On L1 attrition and the linguistic system[1]

Monika S. Schmid, University of Groningen

Abstract

One of the most puzzling observations for linguists is the difference between learning a language from birth and later in life: while all normally developing children can attain full native language proficiency, there is considerable variability in ultimate attainment among older speakers who attempt to acquire a second language (L2). There is an ongoing controversy in linguistic research on whether this discrepancy is due to a maturationally constrained window of linguistic development making language learning difficult or impossible after puberty, or to general cognitive factors linked to the fact that the later an L2 is established, the stronger the competition it has to overcome from the more deeply entrenched first language (L1).

   Studies attempting to resolve this controversy have so far focussed exclusively on the development of L2 skills. New insight may be provided by investigating the first language skills of migrants who have become dominant in the L2 (referred to as L1 attriters). Such speakers have learned their L1 as monolinguals during childhood, and were therefore not impeded by maturational constraints in the acquisitional process. Having lived in an L2 environment for a long period of time, however, their seldom-used L1 shows signs of the influence of their highly active L2.

   A systematic comparison of L1 attriters and L2 learners may therefore be able to shed some light on the question of whether there is a qualitative or merely a quantitative difference between L1 acquisition in childhood and L2 acquisition later in life: If being a native speaker is maturationally constrained, even attrited L1 systems should remain native-like. But if the persistent problems of L2 learners are due to issues such as lack of practice and exposure and

competition between their two language systems, bilinguals who use their second language dominantly should become more similar to L2 speakers.

Introduction

There is a long-standing bias in research on bilingualism towards investigations of later-learned and non-dominant languages. It is assumed that to determine whether these languages are represented or used differently from what we can observe in native speakers will allow us to better understand the nature of language learning or language use, and may eventually provide us with fundamental insights into the human mind. There is good reason for this preference: speakers who have acquired a language sequentially, after the first one has been established, often use it in ways which are persistently different from what can be observed in monolinguals. Most foreign language learners never quite reach fully native speaker levels of proficiency and show variability in the application of some rules or features (e.g. Sorace 2005). They are slower at completing some tasks (such as speeded grammaticality judgments, see the overview in Hulstijn 1997:138) and they may be unable to reliably identify non-targetlike sentences (Birdsong 2006). Some features, for example grammatical gender, remain particularly problematic and these difficulties appear to be more persistent the older the learner is when the second language (L2) was first acquired (Cornips and Hulk 2008, Orgassa and Weerman 2008).

  Indications of apparent differences between L2 and L1 speakers may, however, be deceptive: typically, investigations of SLA compare their target populations with native speakers who are largely or entirely monolingual, making the baseline for comparison highly problematic. Bilingual speech production and processing are cognitively more costly due to the effort needed to manage two systems of linguistic knowledge, particularly since it has been shown that all language systems of a multilingual are always active to some extent (e.g.

Kroll, Bobb, Misra and Guo 2008), and that bilingual processing may be more controlled (i.e. less automatic) than is the case for monolinguals (Abutalebi 2008). These factors may account for some of the differences observed between bilingual and monolingual populations which are often ascribed to age at acquisition.

   In order to shed more light on this issue, it may be helpful to investigate a different population: native speakers, who have acquired their L1 implicitly and as monolinguals during childhood, and therefore possess (or have possessed) native grammatical representations and processing strategies for this language, but who experience the same competition from another language and have to exercise the same control associated with bilingual processing as the L2 populations investigated in SLA research. Should such L1 attriters differ from monolinguals in the same way as L2 users do, that would provide compelling evidence for a straightforward bilingualism effect and against theories which assume L2 acquisition to be governed by a critical period and L2 knowledge to be fundamentally different from L1 knowledge.

   Such a population can be found in speakers who were raised in a monolingual environment, but who emigrated at some point in their lives and use the language learned in emigration on a regular (or predominant) basis. The linguistic development observed in such contexts has been termed first language attrition.

1. The critical period in second language acquisition

There is an unresolved controversy about whether the fact that older L2 learners typically achieve lower ultimate levels of proficiency is due to maturationally constrained biological processes, or to an interaction of general cognitive development and external factors.[2] The issue at stake is whether post-puberty L2 learners' underlying representations of the L2 system and the cognitive processes they apply to using and understanding language are

essentially different from or essentially the same as those of L1 speakers.[3] This controversy is reflected in theoretical accounts of SLA, in particular within those theoretical frameworks which are in some way based on assumptions about the neurocognitive foundations of language learning, knowledge, use and processing. The following sections will briefly explore the controversy about maturational constraints in L2 learning within two such frameworks: generative linguistics and psycho-/neurolinguistics.

Theories from a UG perspective which assume an innate basis for language learning are divided on the question of whether access to the inborn knowledge available to L2 learners in the learning of subsequent languages is subject to some maturational constraint or not. Empirically resolving this matter has proven extremely difficult, since non-targetlike behaviour of L2 speakers almost invariably lends itself to interpretation supporting both points of view – it is very hard to establish whether an 'error' is evidence of a non-targetlike representation or reflects a failure to map the native-like representation to the appropriate surface form due to cognitive limitations, L1 transfer, or other external factors (e.g. Hakwins 2001, Prévost & White 2000).

This problem is aggravated by the nature of the data: the findings reported are often based on offline experiments, such as untimed grammaticality judgments. In such data, it is difficult to establish whether correct responses of an L2 speaker are the result of consciously working out the problem on the basis of explicit grammatical knowledge, or of truly native-like processing strategies.

The controversy on similarities or differences between L1 speakers and (post-puberty) L2 learners is reflected in approaches to linguistic investigation which focus not so much on (the innate basis of) grammar building and the representation of knowledge, but on online processing of linguistic input and its neurocognitive correlates. Here, the competing viewpoints are divided between approaches which assume that post-puberty L2 learners use

processing strategies which are different from those of early learners or native speakers and those which assume that non-native speakers are constrained by more general demands, such as a higher demand on control mechanisms in bilingual processing and resulting limitations of working memory.

Clahsen and Felser (2006) assume that adult L2 learners experience difficulties concerning the integration of information in online processing. They argue that while mature L1 speakers rapidly integrate different levels of linguistic information (lexical, discourse-level, prosodic and structural) when exposed to input, L2 learners overrely on lexical-semantic information and are insensitive to structural syntactic information present in the input – their processing is 'shallower'. Such nonnative-like processing strategies persist in highly advanced learners who are able to perform at native level on a number of behavioural measures (such as grammaticality judgments or language proficiency tests).

A different approach to the neurocognitive basis of L1 and L2 knowledge hinges on the well-established concept that there are different memory systems, subserved by different brain regions: declarative memory, in which (implicitly and explicitly learned) facts and events are represented, and procedural memory, which predominantly concerns skills and habits that have been acquired implicitly (Squire 1992). These two memory systems both play an important part for language. Very broadly speaking, native speakers rely on declarative memory for semantic knowledge and on procedural memory for grammatical knowledge (Paradis 2004, Ullman 2001). It has been argued that adult L2 learners use these memory systems differently: they acquire grammatical knowledge explicitly due to their more highly developed cognitive and analytical skills, and this knowledge is therefore initially based on declarative memory. With increasing skill, it is assumed that more proficient L2 speakers can also build procedurally-based representations for (parts of) grammatical knowledge.

On this view, there is therefore in principle no reason why L2 speakers should not be able to perform in a native-like way on all aspects of the linguistic system. It is, however, assumed that, even when they are behaviourally indistinguishable from native speakers, the memory systems on which this behaviour is based are different from L1 language use. Occasional transgressions from the native target are then due to cognitive limitations of controlling the more explicitly represented knowledge in online speech production or processing.[4]

Critics of the shallow processing hypothesis and of the declarative/procedural approach seek the reasons for the differences found between native and non-native populations in more general principles. One of the main drawbacks pointed out with respect to the studies advanced in support of Clahsen and Felser's argument is that they compare the performance of L2 speakers – that is, of bilingual populations – to monolingual L1 speakers (e.g. Carroll 2006). This makes the baseline for these studies (and, indeed, for most other investigations of SLA) problematic: It has been amply demonstrated that a bilingual individual is not two monolingual individuals in the same mind/person. Once a speaker has acquired two (or more) languages, all of them will to some degree be active and accessed during language processing, and a return to a completely monolingual mode is impossible (e.g. Dijkstra and van der Heuven 2002; Grosjean 2001).

Consequently, bilingual processing can always be assumed to draw more strongly on the coomputational resources and to be more strongly affected by constraints on working memory than monolingual processing. It has recently been shown that monolingual speakers who effortlessly outperform L2 learners on a timed grammaticality judgment task where sentences are presented one word at a time on a computer screen start to converge towards the accuracy scores of the non-natives when the cognitive load is increased through shorter and shorter presentation of each individual word (Hopp 2008). In other words, when the processing demands are increased, monolinguals may also start behaving in a non-native-like way.

Observed differences between L1 and L2 populations may therefore not be down to fundamental differences in representation and processing of the two linguistic systems, but to cognitive demands and limited resources.

The question of the 'difference' or 'identity' of acquisition strategies, knowledge representation and processing mechanisms of pre- and postpuberty language learners to date remains one of the most controversial and most important issues in studies on bilingualism. Researchers cannot directly observe either linguistic knowledge or linguistic processing. They can observe how a speaker or learner reacts to a given stimulus or completes a given task. They may use these results in order to induce with what underlying knowledge or processing strategies the responses seem to be in accordance. However, the findings will almost invariably be compatible with different explanations (and/or open to methodological criticisms).

A potential resolution of this question may be found by abandoning the traditional distinction between L1 and L2 in favour of an investigation of dominant vs. non-dominant languages. Investigations of the non-dominant and lesser-used language of bilinguals should therefore include bilinguals who have acquired their weaker language in different settings and at different ages. This approach allows for a comparison of L2 learners on the one hand and speakers who acquired a language implicitly as monolinguals during childhood but became dominant in a later learned language as adults on the other. The latter type of bilingualism is known as first language attrition.

2. First language attrition as a testing ground for CP hypotheses

The term 'first language attrition' (FLA) refers to a gradual decline in native language proficiency among migrants. As a speaker uses his/her L2 frequently and becomes proficient in it, some aspects of the L1 can become subject to L2 influence or deteriorate due to the

absence of confirming evidence in the input. Like SLA, FLA is mediated by a number of external factors, such as exposure and use (e.g. Hulsen 2000; Schmid 2007), attitude and motivation (Ben-Rafael and Schmid 2007; Schmid 2002) or aptitude (Bylund 2008). However, the overall impact of these factors on ultimate proficiency or fossilization appears to be far less strongly pronounced than what has been found in SLA (Schmid and Dusseldorp forthc.).

