

Frequent native language use supports phonemic and semantic verbal fluency in L1 and L2: An extended analysis of verbal fluency task performance in an L1 language attrition population

International Journal of Bilingualism

1–23

© The Author(s) 2023



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/13670069231193727

journals.sagepub.com/home/ijbNana Lehtinen 

University of Turku, Finland; University of California, Berkeley, USA

Anna Kautto  and

Kati Renvall

University of Turku, Finland

Abstract

Purpose: Verbal fluency (VF) tasks are an efficient tool for exploring lexical retrieval. Attriters generally produce fewer words in semantic categories than monolinguals, but the processes underlying attriters' lexical retrieval remain unclear. Phonemic tasks are scarcely applied in language attrition studies. We aim to identify processes underlying attriters' phonemic verbal fluency (PVF) and semantic verbal fluency (SVF) performance by extending the analysis beyond total scores in the first (L1) and second (L2) language.

Design: We modeled total scores and temporal aspects of task performance in L1 and L2 for the attriters, with L1 use and length of residence (LoR) as predictors. We analyzed the number and types of errors between languages and compared L1 task performance with a matched monolingual group.

Data and analysis: Attriters' ($N=38$) phonemic and semantic task performance in L1 (Finnish) and L2 (English) were modeled, and results were contrasted to a matched monolingual control group ($N=50$).

Findings: Attriters demonstrated strong proficiency in L1 and similar lexical retrieval processes in L1 and L2 after 20 years of immersive L2 exposure. Frequent L1 use supported overall performance but slowed performance down in both languages. Compared with monolinguals, attriters show a disadvantage in SVF, but not due to slower initiation. Instead, attriters rely more on rapid L1 retrieval than monolinguals.

Originality: Prior research on VF tasks in language attrition populations has mainly focused on SVF total scores in L1. Our investigation explores underlying mechanisms of L1 and L2 task

Corresponding author:

Nana Lehtinen, Department of Psychology and Speech-Language Pathology, University of Turku, Turku 20014, Finland.

Email: nanleh@utu.fi

performance in PVF and SVF within a language attrition population and contrasts L1 performance to a matched monolingual group performance.

Implications: This study demonstrates the importance of investigating L2 alongside L1 in language attrition studies for a holistic approach to language processing strategies. Our findings enhance understanding of the processes underlying VF task performance, emphasizing the significance of VF task early-stage performance.

Keywords

Language attrition, verbal fluency task, bilingualism, semantic verbal fluency, phonemic verbal fluency

Introduction

Within the broad field of bilingualism, the gradual, non-pathological change in the native language (L1) due to reduced use and exposure is referred to as L1 language attrition (Schmid, 2011b; Schmid & Köpke, 2017). *Language attrition* is a complex phenomenon that affects all aspects of language and communication. It can be observed as a series of subtle but consistent changes in L1 that are not always directly observable in language performance. L1 attrition primarily occurs during online processing and may lead to reduced fluency, speaker insecurities, reduced accuracy, and overall simplification in L1 use, as well as interference from L2 (e.g., the integration of L2 elements into L1) (Jarvis, 2019, Chapter 21, pp. 241–245; Schmid, 2011b; Schmid & Köpke, 2009). The most investigated domain for language attrition is the lexicon. Fluency and fluency-related phenomena are studied by measuring speech rate, hesitations, filled pauses, repetitions, and self-repairs in free speech. Verbal fluency (VF) tasks are often used for a more controlled approach to lexical fluency (Jarvis, 2019, Chapter 21, pp. 248–249; Schmid & Köpke, 2009). Observed changes in online task performance are not necessarily representations of systematic grammatical restructuring, erosion, or loss of language competence (Sharwood Smith, 2019, Chapter 8, p. 85). However, language attriters tend to apply categorical rules or features with optionality and exhibit greater variation in their L1 performance than monolinguals (Schmid & Cherciov, 2019, Chapter 23, p. 267; Schmid & Dusseldorp, 2010). For a comprehensive overview of language attrition studies and findings, see Schmid and Köpke (2019, pp. 509–541).

Building on previous language attrition research, this study examines the L1 of immigrants who share a language history characterized by late sequential bilingualism, immersion in an L2 environment, and minimal exposure to their native L1 over an extended period (Schmid, 2011b, 2019; Schmid & Jarvis, 2014; Schmid & Köpke, 2017). In this study, we refer to participants who are late sequential bilinguals fully immersed in their second language as “attriters” to differentiate them from other bilingual groups (the term bilingualism is used for the sake of brevity while encompassing multilingualism). We use the more general term “bilinguals” to include individuals with various language backgrounds and sociolinguistic contexts. Nonetheless, we acknowledge the controversy of using the term attriter to refer to a participant group for which evidence of attrition has not been detected (Kasparian & Steinhauer, 2017; Schmid & Köpke, 2017).

Many language attrition studies focus on L1 performance and compare performance between monolingual speakers and L1 attriters (Schmid, 2019; Schmid & Dusseldorp, 2010; Schmid & Yılmaz, 2018). It is widely recognized that languages in the bilingual brain interact continuously in a bidirectional process, altering the cognitive and neural dynamics of both languages and

resulting in shifts in individual language dominance patterns over time (Grosjean, 2013; Gurunandan et al., 2023; Laine & Lehtonen, 2018; Linck & Kroll, 2019, Chapter 9, p. 97; Treffers-Daller, 2019). These shifts are influenced by sociolinguistic parameters such as language exposure and use and can, to some extent, be observed in language performance (Treffers-Daller, 2019; Yılmaz, 2019). The language history of language attrition populations typically involves a rapid decline in L1 input coupled with full immersion in L2 upon immigration. Previous studies show that within language attrition populations, some attriters perform at native levels, while others exhibit noticeable changes in L1 performance (Schmid & Yılmaz, 2018). Examining L2 performance alongside L1 can provide insights into individual factors contributing to L1 attrition or retention, as well as L2 acquisition (Bylund & Ramírez-Galan, 2016; Gurunandan et al., 2023; Laine & Lehtonen, 2018; Runnqvist & Costa, 2012; Schmid & Köpke, 2017; Schmid & Yılmaz, 2018).

To investigate the impact of attriters' shared language history on task performance, many language attrition studies investigate the influence of two variables: length of residence (LoR, i.e., time deprived of native L1 exposure and extensive exposure to L2) and the frequency of L1 use in daily life. It has been shown that LoR predicts task performance when the minimum LoR is less than 10 years, with the effect stabilizing after this period (Opitz, 2011; Schmid, 2011a, 2019). This stabilization has been partially attributed to the initial years of immersion which typically coincide with a period of rapid L2 learning in an immersive environment (Linck & Kroll, 2019, Chapter 9, p. 95). However, while the added prevalence of L2 partly triggers L1 attrition, it is not directly linked to L2 acquisition or level of L2 proficiency (Yılmaz, 2019, Chapter 26, p. 307). Unfortunately, we are not aware of language attrition studies that directly investigate the impact of LoR on L2 performance.

The frequency of L1 use has not been found to systematically impact L1 performance (for an overview, see Schmid, 2019, Chapter 25, pp. 294–295). However, L1 use in different life domains can have different implications for changes in L1 performance. On one hand, frequent use of L1 in a professional setting has been found to have a positive impact on low-frequency word retrieval (Yılmaz & Schmid, 2012), verbal fluency, general proficiency (c-test), grammaticality judgment task, and lexical diversity (Schmid & Dusseldorp, 2010). On the other hand, frequent L1 use in informal interactions with peers (attriters) has been associated with increased variability in the phonemic domain (De Leeuw et al., 2010). Although LoR and frequent L1 use do not systematically predict attrition as separate variables (Schmid, 2019; Schmid & Dusseldorp, 2010; Schmid & Yılmaz, 2018), it has been shown that the impact of LoR can become significant with very little L1 use (de Bot & Clyne, 1994).

Consequently, it has been suggested that attriters who have spent a long time in an L2 environment and use L1 less often exhibit signs of attrition at an individual level. In contrast, speakers who have spent extended time in an L2 environment and who use L1 frequently with peers reflect a contact-induced language change in the community on a larger scale (Schmid, 2011a; Schmid & Köpke, 2017). In the current study, we investigate the impact of LoR and frequency of L1 use on verbal fluency task performance separately, as well as their interactions, to reliably include both extralinguistic variables.

Verbal fluency tasks

VF tasks are widely applied across various populations and languages (Goral, 2004; Olabarrieta-Landa et al., 2017). In a VF task, participants are given a criterion and asked to generate words within a predetermined time frame (typically 60 seconds) (Strauss et al., 2006). The common VF task types are phonemic (PVF) and semantic (SVF).

In a PVF task, participants generate words starting with a specific letter or phoneme, and three trials are typically included for reliable fluency assessment (Oberg & Ramírez, 2006; Strauss et al., 2006; Tombaugh et al., 1999). In studies conducted in English, the letter combination /f/, /a/, and /s/ is frequently used. The letters are initially selected based on their grouping as “easy letters” (Borkowski et al., 1967; Ross, 2003; Strauss et al., 2006), and the frequent use of these letters in the literature allows comparison to other datasets. For a more straightforward letter selection, a language-specific “high-frequency dictionary approach” can be used. In this approach, letters are selected based on their word-initial frequency in the target language, such as /k/, /a/, and /p/ in Finnish (Lehtinen et al., 2021; Mardani et al., 2020; Oberg & Ramírez, 2006; Schmid, 2011a; Tombaugh et al., 1999).

