1. The Critical Period Hypothesis and Second Language Acquisition

Introduction

Much research on L2 acquisition is either implicitly or explicitly informed by the Critical Period Hypothesis, that is, the idea that maturational changes enforce an asymmetry between language acquisition early in life compared to later acquisition. Originally formulated by Penfield & Robert (1959) and Lenneberg (1967), (versions of) the Critical Period Hypothesis have widely been recruited to explain both the success of child L1 acquisition and the comparative failure of adult L2 acquisition. This chapter introduces the concept of critical periods in behavioural and cognitive development and discusses its application to language acquisition. The major part of this chapter deals with critical periods in the context of L2 acquisition. After discussing the characteristics of potential critical periods and advancing a working definition of a critical period (Section 1.3), I review the empirical evidence adduced in favour of and against the existence of bounded age-related constraints on L2 acquisition (Sections 1.3.2 to 1.3.5).

This chapter aims to give a general overview of the various ways in which the investigation of age-related effects in L2 acquisition has been tackled. Some studies representative of each way will be discussed and their respective merits and lacunae outlined. On the basis of this review, I introduce the conceptual basics and rationale of generative approaches to the study of L2 ultimate attainment. These approaches will be discussed with a particular emphasis on the acquisition of phenomena at the interfaces of syntax with other grammatical modules. Several representative studies will be discussed to illustrate theoretical approaches to limitations on adult L2 acquisition that will inform the empirical study of this thesis.

1.1. Critical Periods: Features and biological foundations

A critical period describes a maturationally constrained period of time in which the interaction of extant neurophysiological architecture and relevant external stimuli results in the full development of a particular capacity.\(^1\) Hence, a critical period represents a time of heightened sensitivity to environmental stimuli for acquiring essential aptitudes that will not develop at all or not fully develop if exposure to the relevant input occurs earlier or later than in the biologically determined time window.

\(^1\) In the literature on maturationally governed developmental processes, the terms sensitive or optimal period are often used (for discussion, see Ruben, 1997). A sensitive period denotes a less categorically circumscribed period of heightened sensitivity than suggested by the classic conception of a critical period. In the following, I will retain use of the term critical period for expository convenience (see also Eubank & Gregg, 1999; Hyltenstam & Abrahamsson, 2003), acknowledging, however, that the time span and the slope of the onsets and offsets of critical periods (see Figure 1.1) might be variable and smear out across phenomena.
Generically, then, a critical period is preceded by a period of low sensitivity to a given stimuli and the subsequent critical period encompasses an onset, a peak period and an offset, as illustrated in Figure 1.1. It is succeeded by the subsequent flattening out of the ability to acquire a behavioural process in response to the particular stimuli. Typically, the onset of a critical period is rather abrupt, while the offset is more gradual; however, different geometries of critical periods can be envisaged and have been put forth (see, e.g., Bornstein, 1989, and Birdsong, 2005a for discussion).

Figure 1.1. Geometry of prototypical critical period (adapted from Birdsong, 2005a).

Critical periods have been found to underlie the development of diverse neurosensory capacities in biology (Knudsen, 2004). Critical periods conforming to the geometry in Figure 1.1 have been observed, for instance, in the development of vision in ferals (e.g. Hubel & Wiesel, 1970), haptic perception in mice (Glazewski, Chen, Silva, & Fox, 1996), the mating song in a species of African frogs (Watson & Kelley, 1992), auditory orientation in the barn owl (Knudsen & Knudsen, 1990) and the song of various bird species, e.g. zebra finches and sparrows (Marler, 1990).

Perhaps the closest analogue to the development of human language, the development of species-specific birdsong, has been demonstrated to be contingent on exposure to the song of conspecifics in the first year of life of certain bird species. Further, young birds must be able to hear themselves during that time in order to develop normal song patterns (e.g. Brenowitz, Margoliash & Nordeen, 1997). If these conditions are not met, the birds will develop abnormal song patterns that will not be qualitatively affected by later exposure to conspecific song patterns. Experimentally, it has further been shown that the emergence and the characteristics of birdsong can be modulated by direct neurophysiological intervention (e.g. lesioning brain areas dedicated to song learning) or hormonal manipulation, which suggests that the onset and offset of a critical period for birdsong is neurally founded and hormonally regulated (ibid).
In the human development of hearing, recent studies on auditory perception and speech comprehension of congenitally deaf children who have been fitted with cochlear implants at various ages suggest that age mediates the successful acquisition of auditory discrimination skills. Cochlear implants are surgically implanted devices that transform auditory signals into electrical pulses to stimulate the auditory nerves in individuals suffering from severe to profound hearing loss in order to emulate environmental auditory stimuli, in particular speech. Several studies find a non-linear relation between age of implantation and the development of auditory skills. They suggest that normal speech recognition development is contingent upon auditory input during a constrained critical period in the first few years of life (e.g., McConkey Robbins, Burton Koch, Osberger, Zimmerman-Phillips & Kishon-Rabin, 2004; Sharma, Dorman & Spahr, 2002; for critical discussion, see Harrison, Gordon & Mount, 2005). In view of these critical periods in the acquisition of sensorimotor capacities serving language comprehension, it is natural to ask whether critical periods can also be found for the neurocognitive representations of language.

1.2. The Critical Period Hypothesis and language acquisition

Originally formulated by Penfield & Roberts (1959) and further elaborated by Lenneberg (1967), the idea that a constrained period of heightened neural sensitivity underlies language development has become widely espoused in research on first language acquisition to account for the contrast in speed, ease and success of normal child language development and the failure in later language acquisition when no relevant linguistic input was available during early childhood.

Evidence to underscore the reality of critical periods for language acquisition is furnished by (a) feral children (for review, e.g., Curtiss, 1988), (b) congenitally deaf children acquiring sign language later in life (e.g., Mayberry & Lock, 2003) and (c) differences in auditory discrimination abilities of non-native phonemic contrasts in children at different ages (e.g., Werker & Tees, 2005).

Although a critical period is often ascribed to language acquisition in general, there has been growing awareness that ‘language’ is too coarse-grained a concept to which to attribute maturational effects in toto (e.g., Eubank & Gregg, 1999; Long, 1990; Newport et al., 2001). Seeing that ‘language’ consists of multiple conceptually and neurally distinguishable subcomponents such as syntax, semantics, phonology, etc., each of these components could either be subject to critical periods of potentially different properties, scopes and temporal characteristics or be spared from maturational constraints altogether.

Support for a differentiated conceptualization of multiple critical periods in language acquisition can indeed be found in many of the studies documenting linguistic failure in delayed first language acquisition. For instance, the well-known case of Genie,
who was completely deprived of linguistic input from the age of 1;6 to 13 years (Curtiss, 1988), illustrates dissociations between linguistic components. After rescue, Genie showed quick development of vocabulary and communicative skills, whereas her syntax and intonation remained rudimentary even after prolonged exposure (for further examples, see Pinker, 1994: 290-296; Singleton & Ryan, 2004: 31-60).

More recent studies furnish evidence that even subcomponents of linguistic domains, such as syntax or phonology, can be differentially affected by maturational constraints. In recent work on late-acquired sign language, Mayberry & Lock (2003; see also Mayberry & Lock, 2002) report on a comparative study of congenitally deaf individuals who had no exposure to (signed) linguistic input prior to age six and were tested more than 11 years after the delayed onset of tuition and usage of English as adults. Their linguistic knowledge was compared to three groups of adults who had received spoken or signed English input or spoken non-English (Urdu) input in early childhood and had had comparable lengths of exposure to English. On several tasks, the congenitally deaf individuals with no early exposure to language correctly judged some syntactic violations (ungrammatical auxiliary selection, ungrammatical object placement, ungrammatical conjunction) but performed at chance on identifying ungrammatical structures involving syntactic movement (passives, relative clauses). By contrast, the other groups showed no such asymmetry in behaviour. For phonology, findings of a similar nature indicating different age effects depending on the type of phonological phenomena have been reported in the context of cochlear implantation by Harrison, Gordon & Mount (2005), and in the context of normal child development by Werker & Tees (2005).

In sum, the robust findings that humans with no early exposure to language demonstrate persistent failure in late language acquisition suggest that critical periods constrain language development. However, different domains of language are differentially affected by early deprivation of input, with some showing better recoverability upon delayed exposure than others.

1.3. The Critical Period Hypothesis and second language acquisition

1.3.1. Conceptualizing the Critical Period Hypothesis in L2 acquisition

Conceptualizing a critical period in the context of adult second language acquisition requires severe revisions to the way critical periods are commonly characterized (Eubank & Gregg, 1999). In contrast to the prototypical cases of critical periods, where no relevant external stimuli were available at critical moments, adult L2 acquisition by definition occurs against the background of successful L1 acquisition during the critical period. Hence, an adult L2 learner cannot be said to have missed the critical period for language acquisition outright; at the same time, acquiring the L2 indeed occurs outside the critical
period. Assuming, then, that L1 acquisition entails a particular neurocognitive organization of linguistic knowledge, acquiring an L2 in adulthood requires changes to the extant patterns established for the L1. Irrespective of the conceptualization of these changes, such neurocognitive reorganization implicates much more limited changes in neural architecture and cognitive processes compared to the large-scale organization of linguistic knowledge that underlies the transition from the child’s pre-linguistic state to its mature state (e.g. Lenneberg, 1967: 176). As a consequence, e.g. Eubank & Gregg (1999) caution that the behavioural reflexes of critical period effects observeable in adult second language acquisition should be relatively minor in scope.

Against this background, critical period effects also need to be differentiated from age effects in L2 acquisition that are not maturationally bounded. Late L2 acquisition is likely to be susceptible to gradual effects of cognitive aging that begin in early adulthood. Cognitive aging affects language performance in causing slower and less efficient information processing and reduced working memory capacities (e.g. Park, 2000). In and of themselves, cognitive effects of aging may implicate less successful outcomes in adult L2 acquisition (e.g. Bialystok & Hakuta, 1999; Hakuta, Bialystok, & Wiley, 2003). Crucially, these age effects should be continuous and show a linear relation to age as opposed to the non-linear relation of age and acquisition predicted by a critical period (Figure 1.1).

In addition, the componential organization of language potentially entails that there might be different critical periods for different linguistic subcomponents, with some subcomponents showing no critical periods and others being subject to critical periods of different onsets and offsets. In recent reviews of the literature on critical periods in L2 acquisition, Long (2005) and Singleton (2005) give overviews of the varying ends of the offset that have been proposed for subparts of language. For phonetics and phonology, the proposals range from one year (e.g. Ruben, 1997) to twelve years of age (e.g. Scovel, 1988), for morphosyntax, the suggestions extend from shortly after birth (Hyltenstam & Abrahamsson, 2003) to fifteen years of age (Long, 1990; 2005). Table 1.1 (adapted from Singleton, 2005) presents a schematic overview of some proposals.
An additional issue of debate is what the underlying causes of maturational constraints are. Within neurocognitive research, neurophysiological changes of various kinds have been argued to occur by the end of critical periods affecting cognitive representations and their use (for overview, see, e.g., Hyltenstam & Abrahamsson, 2003; Singleton, 2005). Within socio-affective research, critical period effects have been conceptualized more broadly as age-related effects in terms of motivational, emotional factors (e.g. psychological distance to the target-language community) or environmental factors (e.g. amount and type of L2 input, for overview, see, e.g., Franceschina, 2005). Table 1.2, based on Birdsong (1999a), Hyltenstam & Abrahamsson (2003) and Singleton (2005), sketches some of the factors and underlying causes of age-related effects in L2 acquisition that have been advanced in the literature.

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>End of offset for general CP</th>
<th>End of offset for phonology</th>
<th>End of offset for (morpho-)syntax</th>
<th>End of offset for semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Penfield &amp; Roberts (1959)</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lenneberg (1967)</td>
<td>puberty</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scovel (1988)</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Johnson &amp; Newport (1989; 1991)</td>
<td>puberty or earlier</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long (1990; 2005)</td>
<td>6 or 12</td>
<td>mid-teens</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ruben (1997)</td>
<td>1</td>
<td>4</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Hyltenstam &amp; Abrahamsson (2003)</td>
<td>shortly after birth</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clahsen &amp; Muysken (1986; 1989); Hawkins (2001a)</td>
<td></td>
<td>puberty</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1.1. Summary of proposals for offset of critical periods.
The Critical Period Hypothesis and Second Language Acquisition

<table>
<thead>
<tr>
<th>Type of explanation</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neurophysiological</td>
<td>Penfield &amp; Roberts (1959) Loss of cerebral plasticity</td>
</tr>
<tr>
<td></td>
<td>Lenneberg (1967) Lateralisation</td>
</tr>
<tr>
<td></td>
<td>Pulvermüller &amp; Schumann (1994) Myelination</td>
</tr>
<tr>
<td></td>
<td>Kim et al. (1997) Different spatial cerebral representations</td>
</tr>
<tr>
<td>Cognitive</td>
<td>Clahsen &amp; Muysken (1986); Bley-Vroman (1990) Loss of access to Universal Grammar</td>
</tr>
<tr>
<td></td>
<td>Tsimpli &amp; Roussou (1991); Hawkins &amp; Chan (1997); Hawkins (2001a) Loss of access to parts of Universal Grammar</td>
</tr>
<tr>
<td></td>
<td>Felix (1987) Interference of analytical abilities</td>
</tr>
<tr>
<td></td>
<td>Newport (1990) Rising complexity of analytical ability in adults</td>
</tr>
<tr>
<td>Affective-motivational</td>
<td>Krashen (1985) Blocking effects of ‘affective filter’</td>
</tr>
<tr>
<td></td>
<td>Schumann (1978) Increase in social and psychological distance</td>
</tr>
<tr>
<td></td>
<td>Bialystok &amp; Hakuta (1999) Changes in psychosocial factors (simplified input, motivation, schooling)</td>
</tr>
</tbody>
</table>

Table 1.2. Causes of critical period and age-related effects in language development suggested in previous research.

