Perceived foreign accent in L1 attrition and L2 acquisition: the impact of age of acquisition and bilingualism

Holger Hopp[1] and Monika S. Schmid[2]

[1] Universität Mannheim, [2] Rijksuniversiteit Groningen

Address for correspondence:

Dr. Holger Hopp

Universität Mannheim

Anglistische Linguistik

Schloss EW 266

68131 Mannheim

Tel: ++49-621-1813160

Fax: ++49-621-1812336

E-Mail: hhopp@rumms.uni-mannheim.de

Auxiliary title: Perceived foreign accent in L1 attrition and L2 acquisition

Abstract

To identify the nature of constraints on L2 pronunciation, this study undertakes a direct comparison between 40 late second-language (L2) learners, 40 late L1 attriters, and 20 predominantly monolingual German controls. This investigation can disentangle age and bilingualism effects by contrasting participants with AOAs at birth (monolinguals, L1 attriters) with late L2ers, on the one hand, and bilinguals (L1 attriters, L2ers) with monolinguals, on the other. Group results show that attriters perform indistinguishably from controls, and both differ from the L2ers. However, 80% of all L2ers score within the native (attriter) range. Correlational analyses with background factors further establish effects of use, time spent in Germany and aptitude. We discuss the roles of age and cross-linguistic influence in L2 acquisition.

(120 words)

Introduction

Studies on age effects in second language (L2) acquisition show that pronunciation accuracy in the target language is one of the most difficult skills to acquire for late learners. These investigations have consistently demonstrated that post-puberty learners across different acquisition contexts are detectably different in speech production from monolingual native speakers (NSs) and from early L2ers. The most robust finding is that foreign accent ratings show a negative correlation with Age of Acquisition (AOA), that is, the later exposure to the L2 , the stronger the foreign accent at the endstate of the acquisition process (e.g. Abrahamsson & Hyltenstam, 2009; Flege, Birdsong, Bialystok, Mack, Sung, & Tsukada, 2006). As a consequence, AOA has been taken to be the primary variable determining pronunciation accuracy in L2 ultimate attainment. However, previous research also suggests that the link may not be entirely straightforward. In particular, researchers differ in how they conceptualize age effects in L2 speech production, the crucial question being whether AOA is the *cause* of persisting transfer from the first language (L1), or whether it is merely *associated* with it (Bialystok, 1997; 2001).

Some researchers interpret age effects as a direct reflection of a critical period in the L2 acquisition of articulatory phonology and phonetics which preempts native-like pronunciation in late learners. On this view, behavioural variability in L2 production compared to native speech has been linked to constraints in neurological and fine-motor skills (e.g. Moyer, 1999). These constraints are argued to follow from maturational reductions in cerebral plasticity which categorically prevent the reorganization of the speech production (and comprehension) system from the L1 to the L2 after a certain age (e.g. DeKeyser & Larson-Hall, 2005; Hyltenstam & Abrahamsson, 2003). The location of this cutoff point for attaining nativelikeness is controversial, ranging from shortly after birth (Abrahamsson & Hyltenstam,

2009; Hyltenstam & Abrahamsson, 2003) through age six (Long, 1990), 12 (Scovel 1988) or 15 (DeKeyser, 2000).

Others have interpreted age-related increments in foreign accent to be a consequence of the degree of L1 entrenchment in phonetic categorization. For instance, the speech learning model (SLM; Flege, 2002) hypothesizes that L2 learners are increasingly likely to process L2 phonetic categories as instances of L1 categories the longer the L1 had been spoken before the onset of L2 acquisition (Flege, 1999). On this view, even though differences between L1 and L2 vowel and consonant categories may be detectable in comprehension, the classification of L2 phones as functional equivalents of L1 categories leads to the merging of L1 and L2 categories in speech production.

The entrenchment model implies that factors other than just AOA may impact on (non-) native L2 pronunciation. Previous research has considered a wide variety of predictors such as length of residence in an L2 environment (LOR), chronological age at testing, typological distance of the languages, language use, language aptitude and socio-psychological factors (for an overview, see Jesney, 2004). As has often been noted, however, many of these factors are confounded or co-vary with AOA. LOR, for example, is of necessity longer for earlier learners if they are age-matched to late learners at time of testing (e.g. Flege, 2009; Moyer, 2007; Piske, MacKay, & Flege, 2001). In a similar vein, language use is linked with AOA since early L2ers who arrive in the target-language environment before age 10 tend to adopt the TL as the main language of communication, while late L2ers retain higher degrees of use of the L1 in the target-language environment as well as in contacts to their home community (e.g. Jia, Aaronson & Wu, 2002; Piske, MacKay, & Flege, 2001).

In addition, such investigations face the methodological problem that use of the L1/2[1] can only be measured on the basis of self-reports and may therefore not be reliably estimated.

These problems may to some extent account for the fact that multifactorial analyses of L2 speech production often yield mixed results, with factors other than AOA accounting only for a small amount of variance. For instance, in a meta-analysis of data from 240 L1 Korean and 240 L1 Italian L2ers of English across several tasks and studies, Flege (2009) reports that amount of L1/2 use accounts for less than 10% of the variance in foreign accent rating data.

The controversy on the role of the age effect in SLA is further complicated by findings suggesting that even L2 speakers with AOAs well below the onset of puberty do not invariably score within the native range (e.g. MacKay, Flege, & Imai, 2006; Piske et al., 2001). For instance, Korean children who arrived in the US between the ages of 8 and 9 and who were tested after 3 and 5 years of residence were reliably rated as having a foreign accent compared to age-matched native English children (Flege et al., 2006; see also Flege, Munro, & MacKay 1995; Flege, Yeni-Komshian, & Liu 1999). Similar effects are reported for Italian learners of English by MacKay et al. (2006). On the other hand, Pallier, Dehaene, Poline, LeBihan, Argenti, Dupoux and Mehler (2003), Ventureyra (2005) and Ventureyra, Pallier and Yoo (2004) suggest that speakers who experience sequential monolingualism, i.e. a complete break in L1 input followed by a rapid breakdown of this system and full immersion in another language in all contexts of life (as experienced by international adoptees) may become fully native-like in L2 speech perception even if this language reversal took place as late as age 10 (but see Hyltenstam, Bylund, Abrahamsson, & Park, 2009).

Given the apparently relatively minor contribution of predictors other than age (such as L1/2 use) as well as the fact that even early learners may be perceptibly different from monolingual natives unless a complete language reversal has taken place, one may hypothesize that the fact of being a bilingual speaker in and of itself contributes to perceived non-nativeness. It has long been acknowledged that the endpoint of L2 acquisition cannot be

equated with dual monolingualism (Grosjean, 1998). Rather, it has been argued that multicompetence in more than one language should be taken to be the ultimate goal of L2 acquisition (Cook, 2003). In other words, a proficient bilingual inherently differs from a monolingual by virtue of accessing an integrated language processing system that is partly shared across languages. This in turn implies interactions and cross-linguistic influence in both (or all) languages at multiple cognitive and linguistic levels.

Cross-linguistic interactive effects in bilinguals have been well-documented in cognitive processing (e.g. Bialystok, 2009) and linguistic processing at different levels (e.g. van Hell & Dijkstra, 2002 for the bilingual mental lexicon; Hernandez, Bates, & Avila, 1994 for sentence processing and Cutler, Mehler, Norris, & Segui, 1989 for the structuring of phonetic space). Bilingualism affects strategies and mechanisms of L1 and L2 processing (e.g. Dussias, 2004) as well as the speed of processing in either language (e.g. Hopp, 2010; McDonald, 2000) even if both languages are acquired from early childhood (Foursha, Austin, & van de Walle, 2005; Proverbio, Cok ,& Zani, 2002; Werker & Byers-Heinlein 2008).

For pronunciation, several studies report bidirectional cross-linguistic influence in the speech production of bilinguals. Flege (1987; see also Flege & Hillenbrand, 1984) studied late English-French and French-English bilinguals of different proficiency levels after a period of residence in an L2 environment. The comparison with monolingual controls of either language revealed that the length of voice onset time (VOT) produced by the bilinguals in both their languages diverged from the monolingual norm (i.e. all bilingual speakers had shorter VOTs in English and longer VOTs in French than the controls). The degree of bidirectional influence was modulated by proficiency levels and length of residence in the L2 environment. In a similar vein, Fowler, Sramko, Ostry, Rowland and Hallé (2008) report that simultaneous (2L1) French-English speakers produce VOTs in either language that are different from those produced by monolingual native speakers (Fowler et al., 2008). [2]

Although the bilinguals' VOTs clearly differ between English and French, indicating that phones did not merge across languages, the realization of those phones in one language affects their realization in the other.

These effects point to assimilatory processes in speech production (see also Sancier & Fowler, 1997) and perception (e.g. Sundaraa, Polkaa, & Genesee, 2006). Such cross-linguistic interaction seems to result from the active use of two languages, since the monolingual controls in the Fowler et al. study did not differ in their VOTs depending on whether they were occasionally exposed to the other language or not (but see Au, Knightly, Jun, & Oh,, 2002; Caramazza & Yeni-Komshian, 1973).

Despite these widely-recognized cross-linguistic interactions which affect speech production and comprehension in bilinguals across AOAs, the reference groups in speech production studies typically consist of monolingual native speakers of the target language (e.g. Flege et al., 2006) or native speakers who overwhelmingly use the target language but might have some knowledge of other languages (e.g. Abrahamsson & Hyltenstam, 2009). It may be argued that the choice of a monolingual control group thus serves to move the yardstick of nativelikeness to a point which may, by definition, be out of reach for most bilinguals (see also Birdsong, 2005). It may therefore be more appropriate to investigate whether L2ers can approximate the performance of speakers who have acquired the target language from birth but are also late learners of an L2.

The above implies that the bilingual populations for comparison should be late learner of either the target or the competitor language. In this respect, 2L1 speakers are unsuitable as a reference point in a direct comparison with late L2 learners because they differ in their chronological onset of bilingualism. In 2L1 speakers, assimilatory tendencies in phonetic categories across languages can be observed from the beginning of language acquisition due

to concurrent L1 and L2 input (Fowler et al., 2008). In other words, 2L1 speakers arguably never develop monolingual native categories which can then be affected by the later onset of bilingualism. In this way, they fundamentally differ from late L2 learners whose native (monolingual) phonetic categories affect speech production of a late-acquired language, and vice versa.

Similarly, early (child) L2ers, who are typically compared with late learners in studies on age effects in L2 acquisition, differ in their degree of entrenchment of L1 phonetic categories. Given the comparatively shorter length of exclusive L1 use for child L2ers, the impact of the L1 on a successively-acquired L2 may be quantitatively distinct from the extent of cross-linguistic influenced experienced by late learners (Flege et al. 2006).

The present study therefore introduces a bilingual reference group who shares the chronological asymmetry in the onset of L1 and L2 input characteristic of late L2 acquisition: L1 attriters, i.e. adult long-term emigrants to a non-target language environment whose use of their early-acquired L1 is greatly reduced following emigration. Both these bilingual populations are then compared against a (largely) monolingual reference group.

L1 Attrition

It was pointed out above that bilingual proficiency does not equal the combined proficiency of two monolinguals, and that the monolingual native norm may thus not be a realistic target against which ultimate attainment in the L2 can be measured. Bilingualism and cross-linguistic interaction, however, are not a one-way street. It is widely recognized that the development of L2 knowledge, processing and use is partly influenced by the pre-instantiated knowledge of the L1. At the same time, however, the presence of a developing L2 has

ramifications for the L1, potentially leading to changes in L1 processing and use. This process, known as L1 attrition, is most marked among speakers who experience a drastic and persistant change in linguistic habits and language environment (i.e. long-term migrants in a non-L1 setting) but it is by no means confined to such cases. Multicompetence approaches to bi- and multilingualism assume that the development of a second or foreign language system, even in low-proficiency instructed learners, will to some degree impact on the L1 (e.g. Cook, 2003). That notwithstanding, investigations of L2 effects on L1 have so far usually focussed on speakers who have experienced long-term immersion in an L2 environment and concomitant reductions in L1 use (for an overview see Köpke & Schmid, 2004; Schmid, forthc.).