In SVF tasks, participants generate words within a specific semantic category. Unlike PVF tasks, SVF tasks typically focus on one category. Combining data from semantic categories is complex. As unique life experiences influence semantic memory organization, varied demographic and cultural settings result in diverse category sizes and content between populations (e.g., “items in a supermarket” can evoke varied semantic representations in different demographic groups) (Abwender et al., 2001; Olabarrieta-Landa et al., 2017; Roberts & Le Dorze, 1997; Rosselli et al., 2002; Strauss et al., 2006; Troyer, 2000). The category “animals” has been shown to be relatively neutral regarding culture and language (e.g., Pekkala et al., 2009). The category is widely used in linguistic- and clinical studies and in neuropsychological test batteries across languages (for an overview, see Strauss et al., 2006).

PVF and SVF tasks are commonly used to study various aspects of bilingualism, such as lexical access, vocabulary knowledge, dominance pattern of languages, the role of executive functions, and cross-linguistic fluency strategies (Luo et al., 2010; Marsh et al., 2019; Patra et al., 2020; Roberts & Le Dorze, 1997; Rosselli et al., 2002; Schmid & Köpke, 2009). PVF tasks engage strategic cognitive organization, initiation, inhibition, and maintenance of effort as participants conduct a non-routine search for words based on specific lexical representation (i.e., first letter) without the support of the hierarchical organization of semantic memory (Barry et al., 2008; Santos Nogueira et al., 2016; Strauss et al., 2006). In contrast, the SVF task requires a systematic lexical-semantic search—a relatively automatic process that resembles the everyday use of language (e.g., generating a shopping list). SVF task performance mainly relies on semantic categorization, hierarchical mental lexicon, and memory organization (Luo et al., 2010; Patra et al., 2020; Strauss et al., 2006).

This study in the context of existing literature

VF task performance is typically evaluated by the total number of correct responses (Strauss et al., 2006). A growing body of cross-disciplinary research aims to enhance the analytical power of VF tasks through more detailed data analyses (e.g., Becker & de Salles, 2016; Thiele et al., 2016). Based on these studies, Lehtinen et al. (2021) generated guidelines for a systematic approach to VF task administration and analysis, including scoring and calculating total scores, error analysis, investigating temporal parameters (e.g., words generated in 10-, 15-, or 30-second segments within the total time), and exploring clustering and switching strategies. In this study, we analyze total scores, errors, and temporal parameters in PVF and SVF performance in a Finnish-English language attrition population following these guidelines. Next, we will briefly review previous literature on VF task analysis in bilingual contexts before describing the present study in detail.

Total scores. PVF task performance measured by total score has been found to vary from a comparable performance between mono- and bilinguals (Rosselli et al., 2000, 2002; Soltani et al., 2021)

to bilinguals performing better than monolinguals (Patra et al., 2020). Similar PVF performance between mono- and bilinguals has been interpreted as less language interference in PVF compared with SVF (Rosselli et al., 2000, 2002). Stronger PVF performance has been linked to superior executive functions related to inhibiting language interference while switching between languages (Ljunberg et al., 2013; Luo et al., 2010; Marsh et al., 2019; Patra et al., 2020; Sandoval et al., 2010). However, PVF tasks are scarcely applied in language attrition studies (Jarvis, 2019, Chapter 21, p. 249; Schmid & Köpke, 2009).

In SVF tasks, bilingual populations consistently generate fewer words than monolinguals. This disadvantage in total scores has been attributed to language interference, weaker connections between lexical representations, or smaller vocabulary (Gollan et al., 2002; Rosselli et al., 2000; Sandoval et al., 2010; Schmid & Köpke, 2017; Soltani et al., 2021; Yılmaz & Schmid, 2018). Studies on language attrition populations line up with these findings, with attriters performing more poorly than monolinguals in SVF, though often with a relatively small effect size (e.g., Badstübner, 2011; Dostert, 2009; Opitz, 2011; Schmid & Dusseldorp, 2010; Schmid & Jarvis, 2014; Schmid & Keijzer, 2009; Schmid & Köpke, 2009). It has been shown that SVF total scores have limited predictive power in profiling individual speakers as attriters or monolinguals, and total scores are largely unrelated to LoR frequency of L1 use (Schmid, 2011a; Schmid & Jarvis, 2014; Schmid & Köpke, 2009). Therefore, a more comprehensive approach to VF data analyses is called for.

Errors. Errors are a part of VF task performance in all populations (Crowe, 1998; Sandoval et al., 2010). In bilingual environments, language intrusions are of particular interest (Gollan et al., 2002) as cross-language intrusions are directly connected to language interference (Gollan et al., 2011). Their presence can also indicate language dominance with more frequent language intrusion errors generated in the non-dominant language (Sandoval et al., 2010). In a language attrition population, Badstübner (2011) found that attriters generated more errors than monolinguals in SVF and described these errors as direct L2 intrusions and incorrect lexical items (such as partial recall of an L1 word). They detected no significant difference in the number of errors between the groups in PVF.

Temporal parameters. Analyzing the temporal parameters of VF tasks involves calculating the number of correct words produced within shorter time segments of the total time (e.g., four 15-second time segments within the 60 seconds). Typically, most words are generated in the early segments of the task across all populations. Approximately half of the words are produced in the first 15 seconds, facilitated by a semi-automatic rapid retrieval process that relies on semantic memory. As time progresses, a more effortful retrieval strategy is employed, engaging strategic executive processes, such as monitoring performance to avoid repetitions. This results in slower word-finding with fewer and more infrequent words (Fernaesus et al., 2008; Lehtinen et al., 2021; Sandoval et al., 2010; Venegas & Mansur, 2011). Thus, group differences in the early segments of the task suggest variations in language knowledge and lexical retrieval, while differences in later segments indicate discrepancies in executive control (Fernaesus & Almkvist, 1998; Fernaeus et al., 2008; Gurunandan et al., 2023; Luo et al., 2010).

Studies have consistently shown that bilinguals generate fewer words than monolinguals during the initial stages of a VF task, but the difference between the groups tends to diminish as the trial progresses (Luo et al., 2010; Sandoval et al., 2010). This pattern is interpreted to reflect more language interference in the early stages of the task, where bilinguals may produce high-frequency words they know in both languages (e.g., “cat”) rather than low-frequency words they may only know in one language (e.g., “bobcat”) (Luo et al., 2010; Sandoval et al., 2010). Interestingly,

similar neural activation patterns have been observed during L1 and L2 performance, suggesting a common bilingual effect in both languages (Gurunandan et al., 2023). Regarding language attrition studies, attriters have been shown to retrieve words more slowly than monolinguals in L1 SVF (Jarvis, 2019, Chapter 21, p. 243; Schmid & Jarvis, 2014; Schmid & Köpke, 2009). This finding has been linked to an increased load of inhibiting L2 competitors (Schmid & Jarvis, 2014; Yilmaz & Schmid, 2018) consistent with previous studies on bilingual VF performance (Gollan et al., 2011; Luo et al., 2010; Sandoval et al., 2010). However, highly automatic language skills are considered to be more resistant to language attrition (Goral, 2004; Segalowitz, 1991).

The present study

In the present study, we examine the processes that affect and potentially hinder optimal performance in L1 and L2 PVF and SVF tasks among a group of L1 language attriters who self-report as balanced bilinguals with a slight preference for L2. We conduct a systematic analysis for total scores, errors, and temporal parameters guided by the methodology proposed by Lehtinen et al. (2021). We also investigate how the LoR and frequency of L1 use affect task performance in both L1 and L2. To detect differences between attriters and monolingual speakers, we compare the L1 performance of attriters with a matched L1 monolingual group previously studied by Lehtinen et al. (2021).

Research questions

The research questions and hypotheses for the current study are as follows:

1. How do language attriters perform in PVF and SVF tasks for their first (L1) and second (L2) languages, and to what extent do LoR and frequency of L1 use affect their performance on the measures listed below?
2. How is the performance of language attriters in PVF and SVF tasks in their first language (L1) compared with that of a matched monolingual group across the measures listed below?

Measures:

- a. Total scores
- b. Errors: number, frequency, and distribution of error types
- c. Temporal parameters: performance change during the task (measured as the number of words generated in four 15 seconds segments)
- d. Ability to generate words rapidly in the early stages of the task (measured as the proportion of words generated within the initial 15-second interval of a task relative to the overall word count)

We will be referring to the different aspects of these research questions by referencing their numbers (RQ1 and RQ2) and letters (a, b, c, d) to specify which aspects of the data analysis is relevant.

Hypotheses

As stated above, while both, PVF and SVF task types engage a wide array of cognitive skills, PVF tasks emphasize executive skills, and SVF tasks rely more on semantic categorization, hierarchical mental lexicon, and memory organization. Based on literature, we expect more variation in the

processes required for SVF than PVF in our population, i.e., we expect differences between languages and groups to be more pronounced in SVF than in PVF. For RQ1, we anticipate that attriters will show comparable L1 and L2 performance or stronger performance in their self-reported dominant language (L2) in total scores. We predict shorter LoR and more frequent L1 use will positively affect L1 total scores in PVF and SVF but that these variables will not be strong predictors of performance independently. As participants are healthy, neurotypical adults, we anticipate minimal errors across languages. We hypothesize that attriters will produce more cross-language intrusion errors in their self-reported less dominant language (L1) and that a longer LoR and less frequent L1 use will result in an increased number of errors in L1 but not L2. Regarding temporal parameters, we hypothesize a similar performance profile across languages, and that attriters will more successfully employ rapid retrieval strategies in their self-reported stronger language (L2), particularly in SVF. We also predict that a shorter LoR and more frequent L1 use will facilitate rapid retrieval in L1 but that these variables will not be strong independent predictors of performance.

