



Sequential bilingual heritage children's L1 attrition in lexical retrieval: Age of acquisition versus language experience

Kitaek Kim¹  and Hyunwoo Kim² ¹Seoul National University, Seoul, Republic of Korea and ²Yonsei University, Seoul, Republic of Korea

Research Article

Cite this article: Kim K, Kim H (2022). Sequential bilingual heritage children's L1 attrition in lexical retrieval: Age of acquisition versus language experience. *Bilingualism: Language and Cognition* 25, 537–547. <https://doi.org/10.1017/S1366728921001139>

Received: 23 March 2021
Revised: 10 November 2021
Accepted: 3 December 2021
First published online: 10 January 2022

Keywords:

language attrition; lexical retrieval; heritage speakers; age of acquisition; language experience

Address for correspondence:

Hyunwoo Kim, Yonsei University, 50, Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea
Email: hyunwoo2@yonsei.ac.kr

Abstract

This study investigated the unresolved issue of potential sources of heritage language attrition. To test contributing effects of three learner variables – age of second language acquisition, length of residence, and language input – on heritage children's lexical retrieval accuracy and speed, we conducted a real-time word naming task with 68 children (age 11–14 years) living in South Korea who spoke either Chinese or Russian as a heritage language. Results of regression analyses showed that the participants were less accurate and slower in naming target words in their heritage language as their length of residence in Korea and the amount of Korean input increased. The age of Korean acquisition did not significantly influence their performance. These findings support the claim that heritage speakers' language experience is a more reliable predictor of first language attrition than age of acquisition. We discuss these findings in light of different approaches to explaining language attrition.

1. Introduction

Heritage speakers are characterized as “child and adult members of a linguistic minority who grew up exposed to their home language and the majority language” (Montrul, 2010, p 4; also see Kagan & Dillon, 2013, for discussion of the definition of heritage speakers.). They grow up hearing and speaking the minority language (i.e., the parental/family language) at home while being exposed to the majority language of the society in most other social contexts. Most often, their language dominance shifts with schooling (Montrul, 2008), resulting from the decreased heritage language input and increased majority language input. Consequently, heritage speakers generally become dominant in the majority language while losing linguistic ability in the heritage language (Montrul, 2010). This language loss, commonly referred to as attrition, in heritage speakers has been widely attested in previous studies, which have reported deficiencies in the children's heritage language knowledge in various linguistic domains (Montrul, 2002; O'Grady, Lee & Lee, 2011; Polinsky, 2011).

A central issue in the field of first language (L1) or heritage language attrition concerns what drives the loss of an L1. A matter of continued debate is whether L1 attrition is solely driven by maturational constraints such as age of second language (L2) acquisition, or it is also affected by environmental factors. Previous research has reported mixed findings regarding this issue. Some studies show that the extent of attrition in heritage speakers correlates with onset age of bilingualism (Kim, Montrul, & Yoon, 2009; Montrul, 2002), supporting the critical period effect in L1 attrition. In contrast, other studies have not observed a clear effect of early L2 acquisition (e.g., Montrul & Sánchez-Walker, 2013). Instead, in many cases, language attrition is found to be influenced by the frequency with which children are exposed to an L2 relative to an L1 (e.g., Jia & Aaronson, 2003).

In light of the unresolved issue of underlying sources of heritage language attrition, the current study tested which factor, between an age effect and language experience, better accounts for the extent of heritage language attrition of sequential bilingual children living in South Korea (Age of L2 acquisition = 6–12). We focused on age of L2 acquisition and two L2 experience factors – length of residence and L2-L1 input ratio – and investigated their impacts on heritage speakers' L1 attrition as measured by their lexical retrieval accuracy and speed. These factors are considered highly relevant to language attrition, yet few studies have systematically compared their individual contributions. By disentangling these confounding factors, this study aims to identify driving factors of L1 attrition among heritage speakers, and, ultimately, to further our understanding of the mechanisms of heritage language acquisition.

2. Susceptibility to language attrition: effects of L2 AoA and language experience

An important issue in the study of heritage language attrition is what drives attrition. Two major factors are recognized – the critical period of attrition susceptibility (Bylund, 2009;

© The Author(s), 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Köpke & Schmid, 2004) and external factors related to language experience (Jia & Aaronson, 2003). A question arising from the issue of the critical period of attrition susceptibility concerns whether a younger age at the onset of bilingualism – namely, an earlier exposure to an L2 – leads to greater divergence from monolingual peers (Montrul, 2002). Motivated by the critical period hypothesis of L1 acquisition (Lenneberg, 1967; Penfield & Roberts, 1959), the view of the critical period in L1 attrition predicts that an early age of L2 acquisition (L2 AoA) prevents heritage speakers from fully acquiring their L1, causing an incomplete acquisition and language loss. One of the early evidence supporting the L2 AoA effect comes from Montrul (2002), who reported a reduced L1 ability by heritage speakers with earlier L2 AoA. Montrul examined adult Spanish–English bilinguals living in the United States, testing their morphological and semantic knowledge of imperfect and preterite tenses in Spanish. She divided the participants into three groups according to their onset age of L2 acquisition, involving 16 US-born bilinguals exposed to Spanish and English from ages 0–3 (simultaneous bilingual group), 15 US-born early sequential bilinguals exposed to English beginning at ages 4–7 (early child L2 learner group), and 8 late sequential bilinguals who had their first exposure to English at ages 8–12 (late child L2 learner group). She found that, in both production and comprehension tasks, the simultaneous bilinguals and early child L2 learners – but not the late child L2 learners – differed significantly from a monolingual Spanish control group. These findings are taken to indicate a significant effect of age of L2 acquisition, inspiring subsequent studies claiming a critical age for both L2 acquisition and L1 attrition (Bylund, 2009; Montrul, 2008).

In contrast, other studies have not found a robust effect of L2 AoA on heritage speakers' attrition. For example, Montrul and Sánchez-Walker (2013) tested knowledge of direct object marking (DOM) in Spanish (i.e., morphological marking with preposition *a* on animate direct objects) with Spanish heritage speakers in the United States. In a story retelling task and a picture description task, simultaneous and sequential heritage speakers of L1 Spanish, both children and adults, performed poorly at similar levels, although the adults were more accurate on DOM than the children. Montrul and Sánchez-Walker interpreted these results as suggesting that the participants' poor performance was not solely explained by L2 AoA because both simultaneous and sequential bilinguals, who have different L2 AoA, showed similar production skills in the tasks. Rather, the researchers attributed the results to an interaction of multiple factors, including reduced L1 input, potential attrition in the first generation of immigrants, incomplete L1 acquisition in the second generation, and language transfer from English.