While there has been an increasing interest in L1 attrition over the past two decades or so (see Schmid 2010), virtually all investigations focus on lexical or grammatical features of the attriting language. For these linguistic levels it has been demonstrated that, whereas the cessation of L1 use in childhood can lead to an apparently complete loss of L1 proficiency (Pallier et al. 2003), even severe reductions in L1 contact after puberty entail only relatively minor effects on L1 maintenance. For postpubescent attrition, it has been shown that while attrition may impair lexical access (Schmid & Köpke, 2008), underlying knowledge of grammar is quite resistant to processes of deterioration and L2 interference seems to manifest itself predominantly in optionality at the interface level (Tsimpli, Sorace, Heycock, & Filiaci, 2004; Schmid, 2009). It therefore appears that, in both lexicon and grammar, attrition effects can be ascribed to the increased cognitive load of integrating two linguistic systems and retrieving elements from memory which had not been called upon for an extensive period of time., while underlying grammatical knowledge seems to remain quite stable.

Studies on the late attrition of pronunciation skills or phonetic/phonological perception, on the other hand, are few and far between.[3] Global observations of attriters' pronunciation

suggest that L2 effects may be relatively limited at this level. For example, Giesbers (1997) examines close to an hour of free speech produced by a Dutch native speaker who had been immersed in an Indonesian context for more than thirty years with few opportunities to use his L1. In the entire speech sample, Giesbers finds only 48 instances of clearly non-target pronunciation, 30 of which concern stress and intonation patterns in compounds and sentences. In the absence of control data from monolingual speakers it is, of course, difficult to estimate whether this limited number of 'deviances' differs at all from the native norm – in any case, it can hardly be considered an extensive 'loss'.

However, assessing phonological and phonetic changes on the basis of discrete categories such as 'correct' and 'deviant' is not a straightforward matter, as L2 effects on L1 speech production can take more subtle forms which might not be classifiable as 'errors' but contribute to an overall less-nativelike 'acoustic flavour' in attriters' speech. Flege's (1987) seminal study of bidirectional effects in the speech of French-English and English-French bilinguals discussed above is probably the first formal investigation of such a phenomenon among late learners. Flege's finding that there is bidirectional interference on VOTs among experienced bilinguals is confirmed by Major (1992) in the context of English-Brasilian Portuguese bilingualism. Similar bidirectional effects are reported by de Leeuw (2008) on the lateral phoneme /l/ for late German-English bilinguals and by Mennen (2004) on the suprasegmental level for late Dutch-Greek bilinguals.

Both Major and Mennen investigate five late learners of the L2, four of whom show bidirectional cross-linguistic interference. Yet, in both studies, one of the five participants appears to be exempt from cross-linguistic interference in either of her languages in formal (reading) style.[4] Different suggestions have been made to account for such inter-individual variation in L2 influence on L1. Major appeals to L2 proficiency in order to explain the variable levels of bidirectional interference among his participants, while Mennen points out

that her exceptional speaker has a lower AOA than the others (15 as opposed to 20-25). In a case study of a Brazilian Portuguese – English bilingual, Sancier & Fowler (1997) furthermore observe that the extent of L2 influence on L1 VOTs is considerably stronger after recent exposure to and use of the L1 than at the end of two periods of several months spent in an L2 environment.

Interestingly, neither Flege nor Mennen use terms such as "attrition" or "loss", instead arguing for a "'merging' of the phonetic properties of similar L1 and L2 phones" (Flege 1987, p. 62). Major (1992), on the other hand, interprets the approximation of L1 VOTs to the L2 norm in an immersion setting as straightforward "loss", arguing that "a close correlation exists between VOTs and other aspects of phonological proficiency, including global foreign accent" (p. 191). In other words, the larger the deviation in VOT from the native baseline, the more likely it is that a speaker will not be perceived to be a native, which entails that a change in VOT can contribute to the "loss" of the native-speaker status. In her study on the lateral phoneme /l/ in late German-English bilinguals, de Leeuw (2008) also opts for the term "loss", but she qualifies its implications. While her findings clearly indicate that "at the level of performance, the bilingual migrants no longer conformed to the German monolingual norm" (in that they had a higher F1 frequency and an earlier alignment of the prenuclear rise), she goes on to assert that "that despite these deviations, the German migrants to Canada are still native German speakers." (p. 203).[5]

Irrespective of the extent to which L2 effects on L1 speech production can be phonetically measured, the question arises as to whether these effects are actually perceptible in bilinguals' L1 speech. The only investigation of perceived global foreign accent in a group of L1 attriters we are aware of to date is de Leeuw (2008, see also de Leeuw, Schmid, & Mennen, 2010).[6] In this study, short excerpts of spontaneous speech elicited in a narrative task performed by German migrants (to the Netherlands and to Anglophone Canada) were

rated for native-likeness by predominantly monolingual speakers of German in Germany. These ratings were compared to the ratings received in the same experiment by a control group of predominantly monolingual speakers of German residing in Germany. Overall, group comparisons showed that attriters (L1Aers) were more likely to be perceived as nonnative than controls (for further details of this study see below).

There were, however, also considerable differences among the attriters themselves, with 35% of the speakers being consistently perceived as native-like and less than 25% receiving a clear non-native rating. De Leeuw et al. (2010) attempt to account for the difference in perceived nativeness on the basis of individual background and language use variables. By means of regression analyses they establish that neither AOA (> 17 for all L1Aers) nor LOR (> 15 years for all L1Aers) are significant predictors of perceived foreign accent in their sample. The frequency with which speakers use their L1 in informal settings with other bilinguals (family, friends) does not contribute towards perceived nativelikeness either, nor does the frequency of visits back to the home country. It is only L1 use in settings where little codeswitching is to be expected (in formal, work-related contexts or in distance communication with speakers in Germany) which has some (albeit limited) predictive power.

It therefore seems likely that the development of L1 pronunciation revealed by the phonetic microanalyses reviewed above can eventually lead to a perceived foreign accent in the native language for some experienced bilinguals. Other speakers appear to be spared from these developments and remain perceptibly native-like even after several decades of residence in an L2 environment with little opportunity to use their L1.

On the face of it, the proportion of attriters who do *not* develop a foreign accent in their L1 appears higher than the number of highly successful L2 learners who ultimately do attain nativelikeness: In de Leeuw et al.'s investigation, more than a third of the bilingual

participants were unambiguously rated as native speakers.[7] While the proportion of L2ers identified as achieving nativelikeness varies across studies, most investigations locate it around or below the 5% norm originally proposed by Selinker (1972).

A study of L2 acquisition and L1 attrition which directly compares the two populations on foreign accent may consequently offer an interesting perspective on the impact of age effects in language acquisition. Such a direct comparison allows for matching L2 speakers and native speakers on variables such as (the onset of) bilingualism and its effects, thus isolating age of onset. L1 attriters are native speakers of the target language since they have acquired it from birth. At the same time, for L1 attriters (as for L2 speakers) the target language is no the language they predominantly use. Hence, L1 attriters, like L2 speakers, experience asymmetric bilingualism effects, i.e. cross-linguistic influence from the dominant language onto the weaker one Figure 1 illustrates the rationale of the group comparisons.

--- Insert Figure 1 around here ---

Hypotheses

In this paper, we will present findings from a global foreign accent rating (FAR) experiment, comparing advanced late L2 learners of German, long-term L1 attriters of German and (predominantly) monolingual German control speakers. We aim to assess the relative impact of AOA and bilingualism in late L1 attrition and L2 acquisition. Based on the above rationale, we advance the following hypotheses:

Hypothesis 1: If L2 speech production is constrained by age effects, e.g. a critical period, AOA should be the primary determinant of (non-)nativelikeness in perceived accent. That is, late L2 learners should differ from native controls and late L1 attriters in foreign accent ratings at the group level. Moreover, the foreign accent ratings of individual late L2 learners should fall outside the range delimited by the L1 speakers of the target language with AOAs of 0. Factors other than AOA should account for little or no explained variance in foreign accent ratings.

Hypothesis 2: If L2 speech production is affected by cross-linguistic interference in bilingualism, there should be a substantial overlap in foreign accent ratings for late L2ers and late L1 attriters. In addition, factors other than AOA, e.g. LOR, use, attitudes, etc., should make a contribution towards the proportion of explained variance in foreign accent ratings.

Hypothesis 3: If cross-linguistic interference affects bilingual speech production, there should be differences between bilingual groups depending on language combination.

The Study

Global Foreign Accent Ratings

We conducted a global foreign accent rating experiment on free speech samples from L1Aers and L2ers (see sections Participants and Materials below). Such global assessments of perceived foreign accent in bilinguals have been widely used to make inferences about ultimate attainment in L2 phonology (for an excellent overview, see Jesney, 2004). Typically,

in these studies, phonetically untrained natives judges listen to samples of L2 and native speech and are asked to rate these for degree of native-likeness. Native speaker performance is thus used as the implicit or explicit[8] reference point in judging foreign accent and – concomitantly - the range of the native scores is used as the cut-off criterion for establishing native-likeness in the analysis of the rating results for the L2 speakers.

Despite the wide-spread use of foreign accent ratings, no established or standardized methodology exists for such studies (Jesney, 2004). There is therefore considerable variance between studies regarding the following issues:

- *Measurement scales*. Most studies rate speakers for native-likeness on discrete, Likert-type scales with between 3 and 10 levels, although there are also investigations using sliding scales or magnitude estimations (for an overview, see Jesney 2004, pp. 2f.).

- *Number of native control speech samples*. It has been pointed out (Flege & Fletcher 1992) that native speaker judgments of foreign accent are subject to range effects. In other words, the larger the proportion of native or near-native samples included in the experiment, the stronger the perceived foreign accent of the L2ers. The proportion of control samples included in pervious studies varies from none (E. M. Brennan & J. S. Brennan, 1981; Schmid, 2002) to 50% (Munro & Derwing, 1995), with the majority of studies using somewhere between 10% and 20% (see the overview in Jesney, 2004).

- *Type of speech*. It has been demonstrated that the proportion of L2ers who are perceived to be native-like is higher in studies which use material that was formally elicited, such as word or list reading, than in casual style/free speech (e.g. Major, 1992; Moyer, 1999; Oyama, 1976). Abrahamsson & Hyltenstam (2009) therefore argue that formally elicited samples may reflect "language-like behaviour" rather than actual L2 proficiency (p. 254). A further problem of using formally elicited stimuli is that Flege & Fletcher (1992)

established that raters were more likely to judge samples as native-like after they became familiar with them. This may then impact on experiments that recurrently use the same material produced by different speakers, favouring those participants whose speech samples occur later.

- *Length of samples*. The length of the stimuli used also varies considerably, from single words in reading list style through full sentences to entire paragraphs. In free speech, the clips used are typically 10 to 20 seconds in length (Jesney, 2004). However, it has been established that native speaker judgments are usually made very fast; indeed, Flege (1984) shows that 30 milliseconds may be enough for natives to accurately judge nativeness.

The highly divergent results across such studies, some of which suggest that a sizeable proportion of late L2ers can attain nativelikeness (e.g. Bongaerts, van Summeren, Planken, & Schils, 1997) while others find that such ultimate success is extremely rare (Moyer, 1999), may to some extent be ascribed to these methodological differences. Methodological variance notwithstanding, previous research has demonstrated the impact of a number of independent variables other than AOA on pronunciation accuracy. Perceived foreign accent has been found to correlate with length of residence (e.g.. Asher & García, 1969; Flege & Fletcher, 1992; but see Oyama, 1976[9]), amount of L1/2 use (Yeni-Komshian, Flege, & Liu, 2000) attitude and motivation (Elliott, 1995; Moyer, 1999), and aptitude (Abrahamsson & Hyltenstam, 2008). A gender effect has also occasionally been found, which appears to favour female pre-puberty learners but to disappear (or even reverse) among older learners (Asher & García 1969; Flege et al., 1995; Thompson, 1991). These factors will therefore also be considered in the present study.

Precursor Studies

The populations whose perceived foreign accent was tested in the present investigation were subsamples drawn from two earlier studies: an investigation of L1Aers reported by de Leeuw, Schmid & Mennen (2010) and an investigation of L2ers reported by Hopp (2007).

Precursor study I: L1 attrition of German in an L2 English and L2 Dutch setting (de Leeuw, Schmid, & Mennen, 2010, henceforth DSM). DSM present a study of foreign accent ratings (FARs) of 57 L1Aers. All speakers are long-term migrants who were predominantly monolingual speakers of German prior to migration. Twenty-three of these speakers resided in the Netherlands and had Dutch as a second language, while 34 resided in an English-speaking environment (the Greater Vancouver area, Canada). These speakers had been recruited for a large-scale investigation of the L1 attrition of German (e.g. Schmid, 2007; 2009). Migrants from Germany with an LOR of > 15 years and an AOA of > 17 years were invited to participate in an experiment which purported to investigate language change in Germany since the German reuinification in 1989. No other selection criteria (levels of L1 use, proficiency, etc.) were applied, and the recruitment text (which was circulated through newspapers and other media, German clubs and organisations and personal contacts) explicitly stated that it was of no concern whether German was used daily or virtually never.