Furthermore, research on L1 attrition has so far found an astonishingly small range of variability and a low incidence of non-targetlike use in data even from speakers who claim not to have used their L1 for many decades (in some cases upwards of 60 years, e.g. de Bot and Clyne 1994, Schmid 2002), provided they emigrated after puberty. If, on the other hand, environmental exposure to the L1 ceases before puberty, the L1 system can deteriorate radically. While there are few principled and systematic investigations of FLA specifically investigating the impact of age at onset (AaO), converging evidence suggests the possibility of an age effect which is much stronger and more clearly delineated than the effects which have been found in SLA research. Two studies which consider pre- and postpuberty migrants (Ammerlaan 1996, AaO 0-29 yrs; Pelc 2001, AaO 8-32 years) find that AaO is one of the most important predictors of ultimate proficiency. On the other hand, a number of investigations considering the impact of age among postpuberty migrants fail to find any effect whatsoever on a range of formal tests and in free speech (Köpke 1999, AaO 14-36 yrs; Schmid 2002, AaO 12-29 yrs; Schmid 2007, AaO 17-51 yrs). These findings indicate that early (pre-puberty) and late (post-puberty) L1 attrition have a substantially different impact on possible fossilization and/or deterioration of the linguistic system (for more detailed overviews see Köpke 2004; Köpke and Schmid 2004).

A recent investigation, focussing specifically on the age effect in L1 attrition, lends further substantiation to the assumption that there is a qualitative change around puberty: Bylund

(2008) investigates the L1 of 31 Spanish speakers who emigrated to Sweden between the ages of 1 and 19 years and concludes that "there is a small gradual decline in attrition susceptibility during the maturation period followed by a major decline at its end (posited at around age 12)" (Bylund 2008:22). No further age effects were found among the later attriters.

The strongest indication that an L1 can be extremely vulnerable to attrition if exposure ceases before puberty comes from a study of Korean adoptees in France, who had been between 3 and 10 years old when they were adopted by French-speaking families (Pallier et al. 2003; Ventureyra and Pallier 2004). This investigation could find no trace of L1 knowledge on a range of speech identification and recognition tasks, nor did an fMRI study reveal any differences in brain activation when these speakers were exposed to Korean as opposed to unknown languages (Japanese or Polish). In all respects, the Korean adoptees presented in exactly the same way as the French controls.

It therefore appears that a basic distinction needs to be applied with respect to the classification of bilinguals and language attrition: Depending on the age at which the speaker was removed from the environment in which the language under observation was spoken, bilinguals can be classified as either 'heritage speakers' (speakers who were born to families speaking a language which is different from the language of the environment), 'incomplete learners' (speakers for whom the emigration took place before the age of 12) and 'L1 attriters' (speakers who emigrated at an age above 12).

2.1. Generative accounts of L1 attrition

Generative approaches to L1 attrition often focus on the possibility that the developing linguistic system may show evidence of irrevocable *structural* changes to the actual grammar of a native language. This was highlighted early on in the history of attrition research: "It is crucial to know whether a given example of language loss can be attributed to a change in

how the relevant language is represented in the mind of the user or to a change in the way stable knowledge (competence) is being *used*." (Sharwood Smith 1983:49, his emphasis). In a similar vein, Seliger and Vago define the object of investigation as "the disintegration or attrition of the *structure* of a first language (L1) in contact situations with a second language (L2)" (Seliger and Vago 1991:3, my emphasis). Here, too, the point is made that:

> effects of performance (accessing, processing, control) need to be sorted out from those of competence (tacit knowledge): it is erosion that reaches the level of competence that allows for interesting claims about and meaningful insight into the attrition process. (Seliger and Vago 1991:7)

On this view, the object of attrition research is therefore to determine whether the attriting or attrited linguistic system may show evidence of a representational deficit akin to the one assumed to obtain for non-native speakers, and (if so) how this deficit is constrained by the nature of the human language faculty. This question is at the heart of investigations of L1 attrition from a generative perspective (e.g. Gürel 2002; Montrul 2008; Tsimpli et al. 2004). The tentative conclusion from such investigations appears to be that attrition does not affect uninterpretable features, but that variability may be observed in features that are interpretable at the interface levels (Tsimpli et al. 2004:274; Tsimpli 2007: 85). There therefore seems to be little evidence for an actual restructuring of the language system: the narrow syntax remains unaffected, and the observed variability may be ascribed to the cognitive demands of bilingual processing.

  Most of the studies on the development of the L1 system from a generative perspective investigate speakers for whom the age at onset is located after puberty. The only evidence comparing heritage speakers, incomplete learners and L1 attriters comes from the work of Montrul (2002, 2004, 2008), who suggests that the L1 system of early bilinguals may be similar to that of L2 speakers, while speakers for whom the onset of L2 learning took place

later in childhood pattern with monolinguals in their L1. These findings again appear to point towards a critical period effect for L1 loss, beyond which actual restructuring to the syntactic module will no longer take place.

2.2. Processing and neurocognitive approaches to L1 attrition

Research on L1 attrition has so far availed itself only in very limited ways of recent methodological and experimental developments made in investigations on the neurocognition of language. Online measurements, such as the reaction time paradigm, have been mainly used for investigations of the accessibility of the mental lexicon in L1 attrition (Ammerlaan 1996; Hulsen 2000), and few experimental studies take a psycholinguistic approach to grammatical processing (a notable exception is Köpke's 1999 study). Measurements such as Event-Related Potentials, eyetracking or fMRI, which are today widely used in investigations of SLA, have been largely neglected in L1 research (the only exceptions being the investigation of Korean adoptees mentioned above and Dussias' (2004) study using eyetracking to investigate relative clause attachment in the L1 and L2 of Spanish-English bilinguals).

  It is therefore difficult to draw conclusions as to the processing strategies and memory systems used by L1 attriters for the production and interpretation of grammatical information. Given the debates presented above, the problem is an intriguing one: do L1 attriters come to overrely on semantic information or on declarative (as opposed to procedural) memory in a way which is similar to what has been observed in late L2 learners? Or are the interferences and non-target like responses elicited by a variety of offline tasks documented in the existing literature on L1 attrition mainly the outcome of localized and momentary processing problems or L2 effects on L1, due to cognitive overload caused by the coordination of two linguistic systems? Various theoretical frameworks provide testable hypotheses which could be

corroborated or falsified by means of neurocognitive methods. Such investigations are clearly called for.

2.3. Conclusion

A comparison of L2 learners and L2 attriters for whom the onset of bilingualism (the beginning of exposure to the L2 in the case of the learners, and the leaving of the L1 environment in the case of the attriters) took place beyond the critical period may shed light on questions about the nature of bilingual knowledge and processing: L1 attriters have acquired their L2 in exactly the same way as the monolingual populations typically used as a baseline for comparison in investigations of L2 speakers. If there are indeed representational deficits or differences in processing for late bilinguals due to a maturationally constrained critical period, L1 attriters should pattern together with monolingual reference groups. If, on the other hand, the observed behavioural differences are caused by a more general bilingualism effect conditioned by the limitation of available cognitive resources, L1 attriters and L2 speakers should be similar to each other and differ from the monolinguals.

The present paper will attempt to take a first step in this direction by combining findings from a quantitative investigation of L1 attrition and a case-study of an attriter and an L2 learner. The following research questions will be addressed:

RQ1. *Attrition on controlled tasks*

Do control group language users (predominantly monolingual non-migrants) outperform attriters (long-term migrants) on controlled linguistic tasks in their L1? Specifically, do they perform better on tasks measuring

a) overall proficiency

b) lexical access

c) the ability to identify unacceptable grammatical structures?

RQ2. *Attrition in free speech*

Does free speech elicited from attriters differ from that of controls? Specifically, do

controls have

a) a higher degree of lexical richness

b) a higher ratio of variability in the syntactic structures they employ, and

c) better overall accuracy

than attriters? Furthermore,

d) are attriters perceived to be less native-like in their accent by native listeners?

RQ3. *Structural attrition or online control?*

If evidence for attrition is found in answer to RQs 1 and 2, is it possible to determine

whether this attrition is indicative of representational deficits vs. online problems in the

integration of intact underlying knowledge?

RQ4. *L1 or L2?*

Are L1 attriters fundamentally more similar to non-attrited L1 speakers, or to highly

proficient L2 speakers?

3. Quantitative study

For the purpose of the present study a combined approach was adopted. Data from a large-

scale quantitative investigation of the L1 attrition of German in an English- and a Dutch-

speaking setting (see below) were used as a background, against which qualitative findings

from a case study of two speakers were interpreted.

3.1. Participants:

The quantitative study is based on three populations: German migrants in Vancouver, Canada

(L2 English), German migrants in the Netherlands (L2 Dutch) and a largely monolingual

German reference group in Germany. As far as possible, age, sex, level of education, age at migration and length of residence were controlled across groups (see Table 1, see also Schmid 2007, Schmid and Dusseldorp forthc.).

Table 1: Overview of participant characteristics

	bilingual groups (attriters)				reference group (controls)	
	GECA: Germans in Canada (n=53)		GENL: Germans in NL (n=53)		GECG: Control group (n=53)	
	mean	stdev	mean	stdev	mean	stdev
Age at experiment	63.23	10.92	63.36	9.55	60.89	11.60
Age at migration (min. 17 yrs)	26.13	7.15	29.08	7.53	-	-
Length of residence (min. 15 yrs)	37.09	12.37	34.28	11.13	-	-

The inclusion of Dutch and English as L2 allowed for some interesting comparisons between the groups, based on similarities and differences between the L1 and the L2. German and Dutch are closely related in both lexicon and grammar. However, a number of morphological characteristics are more complex in German than they are in Dutch, most notably in the noun phrase.

While grammatical gender is an inherent property of all nouns in both languages, the German and the Dutch system differ from each other both in inherent gender and in agreement marking: German has three grammatical genders (masculine, feminine and neuter), Dutch has two (common and neuter). Furthermore, the German inflectional paradigm (Table 2) is richer than the Dutch one (Table 3). German marks gender, definiteness and case by inflection on determiner and adjective, while Dutch has only two forms of the definite article, one form of the indefinite article and two forms of inflection on the adjective.

Table 2: The inflectional paradigm of German NPs (*der grosse Berg* (masc.) 'the big mountain', *die grosse Strasse* (fem.) 'the big street', *das grosse Tal* (neut.), 'the big valley')

Case	Masculine definite	Masculine indefinite	Feminine definite	Feminine indefinite	Neuter definite	Neuter indefinite
N	**der** grosse Berg	**ein** grosse**r** Berg	**die** grosse Strasse	**eine** grosse Strasse	**das** grosse Tal	**ein** grosse**s** Tal
G	**des** grosse**n** Berg**es**	**eines** grosse**n** Berg**es**	**der** grosse**n** Strasse	**einer** grosse**n** Strasse	**des** grosse**n** Tal**es**	**eines** grosse**n** Tal**es**
D	**dem** grosse**n** Berg	**einem** grosse**n** Berg	**der** grosse**n** Strasse	**einer** grosse**n** Strasse	**dem** grosse**n** Tal	**einem** grosse**n** Tal
A	**den** grosse**n** Berg	**einen** grosse**n** Berg	**die** grosse Strasse	**eine** grosse Strasse	**das** grosse Tal	**ein** grosse**s** Tal

Table 3: The inflectional paradigm of Dutch NPs (*de grote berg* (common) 'the big mountain', *het grote dal* (neut.), 'the big valley')

	Common definite	Common indefinite	Neuter definite	Neuter indefinite
Sg.	**de** grote berg	**een** grote berg	**het** grote dal	**een** grootØ dal

Neither Dutch nor German mark gender morphologically in the plural. However, German has five different plural allomorphs, some of which can combine with vowel mutation (umlauting) of the stem vowel (e.g. Schmid 2002), while Dutch has two productive plural allomorphs (e.g. Keijzer 2007). English, on the other hand, has no gender agreement or case marking (except on pronouns). Definiteness is marked only on the article, and plural only on the noun. Productive plural allomorphy is limited and phonologically transparent.