The null effect of L2 AoA in some studies can be problematic for the critical period account of L1 attrition. However, such results may also be taken as an epiphenomenon arising from 'false negative' errors (or Type II errors). In particular, one cannot rule out the influence of extraneous variables that can obscure the effect of L2 AoA, such as a task effect. For example, Kang (2011) conducted a longitudinal study with three Korean–English bilingual children in the one to two years after they had returned to South Korea after a two-year stay in the United States (Rita, from age 4;11, Hera, from age 6;09, and Sammy, from age 11;10). When examining the children's English attrition through various tasks, Kang found significant individual differences in the degree of attrition among these children depending on the task type. In grammatical judgment and

elicitation tasks, Rita showed production errors in English passives (e.g., preposition substitutions such as the use of *with* for *by*) while her older sister Hera did not. Instead, Hera's speech rate in English narratives significantly slowed down in the second year after arrival, indicating her decreased speaking fluency in L2 English. These findings give rise to doubts about the appropriateness of the judgment and elicitation tests as adequate indicators of attrition. Kang maintained that the "earliest signs of attrition do not involve the attriters' linguistic competence (such as knowledge of grammar) or comprehension but rather their general processing skills in production such as speech rate" (p. 174) (for a discussion of speech rate as a promising measure of adult heritage language attrition, see Polinsky, 2011). Given this possibility, he proposed other production tasks as more appropriate measures for capturing language attrition, such as a word-naming task developed as part of the Hawai'i Assessment of Language Access (HALA) project (O'Grady, Schafer, Perla, Lee & Wieting, 2009; henceforth the HALA task). Following Kang's suggestion, the current study used the HALA task to investigate heritage speakers' language attrition by measuring their lexical retrieval accuracy and speed in their two languages (Kang, 2011; O'Grady et al., 2009).

In addition to L2 AoA, other factors can also account for L1 attrition. For instance, Jia and Aaronson (2003) emphasized the role of language experience while considering L2 AoA to be an interacting factor. Their longitudinal study included 10 native Chinese-speaking children and adolescents who had immigrated to the United States between the ages of 5 and 16. Over three years, the researchers traced changes in the children's language preferences, language environments, and proficiency in L2 English and L1 Chinese, using various assessment tools including grammaticality judgment tasks, L1 to L2 translation tasks, parental ratings of L1 proficiency change, child interviews, parent interviews, questionnaires, and observations. The results showed that the children who arrived in the United States at age 9 or younger switched their language preference from L1 to L2 within the first year, whereas the older participants with L2 AoA of 10 or older maintained their preference for the L1 for the three years of the study. Moreover, the younger participants became more proficient in the L2 than in the L1, whereas the older participants maintained greater proficiency in the L1. Crucially, the contrasting features between the two bilingual groups extended beyond their L2 AoA to include their language experience. The younger participants were exposed to a significantly richer L2 than L1 environment, whereas the situation was the other way around for the older participants, who received richer L1 than L2 input. Although both participant groups received substantial L2 input through school instruction, the younger group read more L2 than L1 books while the older participants read more L1 than L2 books. In addition, the younger participants had more English-speaking than Chinese-speaking friends while the opposite pattern obtained for the older participants. The researchers attributed these trends to "the influence of language environment on L2 acquisition" (p. 156) rather than neurobiological factors such as age. These findings indicate that AoA may not be a determining factor, but may critically interact with language experience factors to influence heritage speakers' L1 attrition.

One theoretical perspective highlighting the role of language experience in heritage language attrition, particularly in the domain of lexical retrieval and production skills, is the Weaker Links hypothesis (Gollan, Montoya, Cera, & Sandoval, 2008). The basic assumption of this hypothesis is that increased

language experience in a particular language facilitates accessibility to the language system, affording a speaker faster and more accurate activation of linguistic items. Conversely, the hypothesis predicts that the infrequent access to a language system weakens the level of activation or association strength between forms and meanings, ultimately leading to language loss. Since maintaining two linguistic systems places considerable burdens on bilinguals (Jessner, 2003), it is inevitable that bilinguals have a less stable linguistic system in one language over the other. Resting upon this idea, O'Grady et al. (2009) developed the HALA task to assess bilingual and heritage speakers' language strength by measuring their lexical retrieval accuracy and speed in one language over the other through a picture-naming task. The underlying assumption of the task is that more recent and frequent access to words in a certain language would lead to faster and more accurate word production. While the task includes target words with different frequency (for details of the items, see materials in the methods section below), the primary objective of the task is to measure individual bilinguals' differences in their word access speed and accuracy in one language over the other, which can be affected by individuals' language experiences (e.g., Jo, Kim & Kim, 2021; Kang, 2011; O'Grady et al., 2009). In O'Grady et al.'s experiment, English-Korean heritage speakers were prompted to name a picture presented on a computer screen, with items of different frequency. Although the participants showed an overall high word naming accuracy in both of the languages tested, their naming times were faster in English than in Korean, and faster in more frequent than less frequent words. These results led O'Grady et al. to the conclusion that the HALA task successfully captured the relative strength of the bilinguals' languages, at least in the domain of lexical retrieval speed, consistent with the prediction of the Weaker Links hypothesis. The main tenet of the Weaker Links hypothesis and the empirical support provided by O'Grady et al. suggest that language frequency constitutes a highly reliable factor affecting a bilingual or heritage speaker's language strength, allowing us to test whether decreased language experience gives rise to the attrited lexical access and retrieval in that language.

3. Measuring language experience

As outlined in the previous section, the research is inconclusive with respect to the question of whether attrition is best predicted by the critical period of L2 acquisition or by language experience factors. While several studies have investigated the effect of AoA on language attrition, to the best of our knowledge, few have directly compared the relative contributions of L2 AoA and language experience factors. Traditionally, the critical period issue has been addressed by examining the AoA effect on language acquisition/attrition, with L2 AoA determined by identifying participants' age at onset of bilingualism or of arrival in the immigration destination country (assuming no substantial L2 input before arrival). In contrast, various measures have been employed to estimate participants' language experience. To address the issue of how language experience factors compare to L2 AoA in their impacts on L1 attrition, it is important to discuss how previous studies operationalized and measured heritage speakers' language experience.

A speaker's language experience is closely associated with the quantity and quality of input and output (Unsworth, Hulk, & Marinis, 2011). One of the most widely employed language experience variables is the length of exposure or length of

residence (LoR) in the destination country, which is considered an indicator of input quantity. LoR is measured as duration from the onset age of L2 acquisition (often coinciding with the age of arrival in the destination country for heritage speakers) to the learner's current age at the time of testing. Previous studies have found a strong relation between LoR and L2 acquisition. For example, Hopp (2011) examined the effects of internal factors (e.g., age, AoA) and external factors (e.g., length of exposure) on the development of the German determiner phrase (DP) by 60 child L2 learners of German (age range = 3;5–7;0). Results showed that length of exposure had a stronger predictive effect on accuracy than age or AoA. In addition, when length of exposure was controlled for, the observed correlations between accuracy and age or AoA disappeared, leading Hopp to conclude that "the amount of input is critical for mastering the intricate system of DP inflection" (p. 257).

As Unsworth et al. (2014) noted, however, LoR may not in all cases serve as a reliable indicator of amount of input, due to significant variability in the amount of language input bilingual children receive from various sources. For example, if two bilingual children have the same LoR, but one child goes to daycare while the other mostly spends time with L1-speaking parents at home, they would receive different amounts of L2 exposure. To address this problem, Unsworth et al. (2014) proposed CUMULATIVE length of exposure, assessed through examination of the L1-L2 input ratio. She developed a questionnaire to determine the L1-L2 input ratio at home and at school (or daycare), with items asking, for instance, how often each bilingual living in the home speaks each of the languages, with Likert-scale responses ranging from "almost always [the L1]" to "almost always [the L2]." The current study adopted a similar questionnaire to measure language ratio, in addition to LoR, as an indicator of the heritage speakers' language experience.

