Introduction: Ultimate attainment and the critical period hypothesis for second language acquisition

THEO BONGAERTS

1. Preliminaries

For several decades the topic of age effects on ultimate attainment has been high on the agenda of many second language acquisition researchers. A first major evaluation of research in this area was published by Long (1990), who summarized the findings of studies conducted since Penfield and Roberts (1959) and Lenneberg (1967) developed their versions of a critical period hypothesis for language acquisition. Long (1990: 280) argued that the combined findings of the studies conducted to date warranted the conclusion that

The ability to attain native-like phonological abilities in an SL begins to decline by age 6 in many individuals and to be beyond anyone beginning later than age 12, no matter how motivated they might be or how much opportunity they might have. Native-like morphology and syntax only seem to be possible for those beginning before age 15.

In line with most proponents of a critical period for SLA, Long (1990: 280) posited that the decline in abilities is due to incremental (and presumably irreversible) losses of neural plasticity due to brain maturation.

As noted by Singleton (this volume, Singleton and Ryan 2004), critical period hypotheses have come in different forms and guises, but common to all versions is that the concept of a critical period for language acquisition entails that age effects operate within a circumscribed period of time, bounded by an onset which marks the beginning and a terminus which marks the end of a period of heightened sensitivity to ambient language input (Birdsong 1999, 2004). Such time-boundedness should be reflected in SLA data in at least the following two ways: (1) there should be a discontinuity in the slope of the decline in L2-proficiency situated around the terminus of the critical period, and (2) no second language learners starting after the terminus period should demonstrate achievement of native-like levels of ultimate L2-attainment.

Long (1990) ended his paper with a few desiderata for future research. These desiderata included a focus on very advanced learners, the use of multiple elic-
ivation instruments, and studies comparing data on neural processing of second language input and behavioural data from learners who started to acquire a second language at different ages. Since the publication of Long’s survey well over a hundred studies of age effects in second acquisition have been conducted, a sizeable portion of them following up Long’s suggestions.

This is not the place to present a full-length review of the research conducted since 1990. We will, therefore, limit ourselves to a brief survey that focusses on studies that speak to the issue of (a) the (lack of) discontinuity in the slope of the decline around the terminus, and (b) the (im)possibility of native-like attainment by late learners.

1.1. Discontinuity or linearity in the slope of the decline after the terminus?

In the realm of pronunciation, there is an early study by Patkowski (1980, 1990) which is relevant here. Patkowski reported a sharp discontinuity in L2 pronunciation proficiency around an age of acquisition (hence AoA) of 15, which is, incidentally, way past the terminus of the critical period for pronunciation proposed by Long (1990). The observed discontinuity could perhaps be marshalled in support of the critical period hypothesis, but, as argued by Harley and Wang (1997: 34), a regression analysis on the accent ratings suggested that “the overall pattern of decline (i.e. the slope of the regression lines) was not radically different in Patkowski’s pre-15 and post-15 groups” (see Patkowski 1990: 79). In this context, it should be pointed out that Patkowski’s report of a discontinuity in the slope of the decline has not been supported by the outcomes of a number of other studies. For example, Oyama (1976, 1982), while reporting an AoA-related decline, commented that the increase in foreign accent with AoA was “quite linear” and that “no sharp discontinuities” were observed (Oyama 1982: 28), and Flege, Munro, and MacKay (1995) reported that the pronunciation ratings given to Italian learners of English decreased systematically as AoA increased, and, more importantly, that there was “no discontinuity in the ratings at an AoA of 15 years, or at any other AoA” (Flege 1999: 102). Finally, Flege, Yeni-Komshian and Liu (1999) found no discontinuity in the ratings given to Korean learners of English either at an AoA of 12 years or at an AoA of 15 years. What these studies demonstrate instead is a continuing and substantial decline in the pronunciation ratings not only before, but also after the terminus of the critical period (see also Bongaerts 2003).