As for RQ2 on performance between attriters and monolinguals, we predict that attriters will achieve lower total scores in SVF than monolinguals but demonstrate comparable performance in PVF. We expect minimal errors overall, with a lower number of error-free trials in the attriter group. We anticipate that attriters will retrieve words more slowly than monolinguals in the early stages of the task, especially in SVF.

Method

Participants

Data from two healthy participant groups, attriters, and monolingual controls, were analyzed. The University of Turku Ethics committee approved all experimental procedures, and all participants provided informed consent before participating in the study. Exclusion criteria for both groups included (history of) cardiovascular, neurological, psychiatric, developmental language or speech disorders, toxic substance abuse, severe hearing loss, and age over 80 years.

Attriters

The attriter group ($N=38$) consisted of first-generation Finnish immigrants living in Northern California. Participants were recruited through local Finnish groups and associations in Northern California and through introductions from other participants.

The group's age was $M=60.90$ ($SD=8.42$, range 45–79). Education was measured at two tiers: no academic degree ($n=16$) and academic degree ($n=22$). All participants spoke Finnish as their first language (L1) and English as their second language (L2). They had immigrated after puberty (age at immigration $M=26.68$, $SD=7.38$, range 9–48) between 1948 and 1998. All participants had migrated of their own volition as adults or as older children with their families. Participants who arrived with their parents had chosen to stay in the United States as adults.

Participants had lived in an L2 environment for at least 20 years, and the majority had used L2 as their primary professional language during this time (*What has been the primary working language in your work career?* “Only English” 65.8% $n=25$, “both but mostly English” 26.3% $n=10$, “both, equally” 5.3% $n=2$ “both, but mostly Finnish” 2.6%, $n=1$ “only Finnish” 0% $n=0$).

At the time of the interview, all attriters identified themselves as bilinguals (*Do you see yourself as bilingual?* “no” $n=0$, “yes” $n=38$). The majority reported that their language balance had shifted while living in the United States (*Has the balance between your Finnish and English changed during the time you have lived in the USA?* “no” 13.5% $n=5$, “yes” 86.5%, $n=32$).

Most attriters reported that they had experienced a decline in their L1 proficiency (L1) (*Do you think your Finnish language proficiency has changed since you moved to the USA?* “Yes, I think it has become worse” 84.2% $n=32$, “no” 13.2% $n=5$, “yes, I think it has become better” 2.6% $n=1$). In terms of proficiency, the majority perceived their proficiency in L1 and L2 as equal (*Do you feel that you are equally proficient in Finnish and English?* “No, more proficient in English” 27%, $n=10$, “yes” 48.6%, $n=18$, “no, more proficient in Finnish” 24.3%, $n=9$). Most attriters were equally comfortable in speaking L1 and L2 with a higher number of attriters indicating that they felt more comfortable in speaking L2 than L1 (*Do you feel more comfortable speaking Finnish or English?* “English” 26.3% $n=10$, “Finnish,” 5.3% $n=2$, “no preference” 68.4% $n=26$).

When asked about language dominance, attriters reported the following distribution on a 5-point Likert scale (*Language dominance [i.e., the language you are “best” at]?* “1. English” 31.6% $n=12$, “2. mostly English” 13.2% $n=5$, 3. both Finnish and English 23.7% $n=9$, 4. mostly Finnish 21.1% $n=8$, “5. Finnish” 10.5% $n=4$). This suggests that, on average, the perceived dominance in the attriter group was balanced bilingualism with a slight preference for English (L2) ($M=2.66$ [$SD=1.40$, range 1–5]).

Monolingual group

The monolingual group ($N=50$) consisted of Finnish speakers who had always lived in Finland and used Finnish as their only language in their everyday life. Data from the monolingual group for a similar analysis were published previously (Lehtinen et al., 2021) and are referenced here for group comparison. The group’s average age was $M=62.58$ ($SD=7.59$, range 49–79). All participants identified as monolinguals, but due to language policy in Finland, all individuals were exposed to at least one foreign language in a school setting at a young age. In addition, passive exposure to foreign languages, notably Swedish and English, through media is assumed to be substantial in Finland, as television and movies are typically presented in their original language with Finnish subtitles.

Matched demographics

Monolingual participants were recruited to correspond to the age, education, and gender distribution of the attriter group to minimize demographic variables’ potential effect. Statistical analyses verified no significant differences between the groups for age (attriters $M=60.90$, monolinguals $M=62.60$) $z=-0.97$, $p=.332$, education, $\chi(1)=0.49$, $p=.27$, or gender, $\chi(1)=0.43$, $p=.51$.

Data

Data were extracted from a larger dataset consisting of five verbal fluency tasks (L1, L2), a WUG-task (Crystal, 2015) (L1), speech samples elicited via a film retelling task (L1, L2), and free speech samples (L1). In this study, we focused on the participants’ performance in four VF tasks in L1 and L2: three phonemic categories in L1 (Finnish, /k/, /a/, /p/) and L2 (English, /f/, /a/, /s/) and one concrete semantic category, “animals”.

Background information was collected via the Language Attrition Test Battery Sociolinguistic Questionnaire (SQ), introduced and discussed in detail in Schmid (2011b), Schmid and Cherciov (2019, Chapter 23, pp. 267–276), and Schmid and Dusseldorp (2010), and available in multiple languages on www.languageattrition.org. The first author translated the SQ into Finnish (Supplemental Appendix A). An abridged questionnaire for the control group was used to verify

language use in daily life (Supplemental Appendix B). Participants were given the option to fill out the questionnaire before or during the session.

All tasks were completed in a quiet environment (such as participants' home or clinical setting) in one sitting in Northern CA, USA for attriters, and Finland for the control group. Tasks were completed in L1 and L2 within a single session. All tasks were completed in one language and after a short break in the other. The order of languages was randomized, and participants were aware that the interlocutor was bilingual. Tasks were presented in randomized order within each language, with VF tasks presented in the following order: concrete category "animals," (abstract category "emotions"), phonemes "/k/, /a/, /p/" for L1 and "/f/, /a/, /s/" for L2.

This study focuses on three phonemic and one semantic VF tasks. For the phonemic task in L1, we followed the high-frequency dictionary approach (Lehtinen et al., 2021; Mardani et al., 2020; Oberg & Ramírez, 2006; Schmid, 2011a) selecting the two most frequent word-initial consonants of Finnish /k/ (15242 words) and /p/ (10640 words) and the most frequent word-initial vowel /a/ (4361 words) (Kielitoimiston sanakirja, 2021; Leskinen, 1989). For L2 (English), we selected phonemes /f/ (6939 words), /a/ (10360 words), and /s/ (19236 words) (Merriam-Webster, n.d.) based on their frequent use in literature (e.g., Strauss et al., 2006). For the SVF, task we included one semantic category, the culturally and linguistically relatively neutral, concrete category, "animals" (e.g., Pekkala et al., 2009), to limit the complexity of semantic analysis. The administration of the five VF tasks typically lasted for 5–10 minutes; the whole session was completed within 2 hours.

As extralinguistic variables, we examined the LoR in the L2 dominant environment in years ($M = 34.24$ years, $SD = 10.83$, range 20–70) and frequency of L1 use. For frequency of L1 use, we asked the participants to estimate their language use on a 5-point Likert scale, with the majority reporting weekly use of L1 (*How often do you speak Finnish?* "rarely" $n = 1$, "few times a year" $n = 1$, "monthly" $n = 7$, "weekly" $n = 22$, "daily" $n = 7$). We aimed to investigate L1 use for different domains, especially leisure versus professional use. However, only three participants reported using L1 for work regularly, making it unreliable to analyze the impact of professional use of L1.

Procedure

All VF tasks were administered and scored following the procedure outlined in Lehtinen et al. (2021). Briefly, participants were asked to produce as many different words as possible in 60 seconds following the given criteria, with the only restriction being for proper names. Responses were transcribed, and acceptable words were calculated for total scores, with semantically distinctive words calculated as separate items. Errors were excluded from the total score and classified into four categories: repetitions, categorical errors, nonwords, and language intrusions. Utterances in Finnish, a macaronic mixture of Finnish and English (e.g., *päntsit* "pants" [*standard Finnish housut*] see also Virtaranta, 1992), were scored as nonwords. Temporal parameters were analyzed by calculating the number of acceptable words in four 15-second segments (i.e., 0–15, 16–30, 31–45, and 46–60). An example of scored trials is shown in Supplemental Appendix C (L1 Table 1, L2 Table 2).

We utilized R software (R Core Team, 2019) for statistical modeling and data visualization with packages dplyr (Wickham et al., 2019), lme4 (Bates et al., 2015), sjPlot (Lüdtke, 2018), ggplot2 (Wickham, 2016), arsenal, and ggeffects (Lüdtke, 2018). The analysis script is available at https://osf.io/fue3k/?view_only=6b6762f07e2243d6b8548c0992dce9f1. Continuous predictors used as predictors were scaled and centered to sample means in all models.