Where verb phrase morphology is concerned, both German and Dutch inflect verbs for person and number and form their tenses either by vowel mutation (strong inflectional paradigm) or by suffixation (weak inflectional paradigm), as they do in English. Furthermore, there are two tense auxiliaries (*have* and *be*) as opposed to the single English auxiliary *have*.

In the domain of word order, German and Dutch main clauses follow the Germanic verb second rule: if any constituent other than the subject is fronted in the clause, the subject has to appear after the verb. Furthermore, the finite and the non-finite part of the verb (infinitive, participle, particle) 'frame' other constituents, a feature known as discontinuous word order

(DWO). Like in many other Germanic languages, the verb appears at the end of subordinate clauses. With respect to this last point, there is some variation between German and Dutch: In German subordinate clauses which contain a finite auxiliary and a non-finite lexical verb (infinitive or participle), the auxiliary obligatorily follows the lexical verb. In Dutch, this order may vary, with a preference for the auxiliary preceding the lexical verb.

  German, Dutch and English are thus situated on a cline of morphosyntactic complexity (see Table 4). The populations represented in this study therefore allow for a comparison of bilinguals with varying degrees of competition between their L2 and their L1 systems.

Table 4: A comparison of morphosyntactic complexity across the language considered in this study

			German	Dutch	English
Morphology	NP	Case	✓	✗	✗
		Gender	✓	✓	✗
		Plural	✓	✓	✓
	VP	Person	✓	✓	✓
		Strong/weak	✓	✓	✓
		Auxiliaries	✓	✓	✗
Syntax		V2 in MCl	✓	✓	✗
		DWO in MCl	✓	✓	✗
		V_{LAST} in SubCl	✓	✓	✗

✓ = feature overtly present, ✗ = feature not overtly present

3.2. Experimental tasks

All participants completed the following tasks:

1. A C-TEST (CT, see Grotjahn 1987). The C-TEST is a fill-in test where the subject is presented with a text from which parts of words have been removed following a pre-determined schema and asked to complete the missing parts. Each text was between 80

and 100 words in length and contained 20 gaps. The C-TEST score was computed as the number of times a gap was filled in correctly, a high score on the C-TEST reflects high proficiency, with a possible range of 0-100.

2. Two semantic verbal fluency (VF) tasks, where participants are asked to name as many items in a specific lexical category as they can within the space of 60 seconds (Roberts and Le Dorze 1997) . The two stimuli used were 'animals' on the one hand and 'fruit and vegetables' on the other. The final VF measure was an averaged measure of the score on the two individual tasks. A high score on the VF task reflects high proficiency.

3. An untimed Grammaticality Judgment task (GJ), in which participants were presented with sentences on a laptop computer in written and audio format simultaneously. The overall test consisted of 48 items, of which 22 were ungrammatical. The ungrammatical items were classified into various grammatical categories, based on mistakes that had frequently occurred in the data from a previous investigation of L1 attrition of German (Schmid 2002). If the subject had identified a sentence as ungrammatical, s/he was then asked to indicate what the correct alternative would be. Every target-like response on an ungrammatical item was awarded one point, so that the maximum total score on this task was 22, reflecting a high proficiency. This score was then recalculated to a score between 0 (0% correct) and 1 (100% correct).

3.3. Free speech

In addition to these tasks, two sets of free speech samples were gathered. The first set was elicited through a semi-structured autobiographical interview. These interviews were typically 30 – 45 min. in length and measured on average around 2,500 words spoken by the participant. A set of more controlled, largely monological, speech samples was produced by means of the Charlie Chaplin film retelling task described by Perdue (1993), which involves

watching and then narrating a 10-minute excerpt from the silent movie *Modern Times*. These retellings were typically around 10-15 min long and measured on average 1,500 words.

The quantitative analysis of free speech for the purpose of gaining insight into language attrition processes is a methodological challenge, as the distribution of variables not only depends non-trivially on text length (for example with respect to type-token ratios, which tend to decrease non-linearly in longer texts due to the higher frequency of function items) but also on the overall number of obligatory contexts (Schmid 2004). Keeping this in mind, the following variables were established:

	interview	film retelling
Lexical richness: D		
D is a measure of type-token ratios based on random sampling of stretches of 50 words, i.e. it is not sensitive to variation in text length (see McKee, Malvern and Richards, 2000). A high score reflects low type-token ratios, i.e. more lexical diversity.	√	√
non-targetlike language use:		
total errors, standardized per 1,000 words (ERR)	√	√
morphological errors, standardized per 1,000 words (ERRMOR)	√	√
- violations of noun phrase agreement rules (ERR NP)	√	√
- violations of gender agreement rules (ERR GEN)	√	√
- violations of verb phrase agreement rules (ERR VP)	√	√
violations of obligatory verb placement rules, standardized per 1,000 words (ERRSYN)	√	√
overall distribution of syntactic patterns		
% main clauses of the structure XVS, i.e. the verb is preceeded by non-subject item and subject appears after verb according to the Germanic verb-second rule (SVMC)		√
% subordinate clauses (SUBCL)		√

In addition, for those cases where the recording was of sufficient quality, a foreign accent rating (FAR) was established. The FAR is based on ratings by between 19 and 21 native judges per speaker on a 6 point Likert scale (where 6 is definitely non-native and 1 is definitely native; the experiment is described in detail in de Leeuw, Schmid and Mennen, forthc.). For the purpose of the present analysis, the experiment was repeated with 18 additional speech samples. FARs were obtained for 78 speakers: 36 Germans in Canada, 24 Germans in The Netherlands, 17 speakers from the reference group.

3.4. Results

<u>Experimental tasks</u>

Group differences on the experimental tasks were established by means of One-way ANOVAs, which revealed some group differences on the CT and on the VF task, but not on the grammaticality judgment task (see Table 5). Effect sizes (η^2) were rather small, suggesting that while the attriters did, on the whole, achieve lower scores on these two tasks, the group differences were hardly dramatic.

Table 5: Group differences on experimental tasks (One-way ANOVA)

		GECA	GENL	GECG	F (2, 156)	p	η^2
CT	mean	75.26*	77.21	82.21	5.025	.008	.06
	sd.	11.61	13.86	8.90			
VF	mean	20.24*	20.91*	25.09	16.943	.000	.18
	sd.	4.62	4.68	4.67			
GJ	mean	0.82	0.83	0.84	.661	.518	.01
	sd.	0.12	0.10	0.11			

*: difference from control group < .05, **: difference from control group < .001

In response to Research Question 1, the attriters in the population under investigation were indeed outperformed on the C-Test, which is assumed to measure overall proficiency (RQ 1a), and on the verbal fluency task, which is assumed to measure lexical retrieval (RQ 1b). No evidence was found for impaired grammaticality judgments among the attriting groups (RQ 1c).

Free speech

Where overall measures of lexical richness and syntactic diversity were concerned, there was little evidence of attrition in the free speech samples collected. The only variable on which there was a significant difference between attriters and controls was lexical diversity in the film retelling tasks, and the post-hoc procedure reveals that this difference only applies for the GENL group. The overall group comparison for the amount of main clauses in which an element other than the subject preceded the verb was significant, but the Post-Hoc test revealed that this was due to a difference between the GECA and the GENL group; neither group differed from the controls.

   On the other hand, the attriters were consistently different from the controls on all categories of non-target like constructions in both speech samples (with a somewhat higher η^2 in the interview) and furthermore appear to be perceived as less native-like by German listeners, as evidenced by their significantly higher FAR.

Table 6:Group differences on free speech measures (One-way Anova)

			GECA	GENL	GECG	F (2, 156)	p	η^2
Interview	D	mean	98.91	95.34	94.95	1.459	.236	.02
		sd.	13.83	12.56	12.53			
	ERR	mean	7.94*	7.24*	2.71	49.445	< .001	.40
		sd.	4.09	2.58	1.35			
	ERRMOR	mean	2.23**	3.25**	1.11	34.394	< .001	.31
		sd.	1.32	1.57	.95			

ERR NP	mean	1.31**	1.46**	0.44	27.063	< .001	0.51
	sd.	0.81	0.91	0.51			
ERR GEN	mean	0.56**	0.71**	0.24	15.288	<.001	0.41
	sd.	0.37	0.53	0.41			
ERR VP	mean	0.85	1.32**	0.58	11.963	< .001	0.37
	sd.	0.73	0.93	0.62			
ERRSYN	mean	1.18**	.46*	.15	39.809	< .001	.35
	sd.	.93	.40	.28			
Film retelling D	mean	7.45	63.93*	75.35	5.873	.003	.07
	sd.	17.12	15.67	17.90			
ERR	mean	9,11*	9.30*	1.80	18.070	< .001	.19
	sd.	5.91	1.99	2.56			
ERRMOR	mean	3.15*	4.25**	1.15	8.017	< .001	.10
	sd.	2.43	6.23	2.11			
ERR NP	mean	1.59	2.54**	0.90	5.794	.004	0.27
	sd.	1.75	3.43	1.86			
ERR GEN	mean	0.62	1.04(*)	0.46	2.572	.080	0.18
	sd.	1.26	1.68	0.96			
ERR VP	mean	1.11*	0.95(*)	0.18	4.567	.012	0.24
	sd.	1.40	2.52	0.55			
ERRSYN		2.27**	2.07**	.53	7.985	.001	.10
		2.69	3.13	1.06			
Overall distribution of syntactic patterns							
v2 in main cla.		63.76	56.46	6.37	6,141	.003	.08
		1.59	9.60	1.80			
V_{last} in sub. cla.		17.56	17.41	18.89	.929	.397	.01
		5.48	6.52	6.16	.		
					F (2, 76)		
Foreign accent rating (FAR)		3.32**	3,06*	1.89	8.048	.001	.18
		1.23	1.51	.60			

*: difference from control group < .05, **: difference from control group < .001, (*):

difference from control group marginally significant (p = .05-.075)

In answer to RQ2, we can therefore conclude that the free speech elicited from the control group languae users is different from that of the attriters only in some respects. Lexical richness (RQ 2a) appears to have remained largely unaffected over the emigration span, differences were found for only one group in one task. Overall variability in syntactic structures (RQ 2b) is unaffected: attriters and controls use SVX and XVS structures equally, and all groups have the same proportion of subordinate clauses.