Speech samples ranging in length from 12.6 to 17.7 seconds were extracted from longer narratives elicited by a film retelling task (see section Materials below) and interspersed with similar samples from five native German controls. The study employs 19 German listeners (students of phonetics at the University of Trier, Germany) who were presented with the samples in a sound-proof room. Two judgments were invited for each sample: firstly, the raters were asked to classify the speaker as native or non-native, and secondly, to indicate

their confidence in this rating on a three-point scale (certain, semi-certain, uncertain), resulting in an effective six-point Likert scale (where 1 represents the judgment 'certain of native speaker status' and 6 is 'certain of non-native status'). The experiment is replicated in Schmid (2009) with fifteen additional speakers from the same population (two German migrants in Canada, one German migrant in the Netherlands and twelve reference group speakers) and 21 native raters, so that previous FARs are available for 36 L1Aers with English as L2, 24 L1Aers with Dutch as L2 and 17 native controls.

For inclusion in the present study, the ten speakers who had received the lowest (i.e. most nativelike) and highest (most accented) ratings, respectively, from each of the two language groups were selected from this precursor study, resulting in a population of 40 L1Aers. This selection procedure was motivated by our desire to include samples representing the whole range of speakers who had been previously tested. In addition, speech samples from 16 predominantly monolingual reference group speakers were included in the analysis as baseline data.

Precursor study II: Late L2 acquisition of German by L1 speakers of English and Dutch (Hopp, 2007). Hopp (2007) tested 91 advanced late L2 learners of German with English, Dutch and Russian as their respective L1s in a variety of morphosyntactic and semantico-pragmatic domains in off-line and on-line tasks. Participants were recruited in Germany through informal networks, clubs, organizations and advertisements. All participants had AOAs above 11 (av. 15.6) and advanced German after long periods of residence in Germany (av. 13.0 years). As part of the proficiency testing, all participants completed a standardized C-test and supplied samples of speech elicited in a picture-description task. These speech samples ranged in length from 45 to 330 seconds. All speech

samples, including those of four monolingual natives, were rated for nativeness on a 10-point scale (from clearly non-native to clearly native) by three linguistically naïve predominantly monolingual native speakers of German. The rating was broken down into discrete categories, namely, fluency, vocabulary, expression, mistakes and accent. A composite total score was computed from the ratings in each category.

Based on the composite score of these three raters, the ten lowest-scoring and the ten highest-scoring L1 English and L1 Dutch participants, respectively, were selected for the present study. This led to a sample of 40 L2ers, representing the bottom and the top range of late L2ers in the precursor study. In addition, the speech samples from the four native control participants were included in the present analysis.

Participants for the Present Study

The present study compares native speakers of German who are post-puberty emigrants to either Anglophone Canada or the Netherlands (L1Aers, n=40 selected from DSM on the basis of the criteria established above) and late L2 learners of German with English or Dutch as L1s (L2ers, n=40 selected from Hopp, 2007) with native speakers of German living in Germany (NSS, n=20, 16 from DSM and 4 from Hopp, 2007). Matching the L1Aer and L2er groups in terms of language combinations ensures that the type of cross-linguistic influence is similar across the two groups. Including more than one language combination further allows us to assess whether different language combinations affect perceived pronunciation accuracy in different ways.

The selection of L1Aers was restricted to late bilinguals who had emigrated from Germany to Canada or the Netherlands after age 17 in order to ensure that the L1 had been

fully acquired prior to migration, so that the speakers would qualify as attriters rather than incomplete acquirers of the L1 (see Köpke & Schmid, 2004). For this group, the non-target-language (i.e. the L2) has become the language used most frequently in daily life after an immersion period of more than 15 years.

The L2 group consisted of late L1 English and L1 Dutch L2 acquirers of German who first had contact with the L2 after age 11 and have been long-term residents in a German-speaking environment. For the recruitment of these speakers, advanced proficiency in L2 German was applied as a selection criterion. This process of screening for advanced L2ers was applied since, as Long (1993) points out, "[t]here is no value in studying obviously non-native like individuals intensively in order to declare them non-native like." (p. 204)

Finding purely monolingual subjects of most European languages has become increasingly difficult since obligatory foreign language instruction has been implemented in most educational systems. It was therefore impossible to establish a truly monolingual control group, but the selection of NSS for this study was limited to individuals who had not acquired languages other than German before school age, never lived outside Germany, and who did not regularly use a language other than German in their daily lives. Table 1 presents an overview of age, AOA and LOR of the speakers investigated in this study.

--- Insert Table 1 around here ---

Language use and language attitudes

For the speakers investigated by DSM, a large amount of information on L1/2 use in a variety of situations as well as on language and cultural attitudes and preferences was

available, based on their responses to a sociolinguistic, personal background and attitude questionnaire (see Schmid, forthc.). This questionnaire elicited self-reports by means of 110 questions. Some of these were open-ended or categorical (yes/no), but the majority of items were elicited on a five-point Likert scale, where the highest level (1) indicated overwhelming preference for or predominance of German and the lowest level (0) an equally strong preponderance of L2. By means of principal component analysis, Schmid & Dusseldorp (2010) established the following compound variables, which were shown to possess high internal validity:

- PARTNER: frequency of L1 use with the partner (4 questions)
- CHILDREN: frequency of L1 use with children (4 questions)
- FRIENDS: frequency of L1 use with friends with a German migrant background (4 questions)

(These first three items thus pertained to the frequency of informal use of the L1 with other bilinguals, i.e. L1 use in situations where code-switching frequently occurs).

- INTERMEDIATE[10]: frequency of L1 use in situations where the interlocutors are also bilingual, but where code-switching is deemed inappropriate
    - frequency of attending a German church (2 items)
    - frequency of attending a German club (2 items)
- PASSIVE: frequency of passive exposure to target-like L1
    - frequency of exposure to German media (2 items)
    - frequency of visits to Germany
- AFFILIATION: affiliation and identification
    - L1 use for internal speech, such as thinking, dreaming and counting (3 items)

  - language and culture of preference (2 items)
  - perceived importance of L1 maintenance and its transmission to the speaker's children (2 items)
- TOTAL: total frequency of the use of German, calculated as the average of all language use variables

In order to obtain the same information from the L2ers, an online questionnaire including all questions which factored into the calculation of the compound variables and additional questions on language proficiency (such as the use of German for professional purposes) was constructed (see Schmid & Dusseldorp, 2010). All participants were invited to provide this information via a weblink sent to them by email. Unfortunately, not all of the participants in the original study replied, but 31 of the 40 L2ers included in this study (L1 Dutch, n=18; L1 English n=13) filled in the questionnaire.

Figure 2 shows that the L2ers tend to use German more frequently than the L1Aers in all contexts with the exception of the language spoken with children. In terms of affiliation, the L1Aers in the Netherlands appear to have a somewhat stronger bond with the German language and culture than the other groups, who appear to be largely balanced in their attitudes towards their L1 and their L2.

--- Insert Figure 2 around here ---

Materials

All speakers performed a narrative-descriptive task designed to elicit free speech. The L2ers and four of the native controls were asked to describe the *Cookie Theft* picture (Boston

Diagnostic Aphasia Examination) and a cartoon strip, whereas the L1Aers and the remaining 16 control speakers were given the *Charlie Chaplin - Modern Times* film narration task as described by Perdue (1993). From these narratives, speech samples ranging between 10 and 20 seconds in length were extracted. The following criteria were applied to this selection process:

- samples constituted full sentences or clause/intonational units (in order to be recognizable as grammatical structures)
- samples did not contain lexically or grammatically deviant structures, since it has been shown that such 'errors' may adversely affect FARs (McDermott, 1986 ct. after Jesney, 2004)
- samples did not contain borrowings from the L2, nor items such as proper or place names (since the L1Aers with English as their L2 had the tendency to pronounce names such as *Charlie Chaplin* in an English-like fashion, e.g. realizing the /r/ in *Charlie* as a retroflex approximant [ɻ], while German speakers adapted such items to their L1 phonology).

In addition, we attempted to select speech samples which, as far as possible, described different parts of the stimulus. This was done since the two experimental populations (L1Aers and L2ers) had been given different narrative tasks. We wanted to avoid the effect where listeners would be able to categorize speakers on the basis of topic, not accent. Since DSM used samples which all referred to the same scene of the film, while Hopp (2007) had all speech samples rated in their entirety, which meant that the actual excerpts used in the current study were not the same as the ones used in the precursor study.

All samples were normalized to 3db and background noise was reduced in order to eliminate any possibility of perceptible differences between the excerpts from the two precursor studies.

Procedure

We replicated the procedure described by DSM in that German listeners made two judgements for each speech sample. The first binary judgement determined native versus non-native speaker status (in answer to the question 'Is this person a native speaker of German?'). The second judgement expressed the level of confidence on a three-point scale. This resulted in an operative six-point Likert scale: 6=certain of non-native speaker status, 5=semi-certain of non-native speaker status, 4=uncertain of non-native speaker status, 3=uncertain of native speaker status, 2=semi-certain of native speaker status, 1=certain of native speaker status. Hence, a low foreign accent rating (FAR) reflects a speaker who was perceived as native or near-native, whereas a high FAR reflected a speaker who was rated as having a noticeable foreign accent in his or her German speech.

A silent pause of seven seconds followed each sample, and each sample was played only once. During the silent pause, German listeners assessed native- or nonnative-speaker status of the speaker they had heard and indicated how certain they were of this judgement. After the silent pause, the next sample was presented. The total duration of the sequence of 100 samples was 36.06 minutes. The samples were pseudo-randomized, and two lists were created. All 100 speech samples were played from a mediaplayer which automatically levelled volume across samples and inserted the 7-second-silence between stimuli. For logistical reasons, the stimuli were played via a state-of-the art audio system in a lecture theatre at the University of Mannheim.

Listeners

Two groups of listeners took part in the foreign accent assessment in two separate sessions. Seventy-six listeners took part in the first session, and 73 listeners took part in the second. All 149 listeners were first-year students at the Department of English at the University of Mannheim, Germany. They had received no specific phonetic training. Only those listeners who reported not to have been exposed to languages other than German in childhood were retained for analysis. In all, 130 German listeners were analyzed, 68 in the first group and 62 in the second. The German listeners also had good knowledge of English, as is standard in modern-day Germany.

Results

The control group speakers received a mean FAR of 2.36 (stdev .95). The L1Aers received a mean FAR of 2.79 (stdev 1.25), and the mean FAR of the L2ers was 3.94 (stdev 1.46) (see Fig. 3). The group differences were highly significant ($F(2, 98) = 14.033$, $p < .001$, $\eta^2 = .47$). Post-hoc comparisons (Tukey HSD) revealed that the L2ers were significantly different from both the control speakers and the L1Aers ($p < .001$). There was no difference between L1Aers and controls at the group level ($p = .258$).

--- Insert Figure 3 around here ---

The large number of raters (n = 130) did not permit an assessment of inter-rater reliability by means of measures such as the Cronbach α. However, a comparison of the average FARs received by the individual speakers in this experiment with the ones obtained in the precursor studies[11], in which ratings obtained from smaller populations were shown to have excellent realiability, revealed very strong correlations: For the L1Aers, the correlation with the average scores elicited by DSM was r=.839, $p < .001$; for the L2ers the correlation with the scores reported by Hopp (2007) was $r = .620$, $p < .001$. These robust correlations suggest that ratings across listeners and studies are highly consistent, in particular since different extracts from the speech samples by the speakers were rated in the precursor studies. Moreover, differences in presentation and context do not appear to affect the results.

As for L1 effects, a comparison of the control speakers with the bilingual populations subdivided by contact language (attriters in the Netherlands (L1ANL, mean FAR 2.67, stdev. 1.38) and in Canada (L1AEN, mean FAR 3.01, stdev. 1.18), L2ers with Dutch as L1 (L2ANL, mean FAR 3.48, stdev. 1.55) and English as L1 (L2AEN, mean FAR 4.41, stdev. 1.24), see Fig. 4) revealed a significant difference between the five groups of speakers ($F(4, 156) = 8.867$, $p < .001$, $\eta^2 = .52$)

.