Of course, it is ultimately an empirical question to determine the behavioural manifestations, the nature and the timing of maturational disjunctions in (second) language development. Yet, the wide diversity of Critical Period hypotheses differing in scope and temporal characteristics renders it nigh-impossible to address THE Critical Period Hypothesis, since the Critical Period Hypothesis at best denotes a common frame of reference of approaches to L2 acquisition cast in terms of maturation or age effects on language development (e.g. Birdsong, 2004; Singleton, 2005). It is not the aim of this chapter to review all of these various approaches exhaustively or to evaluate their various merits.

Bearing in mind the diversity of maturational accounts, I will address a working definition of a Critical Period Hypothesis in the rest of this thesis. Given this thesis’
themetic focus on syntax and its interfaces and its theoretical commitment to a (generative) cognitive approach to psycholinguistics, I will put aside phonetics and phonology (for discussion, see, e.g., Bongaerts, 1999; Moyer, 1999, 2004) as well as socio-affective and learner-external factors of L2 acquisition. As a first approximation, the broad working hypothesis of critical period effects in L2 acquisition stated in (1) will guide the following sections in this chapter.

(1) Critical Period Hypothesis (working definition)  
There is a critical period for language acquisition terminating at the latest by puberty, beyond which it is not possible to acquire native-like neurocognitive representations and processing of (components of) syntax and its interfaces.

Broadly speaking, one can identify two different strands of research testing for empirical evidence of a critical period, defined as in (1), in L2 acquisition: (A) the comparative approach, which focuses on the juxtaposition of different types of L2 acquisition, namely, child L2 acquisition and adult L2 acquisition, and (B) the ultimate attainment approach, which concentrates on testing the limits of adult L2 acquisition. For each approach, three types of evidence have been sought to substantiate the notion of a critical period (2).

(2) Approaches and evidence: Critical periods in L2 Acquisition

(A) Comparative approach (adult L2 versus child L2 learners)  
1) Child L2 acquirers (age of onset inside critical period) should outperform adult L2 acquirers (age of onset outside critical period) in terms of ultimate attainment.  
2) There should be a non-linear, bounded relationship between age of onset and ultimate attainment conforming to the geometric features of a critical period (Figure 1.1).  
3) Child L2 acquirers should show different developmental sequences compared to adult L2 acquirers if different acquisition mechanisms underlie child and adult language development.

(B) Ultimate attainment approach (adult L2 learners versus natives)  
1) Adult L2 learners should not attain native-like proficiency in the target-language.  
2) There should be qualitative differences in neurophysiological representation and cognitive processing between adult L2 learners and natives.  
3) Since critical period effects pertain generally to adult L2 acquisition, non-native-likeness in the L2 should obtain regardless of L1-TL pairings.

In the following, I will review in summary form the approaches and the types of evidence adduced in favour of maturational constraints in adult L2 acquisition to set the scene for the research reported in this thesis. This summary cannot do justice to the wide-ranging
research and writing on critical period effects in L2 acquisition; for more comprehensive reviews, the reader is referred to, e.g., Birdsong (1999b), DeKeyser & Larson-Hall (2005), Hyltenstam & Abrahamsson (2003), Long (1990; 2005) and Singleton & Ryan (2004). As indicated above, the following discussion will zoom in on (morpho-)syntax and ignore the development of, e.g., phonological knowledge and skills.

### 1.3.2. The comparative approach: Evidence and counterevidence

The comparative approach adheres to the logic that qualitative dissimilarities in language development and attainment should obtain according to the age of L2 acquisition, that is, between child L2 acquirers and adult L2 acquirers. Typically, studies framed within the comparative approach test L2 learners at the endstate across a range of ages of onset of L2 acquisition to investigate whether (a) child L2 learners attain higher levels of attainment than adult L2 learners, and (b) whether the geometric features of a critical period (Figure 1.1) can be detected as a function of age for the phenomena tested.

#### 1.3.2.1. The shape of the age function in L2 acquisition: Johnson & Newport (1989) and its replications

By far the most influential study with this set-up is Johnson & Newport (1989). Johnson & Newport (1989) administered a 276-item aural grammaticality judgement task to 46 Chinese and Korean L2 learners of English whose ages of onset, defined as age of arrival in the United States, varied from 3 to 39 years. All participants had had at least 5 years of exposure to English and had lived in the United States for minimally 3 years; on the basis of these data the participants were considered to have reached ultimate attainment in the TL. The study tested diverse aspects of surface English morphosyntax (e.g. past tense, third person, plural and present progressive marking, pronouns, determiners, particle movement, question formation, auxiliary selection, subcategorization and word order) in order to probe general proficiency in L2 English. The results attest that none of the participants with ages of arrival in the US above 15 years of age, termed ‘Late Arrivals’, score within the native speaker control range, whereas the lowest age group of Early Arrivals (3-7 years) is statistically indistinguishable from the native English controls. Johnson & Newport (1989) plot the total scores of the participants against their ages of onset, finding that the results conform to the non-linear geometric shape of critical periods (Figure 1.1). For participants with ages of arrival in the US below 15 years of age, there is a linear correlation between age of onset and performance (r=-.87), whereas for Late Arrivals (>15 years), age of onset does not predict performance, as indicated by a near-zero correlation coefficient (r= -.16, p>0.05) and a random distribution of the scores. In short, age of arrival predicts performance among the Early Arrivals, whereas age of onset bears no predictive relation to performance among later learners. Johnson &
Newport interpret this discontinuity as evidence in favour of a critical period, arguing that maturational constraints terminating a critical period for language acquisition around puberty (15 years of age) enforce an asymmetry in the relation between age and L2 performance.

The Johnson & Newport study has been reanalysed (Bialystok & Hakuta, 1994) and multiply replicated, either in full (Birdsong & Molis, 2001; DeKeyser, 2000) or in part (Bialystok & Miller, 1999; Flege, Yeni-Komshian & Liu, 1999; Johnson, 1992; Kellerman, 1995; McDonald, 2000). Some support the original interpretation, e.g. DeKeyser (2000), while others challenge both major findings of the Johnson & Newport (1989) study, namely, (a) that younger learners categorically outperform post-puberty L2 learners and (b) that there is a non-linear dependency of accuracy on age of arrival.

As for absolute performance differences, there is a sizeable overlap of the accuracy scores for Early and Late Arrivals in the Johnson & Newport (1989) study and all of its replications, suggesting that there is no categorical disparity in potential L2 attainment as a function of age. Moreover, Birdsong & Molis (2001) tested Spanish learners of English, who had been residents in the US for at least 10 years, on all items of the Johnson & Newport (1989) study and report that one post-puberty learner in their study scored within the native-speaker range and thirteen late arrivals achieved more than 92% accuracy on the task. Similarly, some of the 14 post-puberty Spanish learners of English tested by McDonald (2000) perform at levels indistinguishable from native controls (and Flege et al., 1999 for Korean learners of English; see also Kellerman, 1995 for Dutch learners of English).

As for the shape of the age function, in a reanalysis of the original data from Johnson & Newport (1989), Bialystok & Hakuta (1994) demonstrate that when the group distinctions between Early and Late Arrivals are made at a different age outside the bounds of a critical period, namely at age of arrival of 20 years, the strong correlation of age of arrival and accuracy for the younger learners remains (r=-.87), whereas there is a stronger and near-significant contingency between age of arrival and accuracy for the older learners (r=-.49) (see also Elman et al., 1996 for a different modeling of the Johnson & Newport data). A similar result is reported by Birdsong & Molis (2001) in their replication study. Most of the younger learners (ages of arrival below 16 years of age) perform at ceiling, so that no correlation between age of arrival and accuracy arises for the younger (r=-.24, p=.22) group. Yet, a regression analysis reveals that when the cut-off point between learner groups is set at 27.5 years, the data conform to the geometric asymmetry of the Johnson & Newport (1989) data set.

Other studies report significant correlations between age of arrival and accuracy for all age groups (for overview, see DeKeyser & Larson-Hall, 2005). Testing 57 Hungarian learners of English who had lived in the US for a minimum of 10 years, DeKeyser (2000) finds a significant correlation between age of arrival and accuracy
across all learners ($r=-.63$, $p<.001$), yet not for either group of Early or Late Arrivals using the Johnson & Newport classification (for Early Arrivals: $r=-.04$; for Late Arrivals, $r=-.26$). For 240 Korean learners of English with a minimum length of residence in the US of 8 years, Flege et al. (1999) find significant correlations between age of arrival and accuracy equally for Early Arrivals, defined either as younger than 12 or 15 years, and for Late Arrivals. Bialystok & Miller (1999) also report significant correlations for Early and Late Chinese and Spanish Arrivals. In the Birdsong & Molis (2001) study, only the Late Arrivals showed a correlation between age of arrival and accuracy ($r=-.69$, $p<.0001$, see also McDonald, 2006). For a group of 28 Spanish learners of English with a minimum length of residence of 3 years and a group of Vietnamese learners of English with a minimum period of residence in the US of 9 years, McDonald (2000) finds correlations between age of arrival and accuracy across all age ranges (Spanish: $r=-.61$; Vietnamese: $r=-.59$). Other experimental studies on the age function (e.g. Johnson & Newport, 1991; Lee & Schachter, 1997; Weber-Fox & Neville, 1996) have been subject to similar criticisms and reanalyses.

Further, some studies on the self-assessment of English language proficiency of immigrants to the US in census data (e.g. Chiswick & Miller, 2007; Hakuta et al., 2003; Wiley, Hakuta & Bialystok, 2005) find a monotonically decreasing, yet no discontinuous, relation between age and self-assessed proficiency level. Such failure to find discontinuity suggests that age effects on acquisition are not maturationally bounded, or at least not maturationally bounded in a manner conforming to the geometry of critical periods.

In sum, the Johnson & Newport (1989) study and its replications on the shape of the age function in L2 acquisition do not lend unequivocal support to the reality of a maturationally constrained period underlying language acquisition. Instead, the research in response to Johnson & Newport (1989) furnishes evidence that factors other than age of onset can be recruited to explain a large extent of the performance differences between pre- and postpubescent learners. These factors are either independent of age of acquisition, such as levels of education (Kellerman, 1995), task-modality (aural vs written, Bialystok & Miller, 1999; Johnson, 1992), L1-target-language pairings (Birdsong & Molis, 2001; McDonald, 2000), or potentially confounded with age of acquisition, such as use of the L2 (Flege et al., 1999), extra-grammatical memory limitations (McDonald 2000) or lexical decoding ability in post-pubescent learners (McDonald, 2006). I come back to many of these factors in the discussion of the present experiments in Chapter 9.

Finally, the performance of the subjects, especially the Late Arrivals, was found to vary depending on the grammatical construction types tested (e.g. Birdsong & Molis, 2001; Flege et al., 1999; Kellerman, 1995; McDonald, 2000; 2006). For instance, Flege et al. (1999) reclassify the wide-ranging and diverse set of items used by Johnson & Newport into two classes of (a) ‘rule based’ sentences exemplifying regular
morphological inflection and (b) ‘lexically based’ sentences of irregular aspects of English morphology. Not only do these two subclasses yield different scores for the late learners, with lexically based violations proving much harder, the two classes also pattern differently relative to age of arrival and other factors (e.g. education and use of the L2). These findings underscore that different aspects of linguistic knowledge, in particular, different aspects of morphosyntax, can be differentially affected by age and other factors (see also Lee & Schachter, 1997; Weber-Fox & Neville, 1996).

In sum, research on the shape of the age function in L2 acquisition remains inconclusive with respect to the questions (a) as to whether there is a non-linear relation between age of onset and attainment, (b) as to when the offset of critical periods occurs, (c) as to whether age is the primary cause of declining L2 attainment and (d) as to whether maturation equally affects all areas of (second) language proficiency.

1.3.2.2. Adult L2 versus child L2 developmental sequences

Another comparative approach to the issue of critical periods in language acquisition is advocated by Schwartz (1992; 2003; 2004). Schwartz notes that comparing adult and child L2 acquirers in terms of ultimate attainment might miss underlying similarities in cases when adult L2 acquirers fail to reach levels of attainment on a par with child L2 acquirers. Seeing that limitations on L2 ultimate attainment could be due to multiple causes, age-related or not, Schwartz advocates studying parallels in the development of adult and child L2 acquisition. The logic of this approach holds that if there is symmetry in developmental sequences of adult and child L2 acquisition, there is symmetry in the underlying acquisitional mechanisms and thus evidence against maturationally enforced discontinuity between adult and child acquisition. Note that Schwartz limits the scope of developmental comparisons to grammatical phenomena for which children employ mechanisms specific to language acquisition, for instance, those defined by Universal Grammar (UG). These mechanisms should lead to a unique acquisitional sequence that, e.g., general learning mechanisms or other learning algorithms available beyond an alleged critical period could not emulate. In these cases, parallels in developmental sequences between adult and child L2 acquirers provide evidence of the continued availability of the specific learning mechanisms underlying child language acquisition.