On the other hand, the attriting groups have a larger incidence of errors (RQ 2c) and a stronger perceived foreign accent (RQ 2d) than the reference group. With respect to the errors observed in both tasks, it is of particular interest to notice that the GENL group have a higher average number of all categories errors in NP agreement than the GECA group (in particular gender marking), while the GECA group consistently make more syntactic errors.

Correlations across tasks

In order to establish whether the evidence for attrition that emerged from the analysis above is indicative of actual structural changes in the linguistic knowledge, that is, of an emerging representational deficit (RQ3), it was investigated whether the areas in which the attriters differed from the controls were correlated across tasks. This analysis is based on the assumption that, if the use of a non-target like structure (or the failure to identify such a structure) by an attriter is due to actual erosion of a particular grammatical rule, this non-target like use should recur across tasks, and those speakers who have a relatively high incidence of a particular error should make similar mistakes in other contexts and speech situations.

Pearson correlations were therefore established for the following variables:

1. Grammatical errors in interview and film retelling task

    - CAS: grammatical case in NPs

- GEN: grammatical gender agreement in NPs

- PLU: plural allomorphy and agreement in NPs

- VP: verb phrase morphology (tense, agreement, use of auxiliary)

- XVS: violation of the V2 rule in main clauses

- DWO: violation of the discontinuous word order rule

- SUB: violation of the V_{fin}-last rule in subordinate clauses

Table 7 shows an almost perfect absence of any kind of correlation across the speech samples: none of the pairings are significant, and effect size (r^2) is extremely low in all cases.

Table 7: Pearson correlation of error types across two free speech tasks per group

	GECA		GENL		GECG	
	r^2	p	r^2	p	r^2	p
CAS	0,00	0,99	0,01	0,60	0,00	0,68
GEN	0,02	0,35	0,00	0,76	0,01	0,40
PLU	0,00	0,77	0,03	0,21	0,00	0,80
VP	0,03	0,20	0,03	0,25	0,00	0,64
XVS	0,02	0,36	0,01	0,42	0,01	0,57
DWO	0,02	0,36	0,02	0,32	0,00	0,75
SUB	0,00	0,75	0,08	0,07	0,00	0,88

A similar picture emerges when the distribution of types of errors in the two free speech samples is correlated with the incidence of similar errors in the C-Test or the ability to detect those errors in the context of the grammaticality judgment task. Table 8 gives the outcome of a reliability analysis per group of NP agreement errors in the two free speech samples and the same types of error in the C-Test, and of the emergence of non-targetlike syntactic structures in the two free speech samples compared with the success in detecting such errors in an untimed grammaticality judgment task[5].

Table 8: Reliability analysis (Cronbach's α) of accuracy on grammatical structures across tasks

	GECA	GENL	GECG

NP agreement in free speech and C-Test	,-096	,024	-,307
Errors on V2 rule in free speech and score on GJ	,036	,086	-,206
Errors on discontinuous word order rule in free speech and score on GJ	,116	-,025	-,237
Errors on V_{last} rule in free speech and score on GJ	-,156	,161	-,206

For all groups, the α is very low across the three tasks for all three structures, suggesting that the same individuals do not consistently encounter problems with the same types of structures in different speech situations or across different tasks.

In answer to research question 3, these findings suggest that the poorer performance of the attriters is not the result of an emerging representational deficit: The results are more consistent with an interpretation which locates the source of non-targetlike behaviour on the part of the attriters in problems with the integration of different levels of knowledge in on-line speech production, possibly due to the effort of coordinating two highly active linguistic systems and/or of inhibiting a dominant L2 system.

4. Case study

In order to determine to what degree the effects of bilingualism which were found in the data from the attriters can be considered similar to effects found in L2 speakers, these findings were used as a background, against which data from two other speakers were compared. The first of these speakers is the most advanced attriter of L1 German in my own experience of the field: a 69-year old (at the time of data collection) female born in the city of Düsseldorf in 1926, who will be refered to as 'Gertrud Ullman'. She was born to parents of German-Jewish descent and left Germany in 1939 at the age of 13. She was taken to England on one of the so-called *Kindertransporte* (rescue missions for Jewish children organized by international Jewish organizations) and spent the war years there, living in foster families. After the war

she briefly returned to Germany, where she worked for the US Armed Forces, but as far as possible avoided all contact with Germans. She subsequently emigrated to the United States, where she still lived at the time of the interview in 1996. Since her emigration, she had hardly ever used German at all. 'Gertrud Ullman' was one of 35 speakers who were investigated in a study reported on by Schmid 2002, and her data were consistently the most strongly attrited on all variables measured in that study.

The second speaker is a highly successful L2 learner of German: a US-American born female, who emigrated to Germany in 1993 at the age of 24. She studied and currently teaches at a German university, and uses German predominantly in everyday life. At the time of the data collection, in the summer of 2008, she was 39 years old. This speaker will be referred to as 'Henrietta Jones'.

For the purpose of this investigation, 'Gertrud Ullman' represents a reference point for the limitations of possible attrition in post-puberty migrants, as she is the most strongly attrited such speaker of L1 German investigated to date. 'Henrietta Jones' represents the highly successful L2 learner of German.

4.1. Experimental data

The experimental tasks described above were not completed by 'Gertrud Ullman'[6]. 'Henrietta Jones' performs within the parameters of the native speakers, and even within the range of the control group, on all tasks. She is particularly successful on the grammaticality judgment task (see Table 9 and figures 1-3).

Table 9:Results on experimental tasks: a comparison of L1 German speakers and 'Henrietta Jones'

	German attriters in Canada	German attriters in the Netherlands	German controls	'Henrietta Jones'
CT	75,26	77,21	82,21	73
VF	20,24	20,91	25,09	20
GJ	0,82	0,83	0,84	0.93

These findings confirm the impression of 'Henrietta Jones' as a highly successful L2 learner of German: she is able to perform within native parameters on tasks which require the application of metalinguistic knowledge.

4.2. Free spoken data

At first glance, 'Gertrud Ullman' and 'Henrietta Jones' appear to be using free speech in a manner that is strikingly similar to each other, and markedly different from the speakers investigated in the quantitative analysis above (see Table 10). For each of the variables analyzed with respect to free speech, their results represent outliers[7] (with the exception of the foreign accent rating, where only the rating given to 'Henrietta Jones' qualifies as an outlier).

Table 10: Free speech measures: comparison of group and individual data

	GECA	GENL	GECG	Gertrud U.	Henrietta J.
D	98,91	95,34	94,95	49,60	47,35
ERR	7,94*	7,24*	2,71	34,25	27,78
ERRMOR	3,15*	4,25**	1,15	8,81	13,89
ERRSYN	2,27**	2,07**	0,53	6,12	3,65
FAR	3,32**	3,06*	1,89	4,20	5,70

However, a closer look at the types of errors made by these two speakers reveals substantial and surprising differences, and these differences are particularly pronounced where errors on morphology are concerned. Table 11 shows that NP morphology, in particular gender concord, appears to be far more problematic for 'Henrietta Jones' than the other error categories. 'Gertrud Ullman', on the other hand, performs relatively well on these categories (again in particular on gender concord), but has a disproportional amount of errors on verb phrase morphology.

Table 11: Morphological errors per 1,000 words: comparison of group means and individual data

	GECA	GENL	GECG	Gertrud U.	Henrietta J.
ERR NP morphology	1,31**	1,46**	0,44	3,92	10,69
ERR gender agreement	0,56	0,71	0,24	0,98	6,58
ERR VP morphology	0,85	1,32	0,58	4,90	2,92

It therefore appears that the most persistent problems for the advanced L2 learner investigated here are the morphological distinctions which are not marked in her L1. Conversely, the most advanced attriter appears to find these distinctions relatively unproblematic, but has a disproportionally high incidence of errors on structures which are overtly present in her L2. This finding is in line with the results of the quantitative analysis reported above, where the

L2 Dutch speakers made consistently more morphological errors than the L2 English speakers.

Lastly, let us briefly consider the correlation of problematic areas of linguistic knowledge across different tasks and speech situations. Recall that it was shown above that no such correlations obtained for either the attriters or the members of the reference group. It is not possible to perform the same analysis for the L2 speaker investigated here, as correlations cannot be calculated for a single data point. The distribution of errors across the groups of speakers is therefore visualized in Fig. 4, summarizing the incidence of gender concord errors per 1,000 words of spoken data in the interview (Y-Axis) and the film retelling (X-Axis). It is evident here that the distribution of the three L1 German groups is fairly randomized in this graph: many speakers have a high incidence of gender concord errors in one task, but none in the other (Pearson coefficient = .087, r^2 = .008, p = .292). However, the picture for 'Henrietta Jones' is somewhat different: her high incidence of gender concord errors is symmetric across both tasks. Similar graphic representations were built for all of the variables for which correlations were established above. Interestingly, the symmetrical pattern observed in the data from 'Henrietta Jones' obtained in both speech samples for case marking on noun phrases (Fig. 5) and for obligatory verb placement in main clauses (Fig. 6) and subordinate clauses (Fig. 7). Across tasks, the incidence of NP agreement violations appeared to correlate between her C-Test on the one hand and her interview (Fig. 8) and film retelling (Fig. 9) on the other. However, no such symmetrical pattern was found for plural agreement (Fig. 10) and VP agreement (Fig. 11) between the two speech samples, and while she had the second highest incidence of verb placement orders in both the interview and the film retelling task, she was also among the best participants in detecting these violations in the Grammaticality Judgment task (see Fig. 12 and 13).

5. Discussion and conclusion

The analysis reported above has returned a number of findings which are suggestive in the context of the debate on whether L1 attrition is indicative of structural changes or representational deficits on the one hand, or due to the increased cognitive load of processing a lesser-used language on the other.

  The quantitative analysis firstly established whether the attriting groups under observation here were outperformed by the controls, both on a number of formal tasks and in free speech (Research Questions 1 and 2). With respect to the formal tasks, the control group did indeed achieve higher scores than the attriters on the C-Test and the Verbal Fluency task. However, these differences were not very large, as indicated by the small effect size found for the group comparison. No difference was found for untimed grammaticality judgments, suggesting that the attriters' metalinguistic knowledge has remained unimpaired.

  In free speech, an interesting pattern emerged: while the attriters consistently made more errors than the control group, the overall lexical and syntactic complexity of their speech samples appeared largely unaffected by the attritional process. Lexical diversity was lower in only one of the two samples, and for only one of the attriting groups, and the distribution of syntactic patterns did not differ from the controls for either group.

  These results suggest that the attriters use and comprehend their L1 in essentially the same way as the controls, although they may sometimes encounter interferences from their L2. Taken together with the fact that the average correlations between the number of particular types of errors, as measured by Cronbach's α, were very low across the different tasks, an account which ascribes attritional phenomena to an emerging representational deficit, or a structural change to the L1, seems unlikely (Research Question 3). It appears more plausible that the differences observed between the attriters and the controls are the outcome of

difficulties which the bilingual speakers encounter in the coordination and integration of information on different linguistic levels in their two language systems.

