scores of perceived nativelikeness. When the overall AO–UA comparison between the early (i.e. AO $\leq$ 11) and late (i.e. $\geq$ 12) learners was broken down into a comparison between five smaller cohorts, the 'stretched Z' pattern emerged (see the left panel in Figure 1.1).[6] First, there was a statistically non-significant difference in mean scores between the early childhood group (AO 1–5) and the late childhood group (AO 6–11), suggesting a fairly stable but possibly mildly sloping plateau of UA through the childhood years.[7] A change in slope then occurred, resulting in a steep gradient through the adolescent years, as evidenced by statistically significant differences between the adolescent group (AO 12–17) and adjacent AO groups in both directions. Finally, the slope flattened out, as evidenced by the non-significance of the apparent difference between the early (AO 18–23) and late (AO 24–47) adulthood groups. The pattern for scrutinized nativelikeness (see the right panel in Figure 1.1) was similar: the two childhood groups did not differ in mean scores, but both differed significantly from the adolescent/adult group, resulting in a discontinuous AO–UA function. As there were no learners beyond AO 19 in this part of the analysis, the slope did not reach any flattening stage; therefore, the expected 'stretched Z' materializes instead as a 'stretched 7' pattern, and whether the slope would have dropped, stayed the same or levelled out beyond this point can only be speculated on.

[6] All group comparisons were made with ANOVA and Fisher's LSD post hoc test.

[7] We are aware of Vanhove's (2013: 6–7) critique of our binning AOs together and that non-significant differences between adjacent cohorts may not indicate plateaus or flattened gradients; in fact, we leave open the possibility that what looks like an initial plateau in Figure 1.1 (left panel) may be a slope – in fact, perhaps even with the same gradient as the following slope. However, the flattened curve between the two oldest AO groups is real, obviously, as it cannot possibly continue to decline below zero (where 'zero' means 'passed for native speaker by 0 of 10 native judges').

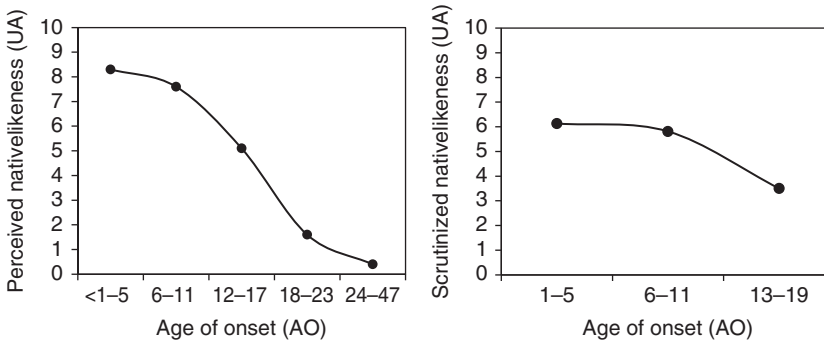

Figure 1.1 The 'stretched Z' and 'stretched 7' patterns (based on data and graphs in Abrahamsson & Hyltenstam 2009).

The initial peak of entirely unconstrained development, as described in (1), above, has been debated, however, and is not always considered a mandatory feature of critical periods (see discussions in e.g. Birdsong 2006; Bornstein 1989; Bruer 2008; Colombo 1982; Long 2013; Munro & Mann 2005). Maturational constraints may be increasingly operative from birth (or earlier), and the initial plateau of the AO–UA curve demonstrated in many studies (e.g. Flege et al. 1999; Granena & Long 2013; Johnson & Newport 1989) may even be interpreted in terms of ceiling effects caused by insufficiently sensitive instruments that have failed to differentiate between degrees of 'non-perceivable non-nativeness' among the earliest learners (see discussion in Hyltenstam & Abrahamsson 2000, 2003a; see, specifically, the discussion in Section 1.4). However, a requirement for claims of a critical period is a discontinuity of the AO–UA function that demarcates the completion of maturation, as described in (2) and (3), above. Rather than a 'stretched Z' or 'stretched 7', this shape of the AO–UA function could thus be depicted as a 'stretched L'. To illustrate, let us return to the data in Abrahamsson (2012). When splitting the overall correlation into two, it is obvious that the inverse AO–UA relation is located at the earlier half of the AO range: the coefficients for the AO 1–15 group were fairly strong (between −.57 and −.60 for GJT and between −.49 and −.53 for VOT when other independent variables were partialled out) but absent for the AO 16–30 group (nonsignificant correlations between −.05 and −.11 for GJT and between −.04 and −.19 for VOT).

To be able to compare the AO–UA function of Abrahamsson (2012) with that of Abrahamsson & Hyltenstam (2009), and for the specific purposes of the present chapter, we have here converted the AO span into approximately the same cohorts as in Figure 1.1 above. As can be seen from Figure 1.2, a clear

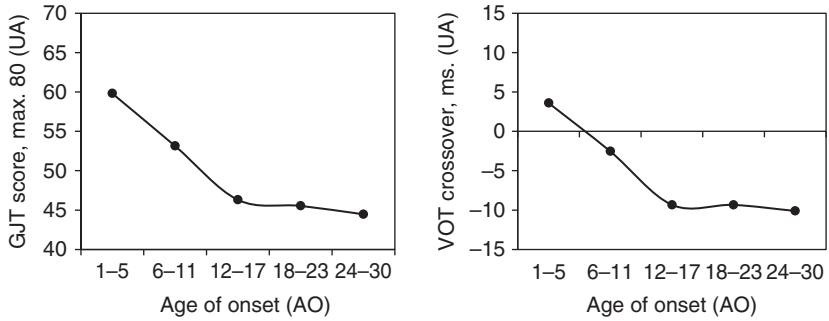

Figure 1.2 The 'stretched L' pattern (based on data reported in Abrahamsson 2012).

discontinuity can be observed between the steep gradient across the three earliest learner groups (all differences in mean scores are statistically significant) and the flat function across the later groups (no significant group differences), and this 'stretched L' pattern applies to both GJT and VOT.[8] The slopes indicate that the UA of grammatical and phonetic intuition decreases throughout childhood up until the early or mid-teens; after that, GJT scores remain stable just above chance level, and the VOT crossover for /p/ and /b/ levels out at approximately −10 milliseconds on the voicing continuum.

Another discontinuity feature, closely related to the differential correlational patterns across the AO range discussed above, is the individual variational patterns of early and late learners. As has been shown by numerous studies (e.g. DeKeyser 2000; Flege et al. 1999; Johnson & Newport 1989), the inter-learner variability in ultimate outcomes tends to be narrow across the earliest AOs but increasingly broader throughout later childhood and into the mid-teens; thereafter, the variation in UA is maximal and the scores are unrelated (or only weakly related) to AO, and must instead be attributed to individual differences in social and psychological circumstances (e.g. aptitude, motivation, education, input, general cognitive abilities, LOR in the L2 setting, and identity formation). As an illustration (and leaving aside at this point the obvious fallacies of self-assessments), consider the scatter plot in Figure 1.3 of the relation between AO and the self-assessed command of L2 Swedish by the 200 learners in Abrahamsson (2012).[9] In a pen-and-paper task, the learners assessed their

[8] These analyses are presented here for the first time, and, again, the reader is referred to Abrahamsson (2012: 195–199) for details on participants and methodology. All group comparisons were performed with ANOVA and Tukey's post hoc test.

[9] These self-assessment data are reported here for the first time, and the reader is referred to Abrahamsson (2012: 195–199) for details on participants and methodology.

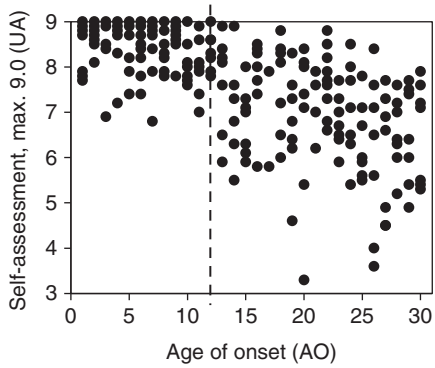

Figure 1.3 The 'megaphone' pattern of variability (based on original and previously unpublished data from the participants in Abrahamsson 2012).

command of ten phonological, grammatical and lexical aspects of L2 Swedish on a nine-grade scale, where '9' had been explicitly labelled 'like a native speaker of Swedish'. First, the AO–UA correlation turns out to be strong across the whole AO 1–30 range ($r = -.63, p < .001$), moderate to strong across the AO 1–15 range ($r = -.50$, $p < .001$), and weak to moderate across the AO 16–30 range ($r = -.32$, $p < .002$). Second, a visual inspection of the distribution indicates that self-assessment scores vary between 7 and 9 (with a tendency to cluster at ceiling) for AO 1 through 12, while becoming increasingly variable (between scores 3 and 9) thereafter, resulting in the characteristic megaphone-shaped pattern demonstrated already by Johnson & Newport (1989: 80; and even more clearly visualized by Pulvermüller & Schumann's [1994: 683] reproduction of the original graph).