--- Insert Figure 4 around here ---

Post-hoc procedures (Tukey HSD) revealed that there were no differences between the two L1A ($p = .915$) or the two L2er ($p = .142$) groups. Further, neither of the two L1A subgroups differed significantly from the controls (L1ANL, $p = .801$; L1AEN, $p = .294$). In contrast, both L2 subgroups received FARs that were significantly lower than those of the

controls (L2ANL, p = .018; L2AEN, p <.001). There was no significant difference between the Dutch L2ers of German and the attriters with Dutch or English as L2 (p = .253 and .754, respectively). However, the English-German bilinguals differed significantly from both groups of attriters (L1ANL p < .001; L1AEN p = .006).

While these group results initially suggest (a) that the L2ers are different from both controls and L1Aers, but (b) that there is no difference between L1Aers and controls, the descriptive statistics presented above also indicate considerable variance within the populations. We therefore converted the average FARs into categorical ratings. Following DSM, we defined a 'clearly native' range with a FAR between 1.0 and 2.5, an average 'uncertain' range (2.5 < FAR < 4.5), and a 'clearly non-native' range (4.5 < FAR < 6). The resulting distribution is presented in table 2 (group differences were significant, $\chi^2 = 18.649$, p = .001).

--- Insert Table 2 around here ---

The scatterplot in Figure 5 depicts the individual FARs in the three groups. As can be seen, 29 L1 attriters fell within the range delimited by the native control group with a FAR of < 3.62, while 11 scored outside this native range. As for the L2ers, 15 scored within the native range, and 32 fell within the range delimited by the L1 attriters with a FAR < 5.46.

--- Insert Figure 5 around here ---

This distribution shows that there are subsamples of bilingual speakers who are perceived differently from the majority of their peers: L1Aers who come to be perceived as non-natives, and L2ers who manage to attain a native-like accent in German. In an attempt to account for such individual differences in FARs, we tested predictors pertaining to speakers' personal background, language habits and attitudes.

There was no difference in the FARs given to male (n = 41, mean FAR 3.32, stdev 1.62) and female (n = 59, mean FAR 3.05, stdev 1.30) speakers (t = .915, p = .363). A correlation of age at testing with FAR initially revealed a significant negative relationship between these two variables (r = -.207, p = .039). It was hypothesized that this correspondence might be due to the fact that the attriting and control population were older on average than the L2ers. This assumption was confirmed by correlations for the individual subgroups, none of which was significant.

It was not possible to test the effects of all background factors, in particular AOA and LOR, in one combined analysis for the two bilingual groups, since they do not impact on L1Aers and L2ers in the same way: A longer period of residence and a lower AOA is assumed to lead to a stronger perceived foreign accent in L1Aers but to less accented speech production in L2ers.. We therefore opted for bivariate correlations within each population between AOA and LOR on the one hand and FAR on the other. These revealed that AOA did not significantly correlate with FAR for either group (L1Aers: FAR vs. AOA r = -.037, p = .823, L2ers: FAR vs. AOA r = .254, p = .114). A partial correlation between FAR and AOA, in which LOR was controlled for, did not reach significance, either (L1Aers: FAR vs. AOA r = -.074, p = .653, L2ers: FAR vs. AOA r = .253, p = .120). The scatterplot in Figure 6 illustrates the lack of a contingency between AOA and perceived accent.

--- Insert Figure 6 around here ---

While there was thus no linear relationship between AOAs and FARs for either of the two groups, a closer look at the data broken down for groups reveals an interesting non-linear, pattern (see Fig. 6a).

--- Insert Figure 6a around here ---

Whereas these scatterplots yet again illustrate the absence of a linear correlation, there is a tendency for the L2ers'ratings to fall outside the bottom right quadrant. Speakers who have an AOA below 17 are distributed across the full range of scores, but only one speaker who was older than that is rated in the native range. Given the overall low number of participants with higher AOAs in either group, however, this tendency would need to be investigated further in a larger sample with a wider distribution of AOAs.

A further correlation established that the degree of perceived foreign accent did not correlate with LOR for the bilingual groups (L1Aers: FAR vs. LOR $r = -.039$, $p = .809$; L2ers: FAR vs. LOR $r = .019$, $p = .906$). A partial correlation between FAR and LOR, in which AOA was controlled for, did not reach significance, either (L1Aers: FAR vs. LOR $r = -.076$, $p = .647$, L2ers: FAR vs. LOR $r = -.012$, $p = .944$). The scatterplot in Figure 7 illustrates the lack of a correspondence between FAR and LOR.

--- Insert Figure 7 around here ---

Further, we ran a correlation between FAR and the length of time spent in a German environment (pre-emigration for the L1Aers and post-emigration for the L2ers). This correlation was marginally significant ($r = -.221$, $p = .049$). The foreign accent ratings received by the speakers in this experiment were thus not significantly modulated by AOA and LOR, but did show some relation to the length of time spent in a German-speaking environment. As in the case of age at testing, it is possible that this effect is largely due to group differences since the attriters spent a longer period in Germany prior to their migration than the L2ers since their moving to Germany.

Next, it was assessed whether self-reported use of German and attitudes towards the L1/2 impacted in any way on FAR. As is evident from table 3, there are two significant correspondences: For the L2ers, the amount of use of German with the partner correlated negatively with FAR, indicating that those speakers who frequently use German in this context were more likely to be perceived as native-like. The averaged total frequency of interactive use of German also reached significance for the L2ers. No such correlations were found for the attriters, and none of the other language use and attitude factors were significant.

--- Insert Table 3 around here ---

Based on the significant correspondence of FAR and use of German with partner discovered here, we hypothesized that a relevant factor might be whether or not the speaker was in a relationship with a native speaker of German. We therefore investigated the distribution of

participants with or without a German partner across the three levels of FAR (native, uncertain and non-native). Table 4 summarizes the count as well as the expected count of this analysis.

--- Insert Table 4 around here ---

For both populations, these differences were not significant (L1Aers: $\chi^2 = .042$, $p = .979$, L2ers: $\chi^2 = 3.516$, $p = .172$). It may, however, be noteworthy that of the 11 L2ers who were rated as native-like, only one had a partner whose native language was not German (while one more was not in a relationship).

Finally, we attempted to investigate the extent to which internal factors, e.g. variation in language aptitude, might be associated with differences in FARs. Language aptitude is commonly defined as the individual, largely innate talent for processing language. It is taken to encompass grammatical sensitivity, phonetic decoding ability and rote and inductive learning ability (e.g. Carroll, 1981). It has been demonstrated to correlate significantly with proficiency and ultimate attainment in adult L2 acquisition (e.g. Abrahamsson & Hyltenstam, 2009; DeKeyser, 2000; Robinson, 2005) and the degree of loss of the L1 in late L1 attrition (Bylund, Abrahamsson, & Hyltenstam, 2010).

Unfortunately, no direct aptitude measures were available for the participants tested in the present study. Given that aptitude scores are typically strongly correlated with proficiency measures of different types (for an overview, see Dörnyei & Skehan, 2003), we used the C-test scores obtained for all participants as part of the general proficiency testing to estimate individual aptitude. In a linear regression analysis, the predictors AOA, LOR and total use

together do not contribute significantly to the total explained variance in the C-test scores (L1Aers: $F(3, 36) = .660$, $p = .582$, $r^2 = .052$, L2ers: $F(3, 27) = .516$, $p = .675$, $r^2 = .054$). We thus surmise that the C-test scores can be taken to provide an indirect and partial measure of language aptitude. For all bilinguals, there is a moderate correlation of the C-test score and FAR at $r = -.490$ ($p < .001$), which also holds for the L1 attriters ($r = -.435$, $p = .005$) and the L2ers ($r = -.472$, $p = .002$) separately. The correlation between C-test score and FAR remains strong at $r = -.428$ ($p < .001$) when AOA, LOR and total use are controlled for, which suggests that the C-test score is a strong predictor of native-like pronunciation, largely independently of other external factors.

--- Insert Table 5 around here ---

Discussion

The findings presented above allow for a number of interesting observations. It was shown that, at group level, late bilingualism leads to stronger foreign accents in the L2 than in the L1. In other words, advanced L2 speakers are overall outperformed in terms of perceived native-likeness by long-term attriters, who learned the language under investigation from birth but have not been using it dominantly for a substantial period of time. Of the 40 L1Aers investigated, 29 (72.5%) scored within the range of the unattrited, pre-dominantly monolingual native controls, while only 13 L2ers (37.5%) fell within this range. This indicates that late bilinguals who start out as native speakers of the target language on average still remain closer to the native benchmark than those who approximate it coming from another language. The group results thus suggest that it is easier to retain an early-

acquired language across an extended period of non-exposure than to attain it from scratch at a later age. At first glance, these results seem to support Hypothesis 1, which holds that L2 acquisition is maturationally constrained.

However, the differences in FARs between the bilingual groups are not nearly as categorical as Hypothesis 1 would predict. In other words, the perceived difference between long-term bilinguals and native speakers in terms of their pronunciation cannot be ascribed entirely to AOA or the sequence in which the languages were acquired: A subset of the attriters (who had acquired German from birth) were perceived to be clearly non-native, while a number of late L2ers did fall within the unambiguously native range. Moreover, there was a sizeable overlap between the L1A and the L2 group – in fact, 32 of all L2ers (80%) fell within the range of perceived foreign accent delimited by the L1 attriters. In a direct comparison of L1 attriters and advanced L2 learners, four fifths of all L2ers were thus rated no worse than native speakers of German in terms of perceived foreign accent. Importantly, this considerable overlap in our study holds true for a population of L2ers who were not pre-screened for native-likeness in speech production. Unlike in most studies on ultimate attainment in L2 speech production (e.g. Abrahamsson & Hyltenstam, 2008; 2009; Bongaerts 1999), inclusion in this study was contingent solely on advanced general proficiency in a C-test (Hopp, 2007). In addition, L2 and native performance overlap in extemporaneous speech which has previously been found to be among the most challenging tasks for late L2ers in L2 global accent studies (e.g., Jesney, 2004).

Contrary to Hypothesis 1, this finding poignantly illustrates that AOA of the target language at birth in and of itself does not ensure sustained native-likeness in speech production. Rather, the late onset of bilingualism affects speech production in such a way that some native speakers lose their perceived native accents after prolonged immersion in a non-target-language environment. In the present study, the ages of emigration in the L1 attriter

group ranged from 17 to 51 years. In view of these late onsets of L2 acquisition, it can safely be assumed that all L1 attriters had acquired German to the monolingual adult standard before emigration. In other words, it is unlikely that any perceived foreign accent in German might have been the result of incomplete acquisition. Instead, the loss of nativelikeness is a consequence of late bilingualism.

In view of the general finding that FARs varied substantially within and across both bilingual populations, we attempted to account for variation of accent ratings in terms of the bilingual experience. We established that the length of time that participants had spent in a bilingual setting (LOR) did not contribute to the FAR. In other words, attriters with longer periods of residence in an L2 environment were not rated to be less native-like than those whose emigration had taken place more recently. In a similar vein, those L2ers who had spent the longest time in the target language environment did not achieve significantly better FAR scores than those who had come to Germany only a few years ago.[12] For L1 attrition, it has often been proposed that the bulk of the development takes place within the first decade after emigration (e.g. de Bot & Clyne, 1994). The speakers investigated here had thus likely reached their ultimate attrition stage during this time, and the long period of residence (> 15 years) stipulated as an inclusion criterion for the present investigation may thus have prevented measurable LOR effects.

Second, the time spent in a German-speaking environment did significantly correlate with FAR scores. Since this correlation only obtained for all bilingual groups, yet not for the L1 attriters and the L2ers separately, it is likely due to the group effect that the L1 attriters were comparatively old at emigration. Nevertheless, the contingency suggests that the degree of entrenchment of German – as expressed in the time spent in a German-speaking environment – affects perceived foreign accent. The overall difference between L1 attriters and L2ers in terms of FAR can thus partially be related to the stronger entrenchment of

German in the L1 attriter group *vis à vis* the L2ers.

We further investigated whether foreign accent is affected by whether participants use the target language on a regular basis, or how they feel about the German language and German culture. We therefore assessed the impact of self-reported frequency of use of and exposure to German and of linguistic and cultural affiliation on FARs. Again, this analysis did not yield any tangible explanatory findings for the L1 attriters. These results are in line with previous findings that have also reported null effects of such factors on L1 attrition across different linguistic levels (e.g. Schmid, 2007; Schmid & Dusseldorp, 2010). Similarly, for the L2ers, which language the participants predominantly used with their children, friends, in clubs or churches or for professional purposes did not influence their perceived nativelikeness. Furthermore, the FARs were not impacted on by the language or culture which the participant preferred.