In a recent survey of studies of ultimate attainment, Birdsong (2004) discusses the results of a number of studies in which participants were asked to provide grammaticality judgements. Birdsong concluded that, if these studies showed a discontinuity in the slope of the decline, this discontinuity did not seem to occur around the terminus of the critical period, but (much) later. Most,
but not all, studies appeared to reveal a continuing decline in test scores with AoA, not only before, but also after the terminus of the critical period.

In sum, it appears that a sharp discontinuity in the slope of the decline around the terminus, which would be consistent with the idea of a critical period for SLA, is often not attested in studies of the acquisition of L2-pronunciation and L2-(morpho)syntax.

1.2. Does post-critical-period second language acquisition inevitably result in non-native-likeness?

As noted by Long (1990, this volume) and Birdsong (2004), the claim in all current versions of the critical period hypothesis is that native-like attainment will not be possible if the start of second language acquisition is delayed past a certain critical age. For pronunciation, this critical age has been put at 12 by Long (1990). Since Long’s survey a number of studies have been conducted that allow an evaluation of this claim. The first study that comes to mind is the study by Flege, Munro and MacKay (1995) referred to above. Flege et al. studied a group of 240 Italian immigrants to Canada, with AoAs ranging from 2 to 23. Recordings of English sentences read out by these immigrants as well as by 24 native English controls were rated for degree of foreign accent by 10 native speakers of English. Flege et al. (1995) reported that 6% of the immigrants with AoAs > 12 performed within the native speaker range. True, none of them had begun learning English after the age of 16, but this age is way beyond the age which has been posited as marking the terminus of critical period for pronunciation. It should be noted, however, that in a similar study, this time with Korean immigrants, none of the participants performed within the native range (Flege, Yeni-Komshian and Liu 1999). Unlike Flege et al. 1995, Bongaerts and colleagues (Bongaerts 1999, Bongaerts et al. 1997, 2000) targeted very advanced, late (post-critical-period) second language learners. On sentence reading tasks, five participants performed at the level of native controls in a study with Dutch learners of English (Bongaerts et al. 1997), three performed in a native-like way in a study with French learners of English (Bongaerts 1999) and two did so in a study with immigrants to the Netherlands with different L1-backgrounds (Bongaerts et al. 2000). The latter two immigrants were a native speaker of German who started to learn Dutch at the age of 14 and a native speaker of English with an AoA of 21 (for more detailed information on these studies see Long, this volume). Incidence of native-likeness has not been reported exclusively in studies using limited speech samples elicited through sentence reading tasks. To give one example, Birdsong (2003) studied 22 native speakers of English who had started to acquire French at or after the age of 18. These learners, as well as 17 native French controls, provided two sets of speech samples: they read aloud 21 words (12 words having the vow-
The incidence of native-like attainment well beyond the critical age in the
domain of morphosyntax seems to be even higher. For recent surveys the inter-

Studies comparing L2-learners’ and native speakers’ linguistic knowledge
across multiple domains of performance are rare. Recent examples are an on-
going study by Hyltenstam and Abrahamsson (see, e.g., Hyltenstam and Abra-
hamsson 2003) and a recently completed study by Marinova-Todd (2003). Wherein in the Hylstenstam and Abrahamsson study none of the L2 learners
obtained native-like scores on all tasks, there were some late learners who per-
formed at a native level on some tasks. Marinova-Todd (2003) studied 30 late
learners of English from a variety of L1-backgrounds who had been immersed
at or after age 16. Marinova-Todd (2003) reported that three subjects performed
within the native range on all nine tasks that were administered (two measures
of pronunciation, two of vocabulary, three of morphosyntax and two of lan-
guage use). None of these three learners had been immersed in English before
the age of 21. Interestingly, two of them were native speakers of a language that
is typologically distant from English (Slovak and Russian). The other learner
was a native speaker of French.

