

The relevance of first language attrition to theories of bilingual development

Monika S. Schmid, University of Essex/Barbara Köpke, Université de Toulouse

Acknowledgement

We are very grateful to the editors of the annual epistemological issues of *Linguistic Approaches to Bilingualism*, Holger Hopp and Tanja Kupisch for both the original invitation to contribute this piece and their thorough, thoughtful and considerate engagement with the text as it developed. Their support has been invaluable, and we could not have done this without them. We are also grateful to Laura Domínguez, Jason Rothman, Rasmus Steinkrauss and three anonymous reviewers for their comments on earlier versions of this article. Any remaining errors and inaccuracies are entirely our own responsibility and due to us not having heeded the excellent advice of these colleagues.

Abstract

Research on second language acquisition and bilingual development strongly suggests that when a previously monolingual speaker becomes multilingual, the different languages do not exist in isolation: they are closely linked, dependent on each other, and there is constant interaction between these different knowledge systems. Theoretical frameworks of bilingual development acknowledge this insofar as they usually draw heavily on evidence of how the native language influences subsequent languages, and how and to what degree this influence can eventually be overcome. The fact that such

crosslinguistic transfer is not a one-way street, and that the native language is similarly influenced by later learned languages, on the other hand, is often disregarded.

We review the evidence on how later learned languages can re-shape the L1 in the immediate and the longer term and demonstrate how such phenomena may be used to inform, challenge and validate theoretical approaches of bilingual development.

Introduction

Bilinguals differ from monolinguals in various ways, not only with respect to the second or weaker language but also to the native or dominant one. Such differences can be observed at all linguistic levels and all stages of bilingual development as a result of interactions between the two linguistic systems.

We refer to any of the phenomena that arise in the native language of a sequential bilingual as the consequence of the co-activation of languages, crosslinguistic transfer or disuse, at any stage of second language (L2) development and use, as *language attrition*¹. First language (L1) attrition is therefore considered to be the process by which a) pre-existing linguistic knowledge becomes less accessible or is modified to some extent as a result of the acquisition of a new language, and b) L1 production, processing or comprehension are affected by the presence of this other language.

We advocate an integrated approach to bilingualism where observations on how bilinguals differ from monolinguals with respect to the processing, use and representation of *each* of their languages inform theories of development. We will first

¹ It should be noted that the term ‘attrition’ has often been criticized as being unfortunate for two reasons: firstly, its connotations are mainly negative and secondly, it implies some kind of permanent erosion. However, the term has been used consistently to refer to the particular and unique process of change described here for close to four decades and thus provides coherence to a field which, in comparison to L2 development is still quite recent. We thus feel that the positives of sticking with ‘attrition’ outweigh the – undoubted – drawbacks. Due to space constraints we limit our discussion here to the attrition of the native language.

discuss the scope of attrition effects and the terminological background. This is followed by a review of some of the previous findings on how processing of the L1 may change as a result of the co-activation of languages. We then turn to a consideration of how different theoretical approaches (specifically, Usage-Based and generative approaches) may integrate such findings into their understanding of bilingual development and use them to predict and model more permanent changes to underlying representations and structures of the L1. Finally, we consider the impact of three factors that have often been proposed to play a role for these processes: crosslinguistic similarity; exposure and use; and the age of onset of bilingualism.

1. Background

Current approaches to bilingual development widely recognise that all languages which co-exist in the same brain are in constant interaction (e.g., Van Hell & Dijkstra, 2002; Shook & Marian, 2013) and that this interaction is at the root of many of the differences which we can observe between bilingual and monolingual speakers. Language processing is more cognitively demanding for bilinguals than for monolinguals since the bilingual has to contend with a number of additional tasks and challenges, among them resisting intrusions or automatisms from any language that has not been selected for use (e.g., Green, 1986, 2011; Abutalebi, Della Rosa, Green, Hernandez, Scifo, Keim, Cappa & Costa, 2011). In addition, bilinguals have to establish and maintain processing routines based on less input – in either language – than monolinguals, which may lead to weaker representations and lower resting activation of linguistic features (Hopp, 2013) and lexical items (Gollan, Montoya, Fennema-Notestine & Morris, 2005).

Crosslinguistic competition and lower levels of activation affect not only the L2 but also the L1. That notwithstanding, theoretical models of bilingual representation and processing rarely acknowledge this as an integral part of becoming bilingual or consider what it may tell us about bilingual development. Instead, such models and theories are usually based on the considerable empirical evidence relating to how learners acquire, process, and use the L2, the linguistic conditions which make the L2 acquisition of particular features more or less difficult, and the external factors which impact on the likelihood of success. Only rarely is it investigated how these aspects of development may impact on the maintenance (in either language) of knowledge once it has been acquired.

As a consequence, and despite the steadily increasing number of studies focussing on L1 attrition (see Schmid, 2016 for a recent bibliography), few theoretical approaches consider whether the process of adapting L1-based structures or processing routines on the basis of evidence in the L2 input may not only feed into L2 grammar-building but also have consequences for those structures and routines in the L1. Many models across the theoretical spectrum are thus built on the tacit assumption that two extremely similar subsystems of information (the L1 and the L2) exist in a state of co-activation in the mind of the bilingual, but that one of them is stable while the other is developing: the L2 grows and changes according to specific mechanisms of human cognition and learning which govern the interaction of different types of knowledge and input, while the L1 remains unaffected.

The fact that control groups in investigations of L2 development are usually (functionally) monolingual compounds this problem: at least some of the observed differences between experimental and control populations which are typically

interpreted as differences between L1 and L2 speakers and interpreted in the context of limitations to L2 development, may, in fact, be differences between bilinguals and monolinguals and thus affect the L1 in equal measure (Hopp & Schmid, 2013). This would indicate that what constrains ultimate success for such phenomena is not so much related to the order of acquisition but to the added cognitive pressure of being bilingual.

We argue that in order to fully understand the nature of bilingual development and to resolve important and fundamental questions about the human capacity for language learning, processing and use, we need to arrive at a better understanding of how the mechanisms that drive and constrain L2 acquisition may also affect already established linguistic knowledge, both in the immediate and in the longer term. Development is not a unidirectional process: what has been learned can also be forgotten, and the principles that govern the process of dismantling and change can tell us as much about the nature of a knowledge system as the ones that govern its acquisition. Comparing the outcome of bilingual development in both learner and attriter systems can thus shed more light on the question of how underlying representations and online crosslinguistic transfer interact, and how language acquisition may be different the second time around (Schmid, 2009, 2014).

2. The scope of attrition effects

L2-to-L1 transfer effects have recently been hailed as one of the main discoveries in bilingualism research (Kroll, Dussias, Bice & Perroti, 2015: 378). Despite such acknowledgements that they form an important and integral part of bilingual development, the view seems to persist that actual language attrition is a rare phenomenon which must somehow go beyond online transfer effects. Attrition has

variously been described as a “special case of variation in the acquisition and use of a language” (Andersen, 1982: 86), occurring in “extreme situations” (Costa & Sebastián-Gallés, 2014: 399) which minimally involve a period of several decades of very limited L1 exposure (e.g., Dussias & Sagarra, 2007) as well as a high level of L2 proficiency (e.g., Kroll & Bialystok, 2013; Tsimpli, Sorace, Heycock & Filiaci, 2005). The fact that many authors (and even more reviewers) raise the question of how to separate the “normal influence between languages in a bilingual or polyglot” from the process referred to as attrition (e.g., Ahlsén, 2013: 1) further suggests the view that there is something unusual about attrition. Attriters are thus often considered to form a limited subset of bilinguals (see Köpke & Schmid, 2004 for an overview).

The underlying perception thus appears to be that ‘true’ attrition effects should go beyond the online manifestations of crosslinguistic transfer which all bilinguals experience, that they should be permanent, irreversible and affect underlying structure: “it is erosion that reaches the level of competence that allows for interesting claims about and meaningful insight into the attrition process” (Seliger & Vago 1991: 7).

This view is based on a fundamental misperception, namely that, in the first instance, all bilinguals follow a process of acquisition which starts out with massive, but entirely one-way, L1-to L2 transfer. As proficiency increases, this transfer is gradually reduced until some kind of stable endstate is reached at which the L2 either fossilizes or converges towards a native-like level and becomes dominant. On this view if, and only if, the latter is the case, language dominance reversal takes place and L1 attrition sets in.

This view was prevalent in early research on language attrition, as for example illustrated by Seliger & Vago’s (1991) model which makes reference to two types of language learner systems first proposed by Ervin & Osgood (1954): Compound

Bilingualism (labelled Compound I Bilingualism by Seliger & Vago), in which the L1 is the source for hypotheses about the target language and both grammars are subserved by a common knowledge base from the L1 (if the learner fails to progress beyond this stage, fossilization will occur). In Coordinate Bilingualism, on the other hand, both languages exist largely independently of each other and traffic in either direction is minimal. Seliger & Vago's model then adds a third stage of development, namely Compound II Bilingualism. This stage resembles the first one in that the two grammars become dependent again, but now it is the more fluent L2 which encroaches on the L1 in a reversal of the direction of transfer (Seliger & Vago 1991:5f.).