4. The present study

This study attempts to determine a primary source of heritage language attrition by examining the relative contributions of L2 AoA, LoR, and L2-L1 input ratio to heritage speakers' ability to retrieve L1 lexical items. To this end, we first measure participants' lexical retrieval accuracy and speed using the HALA task (O'Grady et al., 2009) in their L1 and L2, and then examine the influence of each of the three learner-related variables on their task performance. Based on the attested roles of L2 AoA and language experience in heritage learners' access to L2 vocabulary, we expect all of the factors to affect L1 attrition, albeit to different degrees. If heritage learners experience L1 attrition, they will exhibit reduced ability in lexical retrieval in the L1 task while showing better performance in the L2 task. Therefore, evidence of L1 attrition will be indicated by a significant interaction between language (L1, L2) and the child-internal and -external variables (i.e., AoA, LoR, and L2-L1 input ratio). In the case of an interaction, we further scrutinize the effects of these variables in separate by-language models. We also explore the individual contributions of these variables to L1 attrition by examining their correlations with L1 task outcomes. The specific research questions are as follows:

1. Do heritage speakers' age of L2 acquisition, length of residence, and L2-L1 input ratio account for L1 attrition?
2. Which of these three factors better predicts L1 attrition?

5. Methods

Participants

The study involved 68 children (27 male and 41 female, mean age 12 years, range 11–14) who spoke Korean as an L2. They were native speakers of Mandarin Chinese ($n = 34$) or Russian ($n = 34$), born and raised in mainland China or Russian-speaking countries until they immigrated to Korea. The participants' parents were all foreign-born immigrants who had settled in Korea, primarily for job-related reasons. At the time of testing, the children attended a local elementary school (6th grade) in South Korea. Besides knowledge of their L1 and Korean, all the children had some basic English knowledge as they studied English two to three hours a week as part of the regular school curriculum. Some children reported prior exposure to regional dialects other than Chinese or Russian before arrival in Korea, but they said they had no knowledge about those languages at the time of testing.

We used a survey (Dörnyei & Csizér, 2012) and school records to collate information on participants' language background, including age of L2 acquisition (L2 AoA), length of residence in Korea (LoR), and amount of L2 relative to L1 input (input ratio). All participants reported that their L1 (i.e., Chinese or Russian) was their dominant language, and that their first exposure to Korean occurred upon arrival in Korea. The mean L2 AoA was 9.3 years ($SD = 1.4$; range 6–12 years). There was no significant difference in L2 AoA between the L1-Chinese and L1-Russian groups ($t(66) = 1.051$, $p = .297$, Cohen's $d = 0.255$). Because of the small variability in L2 AoA, following Köpke and Schmid (2004), we divided the participants into earlier L2 AoA (9 years or younger, $n = 31$) and later L2 AoA (10 years or older, $n = 37$) groups and included this division as a categorical variable in the analysis model. Despite the small variability, we also conducted an additional analysis including L2 AoA as a continuous variable.

According to the school records, the children came to Korea at different ages (range 6–12), thus having variation in their LoR (mean 2.6 years, range 1–5). On average, the L1-Chinese group had spent a significantly longer time in Korea than the L1-Russian group ($t(66) = 2.799$, $p = .007$, Cohen's $d = 0.679$). Each participant's LoR was added as a continuous variable to the analysis model (details are given in the data coding and analysis section).

To estimate the amount of L1 and L2 input the children received, we used the language background survey responses. The survey asked participants to report estimated amounts of time they spend using each language on a weekly basis in various contexts (e.g., with parents at home, with siblings at home, with peers at school, and with teachers at school) by rating their usage on a four-point scale, where 0 indicates 'never' and 3 indicates 'always' (cf. Duncan & Paradis, 2020). Almost all participants (65 of 68) reported using their L1 dominantly at home with their parents and siblings, and 52 said they used both the L1 and Korean at home. All of the participants said they used Korean dominantly in every other social context, including at school. By virtue of these language use profiles, the children were characterized as heritage learners, following the definition of heritage speakers as bilinguals from immigrant families who are exposed to an ethnic minority language at home and to a majority language in other social contexts (Montrul & Ionin, 2012; O'Grady et al., 2011). The children's teacher reported that they took all school subjects in the majority language (Korean) and had little difficulty in listening, speaking, reading, and writing

Table 1. Participant information

	L1 group	Mean	Std. Deviation	Range
Age	Chinese	11.8	0.6	11–13
	Russian	12.2	0.8	11–14
Age of L2 acquisition	Chinese	9.1	1.5	6–11
	Russian	9.5	1.3	7–12
Length of residence in Korean (month)	Chinese	37.1	14.5	12–60
	Russian	27.5	13.7	12–60
Mean L1 input	Chinese	1.8	0.9	0–3
	Russian	2.6	0.5	0–3
Mean L2 input	Chinese	1.7	0.7	0.2–3
	Russian	1.5	0.8	0.7–3
L2-L1 input ratio	Chinese	1.9	2.9	0–15
	Russian	0.7	0.6	0–3

in Korean. To quantify the relative amounts of input in the two languages, we divided the mean rating of L2 input by the mean rating of L1 input to obtain the L2-L1 input ratio for each participant. The ratio was significantly higher for the L1-Chinese group than for the L1-Russian group ($t(66) = 2.469$, $p = .016$, Cohen's $d = 0.599$), indicating that the time spent with the L2 relative to the L1 was greater for the Chinese-speaking children than for the Russian-speaking children. The input ratio was included as a continuous variable in the analysis model. Table 1 summarizes the participant information.

Materials

To address the question of whether age of L2 (Korean) acquisition or increased experience with the L2 would affect the heritage speakers' use of their L1 (heritage language, either Chinese or Russian), we tested participants' lexical retrieval ability through the HALA task (O'Grady et al., 2009). This task was designed to determine heritage speakers' language dominance by assessing their word retrieval speed and accuracy in their two languages. The basic assumption underlying the task is that the strength of activation of a particular language is closely associated with the ability to access and retrieve vocabulary items in that language (De Bot, 2004). By assessing the relative facility of lexical retrieval in one language over the other, the task helps determine how strongly a bilingual's lexical activation operates in each language, rendering it a useful tool for measuring degree of attrition (e.g., Kang, 2011). The original version of the task included 43 test items comprising body-part images. To reduce the cognitive load on our young participants, whose attention spans we assumed to be shorter than adults', we adopted the modified version employed by Kang (2011), who included 31 items. The items were divided into two categories based on their relative frequency as measured by the English Lexicon Project (Balota et al., 2007). The task items are listed in English in Table 2.

Procedure

The participants completed the L1 and L2 HALA tasks after the language background questionnaire. For a direct comparison of

Table 2. Items in the HALA task

Frequency	Item (translated in English)
Higher frequency ($k = 17$)	back, leg, ear, lips, eye, mouth, face, nose, fingers, shoulder, foot, stomach, hand, teeth, head, tongue, knee
Lower frequency ($k = 14$)	ankle, forehead, arm, heel, cheek, neck, chin, palm, elbow, thumb, eyebrow, toe, fingernail, wrist

the participants' lexical retrieval ability across the two languages, each participant named the same items in each task. The order of the L1 and L2 tasks was counterbalanced across participants so that half of the participants completed the L1 task first, and the other half completed the L2 task first. There was at least a three-week interval between the two tasks to minimize inter-task interference.

