Introduction

To think 'bilingualism' is usually to think 'second language' (L2). Virtually all research that is conducted on multilinguals focuses on the way in which they acquire and use the languages which are not their 'native' one. Observations gained from that study are then compared to first language (L1) speakers, and attempts are made to pinpoint and explain the differences between the two populations (L1 speakers vs. L2 learners).

The assumption underlying this bias is that once a 'native' language has developed to a certain point, it becomes stable and hence uninteresting. Second language learning, on the other hand, is so obviously impacted on by a large number of other factors, among which crosslinguistic influence, age at acquisition, attitude and motivation, that it is considered a profitable area of research. Bilingual speakers are, in other words, taken to have one language system in which they are identical to monolinguals (their native language) and one which shows influence from that system across all linguistic levels. This view is schematically represented in Fig. 1.

On the other hand, there are cases where the presumed stability of the native language apparently can become compromised: speakers who have lived in a migration environment for a long time can sometimes show signs of 'forgetting' their L1. When they attempt to speak it, they may have difficulties in remembering certain words, they may make what appear to be lexical or grammatical 'mistakes', or they may speak their native language with a foreign accent. This is usually ascribed to a scenario where the
picture represented above has reversed: for these speakers, it is the second language (the language of the country to which they emigrated, and which they speak in daily life) which has become dominant, and which now influences the native language. Originally, this was the view which was assumed to be the case in first language attrition (Fig. 2).

/ insert Fig. 2 about here /

The view of the first language system as something that is stable and invulnerable is extremely pervasive. One’s 'native' or 'mother tongue' may be something that is so intimately linked with early development, identity, cultural roots and so on that any indication that this fundament may be open to change or crosslinguistic influence, and that native speakers may turn into 'foreigners', is often perceived as threatening or shocking. This is evident in the reactions of reproach or even outrage which language attriters are often confronted with. For example, Stolberg & Münch (2010) quote a native speaker of German who had lived in the US for a long time, and relates how one of her correspondents would minutely go over each letter that she had written, pointing out any 'errors' in language use. At some point, her admission that she did not always know how to say certain things in her L1 any more elicited his remonstration that "One does not forget one's mother tongue!"

Similar responses abound in popular culture. In the 5th volume of Alexander McCall Smith’s Botswana Ladies Detective series (The Full Cupboard of Life), the protagonist, Mma Precious Ramotswe, reflects on the case of a fellow Botswana who had gone to live abroad as a young man, and, upon his return many years later, been unable to understand and interact in Setswana anymore. To the native Setswana, he appears to have become a foreigner. This experience deeply shocked Mma Ramotswe, who feels that such a thing is "like forgetting your mother, and as sad" (p. 163). The unforgettable title character of Peter Høeg’s Miss Smilla’s Feeling for Snow compares the erosion of her native Greenlandic to stainless steel snaplinks which had been exposed to the most hostile climate on the planet, the fog of the Arctic Sea, for a few months and had corroded to the point that they could be picked apart with a fingernail (p. 119). And a minor character in Stephen King’s Dead Zone, who escaped from Warsaw as a boy (and alone), calls the fact that he can no longer speak or understand his native Polish "a
shameful thing” (p. 122). These examples illustrate the extent to which the native language is commonly felt to be the incorruptible foundation on which identity is built.

Early investigations of first language attrition consequently assumed that this would be a process that would only occur under the most extreme of circumstances, among speakers who had lived in emigration for a very long time, hardly if ever used their L1, and possibly had particularly strong reasons to reject it and acculturate into their new surroundings (e.g. Andersen 1982, Waas 1995). Such speakers would then, almost by definition, be particularly successful second language learners.

However, psycholinguistic investigations of the bilingual mind have contributed towards a change in perception, and shown that within the continuum of bilingual proficiency, both of the representations shown above are idealizations which do not obtain in reality. In the mind-brain of any speaker of more than one language, a linguistic supersystem is created, where all languages are merged to some degree. The representation and integration of linguistic categories - semantic, phonetic, or grammatical - within this system may not entirely correspond to or overlap with the way they are represented in the mind of a monolingual native speaker. For example, Pavlenko (forthc.) mentions the well-known case of the boundaries and prototypes of color categories which appear to differ for monolingual and bilingual speakers of some languages; and Flege (1987) has shown that some phonemes can take on characteristics of the other language in both of a bilinguals’ languages.

These and other findings strongly suggest that 'extreme' cases of language attrition, where speakers become perceptibly different from natives, are merely the more noticeable end of the full spectrum of what Cook (e.g. 2003) has termed 'multicompetence'. In any such system, even among beginning learners of a foreign language, crosslinguistic traffic can go both ways (see Fig. 3).