We propose that to make such a distinction between online/transient and representational/permanent effects of the L2 on the L1, with only the latter being considered instances of attrition, is both artificial and unhelpful, as they merely represent developmental stages on the same continuum. Attrition effects begin as soon as L2 development sets in, in the first instance as online phenomena of co-activation where production or processing is to some extent affected and subserved by both languages (e.g., difficulties of accessibility, phonetic or grammatical mergers). They may or may not eventually lead to apparent changes to or restructuring of knowledge, processing or production as a result of long-term crosslinguistic interference. There is no meaningful way of establishing two discrete and distinct stages of this continuum, so every bilingual is an L1 attriter.

3. Online effects of linguistic co-activation in the L1

The competition incurred by the co-existence of two languages in the same mind results in a 'bilingual disadvantage' on certain linguistic tasks (e.g., Bialystok, 2009). Such

effects have been studied most frequently in the area of the lexicon, where lexical access or retrieval tasks reveal a robust difference between monolinguals and bilinguals (e.g., Gollan et al., 2005; Sandoval, Gollan, Ferreira & Salmon, 2010). By contrast, top-down tasks such as classification appear to be unaffected, suggesting that the competition effect constraining bilinguals' performance is specific to accessing lexical representations as opposed to meaning (Gollan et al., 2005). A bilingual disadvantage can also be observed in free speech, where both L1-dominant and L2-dominant bilinguals have an increase in disfluencies and decrease in lexical richness as compared to monolinguals (see Bergmann, Sprenger & Schmid, 2015 for an overview).

Bilingual disadvantage effects in lexical access can be observed after rather short periods of immersion: Baus, Costa and Carreiras (2013) found that a cohort of participants engaged in a one semester study-abroad program had slower naming latencies and lower production rates for non-cognates in their L1 at the end of the immersion period than at the beginning. After as little as three years, even unbalanced bilinguals can become faster in recognizing items in the L2 than in the L1 (Frenck-Mestre, 1993).

The effect underlying the bilingual disadvantage in lexical processing and lexical access has been linked to two main factors. The first relates to lower frequency of activation: Gollan et al. (2005) compare the bilingual lexicon to a monolingual one which is composed entirely of low-frequency items. The second factor is cross-linguistic competition and the spread of lexical activation across languages (Marian & Spivey, 2003; Blumenfeld & Marian, 2007). This effect is most pronounced for words which share phonological, orthographical or semantic features (Friesen, Jared & Haigh, 2014). Such similarities may not only influence online bilingual processing and

production in both languages through facilitation or inhibition, they also allow items to map onto each other and thus cause subtle changes in meaning. Effects of transfer and convergence have been described for example in the area of motion events (Bylund, 2009; Hohenstein, Eisenberg & Naigles, 2006; Pavlenko, 2010), specific semantic fields, for example that relating to emotions (Pavlenko, 2003), metaphors, idioms, figurative and non-literal language (e.g., the papers in Heredia & Cieslicks, 2015; Sprenger, Bergmann & Schmid, submitted), or the encoding of manner in speech and gesture (Brown & Gullberg, 2008).

Similar processes of transfer have often been described for the phonetic level: a range of investigations of late bilinguals have found adaptation of phonetic settings in both languages towards those of the other one. This has most frequently been demonstrated with respect to the lag between the release of a voiced or voiceless stop and the onset of the vibration of the vocal fold (voice onset time, henceforth VOT) which characteristically differs between languages (for a recent overview of investigations of VOT in attrition see Chang, 2012). The adaptation of VOT values towards the L2 setting increases with length of L2 experience and proficiency levels (Flege, 1987), is more pronounced in casual than in formal speaking styles (Major, 1992) and is reduced after periods of re-immersion in the L1 (Sancier & Fowler, 1997), suggesting a fluctuating and dynamic interaction of factors such as proficiency, context, and (recent) exposure. Bidirectional crosslinguistic adaptation has also been found with respect to the distribution of the vowel space (Bergmann, Nota, Sprenger & Schmid, 2016; Mayr, Price & Mennen, 2012), the realization of liquids (de Leeuw, Mennen & Scobbie, 2012), rhoticity (Himmel & Kabak, 2016; Ulbrich & Ordin, 2014) and suprasegmentals (Mennen, 2004). This suggests that L1 and L2 sounds are linked at the

system-wide level (Chang, 2012; Mayr, Price & Mennen, 2012), a notion further supported by the fact that the perception of phonological categories in the L1 may become weakened by competing, non-overlapping L2 categories (Tamminen, Peltola, Toivonen, Kujala & Näätänen, 2013).

Bi-directional convergence is most likely to constrain language processing and production in those cases where sufficiently similar features are shared which may then provide a compatible 'launch pad' and 'landing site' for transfer effects (Schmid, 2011). Such similarities are rarer at the morphosyntactic level than in the lexicon and in phonetics/phonology, unless the languages which are studied are closely related. There is, however, evidence that non-selective activation of languages in bilingual processing does occur at the morphosyntactic level: Firstly, it has been shown that the lexicon is influenced by bi-directional CLI not only in single-word processing (as, for example, in naming and in fluency tasks) but also when words are processed in context, even where the sentence provides strong cues as to the language being used (for review, see Kroll & Bialystok, 2013). More importantly, non-selective syntactic activation in bilinguals has also been found, for example in the context of syntactic cross-linguistic priming, leading some researchers to argue for shared syntactic representations of similar structures – that is, structures that share the same word order – in bilingual processing (Bernolet, Hartsuiker & Pickering, 2007; Sanoudaki & Thierry, 2015). In such cases, optionality in interpretation can be affected, as Dussias' studies of relative-clause attachment in Spanish-English bilinguals have shown (Dussias, 2004; Dussias & Sagarra, 2007). Cross-linguistic syntactic activation appears to be modulated by dominance effects (Sanoudaki & Thierry, 2015) and by frequency effects: preferential strategies, such as high vs. low attachment in relative clauses have been shown to be

amenable to input in an intervention study by Dussias et al. (2014, reported in Kroll et al., 2015).

Taken together, the findings reviewed above suggest multiple links and connections between all of a bilingual's languages, which make bilingual language processing less efficient (slower and less accurate) as a result of the added task demand of suppressing and inhibiting the unwanted language. Due to an overall limitation of cognitive resources this increased demand prevents the bilingual to attend to all aspects of language processing with the same depth as monolinguals. This assumption receives further support by studies on L1 reading in bilinguals, which is not only influenced by a cognate facilitation effect (van Assche, Duyck, Hartsuiker & Diependaele, 2009) but also becomes overall less efficient at the paragraph-level, as has been shown in a series of gaze-contingent moving-window paradigm studies (Whitford & Titone, 2015). To compensate, bilinguals may develop more efficient strategies of attention control in both L2 and L1 (Duncan, Segalowitz & Phillips, 2016) as well as potentially better domain-general executive control skills (a matter which is currently the topic of much debate but beyond the scope of the present contribution).

In summary, all linguistic systems co-existing in the same mind are in constant interaction with each other, and this interaction constrains performance, can lead to delayed responses, less efficient processing and the production of intermediate forms. These effects are most pronounced in those instances where similar items or structures exist in both languages, facilitating transfer effects in both directions.

This leads us to the question of the mechanisms underlying such processes of L2-to-L1 transfer, and how they can be captured by current theoretical approaches to bilingual development: how are such attrition effects brought about by the interaction of

the two linguistic systems represented in the mind, can they eventually lead to contact-induced long-term restructuring – and, if so, how?

4. The mechanisms underlying language attrition

Sequential bilinguals are generally assumed to initially bring two things to the process of bilingual development: a) the linguistic knowledge that they have acquired in their L1 and b) a general knowledge of how language works (depending on the theory, such knowledge may be framed in terms of linguistic universals or of domain-general principles governing cognition and interaction). The task of the language learner is to use these two resources in conjunction with L2 input in order to gradually build an L2 system. A crucial part of this development consists of amending the expectations, automatisms and reflexes that guide the use and processing of the L1 where they conflict with the L2.

In the course of this process of acquisition, a number of changes take place which affect the L1. Firstly and most straightforwardly, the amount of L1 use diminishes, potentially constraining the accessibility of items or rules as memory traces decay. Secondly, the L1 begins to exist in a state of co-activation with a competing language system, which may lead to some (or all) of the phenomena described in the previous section. Thirdly, contact-induced change may set in, caused by the repetition and progressive entrenchment of such online transfer phenomena (with *ad hoc* L2-to-L1 mergers gradually establishing themselves as permanent features of the L1).

The interdependency of languages and the bidirectionality of transfer is widely acknowledged by largely holistic theories of L2 development and the bilingual mind, such as the Multicompetence model (e.g., Cook, 2013) or Dynamic Systems Theory

(e.g., Verspoor, de Bot & Lowie, 2011). However, most theories that focus on linguistic (in particular grammatical) structure and predict particular areas of susceptibility to transfer in bilingual development tend to consider the L1 as stable. We argue that an extension of such theories to encompass and account for changes in the L1 as the L2 is being developed and established may provide an opportunity to challenge and validate these theories.