Only one of the individual external factors, namely, the amount of German spoken with the partner, did reach significance. Based on this finding, we established that of those L2ers who were rated as native-like, only one had a partner who was not a native speaker of German (while one other was not currently in a relationship). This finding is interesting in the light of neurobiological investigations on issues such as stimulus appraisal and language learning, which have suggested that emotional involvement may be one of the best predictors of success in L2 learning (e.g. Schumann, 1998). Alternatively, the effect of partner may reflect the high relative frequency of language use in a relationship compared to the overall language use in life. When frequency of use was calculated as the sum total of all its documented component variables, it did in fact show a significant contingency with FARs for the L2ers. In other words, when added across contexts and situations, the relative amount of use of the L2 affects the degree of perceived foreign accent for L2 speakers. No such correspondence was found for the L1 attriters.

These findings suggest that, for speech production, the ability to retain the native status once the language has been acquired, is largely independent of external factors. On the other hand, the ability to attain the native status in speech production after puberty is affected by the overall amount of use of the target language (see also Flege & Liu, 2001; Yeni-Komshian, 2000). Seen in conjunction with the effect of time spent in Germany, the present findings are compatible with Hypothesis 2, and, by consequence, with models of L2 speech production that emphasize the degree of entrenchment of the L1 versus the L2 (e.g. Flege's Speech Learning Model; Flege, 2002). According to these models, increased frequency of use of the L2 loosens the relations of L2 sounds to L1 phonetic and phonemic categories. For the L2ers in this study, greater use of German led to less perceptible foreign accents. However, the use variables in this study were based on global self-estimates collected at one point in time. Such estimates might not give a reliable indication of changes or developments in levels of L1/2 use, and, more importantly, they do not allow for conclusions about the type and quality of L1/2 input. As, for instance, Flege (2009) points out, more extensive data on use and input are needed to gauge the precise impact of external variables.

With respect to Hypothesis 3, the cross-linguistic comparisons in this study also revealed L1/2 effects. Although there were no group differences according to L1/2 within the L1 attriter and the L2 groups, respectively, a comparison of all bilingual groups by language revealed a marginally significant difference between those speakers with Dutch as the L1/2 and English as the L1/2 ($F(1, 78)$ 3.780; $p = .055$). In line with Hypothesis 3, this difference suggests that the amount of cross-linguistic interference in speech production may to some extent be conditioned by phonetic and phonological differences between languages. Interference between German and Dutch, which are typologically and phonologically more closely related than German and English, leads to a less perceptible foreign accent in L1 attrition and L2 acquisition. These differential effects on FAR by language combination

suggest that foreign accent is linked to cross-linguistic interference at the phonetic and phonological level. The cross-linguistic differences observed between the L1 English and the L1 Dutch learners of German hence support models which view L2 speech production as the gradual restructuring of L1 speech categories to the target classification as a function of the relative distance of L1/2 categories and the extent of L2 input.

In this vein, the lack of use effects found here for the L1 attriters would appear to suggest that the L1 categories had been entrenched to a greater degree before the onset of bilingualism, such that late bilingualism has a less pronounced effect on L1 speech production overall. Our finding that length of time spent in a German-speaking environment correlated with FAR indeed strengthens this conclusion. In general, it is further backed up by the observation that advances in acquiring L2 phonology do not go hand in hand with perceptible decrements in L1 speech production ability (e.g. Fowler et al. 2008).[13]

Finally, we attempted to investigate whether the development of (non)native-likeness in late bilinguals is modulated by language aptitude. Such effects have been reported before for late L2ers (e.g. DeKeyser, 2000; Hyltenstam & Abrahamsson, 2009) and prepubescent attriters (Bylund et al., 2010), but they have not yet been investigated in the context of late L1 attrition.

Since no direct measures of aptitude were collected, we resorted to using the scores from a C-test which had been administered to both bilingual populations in order to estimate individual aptitude. Although such global proficiency measures constitute an amalgam of factors, language aptitude has been shown to explain a large degree of variance in proficiency scores (see above). As the C-test scores were not related to any of the external factors measured in this study for either of the bilingual groups, it appears likely that they reflect individual differences in language aptitude to a large degree.

The general proficiency score measured by the C-test proved to correlate strongly with foreign accent in both the L1 attriter and the L2 group. Our finding that there is no effect of L1 use in the L1 attriter group, but a strong correlation between global proficiency scores and foreign accent ratings, is compatible with the assumption that language aptitude may have some protective function in L1 attrition. In other words, higher degrees of language aptitude might mitigate the adverse effects of cross-linguistic influence on L1 speech production even after prolonged periods of non-use.

Conversely, language aptitude would seem to serve a facilitatory function in late L2 acquisition. Interestingly, the total amount of use of the L2 does not correlate with the C-test score (r=.23), our indirect measure of aptitude. This lack of a correlation indicates that use and aptitude independently affect L2 speech production. In other words, the effects of increased use of the L2 on FAR are not mediated by language aptitude; rather, they each appear to impact on FAR independently. What this implies is that increased language use facilitates approximating native-like pronunciation in L2 acquisition across individual variation in terms of aptitude (see also Harley & Hart, 1997). However, the data in the present study cannot address the question of the relative influences of use versus aptitude. In particular, the extent to which above-average language aptitude is a prerequisite for attaining native-like accents cannot be answered in the present context, as the L2 participants in Hopp (2007) were selected on the basis of their advanced proficiency. They might thus represent a skewed sample with above-average language aptitude to start with. Summarizing, the indirect effects of language aptitude across the bilingual groups point to a protective function of aptitude in L1 attrition and a facilitatory function of aptitude in late L2 acquisition.

In general, the direct comparison between late L1 attriters and late L2ers undertaken in the present study magnifies the general impact of late bilingualism on speech production. More particularly, the findings highlight that, irrespective of whether the target language

investigated is the early-acquired L1 or a late-learned L2, speakers who become bilingual after puberty experience bilingualism effects in terms of cross-linguistic influence. The present study thus underlines the oft-made point that the endpoint of bilingualism does not amount to additive monolingualism, whether in L1 or in L2 (Cook, 2003; Grosjean, 1998).

At first glance, the findings from this study do not allow us to reject either Hypothesis 1 or Hypothesis 2; rather, it seems that their interpretation depends on the perspective taken on the data.

Looking at the results from the L1 attrition perspective, we find that L1 attriters do not differ in perceived foreign accent from the native speakers at the group level. Moreover, FARs in L1 attrition are not significantly modulated by external factors such as age of emigration, length of time in an L2 environment, L1/2 use and affiliation. The group findings thus bolster the observation that native speakers can retain their perceived native-speaker-status even after long periods of emigration. Seen in conjunction with the significant differences between L1 attriters and L2ers at the group level, the present study would then appear to support the view that there is an effect of AOA on bilingual speech production.

Looking at the results from the L2 acquisition perspective, we find considerable overlap of the L2ers and the L1 attriters, with only 20% of all L2ers scoring outside the range of native speakers of German. In addition, the FARs in the L2ers are significantly correlated with the levels of use of the L2 and the degree of entrenchment of the L1/2 as measured by the time spent in the TL enviroment. On the basis of these similarities between L1 attriters and L2ers, the present findings would then emphasize the strong bilingualism effects on speech production and the influence of factors other than AOA.

Ultimately, however, we would argue that the interpretation of the findings in the present study is not just a matter of which perspective is preferred. Given the non-trivial

consequences of postulating a critical period in L2 acquisition for our understanding of the neuro-cognitive architecture and the mental processing of language, proponents of a critical period in L2 acquisition have to demonstrate that there is a strong contribution of maturational constraints in L2 development. Experimentally, this translates into the requirement that AOA needs to be shown to exert a strong and (largely) independent predictive role for convergence in bilingualism (e.g. Birdsong, 1999). This entails that late L2ers pattern outside the performance range delimited by native speakers with AOAs at birth. The present study can identify no such predictive role of AOA on perceived foreign accent. Rather, the facts (a) that the bilingual groups overlap to a large extent and (b) that foreign accent in the L2 is correlated significantly with external variables indicate that foreign accent in bilingualism cannot be predominantly related to AOA.

Taken together, our results are compatible with interference models of bilingual speech production (e.g. Flege et al. 2006) and, more broadly, continuity models of L2 acquisition (e.g. Bialystok, 2001; Hopp, 2007; Schwartz & Sprouse, 1996). Detailed differences between the models aside, they propose that cross-linguistic interference constrains L2 acquisition initially but subsides after sustained and sufficient L2 input which leads to the restructuring of the L2 system towards the target language. In these respects, these models can accommodate the finding that cross-linguistic interference in late bilingualism can impact similarly on both first-learned and later-learned languages.

At any rate, our approach of juxtaposing late L2 acquisition and late L1 attrition has wide-ranging methodological implications. A comparison of different bilingual populations introduces a different frame of reference for studies on L2 ultimate attainment (see also Montrul 2008; Tsimpli et al., 2004). Late L1 attriters constitute a control group that matches late L2 learners in terms of the asymmetric onset of bilingualism and a concomitantly lower proportion of input and use of the target language. Directly comparing and contrasting late L2

learners and L1 attriters thus allows us to disentangle effects of AOA from effects which affect bilinguals independently of the age of onset of the target language. We believe that this approach constitutes a methodological advance over the traditional comparisons of late L2 learners and (predominantly) monolingual native speakers. It seems to us that the direct comparison of late L2 acquisition and L1 attrition provides a fruitful line of inquiry which promises to cast new light on questions of age effects and cross-linguistic interference in bilingualism.

Given that late bilingualism has been demonstrated to lead to non-nativelikeness even in the mother tongue, the findings of the present study highlight that the (monolingual) standard of native-likeness in research on L2 acquisition is methodologically unsound. Hence, as far as investigating age effects in (late) L2 acquisition goes, the measure that has been used in virtually all previous research would seem to be unsuitable. In view of this methodological shortcoming, we believe it is also necessary to re-examine and re-evaluate previous research on L2 acquisition that is based on comparisons of L2ers and monolingual native speakers.

Needless to say, the present study is but a first step towards direct comparisons of late L1 attrition and L2 acquisition. Since the data from L1 attriters and L2 learners were collected in two independent studies, we could not ensure direct comparability of, e.g., C-test scores, speech production samples and the range of background information available. In addition, it would be desirable to collect aptitude measures in a direct comparison between late L1 attriters and late L2 learners in future research, and to extend such investigations to encompass a wider range of linguistic features beyond foreign accent.for instance aspects of morphosyntax such as inflection that Bearing these limitations in mind, we hold that the findings nevertheless go some way in illustrating that non-nativelikeness in bilinguals may not primarily be the consequence of late AOAs in L2 acquisition.

Conclusion

This study undertook a direct comparison of perceived foreign accent in late L1 attrition of German and late L2 acquisition of German. We introduced the method of directly comparing L1 attriters and L2 learners to predominantly monolingual natives as a way of disentangling effects of age of onset from effects of bilingualism in speech production. Although the natives and the L1 attriters on the one hand, differ from the L2 learners on the other in perceived foreign accent at the group level, we also found a sizeable overlap between L1 attriters and L2 learners. To explain variation in foreign accent, neither age of onset nor length of residence showed significant correlations. Of other external variables, use of the L2 turned out to be significant to some degree for the L2ers, whereas time spent in a German-speaking environment and an indirect measure of aptitude correlated with foreign accent for all bilingual groups. In line with interference accounts of bilingual speech production, these results illustrate that speech production can perceptibly be affected by cross-linguistic interference in both groups of bilinguals despite the fact that the attriters acquired the target language as their native language from birth, while the L2ers were postpuberty foreign-language learners. We conclude that acquiring a language from birth is not sufficient for ensuring nativelikeness in bilingual speech production. In consequence, nativelikeness - if defined against a predominantly monolingual standard - cannot serve as a performance criterion in investigations of age effects on L2 ultimate attainment.

Acknowledgements

Part of the research presented here was supported by NWO grant 275-70-00. We are grateful to Lysbeth Plas, Bregtje Seton, Dieter Thoma and Gulsen Yilmaz for their help at various stages.

Notes

1 L1/2 use is collapsed as one variable, since the amount of L2 use is linearly dependent on the amount of L1 use and vice versa.

2 In a similar experiment conducted by Caramazza, Yeni-Komshian, Zurif and Carbone (1973) on VOTs, early French Canadian–English bilinguals show evidence only of L1 influence on L2, not vice versa. This might be taken to confirm Flege's assumption that children are able to assign similar sounding phones to new categories, while adults tend to classify and interpret them as instantiations of familiar categories: the early learners investigated by Caramazza et al. may have created different categories for French and English plosives, but the category in their L2 may have been influenced by their L1.