The tasks were administered individually in a classroom after regular classes, following the procedure described by O'Grady et al. (2009). The higher-frequency items were presented first, followed by the lower-frequency items, to alleviate the cognitive burden associated with word naming. The order of items within each frequency category was pseudo-randomized so that no item appeared in the same sequential position across the L1 and L2 tasks. The task items were presented on a computer screen, using Shockwave Flash animation. During each trial, an image appeared on the screen along with a beep sounded, prompting participants to name the pictured body part, which was highlighted on the image by a red circle (see Figure 1). The image remained on the screen for 4000 milliseconds (ms) for higher-frequency words and 4500 ms for lower-frequency words, following Kang (2011). Participants were instructed to name the target image as quickly as possible. When participants failed to provide an answer within 4,000 or 4,500 ms after the picture onset, the experiment automatically advanced to the next trial. The picture for each item was shown to the participant for 4,000 or 4,500 ms until the next picture onset. Each response was audio-recorded during the task and transcribed later. Prior to the task, participants received instruction and worked through six practice items. Including the language background questionnaire, the procedure took approximately 30 minutes.

Data coding and analysis

Participants' responses in the HALA tasks were coded in terms of accuracy and response latency. Accuracy in the L1 task was judged by two trained native speakers of the respective languages (Chinese and Russian), and in the L2 task by a native speaker of Korean. Every correct name for a target image spoken within the picture's duration on screen was given a point. Correction of a wrong response was also accepted and given a point only when made within the time limit. Hypernyms for target words (e.g., *face* for the target word *eye*) were coded as incorrect. For word naming speed, we calculated the duration from the onset of the beep to the onset of the utterance of the correctly named word as the response time (RT). RTs for incorrect responses were excluded from further analyses (14.7% of the L1 task data and 33.7% of the L2 task data).

Our analysis focused on whether the participants' experience with L2 Korean modulates their L1 and L2 word retrieval accuracy and speed in the HALA tasks. For the dependent measures of

word naming accuracies and RTs, we created different types of mixed-effects regression models. Because accuracy was coded as categorical (correct, incorrect), we used logistic mixed-effects regression (Baayen, 2008; Jaeger, 2008) fitted to the proportion of correct responses. To capture the continuous nature of the RT data, we used linear mixed-effects regression (Baayen, Davidson, & Bates, 2008). For each type of dependent measure, we tested the interactive effects of LoR, L2-L1 input ratio, and AoA. Because these measures significantly correlated with one another (all $r_s > .2$), we created separate models including each of the factors as either a continuous (LoR, L2 AoA, and L2-L1 input ratio: centered) or a categorical (L2 AoA: centered and contrast-coded with earlier AoA coded as $-.5$ and later AoA coded as $.5$) fixed factor, instead of including them in the same model. Because the values included as continuous factors had some individual variability, we transformed them to z-scores to adjust for scale biases between participants (Hox, Moerbeek, & Van de Schoot, 2017). We also included language (L1, L2) in the models as a categorical fixed effect (contrast-coded with L1 coded as $-.5$ and L2 coded as $.5$) to explore whether the effects of the three factors emerged differently in the L1 and L2 tasks. In addition, participants' L1 background (Chinese vs. Russian, contrast-coded) was included as a categorical fixed effect to see if the degree of L1 attrition would differ depending on the L1. The models also included the random effects of participant and item, which contained the maximal random-effects structure permitted by the design (Barr, Levy, Scheepers, & Tily, 2013). In the case of a significant interaction between language (L1, L2) and one of the factors, we further conducted separate by-language analyses where we looked at the effect of each variable for each language. For the by-language analyses, the alpha level was corrected to $.025$ ($.05$ divided by 2). The logistic regression modeling was conducted using *glmer*, and the linear regression modeling was conducted using *lmer*, in the *lme4* package in R (R Core Team, 2018).

6. Results

Word naming accuracy

We first report results from the model including accuracy scores as a dependent measure. Participants showed higher accuracy scores in the L1 task ($M = 85.3\%$, $SD = 35.4$) than in the L2 task ($M = 66.3\%$, $SD = 47.3$), indicating their dominance in the L1. Table 3 shows results from each language group.

To scrutinize how different factors modulate the accuracy scores in the L1 and L2 tasks, we created four logistic regression models, each with language (L1 task, L2 task) and respective fixed effects (LoR, L2-L1 input ratio, L2 AoA, L1 background). Table 4 summarizes outputs from these models. (Separate results of the two L1 groups are presented in Appendix A.)

The model including language and LoR as fixed factors showed a main effect of language, with higher accuracy scores in the L1 than in the L2 task. The effect of LoR was only marginal. Crucially, there was a significant interaction between language and LoR, indicating that the effect of LoR emerged differently in the two language tasks. By-language models including LoR as a single fixed factor revealed a robust effect of LoR both in the L1 task ($\beta = -0.70$, $SE = 0.06$, $p < .001$) and in the L2 task ($\beta = 1.22$, $SE = 0.17$, $p < .001$), but in different directions. In other words, as the children spent longer time in Korea, their accuracy scores decreased in the L1 task but increased in the L2 task.



Fig. 1. Procedure of item presentation in the HALA task

Table 3. Mean word naming accuracies in percentage (standard deviations)

	Chinese group		Russian group	
	L1 task	L2 task	L1 task	L2 task
Earlier L2 AoA group	76.5 (42.5)	80.5 (39.7)	88.5 (32.0)	53.0 (50.0)
Later L2 AoA group	87.5 (33.1)	74.6 (43.6)	88.7 (31.7)	56.5 (49.6)
High-frequency items	91.5 (27.9)	90.8 (28.9)	95.5 (20.7)	73.9 (44.0)
Low-frequency items	70.4 (45.7)	61.3 (48.7)	80.3 (39.9)	32.1 (46.8)

Table 4. Results of models for accuracy scores

		β	<i>SE</i>	<i>p</i>
Model including LoR	Intercept	1.97	0.28	<.001***
	Language	-1.48	0.20	<.001***
	LoR	0.02	0.01	.054
	Language \times LoR	1.82	0.20	<.001***
Model including L2-L1 input ratio	Intercept	2.05	0.29	<.001***
	Language	-1.12	0.27	<.001***
	Input	0.21	0.10	.039*
	Language \times Input	3.25	0.65	<.001***
Model including L2 AoA (categorical variable)	Intercept	1.96	0.28	<.001***
	Language	-1.47	0.29	<.001***
	AoA	0.21	0.23	.348
	Language \times AoA	-0.85	0.58	.143
Model including L2 AoA (continuous variable)	Intercept	1.96	0.28	<.001***
	Language	-1.47	0.29	<.001***
	AoA	0.11	0.11	.325
	Language \times AoA	-0.40	0.29	.164
Model including L1 background	Intercept	1.97	0.28	<.001***
	Language	-1.51	0.26	<.001***
	L1	-0.49	0.25	.048*
	Language \times L1	-2.37	0.52	<.001***

*** $p < .001$; ** $p < .01$; * $p < .05$

Similar results were found in the model including language and input ratio as fixed effects. We found a main effect of language and a main effect of input ratio, as well as their interaction. Separate models breaking down this interaction showed that the effect of input ratio emerged both in the L1 task ($\beta = -0.84$, $SE = 0.13$, $p < .001$) and in the L2 task ($\beta = 2.01$, $SE = 0.64$, $p = .002$), yet in

different directions. These results indicate that as the learners received more L2 than L1 input, their accuracy decreased in the L1 task but increased in the L2 task.