In the following, we will exemplify the predictions that may be made for language attrition on the basis of two of the predominant current models of L2 acquisition (SLA), namely Usage-Based and Generative approaches. Since both of these approaches encompass a wide range of different theoretical models, we more specifically exemplify our argument based on one recent theoretical development within each of these approaches, namely the Unified Competition Model (MacWhinney, 2012) and Feature Reassembly (Lardiere, 2009). This limitation in scope should not be taken to imply that other hypotheses or theories might not benefit in equal measure from an application to L1 attrition.

4.1 Entrenchment and competition: the Usage-Based perspective

From the point of view of the theoretical perspective that falls under the umbrella of Usage-Based or Emergentist accounts, language is seen as an emergent and dynamic system. Development is determined by user-independent factors such as frequency, saliency and markedness (Ellis, 2016) and structure is derived from properties and interactions that are inherent to hierarchically organised levels of processing (MacWhinney, *forthc*). Learning occurs as a result of the brain making connections upon encountering an event or structure, and these connections vary in strength as a

function of frequency and entrenchment (Holme, 2013; Langacker, 1987). Events that are encountered frequently will achieve ‘unit status’, which means that they can easily and more or less automatically be evoked as integrated wholes (Langacker, 1987). Such linguistic units are not stored *per se* but exist purely in dynamic and recurring patterns of neurological activity (Langacker, 2009). If this is indeed the case, then linguistic units and structures of the L1 should be amenable to modification and disentanglement as a consequence of either being weakened through disuse (Langacker, 1987:59) or through the activation of similar and competing patterns in an L2.

A central claim of Usage-Based approaches is that processes of language development are governed by domain-general learning mechanisms (Langacker, 2009:628), that all such developmental processes rely on the same principles (e.g., frequency, perceptual saliency, see Holme, 2013) but are shaped and determined by the contexts surrounding acquisition and previously acquired knowledge (MacWhinney, 2012). For instance, infants learn language while learning about the world, they receive strong support from their caregivers, their brain is highly malleable, and they lack linguistic representations in earlier-learned languages. Such factors render child L1 development distinct from later L2 acquisition (MacWhinney, 2012). As such, *any* context of language processing and development – L1 and L2 acquisition, simultaneous bilingualism, code-switching, aphasia – should be capable of being modelled within an overarching theory, and this theory should allow identifying the contribution that other, user-specific or extralinguistic, factors will make in each context (MacWhinney, 2012). One attempt to provide such a theory is MacWhinney’s Unified Competition Model (MacWhinney, 2012, henceforth UCM). According to this model, language development is determined by the availability (frequency) and reliability of cues in the

input, with reliability playing an increasingly strong role in more proficient speakers to the extent that it remains the only factor which plays a role for adult natives (MacWhinney, 2012).

The UCM attributes (positive and negative) transfer in bilinguals to the interplay of risk-generating processes and support processes. Among the support processes are entrenchment, resonance and decoupling: L1 knowledge is first represented in local cortical maps which become more stable, and thus more resistant to modification with increasing age and decreasing plasticity of the brain (MacWhinney, 2008). This is what the L2 learner has to contend with, and it explains why there is an effect of age on potential ultimate success: the longer a speaker was monolingual before the onset of L2 acquisition, the more deeply entrenched and thus the more resistant to modification the pre-existing L1 cortical maps have become. However, L2 encoding can be achieved through the factor of resonance – essentially, the process of linking new information with existing knowledge, for instance through translation equivalents. Inhibition of the L1 in conjunction with resonant activation of the L2 then leads to decoupling, which is further supported if the learner can localize the L1 to those contexts where there is minimal completion (MacWhinney, *forthc.*).

The UCM has not, thus far, been experimentally extended to language attrition, but it proposes a number of highly relevant factors for this process, among them (dis-)entrenchment and negative transfer. The neural connections storing L1 knowledge and the cortical maps that were drawn up in childhood can be assumed to be vulnerable to disuse (disentrenchment), and a highly active L2 system may become a source of negative transfer to the L1. The UCM further lists a set of social factors that can serve as risk or support factors in L2 acquisition; here, the factor pair isolation/participation is

of particular relevance to L1 attrition, predicting that higher levels of L1 use will lead to better maintenance. MacWhinney (forthc.) acknowledges that findings from L1 attrition studies, such as cases of long-term language stability in mature migrants despite disuse of the L1 as well as of catastrophic loss in international adoptees pose a challenge to the UCM, and that studying such cases may help get a better understanding of the interaction of learner-independent and learner-specific features in language development.

Usage-Based approaches to language development and use thus have the potential to provide specific and testable predictions and hypotheses for attrition studies based on factors such as the frequency and reliability of cues, the similarity between languages, and the distribution of background factors. The linguistic mechanisms underlying language attrition should be similar to those proposed for L2 acquisition. With respect to the UCM, this suggests that structures that are identical between L1 and L2 (as far as such can ever exist) should be less problematic to maintain, as should be features that are unique to either language due to the absence of recurring, sufficiently similar competing patterns. What should be most amenable to change would then be those structures which are similar but different. Here, it should be possible for the increasingly deeply entrenched and ingrained L2 constructions to spill back into the L1.

Investigations of L1 attrition thus can and should be used to challenge, verify and validate the hypotheses that cognitive linguistics in general makes about linguistic knowledge and development. These hypotheses should not only hold true for contexts of acquisition and use, but also of maintenance and deterioration in attrition contexts defined by non-production, the absence of input in the L1, and competition from the L2. Researchers within this approach sometimes acknowledge that while frequent exposure

to a language or structure is what drives acquisition and entrenchment, disuse should lead to weakening of memory traces, attrition or forgetting (e.g., Langacker, 1987; MacWhinney, 2008), but this has never been pursued experimentally: While Usage-Based approaches are immensely influential in studies of both L1 and L2 acquisition, they have not, so far, been applied to language attrition. Exploring their predictions in an experimental setting may help validate them, resolve competing theories and accounts, and gain a deeper understanding of the emergent structure of human language (MacWhinney, *forthc.*).

4.2 Interfaces and feature-reassembly: Nativist approaches

The theoretical approach to language development and language knowledge discussed in the preceding section is based on the assumption that language learning is a domain-general skill. By contrast, generative or nativist approaches to SLA assume that some form of domain-specific, innate learning ability also plays a role.² The question to what extent this ability can similarly facilitate L2 learning later in life has been extremely controversially discussed (for a recent overview see Schwartz & Sprouse, 2013). Views have ranged across the entire spectrum from full availability of UG to L2 learners in exactly the same way as in native language acquisition (e.g., White, 2003) to its complete inaccessibility after a certain maturational stage (e.g., Clahsen & Muysken, 1986).

² It is beyond the scope of the present paper to argue for or against one of these paradigms. We do feel, however, that language attrition studies have the potential to feed into this larger discourse in a similar way to the contribution they can make to the theoretical debates within each paradigm, as discussed here: The assumption of an innate learning mechanism inherently predicts a more stable 'endstate' for language development, but in particular for L1 acquisition, than do dynamic, Usage-Based models. Appropriately designed investigations of L1 attrition may therefore be able to contribute to the debate on the existence of an innate Language Acquisition Device, over and above what investigations of L2 acquisition are able to achieve.

The focus of many of the investigations arguing for a particular perspective within this continuum is the *non-convergence* of SLA – the fact that the performance of L2 learners at a stable endstate is less consistent than that of mature, monolingual speakers. However, bilinguals also become non-convergent in the process of L1 attrition, and the question is whether these two types of non-convergent systems are constrained by the same properties and systematicities (e.g., Sorace, 2005).

The challenge for generative approaches to bilingual development in the context of L1 attrition is thus essentially the same as the one outlined for cognitivist/emergentist approaches above: to explain the increased variability and optionality that can be observed in language attrition data, based on the underlying theory of bilingual development. This has, so far, been attempted mainly within the framework of the Interface Hypothesis (e.g., Sorace, 2005; 2011; henceforth IH). This hypothesis assumes that core syntactic features are unproblematic for both L2 learners and attriters. Both near-native L2 speakers and advanced L1 attriters, however, have higher levels of optionality in comparison with monolinguals where phenomena situated at interfaces are concerned. This variability is ascribed to factors such as limitations in working memory, processing capacity or efficiency, and resource allocation (Rothman & Slabakova, 2011). In particular, the model makes a distinction between internal (syntax-semantics) and external (syntax-discourse) interfaces, predicting external interfaces as the locus of emerging optionality in attrition (and residual optionality in SLA). This is the result of the increased task demands of integrating information across a linguistic module (e.g., syntax or semantics) and a non-linguistic one (e.g., discourse) (as opposed to the integration across two linguistic modules, e.g., between the core module and the syntax-semantics interface).