This conclusion is also suggested by the fact that in the area of morphology, the attriters in both the quantitative and the qualitative analysis made more errors on features which are also overtly present in their L2: The L2 Dutch speakers had a much larger average number of errors on NP agreement than the L2 English speakers. In the area of VP agreement, where both English and Dutch share many of the morphological characteristics of German, the incidence of errors was more similar (the GENL group had more errors of this type in the interview, the GECA group in the film retelling task). This distribution suggests that performance in L1 attrition may sometimes diverge from the native norm, due to the competition with L2. These deviations will be most frequent and noticeable where competition between the two linguistic systems is highest – that is, where the systems are similar.

The picture which emerged from the analysis of data produced by a highly successful L2 speaker was rather different. While at first glance she appeared to have a similar incidence of errors as the most advanced L1 attriter, the distribution of these errors diverged from the attrition data in a striking way: the advanced attriter had, overall, the most errors in those areas which had also proven most difficult for the overall group of attriters with English as L2 (verb phrase morphology) – that is, on those features which were most similar between her L1 and L2. For the L2 speaker of German, on the other hand, it was those morphological phenomena which were *not* encoded in her L2 where the most deviations from the native norm were encountered. Interestingly, it was those same areas of grammar where the incidence of errors she made appeared to correlate strongly across different tasks.

Based on these findings, we can hypothesize that the underlying reasons for non-targetlike performance observed in L1 attriters and L2 speakers may be different. In L1 attrition,

divergences from the native norm are caused by similarities between the two linguistic systems. The L1 attriter probably still has the knowledge of a native speaker, but sometimes lacks the control necessary to apply this knowledge in the online speech production process, due to the cognitive demands effected by the competition between the two systems. On the other hand, the L2 speaker may not have the knowledge and processing strategies of the native speaker. However, with increasing practice and proficiency in the L2, she may be able to control her language output, possibly based on non-native like cognitive strategies.

   As these conclusions are to some extent based on the findings from a case study, they should be treated with extreme caution. Large-scale quantitative comparisons of ultimate attainment in L2 acquisition and in L1 attrition are necessary to substantiate the claim that there is a fundamental difference between the L2 of a late learner and the L1 of a late attriter. Such corroborations could provide substantial evidence for the hypothesis that the critical period does indeed differentiate two qualitatively different processes of language learning, and that the non-targetlike language use observed in L2 speakers is not merely the outcome of the higher cognitive load which bilinguals experience.

References:

$$$

Abutalebi, J. 2008. "Neural aspect of second language representation and language control."
   Acta Psychologica 128: 466-478.

Ammerlaan, T. 1996. *You Get a Bit Wobbly.. Exploring bilingual lexical retrieval processes
   in the context of first language attrition.* University of Nijmegen PhD dissertation.

Ben-Rafael, M. and Schmid, M. S. 2007. "Language attrition and ideology: Two groups of
   immigrants in Israel." In Köpke et al. (eds.), 205-226.

Bialystok, E. 1997. "The structure of age: in search of barriers to second language acquisition." *Second Language Research* 13(2): 116-137.

Bialystok, E. and Hakuta, K. 1999. "Confounded age: linguistic and cognitive factors in age differences for second language acquisition." In *Second Language Acquisition and the Critical Period Hypothesis*, Birdsong, D. (Ed.), 161-181. Mahwah, NJ: Erlbaum.

Bialystok, E. and Miller, B. 1999. "The problem of age in second-language acquisition: Influences from language, structure, and task." *Bilingualism: Language and Cognition* 2(2): 127-145.

Birdsong, D. 2006. "Age and second language acquisition and processing: A selective overview." *Language Learning* 56(1): 9-49.

Bley-Vroman, R. 1990. "The logical problem of foreign language learning." *Linguistic analysis* 20(1/2): 3-49.

Bylund, E. S. 2008. *Age Differences in First Language*. Stockholm University PhD dissertation.

Carroll, S. E. 2006. "Shallow processing: a consequence of bilingualism or second language learning?" *Applied Psycholinguistics* 27(1): 53-56.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Clahsen, H. and Felser, C. 2006. "Grammatical processing in language learners." *Applied Psycholinguistics* 27: 3-42.

Clahsen, H. and Muysken, P. 1986. "The availability of universal grammar to adult and child learners - a study of the acquisition of German word order." *Second Language Research* 2(2): 93-119.

Clahsen, H. and Muysken, P. 1989. "The UG paradox in L2 acquisition." *Second Language Research* 5(1): 1-29.

Cornips, L. and Hulk, A. 2008. "Factors of success and failure in the acquisition of grammatical gender in Dutch."*Second Language Research* 24: 267-295.

Curtiss, S. 1988. "Abnormal language acquisition and the modularity of language." In *Linguistics: The Cambridge Survey*, Vol, Newmeyer, F. J. (Ed.), 96-116. II. CUP.

De Bot, K. and Clyne, M. 1994. "A 16 year longitudinal study of language attrition in Dutch immigrants in Australia." *Journal of Multilingual and Multicultural Development* 15(1): 17-28.

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

Dijkstra, T. and van, W. B.Heuven. 2002. "The architecture of the bilingual word recognition system: From identification to decision." *Bilingualism. Language and Cognition* 5(3): 175-197.

Dussias, P. E. 2004. "Parsing a first language like a second: The erosion of L1 parsing strategies in Spanish-English Bilinguals." *International Journal of Bilingualism* 8(3): 355-371.

Flege, J., Yeni-Komshian, G. and Liu, S. 1999. "Age constraints on second-language acquisition." *Journal of Memory and Language* 41: 78-104.

Grosjean, F. 2001. The bilingual's language modes. In *One Mind, Two Languages: Bilingual language processing,* Nicol, J. (Ed.), 1-22. Oxford: Blackwell.

Grotjahn, R. 1987. "How to construct and evaluate a C-Test: A discussion of some problems and some statistical analyses." *Quantitative Linguistics* 34: 219-53.

Hakuta, K., Bialystok, E. and Wiley, E. 2003. "Critical evidence: A test of the critical-period hypothesis for second-language acquisition." *Psychological Science* 14: 31-38.

Hawkins, R. 2001. "The theoretical significance of Universal Grammar in second language acquisition." *Second Language Research* 17(4): 345-367

Hawkins, R. and Chang, C. Y.-H. 1997. "The partial availability of Universal Grammar in

    second language acquisition: the 'failed functional features hypothesis'." *Second*

    Language Research 13(3): 187-226.

Hopp, H. 2007. *Ultimate Attainment at the Interfaces in Second Language Acquisition:*

    Grammar and processing. University of Groningen PhD dissertation.

Hopp, H. 2008. "Inflectional variability in L2 acquisition: grammatical impairment vs.

    processing difficulty." Paper presented at EuroSLA 18, Aix-en-Provence, Sept. 2008.

Hulsen, M. 2000. *Language Loss and Language Processing. Three generations of Dutch*

    migrants in New Zealand. University of Nijmegen PhD dissertation.

Hulstijn, J. J. 1997. "Second language acquisition research in the laboratory. Possibilities and

    limitations. " *Studies in Second Language Acquisition* 19: 131-143.

Johnson, J. S. and Newport, E. L. 1989. "Critical period effects in second language learning:

    the influence of maturational state on the acquisition of English as a second language."

    Cognitive Psychology 21: 60-99.

Köpke, B. 1999. *L'attrition de la première language chez le bilingue tardif: implications pour*

    l'étude psycholinguistique du bilinguisme. Université de Toulouse-Le Mirail PhD

    Dissertation.

Köpke, B. 2004. "Neurolinguistics aspects of attrition." *Journal of neurolinguistics* 17: 3-30.

Köpke, B. and Schmid, M.S. 2004. "First language attrition: the next phase." In Schmid et al

    (Eds.), pp. 1-43.

Köpke, B., Schmid, M.S., Keijzer, M. and Dostert, S. (Eds.). 2007. *Language Attrition:*

    Theoretical perspectives. Amsterdam: John Benjamins.

Kroll, J. F., Bobb, S. C., Misra, M. and Guo, T. 2008. "Language selection in bilingual

    speech: Evidence for inhibitory processes." *Acta Psychologica* 128: 416-430.

Lardiere, D. 1998. "Dissociating syntax from morphology in a divergent L2 endstate

grammar." *Second Language Research* 14(4): 1-26.

Lenneberg, E. H. 1967. *Biological Foundations of Language*. New York: Wiley.

Mayberry, R. I. and Eichen, E. B. 1991. "The long-lasting advantage of learning sign

language in childhood: Another look at the critical period for language acquisition."

Journal of Memory and Language 30: 486-512.

Mayberry, R. I., and Lock, E. 2003 "Age constraints on first versus second language

acquisition: Evidence for linguistic plasticity and epigenesis." *Brain and Language* 87:

369-384.

McConkey Robbins, A., Burton, D.Koch, Osberger, M. J., Zimmerman-Philips, S. and

Kishon, -Rabin, L. 2004. "Effect of age at cochlear implantation on auditory skill

development in infants and toddlers." *Archives of Otolaryngology Head and Neck

Surgery* 130: 570-574.

McKee, G., Malvern, D. and Richards, B. 2000. "Measuring vocabulary diversity using

dedicated software." *Literary and Linguistic Computing* 15(3): 323-337

Montrul, S. 2002. "Incomplete acquisition and attrition of Spanish tense/aspect distinctions in

adult bilinguals." *Bilingualism: Language and Cognition* 5(1): 39-68

Montrul, S. 2004. "Convergent outcomes in L2 acquisition and L1 loss." In Schmid et al.

(eds.), 259-279.

Montrul, S. 2008. *Incomplete Acquisition in Bilingualism. Re-examining the Age Factor*.

Amsterdam: John Benjamins.

Orgassa, A. and Weerman, F. 2008. "Dutch gender in spacific language impairment and

second language acquisition." *Second Language Research* 24: 333-364.

Pallier, C. 2007. "Critical periods in language acquisition and language attrition." In Köpke et

al. (eds.), 155-168.

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

Paradis, M. 2004. *A Neurolinguistic Theory of Bilingualism*. Amsterdam: John Benjamins.

Pelc, L. 2001. *L1 Lexical, Morphological and Morphosyntactic Attrition in Greek-English Bilinguals*. CUNY PhD dissertation.

Perdue, C. 1993. *Adult language acquisition: Cross-linguistic perspectives*. Cambridge: CUP.

Prévost, P. and White, L. 2000. "Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement." *Second Language Research* 16(2): 103-133.

Roberts, P. M. and Le Dorze, G. (1997). "Semantic organization, strategy use, and productivity in bilingual semantic verbal fluency." *Brain and Language* 59: 412-449.

Schmid, M. S. 2002. *First Language Attrition, Use, and Maintenance: The case of German Jews in Anglophone countries*. Amsterdam: John Benjamins.