A crucial property of the developmental sequences approach is that it is framed within L2 acquisition, because comparing adult and child L2 acquirers allows for controlling variables such as bilingualism and L1 transfer. If developmental divergence is found between adult L2 and child L1 acquirers, such divergence might result from asymmetries between monolingual and bilingual acquisition. In particular, L1 transfer might affect the developmental sequences of L2 acquirers. By comparing adult and child L2 acquirers of the same L1, however, these differences are levelled. Hence, by tracking the development of adult and child L2 learners of identical L1 backgrounds, evidence for
The Critical Period Hypothesis and Second Language Acquisition

or against the continued availability of the learning mechanisms that are operative in childhood can be gathered.

A recent example of a study adhering to Schwartz’ (1992; 2003; 2004) logic is Unsworth (2005), which will be discussed in greater detail in Chapters 5.2 and 5.3. Briefly, Unsworth charts the development of obligatory object movement across negation in a cross-sectional study of adult and child English-Dutch Interlanguage. Both for child and adult L2 learners, three successive and implicational stages, i.e. no object movement, optional movement and obligatory movement, are observed across comparable and matched proficiency ranges (for details, see Chapter 5.2). Other studies testing for identity in developmental sequences between adult and child L2 acquisition investigated German word order by L1 Romance speakers (Schwartz, 1992), English particle verbs by L1 Spanish learners (Gilkinson, 2004), adjectival inflection in L2 Dutch (Weerman, Bishen & Punt, 2003) and finiteness and negation in Russian-German Interlanguage (Jordens & Dimroth, 2006; for other studies, see Unsworth, 2005: Chapter 1). However, the focus of these studies on the course of L2 development means that these studies remain agnostic on the outcomes of L2 development. In other words, they do not address the possibility that maturational effects might not affect the paths of L2 development but rather impose limits on the outcomes of L2 development.

1.3.3. The ultimate attainment approach

The ultimate attainment approach focuses on the core characteristic of critical periods, namely, the time-boundedness of acquisitional mastery. More specifically, it seeks to ascertain whether native-like attainment is possible for post-critical-period learners and thus whether there are underlying cognitive and neurophysiological commonalities between native and non-native ultimate attainment. Broadly speaking, research on ultimate attainment has taken two lines that differ in their approach to the study of native-likeness:

---

\(^2\) Terms other than 'ultimate attainment' have been employed to characterize the study of highly advanced L2 learners, such as ‘near-nativeness’ (e.g. Sorace, 2003), the ‘L2 endstate’ (e.g. Lardiere, 1999b) or ‘fossilization’ (e.g. Han, 2004). As has often been noted, none of these terms seems satisfactory. Dismissing the possibility of native-like outcomes in adult L2 acquisition, ‘near-nativeness’ explicitly suggests that native-like performance can at best be approximated in L2 acquisition, while ‘fossilization’ has predominantly been used to refer to divergent L2 grammars that exhibit noticeable non-target-like properties (for discussion, see Long, 2003). The teleological phrasing of the terms ‘ultimate attainment’ and ‘endstate’ denies further L2 development, be it acquisition or attrition, beyond a given point. For want of a better term, I will adopt the terms ‘ultimate attainment’ and ‘L2 endstate’ to identify the idealized limit of L2 acquisition and use them interchangeably, whilst recognizing the idealization inherent in them. Ultimate attainment and the L2 endstate will thus be defined as denoting an interlanguage system after prolonged and sustained exposure and high levels of proficiency that is structurally stable in the sense that further acquisition other than of vocabulary is not likely.
(A) Descriptive studies seek to establish whether comprehensive native-likeness is attainable for adult learners across a range of constructions, spanning different and unrelated facets of general language proficiency.

(B) Linguistic-cognitive studies attempt to identify the similarities and dissimilarities in the mental representations, cognitive processes and neurophysiological underpinnings of native and non-native grammars and typically focus on specific and theoretically well-defined grammatical phenomena.

Whereas the first approach is broader in scope, the two approaches are not mutually exclusive and overlap to a certain extent. I will first review some of the research within the descriptive approach in order to provide an overview as to whether native-likeness has been found to be attainable for post-pubescent L2 learners. This will serve as a benchmark for the subsequent discussion. Since the experiments in this thesis follow the rationale of the second, linguistic approach, I will review these studies in greater detail and outline the models of L2 acquisition that build on these data.

1.3.4. Descriptive studies

Studies aiming to test for native-likeness across a range of target-language aspects have yielded mixed results. Despite pre-screening of participants for native-like appearance in the L2, earlier studies report that none of the adult L2ers tested scored within native-speaker ranges (Coppieters, 1987; Hyltenstam, 1992).

Coppieters (1987) recruited 21 highly proficient adult learners of French of various L1s, who had all arrived in France after age 18 and who had been referred to Coppieters as native-like L2 speakers of French. The participants were presented with a 107-item oral grammaticality judgement task which probed knowledge of different aspects of French grammar, ranging from universal syntactic constraints on extraction to aspectual properties of particular French verbs. In the quantitative analysis of the judgement task results, all non-natives scored at least three standard deviations from the native control group mean. In the analysis of individual construction types, it was found that non-native deviance was least pronounced for universal syntactic constraints and most pronounced for aspectual properties of verb, thus indicating that language-particular and irregular properties are hardest to master. Analyzing the error frequency of adult and child Finnish and Spanish learners of Swedish in written and oral production, Hyltenstam (1992) reports that for all of the adult learners first exposed to Swedish at or above age 7, the composite error frequency was higher than for any of the native controls.

In a partial replication of Coppieters (1987), Birdsong (1992) recruited 20 native English speakers who had started learning French after age 11 and had lived in France for a minimum of three years. A subset of the items from Coppieters (1987) and items exemplifying complex features of French grammar were tested in a judgement task. Of
the 20 participants, 15 fell within the native controls’ judgement range. Similarly, as noted above, some adult learners in the replications of the Johnson & Newport (1989) study scored within the native-speaker range (Birdsong & Molis, 2001; McDonald, 2000). In the most comprehensive study of individual learners to date, Ioup, Boustagui, El Tigi & Moselle (1994) tested two adult English learners of Arabic in a battery of tasks tapping pronunciation skills and grammatical knowledge. The latter was investigated in a grammaticality judgement, an anaphora interpretation and a translation task. Of the two subjects tested, one achieved native-like accuracy on the anaphora interpretation and the translation task throughout, deviating only mildly from native performance on the grammaticality judgement task. Hyltenstam & Abrahamsson (2003) report that one individual adult learner performed indistinguishably from natives in a battery of tasks on L2 Swedish.

In a study that straddles the line between the descriptive and the linguistic approach, White & Genesee (1996) tested whether adult L2 learners of various L1s performed in the native range on judging and producing complex wh-extractions in English. Long-distance wh-extractions are constrained by universal principles such as Subjacency or the Empty Category Principle (e.g. Rizzi, 1990). Eighty-nine learners of ages of onset ranging from birth to over 16 years of age were screened for native-likeness in oral interviews and assigned to two groups depending on whether they were identified as native-like or not (near-native versus non-native). In a timed grammaticality judgement task and a written question formation task, the near-native L2 group performed on a par with the native controls in all but one condition in terms of judgement accuracy, response time and question-formation accuracy. Moreover, there was no age effect within the near-native group, suggesting that age of onset does not affect accuracy or speed of performance in the tasks. Indeed, a comparison of the latest-arriving near-natives (16+ years) with the native controls yielded no significant differences. By contrast, the non-native L2 group displayed significantly lower accuracy and longer reaction times compared to the natives. The White & Genesee (1996) study documents that native-like attainment is possible for post-pubescent learners even for complex syntactic constructions. However, the fact that most of the participants in the White & Genesee (1996) study were native speakers of French, which exhibits almost analogous restrictions on wh-extraction as English, potentially curbs the generalizability of the results (for discussion, see, e.g., Eubank & Gregg, 1999; Long, 2005).

In sum, more recent studies adhering to the descriptive approach indeed furnish evidence that some post-pubescent learners manage to score within the native speaker range across a variety of tasks. Taken as such, these findings indeed provide individual prima facie counterexamples to the categorical claim of the Critical Period Hypothesis that acquisition outside maturationally privileged time span never results in native-like attainment.
Chapter 1

Critics point out that descriptive studies are constrained by potential limitations in scope of their findings and in terms of the validity of the tasks employed (e.g. DeKeyser, 2000; Eubank & Gregg, 1999; Long, 1990, 2005). In terms of scope, descriptive studies spanning a range of linguistic phenomena invite questions about (a) the extent to which native-likeness is pervasive across different areas of linguistic knowledge, such that comprehensive native-likeness is attainable, (b) the extent to which native-likeness is limited to or dependent on specific L1-TL pairings, and (c) the extent to which the nature of non-native linguistic knowledge conforms to that of native speakers.

In terms of task, critical questions are raised as to (a) whether the task is sufficiently difficult to pose a genuine challenge for achieving native-likeness, (b) whether the task probed native-like implicit, linguistic knowledge or whether it allowed for metalinguistic reflection and the application of explicitly taught information and (c) whether native-likeness attested in a particular task generalizes across tasks.

Several recent studies have attempted to address these issues. Marinova-Todd (2003) recruited 30 postpubescent learners of English of various L1s with a minimum residence of five years in the US. Nine tasks covering pronunciation (e.g. reading aloud), morphosyntax (e.g. a replication of White & Genesee, 1996), lexical knowledge (e.g. an adapted version of the Peabody Vocabulary Test) and language use (e.g. narrative coherence) were administered to the 30 non-native and 30 native participants. The global results demonstrate that 3 of the 30 non-natives score within native speaker ranges across all nine tasks and 6 other participants perform in the native range on seven of the nine tasks. Birdsong (2005c) reports on a multi-task study of 22 adult Anglophone learners of French who had spent at least five years in France. They were subjected to seven tasks covering pronunciation and morphosyntax, all of which were designed to incorporate difficult TL phenomena. Although no single learner scored within the native range across all tasks, three of the 22 participants performed in the native range on five or six of the seven tasks. Based on ten tasks covering phonology, grammar and idiomatic knowledge of L2 Swedish by native Spanish speakers, Hyltenstam (2007) reports that some adult learners score within the native ranges on seven out of ten tasks. Seen in conjunction, then, the studies by Marinova-Todd (2003), Birdsong (2005c) and Hyltenstam (2007) suggest that adult learners can in some cases attain wide-ranging native-likeness across tasks and linguistic skills.

Testing adult learner groups of Dutch from distinct L1 backgrounds, van Boxtel (2005) investigates dummy subject constructions that are deemed to be hard to acquire, not least because they exhibit many language-particular properties and are not subject to instruction. In a sentence-preference judgement task and a sentence-imitation task, the performance of 15 L1 German, 15 L1 French and 13 L1 Turkish postpubescent learners of Dutch who were screened for superior proficiency was compared to that of 44 native speakers of Dutch. In the tasks, 3 L1 German, 4 L1 French speakers and one L1 Turkish speaker performed within the native-speaker range and were not statistically
distinguishable from native speakers. In line with White & Genesee (1996), then, van Boxtel (2005) shows that native-likeness is attainable in syntax for late L2 learners and extends this finding to typologically more distant L1-TL pairings such as Turkish-Dutch.

While these studies present an important step forward in documenting native-likeness in late L2 learners, they are in principle incapable of overcoming the inherent limitations of the descriptive approach that aims at documenting native-likeness at the behavioural level.

First, by virtue of being a holistic concept, native-likeness cannot be isolated in one task or even across multiple tasks, so that the goalpost of native-likeness keeps getting moved or redefined (e.g. Birdsong, 2005b; Long, 2005). No matter how wide-ranging a battery of tasks, the possibility of non-native-likeness in other areas or tasks can thus never be excluded.

Second, even a wide-ranging battery of unrelated tasks that probes different (levels of) linguistic skills is not informative with respect to the nature of linguistic knowledge that is acquirable in adult L2 acquisition. In the absence of a theoretical foundation of what it is that needs to be acquired in order to be native-like, i.e. a property theory of linguistic knowledge (see, e.g., Gregg, 1996), identifying native-like behaviour in language tasks does not speak to the issue of whether non-native grammatical representations and processing are equivalent to those of native speakers. Conversely, finding non-native-like behaviour does not imply that the non-native grammatical representations and processing are necessarily epistemologically distinct from those of native speakers (Schwartz, 1987, see also Chapter 9). The issue of qualitative identity between native and non-native competence has been studied within the linguistic-cognitive approach to L2 ultimate attainment.