To address RQ1.a. on overall performance in the attriters group, we used a linear mixed-effects model to examine total scores as a function of task language (L1/L2), task type (PVF/SVF), LoR,

and frequency of L1 use and their interactions. We included participant intercept and slope for task language as a random factor to account for individual differences in proficiency between Finnish and English. We selected predictors for the final model using a model comparison with Bayesian information criterion (BIC) values, choosing the simplest model when the two models did not differ significantly in BIC. To address RQ1b on the number of errors between L1 and L2, we used the Wilcoxon Rank-Sum Test. Due to the limited number of errors in the dataset, we excluded the impact of extralinguistic variables on errors from the analysis. To answer RQ1c on temporal patterns, we fitted a model to examine the attriters' performance changes over time during the 60 seconds. The model considered the number of correct words generated during 15-second time intervals and included task language, time sequence (0–15, 16–30, 31–45, or 46–60), task type, LoR, and L1 use as predictors. Participant intercept with a slope for task type was used as a random factor to account for individual-level differences. To analyze the ability to generate words rapidly in the early stages of the task (RQ1d), we modeled the ratio of correct words generated in the first segment to the total score. We used task language (L1/L2), LoR, and L1 use as predictors, and included participant intercept with a slope for task language as a random factor.

For RQ2a on the overall performance between attriter and monolingual groups, we used a linear mixed-effects model to model total scores as a function of task type (PVF/SVF) and participant group. Participant intercept with a slope for task type was included as a random factor to account for individual-level variation. The number of errors between groups in L1 (RQ2b) was analyzed using the Wilcoxon Rank-Sum Test. For group comparisons between temporal performance profiles between attriters and monolinguals (RQ2c), we fitted a model to investigate performance changes during the task, with the number of correct words generated during 15-second intervals as the outcome variable, and participant group, time sequence, and task type as predictors. We used the same model selection procedure as previously described. To compare the ability to generate words rapidly in the early stages of the task between attriters and monolinguals (RQ2d), we modeled the ratio of correct words generated in the first segment to the total score as the outcome variable, with task type and participant group as predictors, and included participant intercept with a slope for task type as a random factor. The model selection followed the procedure described earlier.

Results

For brevity, we have presented descriptive statistics and complete model summaries in Supplemental Appendix C, with statistics for key findings summarized in the following.

Research question 1

Our first research question was: How do language attriters perform in PVF and SVF tasks for their first (L1) and second (L2) languages, and to what extent do LoR and frequency of L1 use affect performance?

Total scores. To analyze total scores in VF tasks, the model selection procedure based on BIC values (L1 use, LoR, or both, combined with task language and fluency type) suggested the model with L1 use but no LoR as the best fit for the data (Supplemental Appendix C, Table 3 and Model Summary 1). Our analysis showed statistically significant main effects of task language, fluency type, and L1 use, indicating that more words were generated in L1 than in L2 and in semantic compared with phonemic tasks. Frequent L1 use was associated with better overall performance in the tasks across fluency types and task languages (1.46, 95% CI=[0.36, 2.55], $t = 2.62$, $p = .009$).

We found no significant interactions between L1 use and task type and/or task languages. VF performance for attriters in L1 and L2 is visualized in Figure 1.

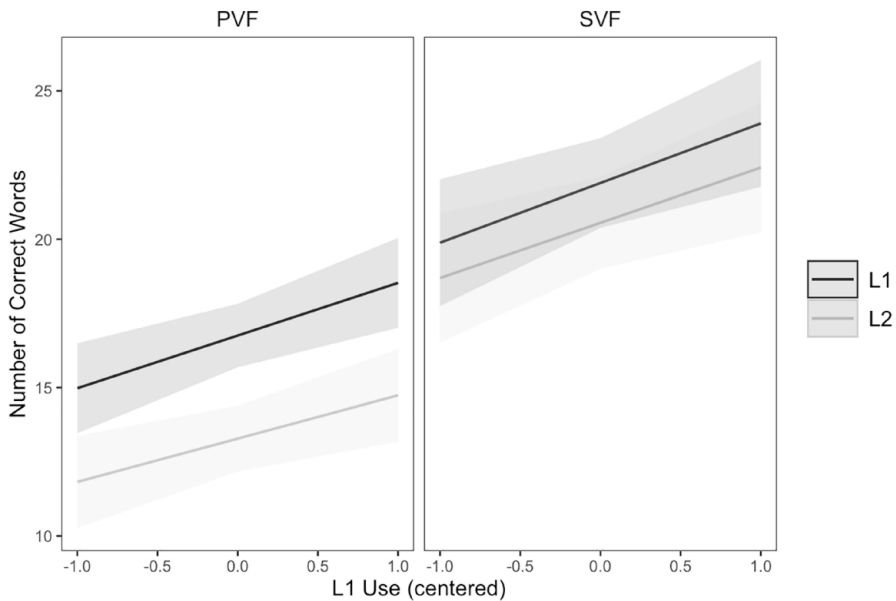


Figure 1. Predicted values of number of correct words in the attriter group in one semantic and three phonemic tasks (Combined) in L1 and L2 with frequency of L1 use as a predictor.

The number of errors, frequency, and distribution of error types. Table 1 presents the descriptive statistics for errors in L1 and L2 for the attriter group. Overall, the number of errors was minimal, which led us to exclude the impact of extralinguistic variables on errors from our analysis.

Table 1. Number of attriters who generate errors, total number of errors, and distribution of error types in one semantic and three phonemic verbal fluency tasks in L1 and L2.

Errors	L1				L2			
	/k/	/a/	/p/	/animals/	/f/	/a/	s	/animals/
<i>n</i> with errors	27 (71%)	22 (58%)	26 (68%)	15 (39%)	22 (59%)	25 (68%)	17 (46%)	17 (46%)
Total nr. of errors	43	40	48	27	41	36	20	26
Mean	1.13	1.05	1.26	0.71	1.11	0.97	0.54	0.70
SD	1.14	1.21	1.25	1.27	1.26	0.96	0.73	1.02
Range	0–5	0–4	0–4	0–6	0–4	0–3	0–3	0–4
Error type								
Repetition	21 (49%)	27 (68%)	34 (71%)	18 (66%)	26 (63%)	16 (44%)	10 (50%)	18 (69%)
Categorical	12 (28%)	12 (30%)	5 (10%)	5 (19%)	10 (24%)	15 (42%)	7 (35%)	6 (23%)
Nonword	10 (23%)	1 (3%)	9 (19%)	0 (0%)	4 (10%)	4 (11%)	2 (10%)	0 (0%)
Intrusion	0 (0%)	0 (0%)	0 (0%)	4 (15%)	1 (2%)	1 (3%)	1 (5%)	2 (9%)

Note. L1 N=38; L2 N=37.

As expected, the most common error type was repetition. There was no significant difference in the number of errors between L1 and L2 in PVF ($W = 5573, p = .1007$) or SVF ($W = 731, p = .741$) in the attriters group.

Performance change during the task. To examine the temporal profile and its association with task attributes (SVF / PVF and L1 / L2), we employed a model to predict the total number of acceptable words produced during four 15-second time windows, with L1 use and LoR as predictors. Using the model comparison procedure described earlier, we determined that a model with L1 use but without LoR was the most parsimonious (Supplemental Appendix C, Table 3 and Model Summary 2). The selected model showed that more words were generated at the beginning of the task than in the later segments in both L1 and L2, with no significant difference between languages, as illustrated in Figure 2. As in RQ1a (Model Summary 1) a main effect of L1 use was detected. However, there were no significant interactions between L1 use and temporal parameters, suggesting that the observed pattern was similar across languages despite the L1 use frequency.

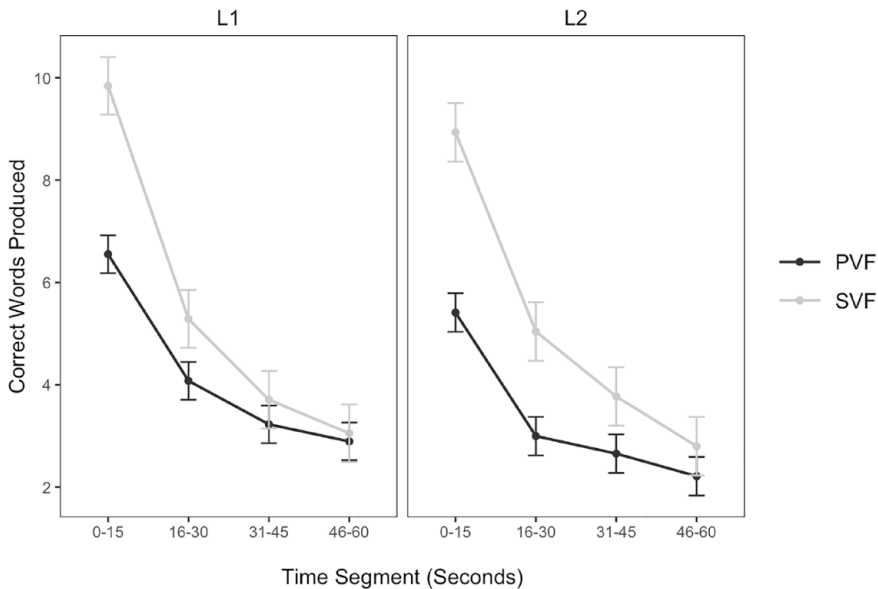


Figure 2. Predicted values of correct words in the four 15-second time segments in one semantic and three phonemic verbal fluency tasks in the attriter group in L1 and L2. Error bars represent 95% confidence intervals.