Schmid, M. S. 2004. "First language attrition: The methodology revised." *International Journal of Bilingualism* 8(3): 239-255.

Schmid, M. S. 2007. "The role of L1 use for L1 attrition." In Köpke et al. (eds), 135-153.

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

Schmid, M. S., Köpke, B., Keijzer, M. and Weilemar, L. (eds). 2004. *First Language Attrition: Interdisciplinary perspectives on methodological issues*. Amsterdam/Philadelphia: John Benjamins.

Schwartz, B. D. and Sprouse, R. A. 2000. "When syntactic theories evolve: Consequences for L2 acquisition research." In *Second Language Acquisition and Linguistic Theory*, Archibald, J. (Ed.), 156-186. Oxford: Blackwell.

Seliger, H. W. and Vago, R. M. 1991. "The Study of First Language Attrition: An Overview."

    In *First Language Attrition*, Seliger, H.W. and Vago, R. M. (eds), 3-15. Cambridge

    University Press.

Sharwood Smith, M. 1983. "On explaining language loss." In *Language Development on the*

    Crossroads, Felix, S. and Wode, H. (Eds.), 49-69. Tübingen: Gunter Narr.

Smith, N. and Tsimpli, I.-M. 1995. *The Mind of a Savant. Language Learning and*

    Modularity. Oxford: Blackwell.

Sorace, A. 2005. "Selective optionality in language development. " In *Syntax and variation.*

    Reconciling the biological and the social, Cornips, L. and Corrigan, K. P., (Eds.) , 55-80

    Amsterdam/Philadelphia: John Benjamins, 55-80.

Squire, L. R. 1992. "Declarative and Nondeclarative Memory: Multiple Brain Systems

    Supporting Learning and Memory." *Journal of Cognitive Neuroscience* 4(3): 232 – 243.

Tsimpli, I. 2007. "First language attrition from a minimalist perspective: Interface

    vulnerability and processing effects." In Köpke et al. (eds.), 83-98.

Tsimpli, I., Sorace, A., Heycock, C. and 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(3): 257-277.

Ullman, M. 2001. "The neural basis of lexicon and grammar in first and second languages:

    The declarative/procedural model." *Bilingualism: Language and Cognition* 4(1): 105-

    122.

Ventureyra, V. and Pallier, C. 2004. "In search of the lost language: The case of adopted

    Koreans in France." In Schmid et al. (eds), 207-221.

White, L. 2003. *Second Language Acquisition and Universal Grammar*. CUP.
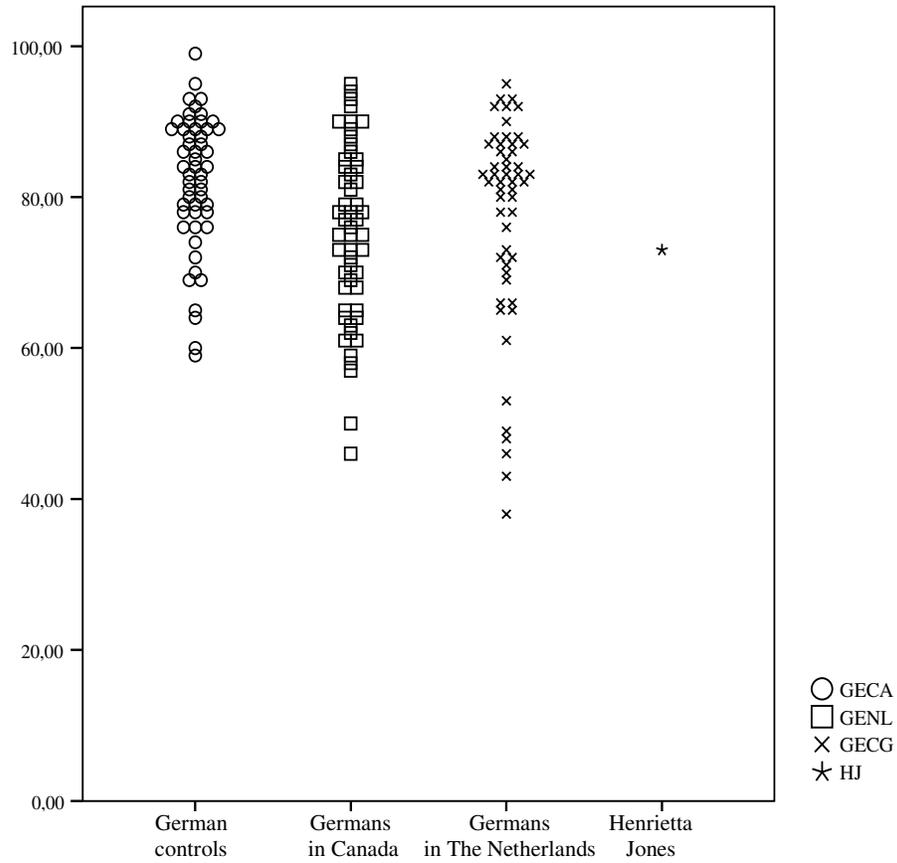
Fig. 1: Scores on the C-Test: comparison of L1 German speakers and Henrietta Jones
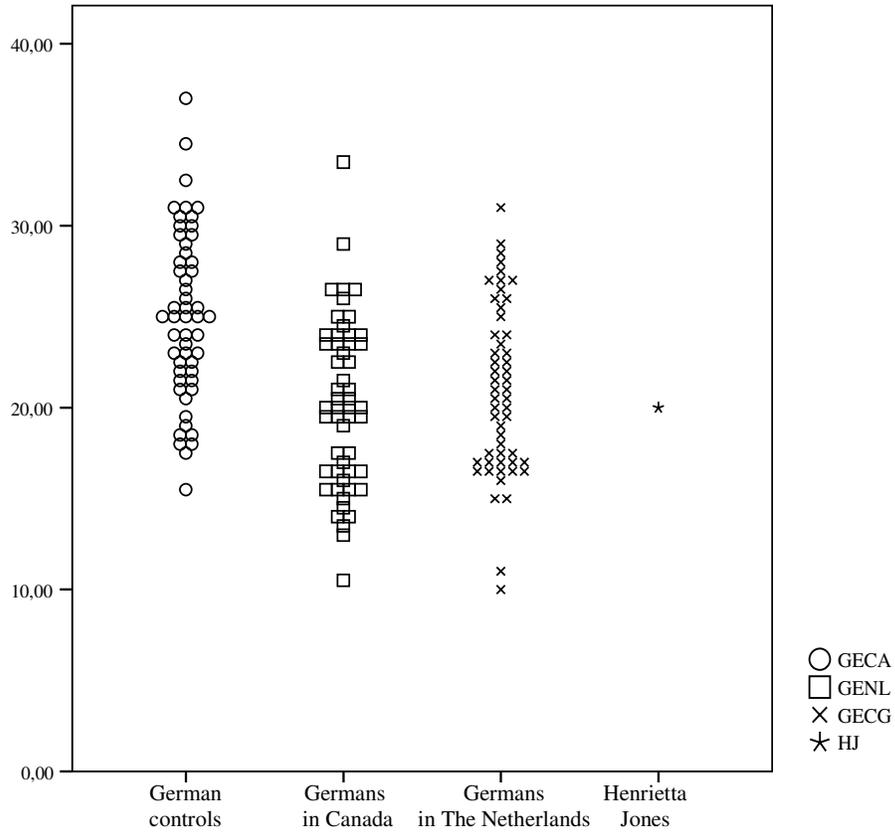
Fig. 2: Scores on the verbal fluency task: comparison of L1 German speakers and Henrietta
Jones
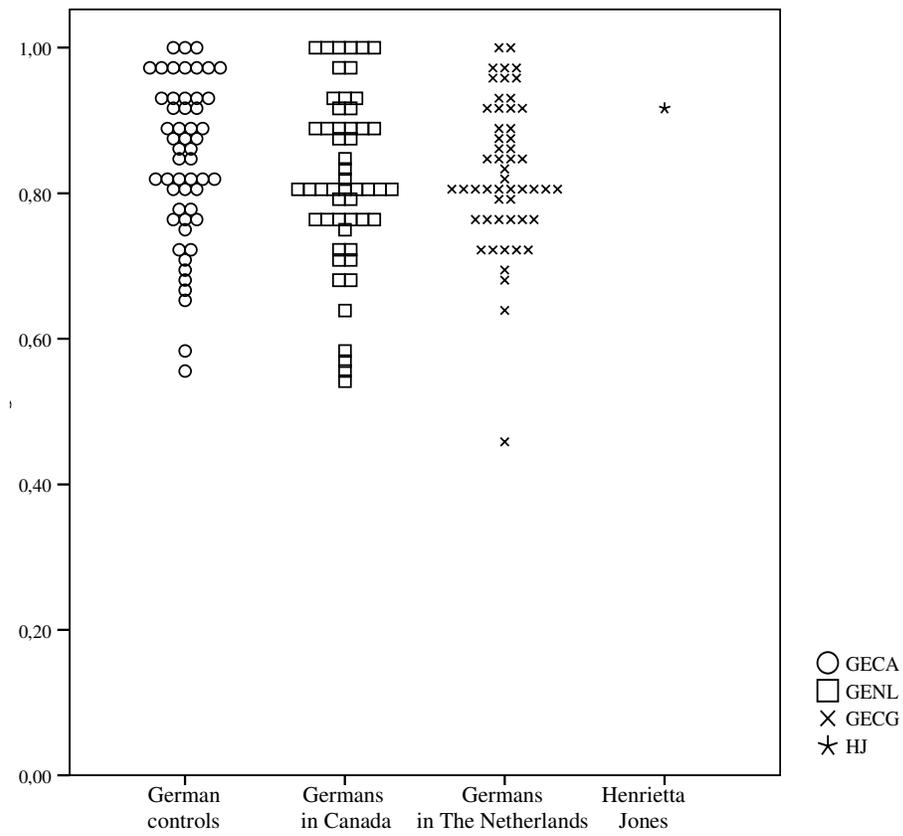
Fig. 3: Scores on the Grammaticality Judgment task: comparison of L1 German speakers and Henrietta Jones
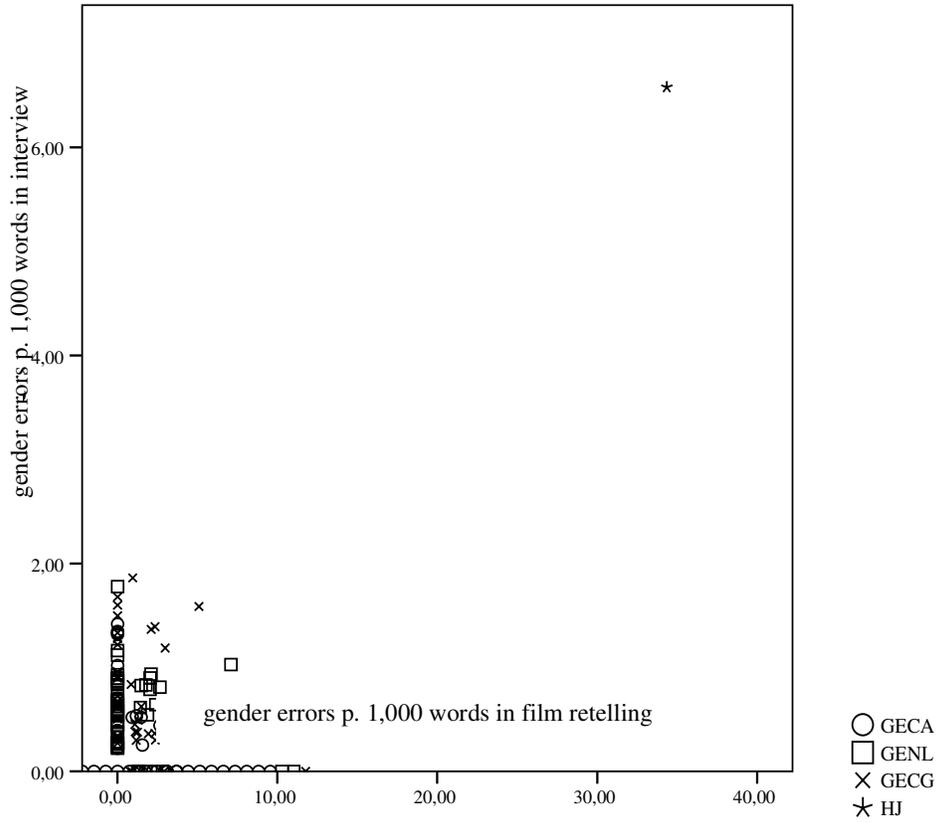
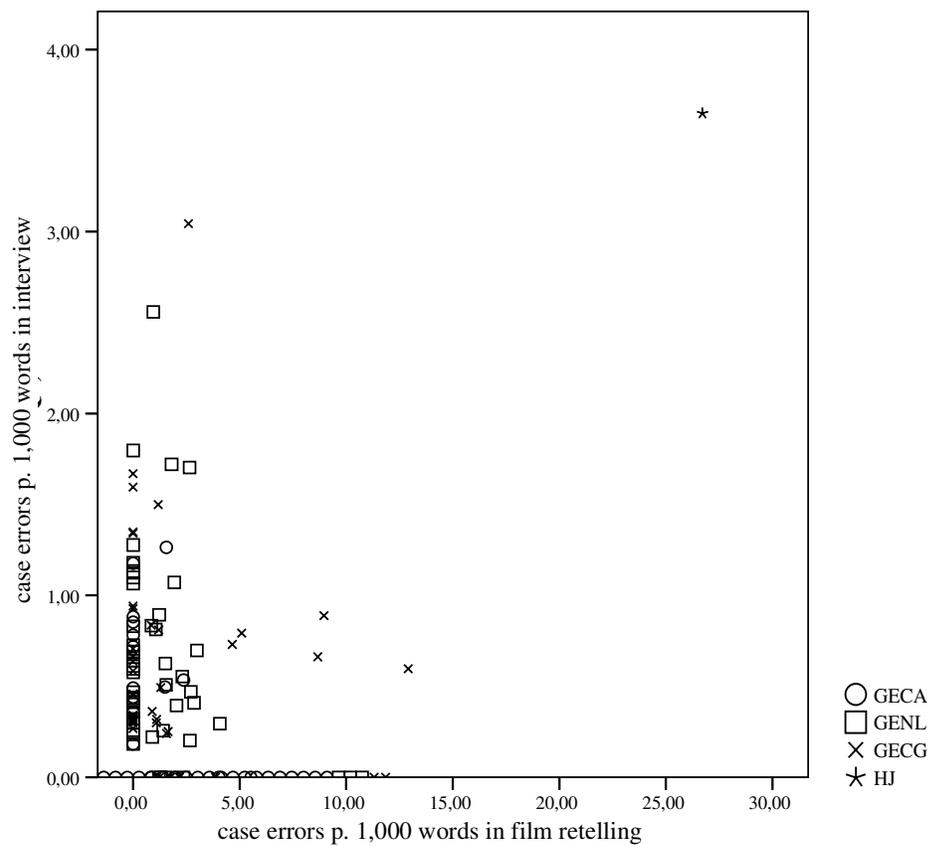Fig. 4: Correlation of gender errors in interview and film retelling