/ insert Fig. 3 about here /

[A]The extent of attrition effects
Largely anecdotal reports of substantial language attrition - such as the ones quoted above - notwithstanding, corpus based quantitative investigations of the attritional process have usually found that the actual degree to which the use of a language has changed is surprisingly low - as long as the speaker did not migrate before the age of 12.
There is considerable cumulative evidence which suggests that this age is, in fact, a turning point for the development of the first language: among younger speakers, it appears that knowledge of the native language can indeed be lost dramatically or even totally, while post-puberty migrants often perform within the native range on a variety of tasks (for an overview see Köpke & Schmid, 2004 and Schmid, 2009). It has therefore been proposed that the two processes - changes to the first language among children versus adolescents/adults - should be rigorously distinguished: in the former instance, it is more appropriate to speak of incomplete acquisition, while the term attrition should be reserved for those speakers whose native language acquisition process has been completed prior to migration.

Many investigations of this latter process (attrition of a language which has been completely acquired) seek to answer the question of whether it can eventually come to affect the structure or (to use a somewhat old-fashioned term) competence in a native language, or whether it is merely a matter of use or performance. However, while it was proposed that "it is erosion that reaches the level of competence that allows for interesting claims about and meaningful insights into the attrition process" (Seliger & Vago, 1991, p. 7), among mature speakers there are to date no documented instances of such radical processes of deterioration. It has been shown that attrition is more liable to affect features which are preferential or optional, such as the use of null vs. overt pronouns, rather than deterministic grammatical relations (Gürel, 2004; Tsimpli, Sorace, Filiaci & Heycock, 2004; Tsimpli, 2007).

Truly structural changes, such as an overall reduction of a complex case system where some of the cases would become radically underused or disappear completely have never been observed in an attritional setting. On the other hand, investigations such as the ones by Schmitt (2004, 2010) suggest that for younger migrants whose acquisitional process was not completed prior to the change in linguistic environment, such a wholesale simplification can, in fact, occur: her investigations of L1 speakers of Russian who were four years old (Schmitt, 2004) and 8-10 years old (Schmitt, 2010) at migration shows a dramatic reduction in the target-like application of all oblique cases, while the nominative seems relatively unaffected by the erosion process.

Schmid (2009) argues that such findings imply that among mature speakers, attrition cannot affect the underlying representation of linguistic knowledge, while for incomplete learners, such a reconstruction is possible. She demonstrates that even the
most strongly attrited case in her overall corpus totaling some 140 attriters, an L1 speaker of German who left her native country at the age of 11 and made no use of German for five decades, still compares favorably with a highly advanced L2 learner of German who speaks it daily on a range of grammatical features as well as with respect to perceived foreign accent.

[A] Psycholinguistic models of language attrition

It therefore seems that the indications of "attrition" that are perceived and remarked on by the attriters themselves and by other speakers of their L1 are usually more linked to problems of (predominantly lexical) access and momentary effects of crosslinguistic influence than to an actual restructuring of the underlying language knowledge. Such effects may be best conceived of and explained within psycholinguistic models, such as Paradis' Activation Threshold Hypothesis (Paradis, 2004). This model holds that any memory trace, such as a word, an expression or a morpheme, has a certain Activation Threshold (AT) associated with it. The higher this threshold, the more effort it takes to recall the item from memory. The AT is determined by frequency and recency: items that have been very frequently called upon have a low threshold, but if the item is not activated for a long time, the threshold will gradually rise.

Based on this hypothesis, Paradis (2007, p. 121) predicts that it is the disuse of a language which will lead to gradual language loss. This prediction is probably the most intuitively and immediately obvious assumption with respect to language attrition. If two migrants who had left their home country at a similar age and lived in migration for a similar length of time show differences with respect to the degree of erosion and L2 influence that is perceptible in their L1, most people will likely assume that they have had different opportunities to use and 'practice' their language skills.

Rather surprisingly, however, such a straightforward correlation between 'amount of L1 use' and 'degree of L1 attrition' appears to be extremely difficult to establish in actual attrition settings. A number of large-scale quantitative investigations have attempted to determine a link between L1 use in a variety of contexts (interactive, passive exposure (books, TV), inner speech) and attriters' performance on a number of different tasks and for different populations (Schmid, 2007; Schmid & Dusseldorp, 2010; van der Kooi, in prep.; Yilmaz, in prep.). These analyses, however, did not yield the expected results: Schmid (2007), for example, finds no impact of the frequency of use of
L1 within the family or with friends on any of the linguistic variables of lexical diversity and fluency in free speech which she has measured. The only L1 use setting which has some impact is the use of the L1 for professional purposes.

Similar results are found in a number of other ongoing investigations, for example a study of the bilingual development of Moroccan and Turkish speakers in the Netherlands (van der Kooi, in prep., Yilmaz, in prep.). Again, the use of the L1 in informal settings plays no role for the task these studies have conducted (a picture naming task), and it is only the use of the L1 at work which is associated with quicker naming times and higher accuracy.