1.3.5. The linguistic-cognitive approach

The linguistic-cognitive approach takes the constructs and categories of linguistic theory as the starting point for investigating similarities and dissimilarities between native and non-native grammars. As a consequence of this conceptual shift, the criterion of ‘native-likeness’ is replaced by the criterion of ‘target-likeness’ or ‘convergence’. Convergence can be defined as conforming to the particular grammatical constraints of the target language under investigation. By recasting the criterion of native-likeness in terms of adherence to grammatical constraints and representations of the target language, similarities and dissimilarities between mature native and endstate non-native grammars can be related to the mental architecture of grammar and language processing, e.g. as described in generative theory.

On the basis of learnability considerations in first language acquisition, generative models of grammar posit the existence of a narrowly restricted set of universal, innate linguistic constraints (Universal Grammar) and a modular organization of the mind that
includes a designated language module (the language faculty). The following section gives a brief overview of the modular organization of grammar.

1.3.5.1. Modularity and interfaces

According to the recent version of the theory, the Minimalist Program (Chomsky, 1995; 2000; 2001), the language faculty consists of autonomous submodules, with a computational syntactic module ($C_{HL}$) connecting the lexicon with the interfaces of the output modules, Logical Form (LF) for interpretation and Phonetic Form (PF) for phonetic spell-out. These, in turn, feed the language external sensorimotor, articulatory-perceptual system and the conceptual-intentional system, respectively. Figure 1.2 illustrates the architecture of grammar in the Minimalist Program.

![Figure 1.2. The organisation of grammar in the Minimalist Program.](image)

A basic requirement on the computational system $C_{HL}$ is that it satisfies the legibility conditions at the interfaces, i.e. it delivers its outputs in a form legible to the interfaces (Chomsky, 2000). It is thus assumed that the computational system generates expressions $\text{Exp} = <\text{Phon}, \text{Sem}>$, where Phon corresponds to information legible to PF and Sem corresponds to information interpretable at LF. Since the interfaces of syntax with other linguistic and extra-linguistic modules will be the focus of investigation in thesis, I briefly outline the characteristics of the syntactic interfaces in the following.

The Lexicon

On Minimalist assumptions, the lexicon stores lexical items in a format that is legible to the computational system. Lexical items consist of three types of features: phonological features (e.g. [-back]), semantic features (e.g. thematic features, [+anim]), and morphosyntactic features (e.g. [-past]) (Chomsky, 1995: 277). Of the three featural types,
only morphosyntactic features, also called formal features, are accessible to the syntactic computation $C_{HL}$. These features encode cross-linguistic parametric variation (e.g. Borer, 1984). Formal features are associated with functional, i.e. inflectional, categories. Formal features come in two types: interpretable features that contribute to semantic interpretation (e.g. categorial features, $\Phi$-features of nominals) and features that are uninterpretable at LF (such as the EPP feature, $\Phi$-features, Case, etc.). Solely the specific requirements of uninterpretable features drive the derivational process. The lexical items selected for expressing a particular proposition are grouped into the lexical array.

**Syntax**

The syntactic derivation of a sentence starts out with $C_{HL}$ recursively concatenating items of the lexical array in the operation *Merge* which yields new syntactic objects. Another syntactic operation is *Move* which rearranges syntactic objects. *Move* is tightly constrained by economy considerations. Only the parametrically specified values of formal features may occasion movement, thus leading to cross-linguistically variant surface syntax. A third operation, *Agree*, checks and deletes the features of functional heads and lexically marked categories that have been conjoined by *Merge* or *Move* against each other, thus locally licensing derivations. Once the uninterpretable features have been checked and deleted by virtue of the three syntactic operations, the syntactic derivation is complete. As by the legibility requirement, the syntactic derivation converges if it contains only features that are interpretable at the interfaces.

**The Syntax-Morphology Interface**

Within the original Minimalist model (Chomsky, 1995), morphology is allocated to the lexicon which stores complete lexical items, including their inflectional and phonological characteristics. In Distributed Morphology (DM) (Embick & Noyer, 2007; Halle & Marantz, 1993; 1994), a non-lexicalist framework partially espoused in the latest versions of the Minimalist Theory (e.g. Chomsky, 2001), the traditional view of the lexicon has been dissolved and, instead, morphological processes are distributed across different components of grammar. Word formation is syntactic or post-syntactic. Derivational morphology takes place in syntax. As for inflection, syntax only operates on bundles of features that are not yet associated with overtly realized morphological inflection. Inflection is inserted post-syntactically in a derivation (Late Insertion) on the branch from syntax to PF. Lexical insertion proceeds by way of competition between vocabulary items in which the item bearing inflectional features, e.g. person, number, case, that best matches the featural array of a terminal syntactic node is chosen.
Chapter 1

The Syntax-PF interface

At the interface with PF, syntax delivers its computational output to the sensorimotor system in a format legible to the articulatory-perceptual system. Chomsky (2000) defines information legible at PF as temporally ordered, with prosodic and syllable structures and endowed with specific phonetic features and information about their relations (Chomsky, 2000: 94). At PF, phonological rule systems independent of syntax then specify how to read relevant pieces of information off the syntactic output and translate them to sensorimotor outputs.

The Syntax-LF interface

At the LF interface, the syntactic output is fed to the conceptual-intentional system for interpretation. On the branch to LF, in what is commonly referred to as covert syntax, the computational system can operate on interpretable features to ensure convergence at LF without affecting outputs at PF. Information legible at LF includes certain arrays of semantic features, event and quantificational structure, etc. (Chomsky, 2000: 94).

The conceptual demarcation of the language faculty into discrete, though interacting, linguistic modules within the Minimalist Program has been employed as the theoretical background of L2 acquisition research. Before I review previous research on L2 ultimate attainment at the interfaces, I outline the rationale of generative L2 acquisition research in the following sections.

1.3.5.2. Generative L2 acquisition: From Poverty of the Stimulus to the difference-oriented approach

The central contention of generative acquisition research has been the stipulation of a domain-specific and genetically encoded blueprint of possible natural-language grammars to account for the speed, effortlessness, relative developmental uniformity and universal success of L1 acquisition in spite of varying, finite and degenerate input (e.g. Crain & Thornton, 1998; Hornstein & Lightfoot, 1981). A large range of linguistic and psychological developmental studies has supported the validity of this claim for native language acquisition (see, e.g., Crain & Lillo-Martin, 1999; Guasti, 2002). In recent decades, the relevance and applicability of Universal Grammar to (adult) L2 acquisition has been demonstrated in a large body of studies (for overview, see, e.g., Hawkins, 2001a; White, 2003b). As in L1 acquisition, the conceptual motivation for postulating innate grammatical knowledge in (adult) L2 acquisition comes from the logical problem of language acquisition.
The logical problem of language acquisition (e.g. Hornstein & Lightfoot, 1981) arises if speakers display subtle and complex grammatical knowledge that could not have been induced from the available input in combination with non-linguistic cognitive learning strategies so that the involvement of pre-given linguistic knowledge is logically implicated. In L2 acquisition, an analogous Poverty of the Stimulus (POS) constellation obtains if two conditions are met (e.g. White, 1989, 2003b):

- A grammatical constraint must not be instantiated in the learner’s L1, so that knowledge of it cannot be accessed via the L1
- This constraint must be underdetermined in the TL input, i.e. neither relevant positive nor negative evidence is available to the learner in the input or instruction

A by-now large number of studies has examined POS phenomena in the areas of distributional syntax and the syntax-semantics interface in adult L2 acquisition (for overview, see Schwartz & Sprouse, 2000; White, 2003b: Chapter 2). For syntax, adult L2 speakers consistently display empirically underdetermined knowledge on restrictions on long distance wh-movement in L2 English (e.g. Martohardjono, 1993; Uziel, 1993; White & Juffs, 1998) and optional word orders in L2 German (Hopp, 2005; Schreiber & Sprouse, 1998), constraints on quantifier binding in L2 Japanese and L2 Spanish (e.g. Kanno, 1996; 1997; Pérez-Leroux & Glass, 1999) and the binding of reflexives in L2 English (e.g. Finer & Broselow, 1986; Hamilton, 1998). At the syntax-semantics interface, adult L2 learners demonstrate subtle knowledge on interpretive contrasts in quantifier extraction and nominal reference in L2 French (e.g. Dekydtspotter, Sprouse, & Anderson, 1997; Dekydtspotter, Sprouse, & Anderson, 1998; Dekydtspotter, Sprouse, & Thyre, 1999/2000; Marsden, 2004; for overview, see Slabakova, 2006). I take these studies as furnishing strong empirical evidence of the involvement of Universal Grammar in adult L2 acquisition and as providing conceptual motivation for employing the architecture and categories of generative grammar in studying IL systems.

POS studies provide compelling evidence against the most straightforward and comprehensive linguistic approach to critical periods, namely, the contention that maturational constraints enforce discontinuity in the availability of Universal Grammar for adult L2 acquisition (e.g. Bley-Vroman, 1990; Clahsen & Muysken, 1986; Meisel, 1997; Neeleman & Weerman, 1997). I will thus not consider this option any further.

Whereas POS-oriented studies place adult L2 acquisition in the generative framework by showing that the same grammatical architecture supports child L1 and adult L2 acquisition, they offer little in the way of delineating systematic differences between L2 and L1 acquisition. Hawkins (2001b) argues that progress in identifying the characteristics and limitations of adult L2 acquisition within generative research can better be gained by what he terms ‘difference-oriented’ studies (see also White, 2000). These are studies that systematically aim at uncovering whether and how components of
grammar or their interactions with one another or with other components of the mind differ in their operation between L1 and adult L2 acquisition.

Difference-oriented studies employ (some version of) the grammatical architecture in Figure 1.2 as the starting point for exploring specific hypotheses about maturational effects in L2 acquisition. Against this backdrop, they investigate whether critical period effects can be located in:

- particular grammatical modules or definable subparts of modules
- the interfaces of grammatical modules
- the interface of grammar with performance systems, e.g., the parser

This thesis is committed to the difference-oriented approach in trying to uncover systematic similarities and dissimilarities between mature native grammars and non-native grammars and in relating them to architectural or procedural properties of Interlanguage systems. The following sections of this chapter will thus be devoted to outlining a number of the linguistic approaches in order to locate age-related asymmetries between native and non-native grammars at ultimate attainment.

### 1.3.6. Guises and causes of non-convergence

#### 1.3.6.1. Limits on L2 ultimate attainment: The Interface Hypothesis

Accounts of limitations on L2 ultimate attainment can be distinguished according to the nature (representational versus computational) of limitations, their cause (impairment versus inefficiency versus L1 interference) and the domain of their applicability (grammatical (sub-)modules versus interfaces). Figure 1.3 gives a schematic overview of where different approaches situate limitations on convergence at ultimate attainment.

![Diagram](image_url)

**Figure 1.3. Overview of approaches to non-convergence at L2 ultimate attainment.**

As for the nature of limitations on ultimate attainment, representational deficit theories assume that non-convergence can be identified in non-target-like grammatical representations, whereas computational approaches locate non-convergence in the
computation of grammatical knowledge. As for causes, three factors have been suggested: (1) impairment, i.e. a particular component of the grammatical or procedural architecture is divergent or incomplete compared to that of native speakers as a consequence of a critical period; (2) inefficiency, i.e. the processes used for computing grammatical representations are less automatized compared to those of native speakers; and (3) L1 interference, i.e. there is L1 influence on grammatical representations or their computation. These three factors potentially coexist or interact in causing non-convergence. As for domains of application, any of the causes could affect grammatical modules or specific parts thereof, their interaction at the interfaces, or the use of different types of grammatical knowledge in real time.

The descriptive studies on L2 ultimate attainment reviewed above already point to differential domains of non-convergence. Studies that focus on syntax (van Boxtel, 2005; White & Genesee, 1996) report a sizeable overlap of late L2 learners and natives. In studies that test a range of constructions, L2ers perform indistinguishably from natives on aspects of word order phenomena of the TL (Coppieters, 1987; McDonald, 2000; 2006), whereas they exhibit differences in inflectional morphology (Johnson & Newport, 1989 and its replications) and semantic aspects of syntax (Coppieters, 1987). Hence, ‘narrow’ syntax seems to be relatively spared from persistent non-convergence, whereas the coordination of syntax with other grammatical information, e.g. morphological, semantic, seems prone to non-convergence. The generative framework in Figure 1.2 captures these findings in its distinction between syntax and its interfaces. Accordingly, Sorace & Filiauci (2006) term the disjunction between target acquisition of syntax and non-target-like interfaces at L2 ultimate attainment the ‘Interface Hypothesis’.

The Interface Hypothesis describes non-convergence at the syntactic interfaces as expressed in residual optionality of target and non-target grammatical forms, persistent L1 effects or indeterminacy in grammatical performance. At present, the Interface Hypothesis does not constitute a model of L2 acquisition or L2 ultimate attainment; rather, it defines the typical locations of non-convergence in adult L2 acquisition. With reference to Figure 1.3, the Interface Hypothesis needs to be specified in terms of domain to answer the question as to whether all interfaces or which interfaces suffer persistent non-convergence, how non-convergence can be defined in terms of its causes (impairment, inefficiency, L1 interference) and its nature (representational, computational).

