L1 attrition as a function of age at onset of bilingualism: L1 attainment of Turkish-English bilinguals in the UK

running head: L1 attrition as a function of age at onset

Abstract

The current investigation aims to provide insights into the controversial debate on the nature of the role 'age at onset of bilingualism' (AaO) plays in human language capacity with a focus on what it entails for L1 attrition. L1 performance of a group of Turkish immigrants (n=57) in the UK with AaO range 7–34 was compared to that of monolingual controls (n=29) across two linguistic properties: structural complexity and L1 accent. Regarding L1 accent development, we propose AaO be taken as a proxy for L1 entrenchment instead of the maturational state of the speaker. In the case of structural complexity, full retention of proficiency prevents us from establishing a relationship with AaO. We suggest that attrition data needs to be better accommodated within such theoretical age accounts by emphasizing that not all areas of linguistic competence are affected by AaO and by detailing the underlying factors in such cases.

keywords: L1 attrition, age at onset of bilingualism, critical period, L1 entrenchment, Turkish immigrants

Tuğba Karayayla

Department of Language and Linguistics, University of Essex, Wivenhoe Park CO4 3SQ *Colchester, Essex United Kingdom*

+44 (0) 1206 873333, tkaray@essex.ac.uk

Monika S. Schmid

Department of Language and Linguistics, University of Essex, Wivenhoe Park CO4 3SQ *Colchester, Essex United Kingdom*

+44 (0) 1206 873333, mschmid@essex.ac.uk

Author's note

We would like to thank Dr Kevin Savage for his support in the data extraction and quantification.

Background to the study

The non-pathological deterioration of the previously acquired native language, i.e. first language (L1) attrition (Köpke & Schmid, 2004) is strongly influenced by the age at which the speaker becomes bilingual, and the impact of this factor appears to be quite pronounced in both production (Bylund, 2009a; Montrul, 2008) and perception (Ahn, Chang, DeKeyser, & Lee-Ellis, 2017) across various linguistic levels. Research on native language change is usually carried out in one of two settings, with little overlap: language development among heritage speakers (HSs) on the one hand and among late bilinguals (LBs, speakers who left their native language environment post-puberty, usually in early adulthood) on the other. There is a notable dearth of studies attempting to fill the gap in the age at onset (AaO) between adult HSs (AaO usually between 0–6 years) and LBs (AaO>12), as well as investigations directly comparing the L1 development between adult HSs and LBs (e.g. Montrul & Sánchez-Walker, 2013).

In heritage language (HL) development there are two important observations. First, HSs often show much greater variability in their use of the HL. While some of them typically score within the range of monolingual speakers (and of native speakers who learned another language later in life), others show accuracy levels below chance, even on features which monolingual children master before the age at which these speakers have become bilingual (Cuza & Pérez-Tattam, 2016). This non-target-like performance is usually observed in the form of simplifications, and reductions especially in morphosyntactic categories such as inflectional morphology (Montrul, 2016) as well as complex syntactic phenomena or properties at the interfaces (e.g. Montrul, 2004; Treffers-Daller, Ozsoy, & van Hout, 2007). Both judgments and the oral performance of Hindi HSs, for example, have been found to be eroded regarding the case system in Hindi in comparison to the performance of late bilinguals in the US (the first generation) included as a control group (Montrul, Bhatt, & Bhatia, 2012).

Second, the level of L1 proficiency HSs may eventually retain has proven extremely difficult to account for. Some HSs ultimately become native-like while others end up with rudimentary skills, and this ultimate success might relate to a combination of factors, such as quantity/quality of input and heritage language instruction (Kupisch & Rothman, 2016; Montrul, 2016; Rothman, 2009). Studies tracing HSs' L1 development longitudinally provide further evidence for the rate and degree of structural erosion experienced by these speakers which can be quite severe (especially in the case of international adoption) even if the

property under investigation had stabilized before the immigration took place (Altenberg, 1991; Isurin, 2000; Montrul, 2008; Schmitt, 2004; Zaretsky & Bar-Shalom, 2007).

Zaretsky and Bar-Shalom (2007), for instance, tested whether AaO and frequency of L1 use would prevent morphosyntactic categories in L1 Russian from attrition in children and adults. They investigated error rates in the narrative and grammaticality judgment task (GJT) performance of ten Russian-English children (aged 4–13) with AaO range 0–6 and ten adults (aged 19–53) with AaO range 4–37. While older AaOs predicted fewer errors in both groups, L1 use did so only in children. Yet, in the case of a Russian-American child whose L1 exposure was reduced considerably after she was adopted at the age of nine, the attrition set in very quickly in the first year of adoption and the participant started to refuse to communicate in the L1 (Isurin, 2000). There is strong evidence suggesting that by adulthood, this knowledge can be completely erased from the brain, indicating the paramount role played by continuous input in maintaining the previously acquired L1 proficiency (Pallier et al., 2003; but see Pierce, Chen, Delcenserie, Genesee, & Klein, 2015 for counterevidence obtained from a fMRI study showing long-lasting effects of early exposure on the maintenance of the perception of Chinese lexical tone contrasts among Chinese adoptees, despite being deprived of continuous L1 input).

It appears, however, that even in such severe cases of L1 loss or incomplete acquisition (as observed in child over-hearers), early exposure might be advantageous in relearning it in adulthood in comparison to second language (L2) learners. These advantages seem to be limited to phonological categories rather than structural ones and are evident in both perception (differentiation of minimal pairs) (Hyltenstam, Bylund, Abrahamsson, & Park, 2009) and production (VOT, pronunciation) (Au, Knightly, Jun, & Oh, 2002; but see Ventureyra, Pallier, & Yoo, 2004 for null results). Despite these long-lasting advantages over L2 learners, HSs are usually perceived as sounding less native-like in comparison to monolinguals (Kupisch et al., 2014), and the voice onset time (VOT) of certain sounds they produce are not always native-like (Lein, Kupisch, & Weijer, 2015).

One likely cause of the changes in the L1 VOT values has been suggested to be cross-linguistic influence (CLI), a process which might lead to the development of an accented L1 (Lein et al., 2015). There is also some evidence showing that the accented speech developed alongside that of an L2 from early on cannot be reversed even after having lived in the home

country for more than 8 years, and this seems to be predominantly determined by the post-puberty ages of return to the home country (Flores & Rato, 2016).

To summarize, findings from these studies conducted with HSs suggest a complex interaction of factors but do point to one straightforward conclusion: early exposure in itself does not constitute a sufficient criterion for becoming target-like in the L1, and neither does age-appropriate development up to puberty. L1 knowledge that was acquired before puberty is unstable and can regress when another language becomes more dominant.

Late bilinguals, on the other hand, differ from HSs in two respects when it comes to their native language. Firstly, AaO ceases to play a role around puberty. Most studies investigating attrition in LBs adopt a minimum threshold for AaO of 15–17 years (following recommendations by De Bot, Gommans, & Rossing, 1991), but even in populations with AaOs below this threshold, –but above or around the onset of puberty– no age effect has been found (Schmid, 2002).

Secondly, the L1 of LBs appears to be much more stable. While the scores attained by attriting populations on virtually any measure of L1 proficiency are almost invariably distributed over a wider range than those of monolingual control populations, with some attriters scoring lower than the worst-performing controls, this does not appear to indicate any systematic impairment to L1 knowledge: error rates usually remain below 5% on any grammatical structure (Montrul, 2008; Schmid, 2013). It has also been shown that such attrition effects in LBs can be reversed by L1 re-exposure through visits to the home country, indicating that attrition does not affect the underlying language system (Chamorro, Sorace, & Sturt, 2016). Unlike in HL development, it has been difficult to obtain a comprehensive picture of the predictors driving language attrition (e.g. de Leeuw, Schmid, & Mennen, 2010). Even in extreme cases of traumatic experiences accompanied with prolonged lack of continuous L1 contact (for around 50 years), an investigation carried out on the structural complexity and morphosyntactic accuracy in the L1 German of the Holocaust survivors with adolescent AaOs (11–17) showed no AaO effect or erosion exceeding the 5% error rate (Schmid, 2012).

Interestingly, however, phonological categories have been found to be more vulnerable to CLI even in late bilingualism (Bergmann, Nota, & Schmid, 2017; de Leeuw et al., 2010; Hopp & Schmid, 2013). Although in these studies the performance of L1 attriters was usually

not different from the controls at the group level, a good percentage of bilinguals (up to 40 %) remained well outside the control range. This suggests that the L1 may be susceptible to the effects of an L2 even in late bilingualism leading to bidirectional transfer. Although some studies showed that maintenance of the L1 accent might be linked to the effects of external factors, such as professional L1 use (de Leeuw et al., 2010) and linguistic aptitude (Hopp & Schmid, 2013), the extent of the contribution of these factors as well as of AaO to this performance is still unclear.