The focus on individual UAs also makes it possible to employ yet another falsification criterion, namely the potential existence of late/adult learners who exhibit nativelike L2 proficiency. The visual inspection of the self-assessment scores in Figure 1.3 suggests not only that AO 12 delimitates the narrow variation among early learners from the clearly wider variation among later learners (as highlighted by the dashed vertical line), but also that this particular AO seems to be the last picket of self-assessed nativelike UA of Swedish; no individual learner with AO beyond 12 assessed their own proficiency as entirely nativelike (i.e. corresponding to a score of 9.0), which is a result consistent with the CPH.

Age turns out to play a role not only for the acquisition of language skills, but also for the retention of them. Findings from the fields of L1 attrition and heritage language acquisition have provided robust evidence that the language abilities of younger speakers are more vulnerable to loss than the abilities of

older speakers: children who experience a reduction, or even a complete cutoff, in L1 contact exhibit greater variability in their linguistic repertoires than adults (Bylund 2009b; Hyltenstam, Bylund, Abrahamsson & Park 2009; Park 2015; Pierce, Klein, Chen, Delcenserie & Genesee 2014; Schmid 2002). Notably, the age function found in situations of language loss and incomplete acquisition has important parallels to findings on age effects in L2 acquisition, and ultimately has important implications for our understanding of critical periods in language development (for extensive reviews and discussions, see Bylund 2009a, in press).

1.4 Nativelikeness as Falsification Criterion

We have thus far limited our discussion to the relative achievements of early versus late language learners, and have referred only in passing to their ultimate proficiency in absolute terms; that is, as compared to native-speaker proficiency. While young L2 learners' superior long-term achievements, including their apparent ability to reach nativelike UA, are indeed highly compatible with a maturational constraints account and the CPH, so is 'the apparent *in*ability of older learners to attain native-like proficiency' (Long 1990: 274; emphasis added). Therefore, one prevailing way to test the CPH has been to focus on the incidence of nativelike UA among late/adult L2 learners. Here, the core of the debate has been to what extent – or *if* – such learners exist, and how they should be relevant to the CPH. Different estimates of nativelike adult learners have been suggested in the literature, Selinker's 'perhaps a mere 5%' (1972: 212) being the most cited one. However, empirical studies vary in their reports, from relatively high occurrences (between 10 and 20 per cent) to no occurrences at all, depending on what learner samples were investigated, what linguistic structures or domains were analysed, what elicitation techniques were employed, or what definitions of nativelikeness were adopted (for an overview of studies and reported nativelikeness rates, see Abrahamsson & Hyltenstam 2009: 252–258).

An example of (what we see as) extreme overestimation is the nativelikeness rate given by Piller (2002), who reported that, among her adult-learner interviewees, '27 out of 73 individuals claimed they had achieved high-level proficiency in their L2 and could pass for native speakers in certain contexts' (p. 186), the conclusion being that these learners are certainly not a negligible minority!' (Piller 2002: 186). Although Piller's ethnographic insider perspective on high-level proficiency may prove to be a promising one (potentially illuminating many as yet unsolved mysteries of seemingly nativelike adult L2 learners), her quantification – approximately 37 per cent – is problematic insofar as it is based entirely on the self-assessed ability to pass for a native speaker 'on occasion' and 'in

certain contexts' (p. 186), by which is meant 'in a limited set of everyday service interactions'. In fact, the participants themselves described their ability to pass for a native speaker as 'a temporary, context-, audience-, and medium-specific performance' (p. 179), which, of course, is too loose a definition of nativelikeness if the goal is to test the CPH.

When adopting the much stricter and, as we see it, only relevant definition of nativelikeness as objectively measured by linguistic instruments, our own data point in the complete opposite direction, a zero occurrence of nativelike late learners being the outcome. For example, when the GJT and VOT results of the 200 learners in Abrahamsson (2012) referred to above were compared to the results of 20 native-speaker controls, it was shown that while more than half (56 per cent) of the participants with AO $\leq$ 6 performed within the range of native speakers on both measures, *no* participant with an AO beyond 13 did. Furthermore, only 2 (ca. 5 per cent) of the 39 learners with AO $\leq$ 6, but more than half (48–55 per cent) of those with AO 11–30, were *non*-nativelike on both measures (the other half being non-nativelike on one of the tests).

However, although these and similar findings (see e.g. Bialystok & Miller 1999; Flege et al. 1999; Johnson & Newport 1989) are fully consistent with what the CPH would predict, the only conclusion that can be drawn from normally distributed learner samples is that nativelike UA is an extremely rare phenomenon in the adult L2 learner population; the results do *not*, however, constitute evidence for the claim that 'all late learners are non-nativelike'. Therefore, in accordance with strict Popperian falsification, Long (1990, 1993) suggested that researchers' focus should be directed to finding the exceptions rather than identifying even more clearly non-nativelike adult learners, the rationale being that one single post-critical period learner with an ultimate proficiency indistinguishable from that of native speakers 'would serve to refute the claim' (1990: 255) – and, ultimately, the CPH.

This research agenda has been solidified by Long (1990, 1993, 2013) as well as by ourselves (e.g. Abrahamsson & Hyltenstam 2009; Hyltenstam & Abrahamsson 2003a) with various suggestions of how to enhance empirical stringency and precision of future research, of which we mention only three here: 1. the screening of highly advanced participants for studies (including only participants whom native speakers consider potentially nativelike); 2. measurements of L2 proficiency 'across the board' of domains and modes (i.e. phonetics/phonology, morphology, syntax, lexis, pragmatics, etc., in both production and perception); and 3. the use of instruments that are sensitive enough to distinguish between near-native and nativelike UA as well as between different degrees of near-native proficiency (specifically, through demanding tests and detailed analyses). Before moving on to the outcomes of our studies, let us first comment on how these methodological requirements have been elaborated on and incorporated into our research programme.

First, screening procedures as a basis for participant selection were employed in different ways. For example, the 24 academically and linguistically successful L2 learners of Swedish (with AO 3–12) in Hyltenstam (1992) had been pre-selected by their senior high-school teachers on the criterion of not being immediately identifiable as non-native speakers through their everyday spoken Swedish. Furthermore, the 20 near-native L2 speakers of Swedish (with AO 4–23) in Hyltenstam & Abrahamsson (2003b) were screened before the study in three steps: first, through self-selection or a word-of-mouth method; then, through initial telephone interviews conducted by a project assistant who, except for being unaware of the specific research aims, was instructed to report to us a preliminary evaluation of the candidates' nativelikeness; and, finally, through confirmatory follow-up telephone interviews, conducted by one of the researchers, with only those candidates whom the assistant had deemed nativelike (or near-native). However, the most rigorous and most objective screening procedures were employed in Abrahamsson & Hyltenstam (2009) (described briefly in Section 1.2). Here, the selection process was carried out in two steps. First, a pool of 195 L2 speakers of Swedish (with AO < 1–47) was collected, consisting of candidates who had responded to prominent advertisements in freely distributed newspapers in the Stockholm area, in which we had explicitly stated that we sought L2 speakers – adult learners in particular – who experienced typically passing-for-native speakers in everyday communication. These self-selected candidates were then (after consent) recorded over the telephone while talking freely for one minute about a specified but common topic. The speech samples, along with samples from 20 native control speakers, were then used in formal screening sessions, where 10 native, linguistically naïve listeners assessed each sample with respect to whether they believed it was produced by a native or a non-native speaker of Swedish. Of the 88 candidates with AO ≥ 12, only 5 were deemed native to the same extent as the 20 native controls (i.e. by at least 9 of the 10 judges); however, 17 late-learner candidates were believed by a majority (i.e. at least 6) of the 10 judges to be a native speaker; of these, 10 were eventually selected (along with 31 easily selected participants with AO ≤ 11 and 15 native controls) for participation in the second part of the study, in which their actual, linguistic UA was scrutinized in detail (see below).

The second recurrent methodological feature that we have incorporated is the broad and global view on what should constitute the UA construct. In accordance with Long (1990, 1993), we have repeatedly argued for an approach that allows for the thorough investigation of learners' ultimate proficiency 'across the board' of L2 structures, phenomena, domains and modes, rather than the hitherto predominant approach, which focuses on limited areas of learners' UA. Hyltenstam (1992) conducted quantitative and qualitative error analyses of the learners' oral and written production,

focusing on grammatical as well as lexical errors, while Hyltenstam & Abrahamsson (2003b) reported on the combined results of three tests of grammatical intuition, grammatical and semantic inferencing, and perceptual abilities in white noise. Again, however, the Abrahamsson & Hyltenstam (2009) study exceeds previous research in terms of the number and variety of measures, including phonetics/phonology (production and perception of VOT, three places of articulation), grammatical intuition (one auditory and one written GJT, covering four different morphosyntactic target features), grammatical processing (as measured by latency times of the auditory GJT), grammatical and semantic inferencing (a written cloze test), phonological perception in noise (highly frequent disyllabic words presented in escalating levels of babble noise), sentence perception in noise (white noise, three levels of speech-to-noise ratios) and lexis (one test of idiomatic expressions and one test of proverbs).