3 For the purpose of the present discussion, we will focus on L2 influence on L1 in late bilinguals, and not consider treatments of second generation speakers or early bilinguals such as El Aissati (1997), Hirvonen (1995) or Vago (1991).

4 Note, however, that Major finds bidirectional interference in casual style even for the speaker whose VOTs are perfectly nativelike in formal style.

5 The findings reviewed here illustrate that proficient bilinguals may experience bi-directional interference on some consonants (plosives and liquids) and on suprasegmentals. However, it remains unknown to what extent similar processes may affect other consonants or vowels. In fact, in Flege's (1987) study, a vowel similarly realized in the two languages (/u/) remained unaffected by the bilingual experience (but see Flege, Bohn, & Jang, 1997). The overall range of phonetic changes effected in the L1 of bilinguals and the question of

whether some phones may be less susceptible to such interferences can therefore only be speculated on at present.

6 Schmid (2002) also reports on a foreign accent rating study of a group of attriters. However, since her investigation merely serves to distinguish degrees of foreign accent within the attriting population and does not compare the ratings against a non-attrited control group her findings will not be discussed here.

7 It should be noted that these migrants had not been pre-selected in any way on the basis of proficiency in either L1 or L2, amount of L1 use, etc.

8 In most studies the raters are unaware which of the speakers are bilingual, but at least one study interspersed a sample from the same native speaker between all bilingual samples in order to provide an explicit reference point (Anderson-Hsieh, Johnson, & Koehler, 1992, cited after Jesney, 2004).

9 The absence of an LOR effect in Oyama's (1976) investigation was interpreted by Krashen, Long, and Scarcella (1979) to be due to the confounding effects of rate of acquisition (which initially favours older learners) and ultimate attainment (which typically favours younger learners in the long term).

10 The label INTERMEDIATE refers to Grosjean's (2001) model of language modes in bilinguals. According to Grosjean, bilingual speech production can be situated on a continuum from fully bilingual mode (where both languages are active and contribute to processing and speech production) to largely monolingual mode (where one language is

almost entirely inactive). Intermediate mode situations are contexts where external stimuli (the environmental language, the knowledge that other interlocutors are also bilingual, etc.) trigger the activation of the language not chosen for communication, but where the speaker has to suppress/inhibit that language.

11 The ratings obtained by Hopp (2007) were converted to the scale as used by DSM to allow for direct comparisons.

12 In fact, one L1Aer with 48 years' residence in the Netherlands received a FAR of 1.40, while an L2er was rated at 1.51 after as little as 6 years of residence in Germany. Conversely, an attriter after 17 years of residence in the Netherlands was rated to be clearly non-native with an average FAR of 5.46, and an L2er was judged to be clearly non-native after 30 years of residence, with a FAR of 5.93.

13 It would be interesting to have speech production data from the same participants in their other language, i.e. the L2s for the L1Aers, and the L1s for the L2Aers, to directly test for an asymmetry in deviance from the native norm in L1 and L2. Unfortunately, such data were not available.

References

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

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

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of *nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. Language Learning,* 42, 529-555.

Asher, J. J., & García, R. (1969). The optimal age to learn a foreign language. *The Modern Language Journal,* 53, 334-341.

Au, T. K., Knightly, L. M., Jun, S. A., & Oh, J. S. (2002). Overhearing a language during childhood. *Psychological Science,* 13, 238–243.

Bialystok, E. (1997). The structure of age: in search of barriers to second language acquisition. *Second Language Research,* 13, 116-137.

Bialystok, E. (2001). *Bilingualism in Development: Language, Literacy, & Cognition.* New York: Cambridge University Press.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition,* 12, 3-11.

Birdsong, D. (1999a). Introduction: Whys and why nots of the critical period hypothesis for second language acquisition. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis.* Mahwah, NJ: Erlbaum.

Birdsong, D. (2005). Interpreting age effects in second language acquisition. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches*. New York: Oxford University Press.

Bongaerts, T. (1999). Ultimate attainment in L2 pronunciation: The case of very advanced L2 learners. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis*. Mahwah, NJ: Erlbaum.

Bongaerts, T., van Summeren, C., Planken, B., &, Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition,* 19, 447-465.

Bot, K. de, & Clyne, M. (1994). A 16 year longitudinal study of language attrition in Dutch immigrants in Australia. *Journal of Multilingual and Multicultural Development,* 15, 17-28.

Brennan, E. M., & Brennan, J. S. (1981). Accent scaling and language attitudes: reactions to Mexican American English speech. *Language and Speech,* 24, 207-221.

Bylund, E., Abrahamsson, N., & Hyltenstam, K. (2010). The role of language aptitude in first language attrition: the case of pre-pubescent attriters. *Applied Linguistics,* 1–22.

Caramazza, A., Yeni-Komshian, G., Zurif, E., & Carbone, E. (1973). The acquisition of a new phonological contrast: the case of stop consonants in French–English bilinguals. *Journal of the Acoustic Society of America,* 4, 421–428.

Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude.* Rowley, MA: Newbury House.

Cook, V. (2003). The changing L1 in the L2 user's mind. In V. Cook (Ed.), *Effects of the second language on the first.* Clevedon: Multilingual Matters.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1989). Limits on bilingualism. *Nature*, 340, 229-230.

Leeuw, E. de. (2008). *When your native language sounds foreign: A phonetic investigation into first language attrition.* Unpublished doctoral dissertation, Queen Margaret University, Edinburgh.

Leeuw, E. de, Schmid, M. S., & Mennen, I. (2010). The effects of contact on native language pronunciation in a migrant context. *Bilingualism: Language and Cognition,* 13, 33-40.

DeKeyser, R., & Larson-Hall, J. (2005). What does the critical period really mean? In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches*. New York: Oxford University Press.

DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition,* 22, 499-533.

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition*. Oxford: Blackwell.

Dussias, Paola E. (2004). Parsing a first language like a second: The erosion of L1 parsing strategies in Spanish-English bilinguals. *International Journal of Bilingualism,* 8, 355-371.

El Aissati, A. (1997). Language Loss Among Native Speakers of Morrocan [sic] Arabic in the Netherlands. Tilburg: Tilburg University Press.

Elliott, A. R. (1995). Field independence/dependence, hemispheric specialization, and attitude in relation to pronunciation accuracy in Spanish as a foreign language. *The Modern Language Journal,* 79, 356-371.

Flege, J. E. (1984). The detection of French accent by American listeners. *Journal of Acoustical Society of America,* 76, 692-707.

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, 47-65.

Flege, J. E. (1999). Age of learning and second-language speech. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis*. Mahwah NJ: Lawrence Erlbaum.

Flege, J. E. (2002). Interactions between the native and second-language phonetic systems. In Burmeister, P., Piske, T., & Rohde, A. (Eds.), *An integrated view of language development: Papers in honor of Henning Wode*. Trier: Wissenschaftlicher Verlag.

Flege, J. E. (2009). Give input a chance! In Piske, T. and Young-Scholten, M. (Eds.), *Input matters in SLA*. Multilingual Matters. Bristol.

Flege, J. E., & Hillenbrand, J. (1984). Limits on phonetic accuracy in foreign language speech production. *Journal of the Acoustic Society of America,* 76, 708-721.

Flege, J. E., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Tsukada, K. (2006). Degree of foreign accent in English sentences produced by Korean children and adults. *Journal of Phonetics,* 34, 153-75.

Flege, J. E., Bohn, O.-S., & Jang, S. (1997) Effects of experience on non-native speakersÕ production and perception of English vowels. *Journal of Phonetics,* 25, 437-470.

Flege, J. E., & Fletcher, K. L. (1992). Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America,* 9, 370-389.

Flege, J. E. & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition,* 23, 527-552

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America,* 97, 3125-3134.

Flege, J. El, Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second- language acquisition. *Journal of Memory and Language,* 41, 78-104.

Foursha, C. D., Austin, J. B., & Van de Walle, G. (2005). Is language processing identical in monolinguals and early, balanced bilinguals? Cascadilla Proceedings Project. Available: http://www.bu.edu/linguistics/APPLIED/BUCLD/proc.html.

Fowler, C. A., Sramko, V., Ostry, D. J., Rowland, S. A., & Hallé, P. (2008). Cross language phonetic influences on the speech of French–English bilinguals. *Journal of Phonetics,* 36, 649–663.

Giesbers, H. (1997): Dutch in Indonesia: Language attrition or language contact. In J. Klatter-Folmer & S. Kroon (Eds.), *Dutch overseas: Studies in maintenance and loss of Dutch as an immigrant language.* Tilburg: Tilburg University Press.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition,* 1, 131–149.

Grosjean, F. (2001). The bilingual's language modes. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing.* Oxford: Blackwell.

Harley, B., & Hart, D. (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. Studies *in Second Language Acquisition,* 1, 379-400.

Hernandez, A. E., Bates, E. A., & Avila, L. X. (1994). On-line sentence interpretation in Spanish-English bilinguals: What does it mean to be 'in-between'? *Applied Psycholinguistics,* 15, 417-446.

Hirvonen, P. A. (1995). Phonological and morphological aspects of Finnish language attrition in the United States. In W. Viereck (Ed.), *Soziolinguistische Variation.* Stuttgart: Steiner.

Hopp, H. (2007). *Ultimate attainment at the interfaces in second language acquisition: Grammar and processing*. Unpublished doctoral dissertation, University of Groningen.

Hopp H. (2010) Ultimate attainment in L2 inflectional morphology: Performance similarities between non-native and native speakers. *Lingua,* 120, 901-931.

Hyltenstam, K., & Abrahamsson, N. (2003). Who can become native-like in a second language? All, some or none? On the maturational constraint controversy in second language acquisition. *Studia Linguistica,* 54, 150-166.

Hyltenstam, K., Bylund, E., Abrahamsson, N., & Park, H-S. (2009). Dominant language replacement; The case of international adoptees. Bilingualism: Language and Cognition, 12, 121-140.

Jesney, K. (2004). The use of global foreign accent rating in studies of L2 acquisition. MS, Language Research Centre University of Calgary.

Jia, G., Aaronson, D., & Wu, Y. (2002). Long-term language attainment of bilingual immigrants: Predictive variables and language groups differences. *Applied Psycholinguistics,* 23, 599-621.

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.* Amsterdam: John Benjamins.

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

Long, M. H. (1993). Second language acquisition as a function of age: Research findings and methodological issues. In K. Hyltenstam and A. Viberg (Eds), *Progression and regression in language.* Cambridge University Press.

Long, Michael H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition,* 12, 251-285.

MacKay, I. R. A., Flege, J. E., & Imai, S. (2006). Evaluating the effects of chronological age and sentence duration on degree of perceived foreign accent. *Applied Psycholinguistics,* 27, 157-183.

Major, R. C. (1992). Losing English as a first language. *The Modern Language Journal,* 76, 190-208.

McDermott, W. L. C. (1986). *The scalability of degrees of foreign accent*. Unpublished doctoral dissertation, Cornell University, Ithaca, NY.

McDonald, J. L. (2000). Grammaticality judgments in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics,* 21, 395-423.

Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics,* 32, 543–563.

Montrul, S. A. (2008). *Incomplete acquisition in bilingualism*. Amsterdam/Philadelphia: John Benjamins.

Moyer, A. (1999). Ultimate attainment in L2 phonology: the critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition,* 21, 81-108.

Moyer, A. (2007). Empirical considerations on the age factor in L2 phonology. *Issues in Applied Linguistics,* 15, 109-129.

Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech,* 38, 289-306.

Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research,* 5, 261-283.

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

Perdue, C. (1993). *Adult language acquisition: Cross-linguistic perspectives*. Cambridge: Cambridge University Press.

Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics,* 29, 191-215.

Proverbio, A. M., Cok, B. & Zani, A. (2002). Electrophysiological measures of language processing in bilinguals. *Journal of Cognitive Neuroscience,* 14, 994-1017.

Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics,* 25, 235-268.

Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics,* 25, 421-436.

Schmid, M. S. (2002). *First language attrition, use and maintenance: The case of German jews in Anglophone countries.* Amsterdam: John Benjamins.

Schmid, M. S. (2007). The role of L1 use for L1 attrition. In B. Köpke, M. S. Schmid, M. Keijzer & S. Dostert (Eds), *Language attrition. Theoretical perspectives.* Amsterdam/Philadelphia: John Benjamins.