In the remainder of this chapter, I will review existing research on L2 ultimate attainment at the syntactic interfaces as summarized in Table 1.3 in order to identify the

---

3 The Interface Hypothesis formulated by Sorace and colleagues further attempts to integrate aspects of non-convergence in adult L2 acquisition with non-convergene in bilingual language acquisition, L1 attrition and diachronic language change (see Sorace, 2005). These linkages go beyond the remit of this thesis, and will thus not be considered further.
scope and the causes of non-convergence and the models of L2 ultimate attainment associated with them.

I review studies that conclude that L2 grammars are representationally impaired in specific submodules (Franceschina, 2005; Sorace, 2003) and studies arguing for procedural constraints on the integration of information across different grammatical modules (Prévost & White, 2000b; White, 2003a). In addition, a study highlighting the effects of L1 transfer as a cause of divergent outcomes in L2 acquisition will be outlined (Sorace, 1993) as well as a study on the coordination of syntactic and discourse information (Belletti et al., 2007).

<table>
<thead>
<tr>
<th>Interface</th>
<th>Phenomenon</th>
<th>Model</th>
<th>Claim</th>
<th>Critical Period</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Syntax-Morphology</strong></td>
<td>Grammatical gender</td>
<td>‘Failed Functional Features’/</td>
<td>L1 determines convergence on morphosyntax</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>‘Representational Deficit’ Hypothesis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Verbal Inflection</td>
<td>Mapping/Morphological Competence</td>
<td>Morphological underspecification</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Missing Surface Inflection Hypothesis</td>
<td>Computational limitations in lexical access</td>
<td>No</td>
</tr>
<tr>
<td><strong>Syntax-Semantics</strong></td>
<td>Aspect, Unaccusativity</td>
<td>Full Transfer/Full Access</td>
<td>L1 interference</td>
<td>No</td>
</tr>
<tr>
<td><strong>Syntax-Discourse</strong></td>
<td>Focus</td>
<td>Interpretable Features</td>
<td>L1 determines convergence on interpretive properties of syntax</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Computational limitations in information integration</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 1.3. Previous approaches to non-convergence at L2 ultimate attainment.
1.3.6.2. The syntax-morphology interface

It has often been observed that inflectional morphology, in particular bound morphemes expressing, e.g., tense, verbal agreement, etc. and closed-class elements, such as determiners and auxiliaries, pose protracted difficulty for adult L2 learners. For instance, early studies on morpheme acquisition orders note that the acquisition of inflectional morphology is a slow and gradual process (for review, see Zobl & Liceras, 1994) and inflectional morphology, alongside phonology, is considered a predominant feature of non-native-likeness (e.g. Long, 1990) and fossilization (e.g. Han, 2004; White, 2003a) among adult L2 speakers.

In generative theory, inflectional morphology is closely linked to parametric variation in syntax. Syntactically, inflectional morphology is represented in uninterpretable features hosted in functional heads, e.g. a [+Tense] feature can be realized in the functional head T and can attract a concomitantly [+Tense]-marked verb, thus occasioning verb raising. Morphologically, inflectional morphology is spelled out according to the language-particular morphological paradigms. It has been a matter of debate whether the abstract relationship between syntax and morphology also implicates an ontological relationship between syntactic representations (e.g. the [+Tense] feature), and morphological forms (e.g. –ed suffixation for past tense in English). Different answers to this questions amount to different conceptualizations of the syntax-morphology interface.

At one end of the spectrum of approaches, the ‘Rich Agreement’ Hypothesis (e.g. Rohrbacher, 1999; Vikner, 1997), based on typological studies, holds that identification of the richness and complexity of morphological paradigms determines the feature strength of functional features in syntax. In acquisition work, corollaries of this approach posit that acquisition of the morphology of inflection precedes or equates acquisition of its associated abstract syntactic features (e.g. Beck, 1998; Clahsen, Eisenbeiss & Vainikka, 1994; Eubank & Grace, 1998; Hawkins & Chan, 1997; Vainikka & Young-Scholten, 1994). For ultimate attainment, these models hold that problems with inflectional morphology index permanent non-convergence in syntax at the L2 endstate (Beck, 1998; Hawkins & Chan, 1997; Tsimpi, 2003; Tsimpi & Dimitrakopoulou, 2007).

At the other end of the spectrum, separationist approaches dissociate syntactic features and morphological spell-outs, arguing that syntactic features are not contingent on morphological realization; in acquisition work, these approaches hold that functional features and their syntactic consequences can be instantiated irrespective of whether their associated morphological forms have been acquired (e.g. Haznedar & Schwartz, 1997; Lardiere, 1998b; 2000; Prévost & White, 2000b). Incomplete acquisition of inflectional morphology does not reflect incomplete acquisition of syntax; rather, non-convergence in inflectional morphology signals incomplete acquisition of (a) morphology or (b) the
mapping between syntax and morphology (for overview, see Lardiere, 2000; White, 2003b: Chapter 6).

With a view to establishing the nature of persistent problems with inflectional morphology in adult L2 acquisition, it is hence essential to consider morphological and syntactic aspects of inflectional morphology to determine whether there is a dissociation between them in endstate L2 acquisition. In the remainder of this section, I will first outline a non-separationist approach to deviance in inflectional morphology in gender marking (Franceschina, 2005); subsequently, I review separationist approaches for tense and verbal agreement (Lardiere, 1998b; Prévost & White, 2000b).

1.3.6.2.1. Representational impairment in syntax? – the ‘Failed Functional Features’/’Representational Deficit’ Hypothesis and grammatical gender

Morphologically, grammatical gender divides nouns into nominal classes. Syntactically, gender triggers agreement between nouns and other syntactic categories (e.g. determiners, adjectives, verbs, participles, pronouns, etc.). Gender agreement is construed as a multiple feature checking process in syntax. In Spanish, for instance, which distinguishes between masculine and feminine gender, complex nominals are multiply marked for gender (4).

(4) las pequeñas mesas blancas
theFEM smallFEM tableSFEM whiteFEM
‘the small white tables’

It is commonly assumed that determiner, noun and adjectives are concatenated in structural configurations that allow for feature checking and deletion under Agree (e.g. Carstens, 2000; Valois, 1991). Gender agreement also underlies the identification of pronouns; for instance, pronoun reference in the question in (5) is identified by gender agreement with the subject of the preceding sentence.

(5) Los dos enchufes que compré estaban fallados. ¿Será posible cambiarlos por unos nuevos?
‘The two plugs I bought were faulty. Could I change them for new ones?’

Errors in gender marking and gender agreement are well attested in intermediate and advanced learners (e.g. Bruhn de Garavito & White, 2000; Hawkins, 2001b; Sabourin, 2003; White, Valenzuela, Kozlowska-MacGregor, Leung & Ben Ayed, 2001). Moreover, L2 learners with L1s instantiating grammatical gender outperform learners with L1s not marking gender; yet even the latter attain relative success in L2 gender marking and agreement (Sabourin, 2003; White et al., 2001).
For gender agreement at L2 ultimate attainment, Franceschina (2001) reports a case study of an adult English learner of Spanish with a length of residence in Spanish-speaking countries of 24 years. In the transcription of spontaneous speech elicited in an interview with this learner, 94 minutes and 1539 clauses are covered. For adjective-noun agreement, 53 errors in 688 contexts (7.8%) are attested, with 41 errors reflecting inaccurate gender agreement and 12 errors owing to inaccurate number agreement. In a larger-scale study, Franceschina (2005) compares knowledge and use of grammatical gender agreement among endstate adult L2 Spanish speakers of L1s that also encode gender and L1 English speakers across a battery of tasks. These tasks comprised an oral interview, several cloze and multiple choice tasks, a grammaticality judgement task and a task in which subjects had to assign nonce words gender based on a variety of morphophonological cues (see Table 1.4). Sixty-eight postpubescent non-native speakers of Spanish participated, although the number of subjects varied across tasks. In total, 53 speakers had L1s that mark grammatical gender (Arabic, French, German, Greek, Italian, Portuguese), the [+gen] group, and 15 speakers had English as their L1, the [-gen] group. Prior to the tasks on gender, the participants were tested for general proficiency on a 43-item multiple choice test and only those that scored within the native range were recruited for the subsequent experiments. Length of exposure and length of residence were not considered as variables, although participants must have had some naturalistic exposure to Spanish and stayed in a Spanish-speaking country at some point.
<table>
<thead>
<tr>
<th>Test</th>
<th>Task</th>
<th>Type of gender agreement</th>
<th>Natives</th>
<th>[+gen] group</th>
<th>[-gen] group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender assignment</td>
<td>multiple choice</td>
<td>det-N</td>
<td>99.62%</td>
<td>99.29%</td>
<td>99.01%</td>
</tr>
<tr>
<td>Interview</td>
<td>spontaneous production</td>
<td>adj-N</td>
<td>100%</td>
<td>100%</td>
<td>90.35%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>pron-N</td>
<td>100%</td>
<td>100%</td>
<td>86.99%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>det-N</td>
<td>100%</td>
<td>100%</td>
<td>92.71%</td>
</tr>
<tr>
<td>Pronoun-noun association</td>
<td>multiple choice</td>
<td>pron-N</td>
<td>91.8%</td>
<td>86.4%</td>
<td>76.3%</td>
</tr>
<tr>
<td>Inserting missing pronouns</td>
<td>fill-in-the-gap task</td>
<td>pron-N</td>
<td>98.9%</td>
<td>96.9%</td>
<td>94.3%</td>
</tr>
<tr>
<td>Cloze test: Missing N or</td>
<td>multiple choice</td>
<td>adj-N</td>
<td>99.4%</td>
<td>98.9%</td>
<td>95.8%</td>
</tr>
<tr>
<td>adj</td>
<td>(missing N)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>adj-N</td>
<td>96.7%</td>
<td>97.8%</td>
<td>75.9%</td>
</tr>
<tr>
<td>Grammaticality Judgement</td>
<td>off-line judgements</td>
<td>det-N-adj</td>
<td>93.2%</td>
<td>94.9%</td>
<td>90%</td>
</tr>
<tr>
<td>Task</td>
<td></td>
<td>det-adj-N</td>
<td>91.9%</td>
<td>93.8%</td>
<td>89.6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>det-N/pron</td>
<td>94.6%</td>
<td>96.8%</td>
<td>81.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>det-N/adj</td>
<td>96.5%</td>
<td>94.2%</td>
<td>86.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>pron-adj</td>
<td>99.5%</td>
<td>98.1%</td>
<td>82.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>pron-pron</td>
<td>98.1%</td>
<td>97.3%</td>
<td>85%</td>
</tr>
</tbody>
</table>

Table 1.4. Overview of results from Franceschina (2005). Results converted to percentage of accuracy. ‘Adj’ is for adjective, ‘pron’ is for pronoun, ‘det’ is for determiner and ‘N’ is for noun.

In a gender assignment task, determiners had to be provided for the 112 nouns used in the other experiments, and all groups scored above 99%. Hence, there were no significant differences between the [-gen] and the [+gen] groups in knowledge of grammatical gender in L2 Spanish. In all tasks probing gender agreement, however, there were significant differences in accuracy between the [+gen] and the [-gen] L2 groups. Table 1.4 gives an overview of the diverse tasks and lists the accuracy of the groups on the

---

Note that interview data were only analysed for fifteen participants, i.e. five native, five [+gen] and five [-gen] participants, so that the production data come from a very limited subsample of the groups, consisting of less than ten percent of the [+gen] subjects and just one third of the [-gen] subjects.
The Critical Period Hypothesis and Second Language Acquisition

respective tasks in Franceschina (2005). For ease of comparison, all scores across tasks have been converted to percentage of accuracy. Although the [-gen] group reached very high levels of accuracy, the [+gen] group consistently outperformed the [-gen] group. The [+gen] group and the native control group significantly differed in accuracy only on the pronoun-noun association task; in all other tasks, their accuracy scores were statistically indistinguishable.

The findings for the [-gen] group suggest that speakers of L1s that do not encode grammatical gender for lexical nouns have persistent problems with gender agreement that cannot be reduced to incomplete knowledge of gender. Franceschina (2005) interprets the findings as evidence of representational impairment in L2 acquisition along the lines of Hawkins and Chan’s (1997) ‘Failed Functional Features’ or ‘Representational Deficit’ Hypothesis (see also Hawkins, 2001a; Tsimpili, 2003). According to Hawkins & Chan (1997), there is a critical period for functional/formal features in syntax, such that postpubescent learners are restricted to the inventory of functional features instantiated in the L1. Since acquiring ‘new’ functional features or reconfiguring their values is maturationally ruled out, these adult learners fail to attain native-like performance on morphosyntax. In lieu of restructuring grammars, adult L2ers emulate the TL settings by modeling the TL system within the (defective) set of functional features of the L1 or by using non-grammatical compensatory strategies to approximate the TL. According to Franceschina (2005), the residual problems of the [-gen] group thus have their origins in incomplete syntactic representation in that the [-gen] learners do not encode a gender feature in their IL-grammars, such that they will categorically fail to converge on the TL. The comparatively high degree of target-like performance of the [-gen] group is attributed to compensatory strategies that involve explicitly memorized inflectional paradigms and the reliance on a default, e.g. masculine gender (see also Hawkins, 2001b).