Finally, on top of screening procedures and across-the-board analyses, studies that aspire to falsify the claim that adult learners cannot become native-like need to employ linguistic instruments that are sufficiently sensitive to be able also to disclose very subtle, non-native features that are imperceptible to the naked ear or that cannot be detected through crude testing techniques. Sorace & Robertson (2001) capture this fact when arguing that 'non-native grammars may exhibit certain subtle features that distinguish them from native grammars' (p. 266), suggesting that 'the empirical investigation of non-native grammars requires sophisticated techniques that are sensitive enough to capture non-overt states [of non-nativeness]' (Sorace & Robertson 2001: 266). We previously created the term *non-perceivable non-nativeness* to denote such non-overt states of non-nativeness that cannot easily be teased apart from states of true nativeness (Hyltenstam & Abrahamsson 2003a), and we have argued that the tests and tasks of nativelikeness studies must be demanding even for very competent – including native – speakers, in order to minimize the risk of ceiling effects and unwarranted claims of nativelikeness. Hyltenstam (1992) conducted the grammatical and lexical error analyses with great precision, revealing extremely low but still significantly different error rates between native, early and late learners, whereas Hyltenstam & Abrahamsson's (2003b) tests were demanding in different ways: the GJT included only structures known to be difficult for L2 learners of Swedish, the cloze test was based on an informationally dense text with every seventh word mechanically removed, and the perception test comprised sentence perception and repetition during escalating white noise (see description below). The Abrahamsson & Hyltenstam (2009) study, finally, included both demanding tests, such as the 80-item GJT that, again, was based on structures known to be difficult for learners and presented in a morphosyntactically and lexically challenging context (i.e. with relatively long sentences, including relatively long and

infrequent words,[10] obligatory syntactic subordination, and with co-referent grammatical constituents kept distantly apart), tasks that tapped with some precision into the participants' fine-phonetic knowledge (e.g. production and categorical perception of VOT, phonological perception in noise, etc.), and tests of idiomaticity of fixed phrases, expressions and proverbs.

So, the crucial question, then, is: did we ever find the single Popperian black swan that would falsify the claim that all swans are white? The answer is a straight *no*, as we, in study after study, have consistently failed to identify Long's (1990) single adult nativelike learner (indistinguishable from a native speaker) that would serve to falsify the CPH and the existence of maturational constraints. First, Hyltenstam's (1992) native control speakers made 1–10 errors, while the earliest L2 learners (AO ≤ 6) made 1–23 errors and the later L2 learners (AO ≥ 7) 13–26 errors; in other words, although there were clear overlaps between native speakers and early learners and between earlier and later learners, there was no overlap in error rates between the native-speaker group and the later-learner group, meaning that no learner with AO 7 or beyond could be identified who performed within the range of native speakers. Secondly, Hyltenstam & Abrahamsson (2003b) failed to produce one single late learner with an overall nativelike command of L2 Swedish; in fact, *none* of the 20 near-natives had results within the range of the native-speaker controls on all three measures; not even among those with low AOs, although 6 learners (including 2 learners with AO 19 and 23, respectively) were nativelike on two measures, some being extremely close on the third. Finally, the Abrahamsson & Hyltenstam (2009) study revealed that one late learner with AO 19 performed within the range of native controls on seven out of the ten linguistic measures, one with AO 17 performed like a native speaker on six measures, while the remaining eight late learners were nativelike on five or fewer measures; the early learners fell within the native range on between two and ten of the measures, although only three individuals were nativelike on all ten measures (their AOs were 3, 7 and 8). Similarly, the detailed analysis of these learners' VOT, reported in Stölten et al. (2014, 2015), revealed that while half of the early learners had nativelike command of all three places of articulation, in both production and perception, *none* of the late learners was nativelike across the board of stops and modes.

Backed by these results, and given that we still subscribe to Long's (1990, 1993) research agenda, one would have thought that the maturational constraints controversy was settled once and for all, were it not for – as it turns out – the controversial result that child learners also exhibit proficiencies outside the native-speaker range. Although early learners with less than nativelike UA have indeed been reported in previous research, usually as single exceptions

[10] Although within the limits of normal adult comprehension.

(cf. Butler 2000; Ioup 1989; Obler 1989), the non-nativeness of (sometimes a majority of) our early learners puts the nativelikeness criterion into perspective. That child learners typically end up as *near*-native rather than nativelike speakers contrasts starkly not only with the layman's experience of child-learner superiority, but also with the predictions of the CPH (at least, the Lennebergian version of it). However, we would like to assert that, on the contrary, differences between near-native and nativelike language behaviour, regardless of age of acquisition, are *definitely* to be expected when the level of linguistic scrutiny is high. Consider, for example, the test of sentence perception in escalating white noise used in Hyltenstam & Abrahamsson (2003b). The task consisted of a recorded coherent text of 29 sentence or phrase units that the participants were to repeat *verbatim*. After the first seven noise-free sentences, successively louder white noise was added after every seventh sentence/phrase unit, until the speech-to-noise ratio was only 5 dB (i.e. almost 1:1), making the speech signal extremely difficult to perceive. As can be seen in Figure 1.4, all participant groups coped fairly well when the speech signal was 15 dB or 10 dB above the noise signal, the native speakers and the AO 4–5 group performing equally well (almost at ceiling with 94–97 per cent correct repetitions). However, at the last noise level, the capability of all learner groups – even the youngest learners – dropped considerably (to only 30–50 per cent correct repetitions), while the native-speaker group remained relatively capable (although they, too, began to experience some difficulty). The point to be made here is that the earliest learners behaved exactly like native speakers – but only up until a certain provoking point, after which their hitherto non-perceivable non-nativeness became observable, and without this last discriminatory noise level, unwarranted claims of entirely nativelike perceptual abilities in early learners would have been made. Our own

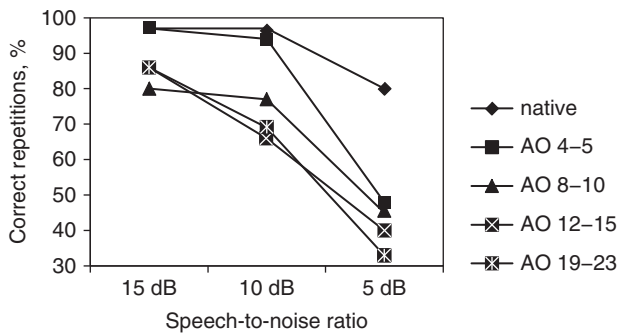

Figure 1.4 Sentence perception in white noise, three speech-to-noise ratios (based on data in Hyltenstam & Abrahamsson 2003b).

interpretation of these and similar data has been that maturational constraints are at work already from birth, and, therefore, that a fully nativelike UA cannot be expected if L2 exposure is delayed, if only minimally so. This view is fully in line with the 'stretched L' pattern discussed earlier, and would explain the initial peak or plateau phase of the 'stretched Z' (or 'stretched 7') pattern observed in studies of lesser scrutiny as a mere ceiling effect (as suggested already by Johnson & Newport 1989; see also the discussion in Hyltenstam & Abrahamsson 2003a: 566–576).

However, the nativelikeness criterion *as such* has been under attack for quite some time now – from the anti-critical period side as well as from the pro-critical period side of the fence. One (anti-CPH) interpretative framework that has gained popularity in recent years is cross-linguistic in nature and rests on the argument that near-native rather than fully nativelike L2 attainment is to be expected, even in very early learners – not because of reduced brain plasticity or as a function of delayed exposure, but simply because L2 learners are bilingual. According to this view, the comparison with monolingual speakers becomes theoretically moot, and it is argued instead that the comparative standard should be the simultaneously acquired bilingual proficiency of 'crib bilinguals' (see e.g. Birdsong 2005; Cook 1999, 2002; Muñoz & Singleton 2011; Ortega 2010, 2013).

A diametrically different (pro-CPH) critique of the nativelikeness criterion rests on the argument that the (possibly) nativelike behaviour of some exceptional advanced late learners is in itself uninformative, imposing no threat to the CPH. This anti-Popperian argument was already put forward by Lamendella (1977), who stated that

> the strawman form of Lenneberg's critical period hypothesis is unreasonable since it fails to take into account *individual differences* in an important respect. If a given person received the first exposure to language at age 26 and learned language in a normal fashion, this one case would not disconfirm the critical period hypothesis as a valid generalization for our species, but might merely reflect the fact that this individual remained more 'plastic' than the average member of our species. Lenneberg's formulation of the critical period hypothesis was actually loose enough to allow for such individual variation, but in the retelling this fact has often been obscured. (Lamendella 1977: 170; emphasis in the original)[11]

[11] Similar explanations of nativelike adult learners as biological exceptions have indeed been discussed by others: Selinker (1972) hypothesized that nativelike adult learners 'have somehow reactivated the latent language structure which Lenneberg describes' (p. 2012); Carroll (1973) suggested that some learners 'have for some reason lost little of their [innate] language acquisition ability' (p. 6); Ioup, Boustagui, El Tigi & Moselle (1994) speculated that in some exceptional learners the neurocognitive change associated with a critical period has not taken place in the usual way; Pulvermüller & Schumann (1994) suggested some individual variation in degree of myelination, making the learning mechanism in some adults continuously apt for language acquisition; in Bley-Vroman (1989), finally, adult nativelike proficiency was given a 'pathological status' (p. 44).