Echoing the above findings, a handful of studies that attempted to fill the AaO gap between adult HSs and LBs demonstrated that L1 proficiency-AaO slope shows a discontinuity around puberty. While participants with AaO over 10-12 were found to be indistinguishable from monolinguals in L1 pronunciation (Yeni-Komshian, Flege, & Liu, 2000), perception of L1 speech sounds (Ahn et al., 2017), general proficiency (Hakuta & D'andrea, 1992), verbal morphology (Silva-Corvalán, 1994), and conceptualization patterns of goal-oriented events (Bylund, 2009a), those with AaO below this range were reported to show more variability with majority of them performing outside the control range.

Taken together, it seems that there is a qualitative change in the stability of the L1 around puberty (age 12). This seems to lead to restructuring in the L1 of the speakers whose first exposure to L2 was before this age, leaving the performance of those with AaO past this age comparatively mildly affected. L1 phonological categories, however, seem to be flexible even beyond this age (e.g. Hopp & Schmid, 2013). Why this should be the case, what exactly happens around this age, to what extent the L1 remains flexible beyond it, and whether there are different outcomes of AaO across linguistic levels are a few of the most important questions for bilingualism research to address.

In a recent theoretical review, Schmid and Köpke (2017) provided a discussion about the relevance of L1 attrition research to theories of bilingualism. As they underline, the fact that the acquired linguistic knowledge might change alongside the development of another language in the brain is not usually acknowledged in theoretical models of bilingual development. Championing an integrated approach to bilingualism, the authors thus argue that "in order to fully understand the nature of bilingual development and to resolve important and fundamental questions about the human capacity for language learning, processing and use, we need to arrive at a better understanding of how the mechanisms that drive and constrain L2 acquisition may also affect already established linguistic knowledge,

both in the immediate and in the longer term" (Schmid & Köpke, 2017, p. 5). It follows from this that more L1 attrition studies need to be conducted looking at the impact of AaO which has been shown previously to be an important factor constraining L2 acquisition. This will allow us to see to what extent this phenomenon is accounted for by relevant previous theoretical approaches to bilingualism.

Theoretical background to age effects in bilingualism

The effect of AaO on language learning has long been a topic of controversial debate especially in second language acquisition (SLA) research. To date, this has been predominantly addressed by two competing accounts.

The first account relates to the maturational state of the learner. Both human and animal developmental behaviour is acknowledged to be sensitive to environmental stimuli, the timing of which is crucial for a proper physiological development (Bornstein, 1989). Early exposure is claimed to be a prerequisite for successful language development as well (Newport, 1990). Studies reporting ultimate attainment in a L2 as a function of AaO are thus often framed within the Critical Period Hypothesis (CPH) beside other maturational explanations (see DeKeyser, 2013; Kinsella & Singleton, 2014 for reviews). The traditional view of the CPH in SLA research predicts loss of capacity for language acquisition past a critical period (usually around puberty) due to gradual maturation in the neural substrates responsible for language learning. This loss of plasticity, in turn, prevents post-puberty learners from attaining native-like proficiency in the L2 (DeKeyser, 2013; Kinsella & Singleton, 2014).

One of the very strong arguments against maturational age effects is the L1 entrenchment view. Proponents of this approach argue that the consistently observed AaO effect that the previous literature reported on L2 learning is not necessarily an indication of an irreversible biological constraint but is a disguised form of the entrenchment of the L1 which causes stabilization in the neural substrates and provides a filter to L2 learning (Pallier, 2007; Pallier et al., 2003). This view thus holds that stronger L1 links (due to increased proficiency with age) entail less strong L2 representations, indicating an inverse relationship between the L1 and L2 proficiency (Yeni-Komshian et al., 2000). As entailed by this inverse relationship, native-like success in the L2 can only be achieved as a result of losing the L1 completely (Pallier, 2007). Studies conducted with adult Korean adoptees to France with the AaO range

3–10 evidenced that if L1 exposure stops completely up to age 10, it is possible to reverse its filtering effects on the L2, and the L2 can override the L1 as a result of a complete reset of the neural substrates (Pallier et al., 2003; Ventureyra et al., 2004 but see Norrman & Bylund, 2016; Schmid, 2012).

Closely related to this, the main L2 performance differences between early and late bilinguals are suggested to be natural consequences of age-related factors rather than of irreversible maturational ones (see Muñoz & Singleton, 2011 for a review). Early learners, for example, might lack a strong sense of the L1 being a component of their identity and thus might show intrinsic motivations towards adapting to the L2 environment more quickly (Köpke, 2007). Furthermore, being enrolled in compulsory education inevitably makes their L2 environment richer than that of the LBs (Jia & Aaronson, 2003). This, in turn, might result in using the L1 less and the L2 more leading to a shift in language dominance (Jia & Aaronson, 1999) and thus to loosened L1 links in the neural substrates facilitating the learning of an L2 (Pallier et al., 2003).

Given the competing nature of these two accounts and conflicting empirical evidence provided, testing their premises in a similar group of bilingual speakers but this time for what happens to their L1 might help resolve the fundamental issue of how to conceptualize age effects. As argued by Schmid and Köpke (2017), a theoretical approach to bilingual development should be capable of predicting both attrition and the acquisition phenomena in that "if the framework fails to predict patterns which can be shown to occur in attrition, this should invalidate the theory in the same manner as would counterevidence from language acquisition studies" (p. 36).

In line with this argument, some L1 attrition researchers investigating age effects interpreted their findings within the premises of the two age accounts mentioned above, which were originally developed for L2 acquisition. In the current study, we follow the perspectives taken by these researchers and use the implications of these accounts on L1 attrition they proposed in order to evaluate our findings. The implication of the CPH on L1 attrition, for example, centres around the idea that while the chances to become native-like in the L2 before the so-called *critical period* (CP) are increased, this will have inevitable consequences for the degree of L1 attrition/maintenance (Bylund, 2009b; Montrul, 2008). During the CP, L1 knowledge is highly susceptible to attrition. This susceptability declines gradually due to

maturation in the neural connections and starts to plateau around the onset of puberty (Bylund, 2009b, 2009a).

Evidence comes from a study conducted with adult Korean adoptees in Sweden with the AaO range 1–10 and Swedish late learners of Korean (see Hyltenstam et al., 2009). The main aim was to trace whether some of the L1 remnants could be recovered by a relearning methodology in the adoptee participants. Both groups of participants were enrolled in a foreign language classroom at a university, learning Korean for an average of three years. As the results showed, while the adoptees were outperformed by the L2 learners in a GJT in Korean, there was no statistical significance between the two groups in the VOT perception test. The individual analysis, however, showed that the performance of some of the adoptees was better than the best-performing L2 learners. Additionally, the best regaining performance belonged to the adoptees with the oldest AaOs.

Based on this, Bylund (2009b) proposes that L1 ultimate attainment of late bilinguals is mainly constrained by maturational age effects (as their AaO is past the CP) which play a determining role over other factors such as L1 contact. In the case of early learners, in line with what Montrul (2008) claims, Bylund sees differences in input conditions and other non-maturational cognitive factors, e.g. language aptitude (see Bylund, Abrahamsson, & Hyltenstam, 2010) resulting in great variability in the linguistic knowledge of early bilinguals within the CP. In a way, these factors are suggested to compensate for the degree of loss that is proposed to occur due to maturational age effects in the first place.

The Interference Hypothesis (IH) or the L1 entrenchment view on the other hand, to date, has usually been applied to severe cases of L1 loss and thus underscores the important role played by continuous L1 exposure upon immigration in maintaining the L1 (Pallier, 2007). Its implications for less severe cases of L1 attrition seem to be most clearly observed in phonetic categories as predicted by the Speech Learning Model (SLM, compatible with the IH). The SLM was originally proposed to account for observed difficulties in the pronunciation of individual sounds experienced by L2 learners, which by default also has predictions for L1 speech production and perception (Flege, 1995; Flege, Schirru, & MacKay, 2003). In this model, both L1 and L2 sounds are assumed to exist in a common phonological space and influence each other (Flege, 1995). Interaction between the languages is proposed to lead to a bidirectional transfer between L1 and L2 sounds, which over time

might result in articulation of both L1 and L2 sounds differently from the monolingual norms (Flege et al., 2003).

Although the SLM does not predict a direct relationship between modifications of this sort and a global foreign accent in the L1 and/or L2, the possibility that changes in the production of L1 and L2 sound categories lead to accented speech in both of the languages has been previously tested and evaluated from the combined perspectives of both the IH and SLM by Yeni-Komshian et al. (2000). In view of their reasoning, advanced L2 learning would predict more changes in the L1 articulatory system, and these changes might cause the L1 to be produced with a foreign accent. In this view, the degree of L2-induced changes is determined by factors such as L1/L2 proficiency levels and amount of L1/L2 use rather than biological age effects (Yeni-Komshian et al., 2000).

It is, however, acknowledged that early bilinguals (with AaO up to puberty) are more likely to establish new categories for the new L2 sounds because the representation of the L1 sounds, i.e. the filtering effect of the L1 on the L2, in their case is not as strong (Yeni-Komshian et al., 2000). For this reason, early bilinguals are considered to be more likely to experience L2-induced sound modifications and develop an accented L1, while this might be observed to a lesser degree in the case of older bilinguals. In the current study, we follow this reasoning.