The Franceschina (2005) study provides a commendable example of a linguistic study testing a clearly delineated grammatical phenomenon across a battery of tasks. However, several points in the selection of participants and in the interpretation of the results undermine the conclusions drawn. In participant selection, the [+gen] and [-gen] groups differ in their make-up according to L1s as well as according to other participant characteristics. As for L1s, the [-gen] group (n=15) consists exclusively of L1 English speakers, while the considerably larger [+gen] group (n=53) comprised speakers of Arabic, French, German, Greek, Italian and Portuguese. More than four-fifths of the [+gen] subjects (44 out of 53) had French, Italian or Portuguese as L1. Typologically, Romance languages show extensive surface similarity in the morphophonology of gender marking. This parallelism allows for the surface transfer of identical morphophonological gender cues for the [+gen] group, so that it is not obvious whether the native-like performance of the [+gen] group is tied to the grammatical encoding of gender or to
surface L1-TL similarities.\(^5\) For lower-level advanced L2 learners of Dutch, Sabourin (2003) finds that surface similarity in gender selection between German and Dutch confers on L1 German speakers a significant advantage over L1 Romance speakers for Dutch gender, even though both German and Romance languages grammatically encode gender. If surface similarity, rather than the abstract grammatical representation of gender, modulates the success on L2 gender marking, the differences between the [+gen] and the [-gen] groups can be accounted for independently of representational impairment in syntax as a consequence of a critical period.

As for interpretation of the findings, Franceschina (2005) argues that two patterns in the results point to L1-specific representational impairment: First, problems with gender agreement are specific to the [-gen] group and do not characterize adult L2 acquisition in general. Second, the disadvantage of the [-gen] group obtains across tasks, irrespective of whether they involve production, comprehension or different types of metalinguistic tasks. Hence, L1-specific problems with gender are robust in that they are not affected by task demands or different computational efforts required.

Upon closer inspection, however, the results from some tasks indicate that performance differs according to task demands for all groups. Consider the pronoun-matching task, in which two types of gender markers were employed: canonical gender markers which show transparent gender by ending in –a for feminine and –o for masculine, and non-canonical gender markers. The results show that canonicity has an effect for all groups, i.e. also for the native control group (see Table 1.5).

\(^5\) As for other biographical variables, participants from the [+gen] and [-gen] groups differ rather substantially in age (20-89) and L2 background which might in themselves account for the attested differences. The [+gen] and [-gen] groups for the two tasks where the starkest asymmetries were found (pronoun-noun association and the cloze test) differed strongly in mean age ( [+gen]: 55.15 years; [-gen]: 40.53 years). Moreover, while participants for the [+gen] group were recruited and tested in Spanish-speaking countries, a number of L1 English (i.e. [-gen]) learners were recruited from the student population at a British university, which, seen in conjunction with the age differences, suggests that the [-gen] group enjoyed considerably shorter exposure to Spanish and shorter lengths of residence in Spanish-speaking countries compared to the [+gen] group. Unfortunately, Franceschina (2005) provides no information about age of onset, length of exposure or length of residence of any group or any individual. Yet, potential differences in these variables undermine the comparability of the groups, with the [-gen] group being at a consistent disadvantage. Franceschina (2005) matches the groups exclusively based on scores in the top quartile of a 43-item multiple choice cloze task, which seems insufficient to ensure comparability across groups.
Table 1.5. Results from pronoun-noun association task in Franceschina (2005) by type of gender marking.

<table>
<thead>
<tr>
<th>Test</th>
<th>Task</th>
<th>Type of gender agreement</th>
<th>Natives</th>
<th>[+gen] group</th>
<th>[-gen] group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pronoun-noun</td>
<td>multiple choice</td>
<td>pron-N/all</td>
<td>91.8%</td>
<td>86.4%</td>
<td>76.3%</td>
</tr>
<tr>
<td>association</td>
<td></td>
<td>canonical gender marking</td>
<td>96.1%</td>
<td>91.4%</td>
<td>81.6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>non-canonical gender marking</td>
<td>87.5%</td>
<td>81.5%</td>
<td>70.9%</td>
</tr>
</tbody>
</table>

The relatively poor performance of even the natives on non-canonical gender marking calls the conceptual basis of representational impairment approaches into question. The line of argumentation of this approach, also explicitly adopted by Franceschina (2005), dictates that non-target-like performance is a direct reflection of non-target-like grammatical representations. For non-canonical gender markers, Franceschina (2005) finds that the accuracy of the native control group is 87.5%; i.e. 12.5% of the native responses are non-target-like. Since this reduced accuracy of the natives can by definition not be attributed to their lack of knowledge of grammatical gender agreement, their decreased performance likely reflects task-related effects or other performance factors depressing accuracy in accessing morphosyntactic gender representations. However, admitting causes other than representational impairment to account for more than negligible deviance voids the impairment approach of its logical consistency. If native deviance is due to factors other than grammatical impairment, such as computational problems in lexical access, L2 deviance could be due to these factors, too, especially if, as in Franceschina’s (2005) data, the margin of overall non-native deviance of the [-gen] group roughly corresponds to the extent of divergence attested for non-canonical gender markers in natives.

In other words, there is no evidence that the divergence of the [-gen] group exclusively relates to representational impairment in syntax, rather than to whatever combination of extra-grammatical factors conspires to depress native behaviour or the behaviour of the [+gen] group in the tasks used. It is instead perfectly plausible that, e.g., extra-grammatical problems in lexical access affecting native speakers are magnified for the L1 [-gen] group, for example, as a consequence of the unavailability of surface analogies in gender, lower degrees of routinization of lexical access for inflection or...
concord in the L1 and possibly as a result of shorter exposure to the TL than the [+gen] group.

In sum, the ‘Failed Functional Features’ Hypothesis (Hawkins & Chan, 1997), adopted by Franceschina, predicts categorical effects of the L1 on L2 ultimate attainment in that L2 grammars are representationally limited to the L1 syntactic repertoire as a result of a critical period. Although robust L1 effects are reported in Franceschina (2005), the selection and grouping of participants as well as the data allow for no definitive conclusions (a) as to whether the behavioural differences between the [+gen] and the [-gen] group stem from grammatical L1 differences and (b) as to whether the deviance of the [-gen] group reflects divergence at the level of syntactic representation or computational problems that also affect natives and [+gen] L2ers, although to a lower degree. These problematic issues in the Franceschina (2005) study illustrate that looking at accuracy rates of suppliance or of judgements of inflection alone cannot distinguish between representational impairment and computational approaches to inflectional difficulty at the L2 endstate (see also Jiang, 2004; Trenkic, 2007). Instead, evidence in this domain should be sought from tasks that disentangle the representation and the processing of morphology.

1.3.6.2.2. Mapping problems at the interfaces? – Morphological Competence, the ‘Missing Surface Inflection Hypothesis’ and tense and verbal agreement

Beyond their morphological marking, tense and verbal agreement serve syntactic functions. Syntactically, abstract tense (finiteness) and agreement features are taken (a) to license nominative case on subjects and (b) to govern verb raising, i.e. movement of finite verbs across adverbs, negation and in questions (Pollock, 1989).

The differences between morphological realization and syntactic repercussions of tense and verbal agreement have been much studied in adult L2 acquisition (e.g. Eubank, 1996; Eubank & Grace, 1998; Herschensohn, 2001; Prévost & White, 2000b). To the best of my knowledge, there has been no systematic investigation of inflectional morphology in the verbal domain in near-native L2 acquisition. In the replication of the Johnson & Newport (1989) study by Flege et al. (1999), Korean speakers of English across ages of onset have more difficulty with irregular inflectional morphology than with regular inflection. This difference points to selective difficulties with morphological marking; yet it does not speak to the relation between morphology and syntax.

Two case studies on fossilized endstate learners who display persistent and pervasive divergence in suppliance of target-like inflectional morphology address the relation between morphology and syntax. Lardiere (1998a; 1998b; 2000; 2005; 2007) analyses oral and written production data from an adult Chinese learner of English, Patty, who was recorded ten years after she had come to the US and again nine years later. There is almost no change in Patty’s suppliance of morphological inflection between
these time points, which suggests that her IL had fossilized. The data show clear contrasts between the morphological marking and the syntactic consequences of finiteness and agreement (Table 1.6). Across the recordings spanning nine years, Patty’s past tense inflection in English remains low (roughly 34% across recordings), whereas she consistently displays 100% accuracy on case marking that is licensed by syntactic finiteness features (Lardiere, 1998a). Similarly, while Patty’s third person singular marking with thematic verbs and the auxiliaries do and have is consistently low (8-17%), her verb placement, which is associated with a syntactic finiteness feature, is virtually perfect (99%) (Lardiere, 1998b). In later grammaticality judgements, Patty robustly disallows verb raising in English and shows target-like command of grammatical adverb placement (Lardiere, 2005).

<table>
<thead>
<tr>
<th></th>
<th>3rd person agreement (lexical verbs)</th>
<th>Past Tense marking (lexical verbs)</th>
<th>Nominative Case</th>
<th>Overt subjects</th>
<th>Verb raising</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lardiere</td>
<td>Production</td>
<td>4.5%</td>
<td>34.5%</td>
<td>100%</td>
<td>98%</td>
</tr>
<tr>
<td>(1998a; 1998b; 2005)</td>
<td>Grammaticality Judgements</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White (2003a)</td>
<td>Production</td>
<td>79.8%</td>
<td>80.5%</td>
<td>100%</td>
<td>98.5%</td>
</tr>
<tr>
<td></td>
<td>Grammaticality Judgements</td>
<td>100%</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1.6. Accuracy (percent) in production and judgements on morphological and syntactic properties. Percentages are averages across all recordings; adapted from Lardiere (1998a; 1998b; 2005) and White (2003a).

White (2003a) reports a case study on a fossilized Turkish learner of English, SD, who had been acquiring English for 10 years, starting at age 40. Since SD showed no improvement in marking inflectional morphology over an 18-month period, she was considered to have attained a fossilized endstate. In spontaneous production, SD’s suppliance of past tense inflection and person marking is around 80% for lexical verbs and near-perfect for auxiliaries and copula. Errors consist mostly of omission (Table 1.6); when inflection is supplied, it is accurate. As for the corresponding syntactic properties, SD displays no errors in case marking; moreover, there are no errors in verb raising in the context of questions and negatives in production or in grammaticality judgements. However, some residual verb raising errors surface in the context of adverbs.
Additional evidence of a difference between the morphological marking of finiteness and agreement, on the one hand, and the syntactic realization of finiteness, on the other hand, is attested in child L2 acquisition (e.g. Haznedar & Schwartz, 1997; Ionin & Wexler, 2002; Tran, 2005) and intermediate adult L2 acquisition (e.g. Herschensohn, 2001; Prévost, 2004; Prévost & White, 2000a). These studies demonstrate that (a) non-finite verbs occur in finite contexts, yet finite verbs do not occur in non-finite contexts, (b) tense and agreement morphology - when present - is correct, and (c) the syntactic reflexes of tense and agreement are consistently effected.

Lardiere (2000) refers to this dissociation between target-like representations of the syntactic features of finiteness and agreement, on the one hand, and deviant morphology, on the other, as a problem of mapping syntactic features to morphological forms, which Prévost & White (2000b) dub the ‘Missing Surface Inflection Hypothesis’ (following Haznedar & Schwartz, 1997). However, these authors assume different underlying causes of non-target-like inflectional morphology: (a) Representational differences in morphology (Lardiere, 2005; 2007), (b) Computational problems in mapping features to form at the syntax-morphology interface (Prévost & White, 2000b).

Lardiere (2000; 2005) stipulates that persistent problems with inflection reflect non-target-like morphological competence, i.e. representational underspecification or non-convergence in the featural composition of morphological markers.

Morphological competence includes, most obviously, the knowledge of which forms ‘go with’ which features. But consider what additional kinds of knowledge are required: What are the conditioning factors and are these phonological, morphosyntactic, semantic or discourse-linked? Are certain forms optional or obligatory, and what constitutes an obligatory context? In which domains are various features expressed, in combination with what other features, and why is supposedly the same feature expressed in some domains in some languages but not others? (Lardiere, 2005: 179)

According to Lardiere, such problems with feature assembly in morphology should be attested predominantly in more complex form-to-function mappings and show distinct L1 effects.

Prévost & White (2000b) argue that non-target-like default morphological forms may surface under processing pressure. Adopting the framework of Distributed Morphology (Halle & Marantz, 1993), they argue that the featural array of lexical items can be underspecified in some respects, e.g. for finiteness. This contrasts with fully specified formal features in syntax. Underspecified lexical items can nevertheless be

---

6 In a third approach, White (2003a), and Goad & White (2004; 2006) argue that L1 transfer at the morphology-phonology interface acts as a filter in preventing the target-like production of unstressed bound morphemes and determiners (the ‘Prosodic Transfer Hypothesis’). Since this approach shifts the explanatory emphasis to the Morphology-PFinterface, I will not consider it further in the present context. For a comparative discussion of the the three approaches, see White (2007).
matched with fully specified syntactic nodes, namely, when there is no competing lexical item bearing a more specified featural array. In the course of acquisition, the featural arrays become specified as a consequence of learning inflectional paradigms, and default forms are used less frequently. Due to increased cognitive demands in adult L2 acquisition, however, defaults are not completely expunged:

> Even when more fully specified forms are acquired, they do not necessarily ‘win’ in the competition for lexical insertion, so that underspecified forms continue to surface. [...] We speculate that this might be due to processing reasons or to communication pressure, in which case one would expect the problem to affect different kinds of language use differentially. (Prévost & White, 2000b: 129)

In consequence, non-target-like use of inflectional morphology should be contingent on the computational demands of the situation or task.