A more theoretically 'to-the-point' critique is represented by different versions of the so-called dual system approach, which actually – directly or indirectly – emphasizes the most prevalent feature of Lenneberg's original CPH, namely the distinction between children's 'automatic acquisition from mere exposure' and adults' 'learning through a conscious and labored effort' (Lenneberg 1967: 176). According to this view, it is entirely irrelevant for the CPH if adults should happen to exhibit nativelike linguistic behaviour, if it could be shown that this behaviour (and the competence underlying it) is handled (and has been arrived at) by a neurocognitive system that is entirely different from that operative in native speakers and early L2 learners. Here, the difference between near-native and native proficiency is thus not a mere quantitative issue, but a qualitative one.

Generally, then, the debate has centred on the validity of nativelikeness as falsification criterion, with one scholarly camp aiming at explaining the *non*-occurrence of nativelikeness without having to subscribe to the CPH, and one camp actually wanting to explain the *possible occurrence* of nativelike adults without having to reject the CPH. In essence, the issue boils down to whether the difference between near-native and native proficiency is actually a marginal quantitative difference that can be attributed to L1–L2 dominance, regulated by shifts in language use and exposure (advocated by the 'bilingualism effects' approach), or a major qualitative difference, originating in fundamentally different underlying learning mechanisms in children and adults (advocated by the 'dual system' approach). We will devote the remaining two sections of this chapter to a discussion of these main alternative theoretical takes on the nativelikeness issue, with examples of how they have been approached by our studies.

1.5 Non-Nativeness – An Effect of Maturation or Bilingualism?

In opposition to the idea of brain maturation rendering the relevant language processing areas less efficient, and thus compromising the attainment of nativelike proficiency, another explanatory framework has gained currency in recent years, the main argument of which is that the low incidence of nativelike attainment in L2 acquisition (demonstrated by e.g. Abrahamsson & Hyltenstam 2009; Granena & Long 2013) is due to the fact that L2 learners are bilingual. This core argument is found in a number of theoretical accounts. For instance, the *Multicompetence Framework* (Cook 1991) emphasizes the unique linguistic competence that emerges from the coexistence of two grammars in one mind. Similarly, the *Speech Learning Model* (Flege et al. 1999; Yeni-Komshian, Flege & Liu 2000) and the *Competition Framework* (MacWhinney 2005) hold that nativelike L2 proficiency is unattainable due to L1 entrenchment. The *Interference Hypothesis* (Pallier et al. 2003) takes this

reasoning one step further, positing that there is actually no loss in neural plasticity during the first decade of life, only a 'stabilization' of L1 neural networks, which can be 'reset if L1 exposure ceases, thus allowing for a nativelike L2 attainment (see also Venturerya et al. 2004) (for an extensive review, see Bylund, Hyltenstam & Abrahamsson 2013). While these frameworks and scholars may differ in the extent to which they invoke bilingualism effects as an explanation, the empirical basis of the argument is in part found in the evidence generated by the long tradition of studies on lexical and syntactic processing showing that bilingual speakers process language differently from monolingual speakers (e.g. Hernández, Bates & Ávila 1994; Kroll & Stewart 1994; Kroll & Tokowicz 2005; Vaughan-Evans, Kuipers, Thierry & Jones 2014). Thus, knowing that linguistic behaviour may differ between monolingual and bilingual speakers, does it really make sense to compare L2 learners who are bilingual with native speakers who are monolingual? And, crucially, is the non-nativelike (or near-native) behaviour attested among L2 learners more appropriately attributed to bilingualism effects rather than age of onset effects?

Even though the question of bilingualism effects *versus* age effects may have far-reaching implications for our understanding of the factors that underlie the incidence of non-nativelike proficiency in L2 speakers, there have been few attempts to address the situation empirically. Instead, bilingualism effects have functioned as an overarching interpretative frame for questioning the biological basis of non-nativelikeness. However, given the potential importance of bilingualism effects, the only way to move forward is to try to formulate specific testable hypotheses that allow us to disentangle age effects from bilingualism effects. Several theoretical accounts of bilingualism effects are based on the communicating-containers metaphor, whereby proficiency in the one language of the bilingual increases if proficiency in the other language decreases. A testable hypothesis would thus hold that individuals who are less proficient in their L1 will have greater chances of attaining nativelikeness in their L2, and vice versa. This reasoning could even be linked to the issue of age effects in L1 attrition: the vulnerability of L1 skills early in life may go hand in hand with the readiness for acquiring L2 skills, such that an individual may end up with nativelike proficiency in at least one of his/her languages, but not both.

Bylund, Abrahamsson & Hyltenstam (2012) conducted one of the first studies to test this particular prediction, investigating L1 Spanish – L2 Swedish adult functional bilinguals in Sweden (n = 30) whose AO of L2 acquisition ranged from 1 to 12 years of age. Two measures were used to assess proficiency in the participants' two languages: GJTs and cloze tests. The bilinguals' performances on these tests were then compared to those of functionally monolingual native groups of speakers of Spanish (living in Chile) and Swedish. The results revealed that at the group level, the bilinguals were outperformed by the monolingual counterparts in both languages and on

both measures. Taken at face value, this finding would suggest that bilingual – and by extension L2 speaker – proficiency is indeed different from monolingual proficiency, and comparisons between the two may be inappropriate. However, the analyses were then taken several steps further, starting with correlational analyses of individual scores. Here, it was found that L1 and L2 performances on the tests were positively correlated [$\Phi = .48$ and $\Phi = .67$, respectively]. More specifically, individuals who scored within the native-speaker range in one language were more likely to also score within the native range in the other language (be it the L1 or the L2). Likewise, bilinguals who exhibited scores outside the native range in one language were also more likely to do so in the other language. This finding is thus in stark contrast with the prediction that nativelike behaviour in one language is accompanied by non-nativelike behaviour in the other language. Next, Bylund et al. (2012) analysed whether the bilingual participants' performance in both languages could be predicted by factors usually associated with nativelike attainment, such as age of onset, frequency of language use and language aptitude. The results from a multinomial logistic regression showed that language aptitude was the only factor that could successfully predict the bilinguals' behaviour, in that high degrees of aptitude were associated with nativelikeness in both languages, and vice versa.

The findings reported in Bylund et al. (2012) add several dimensions of complexity to the age versus bilingualism effects debate. First, they highlight the importance of analysing scores at the individual level when assessing bilingual proficiency, to avoid overgeneralizations about bilinguals necessarily being different from monolingual speakers. Second, the findings show that individual differences, such as an elevated ability to handle language structure, may be a robust determinant of bilingual development. Crucially, this latter finding may introduce individual variation in bilingualism effects, since the possibility exists that the proficiency of the bilingual participants with low-level language aptitude is more likely to be affected by the addition of an extra language. Another important implication of the findings relates to causation and bilingualism effects. A previous study on global pronunciation in Korean–English bilinguals (Yeni-Komshian et al. 2000) reported inverse L1–L2 proficiency levels, suggesting that nativelike proficiency was possible in one language but not in both, adhering to the communicating-container metaphor. However, the fact that it is indeed possible to exhibit nativelikeness in two languages instead suggests that when such inverse relationships are found, they may be correlational rather than causational. Differences in input may instead, by and large, account for inversely related proficiency levels (for further discussion, see Bylund et al. 2013).

Another way to empirically address the possible existence of bilingualism effects in L2 acquisition consists of comparing multiple speaker groups with

different monolingual and bilingual upbringings. This method speaks particularly to the *Interference Hypothesis*, according to which nativelike L2 attainment is possible only if the L1 undergoes a complete loss and the neural networks thus 'reset' to the L2 (Pallier et al. 2003; Ventureyra et al. 2004). In an attempt to investigate this possibility, Norrman & Bylund (2016) assessed phonological perception in three groups of adult speakers: functionally monolingual speakers of L1 Swedish, functionally bilingual speakers of L2 Swedish and functionally monolingual speakers of L2 Swedish. This latter group comprised individuals who were adopted from Latin America to Sweden between the ages of 4 and 8 (matched with the bilingual speakers AOs). Members of this group, who had acquired Spanish as their L1, had according to self-reports undergone an apparently complete L1 loss, and instead led their everyday lives in L2 Swedish. The bilingual group used their L1 Spanish in about 25 per cent of their daily communication. The three groups thus differed in terms of bi-/monolingualism and age of L2 acquisition, as per Table 1.1.

Swedish proficiency was assessed by means of a lexical decision task, in which the contrast of long versus short vowels had been embedded. Using mixed-effects modelling, the results showed unequivocally that the two L2 speaker groups (i.e. the adoptees and the successive bilinguals) did not differ from each other in terms of accuracy and speed, but they were outperformed by the native-speaker control group. These findings thus suggest that the L1 loss experienced by the international adoptees had not afforded them any advantages in terms of L2 ultimate attainment. As such, the findings provide the empirical evidence needed to reject the Interference Hypothesis, which had previously been challenged on grounds of preliminary evidence only (Hyltenstam et al. 2009; for an overview, see Norrman, Hyltenstam & Bylund 2016). However, a potential counterargument here could be that the adoptees did not exhibit nativelike L2 proficiency because they still had L1 remnants that interfered with their acquisition process. Indeed, evidence is accumulating that the L1 is not completely eradicated in international adoptees, but isolated aspects of it may be reactivated through relearning (Hyltenstam et al. 2009; Oh, Au & Jun 2010; Park 2015; Pierce et al. 2014; Singh et al.