Ability to generate words rapidly in the early stages of the task. To further examine performance variation in the attriters group, we examined the quotient of acceptable words produced during the first time segment in all tasks as a function of task language and L1 use (Supplemental Appendix C, Model Summary 3). Our analysis revealed that attriters generated a smaller quotient in L1 (42%) than in L2 (44%) during the first time segment, although the difference between languages was not statistically significant. Moreover, our model showed that attriters who reported using L1 more frequently in their everyday life produced a smaller quotient in both languages than those who reported using L1 less frequently ($-0.02, 95\% \text{ CI} = [-0.04, -0.01], t = -2.52, p = .012$), illustrated in Figure 3.

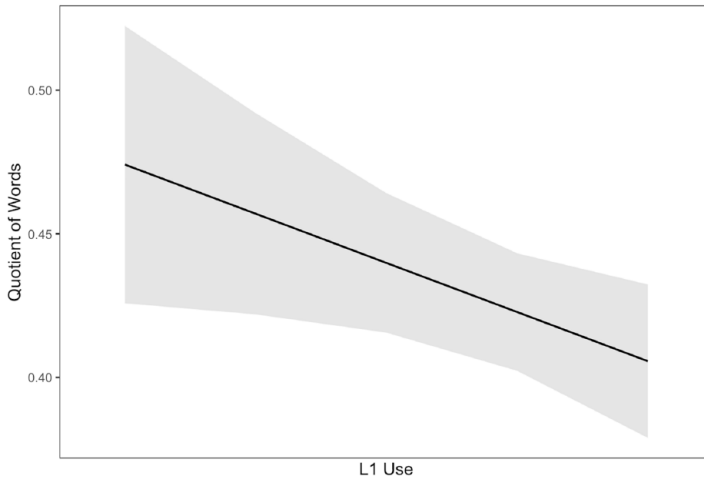


Figure 3. Predicted values of the quotient of correct words in the first 15-second time segment in the attriters group in L1 and L2 (Combined) in all task types (combined) with frequency of L1 use as predictor.

Research question 2

Our second research question was: How is the performance of language attriters in PVF and SVF tasks in their first language (L1) compared with that of a matched monolingual group?

Total scores. For group comparison for overall performance measured in total scores, the results indicated that monolinguals outperformed attriters in the L1 SVF task (3.65, 95% CI=[1.40, 5.91], $t=3.19, p=.002$). The performance of both groups was similar in the PVF task (Supplemental Appendix C, Tables 3 and 4, Model Summary 4). Results are shown in Figure 4.

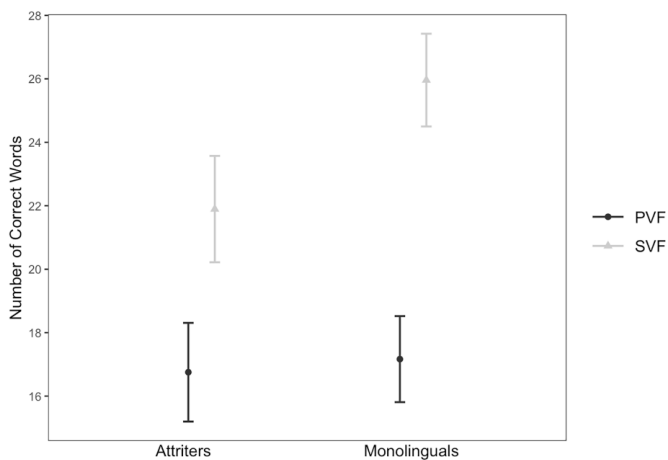


Figure 4. Comparison of predicted values of number of correct words in one semantic and three phonemic verbal fluency tasks (Combined) in L1 between the attriter and monolingual groups using participant group as a predictor.

The number of errors, frequency, and distribution of error types. For the monolingual group, results published previously in Lehtinen et al. (2021) are referenced here. In the PVF task, approximately half of the monolingual participants generated errors in all three trials (/k/ 44% [$n=22$]; /a/ 58% [$n=29$]; /p/ 50% [$n=25$]). Compared with monolinguals, a higher percentage of attriters generated errors in L1 PVF trials /k/ (71% [$n=27$]) and /p/ (68% [$n=26$]), and the percentage was similar in trial /a/ (58% [$n=22$]). Attriters generated significantly more errors than monolinguals in PVF (monolinguals $Md=1$, attriters $Md=1$, $W=9937.5$, $p=.016$). The most common error type was repetitions across tasks in both groups, followed by nonword errors in the attriter group. In SVF more attriters (39% [$n=15$]) than monolinguals (28% [$n=14$]) generated errors. While attriters generated more errors than monolinguals numerically (27 vs. 20), this difference was not statistically significant (monolinguals $Md=0$, attriters $Md=0$, $W=1062.5$, $p=.255$).

Performance change during the task. To investigate the temporal patterns of task performance and their relationship to task type in attriter and monolingual groups, we modeled the total number of acceptable words as a function of the number of acceptable words generated during four 15-second time segments (Supplemental Appendix C, Tables 3 and 4, Model Summary 5). Our analysis showed no significant differences between the two groups in the distribution of words across the four time segments, as demonstrated in Figure 5.

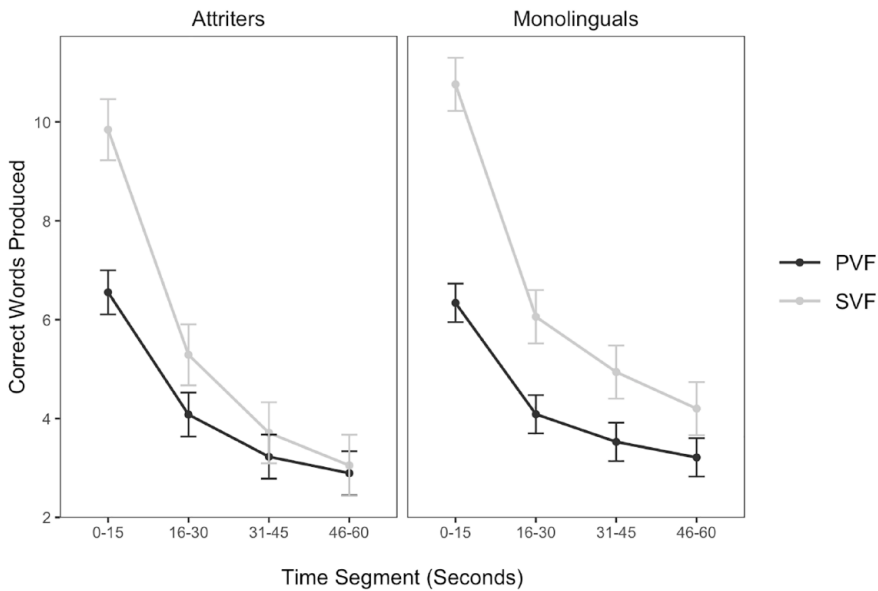


Figure 5. Predicted values of words produced in the four 15-second time segments in one semantic and three phonemic verbal fluency tasks in the attriter and monolingual groups. Error bars represent 95% confidence intervals.

Ability to generate words rapidly in the early stages of the task. We also investigated the performance of attriters and monolinguals during the initial 15-second of all tasks by modeling the quotient of correct words produced as a function of the group (Supplemental Appendix C, Model Summary 6). The analysis revealed that the quotient was higher in attriters (43%) than in monolinguals (41%)

(-0.02 , 95% CI= $[-0.04, 0.00]$, $t=-2.29$, $p=.022$), suggesting that attriters were able to produce more correct words in the initial stages of the task.

Discussion

We set out to explore processes that underlie performance in verbal fluency tasks in Finnish-English mature immersed bilinguals (i.e., language attriters), who self-identify as balanced bilinguals with a slight preference for L2. We analyzed PVF and SVF data in both their first (L1) and second (L2) languages, focusing on total scores, errors, and temporal parameters, and assessed the impact of immersion duration (LoR), frequency of L1 use, and their interactions on task performance. In addition, we contrasted the attriters' performance in L1 verbal fluency tasks to that of monolinguals to identify potential language attrition markers.

Our key findings show that attriters generated more acceptable words in their first language than in their second language across fluency types and that frequent L1 use supported performance in both languages (RQ1a, Model Summary 1, Figure 1). Attriters made very few errors in either language, with the most typical error type being repetitions (RQ1b). The temporal performance profile was similar for both languages (RQ1c, Model Summary 2, Figure 2), and those attriters who used L1 more frequently generated a smaller quotient of words in the initial stage of VF tasks in both languages than those who used L1 less frequently (RQ1d, Model Summary 3, Figure 3).