Hoping to contribute to what we know so far about age effects on language learning capacity and maintenance, the current study first aims to provide an overall picture of L1 proficiency in an immigrant context as a function of AaO by investigating the L1 performance of Turkish-English adult bilinguals in the UK across a wide AaO range (7–34). Secondly, it aims to investigate how well the role, if any, played by AaO can be evaluated within the premises of the theoretical accounts of bilingualism reviewed. Finally, in order to address the underlying sources of the observed asymmetry in the degree of vulnerability to attrition between phonological and structural properties as a function of AaO, we carried out an investigation of two different linguistic skills (structural complexity and global L1 pronunciation) pertaining to these two domains. Given that previous literature linked both L1 and L2 ultimate attainment in similar properties to the impact of additional factors as well, we also tested the effects of L1 proficiency, amount of L1 contact, linguistic identification and cultural affiliation.

The focus of the study

Carrying out this investigation in Turkish is particularly relevant as it has a variety of structures (e.g. complex morphology) which have been previously shown to be the potential loci for erosion and age effects in other languages. Moreover, although previous findings point to a deterioration in the knowledge and use of Turkish spoken in Europe across generations (Arslan, De Kok, & Bastiaanse, 2015; Gürel & Yılmaz, 2011; Huls & van de Mond, 1992; Onar Valk & Backus, 2013; Yılmaz, 2011 among them), we know very little about the role played by AaO in this. This role can only be revealed by looking at the AaO effect in similar features/properties previously shown to be eroded.

One such feature is structural complexity. Turkish is an agglutinative language with complex morphology and employs a variety of complex subordination structures through synthetic processes (Huls & van de Mond, 1992). One way attrition manifests itself is simplifications/reductions in the overall complexity of the linguistic system either because L1 is not activated enough or because of contrastive differences between L1 and L2 (Andersen, 1982; Schmid, Köpke, & Bot, 201.2; Seliger & Vago, 1991). Due to its agglutinative nature, Turkish allows such tendencies to be observed very easily (Huls & van de Mond, 1992).

This might manifest itself first in word formation. As hypothesized by Huls and van de Mond (1992), instead of relying on suffixation —a costly process with each suffix having their own morphological function— one might develop an analytical tendency towards using free morphemes instead. They tested this by using a measure called *agglutination index* based on Lyons (1969 as cited in Huls and van de Mond, 1992) in a small scale study in the L1 performance of two Turkish families (parents and children) in the Netherlands. This measure was calculated as the ratio of the number of morphemes over words per each sentence produced and proven to be a reliable measure in revealing differences between generations.

Similarly, Treffers-Daller et al. (2007) and Onar-Valk and Backus (2013) confirmed that adult HSs avoid complex non-finite clauses and rely on more analytical means by using finite subordination which also structurally resembles the subordination formation in the L2s tested (Dutch or German). Such tendencies, however, were not observed in the performance of late Turkish-Dutch bilinguals (Yılmaz, 2011). As revealed by Treffers-Daller et al. (2007), HSs showed a tendency to avoid relative clauses and three types of verbal complements (*-mA, -DIK, -AcAK)* due to costly functional operations (e.g. agreement morphology between subject

and the verb of the subordinate clause, genitive case-marked overt subjects) required for their formation (see Treffers-Daller et al., 2007 for details). The authors acknowledged, however, that this might be a result of incomplete acquisition rather than attrition as subordination is a relatively late-acquired phenomenon. Monolingual acquisition of relative clauses was reported not to stabilise before the ages around 4–5 (Slobin, 1986) and verbal complements around 5–6 (Aksu-Koç, 1994). These facts are taken into consideration in the selection of our participants.

Another property that we investigate is global L1 accent. To our knowledge, the only study looking at this in L1 Turkish is the one conducted by Stangen, Kupisch, Proietti Ergün, & Zielke (2015). This research looked at whether being bilingual entails sounding less native-like in one or both of the languages of Turkish-German bilinguals as a function of AaO. The AaO range included in the study was 0–9 divided into two groups: one with the AaO range 0–3 and the other with 4–9. As the results showed, the majority of the speakers were perceived as sounding less native-like in either language, and only 3 out of 21 speakers were perceived as sounding native-like in both. AaO, however, did not play a deterministic role in the outcome. Given that the raters linked intonation and pronunciation of some vowels to the degree of accentedness, the authors attributed their findings to bidirectional transfer. There is no direct evidence on the L1 pronunciation of late Turkish bilinguals. It is, therefore, difficult to derive conclusions on what constraints the L1 accent. By including a wide AaO range in the current study, we thus aim to provide some answers to this question.

Hypotheses

Acknowledging the difficulty of disentangling the nature of the role played by AaO in language acquisition and retention as predicted by the CPH and IH, recent SLA investigations suggested controlling for possible confounding factors such as amount of language use and language proficiency either statistically (e.g. Veríssimo, Heyer, Jacob, & Clahsen, 2017) or by establishing a control group that matches the experimental group in terms of confounding factors (e.g. Hopp & Schmid, 2013; Schmid, 2014). In the current study, we control for the effects of confounding factors statistically and take the range of scores delimited by the monolingual controls as the baseline. This means that bilinguals scoring lower than the worst-performing controls in the L1 measures used are considered to have remained below the "control range" and thus experienced some degree of attrition. In view of the logic applied in the studies mentioned above and the implications of the theoretical age accounts for L1

attrition which have been reviewed in the previous sections, we address the following hypotheses:

L1 maintenance is constrained by maturational age effects

If maturational age effects determine the degree of L1 maintenance, AaO should be the primary determining factor accounting for the variability in the L1 performance when the impact of the confounding factors such as L1 contact and proficiency are controlled for. We can, however, only confirm that this role is of a maturational nature, should we find the L1 performance of all bilinguals whose AaO past a certain period (age 12 based on previous reports) to be resistant to attrition (Bylund, 2009b). Additionally, the speakers whose AaO remains below this cut-off point should show greater variation in the degree of their L1 maintenance with many of them potentially remaining outside the control range, and AaO should still make a significant contribution within this group presumably in addition to the effects of other compensatory variables, e.g. L1 use (Bylund, 2009b). This should apply to both structural complexity and global L1 accent performance.

The degree of L1 maintenance is determined by the degree of L1 entrenchment

If on the other hand, age effects are disguised as other factors, such as the degree of L1 entrenchment, frequency of L1 use, and attitudes, some or all of these factors should explain the outcome across the entire AaO range included. In view of the predictions of the SLM and IH for L1 attrition as reviewed above, the L1 accent of bilinguals is expected to show a tendency to deviate from the monolingual norm at the group level. The extent of the deviations. however, might differ between early and late learners due to their differing levels of L1 entrenchment and L2 proficiency (Yeni-Komshian et al., 2000). There is no direct empirical evidence pointing to the degree of L1 attrition on structural properties in traditional cases of attrition from an IH perspective. That notwithstanding, assuming an inverse relationship between L1 and L2 proficiency (Pallier et al., 2003), reduced frequency of L1 use and speaking another language should also result in loosened links in the degree of the L1 entrenchment regarding the structural complexity performance of some of the bilinguals.

Methodology

Participants

The L1 performance of 57 adult Turkish-English immigrant bilinguals (IBs) with the AaO range 7–34, and of 29 monolinguals as a control group (CG) in Turkey was investigated. Although AaO was considered as a continuous variable, we should note that our sample typically represents two AaO groups (AaO<12 or AaO>12) with a roughly equal number of participants in each based on the previous literature which considers age twelve as a cutoff point between early and late bilinguals (e.g. Bylund, 2009b). All bilinguals were born in Turkey and acquired Turkish as their L1[1]. The length of residence (LOR) was set to a minimum of 8 years. Contacts have been made through various Turkish clubs in London and via the snowball technique. Individuals have either been visited in their homes or in public cafes of their choice.

The main criterion of participant selection was their AaO taken as the age of arrival in the UK. As pointed out by Flores, Santos, Jesus, & Marques (2017, p. 797), working with adult HSs "cannot distinguish effects of acquisition from effects of subsequent language attrition". This is especially the case for late-acquired properties like verbal complements and relative clauses which have been reported not to stabilise before the ages 5–6 in monolingual Turkish (Aksu-Koç, 1994; Slobin, 1986). Previous investigations of subordination confirmed that Turkish HSs might not have acquired these structures completely due to an insufficient amount of input (Huls & van de Mond, 1992; Treffers-Daller et al., 2007). The problem this situation might create regarding the interpretation of the exact role of AaO in L1 attrition has repeatedly been pointed out (Bylund, 2009a; Bylund et al., 2010).

Based on the psycholinguistic literature consulted, our primary motivation behind setting the lowest AaO limit to age seven was thus to ensure these late-acquired properties had time to develop age-appropriately before the emigration took place. L1 accent, which is usually assumed to develop earlier than the set age limit (Yeni-Komshian et al., 2000), does not seem to pose any problems in this respect. This allowed us to mainly exclude the possibility that any age effect we could find was due to incomplete acquisition.