Schmid, M. S. (2009). On L1 attrition and the linguistic system. *EUROSLA Yearbook,* 9, 212-244.

Schmid, M. S. (2010). Languages at play: The relevance of L1 attrition to the study of bilingualism. *Bilingualism: Language and Cognition,* 13, 1-7.

Schmid, M. S. (forthc.). *Language Attrition*. Cambridge University Press.

Schmid, M. S, & B. Köpke. (2008). L1 attrition and the mental lexicon, in: Aneta Pavlenko (ed.), *The bilingual mental lexicon*. Clevedon: Multilingual Matters.

Schmid, M. S. and Dusseldorp, E. (2010). Quantitative analyses in a multivariate study of language attrition. *Second Language Research,* 26, 125-160.

Schumann, J. H. (1998): The neurobiology of affect in language. *Language Learning,* 48: Supplement 1.

Schwartz, B. D., & Sprouse, R. A.. (1996). L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research,* 12, 40-72.

Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech.* Cambridge, MA: Newbury House.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching,* 10, 209–232.

Sundaraa, M., Polkaa, L., & Genesee, F. (2006). Language-experience facilitates discrimination of /d-D/ in monolingual and bilingual acquisition of English. *Cognition*, 100, 369-388.

Thompson, Irene. (1991). Foreign accents revisited: The English pronunciation of Russian immigrants. *Language Learning,* 41, 177-204.

Tsimpli, I., Sorace, A., Heycock, C., & Filiaci, F. (2004). First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English. *International Journal of Bilingualism,* 8, 257-277.

Vago, R. M. (1991). Paradigmatic regularity in first language attrition. In H. W. Seliger & R. M. Vago (Eds.), *First language attrition.* Cambridge: CUP.

Van Hell, J., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review,* 9, 780-789.

Ventureyra, V. A. G. (2005). *A la recherche de la langue perdue: etude psycholinguistique de l'attrition de la premiere langue chez des Coreens adoptes en France* [In search of the lost language: a psycholinguistic study of first language attrition among Korean adoptees in France]. Unpublished doctoral dissertation, Ecole des Hautes Etudes en Sciences Sociales/Laboratoire de Sciences Cognitives et Psycholinguistique, Paris.

Ventureyra, V., Pallier, C., & Yoo, H.-Y. (2004). The loss of first language phonetic perception in adopted Koreans. *Journal of Neurolinguistics,* 17, 79-91.

Werker, J. F., & Byers-Heinlein, K. (2008). Bilingualism in infancy: first steps in perception and comprehension. *Trends in Cognitive Sciences,* 12, 144-151

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

Tables

Table 1: Group characteristics

	L1Aers	(n = 40)		L2ers	(n = 40)		NSS	(n = 20)	
	mean	stdev	range	mean	stdev	range	mean	stdev	range
Age	61.88	10.24	37-85	41.15	9.79	24-60	53.25	13.73	24-78
AOA	27.88	8.67	17-51	15.38	4.45	11-27			
LOR	33.98	12.63	16-57	15.88	8.67	3-36			

Table 2: Categorical ratings per population

	CG	L1A	L2er
native	14 (70%)	23 (57.5%)	11 (27.5%)
uncertain	6 (30%)	11 (27.5%)	12 (30%)
nonnative	0	6 (15%)	17 (42.5%)

Table 3: Pearson correlations of self-reported use of and attitude towards German and FAR

	L1Aers			L2ers		
	r	p	n	r	p	n
CHILDREN	.040	.824	34	. 111	. 662	18
PARTNER	.078	.636	39	-.538**	.002	30
FRIENDS	.098	.546	40	-.291	. 112	31
INTERMEDIATE	.227	.160	40	-.179	. 363	28
PASSIVE	-.104	.525	40	. 040	.830	31
WORK	-.089	.585	40	-.058	.772	31
AFFILIATION	-.264	.100	40	-.229	..215	31
TOTAL	.105	.519	40	-.424*	.017	31

*: $p < .05$, **: $p < .01$

Table 4: FAR and native language of partner

		German partner		Non-German partner	
		Count	Exp. Count	Count	Exp. Count
L1Aers	Native	12	11.8	10	10.2
	Uncertain	6	5.9	5	5.1
	Non-native	3	3.2	3	2.8
L2ers	Native	9	7.1	1	2.9
	Uncertain	5	5.7	3	2.3
	Non-native	7	7.1	4	2.9

Table 5: Overview of bivariate Pearson Correlations between FAR and independent variables

	FAR	
	L1A	L2ers
AOA	-.037	.254
LOR	-.039	.019
USE	.105	-.424*
C-test score (Aptitude)	-.435**	-.472**

*: $p < .05$, **: $p < .01$

Figure Captions

Figure 1. Rationale of group comparisons.

Figure 2: Proportion of use of and affiliation with German across experimental populations. A score of 1 represents an overwhelming preference for German, while 0 indicates that Dutch/English is used/preferred almost exclusively in the relevant context.

Figure 3: FARs across populations

Figure 4: FARs across populations and contact languages

Figure 5: Individual FAR scores per group and contact language

Figure 6: AOA and FAR across bilingual populations

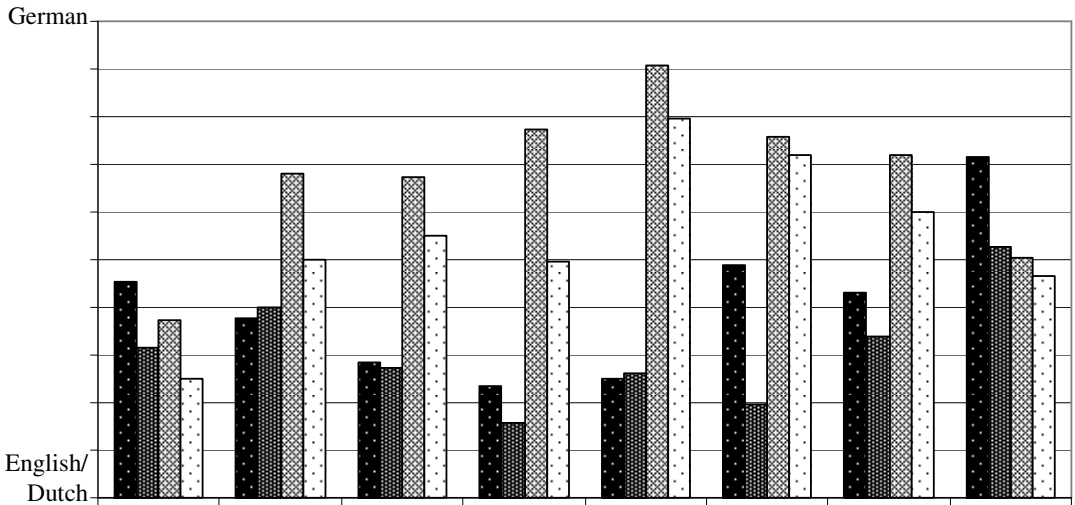
Figure 6a: AOA and FAR for L1Aers (left) and L2ers (right)

Figure 7: LOR and FARs across bilingual populations

Figures

	Natives	L1 attriters	L2 learners
Native speakers (AOA = 0)	✓	✓	✗
Late bilingualism	✗	✓	✓

Figure 1. Rationale of group comparisons.



	CHILDREN	PARTNER	FRIENDS	INTERM	WORK	PASSIVE	TOTAL	AFFIL
L1ANL	0.45	0.38	0.28	0.24	0.25	0.49	0.43	0.71
L1AEN	0.32	0.40	0.27	0.16	0.26	0.20	0.34	0.53
L2NL	0.38	0.68	0.67	0.77	0.91	0.76	0.72	0.50
L2EN	0.25	0.50	0.55	0.50	0.80	0.72	0.60	0.47

Figure 2: Proportion of use of and affiliation with German across experimental populations. A score of 1 represents an overwhelming preference for German, while 0 indicates that Dutch/English is used/preferred almost exclusively in the relevant context.