Table 1.1 *Design of Norrman & Bylund (2016)*

	Successive bilinguals	Adoptees	Controls
Swedish exposure from birth	–	–	+
Monolingualism	–	+	+

2011). The problem with such a counterargument, however, is that it does not attribute any proportions between retained L1 knowledge and attained L2 knowledge: while it is possible, or even likely, that the adoptees studied by Norrman and Bylund unconsciously had retained certain remnants of their L1, it is unreasonable that such a small proportion of remnants would yield the same interfering effect on L2 attainment as would the fluent L1 knowledge possessed by the functional bilingual group. Instead, a more probable factor contributing to their non-nativelike attainment is their delayed exposure to Swedish; in other words, their age of acquisition.

1.6 Fundamental Differences and the Dual System Approach

A general explanatory framework that has surfaced in various shapes over the years is one that stresses the different means by which learners approach the task of learning language, the most renowned distinction, of course, being that between 'acquisition' and 'learning' (as accentuated by Krashen, e.g. 1976). In a discussion of this dichotomy, Lamendella (1977) asserted that 'there is a difference in the nature and internal organization of the neurofunctional systems responsible, correlated with drastically different types of performance abilities in the target language' (p. 176). A recurrent assumption is that the preference for one type of learning over the other is regulated by age and the maturational state of the learner. In their early accounts, Penfield & Roberts (1959) suggested that age 9 marked a shift from 'direct learning from input' to 'analytical learning' (in the Piagetan sense) via the L1, while Lenneberg (1967) pointed to the general qualitative difference between 'automatic acquisition from mere exposure' before puberty and 'learning through a conscious effort' thereafter (reduced brain plasticity being the general cause). In Bley-Vroman's (1989) specific *Fundamental Difference Hypothesis*, originally formulated within the *Universal Grammar* paradigm (and with no specific reference to different neuroanatomical substrates), the relevant distinction was made between 'innate domain-specific mechanisms', available only to children during the critical period, and 'domain-general learning strategies' (part of the general cognitive system used for all kinds of learning), which is what the adult learner must rely on when learning L2 grammatical competence (cf. also DeKeyser 2000).

According to the more recent neurolinguistic versions of the dual system approach, as offered by Paradis (2004, 2009) and Ullman (e.g. 2005, 2015), early and late language development is governed by fundamentally different neurocognitive and neurophysiological systems: while children acquire language 'incidentally' and almost exclusively through *procedural memory* (resulting in *implicit knowledge*), adolescents and adults engage more in 'intentional' learning, relying primarily on *declarative memory* (resulting in

explicit knowledge). The procedural memory system handles the acquisition and control of perceptual and motor skills as well as automatized cognitive abilities, and is associated with rule-governed, hierarchically and sequentially structured, probabilistic and routine-based knowledge, which for language encompasses, for example, articulation, grammatical and phonological structure and regularities, including grammatical aspects of the lexicon. By contrast, the declarative memory system is semantic and episodic in nature and specializes in the learning and knowledge of fact-based information and in the formation of associations between (more or less) unrelated parts of this information; in terms of language, this system hosts vocabulary (i.e. the phonological forms and semantics of lexical items), idioms and irregular grammatical forms. Acquisition via the procedural memory system occurs through long-term practice and immense and continuous exposure, and results in implicit (i.e. intuitive) abilities that are manifested through unconscious, automatic and relatively fast processes, while explicit knowledge, handled by declarative memory, is learned fast and operates mainly on the basis of conscious, willed and relatively slow retrieval mechanisms. Furthermore, the two systems have distinct neuroanatomical substrates, procedural memory being associated with the basal ganglia, the cerebellum and specific frontal cortical areas in the left hemisphere, especially (parts of) Broca's area and premotor cortex, whereas declarative memory is subserved by hippocampus and cortical structures in the temporal lobe. Crucially, while the procedural memory system is fully operative approximately up to age 5, after which it attenuates and becomes less effective, the declarative memory system develops gradually during childhood and is most effective during adolescence and early adulthood. Studies of language impairments and pathologies show that the declarative system can effectively take over a large number of functions that can no longer be handled by a severely deteriorated or attenuated procedural system (see e.g. Ullman & Pullman 2015), and it is assumed that this compensatory role for declarative memory applies to adult L2 learning as well. In resonance with the CPH, then, the main prediction for L2 acquisition holds that the procedural memory system is fully available to the young child but less so to the adult learner, who must instead cope by using declarative memory even when acquiring knowledge outside the realm of this system, such as implicit morphosyntactic and phonological structures and rules.

With neuroimaging research on near-native learners still pending, behavioural data on the compensatory role of explicit learning through declarative memory are offered by studies on the relation between UA and language *aptitude*. Late learners with a near-native (even nativelike) performance on aspects of the L2 seem to draw heavily on general cognitive learning abilities (i.e. declarative memory), as evidenced by their above average (sometimes outstanding) performance on standardized aptitude tests (e.g. DeKeyser

2000).[12] To illustrate, Abrahamsson & Hyltenstam (2008) – when investigating the participants subsequently scrutinized in Abrahamsson & Hyltenstam (2009; see above) – showed that *all* the late near-native learners, all being perceived as native speakers of Swedish by native judges, had above-average scores on the *Swansea Language Aptitude Test* (Meara, Milton & Lorenzo-Dus 2003); the early learners, on the other hand, and the native control speakers, were normally distributed with regard to aptitude scores. In fact, *none* of the late learners had a score below the early-learner average. As high language aptitude scores are indicative of an enhanced ability to reflect (explicitly) on linguistic structure, it may be concluded that, to pass for a native speaker by native listeners, a heightened declarative memory function is a necessary compensatory condition for adult learners. In addition, the study demonstrated a certain role for aptitude also in early learners: An analysis of their performance on the demanding GJT (see above) revealed that among those early learners who scored below the range of native speakers (about half of the early learners), the majority also exhibited below-average aptitude scores, whereas a high aptitude score was typical for those with GJT scores within the native-speaker range.

While language aptitude has been intensely investigated in the field of second language acquisition, it only recently entered the L1 attrition research agenda. In Bylund, Abrahamsson & Hyltenstam (2010), we investigated the impact of language aptitude on 25 child attriters and incomplete learners (AO 1–11) and found that high degrees of aptitude correlated with the ability to correctly judge L1 grammaticality (r = .52). This study also showed that aptitude may not only predict L1 proficiency in situations of reduced contact, but it also regulates the dependence on contact: while no correlation was found between L1 proficiency and self-reported L1 contact across the participant group as a whole, we did establish a significant positive correlation (r = .60) between these two measures among participants whose language aptitude was below average, suggesting that speakers with low degrees of aptitude are more dependent on L1 contact for L1 maintenance/development than are speakers with high degrees of aptitude.[13]

Another feasible way to approximate the dual system hypothesis is to focus on the parallel development of different language domains and components. According to Paradis (2009), 'the availability of procedural memory for acquiring language *as a whole* decreases with age' (p. 24; emphasis added), which can be taken to mean that, while child learners – unconsciously, incidentally and holistically – approach the different levels and sublevels of the target

[12] For discussions on the aptitude construct and how it is operationalized, see Abrahamsson & Hyltenstam (2008); Granena (2013a, b); or Hyltenstam (Chapter 7 of this volume).

[13] Other less researched factors that would compensate for attrition susceptibility are linguistic attitudes and heritage language education (see Bylund & Díaz 2012).

language as interdependent parts of a system, adult learners tend to treat them as independent puzzles. This hypothesis was tested in Abrahamsson (2012) by way of correlations between the two measures of grammatical and phonetic UA. It was shown that, while GJT scores and VOT scores correlated positively among the early (AO 1–15) learners (r = .44, p < .001), they were entirely unrelated among the late (AO 16–30) learners (r = .09, p = .80, n.s.), which possibly indicates that language learning is more sporadic, unsystematic and fragmented in adults, whereas children automatically develop all aspects of all linguistic levels in parallel.

As we see it, the dual system hypothesis is an issue that should have the highest priority in future research, which is why we return to a discussion of it in the following Conclusions section.