Compared with monolingual L1 performance, attriters generated fewer acceptable words in the L1 semantic task than monolinguals, and groups performed similarly in the phonemic task (RQ2a, Model Summary 4, Figure 4). Attriters made more errors than monolinguals in the phonological task, but the number of errors was comparable in the semantic task between groups. Qualitatively fewer attriters than monolinguals generated error-free trials in both tasks, with the most typical error in both groups being repetitions (RQ2b). The temporal performance profile in L1 did not differ between attriters and monolinguals (RQ2c, Model Summary 5, Figure 5). However, attriters generated more correct words in the initial stage of the task compared with monolinguals (RQ2d, Model Summary 6).

In the following, we discuss our findings in relation to earlier literature. We focus on performance between L1 and L2 in the attriter group before moving on to group comparison between the attriter and monolingual groups in L1. Finally, we outline suggestions for future research and address some limitations of this study.

RQ1. Performance between L1 and L2 within the attriter group

Attriters demonstrated strong L1 proficiency by generating higher total scores in L1 than in L2 across tasks. This points in the direction of L1 as the more dominant language, contrary to the attriters' self-reported language dominance (Roberts & Le Dorze, 1997; Rosselli et al., 2002). The small number of errors across tasks did not allow robust statistical analysis of errors, but some observations can be made. While the number of errors was similar across languages, qualitatively, attriters generated more language intrusions in L1 than in L2, suggesting transference from L2 to L1 but not vice versa. As attriters demonstrated stronger overall performance in L1, this contradicts our hypothesis of increased intrusion errors in the non-dominant language (Sandoval et al., 2010). Overall, by generating more correct words and allowing more language intrusions in L1 than L2, attriters demonstrate robust language proficiency but greater flexibility in L1 performance compared with L2, potentially as a marker of cross-linguistic influence of L2 (Schmid & Cherciov, 2019, Chapter 23, p. 267; Schmid & Dusseldorp, 2010; Sharwood Smith, 2019, Chapter 8, p. 85).

For temporal performance profiles, we anticipated similar performance across languages or faster initiation in the more dominant language among the attriters group (Fernaesus & Almkvist, 1998; Fernaeus et al., 2008; Gollan et al., 2011; Luo et al., 2010; Sandoval et al., 2010; Schmid & Jarvis, 2014; Yilmaz & Schmid, 2018). We found comparable temporal word distributions across languages and equally efficient lexical retrieval in the initial stage of both languages, indicating balanced language dominance.

Contrary to our hypothesis, there was no interaction effect for LoR and frequency of L1 use, but frequent L1 use supported overall performance and proportionally slowed down rapid retrieval in both languages. Previous research (e.g., Gurunandan et al., 2023) has shown that similar neural activation patterns occur during L1 and L2 VF performance, suggesting a common bilingual effect for both languages. Our results of similar effect of frequency of L1 use for both languages rather than a specific language may reflect the influence bilingualism has on overall cognitive and neural language processing. Thus, our findings highlight the importance of including both languages, L1 as well as L2 in language attrition studies to account for the general effect of bilingualism.

In addition to the impact of frequency of L1 use, we also aimed to investigate to what extent LoR affects VF performance in L1 and L2. However, our model selection procedure suggested a model without LoR to be the most parsimonious fit to our data and we were not able to directly investigate the effect of LoR on task performance. In a similar vein to the similar impact of frequency of L1 use for both languages, this omission might suggest that LoR was not a meaningful predictor for VF task performance, in either the first or the second language.

RQ2. Performance between the attriter and the monolingual group

Our findings between attriters and monolinguals mirror literature. Monolinguals outperformed attriters in SVF, and groups performed similarly in PVF (e.g., Badstübner, 2011; Dostert, 2009; Opitz, 2011; Schmid, 2011b, 2019; Schmid & Dusseldorp, 2010; Schmid & Jarvis, 2014; Schmid & Keijzer, 2009; Schmid & Köpke, 2009). As expected, there were fewer error-free trials in the attriter group, especially in trials /k/ and /p/, with emphasis on repetitions and nonword errors. Numerically attriters generated more errors in the PVF than monolinguals. Thus, attriters experienced some difficulty in rapid lexical retrievals compared with monolinguals, potentially due to language interference or contact-induced language change that manifested as nonword errors.

Contrary to our hypothesis, performance in the attriter group was not slowed down in the early stages of the task compared with monolinguals. Thus, attriters did not demonstrate markers of language interference in temporal analysis that would explain the differences in overall SVF task performance, as suggested by earlier studies (Rosselli et al., 2000, 2002; Sandoval et al., 2010; Schmid & Jarvis, 2014). Moreover, attriters generated a higher percentage of total words in the first 15-second segment than monolinguals. As word-finding is facilitated by semi-automatic rapid retrieval strategies in the early segments of the task (Fernaesus et al., 2008; Lehtinen et al., 2021; Sandoval et al., 2010; Venegas & Mansur, 2011), our results suggest that attriters relied more on semi-automatic rapid retrieval strategies in L1 than monolinguals, implying not only that these strategies are resistant to language attrition (Gollan et al., 2002; Segalowitz, 1991), but also that in the first 15-second time segment attriters may utilize rapid retrieval more efficiently than monolinguals. Interestingly, this is highlighted by the finding that while frequent L1 use slowed attriters down in the first 15-second segment, they still performed better than monolinguals in the first 15-second segment, with the overall performance resulting in lower total scores in SVF. Investigating the processes underlying these observations is beyond the scope of this study. In the following, we discuss suggestions for future studies and address the limitations of the present study.

Limitations of the present study, and suggestions for future research

As our sample was small, interpretations of the data are preliminary explorations. Future studies with larger datasets are needed to investigate the trends found in our data reliably. This is particularly relevant for error analysis. Due to the small amount of data, we were not able to conduct a robust statistical analysis or investigate the distribution profile of errors within tasks. In future studies, examining the distribution of errors over time may provide valuable insight into the performance differences in the initial stage between attriters and monolinguals.

Regarding error analysis, interpreting fine-grained data on error types from studies using different categorization systems can be challenging. For instance, our findings contrast with Badstübners' (2011) findings, which found a significant difference between monolinguals and attriters only in a SVF task. However, our analysis cannot be directly compared with Badstübners'. They examined combined data from three semantic categories and found that most incorrect lexical items in SVF resulted from L2 transfer in the "things in the kitchen" category (i.e., words related to everyday life). In addition to complexities in combining data from different semantic categories, in environments where L1 is used infrequently (or not at all), words related to everyday life are highly activated in L2. Consequently, Badstübners' (2011) category related to everyday life objects might have prompted more language interference from L2 than the less frequently used category of "animals" in our study.

In addition, we categorized non-standard utterances as nonword errors. These included terms that could be classified as Finglish (a macaronic blend of Finnish and English, as described in Virtaranta, 1992), which might also be considered L2 language intrusions, indicators of substantial language interference, or signs of language evolution within the attriters group in L1. As such, a distinct category for this type of error could have been informative, especially in future studies conducted with larger datasets. We conducted our analysis following Lehtinen et al. (2021) and remain optimistic that this qualitative data analysis can serve as a foundation for future research.

When interpreting our results, it is important to consider the potential impact of shared cognates on language interference. The languages in this study, L1 (Finnish) and L2 (English) have very few cognates. As Schmid and Jarvis (2014) demonstrate, there may be less competition between high-frequency words in Finnish and English compared with language pairs with a larger shared vocabulary. Thus, it is possible that in our dataset, the effect of language interference on rapid lexical retrieval was more subtle, and our analysis only detected it as a within-group variation effect linked to the frequency of L1 use and not at the group level.

Our analysis consisted of two groups: attriters and monolinguals. A comprehensive analysis of the monolingual group VF task performance was previously reported by Lehtinen et al. (2021). They found no effect for age but detected a positive effect of education and gender on VF task total scores. For the present study, the monolingual and attriter groups were matched for age, gender, and education. Although we are confident that closely matched groups effectively control for age, gender, and education in group comparisons, it is possible that these variables influenced the performance in the attriter group in ways not accounted for in our analyses. Controlling for education and gender within the attriter group, especially for the impact of L1 use on rapid retrieval, could have strengthened our data interpretation. We recommend controlling for age and education in future studies to ensure a more comprehensive interpretation of the results.

Limitations of this study include measuring L1 use only as an overall measure of the frequency of use in everyday life. Based on earlier research (e.g., Schmid & Dusseldorp, 2010), our goal was to explore the impact of L1 use for professional purposes on VF performance. Unfortunately, our dataset did not allow for such an analysis. A more detailed approach to L1 use might help explain variation in task performance within the attriter group. However, assessing the influence of

extralinguistic variables on observed phenomena is complex. Language history reports on L1 use typically involve subjective self-reports spanning several years and varying external circumstances that dictate the use of L1 in everyday life. These reports should be considered as the best available approximations, but their comparability between individuals is not straightforward (Bylund & Ramírez-Galan, 2016; Köpke & Schmid, 2004). Furthermore, LoR and L1 use are interconnected in many ways. While the time elapsed from leaving the L1 country is measurable, assessing the level of deprivation of L1 during that time is more challenging, especially in datasets dating to modern times (Schmid, 2019). Despite the limitations, exploring the impact of shared language history in language attrition populations is valuable beyond research interests. Gaining insight into how these factors affect L1 performance can motivate and support individuals in preserving their native language and cultural identity.