When the model included language as a categorical effect and L2 AoA either as a categorical or a continuous variable, there was a main effect of language, with higher accuracy scores in the L1

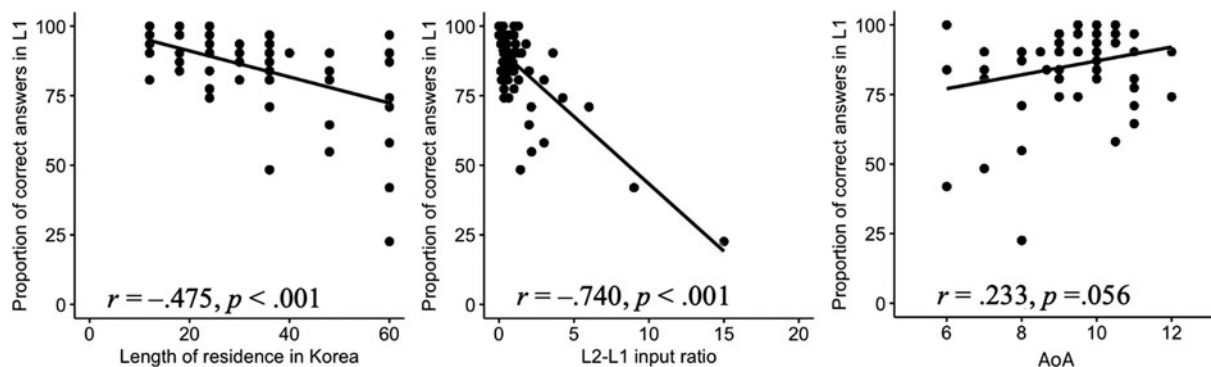


Fig. 2. Correlations between word naming accuracy in the L1 task and learner-related measures: LoR in Korea (left), L2-L1 input ratio (middle), and AoA (right)

than the L2 task. However, the effect of language did not interact with L2 AoA, no matter whether L2 AoA was included as a categorical or continuous variable, indicating that participants had higher accuracy scores in the L1 than the L2 task regardless of L2 AoA. These results run counter to previous findings showing evidence of L2 AoA effects on L1 attrition (e.g., Bylund, 2009; Montrul, 2002; Pallier, 2007). This point will be discussed in detail in the discussion section, where we provide potential accounts for the lack of an interactive effect of L2 AoA in this study.

Finally, the model including language and L1 background showed a significant effect of language, L1 background, and their interaction. Post-hoc analyses by each language revealed a significant effect of L1 background in the L2 task ($\beta = -1.86$, $SE = 0.37$, $p < .001$), with a higher accuracy in the Chinese than the Russian group, but no such L1 background effect emerged in the L1 task ($\beta = 0.57$, $SE = 0.34$, $p = .097$). The increased accuracy in the Chinese versus Russian group in the L2 task is taken to reflect the Chinese group's higher lexical proficiency in the L2 Korean, presumably due to their longer stay in Korea and more amount of Korean input. However, as shown in the result of the L1 task, both groups showed no sign of difference in L1 attrition as a function of their L1 background.

Taken together, the analyses of accuracy scores revealed that participants had a reduced ability to correctly name target words in the L1 as they spent more time in Korea, and as they received more input in the L2 than the L1. However, participants' L2 AoA did not significantly affect their word naming accuracy in the L1 task. These tendencies are reflected in correlation analyses: L2 AoA correlated only marginally while both LoR and input ratio correlated significantly with word naming accuracy scores in the L1 task (see Figure 2). In addition, input ratio was found to have a stronger correlation than LoR with accuracy scores, suggesting that the participants' L1 attrition was better explained by the relative amount of L2 input than by their length of stay in Korea or L2 AoA.¹

¹As visible in Figure 2, two participants in the Chinese-speaking group had extreme L2-L1 input ratio values (7.5 and 15, respectively), which fell beyond two standard deviations from the mean. We thus eliminated them as outliers and reran logistic mixed-effects models with language (L1, L2) as a categorical fixed effect and the L2-L1 input ratio as a continuous fixed effect. As in the analyses of the entire data, the models including these subset data returned a significant interaction between the two factors for both word naming accuracy ($\beta = 1.59$, $SE = 0.30$, $p < .001$) and response times ($\beta = -113.4$, $SE = 15.4$, $p = .003$). Separate models for each language in the analysis of the accuracy data showed a significant effect of input for both L1 ($\beta = -0.38$, $SE = 0.06$, $p < .001$) and L2 ($\beta = 0.91$,

In addition to the models including these variables, we constructed exploratory models adding an additional factor of item frequency either as a categorical or a continuous variable.² Although there was a main effect of frequency in each model (all $ps < .01$), with decreased accuracies as the items were less frequent, we found no significant interaction between frequency and other factors (all $ps > .1$), indicating that item frequency did not modulate the patterns we found in the previous analyses.

Following Hopp (2011), we also conducted partial correlation analyses to investigate correlations between accuracy scores and one of the factors by controlling for another factor. When LoR was controlled for, L2 AoA significantly correlated with the accuracy scores ($r = .255$, $p = .037$). However, when the input ratio was controlled for, the correlation of L2 AoA and the accuracy scores failed to reach significance ($r = .115$, $p = .355$). When L2 AoA was controlled for, the accuracy scores significantly correlated both with LoR ($r = -.484$, $p < .001$) and with the input ratio ($r = -.726$, $p < .001$). These results confirm that LoR and input ratio were influential factors accounting for word naming accuracy in the L1 task, whereas L2 AoA fell short as a reliable determinant of the variability of the accuracy scores.

Word naming speed

Prior to analyzing the RT data, RTs exceeding two standard deviations from the mean were identified as outliers and excluded from further analyses (4.7% of the entire data). Analyses of the remaining data show that the participants named target words more slowly in the L2 task ($M = 1270.7$ ms, $SD = 430.4$) than in the L1 task ($M = 1190.1$ ms, $SD = 404.0$), indicating their L1 dominance, consistent with their word naming accuracy. Table 5 presents results from each language group.

To explore whether participants' L2 experience modulates their RTs, linear mixed-effects models were fitted to the RTs, including language and each of the three measures (LoR, L2-L1 input ratio, L2 AoA, L1 background) as fixed effects. Table 6

$SE = 0.29$, $p = .002$) but in different directions, indicating the decreased word naming accuracies in the L1 task as participants received more L2 than L1 input. By-language models in the analysis of the response time data showed a main effect of input in the L1 task ($\beta = 57.0$, $SE = 9.5$, $p < .001$), with increased response times as participants received more L2 than L1 input, yet there was no significant effect of input in the L2 task ($\beta = -33.1$, $SE = 14.9$, $p = .166$). Overall, the analyses of the subset data yielded the same outcomes as in the total data analysis, suggesting that the data from the two outliers did not affect the influence of L2-L1 input ratio on the L1 word retrieval performance.

²We thank two anonymous reviewers for pointing out the need for these additional analyses.