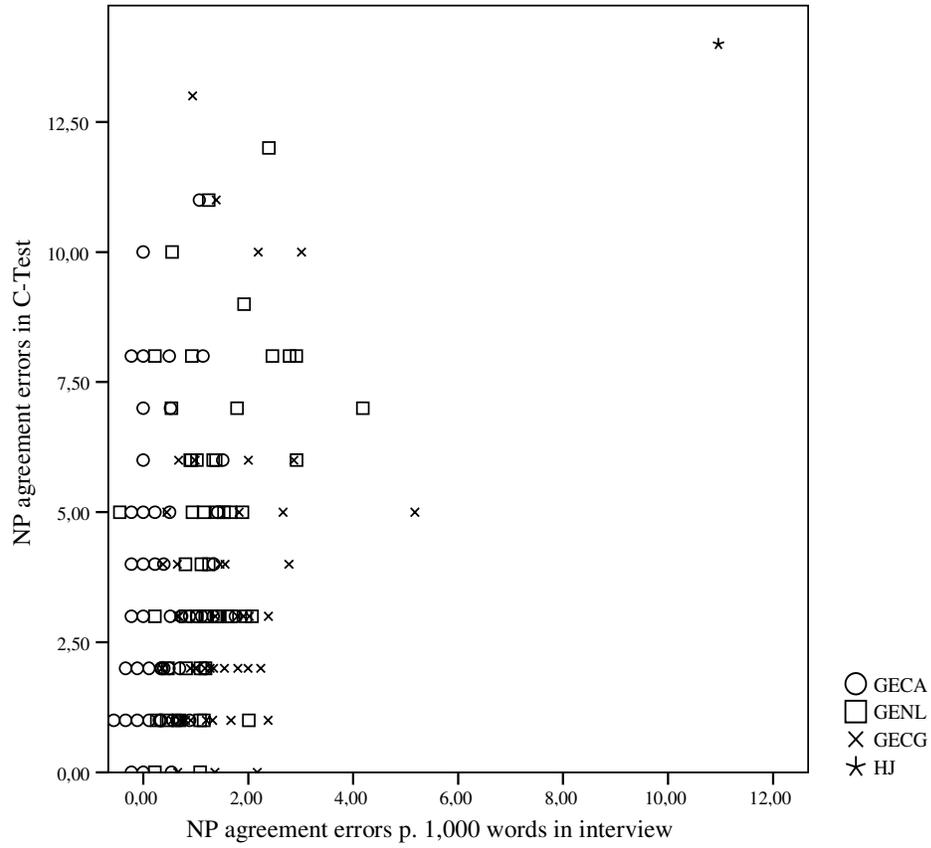Fig. 5: Correlation of case errors in interview and film retelling
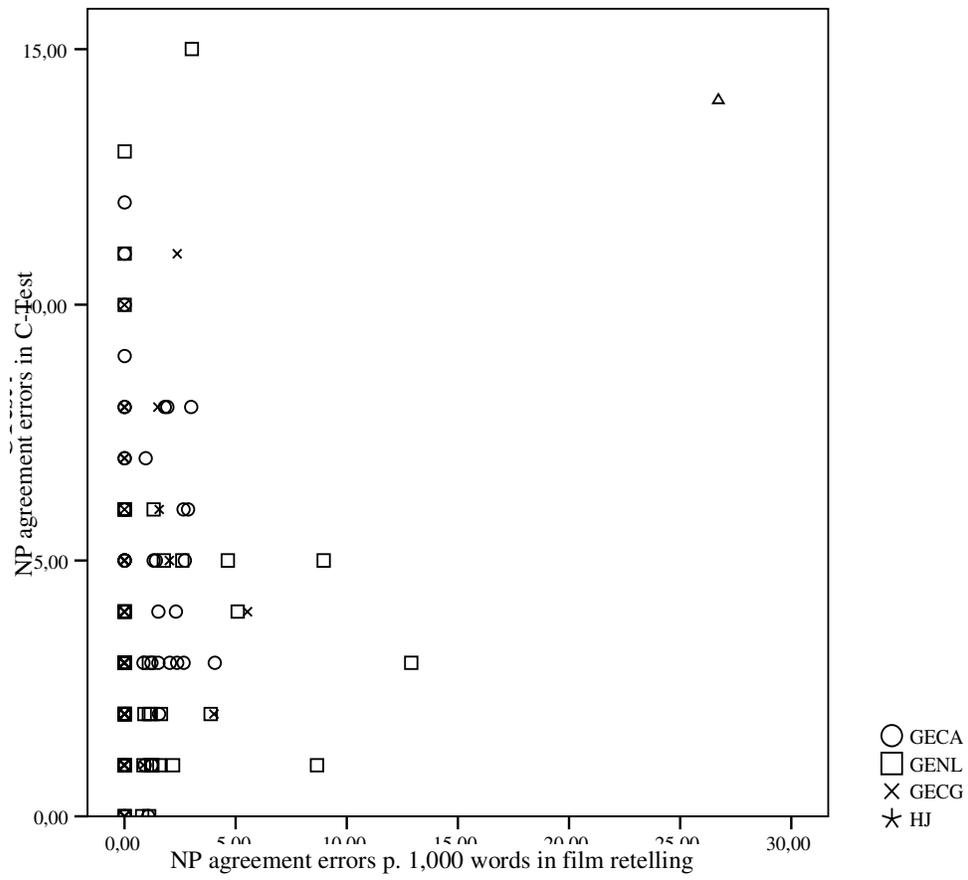
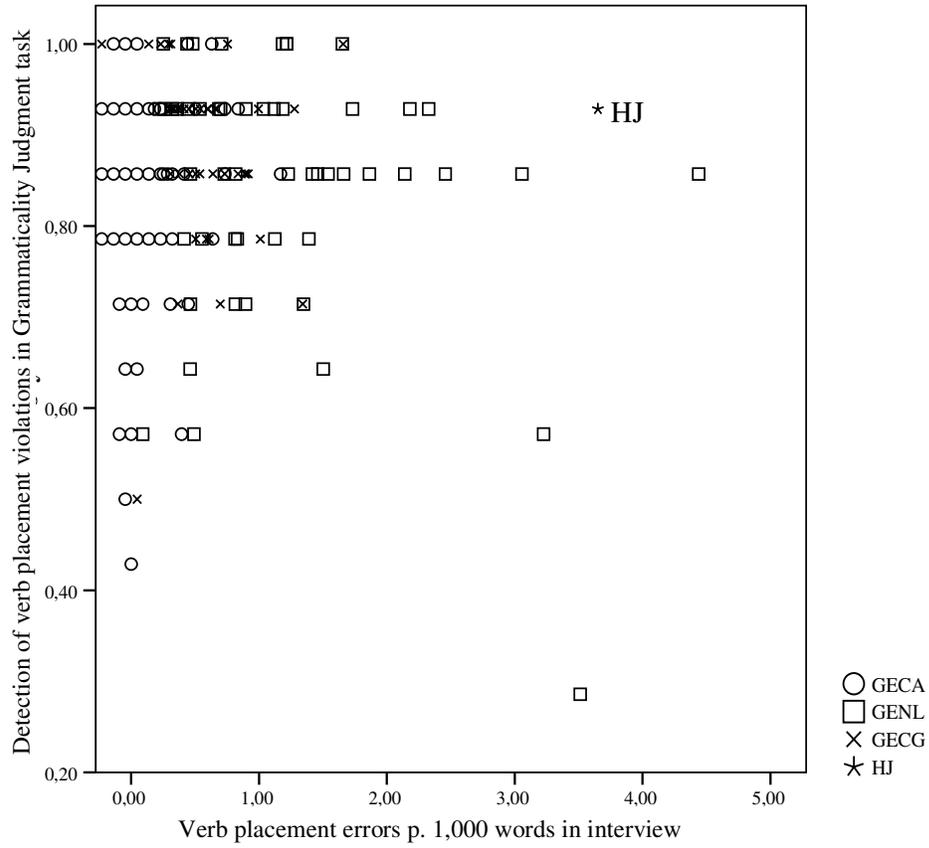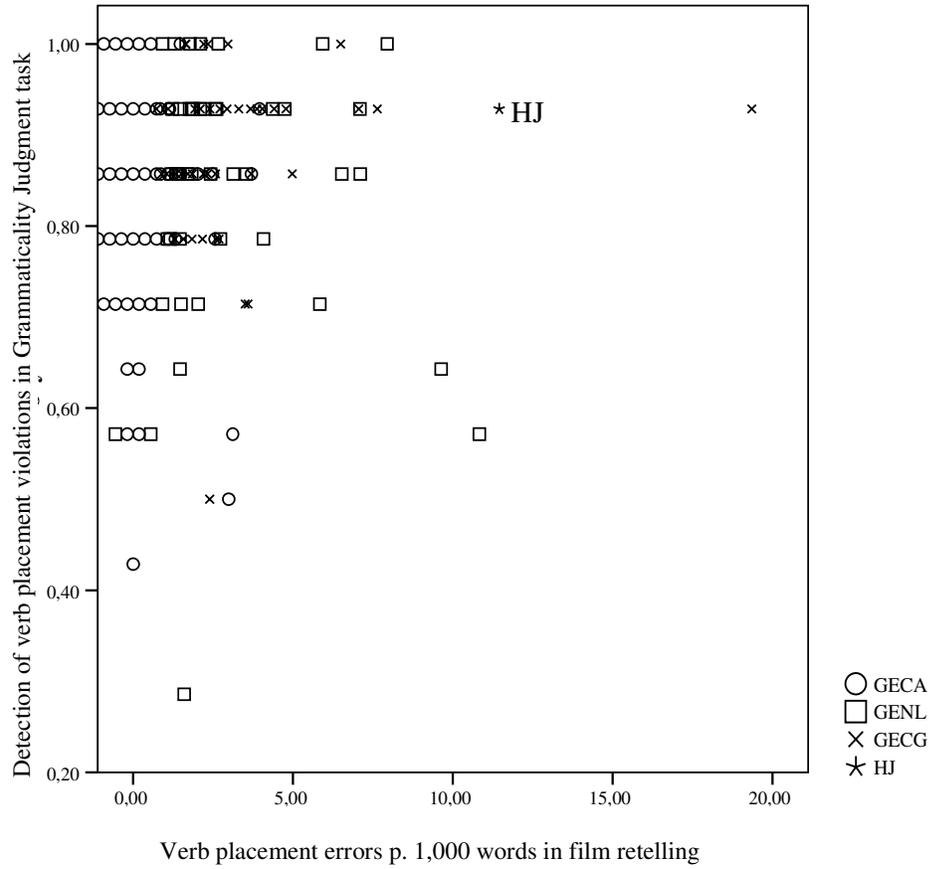Fig. 6 Correlation of XVS errors in interview and film retelling