Personal and linguistic background information of the participants was obtained using a sociolinguistic questionnaire (SQ) adapted from Yılmaz (2013) which was constructed based on the test battery of Monika S. Schmid (https://languageattrition.org/). Following Schmid and Dusseldorp (2010), a principal component analysis (PCA) with varimax (25) rotation was

conducted on the items that asked participants to report the frequency of current L1 contact and, linguistic and cultural preferences on a scale ranging between 0 and 1. Four new composite variables were calculated as the means of the variables included in each component below. Internal consistency was established by a reliability analysis (Cronbach alpha).

– Interactive L1 use (with children, siblings, parents, grandparents in Turkey, other relatives in the UK, in written communication with relatives in the UK and Turkey)

– L1 passive exposure (through TV, radio, and music).

– Linguistic identification (importance given to maintain their L1 and that their children understand and speak Turkish)

– Cultural affiliation (cultural preferences for friends and L1 use with friends and neighbours)

Although participants were also asked to report on their past L1 use (during the first five years upon arrival), it was not possible to conduct a PCA for past L1 use due to either lack of variability in the answers given or a large number of missing values in some of the questions. Instead, Pearson correlations were checked for the variables that were answered by all participants. A mean value of these variables[2] was obtained to represent past L1 use. Table 1 below provides quantitative information on the newly-established variables.

<Table 1>

The CG was representative of the bilinguals regarding the city of birth, gender, age at testing, and educational background. The highest education level was calculated in years by taking into account the last education level completed either in Turkey or the UK. Table 2 provides basic background information about the participants.

<Table 2>

The general L1 proficiency was measured by a 40-item written cloze test (C-Test). An independent sample t-test did not reveal any group differences between bilinguals (M=28.39) and monolinguals (M=31.34) at the group level (t(-2.499)=73.459, p=0.17), although the performance of five bilinguals was not within the control range (see Figure 1 below).

<Figure 1>

Semi-structured interviews

The most suitable data to capture attrition effects in a group of bilinguals with differing AaOs has been suggested to be free speech data (Schmid et al., 2012). This is because it allows "every speaker to employ the full range of her language knowledge" without applying too much cognitive pressure and thus prevents any observations of ceiling effect in the performance of late bilinguals, or of failures in completing the task because it is too demanding for early bilinguals (Schmid et al., 2012, p. 678). Based on this, the current study relied on spoken data collected through a semi-structured interview to naturally detect reductions, if any, in the overall complexity and degree of accentedness in the L1.

In addition to asking participants to share their views on daily topics (see L1 attrition test battery https://languageattrition.org/), four questions, which were originally designed for an earlier investigation of past tense usage in Turkish (Karayayla, forthc.), asked them to tell stories personally experienced or heard from other people that they found interesting, horrifying or amusing. All conversations were, therefore, very spontaneous and rich regarding subordination and many other grammatical structures. Individual recordings lasted from 10 to 35 minutes (M=19.04).

The transcription was done according to CHAT conventions (MacWhinney, 2000). The transcribed data was segmented into AS-Units (Foster, Tonkyn, & Wigglesworth, 2000). Additional criteria were adapted from Berman and Slobin (1994), and Young (1995). Among the principled criteria of data exclusion proposed by Foster et al. (2000), exclusion was carried out at level three. This means that only the AS-Units that included finite or non-finite subordinate clauses together with a main clause, and simple independent clauses were included in the total count to analyse, while other units such as repetitions and errors were excluded. Subordinate clauses were coded by their type and sub-type. Since the number of non-target-like subordination was extremely low, accuracy was not investigated. There were overall 18,351 AS-Units consisting of 25,146 clauses. The pruned speech data consisted of 96,564 words. This data was used to approximate structural complexity and conduct a foreign accent rating experiment.

Structural complexity

Following findings of previous research, we calculated the ratio of total number of morphemes[3] over words (the agglutination index by Huls and van de Mond, 1992) and

counted the number of non-finite relative clauses and three types of verbal complements (-*mA*, -*DIK* and –*AcAK*) per AS-Unit per participant. These non-finite clauses were revealed to occur least frequently in heritage Turkish by Treffers-Daller et al. (2007). Adapting recent SLA methodologies that used similar sub-measures to approximate structural complexity (Lahmann, Steinkrauss, & Schmid, 2016), we Z-transformed the sub-scales and then incorporated them into one single measure of structural complexity by using the reshape package in R. We standardised the final scale by Z-transforming it one more time. A higher score in each sub-component and thus an overall higher score reflected that the speaker did not develop a preference towards using more analytical means or less simple language.

Global foreign accent ratings (FAR)

In order to detect changes in the L1 accent as a function of AaO, we conducted a FAR experiment. Following the procedure and criteria used in de Leeuw et al. (2010), short speech samples (M=16.49 seconds) from the spoken performance of each bilingual and CG speaker (as a response to the same question) were extracted. Particular attention was given to include fully-finished utterances without any code-switching or grammatical mistakes.

Twenty-eight judges with Turkish as their only native language (age range 19–23, M=19.78) were recruited among the first year university students studying foreign language education at the Middle East Technical University in Turkey[4]. The original experiment[5] lasted 52 minutes to complete and took place in a quiet room in the department where the judges listened to the samples through laptop speakers and were asked to rate the degree of the perceived foreign accent of each speaker during the 7 second-long pauses after each sample. The scale used was the final 6-point Likert Scale used in Hopp and Schmid (2013) and ranged between 1=native accent and 6=strong foreign accent. The judges were informed not to confuse the regional accent with a foreign accent during the practice session employed before the experiment[6].

The final score that each speaker had was calculated as the mean value of the ratings given by 28 judges. A higher FAR was an indication that the speaker was perceived to sound less native-like[7].

Analysis of the Data

For each L1 measure we used in the current study, there were multiple responses per subject, which would violate the independence assumption of traditional linear models. Mixed effect

models are considered suitable in such cases (Baayen, Davidson, & Bates, 2008). Given this, we analysed the data by using a linear mixed effects regression modelling with the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015) for the R statistical platform version 3.2.4 (R Core Team, 2016).

We ran separate models for each measure and considered 'participant' and 'rater' (in FAR models) as random factors in order to control for the variability. We log-transformed the dependent variable (FAR) to achieve a normal distribution (see online materials Figure 5, 6 and 7 for the distribution of residuals of the models built). In order to provide a reliable answer to the nature of the role played by AaO and test our hypotheses, it was necessary to statistically control for the effects of external/confounding variables such as L1 contact, education level, and L1 proficiency.

In order, however, not to overfit the data by including too many predictors (Wurm & Fisicaro, 2014), in the final models we only included the ones that predicted the outcome significantly when tested alone. Given that we were particularly interested in the explanatory power of each fixed effect while holding other variables constant, following the suggestions in Wurm and Fisicaro (2014) and the methodology employed by Veríssimo et al. (2017), we included the predictors simultaneously in the final models. The use of this method is also justified if one wishes to control for the correlations between the fixed effects (if any) included as covariates (e.g. Veríssimo et al., 2017).

P-values were obtained by using the *lmerTest* package (Luke, 2017). A comparison of the AIC values of the models showed that addition of random slopes was not justified (Baayen et al., 2008).

Results

Structural complexity

According to Table 3 below, unlike hypothesised, group means do not seem to diverge from each other in any of the sub-components used to approximate the structural complexity score. The statistical model that we conducted on the merged structural complexity Z-score confirmed that the bilinguals did not diverge significantly from the reference group (ß=0.016, SE=0.02, t=0.82).

<Table 3>

As plotted in Figure 2, there was no significant relationship between AaO and the structural complexity ($F(1, 59203)=2.322$, $p=0.127$, ß=-0.00009, $t= -1.524$). Only three bilinguals remained outside the control range. High scores were obtained at all ages. While the lowest scores were obtained by participants with AaOs 8, 9, and 14, the highest scores belonged to participants with AaOs 7, 10, and 13.

Given that participants showed full retention of proficiency regarding the overall structural complexity of their native language within the AaO range investigated here, no further analyses were conducted.

<Figure 2>

Foreign accent ratings

According to Figure 3 below, all monolinguals and a great majority (71.92 %) of bilinguals fell within the range of unambiguous L1 speakers (the control range). In the case of the bilingual group, however, there was a much wider distribution with 16 participants (28.07 %) falling into the non-native range. Four of these were outliers, with one of them being perceived as unambiguously non-native by all raters and the rest having a FAR over four. The AaOs of these outliers were 8, 9, 10 and 13.

<Figure 3>

In order to prevent the influence of the outliers on the dependent measure, the statistical model was conducted without the outlier data points (n=53). This model evidenced that the judges had a tendency to perceive the bilingual participants to sound less native-like in comparison to the monolinguals at the group level (monolinguals mean=1.378, bilinguals mean=1.733, ß=0.355, SE= 0.12, $t=2.79$). This model accounted for 35.5 % of the variance.

There was a significant relationship between AaO and FAR ($F(1, 1482)=138.6$, $p<0.001$, ß=-0.05, $t= -11.77$), which is captured in Figure 4 below. As the fitted line of a high order polynomial function demonstrates, the relationship between AaO and FAR is quite linear until around ages 13–14, and then starts levelling off where the foreign accent is not a function of AaO anymore.