1.7 Conclusions

The aim of this chapter was to show in which ways our research on language acquisition and loss under conditions of early and late language exposure, as well as under conditions of monolingualism and bilingualism, contributes to our understanding of maturational constraints and the critical period for language. The research that was reviewed includes comprehensive designs with a strict selection of participants and native controls, a large battery of testing techniques with demanding language tasks covering a variety of linguistic domains and modes, as well as measures of language aptitude. When summarizing the results of our now-25-year research programme in light of the relevant falsification criteria, it becomes obvious that we have failed to provide the evidence necessary to reject the CPH. Specifically, our results have not shown that AO effects reflect primarily experiential and socio-psychological (i.e. non-maturational) variables; we have not been able to demonstrate a linear AO–UA function that steadily declines across (and way beyond) any alleged critical cutoff ages; we have been unsuccessful in identifying the one single adult learner with an L2 ability indistinguishable from that of native speakers; our data do not indicate that bilingualism, rather than delayed language exposure, is what prevents (both early and late) L2 learners from attaining entirely native-like proficiency, or that monolingualism is a beneficial learning condition; and we have failed to produce evidence that children and adults develop language through the same (implicit/procedural) learning mechanisms. On the contrary, our data clearly suggest:

1. that AO is the strongest predictor of UA, robustly accounting for the lion's share of the variance when the contribution from other variables (e.g. length of residence, L1 use, chronological age, etc.) are partialled out, suggesting that AO potentially conceals the effects of learners' maturational states;

2. that the curve describing the AO–UA function is characterized by discontinuity at theoretically motivated cutoff points (around AO 6–7 and AO 12–13), regardless of whether the function concerns sensitivity to L2 exposure or susceptibility to L1 attrition;
3. that adult learners who are being (subjectively) perceived as nativelike by native speakers actually turn out instead to be *near*-native when their L2 abilities are being (objectively) scrutinized with linguistic instruments;
4. that delayed exposure rather than bilingualism is the cause of non-/near-nativeness, and that monolingualism is *not* a beneficial circumstance (as evidenced by the near-nativeness of international adoptees); and
5. that adults learn and process the L2 through the compensatory use of explicit/declarative memory, evidenced by adult near-native learners' exceptional verbal analytical abilities as measured by aptitude tests and by adults' more scattered, inconsistent and nonparallel development of different L2 domains and abilities.

Even though we refrain from claiming that we have provided definitive evidence of maturational constraints or a critical period, we feel safe to say that the results of our research programme are entirely consonant with such an interpretation. Obviously, using our empirical results to tell the *alternative* narrative – that is, that factors other than AO best explain the variance, that the AO–UA function is linear across the lifespan, that nativelikeness in adults is possible (even a recurrent phenomenon), that bilingualism and language dominance (not age of exposure) is the inevitable cause of non-nativeness, and that children and adults learn language via the same implicit, procedural memory system – would make very little sense. Therefore, suffice it to say at this point that the CPH cannot be rejected on the basis of the results of our studies.

We would like to suggest that the issues in (4) and (5) above should receive top priority in future research. First, while the idea of bilingualism (and not delayed exposure) being the primary cause of the 'inevitable' non-/near-nativeness in both child and adult learners is intuitively appealing, it seems as if empirical enquiries into this issue so far have provided mixed evidence. Importantly, future studies should serve perhaps not to one-sidedly reject or confirm the existence of bilingualism effects, but rather to advance our understanding of the complex interaction of factors that underlie L2 ultimate attainment, on the one hand, and the potential selectivity of bilingualism effects, on the other hand. To the extent that bilingualism effects exist, they do not seem to occur across the board, and when found, a clear causal link may be missing. Large-scale studies including several types of monolingual and bilingual L1 and L2 speakers, covering diverse aspects of linguistic knowledge, are doubtlessly necessary in order to throw further light on these issues. We are currently finalizing such a study at Stockholm University, in which we compare the across-the-board proficiency of monolingual L1 speakers, simultaneous (or

'crib') bilinguals, sequential monolinguals (adoptees) and sequential bilinguals (immigrant child learners), the results of which suggest that bilingualism effects may be less pervasive than was previously thought.

Second, with a neurocognitively motivated distinction between procedural and declarative memory, and with language proficiency measures beyond accuracy, future research should focus on whether adult L2 learning involves the same implicit/procedural mechanisms that govern children's acquisition, or if adults rely instead largely on explicit/declarative mechanisms. With such a dual system approach we will be able to postulate a variety of testable hypotheses, one of the most urgent being whether the near-nativeness demonstrated in early learners is only quantitatively similar to the near-nativeness exhibited by some exceptional adult learners, or if they differ fundamentally in terms of quality, solidity and origin.

By way of illustration, consider the schematic comparison in Figure 1.5 of the age-determined roles of procedural and declarative memory in terms of their diverging emergence, peak efficiency and decline. Small children acquire language implicitly from exposure alone, and exclusively through the procedural memory system up to approximately age 5, after which this system begins to attenuate and the declarative system begins to develop (Ullman 2005). Older children thus experience a gradual shift from implicit acquisition through procedural memory to explicit learning through declarative memory.[14] By contrast, adolescents and young adults, while having to accept a severely attenuated and less efficient procedural system, have at their disposal a full-fledged declarative memory system, which necessarily takes on the task of second language learning; older adults will have to settle with learning language mainly through a declarative memory system that, except for not being optimized for the task, is slowly deteriorating throughout life. With this model, we could investigate the qualitative difference between early-learners' and late-learners' near-nativeness, for example, through *electroencephalography* (EEG) and *event-related potentials* (ERPs), with testable hypotheses concerning both type and location of brain processes. Furthermore, 'nativelike proficiency' or 'nativelike behaviour' – as measured by, for example, accuracy scores or reaction times on grammaticality judgment tests – becomes conceptually irrelevant; instead, the relevant conceptualization of nativelikeness would be in terms of 'nativelike acquisition' or 'nativelike processing'. Thus, if it could be verified that the quantitatively similar (sometimes identical) near-native proficiencies of early and late learners, in fact, rest on qualitatively different foundations, associated with fundamentally different neuroanatomic

[14] Paradis (2004, 2009) and Ullman (e.g. 2005, 2015) agree that all knowledge generated by procedural memory should be implicit in nature, but disagree as to whether all learning through declarative memory is explicit, with Ullman suggesting that declarative knowledge may be either implicit or explicit.

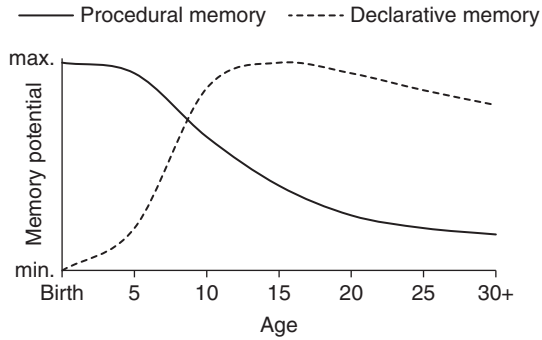

Figure 1.5 Emergence and development of procedural and declarative memory; based on discussions in Paradis (2004, 2009) and Ullman (e.g. 2005, 2015).

substrates, then we will have come extremely far in our search for the explanations for child–adult differences in L2 acquisition. A large-scale EEG/ERP study on the compensatory role of declarative memory in early and late near-native L2 acquisition is currently being launched at Stockholm University, through which we hope to shed light on this issue.

This model also potentially explains the seemingly continuous decline in UA over the lifespan reported in some studies. If older adults have only a steadily declining declarative memory system at their disposal for L2 learning, as can be inferred from Figure 1.5, then a linear AO–UA function across higher AOs should come as no surprise. This decline, however, is entirely unrelated to the rapid attenuation of implicit/procedural memory during later childhood.

Moreover, the dual system model – if it should turn out to be a valid one – would help us explain to language policymakers, school practitioners and private kindergarten entrepreneurs with an 'early foreign language' agenda why their ambition to take advantage of children's superior language learning skills is largely misguided. If implicit acquisition through procedural memory from natural language exposure alone is what young children are good at (and *only* that), and if older children and teenagers are more apt at learning through a fully developed declarative memory system and via explicit language instruction, it could be easily argued that a later introduction to foreign language instruction is to be preferred. The implicit acquisition of grammatical intuition through procedural memory is a long-term activity that requires immense and continuous exposure and practice, while the explicit learning of meta-linguistic knowledge through declarative memory can occur instantly and through explicit instruction. Few school contexts, if any, should be able to provide the quantity and quality of exposure necessary to take advantage of the

maturationally unconstrained abilities of young children. As foreign language learning relies heavily on explicit learning abilities, usually only during a few hours a week, the dual system model has the potential of explaining, quite pedagogically, too, why the most successful policy should be one that introduces foreign languages to older children and adolescents – not young children. For an extensive discussion on the misapplication of the CPH to the foreign language context, see Hyltenstam & Abrahamsson (2001).[15]

Finally, with a theoretically motivated and neurophysiologically well-grounded dual system hypothesis, as represented by the kindred models of Paradis (2004, 2009) and Ullman (e.g. 2005, 2015), we seem to have come full circle. Returning to where we started this chapter, the seemingly conflicting results between the earliest studies showing an initial, short-term advantage for older learners over children and studies showing the long-term superiority of children over older learners, as it happens, are not conflicting at all. As soberly argued by DeKeyser & Larson-Hall (2005), 'the two phenomena can both be explained by the same underlying difference in learning mechanisms: Children necessarily learn implicitly; adults necessarily learn largely explicitly' (p. 103). That is, when interpreted within a dual system framework of the kind presented here, ironically, the empirical findings that were once presented as the first evidence *against* the CPH turn out instead to constitute solid evidence *for* it.

Acknowledgements

Preparation of this chapter was supported by a grant to NA from The Bank of Sweden Tercentenary Foundation (grant no. SAB16-0051:1).