Table 5. Mean word naming speed in milliseconds (standard deviations)

	Chinese group		Russian group	
	L1 task	L2 task	L1 task	L2 task
Earlier L2 AoA group	1232.6 (432.5)	1210.9 (431.7)	1141.5 (352.6)	1384.4 (398.7)
Later L2 AoA group	1193.8 (413.3)	1192.9 (433.6)	1190.5 (405.9)	1359.7 (414.5)
High-frequency items	1170.9 (418.8)	1136.6 (392.1)	1097.0 (358.4)	1310.8 (389.7)
Low-frequency items	1278.5 (420.5)	1328.6 (477.0)	1282.2 (398.7)	1551.0 (411.9)

Table 6. Results of models for response times

		β	SE	<i>p</i>
Model including LoR	Intercept	1278.68	32.23	<.001***
	Language	104.72	17.31	<.001***
	LoR	-17.69	20.26	.387
	Language \times LoR	-149.46	23.28	<.001***
Model including L2-L1 input ratio	Intercept	1279.10	31.90	<.001***
	Language	94.99	20.44	<.001***
	Input	34.30	18.68	.278
	Language \times Input	-284.68	60.62	.003**
Model including L2 AoA (categorical variable)	Intercept	1278.44	32.09	<.001***
	Language	109.37	28.51	<.001***
	AoA	-8.79	39.87	.826
	Language \times AoA	33.85	58.72	.567
Model including L2 AoA (continuous variable)	Intercept	1273.21	31.43	<.001***
	Language	126.77	25.93	<.001***
	AoA	-11.62	18.37	.537
	Language \times AoA	20.59	34.71	.557
Model including L1 background	Intercept	1277.72	31.59	<.001***
	Language	106.82	24.08	<.001***
	L1	38.29	37.76	.314
	Language \times L1	246.84	48.00	<.001***

*** $p < .001$; ** $p < .01$; * $p < .05$

presents the model outcomes. (Separate results of the two L1 groups are presented in Appendix B.)

The model including language and LoR as fixed effects showed a main effect of language, induced by longer RTs in the L2 than the L1 task, and a main effect of LoR, with decreasing RTs as participants spent longer time in Korea. We also found a significant interaction between the two factors, which emerged due to the different directions of the LoR effect between the tasks. To unpack this interaction, we further conducted separate analyses for each task, including LoR as a single fixed effect. Results showed a significant effect of LoR emerging in both the L1 task ($\beta = 58.05$, $SE = 22.08$, $p = .011$) and the L2 task ($\beta = -107.79$, $SE = 23.38$, $p < .001$), yet in different directions. As participants had longer LoRs, their RTs increased in the L1 task but decreased in the L2 task.

Similarly, the model including language and L2-L1 input ratio revealed a main effect of language, qualified by its interaction

with input. Post-hoc models conducted separately for each task showed a main effect of input in the L1 task ($\beta = 125.60$, $SE = 20.92$, $p < .001$), induced by increased RTs with greater L2-L1 input ratio, but there was no main effect of input in the L2 task ($\beta = -72.98$, $SE = 32.89$, $p = .166$). These results suggest that the increased L2 input significantly delayed participants' response in the L1 task but did not affect their response speed in the L2 task.

In the model including language as a categorical factor and L2 AoA either as a categorical or a continuous factor, we only found a main effect of language, with longer RTs in the L2 than the L1 task. There was no main effect of L2 AoA and no interaction between language and L2 AoA, no matter whether L2 AoA was included as a categorical or a continuous variable. The lack of interaction indicates that participants had the same RT patterns in both L1 and L2 tasks, regardless of their L2 AoA, a finding consistent with the findings from the analysis of word naming

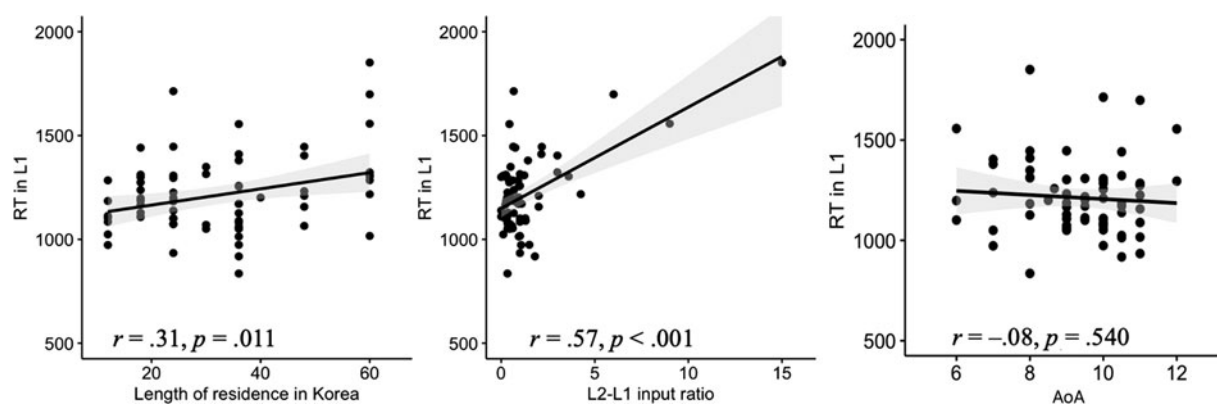


Fig. 3. Correlations between response times in the L1 task and learner-related measures: LoR in Korea (left), L2-L1 input ratio (middle), and AoA (right)

accuracy. We thus found no evidence of effects of L2 AoA in L1 attrition.

When the model included language and L1 background as fixed factors, there was a main effect of language with the longer RTs in the L2 than the L1 task. There was no main effect of L1 background, yet it interacted with language. By-language analyses examining the interaction showed a significant effect of L1 background only in the L2 task ($\beta = 187.31$, $SE = 49.83$, $p < .001$), with shorter RTs in the Chinese- than the Russian-speaking group, but not in the L1 task ($\beta = -70.90$, $SE = 44.33$, $p = .115$). As was the case for the accuracy, the faster naming speed for the Chinese group as compared to the Russian group in the L2 task indicates their superior lexical retrieval ability resulting from their prolonged experience with Korean. At the same time, the lack of difference between the two groups in the L1 task suggests little evidence of the L1 background effect on the participants' L1 attrition.

In summary, the analyses of RTs showed that participants spent longer times responding in the L1 task as they had longer LoRs and increased L2-L1 input ratios. These results suggest that participants had greater difficulty with lexical retrieval in the L1 as they spent a longer time in Korea and received more L2 input, a clear indication of L1 attrition. These patterns are reflected in Figure 3, which shows the correlations between RTs in the L1 task and each of the factors. Participants' RTs in the L1 task positively correlated with LoR and L2-L1 input ratio. Between the two factors, the input ratio showed a stronger correlation than LoR with the RTs. In contrast, L2 AoA did not significantly correlate with the RTs in the L1 task.

When adding item frequency (either as a categorical or a continuous variable) to the models, we only found a main effect of frequency in each model (all $ps < .01$), with slower response times as the items were less frequent. There was no significant interaction between frequency and other factors (all $ps > .1$), suggesting that item frequency did not modulate the effects of the factors investigated in the previous analyses.

As we did for accuracy, we conducted partial correlation analyses to inspect how each factor correlates with RTs when another variable is partialled out in the analysis. Results showed that the partial correlation between L2 AoA and RTs remained insignificant, whether we controlled for LoR ($r = -.074$, $p = .551$) or the input ratio ($r = .057$, $p = .648$). In contrast, when L2 AoA was controlled for, the RTs continued to correlate both with LoR ($r = .307$, $p = .011$) and with the input ratio ($r = .569$, $p < .001$). These results

suggest that the participants' RT profiles in the L1 task can be best explained by the input ratio, followed by LoR, but not by L2 AoA.