Conclusions

The present study showed that attriters, who self-report as balanced bilinguals with a slight preference for L2, demonstrate strong proficiency in L1 with similar lexical retrieval strategies in L1 and L2. Our analysis suggests balanced bilingualism with a subtle emphasis for L1 at the group level after 20 years of immersive exposure to L2, partly contrary to participant self-reports. We showed that frequent L1 use supports overall VF task performance but proportionally slows performance down in the initial stage of the task in both languages without significantly impacting overall performance. Compared with monolinguals, attriters demonstrate an overall disadvantage in SVF, but this disadvantage is not due to a slower initiation profile, as hypothesized. In contrast, attriters rely on rapid retrieval in L1 more than monolinguals. These findings add to our understanding of how attriters and monolinguals approach verbal fluency tasks and highlight the potential importance of early task performance in VF task analysis.

Our findings support the notion of two-way interaction in cognitive language processing, acquisition, attrition and dominance shifts in a bilingual environment. In the future, analyzing processes underlying VF performance in-depth in L1 and L2 using techniques like clustering and switching analysis (Lehtinen et al., 2021; Troyer, 2000; Troyer et al., 1997) could increase our understanding of lexical retrieval strategies in language attrition populations and the role of frequent L1 use in bilingual language processing.

Acknowledgements

The authors thank MPhil Ida Luotonen for her contributions to data analysis and overall support, and Professor Marja-Liisa Helasvuo and Dr Raymond Bertram for scientific advice and valuable comments and suggestions on earlier versions of this manuscript. They would also like to thank Professor Monika Schmid for inspiration and support during this study. Finally, they thank PhD Lotta Weckström for her efforts in recruiting participants, and express their gratitude to the participants of this study for the generous donation of their time and effort in participating in this research.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by The Turku University Foundation, The Finlandia Foundation,

The Migration Institute of Finland, The University of Turku Graduate School, and The Alfred Kordelin Foundation.

ORCID iDs

Nana Lehtinen  <https://orcid.org/0000-0002-3606-1302>

Anna Kautto  <https://orcid.org/0000-0003-4885-9167>

Supplemental material

Supplemental material for this article is available online.

References

- Abwender, D. A., Swan, J. G., Bowerman, J. T., & Connolly, S. W. (2001). Qualitative analysis of verbal fluency output: Review and comparison of several scoring methods. *Assessment*, 8(3), 323–338. <https://doi.org/10.1177/107319110100800308>
- Badstübner, T. (2011). *L1 attrition: German immigrants in the U.S.* [Doctoral thesis, The University of Arizona].
- Barry, D., Bates, M. E., & Labouvie, E. (2008). FAS and CFL forms of verbal fluency differ in difficulty: A meta-analytic study. *Applied Neuropsychology*, 15(2), 97–106. <https://doi.org/10.1080/09084280802083863>
- 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/10.18637/jss.v067.i01>
- Becker, N., & de Salles, J. F. (2016). Methodological criteria for scoring clustering and switching in verbal fluency tasks. *Psico-USF*, 21(3), 445–457. <https://doi.org/10.1590/1413-82712016210301>
- Borkowski, J. G., Benton, A. L., & Spreen, O. (1967). Word fluency and brain damage. *Neuropsychologia*, 5(2), 135–140. [https://doi.org/10.1016/0028-3932\(67\)90015-2](https://doi.org/10.1016/0028-3932(67)90015-2)
- Bylund, E., & Ramírez-Galan, P. (2016). Language aptitude in first language attrition: A study on late Spanish-Swedish bilinguals. *Applied Linguistics*, 37(5), 621–638. <https://doi.org/10.1093/applin/amu055>
- Crowe, S. F. (1998). Decrease in performance on the verbal fluency test as a function of time: Evaluation in a young healthy sample. *Journal of Clinical and Experimental Neuropsychology*, 20(3), 391–401. <https://doi.org/10.1076/jcen.20.3.391.810>
- Crystal, D. (2015). *A dictionary of linguistics and phonetics*. <http://search.credoreference.com/content/title/bkdictling>
- de Bot, K., & 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. <https://doi.org/10.1080/01434632.1994.9994554>
- 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(1), 33–40. <https://doi.org/10.1017/S1366728909990289>
- Dostert, S. C. (2009). *Multilingualism, L1 attrition and the concept of “native speaker”* [Doctoral thesis, Heinrich Heine University Düsseldorf].
- Fernaes, S.-E., & Almkvist, O. (1998). Word production: Dissociation of two retrieval modes of semantic memory across time. *Journal of Clinical and Experimental Neuropsychology*, 20(2), 137–143. <https://doi.org/10.1076/jcen.20.2.137.1170>
- Fernaes, S.-E., Östberg, P., Hellström, Å., & Wahlund, L.-O. (2008). Cut the coda: Early fluency intervals predict diagnoses. *Cortex*, 44(2), 161–169. <https://doi.org/10.1016/j.cortex.2006.04.002>
- Gollan, T. H., Montoya, R. I., & Werner, G. A. (2002). Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology*, 16(4), 562–576. <https://doi.org/10.1037/0894-4105.16.4.562>
- Gollan, T. H., Sandoval, T., & Salmon, D. P. (2011). Cross-language intrusion errors in aging bilinguals reveal the link between executive control and language selection. *Psychological Science*, 22(9), 1155–1164. <https://doi.org/10.1177/0956797611417002>

- Goral, M. (2004). First-language decline in healthy aging: Implications for attrition in bilingualism. *Journal of Neurolinguistics*, 17(1), 31–52. [https://doi.org/10.1016/S0911-6044\(03\)00052-6](https://doi.org/10.1016/S0911-6044(03)00052-6)
- Grosjean, F. (2013). Bilingualism: A short introduction. In F. Grosjean & P. Li (Eds.), *The psycholinguistics of bilingualism* (Vol.5, pp. 5–25). John Wiley.
- Gurunandan, K., Carreiras, M., & Paz-Alonso, P. M. (2023). Verbal production dynamics and plasticity: Functional contributions of language and executive control systems. *Cerebral Cortex*, 33, 740–753. <https://doi.org/10.1093/cercor/bhac097>
- Jarvis, S. (2019). Lexical attrition. In M. S. Schmid, & B. Köpke (Eds.), *The Oxford handbook of language attrition* (pp. 240–250). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198793595.013.20>
- Kasparian, K., & Steinhauer, K. (2017). On missed opportunities and convenient “truths.” *Linguistic Approaches to Bilingualism*, 7(6), 709–714. <https://doi.org/10.1075/lab.00010.kas>
- Kielitoimiston sanakirja. (2021, October 28). [Dictionary of standard Finnish compiled in the institute for the languages of Finland]. <https://www.kielitoimistonsanakirja.fi>
- Köpke, B., & Schmid, M. S. (2004). Language attrition: The next phase. In M. S. Schmid, B. Köpke, M. Keijzer, & L. Weilemar (Eds.), *Studies in bilingualism* (Vol.28, p. 1). John Benjamins Publishing Company. <https://doi.org/10.1075/sibil.28.02kop>
- Laine, M., & Lehtonen, M. (2018). Cognitive consequences of bilingualism: Where to go from here? *Language, Cognition and Neuroscience*, 33(9), 1205–1212. <https://doi.org/10.1080/23273798.2018.1462498>
- Lehtinen, N., Luotonen, I., & Kautto, A. (2021). Systematic administration and analysis of verbal fluency tasks: Preliminary evidence for reliable exploration of processes underlying task performance. *Applied Neuropsychology: Adult*. Advance online publication. <https://doi.org/10.1080/23279095.2021.1973471>
- Leskinen, H. (1989). Tietoja sananalkuisten grafeemien ja grafeemikombinaatioiden yleisyyssuhteista [About the frequency of word-initial graphemes and grapheme combinations]. *Virittäjä*, 93(3), 401–419. <https://journal.fi/virittaja/article/view/38309>
- Linck, J. A., & Kroll, J. F. (2019). Memory retrieval and language attrition: Language loss or manifestations of a dynamic system? In M. S. Schmid & B. Köpke (Eds.), *The Oxford handbook of language attrition* (pp. 87–97). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198793595.013.8>
- Ljunberg, J. K., Hansson, P., Andrés, P., Josefsson, M., & Nilsson, L.-G. (2013). A longitudinal study of memory advantages in bilinguals. *PLOS ONE*, 8(9), e73029. <https://doi.org/10.1371/journal.pone.0073029>
- Lüdecke, D. (2018). *sjPlot—Data visualization for statistics in social science*. Zenodo. <https://doi.org/10.5281/ZENODO.1308157>
- Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition*, 114(1), 29–41. <https://doi.org/10.1016/j.cognition.2009.08.014>
- Mardani, N., Pourjafari, M., Irandegani, M. A., Ahmadi, N., & Baghban, K. (2020). The effect of bilingualism on the processing of clustering and switching in verbal fluency tasks. *Journal of Rehabilitation Sciences & Research*, 7(3), 114–117. <https://doi.org/10.30476/jrsr.2020.86911.1096>
- Marsh, J. E., Hansson, P., Sörman, D. E., & Ljungberg, J. K. (2019). Executive processes underpin the bilingual advantage on phonemic fluency: Evidence from analyses of switching and clustering. *Frontiers in Psychology*, 10, Article 1355. <https://doi.org/10.3389/fpsyg.2019.01355>
- Merriam Webster. (n.d.). [Official SCRABBLE® players dictionary]. Retrieved August 30, 2023, from <https://scrabble.merriam.com>
- Oberg, G., & Ramírez, M. (2006). Cross-linguistic meta-analysis of phonological fluency: Normal performance across cultures. *International Journal of Psychology*, 41(5), 342–347. <https://doi.org/10.1080/00207590500345872>
- Olabarrieta-Landa, L., Torre, E. L., López-Mugartza, J. C., Bialystok, E., & Arango-Lasprilla, J. C. (2017). Verbal fluency tests: Developing a new model of administration and scoring for Spanish language. *NeuroRehabilitation*, 41(2), 539–565. <https://doi.org/10.3233/NRE-162102>