References

Abrahamsson, N. 2012. Age of onset and nativelike L2 ultimate attainment of morpho-syntactic and phonetic intuition. High-level L2 acquisition, learning, and use. Thematic issue of *Studies in Second Language Acquisition* 32/2: 187–214.

Abrahamsson, N. 2013. The critical period hypothesis (CPH), in Robinson, P. (ed.), *The Routledge Encyclopedia of Second Language Acquisition*. London: Routledge.

Abrahamsson, N. & Hyltenstam, K. 2008. The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition* 30/1: 481–509.

[15] For an early and insightful discussion on the difference between 'secondary language acquisition' and 'foreign language learning', see Lamendella (1977).

Abrahamsson, N. & Hyltenstam, K. 2009. Age of onset and nativelikeness in a second language: listener perception versus linguistic scrutiny. *Language Learning* 59/2: 249–306.

Ammerlaan, T. 1996. 'You get a bit wobbly. . .' – Exploring bilingual lexical retrieval processes in the context of first language attrition. PhD dissertation. Nijmegen: Katholieke Universiteit.

Andersen, R. W. 1982. Determining the linguistic attributes of language attrition, in Lambert, R. D. & Freed, B. F. (eds.), *The Loss of Language Skills*. Rowley: Newbury House Publishers. 83–118.

Asher, J. & García, G. 1969. The optimal age to learn a foreign language. *Modern Language Journal* 38: 334–341.

Asher, J. & Price, B. 1967. The learning strategy of the total physical response: Some age differences. *Child Development* 38: 1219–1227.

Bialystok, E. & Miller, B. 1999. The problem of age in second-language acquisition: Influences from language, structure, and task. *Bilingualism: Language and Cognition* 2: 127–145.

Birdsong, D. 1999. Introduction: whys and why nots of the critical period hypothesis for second language acquisition, in Birdsong, D. (ed.), *Second Language Acquisition and the Critical Period Hypothesis*. Mahwah, NJ: Erlbaum. 1–22.

Birdsong, D. 2005. Nativelikeness and non-nativelikeness in L2A research. *International Review of Applied Linguistics in Language Teaching (IRAL)* 43: 319–328.

Birdsong, D. 2006. Age and second language acquisition and processing: a selective overview. *Language Learning* 56: 9–49.

Birdsong, D. & Gertken, L. M. 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* 4/2: 107–133.

Bley-Vroman, R. 1989. What is the logical problem of foreign language learning? in Gass, S. & Schachter, J. (eds.), *Linguistic Perspectives on Second Language Acquisition*. Cambridge: Cambridge University Press. 41–68.

Bornstein, M. 1989. Sensitive periods in development: structural characteristics and causal interpretations. *Psychological Bulletin* 105: 179–197.

Bruer, J. T. 2008. Critical periods in second language learning: distinguishing phenomena from explanation, in Mody, M. & Silliman, E. (eds.), *Brain, Behavior and Learning in Language and Reading Disorders*. New York, NY: The Guilford Press. 72–96.

Butler, Y. G. 2000. The age effect in second language acquisition: is it too late to acquire native-level competence in a second language after the age of seven? in Oshima-Takane, Y. Shirai, Y. & Sirai, H. (eds.), *Studies in Language Sciences* 1. Tokyo: The Japanese Society for Language Sciences. 159–169.

Bylund, E. 2009a. Maturational constraints and first language attrition. *Language Learning* 59: 687–715.

Bylund, E. 2009b. Effects of age of L2 acquisition on L1 event conceptualization patterns. *Bilingualism: Language and Cognition* 12/3: 305–322.

Bylund, E. in press. Age effects in language attrition, in Schmid, M. & Köpke, B. (eds.), *Handbook in Language Attrition*. Oxford. Oxford University Press.

Bylund, E., Abrahamsson, N. & Hyltenstam, K. 2010. The role of language aptitude in first language attrition: the case of pre-pubescent attriters. *Applied Linguistics* 31/3: 443–464.

Bylund, E., Abrahamsson, N. & Hyltenstam, K. 2012. Does first language maintenance hamper nativelikeness in a second language? High-level L2 acquisition, learning, and use. Thematic issue of *Studies in Second Language Acquisition* 34/2: 215–241.

Bylund, E. & Díaz, M. 2012. The effects of heritage language instruction on first language proficiency: a psycholinguistic perspective. *International Journal of Bilingual Education and Bilingualism* 15/5: 593–609.

Bylund, E., Hyltenstam, K. & 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. 69–103.

Carroll, J. B. 1973. Implications of aptitude test research and psycholinguistic theory for foreign language teaching. *International Journal of Psycholingistics* 2: 5–14.

Chiswick, B. R. & Miller, P. W. 2008. A test of the critical period hypothesis for language learning. *Journal of Multilingual and Multicultural Development* 29/1: 16–29.

Choi, J., Broersma, M. & Cutler, A. 2017. Early phonology revealed by international adoptees' birth language retention. *Proceedings of the National Academy of Sciences of the United States of America, PNAS*. 11 July, 114/28: 7307–7312.

Colombo, J. 1982. The critical period concept: research, methodology, and theoretical concerns. *Psychological Bulletin* 91: 260–275.

Cook, V. J. 1991. The poverty-of-the-stimulus argument and multi-competence. *Second Language Research* 7/2: 103–117.

Cook, V. 1999. Going beyond the native speaker in language teaching. *TESOL Quarterly* 33: 185–209.

Cook, V. 2002. Background to the L2 user, in Cook, V. (ed.), *Portraits of the L2 user*. Clevedon: Multilingual Matters. 1–28.

Davies, A. 2003. *The Native Speaker: Myth and Reality*. Clevedon: Multilingual Matters.

DeKeyser, R. 2006. A critique of recent arguments against the critical period hypothesis, in Abello-Contesse, C., Chacón-Beltrán, R., López-Jiménez, M. D. & Torreblanca-López, M. M. (eds.), *Age in L2 Acquisition and Teaching*. Bern: Peter Lang. 49–58.

DeKeyser, R. M. 2000. The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22: 499–533.

DeKeyser, R. & Larson-Hall, J. 2005. What does the Critical Period really mean? in Kroll, J. F. & de Groot, A. M. B. (eds.), *Handbook of Bilingualism: Psycholinguistic Perspectives*. Oxford: Oxford University Press. 88–108.

Dewaele, J.-M. 2017. Why the dichotomy 'L1 versus LX user' is better than 'native versus non-native speaker'. Applied Linguistics, published online January 2017. DOI:https://doi.org/10.1093/applin/amw055.

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

Granena, G. 2013a. Cognitive aptitudes for second language learning and the LLAMA Language Aptitude Test, in Granena, G. & Long, M. (eds.), *Sensitive Periods,*

Language Aptitude, and Ultimate Attainment. Amsterdam: John Benjamins Publishing. 105–129.

Granena, G. 2013b. Reexamining the robustness of aptitude in second language acquisition, in Granena, G. & Long, M. (eds.), *Sensitive Periods, Language Aptitude, and Ultimate Attainment*. Amsterdam: John Benjamins Publishing. 179–204.

Granena, G. & Long, M. H. 2013. Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research* 29/3: 311–343.

Hakuta, K., Bialystok, E. & Wiley, E. 2003. Critical evidence: a test of the critical period hypothesis for second-language acquisition. *Psychological Science* 14/1: 31–38.

Hernández, A., Bates, E. & Ávila, L. 1994. On-line sentence interpretation in Spanish–English bilinguals: what does it mean to be 'in between'? *Applied Psycholinguistics* 15: 416–446.

Huang, B. H. 2014. The effects of age on second language grammar and speech production. *Journal of Psycholinguistic Research* 43: 397–420.

Hyltenstam, K. 1992. Non-native features of near-native speakers. On the ultimate attainment of childhood L2 learners, in Harris, R. J. (ed.), *Cognitive Processing in Bilinguals*. Amsterdam: Elsevier Science. 351–368.

Hyltenstam, K. 2010. Critical periods, in Hogan, P. C. (ed.), *The Cambridge Encyclopedia of the Language Sciences*. Cambridge: Cambridge University Press. 238–240.

Hyltenstam, K. 2012. Critical period, in Chapelle, C. (ed.), *Encyclopedia of Applied Linguistics*. Wiley-Blackwell.

Hyltenstam, K. 2018 (this volume, Chapter 7). Polyglotism – a synergy of abilities and predispositions, in Hyltenstam, K., Bartning, I. & Fant, L. (eds.), *High-Level Language Proficiency in Second Language and Multilingual Contexts*. Cambridge: Cambridge University Press.

Hyltenstam, K. & Abrahamsson, N. 2000. Who can become native-like in a second language? All, some, or none? On the maturational constraints controversy in second language acquisition. *Studia Linguistica* 54/2: 150–166

Hyltenstam, K. & Abrahamsson, N. 2001. Age and L2 learning: the hazards of matching practical 'implications' with theoretical 'facts'. (Comments on Stefka H. Marinova-Todd, D. Bradford Marshall, and Catherine E. Snow's 'Three misconceptions about age and L2 learning'). *TESOL Quarterly* 35/1: 151–170.