In sum, what these studies seem to have shown is that (1) native-like levels
can be attained in a variety of linguistic domains, by individuals who begin to
learn an L2 at ages beyond (sometimes well beyond) a purported critical pe-
riod and that (2) the possibility of attaining native-likeness beyond the critical
period is not restricted to native speakers of languages that are typologically
closely related to the second language to be acquired. Such findings, plus the
finding of a continuing decline of language learning abilities after the terminus
of the critical period that has routinely been attested in group-wise analyses of
data from studies with learners with a variety of AoAs, are hard to accommo-
date with any version of the critical hypothesis for SLA.
Traditionally, studies of ultimate attainment have been inspired by the notion of a critical period for SLA. Unfortunately, this has led to an almost exclusive focus on age of acquisition and to the idea that ultimate attainment is primarily a function of age. Other unfortunate consequences were that research has tended to concentrate on a limited number of L1–L2 combinations, typically typologically related ones, and that only few studies have gathered detailed background information about individual learners. Such information should, in the case of individuals who have settled in another country, minimally include educational experience in the country of immigration, patterns of L1- vs L2-use over a longer period of time (both at home, with people in the neighbourhood, at work), and social-psychological orientations (e.g., instrumental and integrative motivations to learn the language, personal importance of sounding native, perceptions of self in the L2, sense of belonging). One of the very few studies which gathered such detailed information has recently been published by Moyer (2004). Her conclusion was that ultimate attainment was multiply determined, AoA being only one, and not the most important, of the determining variables. There is clearly a great need for more studies of this kind.

2. The contributions to this special issue

The five contributions to this special issue provide a variety of perspectives on the effects of age of acquisition on ultimate attainment in a second language.

Taking a historical perspective, Singleton observes that it would be wrong to speak in terms of the critical period hypothesis, given the fact that the notion of a critical period has been interpreted in so many different ways. He points out that researchers have held different views not only on the timing of the offset and the terminus of the critical period, but also on the language learning capacities that are affected by it and even on its underlying causes. With respect to the capacities affected by a critical period, he shows that, while some believe that all language capacities are affected, others argue that maturational constraints only impact on the acquisition of pronunciation, and still others claim that maturational constraints apply exclusively to implicit learning mechanisms. Regarding underlying causes of the critical period, Singleton observes that, while most researchers have posited neurobiological causes, cognitive-developmental and affective-motivational causes have also been suggested. Singleton’s conclusion strikes a note of pessimism: “the CPH cannot plausibly be regarded as a scientific hypothesis [...] it is like the mythical hydra, whose multiplicity of heads and capacity to produce new heads rendered it impossible to deal with”.

Fifteen years after his first review of the literature, Long surveys the field once again. In the introduction to his contribution, Long clearly states his own position: “My own position was then, and remains, that in terms of absolute
language learning capacity, only young child starters can (not necessarily will) attain native-like proficiency levels". He argues that, while counter-evidence to the critical period hypothesis has been advanced in a number of studies, there are various problems with this counterevidence. The following nine problems are mentioned, each of which is illustrated with reference to one or more studies: (1) confusion of rate and ultimate attainment, (2) inappropriate choice of subjects, (3) measurement of AoA, (4) leading instructions to raters, (5) assessments based on limited samples and/or "language-like" behaviour, (6) use of markedly non-native samples making near-native samples more likely to sound native to raters, (7) unreliable or invalid measures, (8) inappropriate L1–L2 pairings, and (9) faulty interpretation of statistical patterns. Long raises a number of important methodological points that researchers would be wise to heed in future studies. One could take issue, however, with some of the conclusions that Long draws from his literature review. For example, Long claims that native-like attainment by late starters has typically been attested in studies with typologically related L1-L2 pairs using limited samples or tests of language-like behaviour. Studies using a battery of hard tests and more natural language use with late L2-learners with a typologically distant L1 are claimed to have "always unearthed non-native features". However, as pointed out above, there are some researchers that have reported native-like attainment in recent studies using multiple tests of pronunciation (Moyer 1999, Birdsong 2003) and in a study using nine different tests, including some eliciting natural language use (Marinova-Todd 2003). The latter study, which Long was apparently not aware of, reported overall native-likeness in English for three late starters (a native speaker of French, one of Slovak, and one of Russian). As Long observes, such studies are very rare and more studies involving typologically distant L1-L2 combinations, in which multiple tests are used, would be a desideratum for the future.