<Figure 4>

AaO, however, accounted for only 8.5 % of the variance in the outcome. In the next step, we thus tested which other variables besides AaO contributed to the explained variance. The coefficients of the final model, which accounted for 31.7 % of the variance, are provided in Table 4. According to this, the participants with older AaOs (ß=-0.013, SE= 0.004, t=-2.66), those with higher scores in L1 C-Test (ß=-0.014, SE= 0.004, t=-2.94), those with more L1 passive exposure (ß=-0.287, SE= 0.133, t=-2.16) and those who were older at the time of testing (ß=-0.009, SE= 0.005, t=-2.14) were perceived to sound more native-like. None of the other variables including the background education level predicted the outcome.

<Table 4>

Based on the function of the slope in Figure 4 above, we created two subsets of bilinguals with AaO<14 (n=30) and with AaO>13 (n=23). We will call them the early bilinguals (EBs) and the late bilinguals (LBs) respectively. This was crucial to test our hypotheses and see if the reported role played by AaO above remained significant and independent when the effects of the confounding predictors were controlled for. The final model accounting for 29.7 % of the variance showed that the EBs with increased L1 proficiency (C-Test) (ß=-0.015, SE= 0.007, t=-2.118) and more L1 passive exposure (ß=-0.502, SE= 0.206, t=-2.438) sounded more native-like, while AaO (ß=-0.019, SE= 0.029, t=-0.663) ceased to contribute to the explained variance.

For the LBs, on the other hand, the only variable that came back as significant was age at testing (ß=-0.012, SE= 0.005, t=-2.441): older participants were perceived as sounding more native-like. To unravel whether the role played by age at testing was attrition-specific, we checked whether it predicted the variability in the performance of the CG as well. A simple linear regression analysis revealed that older monolinguals were also perceived as sounding more native-like (F(1, 9)=810, p=0.002, ß=0.009, t=3.00). Why this should be the case deserves further empirical scrutiny, but regarding our data, it is safe to say that it is not an attrition-specific variable and the L1 accent of the bilinguals whose AaO is over 13 is resistant to attrition.

Discussion

Overall, our findings showed that while the bilinguals as a group managed to attain a target-like level of proficiency regarding the overall structural complexity of their L1, this was not the case regarding the degree of sounding native-like. While the performance of only three participants (5.2 %) remained below the control range in the structural complexity measure with the rest performing target-like, sixteen participants (28.07 %) fell into the non-native range which significantly distinguished their accent from that of the controls at the group level.

The full retention of proficiency did not allow us to establish a relationship between AaO and the structural complexity scores obtained. The AaO–FAR slope, on the other hand, dropped linearly until it levelled off after AaO 13. In other words, it showed a clear discontinuity with all participants (except for one with AaO=18) past this age falling into the control range. The investigation of this participant showed that he had a heavy regional accent. Although the judges were informed about this, some variability in this participant's grammatical/lexical choices due to the regional accent might have misled their judgment (de Leeuw et al., 2010). In the group whose AaO remained below this cut-off point (AaO<14, n=34 including the four outliers), there was much more variability and the FAR of fifteen participants (44.11 %) fell outside the control range. This part of our observation thus seems to corroborate our first hypothesis, which would predict that the degree of L1 retention is primarily determined by maturational constraints.

We found, however, that AaO was not the only significant predictor explaining the variability in the scores. The outcome was instead a result of an interplay of AaO, amount of passive exposure to L1, level of general L1 proficiency, and biological age. Furthermore, the individual explanatory power of AaO was not any better than that of the other covariates (see the Result section). This seems to run contrary to the expectations about an independent or more significant role played by AaO and thus prevents us from ascribing our findings "fully" to maturational age effects.

The non-native traces in the L1 accent of the bilinguals might rather be a result of the reorganization of the L1 phonetic system under the influence of the L2 assuming that L1 and L2 sound categories exist in a shared system interacting with each other and that the same speech learning mechanisms are active throughout the lifespan (Flege, 1995). In line with the

predictions of the SLM, our statistical findings thus appear promising in showing that L1 sound categories were adaptive even in adulthood presumably under the influence of L2 sound categories and the degree of this influence was not constrained by AaO only, but also by L1 proficiency, frequency of L1 exposure, and biological age. Later, we showed that biological age here did not play an attrition-specific role because older monolinguals were also perceived as sounding more native-like.

The role played by the frequency of L1 (passive) exposure is remarkable, which relates to what Schmid (2007) suggests about the role played by the quality of L1 contact in attrition. Getting exposed to qualitatively native-like input in adulthood on a frequent basis might have helped distinguish between phonetic characteristics of L1 and L2 sounds and prevent L1 sound categories from being modified. This finding parallels previous reports about the protective role of L1 contact in maintaining L1 accent in adulthood (Yeni-Komshian et al., 2000). It is thus plausible to assume that the role played by AaO we reported was quantitative rather than qualitative unlike what the CPH proposes (Flege, 1995). More precisely, as stated in our second hypothesis, the individual contribution of AaO to the explained variance seems to be a result of differing degrees in the L1 entrenchment of our speakers rather than irreversible neurological changes.

The fact all the bilinguals in our study (except for the one that we already discussed above) past AaO 13 were perceived as sounding unambiguously native-like still poses a significant challenge to this explanation. Interference accounts assume an inverse relationship between L1 and L2 proficiency, which should result in at least some of the late bilinguals' L1 accent being perceived as divergent as well. That notwithstanding, these accounts, in general, acknowledge that L2 interference on the L1 might be limited in late bilingualism due to deeply entrenched L1 representations (Pallier, 2007; Yeni-Komshian et al., 2000). Based on our findings, we can speculate that being monolingual at least for this much amount of time results in the representation of L1 categories being deeply entrenched and this makes it quite resistant to external factors and L2 interference.

In our view, this does not necessarily indicate irreversibility or that no interaction between L1 and L2 took place. Phonological drifts might have happened as a result of interactions between L1 and L2 sounds (that our experiment would not capture). Yet, these changes might not have led to deteriorations in the L1 accent if, for example, certain conditions related to frequency and intensity of L1 and/or L2 contact were not met (Chang, 2012).

In general, late immigrants tend to continue using their L1 on a frequent basis and remain mostly L1-dominant (Jia & Aaronson, 1999). It is thus plausible to assume that there might be a certain threshold of L1 use/exposure necessary for the established L1 links not to be weakened upon immigration and this threshold might have been already reached in the case of our late bilinguals. This would explain the null effects of the external variables in this group. In a similar vein, a certain level of intense L2 experience going beyond typical daily L2 use might be necessary for the L2 to affect the deeply entrenched L1. Previous investigations of L1 accent conducted with late German bilinguals, who were reported to be very proficient L2 users and using their L1 less frequently than our participants, provide some support to this explanation (e.g. Bergmann et al., 2017; de Leeuw et al., 2010). As a result, the high levels of L1 retention in this group could be due to availability of the L1 upon immigration rather than age-related reduced susceptibility to attrition.

Distinguishing between the effects of maturational constraints and L1 entrenchment in a group like this is indeed very difficult. In theory, one solution could be to investigate the L1 performance of an additional group of LBs whose L1 contact ceased completely upon immigration (see Schmid, 2012 for details). If the main cause of high levels of L1 retention among late bilinguals (as observed in this study) is due to reduced susceptibility to attrition, which is predicted to be an irreversible process, then no group differences should be obtained and L1 should be retained to a considerable degree even in the group with no prolonged L1 contact. In practice, however, it is extremely difficult to find such comparable groups.

To our knowledge, the only investigation to date has been Schmid's (2012) investigation of age effects in two groups of post-puberty bilinguals (the Holocaust survivors with AaO range 11-15) with and without continuous L1 contact upon immigration. This research demonstrated that the degree of L1 loss in morphosyntactic properties, which was found to be minimal, was better predicted by AaO rather than availability of the L1 upon emigration. It is however not easy to see how this finding can be taken as a direct counter-evidence to our argument above regarding the L1 accent of our participants. This would require further investigations of L1 accent with a similar profile of bilingual participants. Until proven otherwise we therefore argue that our findings, in general, are "more compatible" with an L1 entrenchment view than with a maturational view. We, however, acknowledge that more detailed reports on L1/L2 use and proficiency should be obtained and additional analyses, e.g. acoustic analyses should be carried out to arrive at a more definitive answer.

The picture for the L1 performance in structural complexity was quite different from what we observed in L1 accent as all participants performed fully target-like. Although this is an outcome which was not predicted in our hypotheses, these findings are entirely in line with what Kupisch et al. (2014) found in the performance of adult French-German simultaneous bilinguals (2L1s) who acquired French either in a minority or majority context. While all participants regardless of the context performed target-like in a variety of morphosyntactic categories in controlled tasks, those who acquired French in the minority context had an accented L1 and drifted VOT values. The authors discussed that even if they investigated the morphosyntactic performance in free speech rather than in controlled tasks, their participants would still perform target-like. This is because, as the authors evaluate, speakers are in control of how to express things and might avoid certain structures by compensating for them through other means. The same, however, would be less accurate in pronunciation as it would not be possible to find alternative ways of pronouncing a sound (Kupisch et al., 2014).