Schmid (2007) therefore speculates that for a mature native speaker who has lived in his or her native environment until puberty, the first language knowledge might have been rehearsed to the point where frequent reactivation is no longer necessary to prevent deterioration. She proposes that in such a case it may be that the problems which attriters can experience are not so much due to the fact that their L1 knowledge has become inaccessible due to a high Activation Threshold, but to the increase in competition from the second language. In other words, it is not so much that attriters 'forget' their L1, but that the competing L2 has become too strong to be entirely inhibited.

This explanation hinges on the fact that, in order to access a certain item of knowledge, it is not only necessary that there is enough stimulation to overcome the AT: At the same time, all other items of knowledge which may be in competition with the target have to be inhibited in order to ensure that they will not be activated by mistake. For a bilingual who routinely speaks his or her second language and only infrequently uses the first, this may mean that, while the L1 has not become inaccessible or forgotten, the L2 keeps getting in the way, leading to (voluntary or involuntary) code-switches, mistakes or other signs of crosslinguistic influence.

This explanation can account for the fact that informal use of the L1 (with family and friends), although it is probably the most frequent type of exposure, does apparently not impact on attrition: in this type of situation, all interlocutors are typically bilingual, so that the abovementioned indications of crosslinguistic influence do not impede communication, and little effort needs to be made on inhibiting the L2. On the other hand, speakers who use their L1 for professional purposes (whether as a German language teacher in Canada, a waiter in a Moroccan restaurant or as service staff in a
Turkish supermarket) often find themselves in situations where code-switching may not be considered appropriate. Consequently, they may be more practiced at inhibiting the L2 and keeping it from 'butting in' when they want to speak their first language. More routine experience with inhibiting L2 may then account for their better performance on L1 tasks.

The controversy on the nature of L2 knowledge

The findings presented above on the restructuring of the L1 system in incomplete learners vs. its preservation among attriters may help shed light on one of the most fundamental and controversial issues in the study of multilingual development: whether there are fundamental, underlying and qualitative differences between languages learned early and later in life, or whether these differences are gradient and quantitative. This question is usually addressed in the context of the impact of age on language learning.

While younger learners tend to be more successful in some areas of language learning than older ones, the question of whether or not the level of proficiency which L2 learners can attain is maturationally constrained according to the age at which they start acquiring it is highly contested. Early hypotheses (Penfield & Roberts, 1959; Lenneberg, 1967) assume a strongly biological explanation for the lower levels of L2 proficiency typically witnessed in later learners. An evolutionary advantage is ascribed to restricting highly intensive and cognitively costly learning mechanisms triggered by linguistic input to a certain time window, since no more learning is required once the skill has been acquired and the cognitive resources can be more efficiently deployed for other purposes (e.g. Pinker, 1994) This assumption is corroborated by the finding that computer simulations of human evolution end up with "a critical period for language acquisition, centered in early childhood, as the inevitable outcome" (ibid).

Other views take a less deterministic stance and assume that the fact that older learners rarely if ever attain native-likeness is not due to the total and highly specific 'dismantling' of a certain neurobiological prerequisite for language acquisition but to the fact that these learners have a mature and deeply entrenched L1 system that the L2 has to compete with. Furthermore, it is often proposed that overall cognitive development plays an important role: children acquire their first language entirely implicitly, that is, knowledge-building happens unconsciously based on input which is not cognitively and
metalinguistically analyzed. Older learners, however, bring cognitive and analytical strategies towards language learning, so that at least some of the acquisition process happens explicitly.

The controversy on the impact of age of acquisition in SLA thus centers on whether there is a qualitative or a quantitative difference between early and late L2 learners, or between L1 and L2 speakers. This leads to two different and competing views of the knowledge of late-learned second or foreign languages: one which takes residual optionality to be an indication of an underlying representational deficit conditioned by maturationally constrained limitations to language acquisition, and one which assumes that L2 learners can establish native-like underlying knowledge but fail to apply it consistently, due to the cognitive demands engendered by the competition between the two languages.

These two competing models have important implications for our understanding of the human capacity for language and in particular the degree to which the brain is specialized for language and its acquisition. However, investigations have so far found it difficult to establish conclusively which of the two scenarios applies. The findings from such studies are hardly ever deterministically supportive of one or the other point of view, and therefore invariably lend themselves to interpretations from both points of view.

In order to resolve the issue, additional data for comparison are therefore needed which also differ from native speaker data, but for reasons which are known. As was demonstrated above, language attrition data fulfill this criterion and might therefore prove useful in resolving this question.