At this point, both models are rather vague with respect to the properties and processes of mapping problems at the syntax-morphology interface. For both mapping approaches, it is unclear (a) whether the dissociation of morphological forms and syntactic consequences also characterizes endstate L2 learners who did not fossilize well short of near-nativeness and (b) whether mapping problems are susceptible to L1 effects or whether they affect all L2 learners. The results on gender agreement by Franceschina (2005) seem to suggest that success on inflectional morphology is modulated by L1; however, the way in which L1 representations or computations can affect feature assembly or mapping processes in the L2 is yet to be explained. White (2003a; 2007) hypothesizes that the presence of inflectional morphology in the L1 sensitizes L2 learners to overtly expressed inflection in the L2 (see also Montrul, 2000). It is not clear whether such sensitivity, be it explicit or implicit, makes L2 mastery of inflectional morphology easier across the board or whether L1 effects are limited to particular aspects of L2 morphology, e.g. overt inflection that is similarly realized in L1 and TL. At a microlevel, Lardiere (2005; 2007) argues that acquisition of L2 inflectional marking involves the recalibration of feature matrices from the L1 to the TL settings. If L2 acquisition of inflection operates at the level of feature (re)assembly, L1 effects should be very specific to inflectional morphemes.

Further open questions concern variation across constructions. If difficulties in feature assembly relate to the transparency of the conditions of use of inflectional morphology (Lardiere, 2005), one would expect highly regular uses of inflection to be easier than less regular inflectional marking. Compared to the marking of past tense in English, which is contextually and conceptually conditioned, use of third person singular –s in English is highly regular and transparent; yet, Patty’s performance on third person agreement marking is even worse than on past tense marking (Table 1.6). Similarly, lower-level L2 speakers across L1s perform better on noun plurals than on verbal
agreement marking (Ionin & Wexler, 2002; White, 2007). This approach then needs to
give substance to the notions of feature assembly and mapping in order to clarify which
features are hard to acquire or reconfigure and how the mapping process between syntax
and inflectional morphology works.

As part of the 'Missing Surface Inflection Hypothesis', Prévost & White (2000b)
cite processing pressures as the main reason for deviance in inflectional morphology, yet
they report no psycholinguistic evidence supporting their contention, so that it remains
unclear under what conditions the use of defaults occurs and whether defaults surface for
all types of inflectional morphology (see also Jiang, 2004; 2007). In these respects, the
explanation that non-target-like inflection is due to computational problems in lexical
access tends to be somewhat post hoc (White 2007). As a consequence, the 'Missing
Surface Inflection Hypothesis' cannot make predictions as to the scope of inflectional
difficulty.

In sum, ‘mapping’ approaches assume that deviance at the syntax-morphology interface
is located in morphological competence or the computation of form-function mappings.
Persistent problems in determining (Lardiere, 2000; 2005) or accessing (Prévost & White,
2000b) the featural arrays of inflection are argued to prevail at ultimate attainment. These
accounts can capture dissociations of syntactic and morphological performance in adult
L2 acquisition. Unlike the representational impairment approach sketched in the
previous section, these models do not rely on absolute performance comparisons between
natives and (different) non-native groups in their claims; rather, they appeal to
systematicity of the problems with inflection. This way, they avoid the conceptual
conflation of non-convergence in competence and performance. However, these
approaches have yet to specify the properties and processes involved in mapping
morphology onto syntax, the role of the L1, and they have yet to address the question as
to whether mapping problems are particular to the syntax-morphology interface or extend
to other interfaces.

1.3.6.3. L1 interference? – The ‘Full Transfer/Full Access’ Model and the syntax-
semantics interface

The syntax-semantics interface in adult L2 acquisition has been studied in two ways.
First, studies on scope phenomena investigate structures whose semantic interpretation at
LF is directly read off their syntax (Dekydtspotter & Sprouse, 2001; e.g. Dekydtspotter et
al., 1997; Dekydtspotter et al., 1999/2000; for review, see Slabakova, 2006). A second set
of studies tests semantic distinctions that follow from the interaction of lexically encoded
semantic features and their morphosyntactic reflexes, e.g. aspect. L2 ultimate attainment
studies have focussed on the second type of syntax-semantics interactions and suggest
that the L1 modulates the degree of convergence on the TL. Unlike the ‘Failed Functional
Features’ Hypothesis which assumes categorical L1 effects due to representational deficits incurred by a critical period, studies on the syntax-semantics interface indicate that L2 ultimate attainment is modulated, yet not necessarily determined, by L1-TL differences.

In the early study by Coppieters (1987), non-native and native speakers of French diverge most pronoucnedly on aspectual properties of the *imparfait* versus the *passé composé* which induce interpretive changes in verb meaning (6) or the habituallity of events (7).

(6) Est-ce que tu [as su/savais] conduire dans la neige?  
‘Did you manage/know how to drive in the snow?’

(7) [J’ai très souvent mangé/Je mangais très souvent] de la racine d’arnica après cette histoire.  
‘I often ate [eventive/habitual] arnica root after that event.’

Although Coppieters (1987) does not consider L1 effects systematically, he notes that the degree of divergence from the native controls on aspectual judgements differs according to L1 background: Compared to the French natives, L1 Romance speakers show a relative judgement difference of 19%, L1 Germanic speakers 38%, L1 Farsi speakers 50% and speakers of Asian languages 66.7%. Testing 20 L1 French and 20 L1 Chinese speakers with a minimum residence of three years in an English-speaking country, Cranshaw (1997) uses a variety of oral, written and judgement tasks to investigate the acquisition of English tense and aspect. At the group level, there are significant differences between the native control group and each non-native group. Moreover, the L1 French group performs closer to the English natives than does the L1 Chinese group. L1 differences are also reflected at the level of individual results: Three L1 French speakers but only one L1 Chinese speaker score within the native range across all tasks. These findings indicate that L1 effects constrain ultimate attainment on aspect.

A study by Montrul & Slabakova (2003), however, suggests that aspectual distinction can be acquired to native-speaker levels even by L2ers whose L1 does not match the TL in aspectual distinctions. Sixty-four Anglophone adult L2 speakers of Spanish were recruited in the US for participation; there was no residency criterion, yet all participants had spent between six months and ten years in Spanish-speaking countries. According to a proficiency test and an oral interview, the participants were allocated to three groups: advanced (n=24), superior (n=23) and near-native (n=17). In two judgement tasks on aspectual distinctions in Spanish associated with different tenses, imperfect and preterite, more than 70% of the near-natives perform within the native-speaker range across the two tasks. Unfortunately, the near-native subjects were predominatly university instructors of Spanish, so that “the success of their performance possibly stems from having explicitly learned and taught the distribution of these tenses”
(Montrul & Slabakova, 2003: 387). For lack of a non-native group of a different L1 background, the Montrul & Slabakova (2003) study cannot elucidate whether convergence on aspectual distinctions is modulated by L1 effects.

In a systematic cross-linguistic comparison, Sorace (1993) reports clear effects of L1 influence at ultimate attainment by near-native L1 French and L1 English speakers on morphosyntactic and associated semantic properties of unaccusative verbs in Italian. Intransitive verbs can be semantically divided into unaccusatives and unergatives. In syntax, unaccusatives (e.g. 'arrive') project only one object argument (Burzio, 1986), whereas unergatives (e.g. ‘cry’) have one subject argument. Further, unaccusatives require the auxiliary essere (‘be’) in perfective tenses, unlike unergatives that select the auxiliary avere (‘have’). If unaccusatives occur in the context of raising or control verbs (e.g. the modals potere, ‘to be able to’; dovere, ‘to have to’) that normally take avere, the auxiliary essere may optionally be used (8).

(8) Mario è/ha dovuto andar a casa.
Mario is/has had to go home
‘Mario had to go home.’

Auxiliary selection interacts with clitic choice in restructuring constructions. Clitics like ci (‘there’) can either attach to the embedded verb or climb to the matrix verb. In the former case, auxiliary selection remains optional (9a), whereas clitic climbing obligatorily triggers auxiliary change to essere of the matrix verb (9b).

(9) a. (A casa), Mario è/ha dovuto andarci.
   (Home), Mario is/has had to go there.
   ‘Mario had to go home.’

By contrast, optionality in auxiliary selection exists in French, yet, it is more restricted since only a subset of French unaccusatives take the auxiliary être (‘be’). For raising and control verbs, there is no optionality in auxiliary selection and avoir is the only licit auxiliary. Moreover, French does not display any restructuring phenomena and thus lacks clitic climbing. Finally, English unaccusatives require have and there are no clitics.

Twenty-four English and twenty French near-native speakers of Italian, whose native-likeness was judged impressionistically by Sorace, were tested in a Magnitude Estimation acceptability judgement task. Magnitude Estimation (Bard, Robertson & Sorace, 1996, see also Chapter 6, Experiment 1) requires subjects to make proportional judgements on a self-defined open-ended scale. The acceptance rates are reported as rescaled geometric means ranging from 0 to 10 in Table 1.7.
The Critical Period Hypothesis and Second Language Acquisition

<table>
<thead>
<tr>
<th></th>
<th>L1 French</th>
<th>L1 English</th>
<th>Italian natives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic sentences</td>
<td>essere</td>
<td>3.824</td>
<td>7.231</td>
</tr>
<tr>
<td>(choice optional)</td>
<td>avere</td>
<td>9.420</td>
<td>6.977</td>
</tr>
<tr>
<td>Low clitic (choice</td>
<td>essere</td>
<td>4.065</td>
<td>6.784</td>
</tr>
<tr>
<td>optional)</td>
<td>avere</td>
<td>7.841</td>
<td>6.211</td>
</tr>
<tr>
<td>Clitic climbing</td>
<td>essere</td>
<td>8.525</td>
<td>6.286</td>
</tr>
<tr>
<td>(essere only)</td>
<td>avere</td>
<td>4.285</td>
<td>6.623</td>
</tr>
</tbody>
</table>

Table 1.7. Geometric means of acceptability judgements on auxiliary selection in restructuring constructions (adapted from Sorace, 1993: 39).

The Italian natives show optional acceptance of essere and avere in basic sentences and sentences with unraised clitics; however, for clitic climbing sentences, essere is categorically preferred over avere as the matrix auxiliary. The two non-native groups differ in their judgement patterns: The French group shows a target-like categorical preference for essere in the clitic climbing construction, yet strongly prefers avere in the two optional conditions. By contrast, the English group shows indiscriminate judgements in all conditions. Sorace (1993) interprets the French group as having ‘divergent’ representations and the English group as having ‘incomplete’ representations of the target language (for a different interpretation, see Sorace, 2005). As Papp (2000) and White (2003b) note, the results of the English group can be interpreted in two ways: (a) the English group has determinate representations that allow for optional auxiliary selection throughout, yet they generally make judgements at lower levels of confidence, or (b) they have indeterminate representations that lead to indiscriminate judgements. Without detailed investigation of individual results and group judgement patterns, it is difficult to arbitrate between these interpretations.

Sorace’s (1993) findings illustrate that there are L1 differences at ultimate attainment. However, the endstates of the L1 French and L1 English groups do not just mirror the respective L1 properties, since both groups would then be expected to display a categorical preference for avere across sentence types. Instead, Sorace (1993) argues that a combination of L1 properties and the TL input underlies the asymmetric judgement patterns of the L1 French and L1 English groups. Due to recourse to L1 properties, the L1 French near-natives accommodate the avere/essere alternation within the avoir/être distinction in French and impose a near-categorical preference for avere onto restructuring verbs. In contrast, alternations between avere and essere for unaccusatives in the TL input cannot be accommodated within the more restrictive English morphosyntax of unaccusatives. According to Sorace (1993), the cross-linguistic similarity in the morphosyntactic realization of unaccusativity between Italian and French predisposes French learners of Italian to access distinctions in auxiliary selection for restructuring verbs in the case of clitic climbing, whereas the lack of syntactically
expressed unaccusativity distinctions in English renders auxiliary choice opaque for the L1 English group. However, Sorace (1993) does not elucidate how the TL input and the relevant parts of the L1 grammar conspire in bringing about convergence on obligatory auxiliary selection for clitic climbing.