This explanation might account for the asymmetry we found across our linguistic measures to some extent. Although we counted the number of different types of non-finite clauses which were previously reported to be used infrequently in immigrant Turkish (Treffers-Daller et al., 2007), we did not look at the finite/non-finite clause distribution in general or in specific contexts. In Onar-Valk and Backus' (2013) study, for example, adult HSs compensated for the non-finite clauses by using finite-clauses in reported speech contexts more than they did in other contexts. In that sense, as raised by one of the reviewers as well, we acknowledge that our measure might not have been sensitive enough to detect such compensatory tendencies.

On the other hand, our participants showed full retention of L1 proficiency in the second component that we included in the structural complexity measure as well: the agglutination index. This indicates that none of the participants avoided synthetic costly processes by relying on more analytic means. It follows from this that not all linguistic measures are subject to age effects. It is widely acknowledged in the SLA literature that age effects do not modulate the ultimate attainment in an L2 across the entire range of linguistic domains or even across the properties within the same domain, which is called the selectivity of age effects (e.g. Veríssimo et al., 2017). We can argue that the same holds for L1 attrition and general structural complexity might be something that is not selected by age effects.

Selectivity phenomenon is, in fact, not new to L1 attrition research. Previous research demonstrated external interface-governed structures such as distribution of subject pronouns and differential object marking as potential loci for erosion (Chamorro, Sturt, & Sorace, 2016), which is often framed within the Interface Hypothesis as formulated by Sorace (2011). There is also evidence showing that structures which are not in competition between the L1 and L2 would be fully retained (Gürel, 2004). As Gürel's investigation of L1 attrition in long-term Turkish late bilinguals in Canada exemplified, only the pronoun *o*, the binding domain of which is in competition with the English pronoun s/he, was affected. The binding domains of the other two Turkish pronouns were fully retained. By analogy, high levels of L1 retention across the entire AaO range in the current study might thus relate to the lack of a direct competition between the L1 and L2 structures under investigation (e.g. agglutination, non-finite clauses), and also to the fact that these properties are not governed by external interfaces.

Unlike what Gürel's study revealed above, a recent study investigating the role of AaO in a group of Korean-English pre- and post-puberty learners showed that pre-puberty learners (with AaO up to age 12) failed to perceive L1-specific phonemic contrasts, but did not have problems with the contrasts that are similar to the L2 sounds (Ahn et al., 2017). It, therefore, appears that even phonological competence might be subject to a AaO-determined selective process, but how the level of cross-linguistic similarity/competition between the L1 and L2 influences this outcome might vary based on the language pairs and the linguistic domain. If this is the case, it remains to be seen in the future to what extent there is an overlap between selectivity of attrition and age effects across different linguistic properties with differing levels of competition between L1 and L2. It is also important to employ different methodologies to see whether task demands also play a role in this selectivity.

Conclusion

The present investigation set out to explore the relationship, if any, between AaO and the degree of L1 attrition in the overall structural complexity and the perceived accent. The spoken performance of adult Turkish immigrants in the UK (n=57) with a wide AaO range (7–34) was compared to that of a group of controls (n=29). We formulated our hypotheses based on the premises of two competing accounts: the CPH and the L1 entrenchment view, with a hope that testing these models' capacities in accounting for L1 attrition phenomena might help resolve the fundamental issue of how to conceptualise age effects.

Overall, our findings suggest that L1 accent is sensitive to the effects of external factors and AaO, which we propose to be taken as a proxy for the level of L1 entrenchment instead of the maturational state of the speaker. However, given that we did not have detailed reports on the L2 use or obtained any measures of L2 proficiency, these findings should be taken as preliminary and tentative. It is difficult to claim something similar for the structural complexity performance as all participants performed target-like. One possible explanation for the asymmetry found in the degree of attrition across the two linguistic measures could be the differing levels of competition between the L1 and L2. Even if this was the case, neither of the accounts makes an explicit claim for this, and therefore remains insufficient to account for the findings.

Taken together, if Schmid and Köpke (2017, p. 2) are right in their proposal that L1 attrition findings can "be used to inform, challenge, and validate theoretical approaches of bilingual development", we believe that our findings, despite being preliminary, should be used to inform implications of these models for L1 attrition to accommodate the phenomena such as selectivity and degree of competition between L1 and L2 structures.

Without any doubt, more research needs to be carried out in order to arrive at more definitive answers. We suggest that future researchers should investigate age effects in a number of other properties with differing levels of competition between L1 and L2 by including participants with younger AaOs and a greater variety of language use. This could be achieved using an additional group with a profile similar to that of adoptees whose L1 exposure ceases completely upon immigration either in post or pre-puberty ages. Based on our findings, this is crucial to see these models' limits in accounting for L1 attrition data.

NOTES

1. Effects of any other known native languages such as Kurdish were controlled. Nevertheless, in some cases this was inevitable. One participant learned some Kurdish from her grandmother at the age of 7 but lost the ability to speak the language upon immigration at age 8. Similarly, two participants stated they had a minimum level of Kurdish knowledge. One participant was born in Cyprus where she spent 13 years before her arrival in the UK. She reported not to have acquired/used Cypriot Turkish as she had parents from Turkey.

2. Past L1 frequency of use with "siblings" and "parents" ($r=0.446$, $p=0.01$).

3. The morpheme counts were obtained automatically with the aid of a Turkish morphological parser and a disambiguator developed by Sak et al. (2009) with 96.7 % success rate.

4. An anonymous reviewer pointed out that using monolingual raters instead of raters with L2 English background would have been more appropriate. Our choice of bilinguals, however, was a deliberate one: various studies have found that familiarity with the language background and language combinations of the speakers to be rated can improve inter-reliability and also leads to raters being somewhat more lenient (e.g. Carey, Mannell, & Dunn, 2011), and that even non-native speakers are able to rate speakers reliably (e.g. Xi & Mollaun, 2011). In order to give all of our speakers the 'best' chance of being perceived as natives, we felt that the choice of bilingual raters would be better than choosing speakers entirely unfamiliar with the language that our speakers use in daily life.

5. This investigation was originally designed for a larger project and thus included speech samples of an additional group of UK-born adult HSs (n=31) who were not included here in this study due to concerns about their incomplete attainment.

6. The same anonymous reviewer also suggested that the raters should ideally have had the same region of origin as with the participants in the sample because their choices of rating an accent as "foreign" might be confounded with that participant's heavy "regional accent". The reviewer is, indeed, correct that familiarity with regional accents can affect the accuracy of ratings (e.g. Flege, Frieda, & Nozawa, 1997). This is why we very carefully matched the experimental and the control speakers for region of origin. We can thus assume that regional dialects occur to the same degree in the monolingual and the bilingual groups, and the fact that all monolinguals were unambiguously rated as L1 speakers strongly suggests the absence of a confound as suggested by the reviewer.

7. Note that being perceived as sounding "non-native" does not entail that that speaker is a non-native speaker of Turkish.

References

Ahn, S., Chang, C. B., DeKeyser, R., & Lee-Ellis, S. (2017). Age Effects in First Language Attrition: Speech Perception by Korean-English Bilinguals. *Language Learning*, *67*(3), 694–733. https://doi.org/10.1111/lang.12252

Aksu-Koç, A. (1994). Development of Linguistic Forms: Turkish. In R. A. Berman & D. I. Slobin (Eds.), *Relating Events in Narrative: A Crosslinguistic Developmental Study* (pp. 329–385). Lawrence Erlbaum Associates.

Altenberg, E. P. (1991). Assessing first language vulnerability to attrition. In H. W. Seliger & R. M. Vago (Eds.), *First language attrition* (pp. 189–206). Cambridge: Cambridge University Press.

Andersen, R. W. (1982). Determining the linguistic attributes of language attrition. In R. D. Lambert & B. F. Freed (Eds.), *The Loss of language skills* (pp. 83–117). Rowley, Mass: Newbury House Publishers.

Arslan, S., De Kok, D., & Bastiaanse, R. (2015). Processing grammatical evidentiality and time reference in Turkish heritage and monolingual speakers. *Bilingualism: Language and Cognition, FirstView*, 1–16. https://doi.org/10.1017/S136672891500084X

Au, T. K., Knightly, L. M., Jun, S.-A., & Oh, J. S. (2002). Overhearing a Language During Childhood. *Psychological Science*, *13*(3), 238–243. https://doi.org/10.1111/1467-9280.00444

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/doi:10.18637/jss.v067.i01

Bergmann, C., Nota, A., & Schmid, S. M. (2017). L2 immersion causes non-native-like L1 pronunciation in German attriters. *Journal of Phonetics*, (58), 1–23. https://doi.org/10.16/j.wocn.2016.07.001

Berman, R. A., & Slobin, D. I. (1994). *Relating Events in Narrative: A Crosslinguistic Developmental Study*. Lawrence Erlbaum Associates.

Bornstein, M. H. (1989). Sensitive periods in development: Structural characteristics and causal interpretations. *Psychological Bulletin*, *105*(2), 179–197.