The IH thus makes important testable and falsifiable predictions for the attritional process, in particular that core syntactic features should not be vulnerable to attrition processes. While this is indeed what a number of studies have found (e.g., Perpiñán, 2011; Tsimpli et al.), others do report changes that suggest syntactic representations may become affected (for a recent overview see Domínguez, 2013: ch. 5). Of particular interest here is Iverson's (2012) case study of Pablo, a Chilean Spanish-Brazilian Portuguese speaker who represents an extreme case of L1 shift. Iverson tests a range of phenomena situated at external interfaces (e.g., contexts licensing overt vs. null subject pronouns), internal interfaces (e.g., definite vs. indefinite non-overt subjects) and belonging to narrow syntax (e.g., specific cases of subject-verb inversion). These phenomena are further classified into cases where properties are shared across the two languages and those where they diverge (Iverson 2012: 49-51). A series of timed and untimed grammaticality, acceptability and interpretation judgments shows that, for all linguistic features where there is a divergence between languages, Pablo consistently patterns with a Brazilian Portuguese monolingual control group in both his languages. In other words, for this speaker L1 attrition appears to have effected a restructuring of the underlying grammar to reflect properties of the L2, and this restructuring has affected all three types of phenomena to a similar extent. The fact that the speaker's performance was the same under timing constraints and in situations where this pressure was absent leads Iverson to argue that the divergences from the native norm truly represent a qualitatively different grammar, as opposed to a mere processing issue, a finding not predicted by the IH.

To some extent, it may be possible to explain such findings by interaction effects between the syntactic module and the interfaces (Domínguez, 2013) and to limitations

to computational efficiency, causing integration problems at one interface which may constrain processing on others (Hopp, 2010). However, both Iverson and Domínguez stress that their findings suggest that the interface-based model should be modified to accommodate the possibility that syntactic representations may themselves become impaired.

It is important to note here that many of the studies which seem to suggest restructuring deal with speakers who are exposed to structurally very similar languages or varieties (see also section 5.1 below). In particular, one of the studies reported by Domínguez (2013) illustrates that changes to the distribution of grammatical features of Cuban Spanish (null subjects and postverbal subjects) among speakers immersed in Miami adapt towards the properties of a different variety of the L1 (Peninsular Spanish) also spoken in that environment, not towards the L2, English. By contrast, the distributional properties of L1 Spanish among otherwise similar speakers living in a largely monolingual English-speaking environment in the UK remain much more stable. These findings suggest that L1 attrition may be less a matter of lack of evidence in the input (as originally suggested by Sharwood Smith & van Buren, 1991) and more a case of rapid, contact-induced change or dialectal levelling (Domínguez, 2013).

The assumption that syntactic restructuring can take place as a result of the long-term co-activation of a similar language system does seem more likely than one where changes are caused merely by the absence of confirming evidence in the input (or output). However, it still leaves us with the explanatory challenge of what exactly the mechanisms underlying this process of modification are. As Domínguez (2013) points out, the standard model of language acquisition adopted in generative approaches – a scenario which assumes a stable linguistic environment once an ‘endstate’ has been

reached – is not very well suited to making predictions of what happens if this stability is disrupted. This is particularly true models such as parameter-setting – Domínguez (2013) observes that under such a model, changes should have dramatic, cascading consequences, and such wide-ranging changes are typically not been observed in L1 attrition.

A promising recent alternative approach within generative approaches to L2 acquisition which allows modeling more localised and subtle changes to the L1 is Feature Reassembly (Lardiere, 2005, 2009). Couched within a minimalist framework of grammatical representation (e.g., Chomsky, 1995), this approach rejects earlier theories on variability in L2 acquisition based largely on on-off-switch-like notions of parameter-setting as too simplistic (Lardiere, 2009). It proceeds from the observation that morphological forms across languages can encode multiple grammatical features simultaneously on what seems to be one morpheme, but in fact is often a bundle of multiple features. As defined here, features refer to specific, abstract grammatical functions – such as Case, Definiteness, Logophoricity, Durativity, Evidentiality – encoded on lexical items (Adger & Svenonius, 2011).

For example, the English plural suffix *-s* and the Chinese *-men* both encode plurality. However, the latter is additionally specified for the features [+human] and [+definite] (Ionin, 2013), so that the two plural morphemes cannot be considered fully equivalent: At the featural level, they are distinct, in that only *-men* encodes multiple features that must be satisfied for grammatical use (the noun must be plural as well as human and definite). Similarly, the English past tense *-ed* not only marks tense [+past], but also grammatical aspect [+ perfective] (or [-perfective]) in conjunction with the context, while the past tense suffix in other languages may encode other features, for

example evidentiality in Turkish (Karayayla, *forthc.*) or Quechua (Putnam & Sánchez, 2013).

The level of granularity that features provide has been shown to be very useful in explaining developmental paths and outcomes in different contexts (Lardiere, 2009). The task of all learners is thus to acquire the entirety of the bundle of grammatical features associated with any particular lexical head of the target grammar and assemble them onto the lexical form. For young children, this process is relatively straightforward since they have no previous linguistic knowledge which might get in the way. For sequential bilinguals, on the other hand, cross-linguistic differences between the compositionality of feature bundles on lexical items that ostensibly have similar functions could be the root of differences from monolinguals (Ionin, 2013). Lardiere's Feature Reassembly Hypothesis (FRH) claims just that: L2 acquisition is complicated by the influence L1 feature configurations have on how – and even if – target feature bundles will be acquired in the L2. If the L1 and L2 have the exact same feature configurations for a given property, the FRH does not predict any difficulty in L2 acquisition. However, whenever the assemblies of features are distinct, the task of acquiring the target is potentially problematic and subject to variability due to cross-linguistic influence.

L2 acquisition is thus seen as an incremental process. The learner starts with the feature configurations from her L1 (Full Transfer), and some features may be adjusted to the L2 settings earlier than others (Lardiere, 2009). In other approaches it is furthermore assumed that L2 features may fluctuate between settings, both among populations and within individuals (e.g., Ionin, Zubizarreta & Maldonado, 2008).

Taken together these approaches make it possible to account for input- and activation-based fluctuations to (some) settings in the L1. Feature Reassembly has, so far, been applied to changes in an L1 in two contexts, neither of which are ‘classic’ attrition settings, namely heritage languages and dialect levelling. Both are contexts which can be assumed to facilitate grammatical mergers more than immersed adult bilingualism, in the former due to the fact that the L1 grammatical settings are presumably still developing and hence more malleable, and in the latter because the two varieties in contact are minimally different and thus encourage co-activation and transfer.

Putnam and Sánchez (2013, henceforth P&S) study children who are exposed to Quechua in the home and to Spanish in the environment. They argue that one of the crucial factors for contact-induced change to the L1 is processing for comprehension and for production – that is, processing where input is converted to intake: Largely irrespective of how much input a heritage speaker receives in her L1, the level of activation of individual features or constraints depends on the depth of processing (intake), and if activation of particular (morpho)syntactic expressions is low, this can lead to a decline in the availability of functional features (FFs). Less salient or less frequent elements of the heritage language, as well as ‘fringe’ elements whose change has little impact on other areas of grammar, can become recessive and eventually be dissociated and reassigned to L2 features due to the continued activation of the L2 (see also Cuza & Pérez-Tattam’s 2016 discussion of [+/- strong] features in HL development). In other words, progressive re-assembly may occur from the dominant L2, in effect leading to a grammar that is complete but different from the monolingual target of both L1 and L2, that is, a rule-governed and logical system whose feature

bundles end up being distinct. This implies that the continued activation of the L2 over the L1 not only impacts on parsing strategies but can eventually permanently influence the configurational of formal feature bundles – that is, the amalgamated grammar exists at the level of competence.

P&S propose a four-stage model, where stage 1 corresponds to mainly online re-assembly effects due to high co-activation levels, while at stage 4 the learner struggles to activate L1 features of all types (phonological, semantic and functional) for both production and comprehension.

In the context of L1 attrition, the question then is whether there are mechanisms under which the activation of feature bundles of specific lexical heads (e.g., functional morphology) in the L1 of sequential bilinguals can similarly become weakened, leading to an eventual dissociation and re-assembly towards the L2 settings. If so, it should be determined whether and to what degree these mechanisms may affect both early and late bilinguals – and if not, it should be determined why not. P&S assume that heritage speakers may have only weakly activated some (particularly late acquired or infrequent) L1 features, making them susceptible to competition and replacement by similar FFs from the L2, particularly if a morpheme encoding a similar FF exists in both languages. They explicitly acknowledge that one of the problems in their model is the impossibility at establishing a point within development when acquisition is ‘complete’ – a mechanism that would, presumably, have to be couched in terms of some kind of assumption along the lines of a Critical Period. Is there a stage in L1 development at which L1 features become fully resistant to being weakened by non-activation and competition, and if there is, what does it tell us about maturational constraints?

P&S do not extend their discussion to attrition among mature speakers. The only investigation to date, to the best of our knowledge, which applies the FRH to the adult context is provided by Domínguez & Hicks (2016) in a reassessment of Domínguez' study of null and postverbal subjects in L1 Spanish mentioned above. What is identified here as the main factor driving language change is not so much the influence of L2 English but prolonged and intensive contact with speakers of varieties of Spanish in which null and postverbal subjects are used with different frequencies than in the native variety of the participants. In other words, Domínguez & Hicks ascribe the reason for the change in feature specification not to loss of contact with the L1 or pressure from English, but “to exposure to L1 input with different grammatical properties” (Domínguez & Hicks, 2016: 68). But can we really consider exposure to two different dialects of the same language as something that is qualitatively distinct from exposure to two different languages (and where do two varieties cease to be dialects of each other and begin to be different languages)? How much similarity is necessary for transfer to occur?