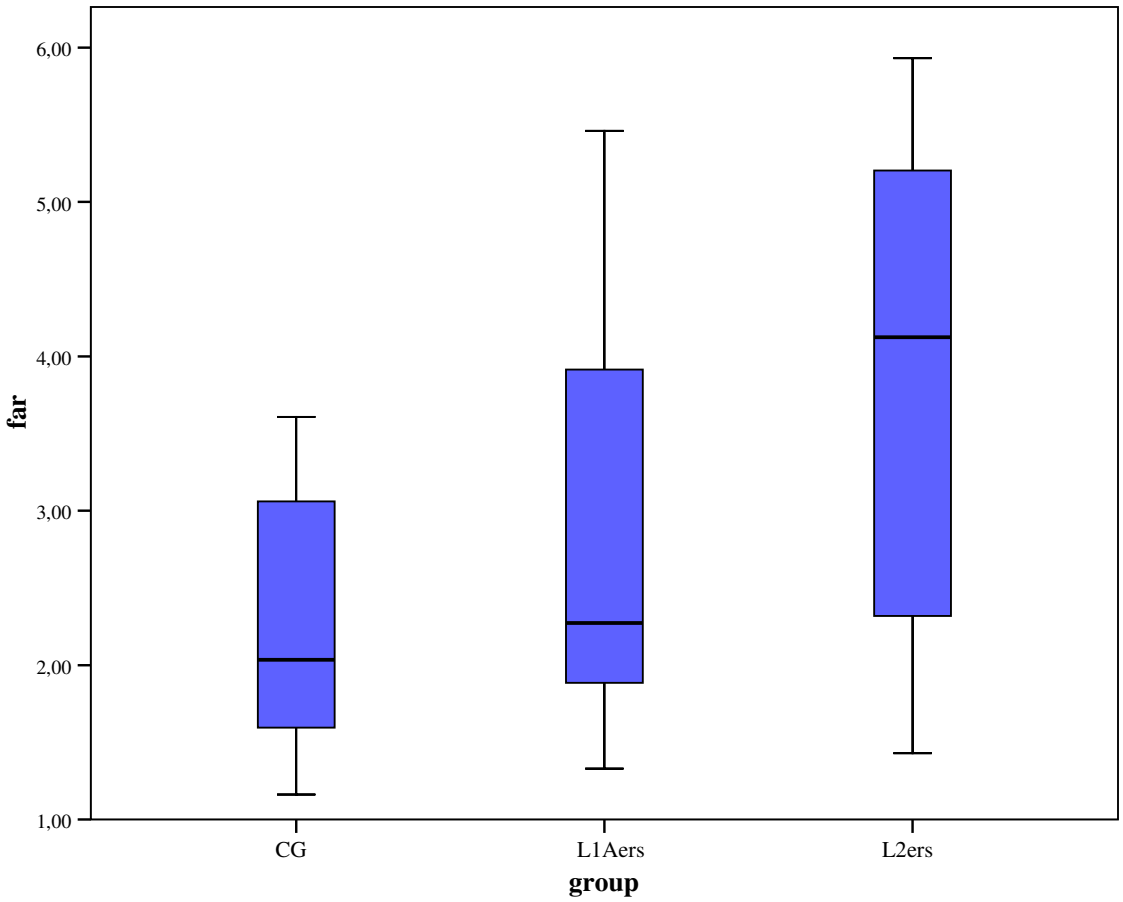


Figure 3: FARs across populations

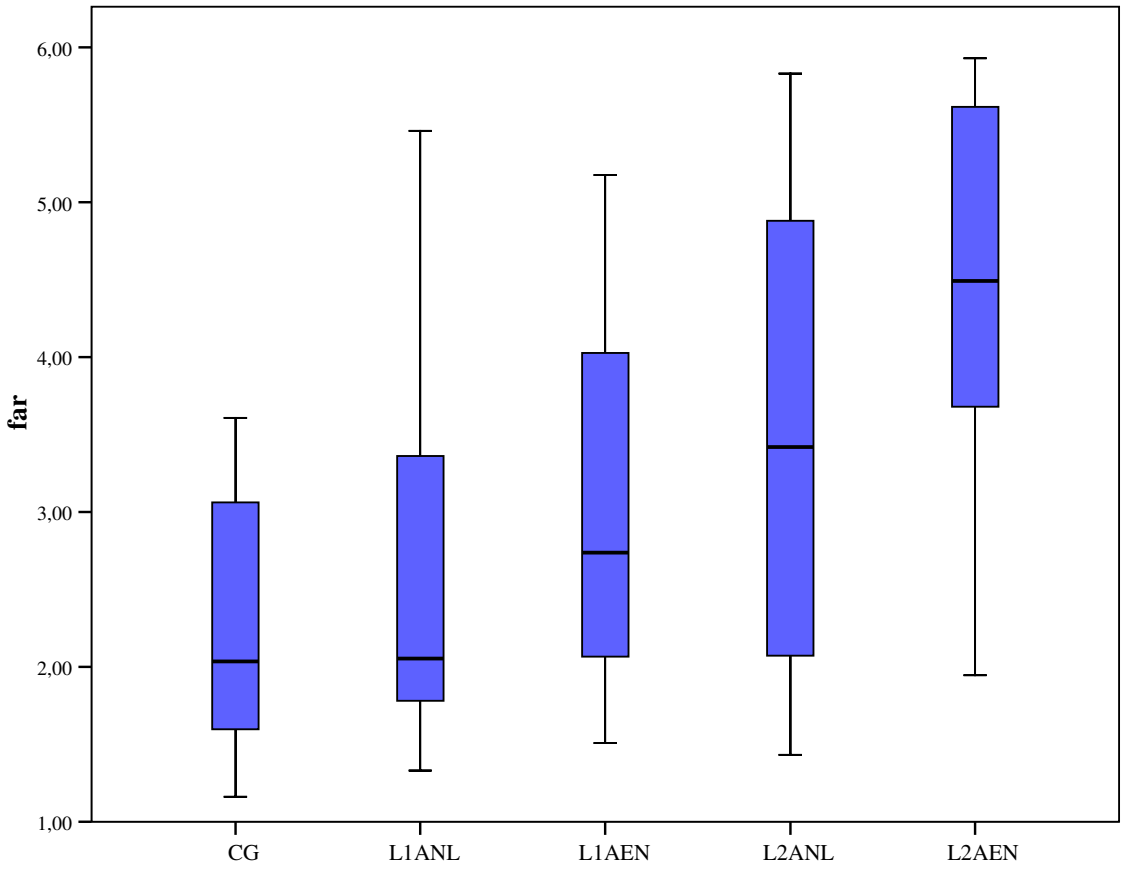


Figure 4: FARs across populations and contact languages

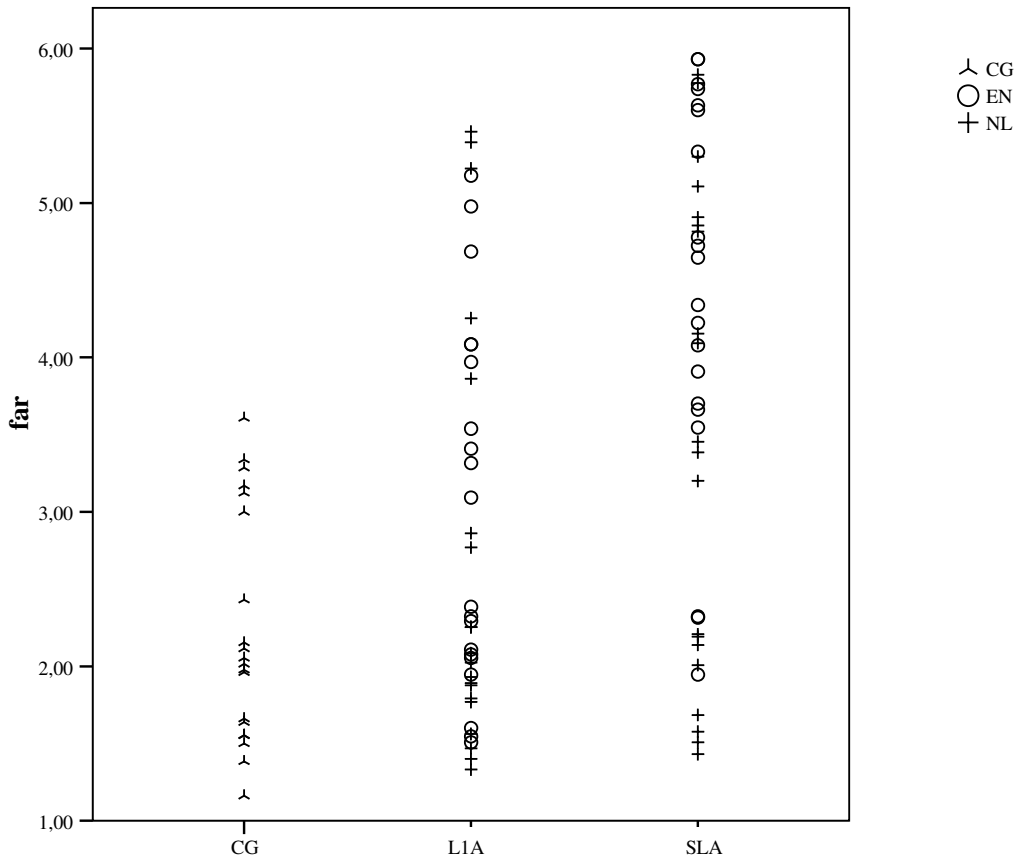


Figure 5: Individual FAR scores per group and contact language

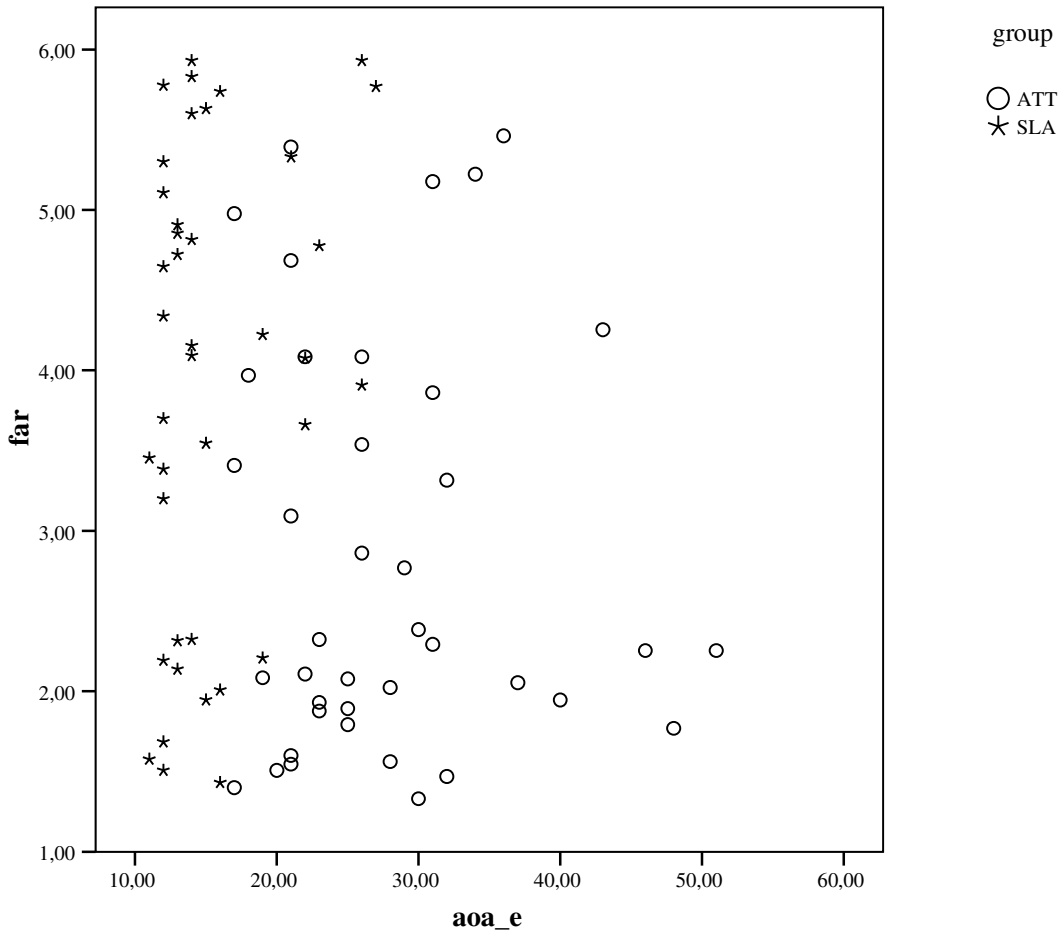


Figure 6: AOA and FAR across bilingual populations

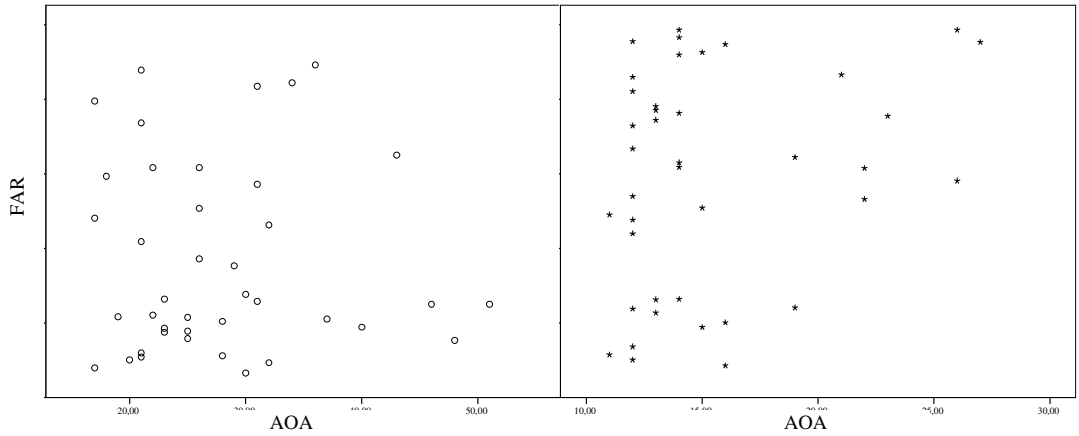


Figure 6a: AOA and FAR for L1Aers (left) and L2ers (right)

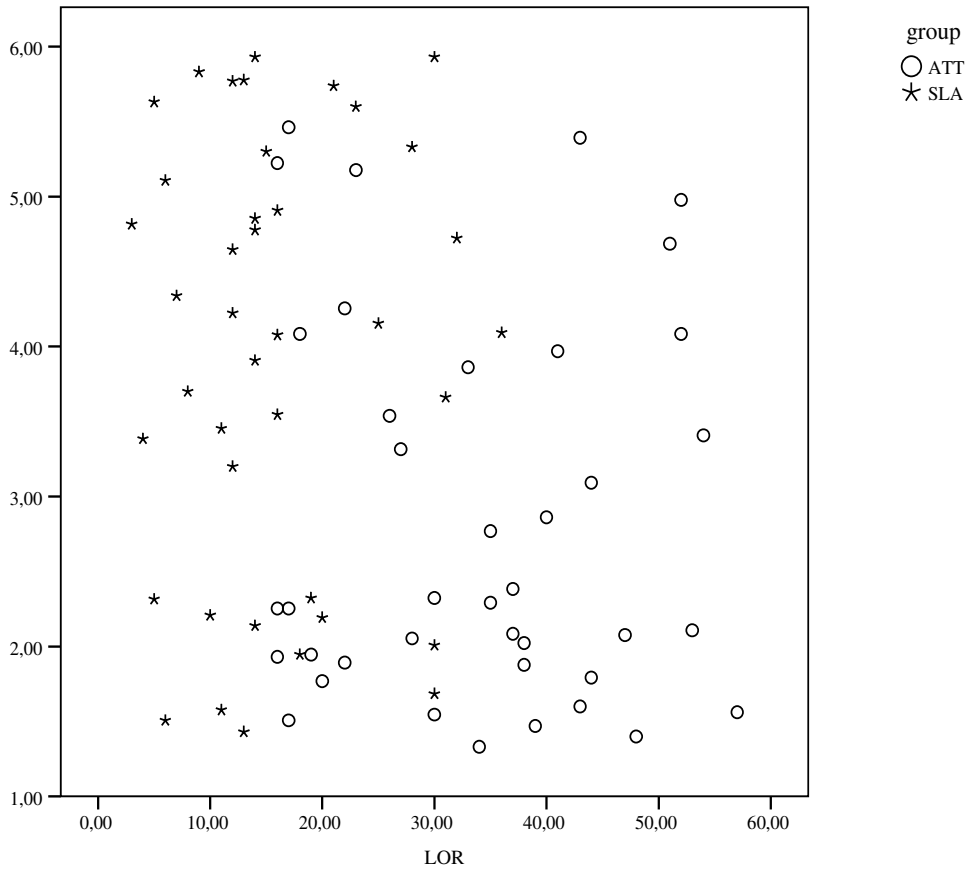


Figure 7: LOR and FARs across bilingual populations