Bylund, E. (2009a). Effects of age of L2 acquisition on L1 event conceptualization patterns. *Bilingualism: Language and Cognition*, *12*(03), 305–322. https://doi.org/10.1017/S1366728909990137

Bylund, E. (2009b). Maturational Constraints and First Language Attrition. *Language Learning*, *59*(3), 687–715. https://doi.org/10.1111/j.1467-9922.2009.00521.x

Bylund, E., Abrahamsson, N., & Hyltenstam, K. (2010). The Role of Language Aptitude in First Language Attrition: The Case of Pre-Pubescent Attriters. *Applied Linguistics*, *31*(3), 443–464. https://doi.org/10.1093/applin/amp059

Carey, M. D., Mannell, R. H., & Dunn, P. K. (2011). Does a rater's familiarity with a candidate's pronunciation affect the rating in oral proficiency interviews? *Language Testing*, *28*(2), 201–219. https://doi.org/10.1177/0265532210393704

Chamorro, G., Sorace, A., & Sturt, P. (2016). What is the source of L1 attrition? The effect of recent L1 re-exposure on Spanish speakers under L1 attrition. *Bilingualism: Language and Cognition*, *19*(3), 520–532. https://doi.org/10.1017/S1366728915000152

Chamorro, G., Sturt, P., & Sorace, A. (2016). Selectivity in L1 Attrition: Differential Object Marking in Spanish Near-Native Speakers of English. *Journal of Psycholinguistic Research*, *45*(3), 697–715. https://doi.org/10.1007/s10936-015-9372-4

Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics*, *40*(2), 249–268. https://doi.org/10.1016/j.wocn.2011.10.007

Cuza, A., & Pérez-Tattam, R. (2016). Grammatical gender selection and phrasal word order in child heritage Spanish: A feature re-assembly approach. *Bilingualism: Language and Cognition*, *19*(01), 50–68. https://doi.org/10.1017/S1366728914000893

De Bot, K., Gommans, P., & Rossing, C. (1991). L1 loss in an L2 environment: Dutch immigrants in France. In H. W. Seliger & R. M. Vago (Eds.), *First language attrition* (pp. 87–98). Cambridge: Cambridge University Press.

de Leeuw, E., Schmid, M. S., & Mennen, I. (2010). The effects of contact on native language pronunciation in an L2 migrant setting. *Bilingualism: Language and Cognition*, *13*(Special Issue 01), 33–40. https://doi.org/10.1017/S1366728909990289

DeKeyser, R. M. (2013). Age Effects in Second Language Learning: Stepping Stones Toward Better Understanding. *Language Learning*, *63*, 52–67. https://doi.org/10.1111/j.1467-9922.2012.00737.x

Flege, J. E. (1995). Speech perception and linguistic experience: Issues in cross-language research. In W. Strange (Ed.), *Second language speech learning: theory, findings, and problems* (pp. 233–272). Baltimore:MD: York Press.

Flege, J. E., Frieda, E. M., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics*, *25*(2), 169–186. https://doi.org/10.1006/jpho.1996.0040

Flege, J. E., Schirru, C., & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, *40*(4), 467–491. https://doi.org/10.1016/S0167-6393(02)00128-0

Flores, C., & Rato, A. (2016). Global accent in the Portuguese speech of heritage returnees. *Heritage Language Journal*, *13*(2), 161–183.

Flores, C., Santos, A. L., Jesus, A., & Marques, R. (2017). Age and input effects in the acquisition of mood in Heritage Portuguese. *Journal of Child Language*, *44*(4), 795–828. https://doi.org/10.1017/S0305000916000222

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: a unit for all reasons. *Applied Linguistics*, *21*(3), 354–375. https://doi.org/10.1093/applin/21.3.354

Gürel, A. (2004). Selectivity in L2-induced L1 attrition: a psycholinguistic account. *Journal of Neurolinguistics*, *17*(1), 53–78. https://doi.org/10.1016/S0911-6044(03)00054-X

Gürel, A., & Yılmaz, G. (2011). Restructuring in the L1 Turkish grammar. Effects of L2 English and L2 Dutch. *Language, Interaction and Acquisition*, *2*(2), 221–250. https://doi.org/10.1075/lia.2.2.03gur

Hakuta, K., & D'Andrea, D. (1992). Some Properties of Bilingual Maintenance and Loss in Mexican Background High-School Students. *Applied Linguistics*, *13*(1), 72–99. https://doi.org/10.1093/applin/13.1.72

Hopp, H., & Schmid, M. S. (2013). Perceived foreign accent in first language attrition and second language acquisition: The impact of age of acquisition and bilingualism. *Applied Psycholinguistics*, *34*(02), 361–394. https://doi.org/10.1017/S0142716411000737

Huls, E., & van de Mond, A. (1992). Some Aspects of Language Attrition in Turkish Families in the Netherlands. In W. Fase, K. Jaspaert, S. Kroon, & ebrary, Inc (Eds.), *Maintenance and loss of minority languages* (pp. 99–117). Amsterdam ; Philadelphia: J. Benjamins.

Hyltenstam, K., Bylund, E., Abrahamsson, N., & Park, H.-S. (2009). Dominant-language replacement: The case of international adoptees. *Bilingualism: Language and Cognition*, *12*(02), 121–140. https://doi.org/10.1017/S1366728908004008

Isurin, L. (2000). Deserted island or a child's first language forgetting. *Bilingualism: Language and Cognition*, *3*(2), 151–167. https://doi.org/null

Jia, G., & Aaronson, D. (1999). Age differences in second language acquisition: The dominant language switch and maintenance hypothesis. In *Proceedings of the 23rd annual Boston University conference on language development* (Vol. 23, pp. 301–312).

Jia, G., & Aaronson, D. (2003). A longitudinal study of Chinese children and adolescents learning English in the United States. *Applied Psycholinguistics*, *24*(01), 131–161. https://doi.org/10.1017/S0142716403000079

Karayayla, T. (forthc.). Effects of first language attrition on heritage language input and ultimate attainment: two generations of Turkish immigrants in the UK. In B. Brehmer, J. Treffers-Daller, & D. Berndt (Eds.), *Lost in Transmission: The role of attrition and input in heritage language development*. John Benjamins.

Kinsella, C., & Singleton, D. (2014). Much More Than Age. *Applied Linguistics*, *35*(4), 441–462. https://doi.org/10.1093/applin/amu032

Köpke, B. (2007). Language attrition at the crossroads of brain, mind, and society. In B. Köpke, M. S. Schmid, M. Keijzer, & S. Dostert (Eds.), *Language Attrition Theoretical Perspectives* (Vol. 33, pp. 9–38). Amsterdam/Philadelphia: John Benjamins.

Köpke, B., & Schmid, M. S. (2004). Language attrition: The next phase. In M. S. Schmid, B. Köpke, M. Keijzer, & L. Weilemar (Eds.), *First Language Attrition interdisciplinary*

perspectives on methodological issues (Vol. 28, pp. 1–45). Amsterdam/Philadelphia: John Benjamins.

Kupisch, T., Lein, T., Barton, D., Schröder, D. J., Stangen, I., & Stoehr, A. (2014). Acquisition outcomes across domains in adult simultaneous bilinguals with French as weaker and stronger language. *French Language Studies*, *24*, 347–376. https://doi.org/doi:10.1017/S0959269513000197

Kupisch, T., & Rothman, J. (2016). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*, 1367006916654355. https://doi.org/10.1177/1367006916654355

Lahmann, C., Steinkrauss, R., & Schmid, M. S. (2016). Factors Affecting Grammatical and Lexical Complexity of Long-Term L2 Speakers' Oral Proficiency. *Language Learning*, *66*(2), 354–385. https://doi.org/10.1111/lang.12151

Lein, T., Kupisch, T., & Weijer, J. van de. (2015). Voice onset time and global foreign accent in German–French simultaneous bilinguals during adulthood. *International Journal of Bilingualism*, 1367006915589424. https://doi.org/10.1177/1367006915589424

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods*, *49*(4), 1494–1502. https://doi.org/10.3758/s13428-016-0809-y

MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk* (3rd ed.). Mahwah NJ: Lawrence Erlbaum Associates.

Montrul, S. (2004). Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and Cognition*, *7*(02), 125–142. https://doi.org/10.1017/S1366728904001464

Montrul, S. (2008). *Incomplete Acquisition in Bilingualism: Re-examining the Age Factor*. John Benjamins Publishing.

Montrul, S. (2016). *The acquisition of heritage languages*. Cambridge: Cambridge University Press.