7. Discussion

This study aimed to address whether heritage children show reduced L1 production speed and accuracy as an early sign of L1 attrition and whether their L1 attrition is better explained by L2 AoA or by the other L2 experience factors of length of residence and L2-L1 input ratio. To address these questions, we administered the HALA tasks and language background survey to L1-Chinese and L1-Russian sequential bilingual heritage children living in Korea. The children showed a reduced ability to correctly name target words and spent longer times responding in the L1 task as they spent more time in Korea and received more input in the L2 than the L1. However, participants' L2 AoA did not significantly affect their word naming accuracy or response times in the L1 task.

These results are difficult to reconcile with the idea of a critical period of attrition – namely, that children are susceptible to language attrition until a certain age. For instance, Köpcke and Schmid (2004) proposed that the period extends until around age 9, which would predict a high degree of attrition for those with an L2 AoA below 9. This prediction was not borne out by the current study, as shown by the absence of an interactive role of L2 AoA in L1 word naming accuracy or response times for both the earlier L2 AoA group (9 and younger) and the later L2 AoA group (10 and older). Similarly, Bylund (2009) suggested a gradual decline in attrition susceptibility during the maturation period, ending at around age 12, which would predict the effect of L2 AoA as a continuous variable. To test this prediction, we entered L2 AoA as a continuous variable in the mixed-effects regression models, but found no interaction effect between L2 AoA and language, with either accuracy or RTs as a dependent measure (all $ps > .1$). All in all, our findings do not provide evidence for the critical period account's claim of an age effect in heritage speakers' attrition.

The study's findings instead align with the claim that language experience is a crucial predictor for L1 attrition (e.g., Jia & Aaronson, 2003). This is consistent with the theoretical position that language frequency determines the extent of accessibility of the language system in bilinguals, featured most prominently in the Weaker Links hypothesis (Gollan et al., 2008). The hypothesis predicts attenuated links between lexical forms and meanings in a

given language as a direct function of a bilingual's less practice or infrequent use of words in the language. We found such an effect in our heritage speakers, who showed less accurate and slower lexical retrieval in the L1 word naming task concomitant with their decreased L1 relative to L2 input and prolonged residence in Korea. It appears that the reduced experience with the L1 may have lowered the activation levels of L1 words in the children's mental lexicon, rendering it more difficult to retrieve target words efficiently during the L1 task. Given this evident sign of L1 attrition, it remains a fruitful avenue for future research to investigate whether the reduced lexical retrieval ability in these children ultimately leads to complete language loss.

The complete loss of linguistic knowledge resulting from a lack of language experience has been attested by several studies on international adoptees and immigrant students, who are severed from L1 input. For example, Ventureyra, Pallier, and Yoo (2004) reported a total loss of L1 phonemic knowledge in their study with L1 Koreans adopted by French-speaking families between 3 and 9 years of age ($n = 18$; age range in testing: 22–36). The researchers assessed the adoptees' perception of L1 phonemic contrasts with Korean voiceless consonants that are difficult to perceive by native French speakers. Their results showed that the Korean phonemic contrasts presented as much of a challenge for the adoptees as it did for native French speakers previously unexposed to Korean, with both groups failing to perceive the difference. Ventureyra et al. interpreted these findings as compelling evidence of the adoptees' total loss of L1 phonetic competence. In the literature with heritage language speakers, it has been reported that the continued experience of heritage language is responsible for the maintenance of the heritage language. For example, Hakuta and D'Andrea (1992) examined the maintenance and loss of Spanish with high school students living in the United States with L1 Spanish background ($n = 308$, mean age 16;4). The results from language tests and linguistic background questionnaires showed that the use of Spanish by their parents at home was the most reliable factor responsible for the maintenance of Spanish proficiency. Unlike these studies, we investigated L1 lexical attrition from young heritage children with a relatively short experience with the L2, yet found a reliable effect of L2 experience on their L1 lexical retrieval. In this regard, the present results furnish compelling empirical evidence for the role of language experience in the early stages of the L1 attrition process.

The influential role of language experience was demonstrated most prominently by Hopp (2011), who showed that length of exposure was a stronger predictor than L2 AoA for accuracy with the German DP among 60 child L2 learners of German (age range = 3;5–7;0). When Hopp controlled for the length of exposure, the observed correlation between age or L2 AoA with accuracy scores disappeared; however, the length of exposure remained significantly correlated with accuracy when age or L2 AoA was controlled for. Similarly, the current study showed that the two language experience variables, LoR and the input ratio, were highly influential factors in word naming accuracy and speed in the L1 task, whereas L2 AoA was not. In addition, when the input ratio was controlled for, the marginal correlation of L2 AoA and accuracy scores failed to reach significance; however, when L2 AoA was controlled for, the accuracy scores still significantly correlated with both LoR and input ratio. Therefore, contra the critical period account, our findings lend credence to the claim that language experience exerts a stronger influence in heritage language attrition. Moreover, given that previous studies have focused mostly on young bilinguals, this study expands the

scope of this research field by corroborating the instrumental role of language experience in heritage language attrition for older participants (age range = 11–14).

It is important to note that between the two language experience variables, the input ratio was more likely than LoR to explain the participants' accuracy scores and RTs in the L1 task. We found stronger correlations between the input ratio and the L1 task outcomes than between LoR and the L1 task outcomes. Moreover, when L2 AoA was controlled for, the accuracy scores in the L1 task correlated with the input ratio ($r = -.726$) more strongly than with LoR ($r = -.484$), as did the RTs (input ratio, $r = .569$; LoR, $r = .307$). This suggests that our estimate of the input ratio was better able than LoR to account for the heritage speakers' language experience because it captured the significant variability among the bilingual children in the relative amounts of language input they received from various sources.

While it seems clear that amount of language exposure affects the linguistic development and loss of the L1 of bilingual children in various domains, we caution against the idea that input alone explains acquisition and attrition. As Paradis and Genesee (1996) discussed in the case of simultaneous bilingual children, the relationship between input quantity and linguistic development is not linear. It is possible that despite significantly less input, some bilingual children may arrive at the same developmental milestones as monolingual children within the same time frame. We therefore conclude that several factors work in concert to affect L1 attrition, with experience factors (LoR and input ratio) being a more reliable predictor than L2 AoA for L1 attrition. In this regard, our study offers a helpful direction for future research by suggesting that any study on language attrition needs to consider the relative amount of L1/L2 input in addition to other relevant factors.