Hyltenstam, K. & Abrahamsson, N. 2003a. Maturational constraints in SLA, in Doughty, C. J. & Long, M. H. (eds.), *The Handbook of Second Language Acquisition*. Oxford: Blackwell. 539–588.

Hyltenstam, K. & Abrahamsson, N. 2003b. Age of onset and ultimate attainment in near-native speakers of Swedish, in Fraurud, K. & Hyltenstam, K. (eds.), *Multilingualism in Global and Local Perspectives. Selected Papers from the 8th Nordic Conference on Bilingualism, November 1–3, 2001, Stockholm Rinkeby.* Stockholm: Centre for Research on Bilingualism, Stockholm University, and Rinkeby Institute of Multilingual Research. 319–340.

Hyltenstam, K., Bylund, E., Abrahamsson, N. & Park, H.-S. 2009. Dominant-language replacement: the case of international adoptees. *Bilingualism: Language and Cognition* 12/2: 121–140.

Ioup, G. 1989. Immigrant children who have failed to acquire native English, in Gass, S., Madden, C., Preston, D. & Selinker, L. (eds.), *Variation in Second Language Acquisition: Vol. 2. Psycholinguistic Issues*. Clevedon: Multilingual Matters. 160–175.

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

Johnson, J. S. & Newport, E. L. 1989. Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21: 60–99.

Krashen, S. 1973. Lateralization, language learning, and the critical period: some new evidence. *Language Learning* 23: 63–74.

Krashen, S. D. 1976. Formal and informal linguistic environments in language acquisition and language learning. *TESOL Quarterly* 10/2: 157–168.

Krashen, S. D., Long, M. A. & Scarcella, R. C. 1979. Age, rate, and eventual attainment in second language acquisition. *TESOL Quarterly* 13: 573–582.

Kroll, J. F. & Stewart, E. 1994. Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language* 33/2: 149–174.

Kroll, J. F. & Tokowicz, N. 2005. Models of bilingual representation and processing: looking back and to the future, in Kroll, J. F. & De Groot, A. M. B. (eds.), *Handbook of Bilingualism: Psycholinguistic Approaches*. Oxford: Oxford University Press. 531–533.

Lamendella, J. T. 1977. General principles of neurofunctional organization and their manifestation in primary and nonprimary language acquisition. *Language Learning* 27/1: 155–196.

Lenneberg, E. 1967. *Biological Foundations of Language*. New York, NY: Wiley.

Long, M. H. 1990. Maturational constraints on language development. *Studies in Second Language Acquisition* 12: 251–285.

Long, M. H. 1993. Second language acquisition as a function of age: research findings and methodological issues, in Hyltenstam, K. & Viberg, Å. (eds.), *Progression and Regression in Language*. Cambridge: Cambridge University Press. 196–221.

Long, M. 2013. Maturational constraints on child and adult SLA, in Granena, G. & Long, M. (eds.), *Sensitive Periods, Language Aptitude, and Ultimate L2 Attainment*. Amsterdam: John Benjamins Publishing Company. 3–41.

MacWhinney, B. 2005. A unified model of language acquisition, in Kroll, J. F. & De Groot, A. M. B. (eds.), *Handbook of Bilingualism: Psycholinguistic Approaches*. Oxford: Oxford University Press. 49–67.

Meara, P., Milton, J. & Lorenzo-Dus, N. 2003. *Swansea Language Aptitude Tests (LAT) v2.0*. Swansea: Lognostics.

Muñoz, C. & Singleton, D. 2011. A critical review of age-related research on L2 ultimate attainment. *Language Teaching* 44: 1–35.

Munro, M. & Mann, V. 2005. Age of immersion as a predictor of foreign accent. *Applied Psycholinguistics* 26/3: 311–341.

Norrman, G. & Bylund, E. 2016. The irreversability of sensitive period effects in language development: evidence from second language acquisition in international adoptees. *Developmental Science* 19/3: 513–520.

Norrman, G., Hyltenstam, K. & Bylund, E. 2016. Long-term language development in international adoptees, in Genesee, F. & Delcenserie, A. (eds.), *Starting Over – Language Development in Internationally-Adopted Children*. Amsterdam: John Benjamins. 125–145.

Obler, L. K. 1989. Exceptional second language learners, in Gass, S., Madden, C., Preston, D. & Selinker, L. (eds.), *Variation in Second Language Acquisition*. Clevedon: Multilingual Matters. 141–149.

Oh, J. S., Au, T. K.-F. & Jun, S.-A. 2010. Early childhood language memory in the speech perception of international adoptees. *Journal of Child Language* 37/5: 1123–1132.

Olson, L. & Samuels, S. J. 1973. The relationship between age and accuracy of foreign language pronunciation. *Journal of Educational Research* 66: 263–267.

Ortega, L. 2010. The bilingual turn in SLA. *Plenary Delivered at the Annual Conference of the American Association for Applied Linguistics*, Atlanta, GA.

Ortega, L. 2013. SLA for the 21st century: disciplinary progress, transdisciplinary relevance, and the bi/multilingual turn. *Language Learning* 63/1: 1–24.

Oyama, S. 1976. A sensitive period for the acquisition of a non-native phonological system. *Journal of Psycholinguistic Research* 5: 261–285.

Oyama, S. 1978. The sensitive period and comprehension of speech. *Working Papers on Bilingualism* 16: 1–17.

Pallier, C., Dehaene, S., Poline, J.-B., LeBihan, D., Argenti, A.-M., Dupoux, E. & Mehler, J. 2003. Brain imaging of language plasticity in adopted adults: can a second language replace the first? *Cerebral Cortex* 13/2: 155–161.

Paradis, M. 2004. *A Neurolinguistic Theory of Bilingualism*. Amsterdam: Benjamins.

Paradis, M. 2009. *Declarative and Procedural Determinants of Second Languages*. Amsterdam: John Benjamins.

Park, H.-S. 2015. Korean adoptees in Sweden: have they lost their first language completely? *Applied Psycholinguistics* 36: 773–797.

Patkowski, M. 1980. The sensitive period for the acquisition of syntax in a second language. *Language Learning* 30: 449–472.

Penfield, W. & Roberts, L. 1959. *Speech and Brain Mechanisms*. New York, NY: Atheneum.

Pierce, L. J., Klein, D., Chen, J.-K., Delcenserie, A. & Genesee, F. 2014. Mapping the unconscious maintenance of a lost first language. *Proceedings of the National Academy of Sciences* 111/48: 17314–17319.

Piller, I. 2002. Passing for a native speaker: identity and success in second language learning. *Journal of Sociolinguistics* 6: 179–206.

Pulvermüller, F. & Schumann, J. H. 1994. Neurobiological mechanisms of language acquisition. *Language Learning* 44: 681–734.

Schmid, M. S. 2002. *First Language Attrition, Use and Maintenance: The Case of German Jews in Anglophone Countries*. Amsterdam: John Benjamins Publishing.

Selinker, L. 1972. Interlanguage. *IRAL* 10: 209–230.

Singh, L., Liederman, J., Mierzejewski, R. & Barnes, J. 2011. Rapid reacquisition of native phoneme contrasts after disuse: you do not always lose what you do not use. *Developmental Science* 14/5: 949–959.

Snow, C. & Hoefnagel-Höhle, M. 1977. Age differences in the pronunciation of foreign sounds. *Language and Speech* 20: 357–365.

Snow, C. & Hoefnagel-Höhle, M. 1978. The critical period for language acquisition: evidence from second language learning. *Child Development* 49: 1114–1128.

Sorace, A. & Robertson, D. 2001. Measuring development and ultimate attainment in non-native grammars, in Elder, C., Brown, A., Grove, E., et al. (eds.), *Experimenting with Uncertainty. Essays in Honour of Alan Davies*. Cambridge: Cambridge University Press. 264–274.

Stevens, G. 1999. Age at immigration and second language proficiency among foreign-born adults. *Language in Society* 28: 555–578.

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

Stölten, K., Abrahamsson, N. & 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* 37/1: 71–100.

Ullman, M. T. 2005. A cognitive neuroscience perspective on second language acquisition: the declarative/procedural model, in Sanz, C. (ed.), *Mind and Context in Adult Second Language Acquisition: Methods, Theory, and Practice*. Washington, DC: Georgetown University Press. 141–178.

Ullman, M. T. 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* (2nd edn.). New York, NY: Routledge. 135–158.

Ullman, M. T. & Pullman, M. Y. 2016. A compensatory role for declarative memory in neurodevelopmental disorders. *Neuroscience and Biobehavioral Reviews* 51: 205–222.

Vanhove, J. 2013. The critical period hypothesis in second language acquisition: a statistical critique and a reanalysis. *PLoS ONE* 8(7), e69172. doi:10.1371/journal.pone.0069172.

Vaughan-Evans, A., Kuipers, J. R., Thierry, G. & Jones, M. W. 2014. Anomalous transfer of syntax between languages. *The Journal of Neuroscience* 34/24: 8333–8335.

Ventureyra, V. A. G., Pallier, C. & Yoo, H.-Y. 2004. The loss of first language phonetic perception in adopted Koreans. *Journal of Neurolinguistics* 17/1: 79–91.

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