Montrul, S., Bhatt, R. M., & Bhatia, A. (2012). Erosion of case and agreement in Hindi heritage speakers. *Linguistic Approaches to Bilingualism*, *2*(2), 141–176. https://doi.org/10.1075/lab.2.2.02mon

Montrul, S., & Sánchez-Walker, N. (2013). Differential Object Marking in Child and Adult Spanish Heritage Speakers. *Language Acquisition*, *20*(2), 109–132. https://doi.org/10.1080/10489223.2013.766741

Muñoz, C., & Singleton, D. (2011). A critical review of age-related research on L2 ultimate attainment. *Language Teaching*, *44*(1), 1–35. https://doi.org/10.1017/S0261444810000327

Newport, E. L. (1990). Maturational Constraints on Language Learning. *Cognitive Science*, *14*(1), 11–28. https://doi.org/10.1207/s15516709cog1401_2

Onar Valk, P., & Backus, A. (2013). Syntactic change in an immigrant language: from non-finite to finite subordinate clauses in Turkish. *Eesti Ja Soome-Ugri Keeleteaduse Ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, *4*(2), 7–29. https://doi.org/10.12697/jeful.2013.4.2.01

Pallier, C. (2007). Critical periods in language acquisition and language attrition. In B. Köpke, M. S. Schmid, M. Keijzer, & S. Dostert (Eds.), *Language Attrition Theoretical Perspectives* (Vol. 33, pp. 155–168). Amsterdam/Philadelphia: John Benjamins.

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*(2), 155–161. https://doi.org/10.1093/cercor/13.2.155

Pierce, L. J., Chen, J.-K., Delcenserie, A., Genesee, F., & Klein, D. (2015). Past experience shapes ongoing neural patterns for language. *Nature Communications*, *6*, 10073. https://doi.org/10.1038/ncomms10073

R Core Team. (2016). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. Vienna Austria. Retrieved from https://www.R-project.org/.

Rothman, J. (2009). Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism*, *13*(2), 155–163. https://doi.org/10.1177/1367006909339814

Sak, H., Güngör, T., & Saraçlar, M. (2009). A stochastic finite-state morphological parser for Turkish. In *Proceedings of the ACL-IJCNLP 2009 Conference short papers* (pp. 273–276). Suntec Singapore: Association for Computational Linguistics. Retrieved from http://dl.acm.org/citation.cfm?id=1667667

Schmid, M. S. (2002). *First Language Attrition, Use and Maintenance : The Case of German Jews in Anglophone Countries*. Amsterdam: J. Benjamins Pub.

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* (Vol. 33, pp. 135–153). Amsterdam/Philadelphia: John Benjamins.

Schmid, M. S. (2012). The impact of age and exposure on bilingual development in international adoptees and family migrants: A perspective from Holocaust survivors. *Linguistic Approaches to Bilingualism*, *2*(2), 177–208. https://doi.org/10.1075/lab.2.2.03sch

Schmid, M. S. (2013). First language attrition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(2), 117–123. https://doi.org/10.1002/wcs.1218

Schmid, M. S. (2014). The Debate on Maturational Constraints in Bilingual Development: A Perspective from First-Language Attrition. *Language Acquisition*, *21*(4), 386–410. https://doi.org/10.1080/10489223.2014.892947

Schmid, M. S., & Dusseldorp, E. (2010). Quantitative analyses in a multivariate study of language attrition: the impact of extralinguistic factors. *Second Language Research*, *26*(1), 125–160. https://doi.org/10.1177/0267658309337641

Schmid, M. S., & Köpke, B. (2017). The relevance of first language attrition to theories of bilingual development. *Linguistic Approaches to Bilingualism*. https://doi.org/10.1075/lab.17058.sch

Schmid, M. S., Köpke, B., & Bot, K. de. (2012). Invited contribution: Language attrition as a complex, non-linear development. *International Journal of Bilingualism*, 1367006912454619. https://doi.org/10.1177/1367006912454619

Schmitt, E. (2004). No more reductions!- To the problem of evaluation of language attrition data. In M. S. Schmid, B. Köpke, M. Keijzer, & L. Weilemar (Eds.), *First Language Attrition interdisciplinary perspectives on methodological issues* (Vol. 28, pp. 299–316). Amsterdam/Philadelphia: John Benjamins.

Seliger, H. W., & Vago, R. M. (1991). The study of first language attrition: An overview. In H. W. Seliger & R. M. Vago (Eds.), *First language attrition* (pp. 3–16). Cambridge: Cambridge University Press.

Silva-Corvalán, C. (1994). The gradual loss of mood distinctions in Los Angeles Spanish. *Language Variation and Change*, *6*(03), 255–272. https://doi.org/10.1017/S095439450000168X

Slobin, D. I. (1986). The acquisition and use of relative clauses in Turkic and Indo European languages. In D. I. Slobin & K. Zimmer (Eds.), *Studies in Turkish linguistics* (pp.

273–294). Amsterdam ; Philadelphia: J. Benjamins Pub. Co. Retrieved from http://site.ebrary.com/lib/universityofessex/Doc?id=10509647

Sorace. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism*, *1*(1), 1–33. https://doi.org/10.1075/lab.1.1.01sor

Stangen, I., Kupisch, T., Proietti Ergün, A. L., & Zielke, M. (2015). Foreign accent in heritage speakers of Turkish in Germany. In H. Peukert (Ed.), *Transfer effects in multilingual language development* (pp. 87–107). Amsterdam: John Benjamins Publishing Company.

Treffers-Daller, J., Ozsoy, A. S., & van Hout, R. (2007). (In)Complete Acquisition of Turkish among Turkish-German Bilinguals in Germany and Turkey: An Analysis of Complex Embeddings in Narratives. *International Journal of Bilingual Education and Bilingualism*, *10*(3), 248–276. https://doi.org/10.2167/beb443.0

Ventureyra, V. A. G., Pallier, C., & Yoo, H.-Y. (2004). The loss of first language phonetic perception in adopted Koreans. *Journal of Neurolinguistics*, *17*(1), 79–91. https://doi.org/10.1016/S0911-6044(03)00053-8

Veríssimo, J., Heyer, V., Jacob, G., & Clahsen, H. (2017). Selective Effects of Age of Acquisition on Morphological Priming: Evidence for a Sensitive Period. *Language Acquisition*, *0*(0), 1–12. https://doi.org/10.1080/10489223.2017.1346104

Wurm, L. H., & Fisicaro, S. A. (2014). What residualizing predictors in regression analyses does (and what it does not do). *Journal of Memory and Language*, *72*(Supplement C), 37–48. https://doi.org/10.1016/j.jml.2013.12.003

Xi, X., & Mollaun, P. (2011). Using Raters From India to Score a Large-Scale Speaking Test. *Language Learning*, *61*(4), 1222–1255. https://doi.org/10.1111/j.1467-9922.2011.00667.x

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*(02), 131–149. https://doi.org/null

Yılmaz, G. (2011). Complex embeddings in free speech production among late Turkish-Dutch bilinguals. *Language, Interaction and Acquisition*, *2*(2), 251–275. https://doi.org/10.1075/lia.2.2.04yil

Yılmaz, G. (2013). *Bilingual Language Development among the First Generation Turkish Immigrants in the Netherlands*. University of Groningen, Groningen. Retrieved from http://www.let.rug.nl/languageattrition/Papers/Yilmaz%20dissertation%202013.pdf

Young, R. (1995). Conversational Styles in Language Proficiency Interviews. *Language Learning*, *45*(1), 3–42. https://doi.org/10.1111/j.1467-1770.1995.tb00961.x

Zaretsky, E., & Bar-Shalom, E. G. (2007). Is selective attrition possible in Russian-English bilinguals? In *BUCLD Online Proceedings Supplement* (Vol. 32).

Table 1: External variables

	interactive L1 use	L1 passive exposure	past L1 use	linguistic identification	cultural affiliation
alpha	0.789	0.649	0.573	0.779	0.603
mean	0.86	0.58	0.96	0.934	0.633
range	0.41–1.00	0.00–1.00	0.65–1.00	0.5–1.00	0.18–0.88
SD	0.125	0.96	0.084	0.115	0.175

Table 2: Basic background information

		age at testing	AaO	length of residence	education in years
controls (n=29)	Mean	32	N/A	N/A	13.44
	St. Dev.	9.75	N/A	N/A	2.49
	range	21–51	N/A	N/A	8–15
bilinguals (n=57)	Mean	35.14	15.54	19.61	13.15
	St. Dev.	7.81	7.48	6.72	2.27
	range	19–58	7–34	9–40	8–15

Table 3: Grammatical complexity sub-variables across groups

	groups	mean	std. dev
agglutination index	controls	0.8253	0.04611
	bilinguals	0.8293	0.05315
number of *-mA* per AS Unit	controls	0.0484	0.02555
	bilinguals	0.0508	0.0323
number of *-DIK* per AS Unit	controls	0.032	0.02054
	bilinguals	0.0363	0.0298
number of *-AcAK* per AS Unit	controls	0.0073	0.00916
	bilinguals	0.0069	0.00809
number of relative clauses per AS Unit	controls	0.0817	0.03653
	bilinguals	0.0689	0.05514

Table 4: Predictors of FAR across bilingual groups

	estimate	SE	t-value	
(Intercept)	1.487366751	0.18921	7.86104	***
AaO	-0.01306183	0.00491	-2.6602	*
C-Test	-0.01360869	0.00463	-2.9407	**
L1 pass exp.	-0.28758268	0.13312	-2.1603	*
age at testing	-0.00991208	0.00463	-2.1426	*