8. Conclusion

This study offers a novel finding that L2-L1 ratio and LoR – but not L2 AoA – predict the early signs of L1 attrition. Our results suggest that what matters more in language attrition is not when a heritage learner is exposed to the majority language, but how much input the learner receives from the majority language relative to the heritage language (Grüter, Hurtado, Marchman, & Fernald, 2014). To our knowledge, few studies have explicitly compared the roles of age and language experience factors on L1 attrition in the domain of spontaneous lexical production. In addition, while previous heritage language acquisition studies have been conducted largely with heritage speakers living in the United States or Europe, the current study explores a new population, L2-Korean children with Chinese and Russian as heritage languages, filling a research gap in this area of study and extending the exploration of the issue of L1 attrition to a relatively understudied context. We also acknowledge some limitations of the study; in particular, by relying on vocabulary accuracy and production speed, the study may bypass attrition at other levels of linguistic knowledge. To better understand the effects of L2 AoA and experience factors on language attrition, further studies need to examine a wider range of linguistic phenomena, as well as measuring accuracy and response times in other linguistic domains. Second, results from our cross-sectional design may fall short of convincingly showing the evidence of attrition since we did not test participants' lexical abilities before any language loss or attrition took place. As a reviewer pointed out, in order for one to argue for attrition, it is essential to ensure

that all participants had stable L1 knowledge in the first place. Although we took different degrees of word naming accuracies and retrieval speed in the L1 over the L2 HALA task as a proxy for L1 attrition, we cannot dismiss the possibility that some participants may have started learning the L2 at different levels of L1 lexical proficiency. Further research must address this point by systematically controlling for children's L1 knowledge before attrition starts or conducting a longitudinal study. Another limitation comes from the limited number of data sets, particularly the small number of items and their restricted semantic categories in the HALA task. We used 31 items that refer to human body parts because they have the same lexical-semantic frequency across languages (Kang, 2011; O'Grady et al., 2009) and hardly constitute cross-linguistic cognates. As a reviewer pointed out, however, this body-part naming test with only 31 items may gloss over potential effects of semantic contents, which points to a need for future studies that include a greater number of items in a broader range of semantic categories. Despite these limitations, we believe that the novel attempt in the current study offers a stepping-stone for future studies by highlighting the importance of considering language experience factors as indicators of language attrition.

Supplementary Material. For supplementary material accompanying this paper, visit <https://doi.org/10.1017/S1366728921001139>

References

- Baayen RH (2008) *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Baayen RH, Davidson DJ and Bates DM (2008) Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390–412.
- Balota DA, Yap MJ, Hutchison KA, Cortese MJ, Kessler B, Loftis B, Neely JH, Nelson DL, Simpson GB and Treiman R (2007) The English lexicon project. *Behavior Research Methods* 39(3), 445–459.
- Barr DJ, Levy R, Scheepers C and Tily HJ (2013) Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68, 255–278.
- Bylund E (2009) Maturation constraints and first language attrition. *Language Learning* 59(3), 687–715.
- De Bot K (2004) Introduction: Special issue on language attrition. *International Journal of Bilingualism* 8(3), 233–237.
- Dörnyei Z and Csizér K (2012) How to design and analyze surveys in SLA research? In Mackey A and Gass S (eds), *Research methods in second language acquisition: A practical guide*. Malden, MA: Wiley-Blackwell, pp. 74–94.
- Duncan TS and Paradis J (2020) Home language environment and children's second language acquisition: the special status of input from older siblings. *Journal of Child Language* 47, 982–1005.
- Gollan TH, Montoya RI, Cera C and Sandoval TC (2008) More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language* 58, 787–814.
- Grüter T, Hurtado N, Marchman VA and Fernald A (2014) Language exposure and online processing efficiency in bilingual development: relative versus absolute measures. In Grüter T and Paradis J (eds), *Input and experience in bilingual development*. Amsterdam: John Benjamins, pp. 15–36.
- Hakuta K and D'Andrea D (1992) Some properties of bilingual maintenance and loss in Mexican background high-school students. *Applied Linguistics* 13, 72–99.
- Hopp H (2011) Internal and external factors in the child L2 development of the German determiner phrase. *Linguistic Approaches to Bilingualism* 1(3), 238–264.
- Hox JJ, Moerbeek M and Van de Schoot R (2017) *Multilevel analysis: Techniques and applications* (3rd ed.). New York, NY: Routledge.
- Jaeger TF (2008) Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59, 434–446.
- Jessner U (2003) A dynamic approach to language attrition in multilinguals. In Cook V (ed.), *Effects of the second language on the first*. Clevedon: Multilingual Matters, pp. 234–247.
- Jia G and Aaronson D (2003) A longitudinal study of Chinese children and adolescents learning English in the US. *Applied Psycholinguistics* 24(1), 131–161.
- Jo K, Kim K and Kim H (2021) Children's interpretation of negation and quantifier scope in L3 English. *Bilingualism: Language and Cognition* 24(3), 427–438.
- Kagan O and Dillon K (2013) Heritage languages and L2 learning. In Gass S and Mackey A (eds), *The Routledge handbook of second language acquisition*. New York: Routledge, pp. 509–523.
- Kang S-G (2011) *English attrition in Korean-English bilingual children*. (Unpublished doctoral dissertation). University of Hawaii at Manoa.
- Kim J-H, Montrul S and Yoon J (2009) Dominant language influence in acquisition and attrition of binding: Interpretation of the Korean reflexive *caki*. *Bilingualism: Language and Cognition* 13(1), 73–84.
- Köpke B and Schmid MS (2004) Language attrition. In Köpke B, Schmid MS, Keijzer M and Weilemar L (eds), *First language attrition: Interdisciplinary perspectives on methodological issues*. Amsterdam: John Benjamins, pp. 1–43.
- Lenneberg E (1967) *Biological Foundations of Language*. New York: Wiley.
- 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 (2008) *Incomplete acquisition in bilingualism: Re-examining the age factor* (Vol. 39). Amsterdam: John Benjamins Publishing.
- Montrul S (2010) Current issues in heritage language acquisition. *Annual Review of Applied Linguistics* 30, 3–23.
- Montrul S and Ionin T (2012) Dominant language transfer in Spanish heritage speakers and second language learners in the interpretation of definite articles. *The Modern Language Journal* 96(1), 70–94.
- Montrul S and Sánchez-Walker N (2013) Differential object marking in child and adult Spanish heritage speakers. *Language Acquisition* 20(2), 109–132.
- O'Grady W, Lee O-S and Lee J-H (2011) Practical and theoretical issues in the study of heritage language acquisition. *Heritage Language Journal*, 23–40.
- O'Grady W, Schafer AJ, Perla J, Lee O-S and Wieting J (2009) A psycholinguistic tool for the assessment of language loss: The HALA project. *Language Documentation and Conservation* 3(1), 100–112.
- Pallier C (2007) Critical periods in language acquisition and language attrition. In Köpke B, Schmid M, Keijzer M and Dosterst S (eds), *Language attrition: Theoretical perspectives* (Vol. 33). Amsterdam: John Benjamins, pp. 155–168.
- Paradis J and Genesee F (1996) Syntactic acquisition in bilingual children: autonomous or interdependent? *Studies in Second Language Acquisition* 18, 1–25.
- Penfield W and Roberts L (1959) *Speech and Brain Mechanisms*. Princeton, NJ: Princeton University Press.
- Polinsky M (2011) Reanalysis in adult heritage language: New evidence in support of attrition. *Studies in Second Language Acquisition* 33(2), 305–328.
- R Core Team. (2018) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Unsworth S, Argyri F, Cornips L, Hulk A, Sorace A and Tsimpli I (2014) The role of age of onset and input in early child bilingualism in Greek and Dutch. *Applied Psycholinguistics* 35(4), 765–805.
- Unsworth S, Hulk A and Marinis T (2011) Internal and external factors in child second language acquisition. *Linguistic Approaches to Bilingualism* 1(3), 207–212.
- Ventureyra VAG, Pallier C and Yoo H-Y (2004) The loss of first language phonetic perception in adopted Koreans. *Journal of Neurolinguistics* 17, 79–91.