Fig. 7: Correlation of V_{last} errors in interview and film retelling

Fig. 8: Correlation of NP agreement errors in C-Test and interview

Fig. 9: Correlation of NP agreement errors in C-Test and film retellig
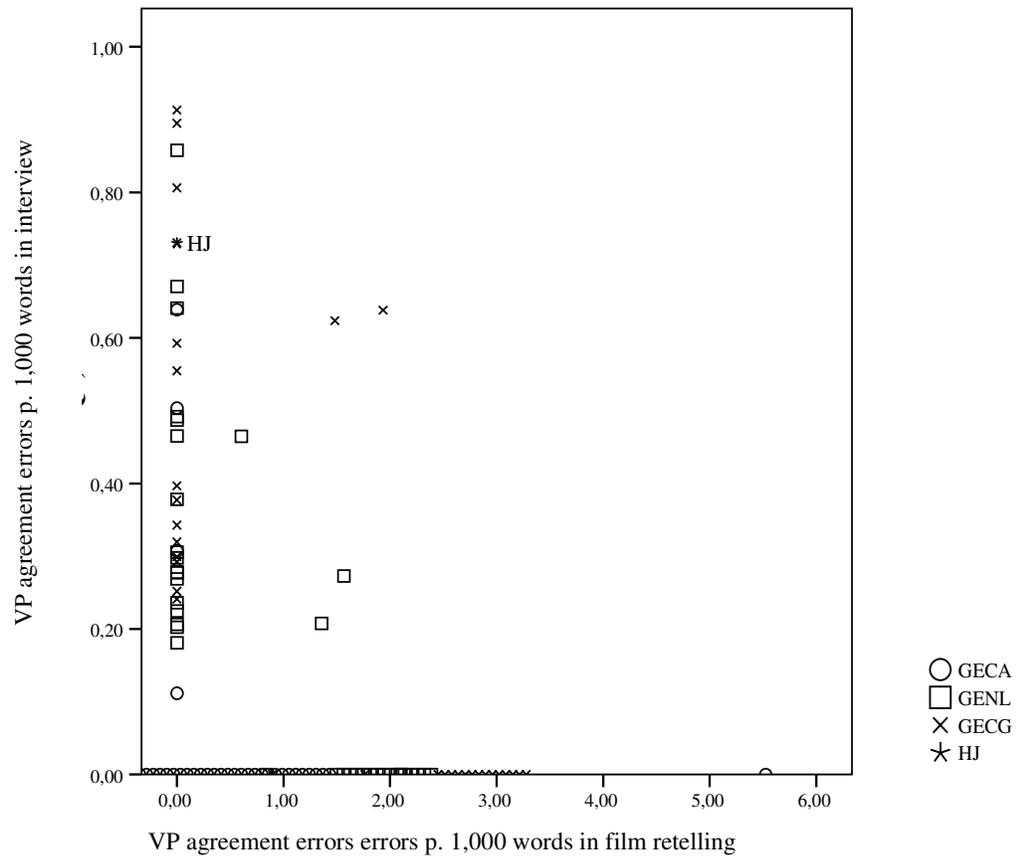
Fig. 10: Correlation of NP plural agreementerrors in interview and film retelling
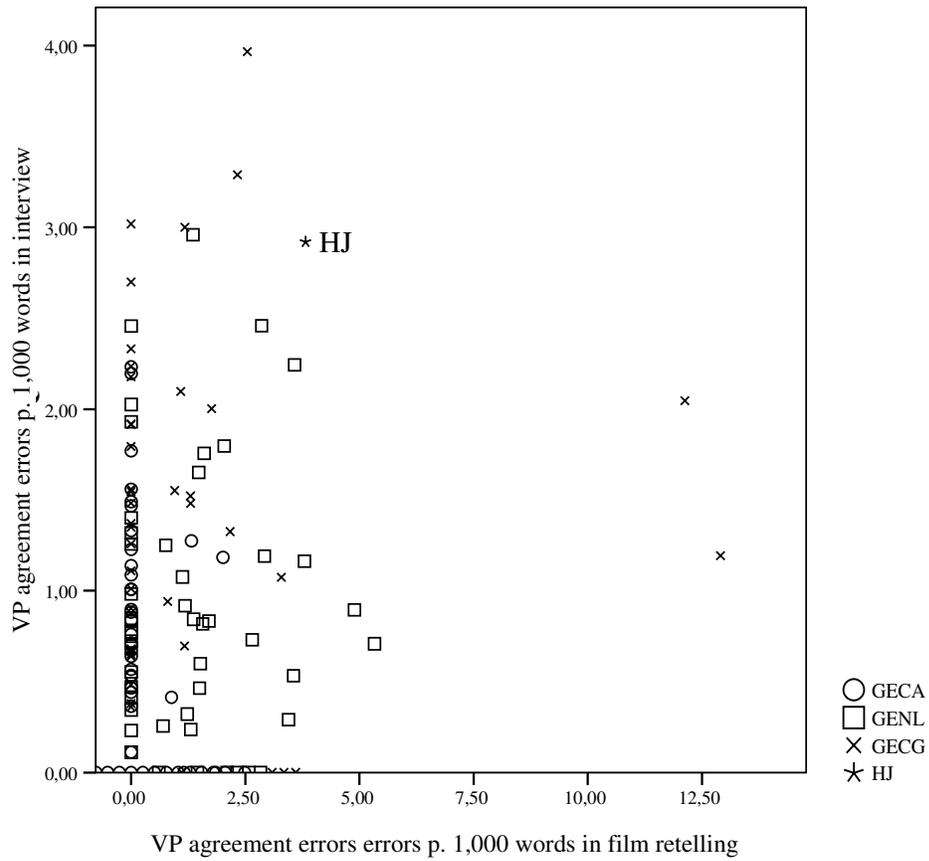
Fig. 11: Correlation of VP morphology errors in interview and film retelling

Fig. 12: Correlation of scores on GJ task and syntactic errors in interview

Fig. 13: Correlation of scores on GJ task and syntactic errors in film retelling

2   Note that while the two positions are often juxtaposed, there is no logical reason to assume that they are mutually exclusive, and that the observed effect might not be due to a combination of both.

3   The baseline for comparison in such studies are typically not early L2 learners but (monolingual) L1 speakers. The logical and methodological problem posed by this practice is discussed below.

4   Note that this view is compatible with representational deficit accounts, in that it assumes that L2 speakers may reach the surface-level target of native-like grammar, but rely on different (extra-linguistic) strategies in order to get there.

5   For the purpose of the reliability analysis, the polarities of the results on the C-Test and the GJ task were reversed by subtracting each individual score from the maximum score, in order to ensure that a high score would indicate low proficiency on all the variables included in the analysis.

6   The investigation reported on in Schmid (2002) was confined to an analysis of free data collected by historians for an oral history project. An experimental investigation of this population was impossible for ethical reasons.

7   An outlier in the current context is a result which is higher than the group mean plus the standard deviation times two or lower than the group mean minus the standard deviation times two.