[A]The potential contribution of investigations of L1 attrition

L1 attriters have acquired their first language as monolingual native speakers. However, among both attriters and L2 learners, overall proficiency has been observed to differ from that of monolinguals, and a higher optionality in the application of target rules has been found for the bilingual populations. While for L2 speakers the reasons for this variable performance (underlying representational deficit or problems in online integration of linguistic knowledge) is highly controversial, there are clear indications that no representational deficit obtains for L1 attriters. Should principled comparisons of L1 attriters, L2 learners and unattrited native speakers reveal that attriters differ
from monolinguals in the same way as L2 users do, that would therefore provide compelling evidence for a bilingualism effect and against theories which assume L2 acquisition to be governed by a critical period and L2 knowledge to be fundamentally different from L1 knowledge.

Extending the comparison to include incomplete learners may further render the question of whether there is a qualitative change in the capacity to establish native-like knowledge of a language around puberty empirically testable: If there are representational deficits or differences in processing for post-puberty bilinguals, L1 attriters should be most similar to a monolingual reference group, since their acquisition was unconstrained by the CP. L2 speakers and incomplete learners should also pattern together, revealing their underlying difference in knowledge and processing. If, on the other hand, a more general bilingualism effect, conditioned by the limitation of available cognitive resources, entrenchment of linguistic knowledge and proficiency level is responsible, L1 attriters, incomplete learners and L2 speakers should all differ from largely monolingual native speakers, and this difference should be modulated by factors such as amount of exposure to the non-dominant language and proficiency in that language.

Including L1 attriters in the bilingual equation can therefore help us gain insight into how dominant and non-dominant languages are processed in the bilingual mind: these speakers have native-like underlying representations in combination with competition from a highly active and strongly entrenched L2 system. Such comparisons therefore have the potential of providing insight into the nature of bilingual language acquisition and knowledge.

[A]Conclusion
L1 attrition has for a long time been a kind of Cinderella of multilingualism studies. It was considered to be a relatively rare case of language development, one which would only occur after a reversal of language dominance had taken place and the L1 had been used not at all or only minimally for a long period of time. Speakers who would experience L1 attrition were thought to have, to all intents and purposes, become native speakers of their second language and L2 speakers of their first.

The findings from attrition studies, which typically revealed even those speakers who had the longest migration period one might reasonably expect to occur within the
lifespan of an adult human being (five to six decades) to have remarkably few signs of
deterioration of their first language knowledge, were therefore often regarded with
suspicion. However, cumulative evidence is reinforcing the impression that a completely
acquired native language system can, in fact, be extremely stable and impervious to
'atrophy' or deterioration in any structural sense. Whatever indications of L2 influence
there are to be detected are probably the outcome of a temporary or momentary
'winning out' of the L2 system in the process of the formulation of utterances.

This implies that the phenomenon which we refer to as 'attrition' is not limited to
the relatively rare case of the long-term migrant who has not had occasion to use his or
her first language in decades: crosslinguistic interference, not only from L1 to L2 but
also from L2 to L1, is something which all bilinguals experience to some degree. In L1
attriters, the process is merely more noticeable than in L2 learners who continue to use
their L1 dominantly.

Based on this realization, studies of L1 attrition have the potential of helping us
resolve one of the fundamental questions about the nature of language learning:
whether languages learned later in life are underlyingly represented in the same manner
as native languages or not. I therefore propose that investigations of multilingual
development should shift their focus of investigation from comparing L1 vs. L2 to
comparing dominant vs. non-dominant languages. If language attriters can ever truly
become like a 'foreigner', in the way that Mma Ramotswe felt it about her compatriot,
that indicates that those foreigners, like the attriters themselves, probably also do not
have a representational deficit in their L2. If, on the other hand, qualitative differences
persist between attriters and L2 learners, then those differences can probably be
ascribed to a representational deficit on the part of the L2 speakers.

[Cross-references]

SEE ALSO: Attitudes and Motivation in Bilingual Education; Bilingual and Monolingual
Language Modes; Bilingualism and Cognition; Code Mixing; Dynamics of
Multilingualism; Early Bilingualism; Grammar Attrition; Inhibition and Control in
Second Language Acquisition; Multicompetence; Multilingualism and Attitudes;
Multilingualism and Emotions; Native Speaker; Second Language Representation in the
Brain
[References]


[Suggested Readings]


[Figure headings]

Fig. 1: L1 influence on L2 in SLA, from Schmid & Köpke (2007:2), their Fig. 1. Used with kind permission by John Benjamins.

Fig. 2: L2 influence on L1 in attrition, from Schmid & Köpke (2007:2), their Fig. 2. Used with kind permission by John Benjamins.

Fig. 3: An integrated view of bilingualism, from Schmid & Köpke (2007:3), their Fig. 3. Used with kind permission by John Benjamins.