To summarize, the two studies just discussed do allow for the possibility of an L1 grammar to be modified through a process that is similar to the one described in L2 acquisition by the Feature Reassembly approach. However, they suggest that such a modification will occur only if a) the exposure to a different linguistic system takes place at a relatively early age or b) the speaker is exposed to a variety of the L1 in which the feature under observation is differently distributed or realized than in the native variety. The question which remains open is if, and under what conditions, similar processes of change and restructuring are also possible among mature native

speakers through exposure to a different L2. An answer to this question would significantly benefit our understanding of developmental processes and constraints.

4.3. Summary

The two approaches to bilingual development discussed above proceed from different assumptions about the architecture of the human language faculty. Both, however, are compatible with the view that the knowledge of a native language, once acquired, does not have to be stable but is susceptible to processes of change and adaptation in the bilingual mind. What these processes will look like and how the two languages will interact with each other depends on a number of key factors. Interestingly, the factors assumed to play a role appear to be the same for both frameworks: First, similarity between the two linguistic systems facilitates interaction; second, frequency of exposure and co-activation is predicted to play a role; and, third, the age at which the speaker becomes bilingual appears to be an important factor.

In the following, we will address these three factors and illustrate, based on recent findings, what is known about how they influence the attritional process.

5. Factors driving the attritional process

5.1 Crosslinguistic similarity

As was pointed out above, processes of change in language attrition are the most productive in those instances where the two languages are sufficiently similar to allow some kind of spillover. Such crosslinguistic similarities are assumed to be of importance both by Usage-Based approaches – as this is where competition between two alternatives may occur – and by FR approaches, under which the activation of some

features of lexical items, which have become weakened in the L1 through co-activation of the L2, might eventually lead to those features being adapted towards their L2 settings.

Where linguistic areas other than morphosyntax are concerned the impact of crosslinguistic similarity on language attrition has often been demonstrated, for example in the context of cognates and lexical accessibility, or in the adaptation of similar phones towards the setting of the other language (see above, section 3). The main difference between these areas and morphosyntax is that the lexicon and the phonetic repertoire can tolerate a certain amount of variance, while inflections or word order are discrete. This may mean that a speaker becomes quite variable in the pronunciation of, for example, a particular phoneme or the use of any given semantic field without this constituting an outright violation of the target norm. In particular for speakers who are part of an immersed L1 community, such minor changes may foster and accelerate a gradual process of linguistic change and adaptation throughout the bilingual social network. It is more difficult to predict how such changes may take place in morphosyntax.

Two factors are probably of importance here. The first is formal (near)-equivalence and the second is the status of the resulting construction in the attriting language – is it an acceptable (although possibly dispreferred) option, or does it constitute an actual violation? Consider the English and German simple and periphrastic past tense: both are formed identically (through suffixation of the verb stem or *ablauting* vs. auxiliary/participle constructions, respectively), but while both German tenses only locate an event in the past, the English is furthermore specified for aspect. For an L1 English-L2 German learner who had successfully acquired this distinction, the aspectual

specification of the English system might become weakened due to co-activation. She might then use some of the forms infelicitously (e.g., “I have worked in the US in the 1960s”). Self-reports from attriters suggest that they are, at least initially and often for a long time, sensitive to such violations in their own output and that of others, which may constitute a barrier towards such online fluctuations taking hold. A strong resilience of grammaticality intuitions is also suggested by recent neurolinguistic investigations, which find that brain responses to morphosyntactic violations are extremely stable in L1 attrition (e.g. Bergmann et al., 2015a; but see also Kasparian, 2015). This may be one reason why error rates in attrition studies rarely exceed 5% of obligatory contexts (Montrul, 2008: 265). In the reverse case, however (that is, L1 German – L2 English), there is nothing to prevent the German attriter from using the periphrastic past to refer to perfective states and the simple past to ongoing ones, as both are perfectly acceptable (though not conventionalized) options, so the statistical distribution of the use of each tense may change, but this will not lead to perceptible anomalies.

Previous research suggests that linguistic changes that create unconventional but not ungrammatical distributional patterns may indeed be more common in the course of language attrition. For example, in the study by Domínguez (2013) discussed above, contact between Peninsular and Cuban Spanish led to a variety in which the distribution of overt and postverbal subjects shifted to resemble the environmental variety, and similar findings have been reported in a number of other studies of null and overt pronouns (for a recent overview see Gürel, *forthc.*). In the same vein, Jackson, McDermott & Schmid (2011) conducted an analysis of word order in free speech from L1 German speakers immersed in an L2 English or an L2 Dutch setting. It was hypothesized that long-term immersed German-English speakers might fail to apply the

verb-second (V2) rule in main clauses consistently, due to the high number of English XSVO constructions they were exposed to on a daily basis, but that German-Dutch bilinguals would show no such change, Dutch also being a V2 language. A somewhat higher proportion of inaccurate constructions (4.3%) was indeed observed in the English-German population than among the Dutch-German (2.57%) and the German monolingual speakers (2.67%), but as V2 was still targetlike in over 95% of all main clause utterances it would be a stretch to claim that actual restructuring had taken place. On the other hand, the distribution of different constituents occupying the pre-verbal position in the correct sentences was exactly the same between the German-English bilinguals and the controls, but had shifted towards a lower incidence of subjects and a higher proportion of adverbial, temporal and prepositional phrases in the German-Dutch speakers, approximating the distribution of these elements in monolingual Dutch.

Under certain conditions, however, it seems that constructions may take hold although they are not licensed by the L1 grammar. A number of studies have demonstrated how grammatical intuitions may change for speakers of closely related languages. Iverson's investigation of a Chilean Spanish-Brazilian Portuguese speaker (Iverson 2012) is an interesting case in point, as it shows a complete shift towards the settings of the L2. For example, Pablo consistently rejects SV inversion in declarative sentences wherever BP does not allow it, both in cases that are optional and those that are obligatory in Spanish. A similar tendency, although somewhat less pronounced, is reported by Ribbert & Kuiken (2010) with respect to the Dutch and German complementizer-infinitive construction *om – zu.../om – te...*. This construction follows a complex pattern with different constraints requiring and restricting its application in the two languages and is often problematic for bilinguals (Brons-Albert, 1994). Ribbert and

Kuiken show that the immersed Germans' intuitions have remained largely intact, with the exception of the category which is ungrammatical in German but optional in Dutch, and which their L1 German-L2 Dutch participants failed to reject in a third of the cases. In both of these studies, similarity between languages and constructions is likely to be a main factor which facilitated a change to the attriters' sense of what is and is not grammatical in their L1.

These findings seem to suggest that the disentanglement or restructuring of linguistic patterns in the process of L1 attrition may, for most speakers, be kept in check by more or less intact grammaticality intuitions. Exceptions to this mainly appear to occur in situations where there is either a very close correspondence between forms, a highly productive pattern of co-activation of the languages, or both.

5.2 Exposure and co-activation

One of the most compelling and most often invoked factors in language attrition is the frequency with which a speaker uses the attriting language. This factor is assigned particular importance in Usage-Based approaches (see above) as well as psycholinguistic studies of linguistic accessibility and activation thresholds (e.g., Paradis, 2007), which assume that the retrieval of any linguistic item will depend on frequency and recency of its prior activation.

The relationship between use and attrition was taken to be axiomatic for a long time. The only early studies which took this factor into account revealed conflicting findings, possibly due to inconsistencies in methodology (for a discussion see Köpke & Schmid, 2004). The first more in-depth investigations of a range of measures of lexical diversity and fluency as well as overall accuracy, however, failed to find any consistent

relationship with a broad range of language exposure and use measures (Schmid, 2007; Schmid & Dusseldorp, 2010). The absence of an effect of exposure, and in particular of the frequency of use in informal contexts, has since been replicated across a host of investigations of attrition across a range of linguistic levels, such as perceived foreign accent (Hopp & Schmid, 2013; de Leeuw, Schmid & Mennen, 2010) or lexical diversity in free speech (Yılmaz & Schmid, 2012; Schmid & Jarvis, 2014) to syntactic complexity (Yılmaz, 2011). The only studies which do find an effect of L1 use are Opitz (2013), who found a weak negative correlation between frequency of use and scores on a C-Test, Bergmann, Nota, Sprenger and Schmid (2016), who find a higher level of use led to a more native-like perceived accent, and Köpke (1999), who reports an overall effect of L1 use on lexical errors in a picture description task and accuracy in grammaticality judgments.