A model that explicitly considers the interaction between L1 transfer and TL input as potentially curtailing convergence in adult L2 acquisition is the ‘Full Transfer/Full Access’ model by Schwartz & Sprouse (1996). According to this model, the full (abstract) L1 grammar constitutes the initial state of L2 acquisition (Full Transfer). In L2 development, learners restructure their grammars to accommodate the TL input making use of the same cognitive processes as children acquiring the L1 (Full Access). Since the ‘Full Transfer/Full Access’ model holds that there are no critical period constraints enforcing asymmetries between non-native and native grammars, L2 ultimate attainment can in principle come to resemble ultimate attainment in L1 acquisition. At the same time, Schwartz & Sprouse (1996) argue that non-convergence on the TL may occur because the evidence in the TL-input to restructure the L1 grammatical configurations to the TL settings is absent, ineffective (e.g. negative evidence, Schwartz, 1993), obscure, very complex or rare or ambiguous (Schwartz & Sprouse, 1996: 42). In terms of learnability, then, certain features of the TL might not be detectable or informative to enforce restructuring to the TL grammar (see also Schwartz, 1998). On this account, either gaps and ambiguities in the TL input can lead to the retention of particular properties of the L1 grammar because there is no unequivocal evidence that forces abandoning the L1-based accommodation of the TL, or the TL input may occasion restructuring to a non-TL grammar that is also compatible with the TL input. The subtle and empirically underdetermined differences in lexical semantics between imparfait versus the passé composé in the Coppieters (1987) study may be a case in which no evidence is available for restructuring; the finding that L1 French learners retain their L1-based preference for avoir with Italian unaccusatives (Sorace, 1993) illustrates that the TL does not force IL grammars to restructure if the TL data can be accommodated within the L1 grammar.

Assuming that the L1 grammatical representations initially act as a kind of perceptual ‘filter’ in the analysis of the TL input (e.g. Brown, 2000; Goad & White, 2006; Schwartz & Sprouse, 1996), grammatical restructuring to the TL likely proceeds differently for learners of different L1s. Depending on L1 properties, the salience of grammatical properties in the TL input (Goad & White, 2006), and, in turn, the sensitivity to the TL input will vary (e.g. Carroll, 2001; Lardiere, 2007). Although these L1 effects are far from well understood, they appear to implicate different outcomes in L2 acquisition.

In the area of aspectual distinctions, grammatical restructuring to the TL requires comparatively less revision to the L1 grammatical representations for speakers of an L1 that has underlyingly similar, if differently aligned, semantico-syntactic properties
compared to speakers of an L1 that expresses these properties in more complex and different ways. In this vein, L1 Romance learners attain greater mastery of the French (Coppieters, 1987) or English (Cranshaw, 1997) tense-aspect system than, for instance, speakers of Asian languages which encode aspect differently.

Further L1 effects on ultimate attainment have been attested for (a) the domain of Dutch expletive subjects, where L1 German and L1 French speakers outperform L1 Turkish speakers (van Boxtel, 2005), and (b) general proficiency measures, on which native-like performance is reported for a subset of L1 Spanish (Birdsong & Molis, 2001; McDonald, 2000) and L1 Dutch (Kellerman, 1995) speakers, yet not for for L1 Hungarian (DeKeyser, 2000) or L1 Korean/Chinese speakers (Johnson & Newport, 1989) in the acquisition of English.

These studies adduce evidence that the native language modulates the potential for attaining native-like grammatical representations in the L2. At the same time, the effects of the first language on L2 ultimate attainment have been described at the rather coarse-grained level of typological similarity, and very few attempts have been undertaken to give substance to the workings of cross-linguistic influence at the L2 endstate. Major unanswered questions include where L1 effects apply, i.e. at the level of grammatical representation, processing or both, how restructuring proceeds, what the threshold level of properties in the TL input is to engender restructuring, and how L1 and TL representation can co-exist in IL grammars.

1.3.6.4. Impairment of interpretable features or processing overload? – The ‘Interpretable Features’ Hypothesis and the syntax-discourse interface

The discourse-relatedness of syntax is encoded in terms of topic or focus which designate the information status of constituents. At the syntax-discourse interface, these information-structural conditions impose constraints on syntax by restricting syntactic options to particular discourse interpretations.

The relation between the order of subjects and focus in pro-drop languages, such as Italian and Spanish, exemplifies the interplay between syntax and discourse. In pro-drop languages, subjects can optionally surface in preverbal and in postverbal position (Burzio, 1986). SV and VS orders are available for all types of verbs; however, the base order varies according to verb type. Unaccusatives project a VS base order, while unergative verbs project an underlying SV order (see Section 1.3.6.3). At the syntax-discourse interface, overt subjects in preverbal position are canonically interpreted as given information, denoting topics, whereas postverbal subjects receive focus, expressing new information (e.g. Belletti, 2001). The VS order in (11a) constitutes a felicitous answer to the question in (10), whereas the SV order violates the discourse-to-syntax mappings and sounds marked (indicated by #).
Research on the L2 acquisition of discourse-related syntax has shown differences in the acquisition of the syntax, on the one hand, and the discourse conditions of syntax, on the other. Hertel (2003) reports that lower-level intermediate English learners of Spanish show awareness of the syntactically based VS preference for unaccusative verbs as opposed to unergatives, yet that they perform worse on identifying the connection between focus and VS orders in written production tasks where focus structure was manipulated by wh-questions. Lozano (2006) reports that advanced L1 Greek and English learners of Spanish equally appropriately differentiate VS orders according to verb class but not according to focus in a contextualized judgement task (see also Belletti & Leonini, 2004 for French and German learners of Italian).

At the near-native level, Belletti, Bennati & Sorace (2007) report a series of experiments with 17 English near-native speakers of Italian, who were screened for native-likeness on the basis of the tests in White & Genesee (1996). Two elicitation experiments and a spontaneous production task were employed to test for the contingency between word order and focus. In spontaneous production, there are no differences between near-natives and native controls in the production of null subjects (52% vs 59%) or postverbal subjects (16% vs 15%), which suggests that near-natives have come to acquire the syntactic expressions of subjecthood in pro-drop languages. In an elicitation task that prompted answers including verbs of different classes (unergative, unaccusatives, transitives, existential-'there’ constructions) by designating the subject as the focus exponent (e.g. Who called?), the near-natives produce significantly fewer postverbal subjects compared to the native controls across almost all verb types (Table 1.8).
In a second task, all-focus sentences were elicited by asking subjects to complete the sentence fragment *Have you heard that ...* by using eventive verbs and NPs presented in isolation. The givenness of subject NPs was manipulated by definiteness, i.e. definite NPs denoted given referents and indefinite NPs denoted new referents. Although the near-natives are sensitive to definiteness differences, Table 1.8 again shows that their overall production of postverbal subjects remains significantly below native-speaker levels.

Belletti et al. (2007) attribute the attested dissociation between target-like syntactic knowledge of postverbal subjects and failure to identify the discourse conditions of their use to residual effects of L1 discourse strategies. They suggest that the canonical English SV order predominates in the L2 forcing SV orders in all constructions other than the existential *there*-construction which is also VS in English. Other studies on near-native L2 speakers also point to a split between target-like acquisition of syntax and failure to acquire discourse-related aspects of syntactic options in the TL. In a truth-value judgement task, Sorace & Filiaci (2006) find that near-native L1 English speakers of Italian allow overt pronouns to corefer to intrasentential topics (subjects) while natives prefer extrasentential non-topical referents. Across several off-line comprehension and production tasks, Valenzuela (2006) reports that L1 English near-native speakers of Spanish acquire the syntax of clitic left dislocation constructions; yet, they fail to associate object topics in clitic left dislocation constructions with an interpretive specificity requirement.

In the 'Interpretable Features' Hypothesis, Sorace (2000; 2003) formalizes non-convergence in the mapping of discourse and syntax by hypothesizing that interpretable features such as [+focus] etc. are subject to maturationally enforced impairment in adult L2 acquisition. By contrast, uninterpretable features that affect morphosyntactic operations are spared from maturational effects, such that ‘narrowly’ syntactic features of the TL that underlie parametric differences in word order are fully acquirable (contra,
e.g., Hawkins & Chan, 1997). As a consequence of a critical period affecting interpretable features, the L1 discourse-to-syntax mapping is never fully expunged from the L2 grammar and continues to coexist with the TL configuration. Residual optionality of interpretable features in endstate L2 grammars is expressed in indeterminacy between L1 and TL options at the syntax-discourse interface.

However, as Sorace (2005), inter alii, notes, the findings might follow from a higher-level generalization. Identifying the discourse-relatedness of syntax is inherently more complex than computing syntax on its own. The disjunction between convergence on syntax and continued divergence on interface aspects of syntax might thus reflect the additional computational complexity required by interface mappings rather than a deficit specific to the syntax-discourse interface (e.g. Sorace, 2005; Sorace & Filiaci, 2006). Consider postverbal subjects in this respect. Target-like performance requires learners to derive the syntactic VS order and then constrain the applicability of VS orders to the appropriate information-structural conditions. The composite nature of this task might tax or exceed the processing resources of L2 speakers, such that L2 speakers manage to compute the syntax, yet not the discourse-relatedness of postverbal subjects. Parallels are attested in research on monolingual L1 acquisition. Children show significant developmental delays in acquiring pronoun interpretation and discourse constraints on optional syntactic movement. Such delays have been attributed to the limited processing resources in younger children that curtail efficient computation of syntactic and discourse information (e.g. Avrutin, 1999; Krämer, 2000; Reinhart, 2006).

It is worth noting, though, that the generalization that interface performance correlates with the computational demands of particular constructions does not capture the data from Belletti et al. (2007). Since unaccusatives project a base VS order, the identification of discourse constraints for unaccusatives should overall be less costly than for unergative verbs, for which the VS order is syntactically derived. If difficulty with the discourse-relatedness of VS orders were a product of the combined processing effort required by computing the syntax and the discourse relations, one would, ceteris paribus, expect a difference in performance according to verb type: The data in Belletti et al. (2007), however, do not show differences in suppliancy of VS orders depending on verb type (Table 1.8). It seems, then, that computational complexity needs to be defined more specifically for individual interface mappings.

In sum, a growing body of research converges in reporting disjunctions between the successful acquisition of syntax and persistent difficulties in relating syntax to discourse conditions. On Sorace’s (2003) ‘Interpretable Features’ Hypothesis, this dissociation reflects local grammatical impairment at the level of interpretable features that map discourse information onto syntax. Alternatively, non-convergence observed at the syntax-discourse interface might reflect the increased computational complexity inherent in mappings between grammatical modules. The latter approach also resonates with accounts given in computational terms for difficulties at the syntax-morphology
interface (Prévost & White, 2000b). Yet, as with the ‘Missing Surface Inflection’ Hypothesis, on-line data from L2 processing are necessary to substantiate the claim that computational limitations underlie non-convergence at interfaces.

1.4. Summary

This chapter reviewed proposals that a neurocognitive critical period constrains L2 acquisition which offer a principled explanation of differences in outcome between child L1 and early L2 acquisition, on the one hand, and adult L2 acquisition, on the other hand. However, descriptive studies on the geometry of the age function and on the attainability of native-like morphosyntax in late L2ers have proven empirically inconclusive to date. Moreover, they remain theoretically somewhat unsatisfactory, since neither can explain linguistic patterns of non-convergence at ultimate attainment. Research against the backdrop of linguistic theory can more systematically identify the locations and causes of non-convergence. In the preceding review of studies and approaches to residual non-convergence, I identified some common themes and common lacunae in the generative study of L2 ultimate attainment.

First, a greater degree of convergence at L2 ultimate attainment is attested for syntactic properties (van Boxtel, 2005; White & Genesee, 1996), whereas associated interface properties are subject to relatively more difficulty even after prolonged exposure to the TL. Asymmetric convergence on syntax and the interface has been shown for morphology (Franceschina, 2005; Lardiere, 2007; White, 2003a), for discourse (Belletti et al., 2007; Sorace & Filiaci, 2006) and for interpretation (e.g. Coppieters, 1987; Valenzuela, 2006). These interfaces have thus far been studied in isolation. In consequence, causes of non-convergence like representational differences, L1 effects, computational problems have been separately proposed for each interface, which leaves open the question as to whether these causes apply generally across interfaces or whether different causes interact across interfaces. Addressing this question, this thesis reports experiments on the interfaces of syntax with morphology, semantics, discourse and the lexicon.

Second, non-convergence on interface properties has been shown to be related to L1 influence in cross-linguistic studies at the syntax-morphology interface (e.g. Franceschina, 2005), the syntax-lexicon interface (Sorace, 1993) and the syntax-discourse interface (Belletti et al., 2007). Not least because the question whether native-like behaviour is attainable at all has taken centre stage in research on L2 ultimate attainment so far, the role of L1 influence is only beginning to be explored (Franceschina, 2005; van Boxtel, 2005). It is an open question as to how the L1 affects endstate L2 grammars, in particular, whether identity between the L1 and the TL in abstract grammatical properties, e.g. gender, determines convergence (e.g. Franceschina, 2005), or whether more indirect L1-TL correspondences in typological similarity of grammatical properties
modulate ultimate attainment (Sorace, 1993). This thesis tests advanced and near-native speakers of L1s that differ systematically in typological relatedness and identity in abstract grammatical properties.

Finally, non-convergence has been located at different levels of grammar or computation. In particular for the syntax-morphology and syntax-discourse interfaces, representational impairment in grammar (e.g. Franceschina, 2005; Sorace, 2003) or computational inefficiency in the coordination of grammatical knowledge (e.g. Prévost & White, 2000b; Sorace, 2005) have been argued to underlie non-convergence at L2 ultimate attainment. On the basis of the off-line studies on L2 ultimate attainment available so far, the data can at best indirectly arbitrate between the two options. Research on the L2 processing of interface phenomena is needed. To this end, this thesis compares and contrasts off- and on-line performance at the interfaces. As the logical next step, the following chapter considers the link between representation and processing at L2 ultimate attainment by reviewing previous research on L2 processing.