In his contribution, Birdsong takes issue with the way criteria of native-likeness and non-native-likeness have sometimes been used in studies on the critical period for SLA. He argues that it would be inappropriate to insist that, for the critical period hypothesis to be falsified, late L2-learners would have to be identified who behave like monolingual native speakers in every nook and cranny of the target language. It is suggested that not all departures from native-likeness are indicative of defective learning mechanisms. To illustrate his point, Birdsong reviews some routinely observed effects of the L2 on the L1 and of the L1 on the L2 of bilinguals. There are studies showing that the L1 of a bilingual is not identical to the L1 of a monolingual, but no researcher would view such differences between bilinguals and monolinguals as shortcomings in the language learning abilities of bilinguals. Another effect of bilingualism is that there are cognitive costs associated with maintaining the L1, which are reflected in L2-processing which is often less efficient than that of monolinguals.
Birdsong concludes his contribution by suggesting that, while native-likeness for late learners has been attested across multiple domains of L2-performance, this may be out of reach in tasks involving language processing, for example in parsing and lexical retrieval tasks. To explore this possibility, such tasks should be administered to late learners whose L2 is their dominant language, and, one might add, who have not used their L1 and have been deprived of L1-input for a long period of time.

Common to nearly all versions of the critical period hypothesis is the assumption that child/adult differences in ultimate attainment are due to changes in the neuronal structure of the brain as one gets older. However, as Stowe and Sabourin observe in their contribution, there seems to be a lack of consensus on some fundamental questions, for example on the nature of the neuronal changes and on the time course over which these changes take place. Neuroimaging studies could provide useful information concerning these questions. Positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) provide evidence as to where language is processed in the brain and event related potentials (ERP) and magnetic encephalography (MEG) as to when various processes take place. In a first review of PET and fMRI studies of language comprehension and language production in bilinguals, Abutalebi et al. (2001: 179) concluded: “Consistent results indicate that attained proficiency, and maybe language exposure, are more important than the age of acquisition as a determinant of the cerebral representation of languages in bilinguals/polylnguists”. Since Abutalebi et al.’s review, many new studies have been conducted. Stowe and Sabourin provide an updated survey of PET and fMRI studies as well as a review of ERP and MEG studies, which were not included in Abutalebi et al. Stowe and Sabourin observe that the results of the where studies lead to two main, be it tentative, conclusions: (1) the same brain areas are used for processing both first and second language, even by late learners, although late learners are more likely to draw on additional resources during L2-processing, and (2) that these areas are not necessarily employed as efficiently during L2-processing, even for languages learned early in life. The results of the when studies, while being in general consistent with the above conclusions, also point to a role of factors such as the type of information being processed, the amount of early input or input over the lifetime, proficiency, and similarity between L1 and L2. With respect to the role of age of acquisition, Stowe and Sabourin’s conclusions are largely in line with those of Abutalebi et al. (2001): factors such as the ones mentioned above need to be considered more seriously before any firm conclusions on the role of age of acquisition can be made.

Finally, Van Boxtel et al. report the results of an empirical study on ultimate attainment of L2-syntax by very advanced late learners of Dutch as a second language from three different L1-backgrounds: German, French and Turkish.
They compared the learners’ performance with the performance of Dutch native speakers on two tasks: a sentence preference task and an elicited imitation task. For these tasks, dummy subject constructions were used, for two reasons: (1) they are known to be very difficult to acquire for second language learners, and (2) very little is known about these constructions, which is why they are hardly covered in Dutch grammars and textbooks. Therefore, no explicit rules about dummy constructions are available to learners of Dutch and, consequently, these constructions can only be acquired by applying implicit learning mechanisms to the input. Thus, they provide an excellent test case for the critical period hypothesis. Under this hypothesis, native-like attainment of such constructions would be out of the reach of late L2-learners. The results of the study show, however, that native-like attainment by late starters is possible for these constructions, and that native-like attainment is not the exclusive domain of native speakers of languages that are typologically not very distant from Dutch (German and French). There was one native speaker of Turkish who performed within the native range on both tasks.

References


Introduction