These findings suggest that what is important is not so much the frequency of use of the L1 but the mode in which it is activated and used (Grosjean, 2001). Processing-based accounts, Emergentism and Usage-Based Theory as well as the Nativist approaches represented here by the Feature Reassembly Hypothesis all rely to some extent on the notion of co-activation of linguistic systems. It is this co-activation, rather than the use of one language and disuse of the other, that leads to mergers at the online level which may or may not then spill over to the underlying representation. Frequent use of the L1 within a community where code-switching is the norm may thus trigger an accelerated process of contact-induced L1 change (Grosjean & Py, 1991).

This suggests that a much more prominent role should be given in investigations of bilingual development to the notion of code-switching: Language systems are probably never co-activated more closely than in speakers who are dense code-switchers (Green,

2011). On the other hand, bilingual speakers who frequently use one of their languages in a setting in which it is inappropriate to code-switch, for instance at work or with monolinguals, especially professionals such as (simultaneous) interpreters (Woumans, Ceuleers, van der Linden, Szmalec & Duyck, 2015), have extensive practice in minimizing co-activation and crosslinguistic interference, in addition to a more developed monitoring system.

Several studies have found that it is mainly speakers who use the L1 regularly in professional contexts who show relatively minimal attrition effects, while there is no such benefit among speakers who use the L1 mainly within the family or with friends (de Leeuw, Schmid & Mennen, 2010; Schmid, 2007; Schmid & Dusseldorp, 2010). Speakers who use their L1 professionally perform better on a verbal fluency task and have higher lexical diversity and a lower error-rate in free speech than those who do not use it regularly at all as well as those who use it mainly in social settings (Schmid & Dusseldorp, 2010). In addition, professional language use has been associated with a more native-like perceived foreign accent (de Leeuw, Schmid & Mennen, 2013) and shorter naming latencies for low-frequency lexical items (Yılmaz & Schmid, 2012). At the other end of the spectrum in Iverson's (2012) highly attrited speaker, Pablo (see section 4.2 above), who himself claims not to actually speak Portuguese. In all his daily interactions, he speaks Spanish while his family, friends and colleagues answer in Portuguese. It is likely that it is this unusual and productive co-activation which has produced the sweeping change to his L1 grammar.

Such findings can be accounted for on the basis of Green's model of inhibitory control (e.g., Green 1986; 2011) as well as Grosjean's language mode model (Grosjean, 2001), both of which are integrated in the recent control process model of code-

switching as proposed by Green & Li Wei (2014): in professional contexts, many external stimuli contribute to keeping the activation level of the L2 high, so the speaker has to exert effort to suppress it (for example in a language teacher, who receives constant L2-influenced input from her students but has to ensure that her own use of the L1 remains target-like), contributing to the development of enhanced inhibitory control. All in all, the interaction between L1 use and L1 attrition is therefore a more complex one than often assumed, and more fine-grained analyses are necessary to establish a clearer picture.

5.3 The role of AoA

Thus far, our discussion has focused largely on L1 attrition among mature, sequential bilinguals. It should, however, be acknowledged that the role of AoA for L1 attrition is an important consideration for any theoretical approach, since findings indicate that there are massive differences between the developmental outcomes of early and late bilinguals where the language they were exposed to first is concerned.

It is a well-established finding that the developmental trajectory of the home language among learners exposed to more than one language in childhood is extremely complex: It varies per child and per situation, is affected by a wide range of factors (such as the cultural context, literacy etc.), is often characterised by non-linearities, spurts and reversals in development etc., and its outcome is impossible to predict (e.g., De Houwer, 2009; Kupisch & Rothman, 2016; Montrul, 2008; Pascual y Cabo & Rothman, 2012; Scontras, Fuchs & Polinsky, 2015). Of particular relevance here is the observation that, despite the fact that such early bilinguals or Heritage Speakers (HSs) are exposed to their home language from birth, as adolescents or adults many of them

share more characteristics with adult L2 learners than with monolinguals (e.g., Montrul, 2008; Montrul, Davidson, de la Fuente & Foote, 2014). They often fail to establish target-like representations of grammatical categories, even if they had been exposed solely or largely to the home language up to and beyond the age at which these categories are typically mastered in monolingual L1 acquisition (e.g., Cuza & Perez-Tattam, 2016; Karayayla, *forthc.*; Montrul et al., 2014; Polinsky, 2008), suggesting the necessity for either an extended period of entrenchment or some kind of maturational stabilization effect after the rule has been acquired in order to decrease vulnerability to erosion (Schmid, 2012). By comparison, attrition effects observed among post-puberty bilinguals in the L1 are typically limited, and such attriters tend to be much more similar to monolinguals than to L2ers (see above, see also Schmid, 2014).

Findings such as these have important implications for our understanding of L1 and L2 grammars, their interaction, and the development of these phenomena across the lifespan. They tentatively suggest that:

1. for many bilinguals (some) L1 grammatical features are susceptible to dissociation and re-assembly
2. for most individuals this susceptibility decreases with a higher age of onset of bilingualism, but in some relatively rare individual cases restructuring may take place beyond that age

Treating heritage language development and L1 attrition as different developmental contexts carves up a continuous spectrum of L1 development into artificially distinct categories. This has resulted in the emergence of two related but mainly independent fields of research: investigations of attrition, which focus largely if not exclusively on post-puberty bilingualism, and studies of heritage speakers, which consider populations

that became bilingual roughly between birth and school age. Extremely few studies examine the AoA range in between these two or make comparisons across age groups.³ A more holistic approach which considers changes to the native language across the entire AoA spectrum may allow us to gain important insights into aspects of bilingual development at different ages: The many comparisons of the L2 development of bilinguals of all AoAs and the extensive discussions surrounding the role of AoA for ultimate success have been among the most challenging, interesting and informative contributions to our understanding of L2 acquisition. A similarly integrative and integrated perspective on how the age at which a speaker becomes bilingual will affect L1 development is necessary in order to complement this understanding of how and why the age of learning plays such an important role for ultimate success in L2 learning (see Schmid, 2009, 2014).

Conclusion

We have argued in the above that expanding the perspective on bilingual development to fully acknowledge and include changes to the L1 has the potential to broaden the scope of the discussion, refine theoretical models and gain insight into phenomena that are as yet not fully understood. In order to achieve this, we have challenged a number of preconceived notions:

³ We fully acknowledge that the practical reasons for distinguishing L2-to-L1 effects between HSs on the one hand and late bilinguals on the other are obvious and compelling. Firstly, it can be difficult to design linguistic tasks that are not either too difficult for HSs or too easy for attriters, making it taxing to investigate both groups within the same study or with the same methodology. Secondly, unlike in investigations of L1 attrition among late bilinguals, the baseline is very hard to establish for HSs: we can assume with a reasonable amount of confidence what the L1 knowledge of a post-puberty learner would have looked like at the time of L2 onset, but if another language is thrown into the developmental mix at such an early stage, all bets are off.

- the notion that ‘linguistic development’ is invariably, or even typically, unidirectional, involving a series of changes in which linguistic knowledge increasingly comes to resemble that of the mature, stable native speaker
- the notion that there is such a thing as ‘complete development’ in the native language, and that after this stage has been reached, the processes and forces which drive (first or second) language acquisition will cease to operate on this language system
- the notion that there is a difference between ‘normal’ processes of crosslinguistic interference from L2 to L1 which affect all bilinguals, and some ‘abnormal’ process of language attrition which affects only a subset of long-term, immersed L2 speakers who make little use of their L1.

Counter to these notions, we propose that all of a bilingual’s languages are in a constant state of flux and adaptation, and that the same processes and principles which drive and guide the acquisition of a language will also come to bear, in similar ways, on linguistic knowledge that has already been developed. Investigations of how grammatical processes that had previously been mastered may change or deteriorate are as informative about the nature of human linguistic knowledge as those that focus on their acquisition. Both should be capable of being accounted for within the same theoretical framework – and if the framework fails to predict patterns which can be shown to occur in attrition, this should invalidate the theory in the same manner as would counterevidence from language acquisition studies.

Acknowledgements

We are grateful to the editorial board of LAB for commissioning this epistemological review and to the editors, Tanja Kupisch and Holger Hopp, for their patience and extensive input on various previous drafts. We would also like to thank Laura Domínguez, Jason Rothman and Rasmus Steinkrauss and three anonymous reviewers for their input and helpful comments. All of these have been of enormous benefit to this chapter. All remaining inaccuracies are entirely our own.

References:

- Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., Keim, R., Cappa, S. R. & Costa, A. (2011). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cerebral Cortex*, 22(9): 2076-86.
- Adger, D., & Svenonius, P. (2011). Features in minimalist syntax. In C. Boeckx (Ed), *The handbook of linguistic minimalism* (pp. 27-51). Oxford: Blackwell.
- Ahlsén, E. (2013). Language attrition. In C.A Chappelle (Ed.), *The encyclopedia of applied linguistics*. Oxford: Blackwell. DOI: 10.1002/9781405198431.wbeal0477
- Andersen, R. W. (1982). Determining the linguistic attributes of language attrition. In R. D. Lambert & B. F. Freed (Eds.), *The loss of language skills* (pp. 83-117). Rowley, MA: Newbury House.
- Baus, C., Costa A. & Carreiras M. (2013). On the effects of second language immersion on first language production. *Acta Psychologica*. 142(3), 402-409.
- Bergmann, C., Meulman, N., Stowe, L. A., Sprenger, S. A. & Schmid, M. S. (2015a) Prolonged L2 immersion engenders little change in morphosyntactic processing of bilingual natives. *Neuroreport* 26, 1065-1070.
- Bergmann, C., Sprenger, S. A. & Schmid, M. S. (2015b). The impact of language co-activation on L1 and L2 speech fluency. *Acta Psychologica*, 161, 25-35.
- Bergmann, C., Nota, A., Sprenger, S. & Schmid, M.S. (2016). L2 immersion causes non-native-like L1 pronunciation in German attriters. *Journal of Phonetics*. 58, 71-86.
- Bernolet, S., Hartsuiker, R. J. & Pickering, M. J. (2007). Shared syntactic representations in bilinguals: Evidence for the role of word-order repetition.

- Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(5), 931-949.
- Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, 12(1), 3-11.
- Blumenfeld, H. K. & Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: examining proficiency and lexical status using eye-tracking. *Language and Cognitive Processes*, 22(5), 633-660.
- Brons-Albert, R. (1994). Interferenzfehler in der Muttersprache von in den Niederlanden lebenden Deutschen. In B. Spillner (Ed.), *Nachbarsprachen in Europa* (pp. 96-104). Frankfurt: Peter Lang.
- Brown, A., & Gullberg, M. (2008). Bidirectional crosslinguistic influence in L1-L2 encoding of manner in speech and gesture. *Studies in Second Language Acquisition*, 30(2), 225-251.
- Bylund, E. (2009). Effects of age of L2 acquisition on L1 event conceptualization patterns. *Bilingualism: Language and Cognition*, 12(3), 305-322.
- Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics*, 40(2), 249-268.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Clahsen, H. & Muysken, P. (1986). The availability of universal grammar to adult and child learners - a study of the acquisition of German word order. *Second Language Research*, 2(2), 93-119.
- Cook, V. (2013). Multicompetence. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 3768-3774). Oxford: Wiley-Blackwell.
- Costa, A. & Sebastian-Galles N. (2014). How does the bilingual experience sculpt the brain?. *Nature Reviews Neuroscience*. 15, 336–345.
- Cuza, A. & Pérez-Tattam, R. (2016). Grammatical gender selection and phrasal word order in child heritage Spanish: A feature re-assembly approach. *Bilingualism: Language and Cognition*, 19(1), 50-68.
- De Houwer, A. (2009). *Bilingual First Language Acquisition*. Clevedon: Multilingual Matters.

- de Leeuw, E., Mennen, I. & Scobbie, J. (2012). Singing a different tune in your native language: L1 attrition of prosody. *International Journal of Bilingualism*, 16(1), 101-116.
- de Leeuw, E., Schmid, M. & Mennen, I. (2010). The effects of contact on native language pronunciation in an L2 migrant context. *Bilingualism: Language and Cognition*, 13, 33-40.
- Domínguez, L. & Hicks, G. (2016). Synchronic change in a multidialectal community: evidence from Spanish null and postverbal subjects. In A. Cuza, L. Czerwionka & D. Olson (Eds.), *Inquiries in hispanic linguistics: from theory to empirical evidence* (pp. 53-72). Amsterdam: John Benjamins.
- Domínguez, L. (2013). *Understanding Interfaces: Second language acquisition and first language attrition of Spanish subject realization and word order variation*. Amsterdam: John Benjamins.
- Duncan, H. D., Segalowitz, N. & Phillips, N. A. (2016). Differences in L1 linguistic attention control between monolinguals and bilinguals. *Bilingualism: Language and Cognition*, 19(1), 106-121.
- Dussias, P. E. (2004). Parsing a first language like a second: The erosion of L1 parsing strategies in Spanish-English Bilinguals. *International Journal of Bilingualism*, 3, 355-371.
- Dussias, P. E. & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish-English bilinguals. *Bilingualism: Language and Cognition*, 10(1), 101-116.
- Dussias, P. E., Perrotti, I., Brown, M. & Morales, L. (2014). Re-learning to parse a first language: the role of experience in sentence comprehension. *27th CUNY Conference on Human Sentence Processing, Columbus (OH)*.
- Ellis, N. C. (2016). Salience, cognition, language complexity, and complex adaptive systems. *Studies in Second Language Acquisition*, 38(2), 341-351.
- Ervin, S. & Osgood, C. E. (1954). Second language learning and bilingualism. In C. E. Osgood & T. A. Sebeok (Eds.), *Psycholinguistics* (pp. 139-146). Bloomington, IN: Indiana University Press.

- Flege, J. E. (1987). The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, 15(1), 47-65.
- Frenck-Mestre, C. (1993). Use of orthographic redundancies and word identification speed in bilinguals. *Journal of Psycholinguistic Research*, 22(4), 397-410.
- Friesen, D. C., Jared, D. & Haigh, C. A. (2014). Phonological processing dynamics in bilingual word naming. *Canadian Journal of Experimental Psychology*, 68(3), 179-193.
- Gollan, T. H., Montoya, R. I., Fennema-Notestine, C. & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition*, 33(7), 1220-1234.
- Green, D. W. (1986). Control, activation and resource: A framework and a model for the control of speech in bilinguals. *Brain and Language*, 27(2), 210-223.
- Green, D. W. (2011). Language control in different contexts: the behavioral ecology of bilingual speakers. *Frontiers in Psychology*, 2, 103.
- Green, D. W. & Wei, L. (2014). A control process model of code-switching. *Language, Cognition & Neuroscience*, 29(4), 499-511.
- Grosjean, F. (2001). The bilingual’s language modes. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 1-22). Oxford: Blackwell.
- Grosjean, F. & Py, B. (1991). La restructuration d’une première langue: l’intégration de variantes de contact dans la compétence de migrants bilingues. *La Linguistique*, 27, 35-60.
- Gürel, A. (forthc.). Null and overt pronouns in attrition. In M. S. Schmid & B. Köpcke (Eds), *The Oxford handbook of language attrition*. Oxford: OUP.
- Heredia, R. R. & Cieslicka, A. B. (2015). *Bilingual Figurative Language Processing*. Cambridge: CUP.
- Himmel, M.-C. & Kabak, B. (2016). Adaptation or attrition? L1 rhoticity in American English-German late bilinguals. Paper presented at the 3rd International Conference on Language Attrition, University of Essex, July 2016.
- Hohenstein, J., Eisenberg, A. & Naigles, L. (2006). Is he floating across or crossing afloat? Cross-influence of L1 and L2 in Spanish–English bilingual adults. *Bilingualism: Language and Cognition*, 9(3), 249 – 261.

- Holme, R. (2013). Emergentism, connectionism and complexity. In J. Herschensohn & M. Young-Scholten (Eds), *The Cambridge handbook of second language acquisition* (pp. 605-626). Cambridge: CUP.
- Hopp, H. (2010) Ultimate attainment in L2 inflectional morphology: Performance similarities between non-native and native speakers. *Lingua* 120: 901-931.
- Hopp, H. (2013) Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, 29(1), 33-56.
- Hopp, H. & Schmid, M. S. (2013). Perceived foreign accent in L1 attrition and L2 acquisition: the impact of age of acquisition and bilingualism. *Applied Psycholinguistics*, 34(2), 361-394.
- Ionin, T. (2013). Morphosyntax. In J. Herschensohn & M. Young-Scholten (Eds), *The Cambridge handbook of second language acquisition* (pp. 505-528). Cambridge: CUP.
- Ionin, T., Zubizarreta, M. L. & Maldonado, S. B. (2008). Sources of linguistic knowledge in the second language acquisition of English articles. *Lingua*, 118(4), 554-576.
- Iverson, M. (2012). *Advanced language attrition of Spanish in contact with Brazilian Portuguese*. PhD dissertation, University of Iowa.
- Jackson, C., McDermott, L. & Schmid, M.S. (2011). Changing syntactic preferences in L1 attriters of German. *Paper presented at ISB 8*, Oslo, June 2011.
- Karayayla, T. (forthc.). Effects of first language attrition on heritage language input and ultimate attainment: two generations of Turkish immigrants in the UK. In B. Brehmer, D. Berndt & J. Treffers-Daller (Eds), *Lost in transmission*. Amsterdam: John Benjamins.
- Kasparian, K. *The Case of the Non-native-like First-Language: Neurophysiological investigations of first-language attrition and second-language processing*. PhD Dissertation, Montreal: McGill.
- Köpke, B. (1999). *L'attrition de la première langue chez le bilingue tardif : implications pour l'étude psycholinguistique du bilinguisme*. PhD Dissertation, Toulouse: Université de Toulouse-Le Mirail.
- Köpke, B. & Schmid, M. S. (2004). First language attrition: The next phase. In M. S. Schmid, B. Köpke, M. Keijzer, M. & L. Weilemar (Eds.), *First language attrition:*