- Opitz, C. (2011). *First language attrition and second language acquisition in a second language environment* [Doctoral thesis, Trinity College, University of Dublin]. <http://hdl.handle.net/2262/78154>
- Patra, A., Bose, A., & Marinis, T. (2020). Performance difference in verbal fluency in bilingual and monolingual speakers. *Bilingualism: Language and Cognition*, 23(1), 204–218. <https://doi.org/10.1017/S1366728918001098>
- Pekkala, S., Goral, M., Hyun, J., Obler, L. K., Erkinjuntti, T., & Albert, M. L. (2009). Semantic verbal fluency in two contrasting languages. *Clinical Linguistics & Phonetics*, 23(6), 431–445. <https://doi.org/10.1080/02699200902839800>
- R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Roberts, P. M., & Le Dorze, G. (1997). Semantic organization, strategy use, and productivity in bilingual semantic verbal fluency. *Brain and Language*, 59(3), 412–449. <https://doi.org/10.1006/brln.1997.1753>
- Ross, T. P. (2003). The reliability of cluster and switch scores for the Controlled Oral Word Association Test. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, 18(2), 153–164.
- Rosselli, M., Ardila, A., Araujo, K., Weekes, V. A., Caracciolo, V., Padilla, M., & Ostrosky-Solí, F. (2000). Verbal fluency and repetition skills in healthy older Spanish-English bilinguals. *Applied Neuropsychology*, 7(1), 17–24. https://doi.org/10.1207/S15324826AN0701_3
- Rosselli, M., Ardila, A., Salvatierra, J., Marquez, M., Luis, M., & Weekes, V. A. (2002). A cross-linguistic comparison of verbal fluency tests. *International Journal of Neuroscience*, 112(6), 759–776. <https://doi.org/10.1080/00207450290025752>
- Runnqvist, E., & Costa, A. (2012). Is retrieval-induced forgetting behind the bilingual disadvantage in word production? *Bilingualism: Language and Cognition*, 15(2), 365–377. <https://doi.org/10.1017/S1366728911000034>
- Sandoval, T. C., Gollan, T. H., Ferreira, V. S., & Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition*, 13(2), 231–252. <https://doi.org/10.1017/S1366728909990514>
- Santos Nogueira, D., Azevedo Reis, E., & Vieira, A. (2016). Verbal fluency tasks: Effects of age, gender, and education. *Folia Phoniatrica et Logopaedica*, 68(3), 124–133. <https://doi.org/10.1159/000450640>
- Schmid, M. S. (2011a). Contact x time: External factors and variability in L1 attrition. In M. S. Schmid & W. Lowie (Eds.), *Studies in bilingualism* (Vol.43, pp. 155–176). John Benjamins Publishing Company. <https://doi.org/10.1075/sibil.43.11sch>
- Schmid, M. S. (2011b). *Language attrition*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511852046>
- Schmid, M. S. (2019). The impact of frequency of use and length of residence on L1 attrition. In M. S. Schmid & B. Köpke (Eds.), *The Oxford handbook of language attrition* (pp. 287–303). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198793595.013.24>
- Schmid, M. S., & Cherciov, M. (2019). Introduction to extralinguistic factors in language attrition. In M. S. Schmid & B. Köpke (Eds.), *The Oxford handbook of language attrition* (pp. 265–276). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198793595.013.22>
- 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., & Jarvis, S. (2014). Lexical access and lexical diversity in first language attrition. *Bilingualism: Language and Cognition*, 17(4), 729–748. <https://doi.org/10.1017/S1366728913000771>
- Schmid, M. S., & Keijzer, M. (2009). First language attrition and reversion among older migrants. *International Journal of the Sociology of Language*, 2009(200), 83–101. <https://doi.org/10.1515/IJSL.2009.046>
- Schmid, M. S., & Köpke, B. (2009). L1 attrition and the mental lexicon. In A. Pavlenko (Ed.), *The bilingual mental lexicon: Interdisciplinary approaches* (pp. 209–238). Multilingual Matters.
- Schmid, M. S., & Köpke, B. (2017). The relevance of first language attrition to theories of bilingual development. *Linguistic Approaches to Bilingualism*, 7(6), 637–667. <https://doi.org/10.1075/lab.17058.sch>

- Schmid, M. S., & Köpke, B. (Eds.). (2019). *The Oxford handbook of language attrition* (1st ed.). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198793595.001.0001>
- Schmid, M. S., & Yılmaz, G. (2018). Predictors of language dominance: An integrated analysis of first language attrition and second language acquisition in late bilinguals. *Frontiers in Psychology, 9*, Article 1306. <https://doi.org/10.3389/fpsyg.2018.01306>
- Segalowitz, N. (1991). Does advanced skill in a second language reduce automaticity in the first language? *Language Learning, 41*(1), 59–83. <https://doi.org/10.1111/j.1467-1770.1991.tb00676.x>
- Sharwood Smith, M. (2019). Language attrition as a special case of processing change: A wider cognitive perspective. In M. S. Schmid & B. Köpke (Eds.), *The Oxford handbook of language attrition* (pp. 72–87). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198793595.013.7>
- Soltani, M., Moradi, N., Rezaei, H., Hosseini, M., & Jasemi, E. (2021). Comparison of verbal fluency in monolingual and bilingual elderly in Iran. *Applied Neuropsychology: Adult, 28*(1), 80–87. <https://doi.org/10.1080/23279095.2019.1594234>
- Strauss, E., Sherman, E. M. S., Spreen, O., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). Oxford University Press.
- Thiele, K., Quinting, J. M., & Stenneken, P. (2016). New ways to analyze word generation performance in brain injury: A systematic review and meta-analysis of additional performance measures. *Journal of Clinical and Experimental Neuropsychology, 38*(7), 764–781. <https://doi.org/10.1080/13803395.2016.1163327>
- Tombaugh, T. N., Kozak, J., & Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists, 14*(2), 167–177.
- Treffers-Daller, J. (2019). What defines language dominance in bilinguals? *Annual Review of Linguistics, 5*(1), 375–393. <https://doi.org/10.1146/annurev-linguistics-011817-04555>
- Troyer, A. K. (2000). Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Neuropsychology, 22*(3), 370–378. [https://doi.org/10.1076/1380-3395\(200006\)22:3;1-V;FT370](https://doi.org/10.1076/1380-3395(200006)22:3;1-V;FT370)
- Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology, 11*(1), 138–146. <https://doi.org/10.1037/0894-4105.11.1.138>
- Venegas, M. J., & Mansur, L. L. (2011). Verbal fluency: Effect of time on item generation. *Dementia & Neuropsychologia, 5*(2), 104–107. <https://doi.org/10.1590/S1980-57642011DN05020008>
- Virtaranta, P. (1992). *Amerikansuomen sanakirja. A dictionary of American Finnish*. Siirtolaisuusinstituutti, Institute of Migration. <https://urn.fi/URN:ISBN:951-9266-43-7>
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. <https://ggplot2.tidyverse.org>
- Wickham, H., François, R., Henry, L., & Müller, K. (2019). *dplyr: A grammar of data manipulation*. <https://CRAN.R-project.org/package=dplyr>
- Yılmaz, G. (2019). L1 attrition, L2 development, and integration. In M. S. Schmid & B. Köpke (Eds.), *The Oxford handbook of language attrition* (pp. 303–313). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198793595.013.25>
- Yılmaz, G., & Schmid, M. S. (2018). First language attrition and bilingualism: Adult speakers. In D. Miller, F. Bayram, J. Rothman, & L. Serratrice (Eds.), *Studies in bilingualism* (Vol. 54, pp. 225–250). John Benjamins Publishing Company. <https://doi.org/10.1075/sibil.54.11yil>
- Yılmaz, G., & Schmid, M. S. (2012). L1 accessibility among Turkish-Dutch bilinguals. *The Mental Lexicon, 7*(3), 249–274. <https://doi.org/10.1075/ml.7.3.01yil>

Author biographies

Nana Lehtinen is a Speech-Language Pathologist and a doctoral student at the University of Turku. Her research focuses on native language change and attrition in immigrant communities, particularly in the Finnish language in California, with the aim of providing insight into how immigrant settings influence the evolution

of native languages. Additionally, she teaches Finnish at UC Berkeley, CA, USA enriching students' language and cultural understanding.

Anna Kautto is a researcher at Åbo Akademi University and the University of Turku, examining how linguistic skills intersect with non-linguistic information processing.

Dr. Kati Renvall, a Speech-Language Pathologist with a Ph.D. in psychology and holds the position of Senior Lecturer at the University of Turku. She also serves as a Docent at the University of Helsinki, Finland, and holds the title of Honorary Senior Research Fellow at Macquarie University, Australia. Her research primarily focuses on cognitive neuropsychology of language, aphasia, word-finding difficulties (anomia), and the assessment and treatment of adult neurogenic language disorders.