- Interdisciplinary perspectives on methodological issues* (pp. 1-43). Amsterdam: John Benjamins.
- Kroll, J. F. & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and control. *Journal of Cognitive Psychology*, 25(5), 497-514.
- Kroll, J. F., Dussias, P. E., Bice, K. & Perroti, L. (2015). Bilingualism, mind, and brain. *Annual Review of Linguistics*, 1, 377-394.
- Kupisch, T. & Rothman, J. (2016). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*, doi: [10.1177/1367006916654355](https://doi.org/10.1177/1367006916654355).
- Langacker, R. W. (1987). *Foundations of Cognitive Grammar: Theoretical Prerequisites*. Stanford, CA: Stanford University Press.
- Langacker, R. W. (2009). A dynamic view of usage and language acquisition. *Cognitive Linguistics*, 20(3), 627-640.
- Lardiere, D. (2005). On morphological competence. *Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference (GASLA 2004)* (178-192). Somerville, MA: Cascadilla Proceedings Project.
- Lardiere, D. (2009). Some thoughts on a contrastive analysis of features in second language acquisition. *Second Language Research*, 25(2), 173-227.
- MacWhinney, B. (2008). A Unified Model. In N. Ellis & P. Robinson (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 341-371). Lawrence Erlbaum Press.
- MacWhinney, B. (2012). The logic of the unified model. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 211-227). New York: Routledge.
- MacWhinney, B. (forthc.). Attrition and the competition model. In M. S. Schmid & B. Köpcke (Eds.), *The Oxford handbook of language attrition*. Oxford: OUP.
- Major, R. (1992). Losing English as a first language. *The Modern Language Journal*, 76(2), 190-208.
- Marian, V. & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition*, 6(2), 97-115.

- Mayr, R., Price, S. & Mennen, I. (2012). First language attrition in the speech of Dutch–English bilinguals: The case of monozygotic twin sisters. *Bilingualism: Language and Cognition*, 15(4), 687–700.
- Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics*, 32(4), 543-563.
- Montrul, S. (2008). Second language acquisition welcomes the heritage language speaker. Opportunities of a new field. Review article. *Second Language Research* 24(4), 487-506.
- Montrul, S., Davidson, J., de la Fuente, I. & Foote, R. (2014). Early language experience facilitates gender agreement processing in Spanish heritage speakers. *Bilingualism: Language and Cognition*, 17(1), 118-138.
- Opitz, C. (2013). A dynamic perspective on late bilinguals' linguistic development in an L2 environment. *International Journal of Bilingualism*, 17 (6), 701-715.
- Paradis, M. (2007). L1 attrition features predicted by a neurolinguistic theory of bilingualism. In B. Köpcke, M. S. Schmid, M. Keijzer & S. Dostert (Eds.), *Language attrition. Theoretical perspectives* (pp. 121-133). Amsterdam: John Benjamins.
- Pavlenko, A. (2003). "I feel clumsy speaking Russian": L2 Influence on L1 in narratives of Russian L2 users of English. In V. Cook (Ed), *Effects of the second language on the first* (pp. 32-61). Clevedon, UK: Multilingual Matters.
- Pavlenko, A. (2010). Verbs of motion in L1 Russian of Russian-English bilinguals. *Bilingualism: Language and Cognition*, 13(1), 49-62.
- Perpiñán, S. (2011). Optionality in bilingual native grammars. *Language, Interaction, Acquisition*, 2(2), 312-341.
- Polinsky, M. (2008). Gender under incomplete acquisition: Heritage speakers' knowledge of noun categorization. *Heritage Language Journal*, 6(1), 40-71.
- Putnam, M. T. & Sánchez, L. (2013). What's so incomplete about incomplete acquisition? A prolegomenon to modeling heritage language grammars. *Linguistic Approaches to Bilingualism*, 3(4), 478-508.
- Ribbert, A. & Kuiken, F. (2010). L2-induced changes in the L1 of Germans living in the Netherlands. *Bilingualism: Language and Cognition*, 13(1), 41-48.

- Rothman, J. & Slabakova, R. (2011). The mind-context divide: On acquisition at the linguistic interfaces. *Lingua*, 121(4), 568-576.
- Sancier, M. L. & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, 25, 421–436.
- Sandoval, T. C., Gollan, T. H., Ferreira, V. S. & Salmon, P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition*, 13(2), 231-252.
- Sanoudaki, E. & Thierry, G. (2015). Language non-selective syntactic activation in early bilinguals: the role of verbal fluency. *International Journal of Bilingual Education and Bilingualism*, 18(5), 548-560.
- Schmid, M. S. (2007). The role of L1 use for L1 attrition. In B. Köpcke, M. S. Schmid, M. Keijzer & S. Dostert (Eds), *Language attrition. Theoretical perspectives* (pp. 135-153). Amsterdam: John Benjamins.
- Schmid, M. S. (2009). On L1 attrition and the linguistic system. *EuroSLA yearbook 9*: 212-244.
- Schmid, M. S. (2011). *Language Attrition*. Cambridge: CUP.
- Schmid, M. S. (2012). The impact of age and exposure on forgetting and retention of the birth language in international adoptees: a perspective from Holocaust survivors. *Linguistic Approaches to Bilingualism*, 2(2), 177-208.
- Schmid, M. S. (2014). The debate on maturational constraints in bilingual development: A perspective from first-language attrition. *Language Acquisition*, 21(4), 386-410.
- Schmid, M. S. (2016). Language attrition timeline – annotated bibliography. *Language Teaching*, 49(2), 186-312.
- Schmid, M. S. & Dusseldorp, E. (2010). Quantitative analyses in a multivariate study of language attrition: The impact of extralinguistic factors. *Language Research*, 26(1), 125-160.
- Schmid, M. S. & Jarvis, S. (2014). Lexical access and lexical diversity in first language attrition. *Bilingualism: Language and Cognition*, 17(4), 729-748.
- Schwartz, B. D. & Sprouse, R. A. (2013). Generative approaches and the poverty of the stimulus. In J. Herschensohn & M. Young-Scholten (Eds), *The Cambridge handbook of second language acquisition* (pp. 137-158). Cambridge: CUP.

- Scontras, G., Fuchs, Z. & Polinsky, M. (2015). Heritage language and linguistic theory. *Frontiers in Psychology*, 6, 1545.
- Seliger, H. W. & Vago, R. M. (1991). The study of first language attrition: An overview. In H.W. Seliger & R.M. Vago (Eds.), *First language attrition* (pp. 3-16). Cambridge: CUP.
- Sharwood-Smith, M. A. & Van Buren, P. (1991). First language attrition and the parameter setting model. In H. W. Seliger & R. M. Vago (Eds.), *First language attrition* (pp. 17-30). Cambridge: CUP.
- Shook, A. & Marian, V. (2013). The bilingual language interaction network for comprehension of speech. *Bilingualism: Language and Cognition*, 16(2), 304-324.
- Sorace, A. (2005). Selective optionality in language development. In L. Cornips & K. P. Corrigan (Eds.), *Syntax and variation. Reconciling the biological and the social* (pp. 55-80). Amsterdam: John Benjamins.
- Sorace, A. (2011). Pinning down the concept of “interface” in bilingualism. *Linguistic Approaches to Bilingualism*, 1(1), 1-33.
- Sprenger, S. A., Bergmann, C. & Schmid, M. S. (submitted). The effect of fixed expressions on speech fluency in monolingual and bilingual speakers of German. **\$update before publication\$**
- Tamminen, H., Peltola, M.S., Toivonen, H., Kujala, T. & Näätänen, R. (2013). Phonological processing differences in bilinguals and monolinguals. *International Journal of Psychophysiology*, 87(1), 8-12.
- Tsimpli, I., Sorace, A., Heycock, C., & Filiaci, F. (2005). First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English. *International Journal of Bilingualism*, 8(3), 257-277.
- Ulbrich, C. & Ordin, M. (2014). Can L2-English influence L1-German? The case of post-vocalic /r/. *Journal of Phonetics*, 45C, 26-42.
- Van Assche, E., Duyck, W., Hartsuiker, R.J. & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in sentence context. *Psychological Science*, 20, 923-927.
- Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review*, 9(4), 780-789.

- Verspoor, M., De Bot, K. & Lowie, W. (Eds.). (2011). *A Dynamic Approach to Second Language Development: Methods and techniques* (Vol. 29). Amsterdam: John Benjamins.
- White, L. (2003). *Second Language Acquisition and Universal Grammar*. Cambridge: CUP.
- Whitford, V. & Titone, D. (2015). Second-language experience modulates eye movements during first- and second-language sentence reading: Evidence from a gaze-contingent moving window paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 1118-1129.
- Woumans, E., Ceuleers, E., Van der Linden, L., Szmalec, A. & Duyck, W. (2015). Verbal and non-verbal cognitive control in bilinguals and interpreters. *Journal of Experimental Psychology-Learning Memory and Cognition*, *41*(5), 1579–1586.
- Yilmaz, G. (2011). Complex embeddings in free speech production among late Turkish-Dutch bilinguals. *Language, Interaction and Acquisition*, *2*(2), 251-275.
- Yilmaz, G. & M. S. Schmid (2012). L1 accessibility among Turkish-Dutch bilinguals. *The Mental Lexicon*, *7*(3), 249-274.